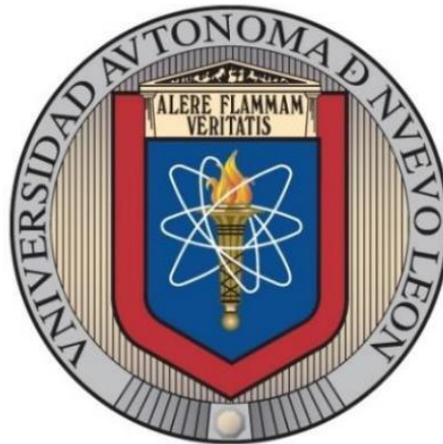


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA



**PREDICCIÓN DE INTENSIDADES EMOCIONALES EN UNA
RED SOCIAL**

POR

FERNANDO MANUEL RODRÍGUEZ ALDAPE

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE
DOCTOR EN INGENIERÍA CON ORIENTACIÓN EN TECNOLOGÍAS
DE LA INFORMACIÓN**

ENERO, 2018

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO



**PREDICCIÓN DE INTENSIDADES EMOCIONALES EN UNA
RED SOCIAL**

POR

FERNANDO MANUEL RODRÍGUEZ ALDAPE

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE
DOCTOR EN INGENIERÍA CON ORIENTACIÓN EN TECNOLOGÍAS
DE LA INFORMACIÓN**

ENERO, 2018

Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica
Subdirección de Estudios de Posgrado

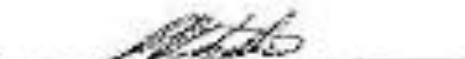
Los miembros del Comité de Tesis recomendamos que la Tesis «Predicción de intensidades emocionales en una red social», realizada por el alumno Fernando Manuel Rodríguez Aklago, con número de matrícula 1183110, sea aceptada para su defensa como requisito parcial para obtener el grado de Doctorado en Ingeniería con Orientación en Tecnologías de la Información.

El Comité de Tesis



Dra. Sara Elena Garza Villareal

Asesora



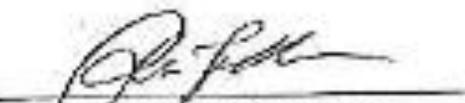
Dr. Héctor Gibrán Ceballos Cauceño

Revisor



Dr. Hugo Juli Eschante Calderas

Revisor



Dra. Sara Elisa Schaeffer

Revisora



Dr. Francisco Torres Guerrero

Revisor

M. D.


Dr. Simón Martínez Martínez

Subdirección de Estudios de Posgrado



San Nicolás de los Garza, Nuevo León, enero 2018

A mi familia, quienes me han brindado su apoyo y amor incondicional.
A mis amigos, que en algún momento me han preguntado por mi avance
y por el tema de mi tesis.
A mi asesora de tesis, que me ha dedicado de su tiempo personal
y siempre ha estado allí.

ÍNDICE GENERAL

Agradecimientos	XII
Resumen	XIII
1. Introducción	1
1.1. Definición del problema	4
1.2. Motivación y justificación	7
1.3. Protocolo	10
1.3.1. Hipótesis	10
1.3.2. Preguntas de investigación	10
1.3.3. Objetivos	11
1.4. Contribuciones	12
1.5. Organización del documento	14
2. Marco Teórico	16
2.1. Teoría de emociones	16
2.1.1. Emociones básicas	18

2.2. Redes sociales en línea	19
2.2.1. Twitter	21
2.3. Aprendizaje máquina	22
2.3.1. Regresión lineal	24
2.3.2. Máquinas de soporte vectorial	25
2.4. Series de tiempo	26
2.5. Métricas de evaluación	31
2.5.1. Clasificación	31
2.5.2. Regresión	32
2.6. Acuerdo de intercorrelación	33
2.7. Resumen	34
3. Estado del Arte	35
3.1. Predicción, representación, detección y medición de emociones	35
3.1.1. Representación de emociones	36
3.1.2. Detección y medición de emociones	38
3.1.3. Datos históricos y ventana de tiempo	46
3.1.4. Predicción de emociones	47
3.2. Trabajos relacionados	52
3.3. Resumen	59
4. Modelo	61

4.1. Tarea de predicción	62
4.2. Intensidad emocional	63
4.3. Descripción del modelo	65
4.4. Función de predicción	68
4.5. Fases en la predicción de intensidad emocional	69
4.5.1. Detección de emociones	70
4.5.2. Medición de variables de entrada	73
4.5.3. Predicción	74
4.6. Resumen	75
5. Experimentos y Resultados	77
5.1. Repositorio	77
5.2. Detección de emociones	79
5.2.1. Configuración	80
5.2.2. Resultados	85
5.2.3. Discusión	86
5.3. Predicción de intensidad emocional	88
5.3.1. Configuración	88
5.3.2. Resultados	96
5.3.3. Discusión	105
5.4. Comprobación de hipótesis	108

5.5. Resumen	110
6. Conclusiones y Trabajo futuro	112
6.1. Resumen	112
6.2. Comentarios finales	113
6.2.1. Clasificadores automáticos	114
6.2.2. Modelo	115
6.3. Respuesta a las preguntas de investigación	116
6.4. Contribuciones	118
6.5. Posibles aplicaciones	118
6.6. Trabajo futuro	119
6.6.1. Clasificadores automáticos	120
6.6.2. Modelo	120

ÍNDICE DE FIGURAS

2.1. Emociones básicas según Ekman (1992).	19
2.2. Máquinas de soporte vectorial (ejemplo).	27
3.1. Representación de emociones.	38
3.2. Detección de emociones.	39
3.3. Predicción de emociones.	52
4.1. Metodología propuesta para la detección de emociones	71
5.1. Prototipo para etiquetado manual	82
5.2. Error absoluto medio de los modelos más relevantes.	100
5.3. Correlación para los modelos más relevantes.	102
5.4. Resumen de comparación entre modelos.	103

ÍNDICE DE CUADROS

2.1. Emociones básicas según varios autores.	20
3.1. Características empleadas por Nguyen <i>et al.</i> (2012).	50
3.2. Comparativa de otros trabajos relacionados con el propuesto (I).	54
3.3. Comparativa de otros trabajos relacionados con el propuesto (II).	55
3.4. Comparativa de otros trabajos relacionados con el propuesto (III).	56
4.1. Resumen de conceptos relevantes en la tarea de predicción	63
5.1. Cantidad de evaluaciones y evaluadores que participaron.	83
5.2. Resumen de clasificaciones de comentarios (preliminar).	83
5.3. Resumen de clasificaciones de comentarios (final).	84
5.4. Evaluación de clasificadores de emoción	86
5.5. Resumen de configuraciones.	89
5.6. Comparación de modelos utilizando error absoluto medio.	99
5.7. Comparación de modelos utilizando correlación.	101
5.8. Resultados de predicción de transición emocional.	103

5.9. Resultados comparativos para predicción de transición emocional. . .	104
---	-----

AGRADECIMIENTOS

A mi asesora de tesis, la Dra. Sara Elena Garza Villarreal, que gracias a su supervisión, extrema paciencia y buena disposición ha sido posible la realización de este trabajo; sin su apoyo, hubiese sido imposible.

A mis revisores de tesis, la Dra. Satu Elisa Schaeffer, el Dr. Francisco Torres Guerrero, el Dr. Hugo J. Escalante Balderas y el Dr. Héctor G. Ceballos Cancino, quienes dedicaron tiempo y esfuerzo en revisar la tesis y nos apoyaron con su consejo durante este tiempo.

A la Universidad Autónoma de Nuevo León y a la Facultad de Ingeniería Mecánica y Eléctrica por proporcionarme un lugar para trabajar y aprender, y porque me apoyaron económicamente con los gastos para estudiar este doctorado.

A las personas que me permitieron utilizar sus cuentas de Twitter para obtener el repositorio de comentarios.

A mis amigos, quienes estuvieron al pendiente y me preguntaban sobre el avance de mi tesis, e incluso me dieron palabras de aliento para seguir adelante e ideas para experimentar.

A mi madre, Ana Cristina Aldape, que siempre me apoyó y estuvo al pendiente durante este tiempo.

A toda mi familia, que ha estado siempre ahí y me ha apoyado incondicionalmente en cada decisión. Sin ellos, este objetivo habría sido impensable.

RESUMEN

Fernando Manuel Rodríguez Aldape.

Candidato para obtener el grado de Doctorado en Ingeniería con Orientación en
Tecnologías de la Información.

Universidad Autónoma de Nuevo León.

Facultad de Ingeniería Mecánica y Eléctrica.

Título del estudio: PREDICCIÓN DE INTENSIDADES EMOCIONALES EN UNA RED
SOCIAL.

Número de páginas: 144.

OBJETIVOS Y MÉTODO DE ESTUDIO: Las emociones juegan un papel fundamental en la toma de decisiones y en el comportamiento humano. Actualmente, es común ver que los usuarios de redes sociales plasmen en sus publicaciones contenido emocional, el cual puede ser estudiado con diversos fines. En este trabajo, se propone un modelo para la predicción de emociones en redes sociales. Específicamente, este modelo predice, para un usuario, la proporción de comentarios que publicará con una emoción en particular para un periodo de tiempo futuro; a esta proporción la definimos como *intensidad emocional*. Contrario a otros modelos que se enfocan en una sola emoción, el modelo propuesto considera un esquema básico de emociones y utiliza todas las emociones para la predicción de la intensidad de cada emoción. El modelo, además, utiliza tres fuentes distintas de información: (1) proveniente del

usuario, (2) proveniente de los contactos del usuario y (3) proveniente del entorno. La predicción finalmente se hace mediante regresión lineal.

CONTRIBUCIONES Y CONCLUSIONES: Encontramos que nuestro modelo, aplicado sobre un conjunto de datos extraído de Twitter, produce mejores resultados (hasta en un 35 %) comparado con modelos como ARIMA y mejores resultados en comparación con otros 27 modelos base y un método del estado del arte. Entre algunas aplicaciones que se pueden dar a una herramienta de este tipo, se encuentra la predicción de caídas en el nivel de alegría (o aumento de tristeza) del usuario y, en consecuencia, la generación de recomendaciones para ayudar al bienestar emocional del usuario (por ejemplo, de recuerdos alegres). Las emociones positivas contribuyen a un estado emocional favorable para el bienestar físico y mental, y por lo tanto, a mejorar la salud Kim *et al.* (2014).

Las siguientes contribuciones son, además, las principales de esta tesis:

- Introducción del concepto de *intensidad emocional*. Si bien existen trabajos que predicen estados emocionales para los usuarios en una red social, el modelo que se propone predice en específico la proporción de comentarios que serán escritos con una emoción en particular. Hasta donde sabemos, es el primer trabajo que propone este concepto.
- Un modelo de predicción de emociones. Este modelo permite pronosticar emociones básicas de los usuarios a partir de su información personal, la información de sus contactos (“contagio emocional”) e información del medio ambiente. El modelo en sí mismo considera que una emoción tiene influencia sobre otras.
- La utilización de emociones básicas para un modelo predictivo. Se emplean las emociones de alegría, enojo, miedo y tristeza. Esto permite trabajar con un mayor nivel de granularidad — contrario a esquemas que solo manejan etiquetas como “positivo” o “negativo”. La anterior representación numérica o abstracta (positivo y negativo) pudiera ser suficiente para el área comercial

o de mercadotecnia, en donde se busca tener una noción de la aceptación de un producto, por ejemplo, pero pudiera no serlo en un área médica, en donde se busca analizar patrones de agresividad, depresión, estrés, etcétera.

- El considerar interdependencia de emociones para la predicción. Otros trabajos utilizan solo datos históricos de una única emoción sin tomar en cuenta la existencia de otras. En modelos donde se registran las emociones día con día o por intervalos de tiempo, con frecuencia se hace uso de una sola emoción. Proponemos, por tanto, emplear varias emociones y asumimos que estas interactúan.
- La generación de recursos para el idioma español. Se generó un conjunto de comentarios etiquetados de acuerdo con las emociones que expresan (alegría, tristeza, miedo y enojo).
- Una metodología para detectar (reconocer) emociones en texto. Esta metodología conjunta técnicas de clasificación de sentimiento en español con otras técnicas de aprendizaje máquina.

Firma de la asesora: _____

Dra. Sara Elena Garza Villarreal

CAPÍTULO 1

INTRODUCCIÓN

Las emociones juegan un papel fundamental en la toma de decisiones. Por ejemplo, Small y Verrochi (2009) han encontrado, para una campaña de publicidad, que mostrar un rostro que refleja tristeza tiene un mayor impacto en las donaciones de caridad que mostrar un rostro que refleja felicidad o uno que no refleja ninguna emoción. También se han implementado agentes inteligentes cobradores que promueven ansiedad en los deudores (relacionada con la emoción del miedo) para crear un deseo por pagar y aeromozas que generan un ambiente de buen ánimo en un avión (Tsai *et al.*, 2011a). Al afectar el organismo, la influencia que ejercen las emociones puede llevar al ser humano a ejecutar acciones pequeñas, tales como comprar un producto (Fenton-O’Creivy *et al.*, 2011), o acciones más radicales, como lastimarse a sí mismo (Desmet y Hoste, 2013; Pisani *et al.*, 2013), o incluso iniciar una guerra (Shepherd *et al.*, 2013).

A pesar de que las emociones son relevantes para el comportamiento del ser humano, hasta ahora se ha visto que son difíciles de comprender y de definir. Existen varias teorías sobre el surgimiento de las mismas y sobre cómo estas nos afectan, pero actualmente no existe un consenso sobre cómo definir una emoción, e incluso se han dado diferentes definiciones según la cultura. Asimismo, se han concebido diferentes orígenes para las emociones y su desarrollo en las personas — por ejemplo, se encuentran la evolución y la selección natural, la sociedad y la cultura, o los estímulos

y las reacciones (Johnson, 2009). Se tiene también un debate sobre cuáles son las emociones que somos capaces de experimentar tanto seres humanos como animales, y lo que son las *emociones básicas* o elementales, de las cuales se derivan las emociones *complejas*, como los celos o la desesperanza; los autores difieren dependiendo de las corrientes y de diversos estudios realizados. Por ejemplo, Matsumoto (2007) define 22 emociones básicas, mientras que Ekman (1992) solamente seis, e Izard (2007), a quien tomamos de base en esta tesis, cuatro: alegría, tristeza, enojo y miedo. Todo lo anterior nos lleva a que el estudio de las emociones no solamente resulte complicado, sino retador. En ese sentido, este es actualmente un campo de estudio abierto en las áreas de psicología, ciencias sociales y, recientemente, las ciencias computacionales. En esta área en particular, las emociones han sido estudiadas a través de disciplinas como el cómputo afectivo (Cambria, 2016) y la minería de datos (Tan *et al.*, 2006).

Un medio por excelencia para expresar emociones hoy en día es el Internet, que en nuestros tiempos se ha convertido en una herramienta clave para la búsqueda de información y comunicación. Se ha observado que los usuarios pasan cada vez más tiempo en sitios como Twitter¹, Facebook² y YouTube³, entre otros (Chmiel *et al.*, 2011a). La sofisticación también ha ido en aumento y ha permitido la aparición de espacios en donde los usuarios pueden intercambiar opiniones, información, imágenes, videos y enlaces a otras páginas, además de poder estar conectados con otros usuarios (Liu, 2010; Rodríguez Aldape, 2013). A este tipo de espacios se les conoce como foros, comunidades o *redes sociales en línea* (en inglés, *social networking sites*). Los *microblogs*, redes sociales en línea donde los mensajes publicados son cortos, han sido recientemente objeto de una numerosa cantidad de investigaciones. Por ejemplo, Pak y Paroubek (2010) fueron los pioneros en realizar análisis de sentimiento sobre Twitter (es decir, en estudiar texto subjetivo sobre este microblog). Otros microblogs y redes sociales en línea incluyen Plurk⁴, Weibo⁵, Facebook, Instagram⁶, YouTube

¹Disponible en <http://www.twitter.com>.

²Disponible en <http://www.facebook.com>.

³Disponible en <http://www.youtube.com>.

⁴Disponible en <http://www.plurk.com>.

⁵Disponible en <http://www.weibo.com>.

⁶Disponible en <http://www.instagram.com>.

y Tumblr⁷, que al momento de la escritura de esta tesis se encuentran vigentes.

Las redes sociales en línea se han utilizado exitosamente como herramientas de mercadotecnia (Rodríguez Aldape, 2013), principalmente debido a la continua generación y propagación de contenido por parte de los usuarios. En cuanto al contenido generado, se ha encontrado evidencia de que, además de opiniones e información, en las redes sociales en línea se pueden expresar y transmitir emociones de un usuario a otro (Bao *et al.*, 2012; Bispo y Paiva, 2009; Bliss *et al.*, 2012; Chmiel *et al.*, 2011a; Hoogendoorn *et al.*, 2011; Kim *et al.*, 2014). A la transmisión de emociones de una persona a otra se le conoce como *contagio emocional*. De acuerdo con Hatfield y Cacioppo (1994), el contagio emocional es la tendencia de converger emocionalmente a través de imitar y sincronizarse automáticamente con los movimientos, posturas y vocalizaciones de otra persona. Se ha mostrado en ciencias de la administración que, mediante el contagio emocional, se puede mejorar la cooperación, reducir el conflicto e incrementar la productividad percibida en grupos y organizaciones (Barsade, 2002). Este contagio emocional se da en las redes sociales en línea a partir de la lectura de los comentarios publicados en ellas.

El estudio de las emociones (por ejemplo, del contagio emocional), aunado al crecimiento del Internet, ha propiciado la investigación relacionada con la representación, detección, medición, análisis y *predicción* de emociones en medios sociales. Aún cuando los trabajos han florecido, como veremos en la sección 1.1, también existen retos por abordar y áreas de oportunidad. Todo ello ha empujado hacia el desarrollo de nuevas ideas.

⁷Disponible en <http://www.tumblr.com>.

1.1 DEFINICIÓN DEL PROBLEMA

A continuación, se define el problema de investigación, comenzando por describir el contexto general y los antecedentes más relevantes que enmarcan este contexto. Más adelante, se definen el objeto, variables de estudio y alcance de la tesis.

CONTEXTO PROBLEMÁTICO.- Uno de los problemas de investigación modernos que se busca resolver es la comprensión de fenómenos sociotecnológicos de gran escala en donde interactúan personas y sistemas conectados a través de una red. Por ejemplo, se desea comprender el comportamiento de una población ante una situación como una crisis económica mundial (Bennett y Segerberg, 2011; Bollen *et al.*, 2011) o durante el proceso de escape en una evacuación (Tsai *et al.*, 2011b). También se ha investigado el nivel de felicidad en una población (Dodds *et al.*, 2011) y la propagación de creencias (Acemoglu *et al.*, 2010; Lynch, 1996).

Desde que Picard (1995) propuso el estudio de las emociones humanas en el área de las ciencias computacionales (llamado *computación afectiva*), han surgido numerosos modelos de emociones. Puesto que el objetivo de estos modelos computacionales es generar una respuesta apropiada a las emociones (Cambria *et al.*, 2012), el área de cómputo afectivo incluye el estudio y desarrollo de sistemas (así como dispositivos) que puedan reconocer, interpretar, procesar y simular estas. Algunos dispositivos contemplados en los trabajos alusivos al área abarcan teléfonos móviles (Maxhuni *et al.*, 2017) e, inclusive, el *internet de las cosas* (Kajiwara *et al.*, 2017).

Las redes sociales en línea como Twitter y Facebook también han hecho posible tanto el estudio de usuarios individuales (Qiyao *et al.*, 2016), como el de grandes poblaciones (Dodds *et al.*, 2011). En ese sentido, se han podido analizar de manera colectiva, por ejemplo, los cambios emocionales debido al clima en una cierta población (Hannak *et al.*, 2012) o el aumento y disminución de emociones (Nguyen *et al.*, 2012). Otros trabajos han estudiado las emociones dentro de las redes sociales con

finés específicos, tales como la prevención de estrés en adolescentes (Li *et al.*, 2015; Zhao *et al.*, 2015), la detección de bipolaridad (Cochran *et al.*, 2017) y la prevención de la depresión (Moore *et al.*, 2012). Sin embargo, algunos de estos estudios se basan en conjuntos de emociones limitadas, como la felicidad (Biever, 2010; Bliss *et al.*, 2012), o emociones etiquetadas simplemente como positivas o negativas (Mogadala y Varma, 2012), por lo que otros patrones en emociones (como aquellos relacionados con la tristeza o el enojo) no han sido descubiertos, y esquemas más sofisticados de emociones no han sido incorporados. Por ejemplo, desde los años 70, Ekman (1992) encontró evidencia de la existencia de seis emociones básicas: enojo, repulsión, miedo, alegría, tristeza y sorpresa. Estudios más recientes han propuesto la existencia de solo cuatro emociones básicas: alegría, enojo, miedo y tristeza (Izard, 2007). Si se incorporaran estos esquemas al análisis y predicción de emociones en redes sociales, podrían brindar una mayor información al respecto.

Asimismo, algunos trabajos que predicen el *estado emocional* de los usuarios en redes sociales en línea (es decir, que predicen la emoción o emociones presentes en el usuario) etiquetan al usuario con una sola emoción (Zhang *et al.*, 2010), cuando esta visión podría generalizarse al considerar el estado emocional como un conjunto de emociones con diferentes intensidades. Otros trabajos se enfocan en una sola emoción, como la alegría (Dodds *et al.*, 2011). Es decir, por una parte predicen solamente una emoción. Por otra parte, al considerar las variables de entrada para hacer la predicción (pronóstico), utilizan solamente la misma emoción. Lo anterior implica que se omiten los valores del resto de las emociones (es decir, que el valor futuro de una emoción depende exclusivamente de los valores anteriores de esta misma emoción y no del resto de las demás emociones del esquema considerado). Sin embargo, se ha visto que las emociones afectan la percepción y tienden a permanecer un cierto tiempo en el ser humano, por lo que un evento del presente puede ser mermado o aumentado dependiendo de las distintas emociones que se hayan experimentado en el pasado (DeWall *et al.*, 2016; Lerner *et al.*, 2003; Verduyn y Lavrijsen, 2015; Yin *et al.*, 2013).

Otra limitante de los trabajos actuales es que la mayoría de estos (nos referimos a los que trabajan con texto) están confinados al idioma inglés, pues también la mayoría de los recursos disponibles (diccionarios, lexicones y otras herramientas) están solamente disponibles en dicho idioma.

Es por lo anterior que proponemos la creación de un modelo de predicción de emociones — específicamente, de *intensidad emocional* — que incluya información proveniente (1) del usuario, (2) de sus contactos y (3) del entorno o medio ambiente. Para comprender este modelo, definimos a continuación el objeto y variables de estudio.

OBJETO DE ESTUDIO.- El objeto de estudio en esta tesis son las *redes sociales en línea*, puesto que se extraen y utilizan los *comentarios*, *relaciones* entre usuarios e *información temporal* de una de estas redes (de un microblog, específicamente) para hacer la predicción de intensidades emocionales.

VARIABLES DE ESTUDIO.- Dentro del modelo predictivo propuesto, se contemplan dos tipos de variables de estudio. La primera de ellas es la *intensidad emocional* (extensamente descrita en el capítulo 4), que consiste en la proporción de comentarios que escribe el usuario con una emoción específica para un periodo de tiempo en particular (llamado *ventana de tiempo*). Por ejemplo, si diez de cien comentarios en un día contienen alegría, la intensidad emocional del usuario para este día con la emoción de alegría se calcula como $\frac{10}{100} = 0.1$.

Note que un usuario tendrá tantas intensidades emocionales como emociones se contemplen; por ejemplo, en un esquema de cuatro emociones básicas, habrá cuatro intensidades emocionales por usuario en una ventana de tiempo. En consecuencia, el estado emocional del usuario estará caracterizado por una serie de intensidades emocionales, de acuerdo con el esquema de emociones que se esté manejando.

La intensidad emocional es la variable objetivo a predecir. Este mismo tipo de variable, con diferentes conjuntos de usuarios, conjuntos de emociones y conjuntos

de ventanas de tiempo, será considerado como entrada para hacer la predicción para ventanas de tiempo futuras. El segundo tipo de variable es el *día de la semana*. Esta variable también se explica con mayor detalle en el capítulo 4.

En cuanto al alcance de la tesis, la predicción de intensidades emocionales se hace mediante los siguientes supuestos y restricciones:

1. Utilizamos Twitter como caso de estudio. Se emplea esta red debido a su popularidad, a que los datos son públicos, a la cantidad de datos que hay disponibles, a la existencia de datos en español y a la existencia de herramientas (por ejemplo, el API⁸ de Twitter) para extraer datos de la red. Se utiliza un repositorio en particular, descrito en el capítulo 5.
2. Como simplificación al marco psicológico, tal como se explica más adelante en el capítulo 2, consideramos como sinónimos los términos de *emoción*, *humor* y *sentimiento*, aunque estos términos tienen diferentes connotaciones en contextos más especializados.
3. Asumimos que *pronóstico* es una subtarea de *predicción*, aunque existen diversas opiniones al respecto en diferentes áreas. Estaremos utilizando estos dos términos de manera intercambiable.

1.2 MOTIVACIÓN Y JUSTIFICACIÓN

Se expone la motivación para el desarrollo de esta tesis en tres áreas principales: (1) el interés comercial y de aplicaciones prácticas, (2) el interés por parte de la sociología de comprender las emociones y (3) el diseño de mejores interfaces humano-computadora.

En cuanto al interés comercial y de aplicaciones prácticas, se pueden utilizar modelos de predicción de emociones de una sociedad en sistemas de recomendación

⁸Disponible en: <https://developer.twitter.com/>.

(García-Cumbreras *et al.*, 2013; Niu *et al.*, 2013), en videojuegos (Cai *et al.*, 2013; Ng *et al.*, 2012; Popescu, 2012), en productos comerciales (Fung, 2012; Walter, 2012) y, en general, en dispositivos con interfaces humano-computadora para proporcionar al usuario una mejor experiencia (Botella *et al.*, 2012; Kim y Kwon, 2010; Kirby *et al.*, 2010; Riva *et al.*, 2012). Dentro del diseño de interfaces, se puede incluir también el uso de modelos emocionales en robots para facilitar su inclusión en las actividades de los seres humanos, ya que lucirían más naturales, cálidos o “humanos” (Kim y Kwon, 2010; Kirby *et al.*, 2010). Lo anterior resulta interesante, pues existe también un punto a partir del cual se entra en una región de percepción llamada *el valle de extrañeza* (en inglés, *uncanny valley*), en la cual a un humano le resulta “escalofriante” la semejanza que puede tener un robot con su especie, y este tema continúa siendo estudiado (Ho y MacDorman, 2017). Por tanto, se podría analizar si la introducción de las emociones contribuye a expandir esta región (hacerlos más escalofriantes) o a reducirla (hacer que el robot pueda producir calidez y agrado al usuario).

En cuanto a la sociología, esta actualmente busca comprender cómo se manifiestan las emociones en una sociedad y la relación que tienen entre sí (Aknin *et al.*, 2011; Brown *et al.*, 2011; Casti, 2010; Kagan *et al.*, 2013). Conocer anticipadamente las emociones de una persona puede crear sistemas que le ayuden a esta persona en aspectos de su vida — por ejemplo, en momentos en que decae la felicidad; el sistema generaría una emoción positiva cuando se anticipa este decaimiento (Botella *et al.*, 2012; Ho y Spence, 2013; Riva *et al.*, 2012).

En cuanto al punto de vista psicológico, la creación de un modelo de predicción de emociones podría contribuir a diferentes tipos de análisis. Por ejemplo, el modelo puede aplicarse a grupos de usuarios con el fin de estimar las emociones de una población y detectar patrones en los cambios de las mismas.

Entre las aplicaciones prácticas de este modelo se pueden mencionar:

- La creación de visualizaciones gráficas que indiquen cambios y flujos de las

emociones a nivel de una población desde una perspectiva geográfica (mediante el uso de mapas).

- El análisis matemático y estadístico del comportamiento de las emociones en los seres humanos.
- La observación del efecto que tienen los eventos sociales (por ejemplo, culturales, económicos) en las emociones.
- La creación de interfaces de computadora que puedan anticipar las emociones de un grupo de personas y se adapten a ellas a través, por ejemplo, de ofrecer apoyo afectivo.
- La detección del surgimiento de emociones negativas en usuarios específicos para identificar actividades de acoso cibernético y mal comportamiento.
- La detección de problemas de salud del individuo, como la depresión, o problemas colectivos, como tensión social.
- La detección de usuarios falsos (*bots*) que no muestran emociones o lo hacen de forma errante.
- La detección de contenido ofensivo a partir de las reacciones de los usuarios que lo observan.

Predecir las emociones nos permitiría tomar acciones para dar una mejor experiencia al usuario. Mediante los sistemas de recomendación, se le podría sugerir contenido que pueda ayudarlo en alguna depresión o que, incluso, pueda prevenirla. Recientemente, Facebook realizó un estudio en el que filtraba contenido emocional a los usuarios, permitiendo para ello pasar un porcentaje mayor de emociones negativas o positivas y observando los cambios en las emociones de quien los percibía (Kramer *et al.*, 2014). De esta forma se podría, por ejemplo, introducir mayor cantidad de contenido positivo cuando se anticipa una etapa de estrés o depresión. Comercialmente hablando, se podrían analizar los estados emocionales del usuario

asociándolos a comportamientos de compra, de manera que un comerciante pueda esperar el momento adecuado para realizar una oferta de venta. A gran escala, en una campaña política se podría anticipar una mala opinión y tratar de cambiarla antes de que suceda. Se ha visto que las redes sociales en línea predicen bien la opinión pública global acerca de los candidatos políticos (O'Connor *et al.*, 2010).

1.3 PROTOCOLO

A continuación se describe el protocolo que se utilizará en la tesis. Este incluye los objetivos, preguntas de investigación e hipótesis.

1.3.1 HIPÓTESIS

Se describe la hipótesis general a comprobar.

- Es posible crear un modelo para predicción de intensidad emocional en una red social en línea mediante técnicas de procesamiento de texto en español y aprendizaje máquina.

A partir de esta hipótesis se generan varias preguntas de investigación que deben ser constestadas con el fin de comprobarla. Es decir, debemos comprobar primero si podemos trabajar con el idioma Español para detectar emociones, utilizar minería de texto y aprendizaje máquina para hacer la predicción de la intensidad emocional. Estas preguntas se enumeran en la Sección 1.3.2.

1.3.2 PREGUNTAS DE INVESTIGACIÓN

Se describe la pregunta principal en la que se concentra este trabajo, y de esta se desprende una serie de preguntas que al resolverse contribuirán a su consecución.

- ¿Cómo puede crearse un modelo de predicción de intensidad emocional en una red social en línea utilizando técnicas de procesamiento de texto en español y aprendizaje máquina?
 - ¿Es posible utilizar minería de texto en español para este propósito?
 - ¿Es posible detectar emociones básicas en redes sociales a través del análisis de texto en español?
 - ¿Es posible medir la intensidad emocional de los usuarios para cada emoción básica?
 - ¿Es posible crear un modelo de predicción de intensidad emocional de los usuarios utilizando información que procede de ellos, de sus contactos y del medio ambiente?
 - ¿Es posible utilizar el modelo para predecir la intensidad emocional de los usuarios para cada emoción básica?

1.3.3 OBJETIVOS

Se menciona el objetivo principal y los objetivos secundarios que se derivan de este.

Objetivo principal: *Crear un modelo de predicción (pronóstico) de intensidad emocional de los usuarios utilizando técnicas de procesamiento de texto en español y aprendizaje máquina.*

Objetivos secundarios:

- Crear clasificadores automáticos para reconocer las emociones básicas: alegría, enojo, miedo y tristeza.
- Definir el concepto de *intensidad emocional*.

- Definir un modelo de predicción de intensidad emocional que tome en cuenta información proveniente del usuario (por ejemplo, comportamiento y datos personales), de sus contactos y del medio ambiente.
- Analizar ventanas de tiempo con diferentes granularidades para la predicción.
- Predecir las intensidades emocionales de los usuarios en una red social para las cuatro emociones básicas.

1.4 CONTRIBUCIONES

Esta tesis aporta las siguientes contribuciones, que consideramos son las más importantes:

- Un modelo de predicción de emociones. El modelo propuesto permite predecir emociones básicas futuras de los usuarios a partir de su información personal, de sus contactos y del medio ambiente. Es el primer modelo de predicción (o pronóstico) de emociones básicas, dentro de nuestro conocimiento, que además ha sido probado utilizando comentarios en español, aunque el idioma no representa una limitación para su aplicación. El modelo en sí mismo considera que una emoción puede tener influencia sobre otra en el futuro. Parte de información obtenida a partir del usuario, sus contactos y del entorno en intervalos de tiempo o sucesos que denominamos ventanas de tiempo. Estas ventanas de tiempo pueden representar minutos, horas días, etc. o eventos en los que algún suceso tiene lugar como por ejemplo la publicación de un nuevo comentario. El modelo en sí no está limitado a una ventana siguiente si no que puede utilizarse para predecir la intensidad emocional en una ventana más distante en el futuro.
- La introducción del concepto de *intensidad emocional*. Si bien existen trabajos que predicen estados emocionales para los usuarios en una red social en línea,

el modelo que se propone predice en específico la proporción de comentarios que serán escritos con una emoción en particular. Hasta donde sabemos, es el primer trabajo que propone este concepto.

- La utilización de emociones básicas. La mayoría de los trabajos actuales utilizan representaciones numéricas que se interpretan según qué tan positivo o negativo es un comentario. En trabajos más recientes, se utilizan vectores en los que cada componente representa una característica de las emociones; la intención de asignar valores numéricos a sus componentes es representar eficientemente todo el espectro de emociones humanas (Cambria y Hussain, 2012). Nosotros utilizamos las cuatro emociones básicas (alegría, enojo, miedo y tristeza), lo cual nos permite analizarlas de forma individual y obtener resultados que se interpreten de una forma más intuitiva. Esto difiere de otros trabajos; por ejemplo, cuando se reportan resultados en los que una opinión se clasifica como negativa por haberse obtenido un número negativo, en estos trabajos no se está especificando si dicha emoción está relacionada con ansiedad, estrés, depresión o alguna otra condición de cuidado para la cual se requiera una acción correctiva. La representación numérica o abstracta generalmente es suficiente para el área comercial o de mercadotecnia, pero pudiera no serlo — por ejemplo — en un área médica.
- El considerar interdependencia de emociones para la predicción. Otros trabajos toman en cuenta datos históricos de una única emoción sin considerar la existencia de otras. Por ejemplo, en modelos donde se registran las emociones día con día o por cada ventana de tiempo, con frecuencia se emplea una sola emoción (positivo, negativo o neutro), o alguna emoción básica representada como vector. Asumimos que las emociones interactúan y que esta interacción, de conocerse, mejoraría la calidad en la predicción de una emoción futura.

Otras contribuciones relevantes son:

- La incorporación de *contagio emocional* a través del uso de intensidades emo-

cionales de contactos en la red social.

- La creación de clasificadores automáticos de emociones. A través de análisis de texto, creamos clasificadores para las emociones básicas. Asimismo, es posible establecer un punto de partida para otros clasificadores más sofisticados de estas emociones en idiomas como el español. Al incorporar métodos más complejos de clasificación, se podrá tener una referencia con la cual comparar resultados. Nuestro trabajo sería de los pocos que versan sobre detección de emociones básicas en español.
- Un repositorio de comentarios en español con etiquetas de emociones. Los comentarios clasificados manualmente durante los experimentos podrán ser utilizados en otros estudios. Para esta clasificación manual, participaron 141 evaluadores, quienes estuvieron etiquetando los comentarios de forma independiente para cada emoción básica — es decir, los comentarios recibieron una calificación para cada emoción básica estudiada. Este proceso tomó aproximadamente un año.

1.5 ORGANIZACIÓN DEL DOCUMENTO

La tesis se organiza de la siguiente forma: en el capítulo 2, se explican los conceptos básicos que deben entenderse para poder obtener una mejor comprensión del modelo propuesto y su validación. Se incluyen conceptos como aprendizaje máquina, regresión lineal y series de tiempo; también se explican de manera básica las emociones humanas desde los puntos de vista computacional y psicológico. En el capítulo 3, se expone el estado del arte relacionado con el trabajo; se describen los métodos actuales para análisis de sentimiento y detección de emociones en diferentes tipos de contenido, como música, voz, video, texto y fotografías. Asimismo, se discuten los principales trabajos que se relacionan de forma directa con la predicción de emociones a futuro o pronóstico de emociones básicas.

En el capítulo 4, se describe el modelo propuesto: sus características y aquello que lo hace diferente con respecto a otros modelos. Se habla de los principales aspectos, como la ventana de tiempo, los modelos base y la regresión lineal — que es utilizada como método principal durante la predicción. En el capítulo 5, se describe la metodología de validación del modelo, se muestran resultados para la detección de emociones y la predicción de intensidades emocionales, y se provee una discusión de los hallazgos más relevantes.

En el capítulo 6, se exponen las conclusiones — con base en los resultados obtenidos — sobre la creación de clasificadores automáticos y el modelo propuesto. También se discuten los problemas que se presentaron y la forma en que se solucionaron, las posibles mejoras que se pueden realizar y aplicaciones prácticas.

CAPÍTULO 2

MARCO TEÓRICO

En este capítulo, se describen conceptos y teoría básica para la comprensión de capítulos subsecuentes. Se cubrirá lo que son las emociones y por qué son importantes (sección 2.1). También se describirán algunos de los métodos de aprendizaje máquina que fueron utilizados en la experimentación (sección 2.3), así como métricas de evaluación para clasificación y regresión (sección 2.5), y métodos para etiquetado de conjuntos de datos (sección 2.6).

2.1 TEORÍA DE EMOCIONES

Actualmente, debido a la complejidad de las emociones y a la diversidad cultural que hay en el mundo, no existe una definición globalmente aceptada de qué es una emoción. Compilaciones relativamente recientes (Johnson, 2009; Lewis *et al.*, 2010) — basadas en trabajos clásicos (Frijda, 1986; Solomon, 1977) — organizan las teorías de emociones en tres categorías principales: (1) teorías evolucionistas, (2) teorías sociales y culturales y (3) teorías sobre el proceso de emoción. Se menciona que las emociones pueden entenderse de dos formas: como un estado y como un proceso. Ver las emociones como un estado significa que la persona puede experimentar la emoción y esta puede influir en estados futuros y en el comportamiento de la persona. En cambio, ver las emociones como un proceso implica que existe un

estímulo que se recibe y provoca una reacción física en la persona. No obstante, esta última visión de las emociones no toma en cuenta los calificativos subjetivos con los que identificamos estos procesos, tales como “alegría” o “enojo”; más bien toma en cuenta reacciones físicas, como gestos faciales, conductancia de la piel y ritmo cardiaco, entre otras.

Johnson (2009) describe las tres teorías principales sobre las emociones. Resumimos esta perspectiva:

- Las teorías evolucionistas explican la razón por la que las emociones están presentes en los seres humanos. Se basan principalmente en la historia de la humanidad, de los animales en general y de la selección natural.
- Las teorías sociales y culturales se basan en la idea de que las emociones son una construcción social, y que son aprendidas durante la infancia de las personas debido a la interacción social. La idea principal de esta corriente es que las emociones están tan arraigadas a la cultura social que es mejor estudiarlas desde este punto de vista.
- Las teorías del proceso de emoción se centran en la descripción del proceso mismo como tal, comenzando con el análisis de un estímulo que puede ser interno o externo en la persona y continuando con la reacción física que se presenta. Entre las teorías de este tipo podemos mencionar: teorías cognitivas, teorías no cognitivas, y teorías de retroalimentación somática.

En la literatura, con frecuencia se utilizan los términos *emoción*, *humor* y *sentimiento* (en inglés, *emotion*, *mood* y *sentiment*, respectivamente). En general, es aceptado que la emoción es de duración corta y puede durar solo unos segundos (Johnson, 2009). Además, la emoción está dirigida hacia un estímulo que la provoca, por ejemplo, una acción desagradable de una persona. El humor, por otro lado, tiene una duración más larga y puede ser provocado por factores internos y externos de la

persona, como factores psicológicos o físicos (Johnson, 2009). El humor no está dirigido hacia un objeto en particular. Los sentimientos son mucho más duraderos, incluso pueden permanecer durante años; se hace referencia, por ejemplo, al amor que una persona puede sentir por otra. Munezero *et al.* (2014) hacen distinción entre un conjunto de términos más amplio para definir emociones: *afecto*, *sensaciones*, *emociones*, *sentimientos*, y *opinión*.

En esta tesis, cuando utilizamos la palabra *emoción*, hacemos referencia tanto a la emoción, humor o sentimiento que pudieron manifestarse al momento de la publicación de comentarios de los usuarios en la red social. Como se plantea en el capítulo 4, la *intensidad emocional* es una medida de “qué tanto” se encuentra presente una *emoción* particular durante una ventana de *tiempo* en la persona (*usuario*).

2.1.1 EMOCIONES BÁSICAS

Dentro de las teorías evolucionistas, Plutchik (1984) propone que las emociones han evolucionado debido a la selección natural y que están presentes en todos los animales, incluyendo los seres humanos. Su teoría afirma que existen ocho emociones básicas compartidas por todos: enojo, miedo, tristeza, repulsión, sorpresa, anticipación, confianza y alegría. Ekman (1992), por otro lado, basado en estudios de otros investigadores que defienden ideas evolucionistas sobre las emociones, encontró evidencia a través de estudios en expresiones faciales de que existen solo seis emociones básicas: alegría, enojo, miedo, tristeza, repulsión y sorpresa (ver figura 2.1). En un estudio más reciente, sin embargo, se encontró evidencia, a través de estudios realizados con resonancia magnética en el cerebro, de la existencia de solo cuatro emociones básicas (Izard, 2007): alegría, tristeza, enojo y miedo. Podemos notar que, a diferencia del modelo de Ekman, este otro modelo fusiona la sorpresa con el miedo y la repulsión con el enojo. El cuadro 2.1 muestra un resumen de conjuntos de emociones básicas propuestas por diversos autores. Algunos autores utilizan el término felicidad para describir una emoción básica mientras que otros

utilizan alegría tal como se ve en el cuadro 2.1; a lo largo de esta tesis hacemos uso de ambos términos indistintamente.

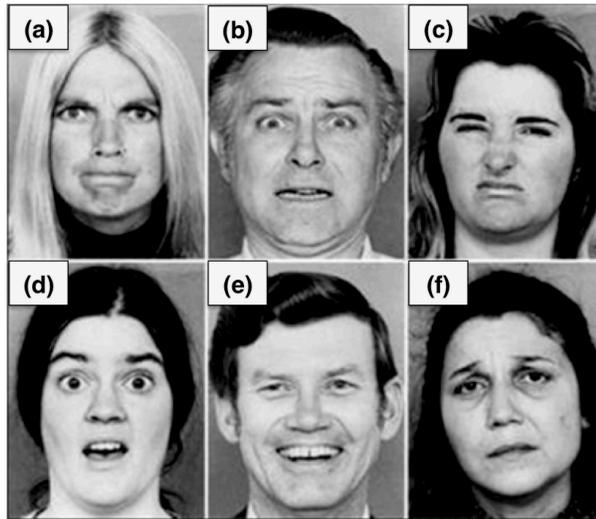


Figura 2.1: Emociones básicas de Ekman: (a) enojo, (b) miedo, (c) repulsión, (d) sorpresa, (e) alegría y (f) tristeza. Imagen original obtenida de <https://autorrealizarte.com/wp-content/uploads/2014/05/emocionesbasicas.png> (modificada para esta tesis).

2.2 REDES SOCIALES EN LÍNEA

El crecimiento del Internet ha permitido la aparición de espacios en donde los usuarios pueden intercambiar opiniones, información, imágenes, videos y enlaces a otras páginas con contenidos diversos, permitiéndoles así estar en contacto con otros usuarios (Liu, 2010; Rodríguez Aldape, 2013). A estos espacios se les conoce como *redes sociales en línea* (en inglés, *social networking sites*), y los usuarios cada vez pasan una mayor cantidad de tiempo sumergidos en ellas (Chmiel *et al.*, 2011a). Algunas de las más populares hoy en día son Twitter, Facebook, YouTube y Google+¹. En años recientes, estas redes sociales han sido utilizadas como herramientas de mercadotecnia con mucho éxito gracias al contenido generado por sus usuarios (Rodríguez Aldape, 2013). Además de opiniones e información, se ha encontrado

¹Disponible en <https://plus.google.com>.

Cuadro 2.1: Emociones básicas según varios autores.

Autor	Emociones	Emociones básicas
Tomkins y Mc Carter (1964)	6	deseo, felicidad, interés, sorpresa, curiosidad, tristeza
Frijda (1988)	6	deseo, felicidad, interés, sorpresa, curiosidad, tristeza
Ekman (1992)	6	odio, repulsión, miedo, alegría, tristeza, sorpresa
Parrott (2001)	6	odio, miedo, alegría, amor, tristeza, sorpresa
Plutchik (2001)	8	aceptación, odio, anticipación, repulsión, alegría, miedo, tristeza, sorpresa
Matsumoto (2007)	22	alegría, anticipación, odio, repulsión, tristeza, sorpresa, miedo, aceptación, timidez, orgullo, aprecio, tranquilidad, admiración, desprecio, amor, felicidad, excitación, arrepentimiento, alivio, incomodidad, respeto, gusto
Izard (2007)	4	alegría, enojo, miedo, tristeza

evidencia de que en estos lugares se pueden expresar y transmitir emociones de un usuario a otro (Bao *et al.*, 2012; Bispo y Paiva, 2009; Bliss *et al.*, 2012; Chmiel *et al.*, 2011a; Hoogendoorn *et al.*, 2011; Kim *et al.*, 2014).

En las redes sociales en línea, los usuarios tienen las siguientes capacidades:

1. Publicar comentarios utilizando texto, imágenes y video.
2. Crear y eliminar relaciones de amistad.
3. Navegar entre las publicaciones de sus amigos.
4. Definir parámetros de privacidad.

5. Navegar entre publicaciones de otras personas, siempre y cuando su privacidad lo permita.
6. Recibir publicidad de acuerdo a sus intereses a cambio del servicio gratuito.

Existe evidencia de que en las redes sociales se transmiten emociones de una persona a otra a través de las publicaciones. El fenómeno de transmisión de emociones se conoce como *contagio emocional*; de acuerdo con Hatfield y Cacioppo (1994), el contagio emocional es la tendencia de imitar y sincronizar automáticamente movimientos, posturas y vocalizaciones con los de otra persona y, consecuentemente, converger de manera emocional. Este contagio se da en las redes sociales a partir de la lectura de los comentarios publicados en ellas. Desde un punto de vista psicológico, Hoffman (2002) propone varios mecanismos por los cuales podría iniciarse el contagio emocional. Uno de estos es la asociación a través del lenguaje, en donde descripciones escritas o habladas de una situación provocan que el observador piense en una situación comparable donde reaccionó de una forma similar. La imaginación de dicha situación, de esta forma, induce una experiencia emocional congruente en el observador. Otro mecanismo descrito por Hoffman (2002) propone que el observador se imagine a sí mismo en dicha situación para experimentar así una emoción similar a la descrita.

Recientemente, ha surgido el concepto de *redes sociales afectivas*, las cuales tienen en consideración el estado emocional de sus usuarios con el objetivo de ofrecerles un servicio que se adapte mejor a ellos o intentan inducir un estado emocional más favorable a través, por ejemplo, de recomendaciones de contenidos multimedia (Kim *et al.*, 2014).

2.2.1 TWITTER

Una de las redes sociales en línea más populares hasta el momento es Twitter. El servicio que esta red provee permite a los usuarios publicar mensajes cortos de 140

caracteres máximo. A estos mensajes se les conoce como *tuits* (en inglés, *tweets*). Cada usuario está habilitado para *seguir* a otros usuarios o para *ser seguido* por otros usuarios por lo que según el tipo de relación se tienen dos listas asociadas a cada usuario. A la primer lista se le conoce como *siguiendo* (el término utilizado en inglés es *followers*) y se dice que el usuario *sigue* a los usuarios de la lista. A la segunda lista se le conoce como *seguidores*, el usuario *es seguido* por aquellos que se encuentran en esta lista y él se encuentra en la lista *siguiendo* de ellos. En general, un usuario es *contacto* de otro si el primero se encuentra en cualquiera de estas dos listas del segundo. Los tuits publicados por quienes están en la lista *siguiendo* del usuario van apareciendo en su *línea de tiempo* (en inglés, *timeline*), la cual puede verse como la página personal del usuario en Twitter. El usuario también puede ser *mencionado* en los tuits de otros usuarios, independientemente de si los sigue o no; para hacer una mención, se coloca en el tuit el nombre del usuario (que empieza con “@”, por ejemplo @soyusuario).

2.3 APRENDIZAJE MÁQUINA

A través de las técnicas de *aprendizaje máquina*, las computadoras adquieren la capacidad de modificar su propio comportamiento para resolver problemas. En la mayoría de los casos prácticos, se puede crear un modelo poco preciso y, al permitir que la computadora aprenda, se puede obtener un modelo más general. En la actualidad, el área de aprendizaje máquina es muy vasta, y se relaciona con otras áreas como inteligencia artificial, cómputo evolutivo, control adaptativo, estadística y otras. En ocasiones, a todo este conjunto de áreas se le llama también aprendizaje máquina.

Existen dos vertientes principales para el aprendizaje máquina: el aprendizaje *supervisado* y el aprendizaje *no supervisado*. El aprendizaje supervisado se basa en *ejemplos* o *instancias*. En este proceso, las computadoras cuentan con un *conjunto de entrenamiento*, el cual consta de *vectores de características* que modelan objetos

de forma abstracta y que están cada uno asociado a un valor o *etiqueta* (también conocido como *clase* o *categoría*) mediante una función desconocida; cuando se trata de predecir el valor numérico del vector de características, a la tarea de aprendizaje se le llama *regresión*, y cuando se trata de predecir la clase, se le llama *clasificación* o categorización. Por ejemplo, una computadora puede aprender a conducir un automóvil a partir de ejemplos proporcionados por seres humanos. Pomerleau (1991), dentro de este contexto, asoció imágenes del camino (tomadas a través de una cámara) a las decisiones del conductor sobre acelerar, frenar o girar el volante; de esta manera, la computadora pudo aprender a conducir por sí misma sin que se le especificaran todas las reglas de manejo posibles.

En el aprendizaje supervisado, es necesario evaluar si el modelo generado ha aprendido correctamente. Para este fin, se utiliza un conjunto de datos aparte que no ha sido usado en el entrenamiento, al cual se le denomina *conjunto de prueba*. Con los resultados del modelo sobre este conjunto de prueba, se puede saber si el modelo ha aprendido o, por el contrario, si solo ha memorizado las respuestas correctas. Un método de validación que se utiliza es el método de *k-dobles* o *validación cruzada*, en donde los datos se dividen en k conjuntos, eligiendo uno como conjunto de prueba y los restantes como conjunto de entrenamiento. Este proceso se repite para cada uno de los k conjuntos y se obtiene un promedio de precisión, exhaustividad u otras métricas de error de cada conjunto de prueba.

A diferencia del aprendizaje supervisado, el aprendizaje no supervisado no cuenta con instancias de entrenamiento que tengan una clase o valor determinado a través de una función. La aplicación principal de estos métodos es la de agrupar los datos de entrada en subconjuntos que tengan un significado especial y útil. Por ejemplo, es posible organizar una colección de documentos y agruparlos según los temas que tratan (Xie y Xing, 2013). Este problema de agrupamiento se conoce en inglés como *clustering* (Xu y Wunsch, 2005) y es la principal técnica del aprendizaje no supervisado.

En el área de análisis de sentimiento o minería de opinión, el problema principal es la clasificación de documentos en clases como *positivo*, *negativo* y *neutro*. La clasificación se puede llevar a cabo desde una perspectiva de aprendizaje supervisado (Liu, 2010, 2012; Medhat *et al.*, 2014; Pang *et al.*, 2008), y asimismo existen métodos de agrupamiento para dividir los documentos en positivos y negativos (Li y Liu, 2012). Sin embargo, el método más común para este tipo de problemas sigue siendo el aprendizaje supervisado.

2.3.1 REGRESIÓN LINEAL

La regresión lineal es un método estadístico que busca obtener la línea recta que mejor se adapte a los datos de entrada o conjunto de entrenamiento. Este modelo, en términos matemáticos, puede expresarse como

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \epsilon, \quad (2.1)$$

donde las x_i representan las variables independientes o datos a los que se desea ajustar el modelo lineal, los parámetros b_i son coeficientes que miden la influencia que ejercen cada una de las variables en el modelo final y ϵ es un error de ajuste del modelo. En el error se incluyen todos los factores que afectan al modelo de forma desconocida o que no se pueden controlar. En un modelo de regresión lineal, es deseable que ϵ sea una variable aleatoria y que el valor esperado $E[\epsilon] = 0$. De esta forma tenemos que,

$$\begin{aligned} E[Y|X_1 = x_1, \dots, X_n = x_n] &= l + E[\epsilon|X_1 = x_1, \dots, X_n = x_n] \\ &= l \end{aligned} \quad (2.2)$$

donde $l = b_0 + b_1x_1 + \dots + b_nx_n$. El valor esperado del modelo lineal es igual al valor real de la variable dependiente y . Además de la hipótesis en la que la media del error es igual a 0, se establecen también las siguientes hipótesis:

1. La varianza de ϵ es constante: $\text{Var}(\epsilon|X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \sigma^2$.

2. La distribución de ϵ es normal, media igual a 0 y desviación estándar igual a σ .
3. Los valores del error son independientes para cada uno de los valores de x_i .

La regresión lineal es utilizada para estimar el valor de una variable dependiente y cuando no se tiene suficiente información de las variables independientes. Asimismo, estos modelos con frecuencia se utilizan para predecir valores futuros o para identificar una tendencia. Por ejemplo, se pueden usar para determinar una tendencia acerca del sentimiento de una población con el fin de predecir de forma indirecta los niveles de consumo en el mercado (Koshal *et al.*, 2015).

Uno de los métodos más comunes para estimar los valores de los parámetros b_i es el método de *mínimos cuadrados*. Si definimos la suma del error cuadrático como

$$\text{SSE} = \sum_{i=1}^n \epsilon^2 = \sum_{i=1}^n (y_i - (b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n))^2, \quad (2.3)$$

entonces la solución consiste en encontrar los valores de b_i que minimicen el valor de SSE. Podemos notar que los valores de los parámetros b_i pueden variar dependiendo de la muestra particular con la que se está trabajando, por lo que también pueden considerarse variables aleatorias y con base en esto es posible encontrar un intervalo de confianza para estos valores.

La regresión lineal se ha utilizado tradicionalmente para predecir valores futuros o tendencias de una variable dependiente, por ejemplo, predecir tendencias de consumo eléctrico en una ciudad (Bianco *et al.*, 2009), tráfico vehicular (Sun *et al.*, 2003), la bolsa de valores (Altay y Satman, 2005) y cambios climáticos (Cao *et al.*, 2015).

2.3.2 MÁQUINAS DE SOPORTE VECTORIAL

Las máquinas de soporte vectorial son un método de clasificación, es decir, un método que asigna automáticamente una categoría dado un vector de característi-

cas, el cual representa un objeto; se pueden usar, por ejemplo, para detectar, en una muestra de sangre, la existencia de determinada enfermedad. En este tipo de problemas, las máquinas de soporte vectorial han mostrado tener una precisión competitiva con otras técnicas utilizadas tradicionalmente, como las redes neuronales. Con frecuencia, en el área de cómputo afectivo o análisis de sentimiento, se utilizan para clasificar el sentimiento (positivo o negativo) de textos. La técnica en sí consiste en asociar los vectores de características a un espacio n -dimensional y encontrar un hiperplano que pueda separar los objetos en dos categorías, de forma que un parámetro llamado *margen* (denotado por m) se maximice. El margen es la distancia entre el hiperplano y los objetos de las clases a diferenciar.

Maximizar el parámetro m es un problema de programación cuadrático que se puede resolver mediante el uso de multiplicadores de Lagrange. La máquina de soporte vectorial encuentra el hiperplano óptimo utilizando el producto punto con funciones en el espacio de características llamadas *núcleos* (en inglés, *kernels*). La solución del hiperplano óptimo puede ser escrita como la combinación de unos pocos puntos de entrada que son llamados *vectores de soporte*. La figura 2.2 muestra un ejemplo visual de cómo operan las máquinas de soporte vectorial. Vemos que el problema principal es encontrar el hiperplano de clasificación, que en la figura se representa como una línea recta. Se pueden observar dos clases representadas por puntos circulares y de forma triangular. El parámetro a optimizar m define los márgenes para la clasificación. Las máquinas de soporte vectorial actualmente son utilizadas en clasificación de texto, detección de idiomas, minería de opinión y pronóstico de sentimiento (Mogadala y Varma, 2012; Moraes *et al.*, 2013).

2.4 SERIES DE TIEMPO

Cuando se toman mediciones sobre un fenómeno en tiempos regulares como segundos, minutos y horas, se obtiene una *serie de tiempo* (Durbin y Koopman, 2012). Las series de tiempo se aplican, por ejemplo, si se necesita predecir cambios

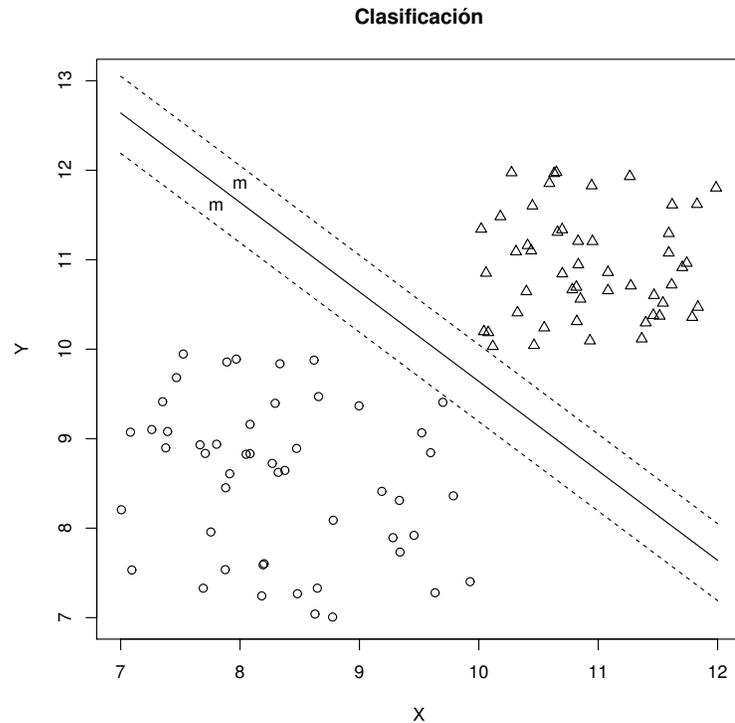


Figura 2.2: Ejemplo de clasificación con máquinas de soporte vectorial. En la gráfica se muestra la recta de separación y los límites para el margen denotado como m . Las clases se representan visualmente a través de círculos y triángulos.

climáticos o atmosféricos con el fin de tomar debidas precauciones (Nordhaus y Moffat, 2017). Se miden también la mortalidad, crecimiento demográfico, tendencias en la propagación de enfermedades y tendencias de crecimiento económico y político (Granger y Hatanaka, 2015; Granger y Newbold, 2014), entre otros. El principal objetivo de mantener un registro de mediciones de estos indicadores es precisamente el de predecir sus valores a futuro. Esta predicción es la que permite tomar decisiones más acertadas en el presente. Partiendo de una serie de tiempo, es posible aplicar métodos estadísticos para obtener un modelo aproximado de los datos mediante el cual se realicen inferencias sobre el futuro (Durbin y Koopman, 2012). El método clásico para el análisis de series de tiempo incluye cuatro componentes básicos:

Tendencia: denotada por T , representa un cambio que se mantiene de forma relativamente constante — por ejemplo, las ventas de una compañía pueden ir a

la baja durante todo el año. La característica principal de la tendencia es que se consideran plazos largos en donde se observa un crecimiento, disminución, o mantenimiento constante de los valores observados.

Fluctuación cíclica: denotada por C , indica la presencia de un ciclo repetitivo — por ejemplo, el horario de mayor tráfico durante la semana se repite constantemente semana tras semana. La fluctuación puede ocurrir en diferentes intervalos de tiempo, como semanas, días y meses.

Variación estacional: denotada por E , es una fluctuación cíclica que se repite con respecto a la tendencia y que también ocurre en plazos de tiempo más prolongados. Por ejemplo, la venta de hospedaje en hoteles en áreas turísticas que tienen temporadas altas o bajas durante el año.

Variación irregular: denotada por I , es un evento que sucede y que no se puede predecir o medir — por ejemplo, desastres naturales como los terremotos que pueden ocurrir sin que nos demos cuenta de que están a punto de suceder.

Teniendo en cuenta lo anterior, podemos escribir una serie de tiempo Y como la suma de estas componentes:

$$Y(t) = T(t) + C(t) + E(t) + I(t), \quad (2.4)$$

donde t representa el tiempo en que fue tomada la medición. A este modelo se le conoce como modelo *aditivo*, debido a que las componentes se suman. De igual forma, existe el modelo *multiplicativo*, en el que las componentes se multiplican, y el modelo *mixto*, en el que existen tanto multiplicaciones como sumas de las componentes.

Las series de tiempo también se pueden clasificar en dos categorías: (1) estacionarias y (2) no estacionarias. Las series de tiempo estacionarias son series en las que existe una fluctuación que permanece constante con respecto al tiempo. Estas series tienen una media y una varianza constantes. Por ejemplo, si midiéramos los niveles del mar cada hora, podríamos notar a lo largo del día que tenemos una fluctuación constante debido a que sus valores oscilan dentro de un rango definido y, por

lo tanto, la serie parece ser igual. Las series no estacionarias son series que tienen una tendencia cambiante con respecto al tiempo, por ejemplo, el ciclo de vida de un producto que al lanzarse puede tener muchas ventas y, posteriormente, comenzar a decrecer.

2.4.0.1 MODELOS ARIMA

En los años 70, Box *et al.* (2015) propusieron un método más sencillo para analizar series de tiempo, el cual utilizaba información solo de la variable dependiente y el tiempo, dejando de lado las variables explicativas. Esto fue de ayuda para los investigadores que tuvieron la oportunidad de analizar los datos sin variables explicativas para crear un modelo. La única información que se utiliza es el pasado histórico de la variable. Una de las desventajas obvias de este método es que se excluye información exógena que podría ser de mucha importancia para el modelado. Por ejemplo, al estudiar la tasa de mortalidad, no se consideran variables como la aparición de epidemias o de nuevas vacunas.

Las siglas *ARIMA* significan en inglés *AutoRegressive Integrated Moving Average*, es decir, *Modelos Autorregresivos Integrados de Medias Móviles*. Un modelo autorregresivo es aquel en el que la variable dependiente es explicada por las observaciones de ella misma en periodos anteriores, añadiendo un término que representa un error. Los modelos autorregresivos se denotan por $AR(p)$, donde p representa el orden del modelo que indica la cantidad de observaciones pasadas consideradas dentro del modelo. En términos matemáticos, una serie de tiempo de tipo $AR(p)$ está dada por:

$$Y(t) = C_0 + C_1Y(t-1) + C_2Y(t-2) + C_3Y(t-2) + \dots + C_pY(t-p) + \epsilon(t). \quad (2.5)$$

El término de error para las series de este tipo se denomina *ruido blanco*, el cual es una serie de tiempo por sí misma que cumple las siguientes características:

- $\mu = 0$,

- σ^2 es una constante,
- $\text{Cov}(X_1, X_2) = 0$ para cualesquiera dos observaciones X_1 y X_2 de la serie X .

Es común emplear valores bajos para p , como AR(1) o AR(2) o, en caso de que la serie muestre alguna periodicidad, en ocasiones es conveniente emplear dicho valor aunque sea mayor. Por ejemplo, se utiliza AR(7) para ciclos que se repiten durante los 7 días de una semana.

Los modelos de medias móviles se denotan por MA(q), donde q es la cantidad de términos de error en el pasado. Estos modelos explican la variable dependiente a partir de un promedio μ y una sucesión de errores ponderados de estimación en el pasado. Es decir, MA(q) está dado por

$$Y(t) = \mu + C_1\epsilon_t + C_2\epsilon_{t-1} + C_3\epsilon_{t-2} + \dots + C_q\epsilon_{t-q}, \quad (2.6)$$

donde las C_i son constantes.

Existe un modelo mixto llamado ARMA(p, q), el cual es una combinación de los modelos AR(p) y MA(q). El modelo toma en cuenta los valores pasados de la variable y de los errores que se hayan cometido en su estimación.

Los modelos AR(p) y MA(q) se aplican en situaciones en las que la serie de tiempo es estacionaria. Sin embargo, existen situaciones en la vida cotidiana en las que esta característica no se cumple. Cuando es así, las series *no estacionarias homogéneas* tienen la cualidad de transformarse en series estacionarias a través de una diferenciación, y posteriormente se pueden aplicar los modelos AR(p) y MA(q). En términos matemáticos, si la serie Y_t no es estacionaria y

$$\begin{aligned} \Delta Y_t &= Y_t - Y_{t-1}, \\ \Delta^{n+1} Y_t &= \Delta^n Y_t - \Delta^n Y_{t-1} \end{aligned} \quad (2.7)$$

corresponde a una serie ARMA(p, q), decimos que la serie Y_t sigue un proceso ARIMA(p, d, q), donde d se conoce como el orden de integración de la serie.

2.5 MÉTRICAS DE EVALUACIÓN

Existen diferentes métricas para evaluar la calidad obtenida por los métodos que resuelven problemas de clasificación y regresión. Las métricas utilizadas en esta tesis se describen a continuación.

2.5.1 CLASIFICACIÓN

La manera convencional de evaluar los resultados de la clasificación consiste en calcular *precisión*, *exhaustividad* y *valor-F* (F_1). El objetivo de la precisión, por un lado, consiste en evaluar la pertenencia de los elementos clasificados (“son todos los que están”) y se calcula como

$$p = \frac{vp}{vp + fp}, \quad (2.8)$$

donde vp es la cantidad de elementos correctamente clasificados (verdaderos positivos) y fp es la cantidad de elementos que se dijo pertenecían a la clase y realmente no pertenecían (falsos positivos).

El objetivo de la exhaustividad, en contraparte, consiste en evaluar la completez (“están todos los que son”) y se calcula como

$$e = \frac{vp}{vp + fn}, \quad (2.9)$$

donde fn es la cantidad de elementos que se dijo no pertenecían a la clase y realmente sí pertenecían (falsos negativos).

Por último, el valor-F (F_1) integra ambas caras de la moneda al obtener la media armónica entre precisión y exhaustividad, dando para ello la misma importancia a ambas:

$$F_1 = 2 \cdot \left(\frac{p \cdot e}{p + e} \right). \quad (2.10)$$

2.5.2 REGRESIÓN

Debido a que nuestro problema de predicción se puede abordar con técnicas de regresión, discutimos métricas de evaluación para esta tarea. Para medir el parecido de la predicción con la realidad, existen métricas utilizadas frecuentemente en el área de minería de datos. Algunas de estas métricas son la *correlación* y el *error absoluto medio* (Chai y Draxler, 2014; Willmott y Matsuura, 2005), que en inglés se conoce como *mean absolute error*. Una correlación indica la dependencia lineal que tienen una o más variables dadas (Crawford, 2006); es decir, si la correlación es igual a uno podemos afirmar que la variable de salida o dependiente se puede representar como una combinación lineal de las variables de entrada o independientes.

El error absoluto medio indica qué tan diferentes son los valores predichos por el modelo con respecto a los reales. El valor ideal para esta métrica es cero, lo cual quiere decir que el valor real se ha predicho con exactitud. Denotamos la correlación con las siglas COR y el error absoluto medio como MAE. En términos matemáticos, si tenemos una colección de n predicciones x_1, \dots, x_n y n observaciones y_1, \dots, y_n y definimos $\epsilon_i = y_i - x_i$ como el error para la observación i ,

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\epsilon_i|, \quad (2.11)$$

y

$$\text{COR} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (2.12)$$

Cuando $\text{MAE} = 0$, implica que $|y_i - x_i| = 0$ y por lo tanto $x_i = y_i$. Al sustituir este valor en la fórmula para la correlación, obtenemos $\text{COR} = 1$. Otras métricas para comparar modelos de predicción o regresión son la raíz del error cuadrático medio (RMSE), el error absoluto relativo (RAE) y la raíz del error relativo cuadrático (RRSE). Estas métricas se encuentran descritas en las ecuaciones 2.13, 2.14 y 2.15:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \epsilon_i^2}, \quad (2.13)$$

$$\text{RAE} = \frac{\sum_{i=1}^n |\epsilon_i|}{\sum_{i=1}^n |\bar{y}_i - y_i|}, \quad (2.14)$$

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^n \epsilon_i^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2}}. \quad (2.15)$$

Las métricas anteriormente descritas indican magnitud del error generado por el modelo. Como en el caso de MAE, el valor ideal es de cero, por lo que a mayor error, más grande será el valor generado por estas métricas.

2.6 ACUERDO DE INTERCORRELACIÓN

Al estar haciendo etiquetado manual para generar instancias de entrenamiento, existe la dificultad de escoger una etiqueta global que represente coherentemente el voto colectivo de los evaluadores (por ejemplo, una opción es escoger la etiqueta votada por la mayoría). Este problema es de especial importancia cuando se desea emular el etiquetado de un experto. Para abordar este problema, Qing *et al.* (2014) proponen un método llamado *acuerdo de intercorrelación* (en inglés, *intercorrelation agreement*) para seleccionar la etiqueta más apropiada para un comentario a partir de las evaluaciones de los usuarios. La calidad (o *confiabilidad*) de un evaluador i , denotada por ω_i , está determinada por

$$\omega_i = \log \frac{(|K| - 1) \cdot \rho_i}{1 - \rho_i}, \quad (2.16)$$

donde ρ_i representa la precisión que tiene el evaluador i al elegir la etiqueta correcta para el comentario y K representa el conjunto de etiquetas que pueden ser asignadas al comentario. El valor de ρ_i debe ser estimado para cada evaluador; Qing *et al.* (2014), para este fin, proponen

$$\rho_i = \frac{|j \in J : a_{ij} = \text{SPR}(A)_j| + 0.5}{|j \in J : i \text{ etiqueta a } j| + 1}, \quad (2.17)$$

donde j es un comentario del conjunto J que se va a etiquetar, A es el conjunto de todas las asignaciones realizadas por los evaluadores y a_{ij} es la etiqueta seleccionada por el evaluador i para j . El operador SPR hace referencia a un método llamado

regla de pluralidad simple (en inglés, *simple plurality rule*), en el que la etiqueta elegida es simplemente la más común asignada por los evaluadores. Por lo tanto, $\text{SPR}(A)_j$ representa la etiqueta más elegida para el comentario j . De esta forma, la precisión del evaluador está dada por la razón de etiquetas que asigna a cada comentario y que es igual a la que fue asignada con mayor frecuencia entre el total de sus evaluaciones. Es decir, entre más esté de acuerdo con los demás usuarios, mayor será su confiabilidad y, por lo tanto, sus votos contarán más en la selección final de la etiqueta. Finalmente, cada evaluador asigna una etiqueta con un peso ω_i y la etiqueta que acumule una mayor puntuación se elige como etiqueta definitiva.

2.7 RESUMEN

En este capítulo, exploramos brevemente las teorías principales acerca de las emociones en los seres humanos y cómo estas han pasado a expresarse también en las redes sociales en línea. Esto ha posibilitado el estudio de las emociones a nivel masivo mediante la extracción de comentarios de usuarios que se encuentran en ciudades enteras. Asimismo, explicamos brevemente las técnicas de aprendizaje máquina que se utilizan en el desarrollo de esta tesis, tales como las máquinas de soporte vectorial, las series de tiempo (ARIMA) y la regresión lineal. Para la creación de clasificadores automáticos mediante técnicas de aprendizaje supervisado, es necesario contar con un repositorio de comentarios clasificado (etiquetado) previamente. Con tal fin, hemos recurrido al trabajo de Qing *et al.* (2014) para obtener las etiquetas de un comentario a partir de “votos” realizados por usuarios no expertos. Vimos las métricas más utilizadas para comparar modelos de regresión, siendo uno de los más importantes el error absoluto medio o MAE.

CAPÍTULO 3

ESTADO DEL ARTE

En este capítulo se exponen los principales trabajos que se han llevado a cabo con el fin de predecir emociones. Veremos cómo están agrupados y las principales tendencias, enfocándonos con mayor detalle en los trabajos que están directamente relacionados con el modelo propuesto. El capítulo se divide en dos partes principales:

1. Predicción de emociones (sección 3.1). Se explica de forma general el proceso que con mayor frecuencia se sigue en la predicción de emociones y tareas relacionadas.
2. Modelos de predicción relacionados a este trabajo (sección 3.2). Se explican los trabajos que se relacionan de una forma más directa con el modelo propuesto y cuáles son las diferencias más importantes con respecto a este modelo.

3.1 PREDICCIÓN, REPRESENTACIÓN, DETECCIÓN Y MEDICIÓN DE EMOCIONES

La predicción de emociones se asemeja al proceso de pronóstico o predicción de series de tiempo, en donde se cuenta con información histórica de una o más

variables relacionadas para predecir el valor futuro de una variable objetivo (Box *et al.*, 2015; Granger y Newbold, 2014). Se compone de tres procesos importantes: (1) representación, detección y medición de la emoción, (2) almacenamiento de datos históricos sobre las mediciones y (3) predicción de la emoción mediante un modelo. En las siguientes subsecciones se describe con mayor detalle cada uno de estos procesos.

Antes de detectar una emoción, es necesario elegir una forma de representarla, de manera que una computadora pueda trabajar con ella. Posteriormente, se procede a detectar la emoción — es decir, a reconocerla dentro del repositorio utilizado. Esto permitirá, más adelante, llevar a cabo la predicción. A continuación se explican los procesos de representación y detección.

3.1.1 REPRESENTACIÓN DE EMOCIONES

Al trabajar con emociones mediante herramientas computacionales, se necesita una forma de representar las emociones de manera que puedan ser procesadas mediante alguna técnica de aprendizaje máquina, análisis numérico, método probabilista u otro método. En la actualidad, existen varias representaciones. Estas se pueden agrupar en dos áreas principales: (1) categorías o etiquetas y (2) vectores multidimensionales.

En la representación por categorías, la detección se centra en identificar la existencia o ausencia de ciertas emociones, tales como emociones básicas (por ejemplo, alegría, enojo y miedo), emociones complejas (por ejemplo, celos, culpa y orgullo) o emociones abstractas (por ejemplo, positivas, negativas o neutras). Una cantidad considerable de trabajos utiliza emociones abstractas como parte de un mismo conjunto continuo o discreto de números, en donde los valores más bajos corresponden a una emoción negativa, los números que se encuentran al centro de la escala corresponden a una emoción neutra y los valores más altos corresponden a una emoción

positiva (Nguyen *et al.*, 2012; Zhang *et al.*, 2010). Con frecuencia, dicho conjunto de números está centrado en el cero con el fin de facilitar operaciones de suma y resta de emociones (conveniencia matemática). Esta representación es común en trabajos actuales debido a su simplicidad. Asimismo, las categorías o etiquetas (emociones básicas o complejas) también pueden tener una representación numérica mediante la cual se indique su intensidad (es decir, un rango) o pueden representarse mediante un conjunto de valores discretos que indiquen la presencia (uno) o ausencia (cero) de dicha emoción.

En el tipo de representación vectorial, cada una de las componentes del vector tiene una interpretación concreta acerca de la emoción o situación. Se han propuesto hasta ahora representaciones que utilizan dos, tres o hasta cuatro componentes en los vectores. WordNet Affect (Strapparava *et al.*, 2004) es un diccionario de palabras en inglés que asigna vectores con dos componentes: valencia y excitación. La valencia indica qué tan positiva o negativa es la emoción, mientras que la excitación indica qué tan tensa o relajada hace sentir a la persona. Mehrabian (1996) propone el uso de tres componentes: placer, excitación y dominación. La dominación hace referencia a la sensación de la persona de tener el control e influir a su alrededor.

Cambria *et al.* (2012) proponen el reloj de arena de las emociones (en inglés, *Hour Glass*) en el área de cómputo afectivo y particularmente en el área de interfaces humano-computadora. Utilizan vectores de números reales de cuatro dimensiones: placer, atención, sensibilidad y aptitud. El placer indica qué tanto se encuentra la persona disfrutando el momento, la atención indica qué tanto se muestra interesada en la actividad o información que se le muestra, la sensibilidad indica qué tan cómoda se encuentra con la interfaz y la aptitud indica qué tan hábil es con la interfaz. La figura 3.1 resume y organiza los tipos de representación para las emociones. El trabajo propuesto, por su parte, se fundamenta en las investigaciones de Plutchik (2001), quien propone un modelo para representar las emociones basándose en las *similitudes* entre ellas.

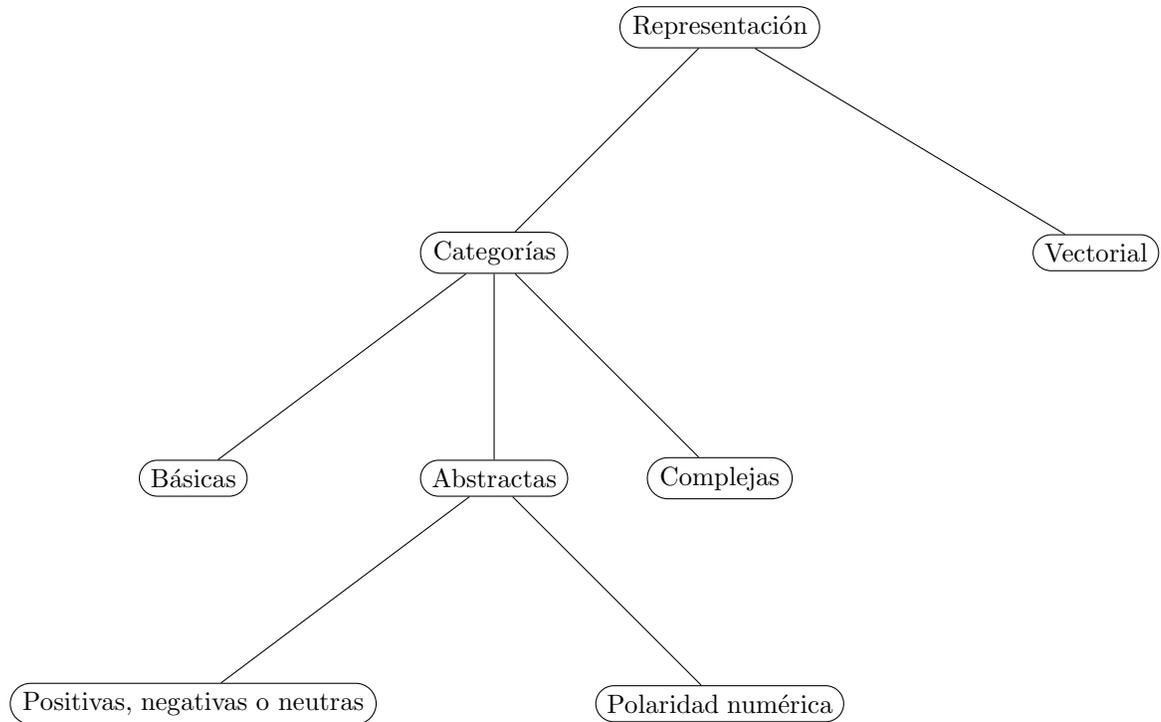


Figura 3.1: Representación de emociones. Se muestran las representaciones de emociones que se utilizan con mayor frecuencia actualmente desde un punto de vista computacional.

3.1.2 DETECCIÓN Y MEDICIÓN DE EMOCIONES

La detección de la emoción es el proceso por el que se reconoce la emoción que se está observando en un determinado momento. Para que una emoción pueda ser detectada, debe existir un medio en el cual se transmite; uno de los medios más estudiados actualmente es el *texto*, debido a la explosión que han tenido las redes sociales. Sin embargo, existen áreas de estudio como el análisis de sentimiento y el cómputo afectivo que estudian otros medios, como fotografías, música, imágenes de resonancia magnética, video y voz, por mencionar algunos. También puede ser la misma persona quien reporte sus cambios de emoción para el análisis. Otros estudios recientes analizan el tacto como una forma de transmitir emociones (por ejemplo, mediante el vibrador de un celular). La detección de la emoción como tal no está limitada a un individuo, sino que puede ser a nivel grupo, ciudad o hasta país. En las siguientes secciones se describen los métodos más comunes de detección

de emociones para diferentes medios de transmisión; estos métodos se resumen en la figura 3.2.

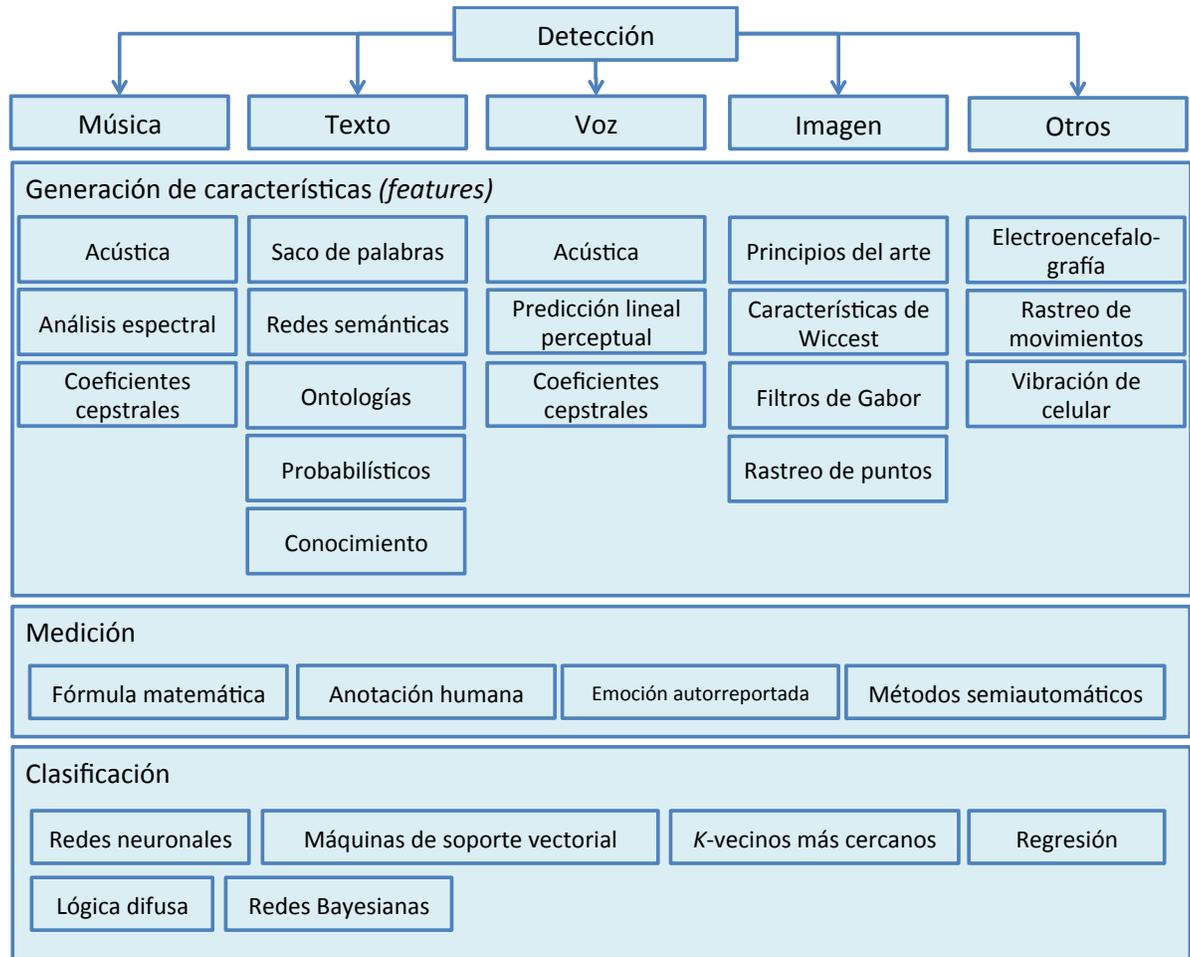


Figura 3.2: Detección de emociones. Se muestra un resumen de herramientas o métodos utilizados frecuentemente para detectar emociones en diferentes contextos, tales como texto, música y voz.

3.1.2.1 TEXTO

Uno de los métodos más simples para detectar emociones en texto es el paradigma de saco de palabras — en inglés, *bag of words* (Wilks, 1997) — el cual divide el texto en unidades simples que pueden ser palabras, n -gramas o frases (Cavnar *et al.*, 1994; Kennedy y Inkpen, 2006) y agrupaciones de palabras por su raíz etimológica

(Moral *et al.*, 2014). A cada una de estas unidades le es asignada una valencia o peso indicando qué tan positiva o negativa es la emoción presente. Los primeros trabajos relacionados con análisis de sentimiento en redes sociales utilizaron este método dividiendo los comentarios en palabras simples (Moreno-Ortiz y Hernández, 2013; Pang *et al.*, 2002; Taboada *et al.*, 2011; Turney, 2002). SentiWordNet (Baccianella *et al.*, 2010) y ANEW (Nielsen, 2011) son repositorios de palabras con anotaciones de emociones que utilizan este enfoque. Al final, el sentimiento (positivo o negativo) está dado por la suma de todas las valencias o pesos que aparecen en el texto.

Otros trabajos incluyen información relacionada con la estructura del texto, como las frases y los modificadores de valencia, los cuales sirven para potenciar o disminuir la valencia de una palabra — por ejemplo “muy bonito” (Kennedy y Inkpen, 2006; Ortiz *et al.*, 2010a,b; Wilson *et al.*, 2005). Una palabra que existe dentro de una frase puede tener una valencia (peso) diferente si se encuentra aislada. Existe un área dentro del análisis de sentimiento dedicada a la detección de emociones basada en diccionarios de palabras (o *lexicones*), a los cuales se les ha asignado una valencia previamente (Bosco *et al.*, 2013; Taboada *et al.*, 2011). Un resumen de los últimos años en cuanto al análisis de sentimiento se puede encontrar en el trabajo de Medhat *et al.* (2014).

Más recientemente, se ha extendido el uso de ontologías para analizar el sentimiento utilizando conocimiento de objetos y sus atributos (Balahur *et al.*, 2011; Baldoni *et al.*, 2012; Borth *et al.*, 2013; Chen *et al.*, 2016; Fathy *et al.*, 2017; Grassi, 2009; Kontopoulos *et al.*, 2013; Zhang *et al.*, 2013). Por ejemplo, para comparar el sentimiento público hacia un celular con el de otro, se pueden comparar también los sentimientos con respecto a las características de estos (por ejemplo, batería y cámara). Uno de los repositorios de conocimiento más amplios en estos momentos es ConceptNet (Liu y Singh, 2004), el cual fue creado por el proyecto Open Mind Common Sense (Singh *et al.*, 2002), cuyo objetivo es crear una base de conocimiento compartida y creada por el público en general a través del Internet. Dicho proyecto

pertenece al Massachusetts Institute of Technology (MIT)¹.

Existen trabajos en los que se ha hecho el esfuerzo de combinar información proveniente de ConceptNet y WordNet Affect (Strapparava *et al.*, 2004), que es un recurso que agrega información sobre emociones a un conjunto de palabras extraídas de WordNet (Miller, 1995). AffectNET (Cambria y Hussain, 2012) es el producto resultante de combinar estas dos herramientas. Actualmente, SenticNET (Cambria y Hussain, 2015; Cambria *et al.*, 2016, 2010) es una de las bases de conocimiento más avanzadas que se ha ido mejorando en los últimos años. Utilizando un conjunto de resúmenes o comentarios sobre películas de cine — anotados previamente como positivos, negativos o neutros por Pang y Lee (2005)—, los trabajos con SenticNET en la versión más reciente arrojan una precisión de 90.1 % considerando solo sentimiento positivo o negativo.

3.1.2.2 MÚSICA

Los trabajos con música se basan en la obtención de características a partir de propiedades acústicas como el ritmo, el timbre, niveles de intensidad (volumen) y otras propiedades que corresponden a un análisis espectral, tales como flujo espectral, ancho de banda, centroide y los coeficientes cepstrales — los cuales resultan de dividir la música en fragmentos y aplicar métodos matemáticos como la transformada de Fourier y la escala de Mel (Kim *et al.*, 2010; Li y Ogihara, 2004; Liu *et al.*, 2003; Lu *et al.*, 2006; Yang y Chen, 2012). A partir de estas características, mediante técnicas de aprendizaje máquina, se puede obtener una emoción representada, ya sea en forma de categoría o de vector (Han *et al.*, 2010; Yang *et al.*, 2008). También se pueden utilizar combinaciones de características de audio y texto para hacer la detección de emociones, como en el caso de las letras musicales de canciones (Hu y Downie, 2010). Existe un área llamada *recuperación de información en música* (en inglés, *music information retrieval*) que se encarga de estudiar las características más útiles

¹Véase <https://www.media.mit.edu/research/groups/5994/open-mind-common-sense>.

para analizar música e incluso poder hacer comparaciones entre diferentes canciones o sonidos para encontrar similitudes; esto último tiene mucha aplicación para crear sistemas de recomendación de música (Schedl *et al.*, 2014; Sturm, 2014; Van den Oord *et al.*, 2013).

3.1.2.3 FOTOGRAFÍAS O IMÁGENES

El uso de imágenes se ha popularizado en las redes sociales. Existe ahora un área emergente llamada *clasificación de imágenes afectivas* (en inglés, *affective image classification*), la cual se centra en la extracción de emociones a partir de imágenes (Wang *et al.*, 2013a,b; Zhao *et al.*, 2014b). Incluso hay trabajos en los que se busca predecir la respuesta emocional de un usuario ante la percepción de una imagen emotiva (Chen *et al.*, 2014). Para encontrar emociones en fotografías, existen dos modalidades principales: (1) detectar la emoción que expresan las personas que aparecen en la misma y (2) detectar la emoción de la imagen como un todo, por ejemplo paisajes y figuras que evocan emociones (You *et al.*, 2015; Zhao *et al.*, 2016).

Al analizar la imagen para buscar las emociones que expresan las personas que aparecen en ellas, se siguen comúnmente varios pasos: (1) detectar las caras de los participantes, (2) detectar los cuerpos de los participantes, (3) obtener características de las caras y cuerpos para procesarles por técnicas de aprendizaje máquina y (4) hacer la clasificación o estimación de la emoción (Mou *et al.*, 2015; Patwardhan y Knapp, 2016a,b).

El problema de detectar caras en una fotografía ha sido estudiado desde los años 60. En el trabajo de Zafeiriou *et al.* (2015), se explican avances recientes y parte de su historia. Una vez reconocidas las caras en la imagen, se procede a evaluar sus características, como simetría, distancia entre puntos importantes, labios levantados, posición de la nariz y posición de las cejas — junto con otras características que sirven para clasificar la emoción como tal (Sariyanidi *et al.*, 2015). De igual forma,

existen trabajos en los que se considera la detección de cuerpo, manos, piernas, brazos, torso y movimiento para reconocer emociones en las personas (Kleinsmith y Bianchi-Berthouze, 2013; Patwardhan y Knapp, 2016b; Stathopoulou y Tsihrintzis, 2011).

Al considerar la imagen como un todo, se analizan propiedades tales como color, textura e iluminación, entre otras. Comúnmente se utilizan filtros de Wiccest (Geusebroek, 2006) y Gabor (Bovik *et al.*, 1990) para obtener propiedades de las imágenes y utilizarlas como características durante el entrenamiento con técnicas de aprendizaje máquina (Yanulevskaya *et al.*, 2008). Los filtros de Wiccest analizan sombras, degradado de colores e iluminación. Los filtros de Gabor se utilizan para obtener propiedades de las texturas en la imagen y de patrones regulares que se repiten en determinadas direcciones dentro de una región de la imagen. Otros trabajos utilizan los principios del arte (en inglés, *principles of art*), que incluyen balance, énfasis, armonía, variedad, graduación y movimiento. En el trabajo de Zhao *et al.* (2014a), se explican con más detalle cada una de las características obtenidas mediante los filtros de Gabor.

Entre los repositorios de imágenes disponibles para experimentación, se encuentra el Sistema Internacional de Imágenes Afectivas (en inglés, *International Affective Picture System* o IAPS) por Lang *et al.* (1997). Este repositorio cuenta con 1,182 imágenes con información afectiva representada como vector de tres componentes: placer, excitación y dominación. Otros repositorios importantes son el LFW por sus siglas en inglés de *Labeled Faces in the Wild* (Huang *et al.*, 2007) y el AFEW (Dhall *et al.*, 2011), por sus siglas en inglés de *Acted Facial Expressions in the Wild*. El primero contiene imágenes con caras etiquetadas con el nombre de la persona y el segundo contiene pequeños videos en los que aparecen actores de películas imitando alguna emoción humana. Cada pequeño video está etiquetado con dicha emoción utilizando siete emociones básicas: alegría, enojo, miedo, tristeza, repulsión, sorpresa y neutral.

El proyecto llamado *El Reto de Reconocimiento de Emociones en la Naturaleza* (en inglés, *Emotion Recognition in the Wild Challenge*)² (Dhall *et al.*, 2016), se lleva a cabo anualmente haciendo la invitación a investigadores de todo el mundo para que mejoren las técnicas de reconocimiento de emociones. Existe también en la Universidad de Amsterdam otro reto similar llamado AVEC (por sus siglas en inglés, *Audio/Visual Emotion Challenge*), el cual es un evento que se lleva a cabo anualmente y se enfoca en la detección de depresión y emociones contando con dos subcategorías para cada tipo; al inscribirse, los investigadores reciben un repositorio de datos con imágenes y clips de video para comenzar sus experimentos (Valstar *et al.*, 2016). Algunos trabajos han utilizado este repositorio para procesar las imágenes con el fin de detectar depresión (Pérez Espinosa *et al.*, 2014). Dado que hoy en día los usuarios de redes sociales como Facebook e Instagram comparten muchas fotografías, una aplicación del trabajo anteriormente citado sería la de encontrar usuarios que hayan expresado depresión en tiempos prolongados y poder tomar una acción que los ayude a superar el problema.

ChaLearn³ es una organización que lanza retos en el área de aprendizaje máquina. Los ganadores de los retos pueden obtener premios en efectivo, lo cual hace aún más atractivos estos retos para diversos competidores. Entre las áreas en las que se puede competir se encuentran: detección de gestos y posturas, clasificación de expresiones reales vs. imitadas, diseño de técnicas de aprendizaje máquina que no necesiten intervención humana para comenzar a aprender, etcétera. Wan *et al.* (2017) hacen un resumen de los resultados de la competencia más reciente que se llevó a cabo en 2017.

²Disponible en: <https://sites.google.com/site/emotiwchallenge/>.

³Disponible en <http://www.chalearn.org/>.

3.1.2.4 VOZ

La detección de emociones en voz involucra (1) preprocesamiento de señales, (2) extracción de características, (3) reconocimiento de patrones y (4) clasificación de la emoción (Dai *et al.*, 2015). En el preprocesamiento, la señal se transforma de manera que sea más fácil extraer las características acústicas. Algunas de las propiedades de la voz que se procesan al comienzo son: tono, armónicos, tracto vocal y energía, entre otras (Ververidis y Kotropoulos, 2006). Otras propiedades que se procesan para obtener características son: frecuencia, volumen, coeficientes cepstrales, inestabilidad (*jitter*), brillo (*shimmer*), relación ruido-armonía (Teixeira *et al.*, 2013), pendiente espectral y flujo espectral (Ringeval *et al.*, 2014).

Mediante un análisis de estas propiedades de la voz, se extraen las características para ser procesadas por herramientas de aprendizaje máquina, como k -vecinos más cercanos, máquinas de soporte vectorial, redes neuronales (Ververidis y Kotropoulos, 2006), modelos de Markov (Nwe *et al.*, 2003) y árboles de decisión (Yacoub *et al.*, 2003), entre otras. El Ayadi *et al.* (2011) y Ringeval *et al.* (2014) profundizan más sobre las diferentes características y métodos de clasificación de emociones en voz. El reto de *Reconocimiento de Emociones en la Naturaleza* (Dhall *et al.*, 2016) incluye, en la sección de voz y video, un reto que se lleva a cabo anualmente⁴.

3.1.2.5 VIDEO

El procesamiento de video para la detección de emociones es el más complejo, ya que utiliza técnicas de detección de emociones para música, imágenes y voz al mismo tiempo. Por ejemplo, Ringeval *et al.* (2014) emplean detección de movimiento en los labios a la vez que analizan la voz. Una de las características que se tiene que tomar en cuenta en los videos es el tiempo. Las caras y voces que aparecen en el video van cambiando y, por lo tanto, hacer un rastreo de estos cambios ayuda en la

⁴Disponible en: <https://sites.google.com/site/emotiwchallenge/>.

predicción de emociones (Cohen *et al.*, 2003; Michel y El Kaliouby, 2003; Noroozi *et al.*, 2017; Zhalehpour *et al.*, 2016). En trabajos recientes, las redes neuronales recurrentes están siendo utilizadas con buenos resultados (Ebrahimi Kahou *et al.*, 2015; Kahou *et al.*, 2013; Khorrami *et al.*, 2016). El reto de *Reconocimiento de Emociones en la Naturaleza* (Dhall *et al.*, 2016) incluye también un apartado para identificar emociones en video. Los videos están etiquetados con emociones y tienen voz, música e imagen. Asimismo, el ya mencionado reto AVEC (Valstar *et al.*, 2016) tiene una subcategoría para reconocer emociones en videos.

3.1.3 DATOS HISTÓRICOS Y VENTANA DE TIEMPO

El almacenamiento de datos históricos es un proceso en el que se toman mediciones relacionadas con la emoción que se desea predecir. Estas mediciones se han llevado a cabo tradicionalmente en ventanas de tiempo definidas e invariables — por ejemplo, minutos (Zhao *et al.*, 2015), horas (Nguyen *et al.*, 2012), días (Hu *et al.*, 2015) y semanas (Li *et al.*, 2015), entre otras. En algunos trabajos, la ventana de tiempo utilizada no necesariamente coincide con un intervalo de tiempo definido, sino que utilizan eventos puntuales o sucesos como ventana para el cambio en las mediciones. Por ejemplo, en un microblog, cada vez que un usuario publica un comentario puede significar un cambio en la emoción actual y, por lo tanto, requiere de una nueva medición (Mogadala y Varma, 2012). Estas mediciones secuenciales que se llevan a cabo son almacenadas para posteriormente ser procesadas en los métodos de predicción. Las mediciones no se limitan simplemente a datos de la emoción como tal, sino que pueden incluir otro tipo de variables, como temperatura ambiental, clima, actividad física, humedad atmosférica (Hannak *et al.*, 2012), emociones de otros (por ejemplo, amigos, familiares o ciudades vecinas), contenido publicado en redes sociales, conversaciones telefónicas y noticias relevantes, por mencionar algunas.

La mayoría de los trabajos actuales se enfocan principalmente en información endógena de la red social, que incluye las emociones de los amigos, la cantidad de

amigos y el comportamiento (Hu *et al.*, 2015; Li *et al.*, 2015; Mogadala y Varma, 2012; Nguyen *et al.*, 2012; Zhao *et al.*, 2015). También se pueden incluir ciertos indicadores como la edad, sexo, historial clínico, consumo de sustancias y día de la semana actual. La cantidad de información que se desee almacenar para hacer la predicción depende del contexto en donde se lleve a cabo dicha actividad. Entre más información histórica se tenga, mejorarán los resultados, pero solo hasta un cierto punto en donde el agregar más información no genera ningún beneficio o — al contrario — perjudica la precisión (Nguyen *et al.*, 2012). Por ejemplo, Nguyen *et al.* (2012) obtuvieron en sus experimentos que, a partir de cuatro horas históricas, la precisión comienza a decaer, y Hu *et al.* (2015) realizando mediciones semanalmente, obtuvieron los mejores resultados con dos meses de mediciones. En este último experimento, la comparación se hizo con hasta tres meses de mediciones históricas.

3.1.4 PREDICCIÓN DE EMOCIONES

Finalmente, en la predicción de emociones, se generan modelos a partir de los valores históricos para estimar el valor de una ventana de tiempo futura, según la representación utilizada para la emoción. Algunos de los métodos más comunes en la actualidad para este fin son: (1) métodos de aprendizaje máquina; (2) regresión lineal, no lineal, univariable o multivariable; (3) series de tiempo y (4) funciones matemáticas o estadísticas. La figura 3.3 muestra un resumen de las etapas en la predicción y los enfoques principales. En general, los trabajos actuales llevan a cabo mediciones y predicciones de emociones utilizando dos enfoques: grupal o individual. A nivel grupo, la medición de la emoción requiere algún tipo de agregación para ser procesada y la predicción también será llevada a cabo de la misma manera (Nguyen *et al.*, 2012). La elección depende del problema que se esté atacando. Asimismo, cuando la predicción es numérica, los trabajos actuales predicen la emoción de dos maneras: a través de predecir un cambio (aumento o disminución) o predecir el valor como tal. Cuando predicen un cambio, se enfocan en determinar si el nivel de

positividad (o negatividad) cambia con el tiempo — es decir, si la emoción se vuelve más positiva o más negativa (Mogadala y Varma, 2012; Nguyen *et al.*, 2012).

Entre los métodos de aprendizaje máquina, el más común al momento de la escritura de esta tesis emplea máquinas de soporte vectorial; esto, debido a su especialidad para hacer clasificaciones basadas en características. Por lo tanto, la selección de características es tan importante como la aplicación del método en sí. Las series de tiempo son utilizadas con menos frecuencia (hasta ahora) para predecir emociones. O'Connor *et al.* (2010), de cara a las elecciones presidenciales del 2009, usan el método del promedio móvil para pronosticar la emoción con respecto a un político. Determinan la emoción en los comentarios mediante lexicones y agrupan los valores empleando todos los comentarios publicados en ventanas de tiempo de un día. El periodo histórico empleado consta de 30 días y predicen el día siguiente con una correlación de 75 %.

Nguyen *et al.* (2012) utilizan una razón r para determinar la emoción en una ventana de tiempo dada mediante la fórmula

$$r = \frac{P}{P + N}, \quad (3.1)$$

donde P y N representan la cantidad de comentarios positivos o negativos respectivamente en la ventana de tiempo. Los autores aplican la fórmula para todos los comentarios publicados en dicha ventana de tiempo sin importar los usuarios que los hayan publicado. También proponen que, para hacer una predicción de emociones, es necesario considerar tres parámetros importantes: (1) cantidad datos históricos, (2) tiempo de respuesta en el futuro y (3) duración del efecto en la respuesta. La cantidad de datos históricos es la cantidad de información utilizada para la predicción; como ya hemos mencionado en la sección anterior, estos datos son necesarios. El tiempo de respuesta en el futuro hace referencia a que un acontecimiento no necesariamente impacta en el tiempo inmediato siguiente, sino que puede tardar un cierto periodo de tiempo antes de comenzar el efecto. La duración del efecto es el tiempo que un evento en particular continúa generando emociones en el usuario. Por ejemplo, una discusión puede generar enojo durante varias semanas. Asimismo,

Nguyen *et al.* (2012) utilizan máquinas de soporte vectorial para hacer la predicción de un cambio emocional y las características empleadas por ellos se muestran en el cuadro 3.1. Se obtuvo un nivel de precisión de 85 % al predecir cambios con la ecuación 3.1.

Mogadala y Varma (2012), valiéndose de regresión lineal con máquinas de soporte vectorial, predicen la transición en las emociones de los usuarios. Las características consideradas son: palabras emotivas (lexicón), emoticones, signos de interrogación y de admiración, retuiteos y menciones. Entre cada comentario miden la diferencia en la emoción y con los datos resultantes llevan a cabo la regresión. Reportan un MAE = 1.932 como mejor resultado en sus experimentos. Zhao *et al.* (2015) miden el estrés utilizando cuatro niveles: *nulo*, *liviano*, *moderado* y *fuerte*. Emplean un método logístico comparándolo con máquinas de soporte vectorial para predecir el nivel de estrés de un solo usuario. Las características consideradas son similares a las de Nguyen *et al.* (2012): comentarios de amigos y respuestas del usuario a los mismos, cantidad de comentarios gustados (en inglés, *liked*), cantidad de comentarios reenviados, comentarios que incluyen frases como “no te preocupes” o “todo estará bien”, medidas de inactividad del usuario ante comentarios de sus amigos y tasa de contestación a comentarios. Se detecta el estrés de los usuarios con una precisión de 76 % (valor-F = 72 %). Kajiwara *et al.* (2017) utilizaron datos médicos de la persona como características; entre estos datos se encuentran: peso, grasa corporal, edad, masa muscular y presión sanguínea máxima, entre otras. En su trabajo detectan emociones abstractas (positivas y negativas). En sus experimentos obtuvieron una precisión promedio de 72 % (valor-F = 72 %).

Li *et al.* (2015) predicen el estrés de usuarios adolescentes utilizando análisis de series de tiempo, en concreto el método SVARIMA (promedio móvil integrado autorregresivo vectorial estacional), el cual considera una serie de tiempo (los niveles de estrés medidos a lo largo de varias semanas) y varias funciones extraídas a partir de los mismos datos, como promedio, mínimo, máximo, cantidad de comentarios que revelan estrés, cantidad total de comentarios y razón de comentarios con estrés sobre

Cuadro 3.1: Características empleadas por Nguyen *et al.* (2012).

Tipo	Descripción
Comentarios	Sentimiento del comentario
	Cantidad de comentarios marcados como favoritos
	Cantidad de retuits
Usuario	Cantidad de usuarios que sigue
	Cantidad de seguidores
	Cantidad de comentarios publicados
	Cantidad de listas a las que pertenece el usuario
Razón de sentimiento	Positivos : negativos
	Positivos : (positivos + negativos)
	Negativos : (positivos + negativos)
	Neutral : (positivos + negativos)
	Todos : (positivos + negativos)
	Neutral : todos
Dinámica	Derivadas de primer y segundo orden

el total. A través de dicho método logran predecir la siguiente ventana de tiempo con un valor $MAE = 0.254$. Asimismo, compararon el uso de SARIMA (promedio móvil integrado autorregresivo estacional) y en todos los casos obtuvieron mejoras con medidas de errores más bajos.

Al hacer la predicción, es posible proponer otros modelos matemáticos. Zhang *et al.* (2010), en su modelo, utilizan tres funciones:

1. Función de correlación temporal: expresa la forma en que las emociones propias del usuario lo afectan más adelante.
2. Función de correlación social: expresa el efecto que tienen en el usuario las interacciones pasadas con otros usuarios.
3. Función de correlación de atributos: expresa la forma que los atributos del

usuario tienen un efecto en sus emociones.

Por ejemplo, la función de correlación temporal empleada por Moodcast (Zhang *et al.*, 2010) está dada por:

$$\frac{1}{Z_1} \exp \left\{ \sum_{v_i \in V} \sum_{x_{ik}^t \in X} \alpha_k f_k(x_{ik}^t, y_i^t) \right\}, \quad (3.2)$$

donde Z_1 es un factor de normalización, v_i representa un usuario, V es el conjunto de todos los usuarios, x_{ik}^t representa un cambio en el atributo k del usuario i en durante el tiempo t , X es el conjunto de todos los cambios que se han dado en los atributos de los usuarios y α_k es el peso o importancia para la función f_k . La función f_k utilizada por los autores es una función binaria que evalúa si la presencia del atributo existe durante el tiempo t para el usuario i . En caso afirmativo, $f_k = 1$; de lo contrario, $f_k = 0$.

En sus experimentos, Zhang *et al.* (2010) utilizan dos redes sociales: (1) una *red social móvil* (en inglés, *mobile social network*) de 30 usuarios voluntarios que compartieron sus mensajes de celular, localización GPS, señal de redes inalámbricas, calendario, alarmas e información de sus emociones y (2) LiveJournal, que es una red social en Internet en la que los usuarios pueden compartir sus vivencias. El enfoque de detección de emociones es semi-automático, pues se hace uso de las etiquetas que la misma red social proporciona a sus usuarios; estas etiquetas fueron catalogadas como positivas, negativas o neutras. En el caso de la red social móvil, se utilizaron cuatro días de periodo histórico, mientras que para LiveJournal se utilizaron tres meses. Moodcast obtuvo una precisión de 55.66 % (valor-F = 53.31 %) para la red social real, mientras que para LiveJournal obtuvo una precisión de 52.52 % (valor-F = 62.17 %).

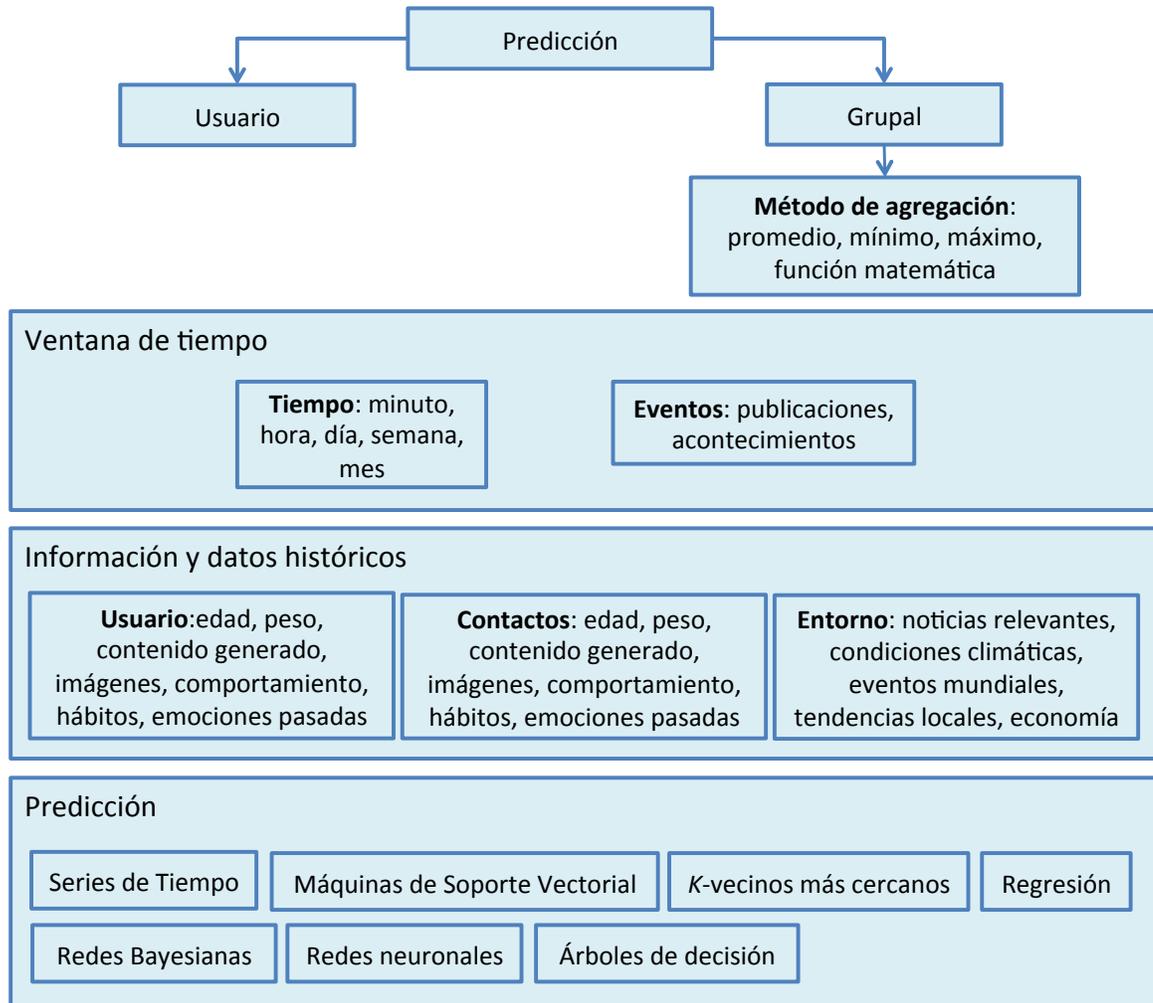


Figura 3.3: Predicción de emociones. Se muestran de forma general las diferentes opciones que se han investigado con más frecuencia a la hora de predecir emociones en redes sociales.

3.2 TRABAJOS RELACIONADOS

Los cuadros 3.2, 3.3 y 3.4 muestran un resumen comparativo de los principales trabajos acerca de predicción de emociones que se relacionan directamente con el modelo propuesto. Con el fin de identificar estos trabajos en dichos cuadros, para cada uno de ellos (denotado por T_i en el cuadro correspondiente) se lista su título, autor y año:

- T_1 .- *Predicting Collective Sentiment Dynamics from Time-series Social Media* (Nguyen *et al.*, 2012).
- T_2 .- *MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model* (Zhang *et al.*, 2010).
- T_3 .- *Predicting Teenager's Future Stress Level from Micro-blog* (Li *et al.*, 2015).
- T_4 .- *Twitter User Behavior Understanding with Mood Transition Prediction* (Mogadala y Varma, 2012).
- T_5 .- *Teenagers' Stress Detection Based on Time-Sensitive Micro-blog Comment/Response Actions* (Zhao *et al.*, 2015).
- T_6 .- *A Mood Prediction System for Preventing Diseases Using Biological Information and Weather Information* (Kajiwara *et al.*, 2017).
- T_7 .- *ULM: A user-level model for emotion prediction in social networks* (Qiyao *et al.*, 2016).
- T_8 .- *Joint user-interest and social-influence emotion prediction for individuals* (Rui *et al.*, 2017).
- T_9 .- *Predicting Depression of Social Media User on Different Observation Windows* (Hu *et al.*, 2015).

El primer punto relevante en cuanto a los cuadros comparativos es situar el modelo propuesto dentro de la taxonomía establecida para predicción, representación, detección y medición de emociones (figuras 3.1, 3.2 y 3.3). El segundo es propiamente comparar el modelo propuesto con algunos de los trabajos más cercanos dentro del estado del arte, ahora que se ha presentado una panorámica del mismo.

POSICIÓN DEL MODELO PROPUESTO DENTRO DEL ESTADO DEL ARTE.- En cuanto a la representación de una emoción, nuestro trabajo adopta el esquema binario

Cuadro 3.2: Comparativa de otros trabajos relacionados con el propuesto. La palabra “interdependencia” se refiere a que en el trabajo se considera la posibilidad de expresar varias emociones al mismo tiempo. La “emoción global” se refiere al uso de un número natural o real que indica el nivel de emoción sin ser más específico.

		N	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9
Tipo de emociones	BÁSICAS										
	Alegría	✓									
	Enojo	✓									
	Tristeza	✓									
	COMPLEJAS										
	Depresión										✓
	Estrés				✓		✓				
	ABSTRACTAS										
	Emoción global					✓					
	Positivo		✓	✓				✓	✓	✓	
Negativo		✓	✓				✓	✓	✓		
Neutro			✓				✓	✓			
Interdependencia	✓										
Representación	CATEGORÍA										
	Peso numérico				✓						✓
	Etiqueta de intensidad						✓	✓			
	Binario	✓	✓	✓					✓	✓	
	Numérico					✓					

(por ejemplo, el comentario contiene alegría o no la contiene). No obstante, en cuanto a la representación del *estado emocional* de un *usuario* en una red social en línea, nuestro trabajo utiliza un esquema vectorial con intensidades; hasta donde tenemos conocimiento, dicha representación sería nueva, pues las representaciones vectoriales conciernen emociones individuales, no etiquetas (estados) para usuarios. En cuanto al tipo de emociones manejadas, como ya se ha mencionado previamente, el trabajo propuesto utiliza el esquema de cuatro emociones básicas y considera interdependencia entre ellas. Con respecto a la detección de emociones, el trabajo propuesto se basa en el uso de lexicones y aprendizaje máquina. Con respecto a la tarea de predicción, se pueden manejar diferentes granularidades para las ventanas temporales y se predice un valor real (regresión); el nivel manejado es el individual, pues las

Cuadro 3.3: Comparativa de otros trabajos relacionados con el propuesto. La detección “autorreportada” indica que el propio sujeto de estudio reporta con una cierta frecuencia su estado emocional a los investigadores. En el tipo de predicción, la clasificación “cambio” indica que lo que se predice en el trabajo es el cambio de emoción con respecto a una medición anterior (puede subir, bajar o permanecer igual). La categoría “valor real” indica que, efectivamente, lo que se predice en el trabajo es la emoción del usuario o grupo de usuarios.

		N	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9
Detección de emociones	LEXICÓN	✓	✓	✓	✓	✓			✓	✓	✓
	Lingüística										✓
	Ontologías										
Detección de emociones	CLASIFICACIÓN										
	Autorreportada						✓	✓			
Granularidad	Aprendizaje Máquina	✓	✓	✓	✓	✓			✓	✓	✓
	INTERVALO										
	Minuto	✓					✓				
	Hora	✓	✓								
	Día	✓		✓				✓			
	Semana				✓						
	Mes										✓
	EVENTO										
	Tuit					✓					
	Toda la historia								✓	✓	
Tipo de predicción	NIVEL										
	Grupal		✓								✓
	Individual	✓		✓	✓	✓	✓	✓	✓	✓	
Tipo de predicción	PREDICCIÓN										
	Valor continuo	✓		✓	✓		✓	✓	✓	✓	
	Cambio		✓			✓					

predicciones se hacen a nivel usuario.

DIFERENCIAS CON LOS TRABAJOS MÁS CERCANOS.- Trabajos como los de Nguyen *et al.* (2012), Zhang *et al.* (2010), Li *et al.* (2015) y Mogadala y Varma (2012) utilizan emociones abstractas. En este sentido, nuestro trabajo es de los primeros en pronosticar emociones básicas. En el caso de Mogadala y Varma (2012), la emoción es representada por un número real que indica una *emoción global* del usuario. Aun-

Cuadro 3.4: Comparativa de otros trabajos relacionados con el propuesto. Los tipos de modelo indican si se ha utilizado algún método de aprendizaje máquina que emplea información o características (en inglés, *features*) para predecir la emoción o, si por el contrario, se ha diseñado un modelo matemático propio. En este último caso, los autores han tenido que diseñar métodos de ajuste de parámetros para sus modelos a través de algoritmos computacionales y posteriormente comparar los resultados con otros métodos de aprendizaje máquina tradicionales. MAPE = Error porcentual absoluto medio (en inglés, *mean absolute percentage error*).

	N	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	
Modelo	Basado en aprendizaje máquina	✓	✓		✓	✓	✓			✓	
	Matemático o probabilístico			✓				✓	✓		
Método de aprendizaje	Máquina de soporte vectorial		✓	✓		✓	✓				
	Regresión logística		✓				✓			✓	
	Árbol de decisión		✓								
	Regresión lineal	✓				✓					
	Serie de tiempo				✓						
	Clasificador bayesiano			✓			✓	✓			
	Regresión isotónica					✓					
	Proceso gaussiano						✓				
	Bosque aleatorio							✓			
	Algoritmo propio			✓					✓	✓	
	Medición del error	CLASIFICACIÓN									
		Precisión		✓	✓			✓	✓	✓	✓
Exhaustividad			✓	✓			✓	✓	✓	✓	
Valor-F			✓	✓			✓	✓	✓	✓	
REGRESIÓN											
Correlación		✓			✓	✓				✓	
MAE		✓			✓	✓				✓	
RMSE				✓	✓						
RRSE				✓	✓						
RAE				✓	✓						
MAPE				✓							

que esta sería una forma aún más simple de entender y representar las emociones humanas, ya que no se tiene claro en qué momento una emoción deja de ser positiva para ser neutra, el esquema de emociones básicas manejado en nuestro trabajo es todavía más granular y se puede aplicar a otros contextos (como los mencionados en el capítulo 1).

Asimismo, ninguno de los trabajos relacionados considera la coexistencia o interdependencia entre las distintas emociones en el usuario para hacer la predicción, es decir, no toman todas las emociones al mismo tiempo o se enfocan en una sola. Por ejemplo, en los trabajos de Li *et al.* (2015) y Zhao *et al.* (2015) el principal problema es detectar el estrés⁵ en usuarios adolescentes; por tanto, se concentran en una sola emoción y dejan de lado las demás. Un caso similar es el de Hu *et al.* (2015), ya que de igual forma utilizan solo una emoción para detectar depresión⁶. El no tomar en cuenta la interdependencia entre emociones es una de las principales diferencias con respecto al modelo propuesto, pues en este se considera que en un momento dado el usuario pueda experimentar varias emociones a la vez y esto interviene en intensidades emocionales futuras.

En cuanto a la detección de emociones, consideramos un esquema binario (la emoción está presente o ausente). Sin embargo, en otros trabajos se han utilizado etiquetas — por ejemplo: *malo*, *poco malo*, *normal*, *poco bueno* y *bueno* (Kajiwara *et al.*, 2017; Zhao *et al.*, 2015) y números reales (Mogadala y Varma, 2012). En el caso de Li *et al.* (2015), se manejan las etiquetas *sin estrés*, *muy débil*, *débil*, *moderado*, *severo* y *muy severo*, y se les asigna un peso numérico con el fin de obtener una serie de tiempo y llevar a cabo una predicción mediante métodos como ARIMA y SVARIMA.

En todos los trabajos relacionados mencionados se observa una tendencia muy marcada en la utilización de lexicones y de técnicas de aprendizaje máquina. Solamente Zhao *et al.* (2015) y Kajiwara *et al.* (2017) se apartan del grupo, ya que no utilizan un método de detección, sino que obtienen los datos a partir del propio usuario, quien va informando a lo largo del tiempo su estado emocional. Suponemos que esto es debido a la gran facilidad con la que se pueden obtener repositorios de información grandes en Internet de forma gratuita y los métodos de aprendizaje máquina que son más sofisticados cada vez. Nuestro trabajo tiene la misma tendencia

⁵El estrés puede ser considerado como una emoción (Verduyn y Lavrijsen, 2015).

⁶La depresión puede ser considerada también como una emoción (Spielberger y Reheiser, 2009).

de aprovechar las herramientas computacionales que se encuentran disponibles.

En cuanto a la granularidad para la ventana de tiempo, el trabajo propuesto — como se detalla en el capítulo 5 — analiza tres posibilidades: *días*, *horas* y *minutos* (con ventanas respectivas de un día, una hora y 15 minutos), mientras que los trabajos relacionados mencionados analizan un solo tipo de ventana de tiempo, ya sea un mes (Hu *et al.*, 2015), una semana (Li *et al.*, 2015), x cantidad de días (Kajiwara *et al.*, 2017; Zhang *et al.*, 2010), x cantidad de horas (Nguyen *et al.*, 2012) o x cantidad de minutos (Zhao *et al.*, 2015). Asimismo, otros utilizan eventos puntuales, tales como el momento mismo de la publicación de un comentario (Mogadala y Varma, 2012), o incluso toda la historia disponible (Qiyao *et al.*, 2016; Rui *et al.*, 2017). El uso de una ventana de tiempo en particular tiene implicaciones importantes que se deben considerar. Por ejemplo, se ha observado que las personas publican más comentarios y con más emociones positivas en los fines de semana (Dodds *et al.*, 2011). El tipo de ventana de tiempo a elegir depende de la aplicación que se esté buscando dar al modelo, por lo tanto no existe un tipo de ventana que sea más adecuado.

En cuanto a los modelos de predicción, también se aprecia una inclinación hacia utilizar técnicas de aprendizaje máquina a partir de casos previamente etiquetados; esto es, aprendizaje supervisado (Hu *et al.*, 2015; Kajiwara *et al.*, 2017; Li *et al.*, 2015; Mogadala y Varma, 2012; Nguyen *et al.*, 2012; Zhao *et al.*, 2015). Sin embargo, también se utilizan otro tipo de modelos en los que se propone un método probabilístico o matemático (Qiyao *et al.*, 2016; Rui *et al.*, 2017; Zhang *et al.*, 2010). Entre los métodos de aprendizaje máquina, se encuentran las máquinas de soporte vectorial, regresión logística, árboles de decisión, regresión lineal, series de tiempo, clasificadores bayesianos, regresión isotónica, proceso gaussiano, bosque aleatorio o algoritmos computacionales desarrollados por los autores, como en el caso de Zhang *et al.* (2010), Qiyao *et al.* (2016) y Rui *et al.* (2017).

La predicción final se hace en uno de dos niveles, según los trabajos actuales: a nivel de grupo — por ejemplo ciudad, estado, comunidad (Hu *et al.*, 2015; Nguyen

et al., 2012) — o a nivel individual (usuario). En los trabajos a nivel grupo, lo que se busca predecir es la emoción de una comunidad en conjunto. Por esta razón, debe existir un método de agregación que permita obtener un solo valor para el grupo. Por ejemplo, Hannak *et al.* (2012) evalúan las emociones de una ciudad a través de analizar los comentarios de todos los usuarios pertenecientes a dicha ciudad y ponderar sus valores. Una vez decidido el nivel de análisis para la predicción, los trabajos actuales emplean dos formas de medición: (1) el valor real de la emoción o emociones (como en nuestro trabajo) y (2) el cambio que se ve reflejado numéricamente (Mogadala y Varma, 2012; Nguyen *et al.*, 2012). Por “cambio”, nos referimos a una diferencia que se presente entre una medición y otra. El objetivo de estos trabajos es predecir dicha diferencia.

Como medidas de error, en los trabajos relacionados se utilizan con mayor frecuencia las métricas estándar relacionadas con aprendizaje máquina: correlación, error absoluto medio, error absoluto medio cuadrático, error absoluto relativo (Hu *et al.*, 2015; Li *et al.*, 2015; Mogadala y Varma, 2012) y, en el caso de clasificación por categorías, se utilizan precisión, exhaustividad y valor-F (Kajiwara *et al.*, 2017; Nguyen *et al.*, 2012; Qiyao *et al.*, 2016; Rui *et al.*, 2017; Zhang *et al.*, 2010; Zhao *et al.*, 2015).

3.3 RESUMEN

En este capítulo, analizamos los trabajos que se relacionan de forma directa con el nuestro y exploramos las diferentes áreas en las que se aplica teoría sobre representación, detección y clasificación de emociones: voz, imagen, video y música. Observamos que la mayoría de los trabajos existentes trabajan una sola emoción, como la alegría, el estrés o la depresión (Dodds *et al.*, 2011; Li *et al.*, 2015; Zhao *et al.*, 2015). Nuestro modelo incorpora una mejora en la que se considera la interacción entre varias emociones en un usuario con el fin de aumentar la precisión en la predicción futura. Los principales métodos hasta ahora están relacionados con el

aprendizaje máquina y los repositorios existentes en Internet. Vimos que las ventanas de tiempo son variadas en los diferentes trabajos relacionados y que su elección depende más bien de la aplicación práctica que se le vaya a dar al modelo. Como una de las contribuciones principales de nuestro trabajo podemos mencionar la utilización de emociones básicas en un modelo de predicción a futuro. Los demás trabajos se basan por lo general en emociones numéricas o abstractas. Otro tipo de representación como la vectorial (Cambria y Hussain, 2012) no se ha empleado para predicción de emociones futuras de usuarios, por lo que eso podría suponer un problema de investigación. Vimos que la predicción se puede llevar a cabo a nivel de usuario o a nivel grupo (Hu *et al.*, 2015; Nguyen *et al.*, 2012) a través de un método de agregación. En nuestro modelo, hacemos la predicción a nivel de usuario.

CAPÍTULO 4

MODELO

En este capítulo se explica el modelo propuesto para predecir la intensidad emocional de un usuario en una ventana de tiempo futura. La intensidad emocional, tal como se mencionó en el capítulo 1, consiste en la proporción de comentarios que escribe el usuario con una emoción en particular para un periodo de tiempo llamado *ventana de tiempo*. Se asume que esta intensidad depende de tres tipos de información o características: (1) del usuario — por ejemplo, sus emociones —, (2) de los contactos del usuario dentro de la red social y (3) del ambiente — por ejemplo, el día de la semana. Se asume también que la intensidad de una emoción se puede predecir utilizando tanto intensidades previas de la misma emoción como intensidades previas de las demás emociones. Para hacer la predicción, se emplea una *función de predicción* que tiene como objetivo obtener el valor futuro de la intensidad emocional a partir de la información histórica. En las secciones siguientes se describen estas características con mayor detalle.

En la sección 4.1, se comienza por establecer un marco de referencia (*vocabulario*) para la tarea que se está abordando. En la sección 4.2, se define y ejemplifica el concepto de intensidad emocional para dar paso a la descripción formal del modelo en la sección 4.3 y la función de predicción en la sección 4.4. Por último, en la sección 4.5 se considera la predicción de intensidades emocionales como un proceso y se describe cada una de sus fases: detección de emociones en comentarios, medición

de variables de entrada y, por último, propiamente la predicción de intensidades.

4.1 TAREA DE PREDICCIÓN

Nos referimos a *predicción* en el sentido de *pronóstico*, es decir, la estimación de valores futuros para una *variable objetivo*, dado un conjunto de *variables de entrada*. Cabe destacar que la relación entre los términos “predicción” y “pronóstico” es, hasta cierto punto, difusa. En algunas disciplinas se consideran sinónimos; en otras, *predecir* significa estimar valores futuros sin información (“corazonadas”) y *pronosticar* significa hacerlo con información. Otras definiciones sitúan el pronóstico como una subtarea de la predicción, siendo esta concepción a la que nos apegamos.

Normalmente, los valores de las variables de entrada son históricos, es decir, abarcan un *periodo* $\{t_1 \dots t_k\}$ de k unidades de tiempo y a cada unidad t_i le llamamos *ventana de tiempo*. De igual manera, nos referimos a k como el *tamaño del periodo histórico*. La *ventana de tiempo predecida*, como su nombre lo indica, corresponde a la unidad de tiempo t_{k+p} para la cual se hará la predicción, considerando que t_k es la última ventana para la cual se tienen datos históricos disponibles. Formalmente, también podemos denotar la ventana de tiempo predecida con el valor de la variable p , donde $p \geq 1$; note que si $p > 1$, existe *holgura* entre el periodo histórico y la predicción.

También es importante mencionar que existen diferentes tipos de *granularidad* temporal (minutos, horas, días, meses, bimestres, . . .); cabe destacar que la granularidad de la variable objetivo no necesariamente coincide con la de las variables de entrada. Por ejemplo, la granularidad de estas puede estar expresada en días, mientras que la variable objetivo puede predecirse para la siguiente hora. Por simplicidad, la granularidad de todas nuestras variables es la misma.

Los conceptos antes mencionados se resumen en el cuadro 4.1.

Cuadro 4.1: Resumen de conceptos relevantes en la tarea de predicción

Concepto	Definición
Variable objetivo	Variable a predecir (dependiente).
VARIABLES DE ENTRADA	VARIABLES PARA REALIZAR LA PREDICCIÓN (INDEPENDIENTES).
Ventana de tiempo	Unidad t_i de tiempo.
Periodo	Secuencia de k ventanas de tiempo.
Tamaño del periodo histórico	Cantidad de ventanas de tiempo utilizadas para predecir.
Ventana de tiempo predecida	Considerando que t_k es la última ventana de tiempo histórica, es la ventana t_{k+p} para la cual se hará la predicción.
Holgura	Cantidad de ventanas de tiempo que pasan entre t_{k+1} y la ventana de tiempo predecida, considerando que t_k es la última ventana de tiempo histórica.
Granularidad	Forma de medir el tiempo (segundos, días, meses, etcétera).

4.2 INTENSIDAD EMOCIONAL

La variable objetivo es la *intensidad emocional* de un usuario sobre un conjunto de emociones básicas. Definimos ξ como el conjunto de emociones consideradas para realizar la predicción. Dado que trabajos recientes han encontrado evidencia para sustentar el esquema de cuatro emociones básicas — alegría, enojo, miedo y tristeza — (Jack *et al.*, 2014), utilizamos este esquema de emociones. Sin embargo, el modelo tiene la capacidad de ajustarse a otros esquemas.

Asumimos que los usuarios manifiestan emociones a través sus publicaciones escritas, de forma que estas emociones pueden ser detectadas automáticamente. Teniendo esto en cuenta, definimos la intensidad emocional $I_{u,e}$ de un usuario u para la emoción $e \in \xi$ como la razón de comentarios que contienen e con respecto al total

de comentarios publicados por u en una ventana de tiempo t . Es decir,

$$I_{u,e}(t) = \frac{|\zeta_{u,e}(t)|}{|\zeta_u(t)|}, \quad (4.1)$$

donde $\zeta_{u,e}(t)$ y $\zeta_u(t)$ representan, respectivamente, el conjunto de comentarios de u que contienen la emoción e y el conjunto de comentarios publicados por u en la ventana en la que se llevan a cabo las mediciones. Estas ventanas de tiempo pueden tener diferentes granularidades, tales como un día, una semana o una hora, o pueden corresponder a algún tipo de evento puntual, como el momento mismo de la publicación de un comentario o el surgimiento de alguna noticia o acontecimiento externo a la red social. En general, la ventana de tiempo t representa los momentos de particular interés en el estudio. Por lo tanto, la intensidad emocional de un usuario es una función que depende de los comentarios publicados por el usuario, la emoción presente en los mismos y la ventana de tiempo. Dado que $\zeta_{u,e}(t) \subset \zeta_u(t)$, y que $\forall t \quad |\zeta_{u,e}(t)| \geq 0$, tenemos que $\forall t \quad I_{u,e}(t) \in [0, 1]$.

Para comprender mejor el concepto de intensidad emocional, supongamos que $I_{u,A}(t) = 2/5$, donde A representa alegría. Esto significa que el usuario u publicó cinco comentarios en la ventana t , y dos de estos comentarios expresaron alegría. Por tanto, la intensidad emocional para u con esta emoción en específico (alegría) es de 0.4 para la ventana analizada.

Note, asimismo, que la representación del *estado emocional* de un usuario, entonces, está dada por un vector de pesos

$$I_u \vec{t} = (I_{u,1}(t) \dots I_{u,|\xi|}(t)), \quad (4.2)$$

donde cada peso está dado por la intensidad $I_{u,j}(t)$ y $1 \leq j \leq |\xi|$. La intensidad emocional es importante en nuestro modelo ya que, como se mencionó anteriormente, el objetivo principal del mismo es predecir $I_{u,e}(t_{k+p})$ a partir de información conocida para el periodo histórico t_1, t_2, \dots, t_k .

4.3 DESCRIPCIÓN DEL MODELO

Se propone un modelo para la predicción de la intensidad emocional futura de un usuario¹. Como ya hemos mencionado, este modelo toma en cuenta variables de entrada que comprenden información del usuario y de su entorno para hacer la predicción. Esta información incluye: (1) datos relevantes acerca del usuario, (2) datos relevantes acerca de los contactos del usuario y (3) datos relevantes acerca del entorno, los cuales pueden ser internos o externos a la red social. Para realizar la predicción, es necesario contar con datos históricos de estas variables y asociar las variables de entrada con la variable objetivo. Esto último puede hacerse mediante una función. Por tanto, $I_{u,e}(t)$ sería el resultado de la *función de predicción de intensidad emocional* para la emoción $e \in \xi$. Esto es,

$$I_{u,e}(t+p) = f(\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_m, \mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n, \mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_q), \quad (4.3)$$

donde $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_m$ es una serie de m matrices con la información histórica del usuario, $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ es una serie de n matrices con la información histórica de sus contactos y $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_q$ es una serie de q matrices con la información sobre el entorno, f es la *función de predicción* y $t+p$ representa la ventana de tiempo predecida (recordando que $p \geq 1$). Las matrices del usuario \mathbf{U}_i son matrices de tamaño variable cuyo contenido depende del usuario propiamente.

Definimos la matriz \mathbf{I} como una *matriz de intensidades emocionales* de tamaño $k \times |\xi|$, en la que sus componentes son intensidades emocionales históricas. Cada columna de la matriz \mathbf{I} corresponde a una de las emociones consideradas (alegría, enojo, miedo y tristeza) y cada fila corresponde a una ventana de tiempo t_i de las mediciones. En términos matemáticos, la matriz \mathbf{I} está dada por

¹La versión final del modelo que se propone es resultado también de experimentación preliminar con otros modelos y procesos de selección de variables (*feature selection*); resultados de estos procesos no se reportan por cuestiones de espacio y relevancia.

$$\mathbf{I} = \begin{bmatrix} I_{u,1}(t_1) & \cdots & I_{u,|\xi|}(t_1) \\ \vdots & \ddots & \vdots \\ I_{u,1}(t_k) & \cdots & I_{u,|\xi|}(t_k) \end{bmatrix}, \quad (4.4)$$

donde $I_{u,j}(t_i)$ es la intensidad para la emoción con $1 \leq j \leq |\xi|$ y $1 \leq i \leq k$, medida a partir de los comentarios publicados en la ventana de tiempo t_i . Se podría considerar otro tipo de información proveniente de los usuarios, por ejemplo, su peso corporal medido en cada ventana de tiempo, edad, historial clínico u otro tipo de datos similares. Sin embargo, eso está fuera del alcance de esta tesis, ya que por el momento no se tiene acceso a esa información. El estudio de tales variables quedaría como una mejora para un trabajo futuro.

Definimos las matrices \mathbf{U} , \mathbf{A}_f y \mathbf{A}_m como tres tipos de matrices de intensidades emocionales. En cada una de estas matrices, las intensidades son medidas a partir de información distinta. Para la matriz \mathbf{U} , las intensidades son obtenidas a partir de comentarios publicados por el *usuario*. Para la matriz \mathbf{A}_f , las intensidades se obtienen a partir de los comentarios publicados por los usuarios que sigue. Para la matriz \mathbf{A}_m , las intensidades se obtienen a partir de las menciones (en inglés, *mentions*) hacia el usuario. Las menciones tienen la peculiaridad de generar una notificación de mensaje hacia el usuario, provocando que este les brinde atención. Por lo tanto, la intensidad emocional se mide tomando en cuenta solo los comentarios en donde se efectúa una mención.

Las matrices \mathbf{A}_f y \mathbf{A}_m son nuestra propuesta para incorporar el *contagio emocional* (definido en el capítulo 2) en la implementación de nuestro modelo. Dado el impacto que tiene el contagio emocional en las emociones, consideramos que es importante tomarlo en cuenta para llevar a cabo una predicción de la intensidad emocional. Al existir en Twitter estos dos mecanismos de interacción entre usuarios, tomamos la decisión de incluir una matriz para cada uno, \mathbf{A}_f y \mathbf{A}_m .

Las matrices del entorno \mathbf{E}_i , son matrices de tamaño variable, cuyo contenido

depende del entorno del usuario (por ejemplo temperatura, presencia de lluvia, día de la semana). Existe amplia evidencia de que estos factores afectan a las emociones; por ejemplo, Dodds *et al.* (2011) mostraron que el día de la semana tiene un impacto en la expresión de la felicidad (con picos en fines de semana y la menor expresión de felicidad en el día martes). Hannak *et al.* (2012) encontraron que la humedad atmosférica tiene un efecto negativo en la cantidad de comentarios positivos publicados en Twitter. Definimos la matriz \mathbf{E} como una matriz de tamaño 1×7 cuyas componentes son todas iguales a cero excepto aquella cuyo número de la columna sea igual al número del día de la semana del día al que pertenece la ventana de tiempo predecida t_{k+p} (consideramos que la semana empieza en domingo). Por ejemplo, el día lunes quedaría representado como $(0,1,0,0,0,0,0)$ y el viernes quedaría representado como $(0,0,0,0,0,1,0)$.

Partiendo de la información disponible, nuestro modelo propuesto quedaría expresado como

$$I_{u,e}(t+p) = f(\mathbf{U}, \mathbf{A}_f, \mathbf{A}_m, \mathbf{E}). \quad (4.5)$$

Una de las principales aportaciones del modelo propuesto frente a otros es que en nuestro modelo es posible incluir mediciones históricas de otras emociones; por ejemplo, si se desea predecir la alegría del usuario en un día futuro, nuestro modelo permite incluir información sobre emociones como enojo o tristeza de ventanas de tiempo anteriores para hacer la predicción. Otros trabajos actuales han utilizado solo la misma emoción para la predicción. Por ejemplo, han utilizado datos históricos de solo la alegría para predecir alegría. El modelo propuesto en sí no está limitado a datos extraídos a partir de la red social — por ejemplo, se podrían agregar el peso corporal, la temperatura ambiental, o incluso el historial médico del usuario. Sin embargo, dado el alcance de la tesis, trabajamos solo con las matrices ya definidas anteriormente.

4.4 FUNCIÓN DE PREDICCIÓN

La función de predicción f tiene la finalidad de obtener un valor de intensidad emocional futuro a partir de la información que se tiene del usuario, sus contactos y el entorno. En este sentido, nos estamos enfrentando a un problema de pronóstico (en inglés, *forecasting*). Esta función no es trivial y no se conoce una expresión directa de la misma; por tal motivo, la única forma de obtener una expresión para f es a través de la obtención de un modelo de los datos con métodos computacionales apropiados. Como se mencionó en el capítulo 2, algunos de los métodos más utilizados en la actualidad para este fin son: (1) técnicas de aprendizaje máquina; (2) regresión lineal, no lineal, univariable o multivariable y (3) análisis de series de tiempo. Entre los métodos de aprendizaje máquina se encuentran la regresión lineal², regresión con máquinas de vectores de soporte, redes neuronales, lógica difusa, método de k -vecinos más cercanos y redes bayesianas, entre otros. Dado que la intensidad emocional puede ser medida para un usuario en una ventana de tiempo dada, es posible crear repositorios con datos reales, lo cual habilita el uso de métodos de aprendizaje máquina supervisados.

Debido a lo exhaustivo de los experimentos realizados en esta tesis, optamos por utilizar como método de aprendizaje supervisado principal la regresión lineal. La regresión lineal tiene la ventaja de ser un método muy rápido para obtener un modelo de los datos y, al utilizar solo coeficientes reales, permite una interpretación más intuitiva del modelo final. Como se explicó anteriormente, es un método computacional para la obtención de un modelo de los datos de una población a partir de una colección limitada de datos llamada *muestra*.

En el modelo propuesto, cada componente de las matrices de intensidad emocional y las matrices del entorno que se utilizan como parámetros de la función f

²Aunque la regresión lineal es un método propiamente estadístico, debido a sus características y su *modus operandi*, también se le puede considerar dentro de la gama de técnicas supervisadas de aprendizaje máquina.

son consideradas variables de entrada. Por lo tanto, al aplicar el modelo de regresión lineal existe un coeficiente β_i para cada componente de las matrices. Definimos el vector columna ν_j como la j -ésima columna de la matriz de intensidad emocional \mathbf{I} y, asimismo, $\nu_{U,j}$, $\nu_{f,j}$ y $\nu_{m,j}$ como las j -ésimas columnas de las matrices \mathbf{U} , \mathbf{A}_f , \mathbf{A}_m respectivamente con $1 \leq j \leq |\xi|$ y definimos ν_E como el vector columna correspondiente para \mathbf{E} . Definimos los vectores columna $\beta_{U,j}$, $\beta_{m,j}$, $\beta_{f,j}$ y β_E como vectores de coeficientes cuyas componentes son parámetros que deben estimarse para obtener el modelo lineal. Aplicando estas definiciones en la ecuación 4.6, nuestro modelo propuesto quedaría de la siguiente manera:

$$I_{u,e}(t+p) = \beta_0 + \sum_{j \in |\xi|} (\beta_{U,j} \cdot \nu_{U,j} + \beta_{f,j} \cdot \nu_{f,j} + \beta_{m,j} \cdot \nu_{m,j}) + \beta_E \cdot \nu_E + \epsilon, \quad (4.6)$$

donde la operación del producto escalar de dos vectores está denotada por “ \cdot ”. De esta forma, el problema se reduce a estimar las componentes de los vectores $\nu_{U,j}$, $\nu_{f,j}$, $\nu_{m,j}$ y ν_E , así como los parámetros β_0 y ϵ .

4.5 FASES EN LA PREDICCIÓN DE INTENSIDAD EMOCIONAL

En esta sección se describe el proceso de predicción que se lleva a cabo al aplicar el modelo propuesto. En él se incluyen como variables de entrada las matrices de intensidades emocionales del usuario y la de sus contactos (\mathbf{U} , \mathbf{A}_f y \mathbf{A}_m , respectivamente), y la matriz de variables del entorno \mathbf{E} . Comenzamos explicando el proceso de detección de emociones para generar las matrices de intensidad; posteriormente, describimos la medición de la misma en las ventanas de tiempo, y finalmente describimos también la predicción.

4.5.1 DETECCIÓN DE EMOCIONES

Para predecir la intensidad emocional en los comentarios futuros del usuario — es decir, para poder determinar con cuánta alegría, tristeza, enojo y miedo escribirá sus próximos comentarios — necesitamos primero poder reconocer emociones en texto. Para ello, asumimos que los usuarios transmiten sus emociones a través de los comentarios que publican dentro de la red social. En el área de análisis de sentimiento, se ha analizado anteriormente el problema de detectar emociones en texto; por ejemplo, algunos trabajos han intentado detectar emociones básicas (Sidorov *et al.*, 2012) y otros han definido emociones abstractas positivas, neutras y negativas, es decir una polaridad (Nguyen *et al.*, 2012; Zhang *et al.*, 2010).

En nuestro modelo, definimos la matriz de intensidad \mathbf{I} que se obtiene a partir de los comentarios publicados. La intensidad emocional está dada por la ecuación 4.1, donde la cuenta de comentarios se lleva a cabo utilizando algún criterio específico en una ventana de tiempo t . Por lo tanto, nuestro principal enfoque para la obtención de datos acerca del usuario y de sus contactos es a través de la categorización de texto en una o más de las emociones consideradas en ξ . Cabe mencionar que la detección de emociones normalmente se da por sentada en otros trabajos, es decir, se parte del problema de predicción; este trabajo, por el contrario, aborda también esta tarea, pues en el idioma español se carece de métodos específicos para la detección de emociones en texto.

La metodología propuesta para detectar emociones en texto se muestra en la figura 4.1. Se parte de un conjunto \mathbb{C} de comentarios candidatos, los cuales se dan como entrada a un *clasificador de sentimiento*. Este clasificador tiene como fin detectar comentarios subjetivos, ya que este tipo de comentarios probablemente contendrán emociones (contrario a los comentarios objetivos, que suelen contener hechos) y servirán para entrenar clasificadores específicos de emoción. Los comentarios subjetivos encontrados son manualmente *etiquetados*, es decir, diferentes evaluadores asocian cada comentario con las diferentes emociones que expresa y estos comentarios etique-

palabras y frases) y ha demostrado aventajar otros clasificadores para este mismo idioma (Rodríguez Aldape, 2013); sin embargo, no descartamos el uso de otros clasificadores para otros idiomas como trabajo futuro.

En cuanto al etiquetado manual, consideramos que cada comentario puede expresar varias emociones; por tanto, puede contener varias etiquetas. Para determinar las etiquetas de cada comentario, nuestra metodología combina ideas que han sido propuestas para, respectivamente, aproximar el criterio de un experto y producir un puntaje global cuando se tienen diferentes evaluadores. En cuanto a lo primero, Snow *et al.* (2008) encontró que a partir de múltiples evaluaciones (siete) por parte de no expertos, se puede aproximar el criterio de un experto. En cuanto a lo segundo, Qing *et al.* (2014) propone un método llamado *acuerdo de intercorrelación* (en inglés, *intercorrelation agreement*), previamente explicado también en el capítulo 2, para ponderar los puntajes asignados por diferentes evaluadores.

El acuerdo de intercorrelación, en nuestro caso, se aplicaría de la siguiente manera. Considerando un conjunto de etiquetas $L = \{a, p\}$ que indican la ausencia o presencia de cada emoción $e \in \xi$ en un comentario c , la función de etiquetado consiste en asignar a c , dada e , la etiqueta que obtenga el mayor puntaje:

$$f_A(c, e) = \operatorname{argmax}_{l \in L} \sum_{r \in R} \omega_r \cdot \delta(l_r, l), \quad (4.9)$$

donde R es el conjunto de evaluadores, ω_r es la confiabilidad del evaluador r (ver ecuación 2.16, página 33) y δ es la delta de Kronecker³. De esta manera, por ejemplo, si la etiqueta “presente” (p) es la que predomina para la emoción de tristeza en el comentario i , entonces este comentario será etiquetado como una instancia positiva (ejemplo positivo) para la emoción de tristeza.

Con los comentarios etiquetados, es posible recolectar instancias para entrenar clasificadores binarios de emoción (uno para cada emoción del conjunto ξ). En este caso, utilizamos máquinas de soporte vectorial — un tipo de clasificador que

³La delta de Kronecker $\delta(i, j)$ es una función que da como resultado 1 cuando $i = j$ y 0 en caso contrario.

ha demostrado encontrarse bien acoplado para la clasificación de texto (Aggarwal y Zhai, 2012; Medhat *et al.*, 2014; Ye *et al.*, 2005). Para los vectores de características, se usan frecuencias de unigramas; como veremos más adelante, estas características dan resultados satisfactorios para la tarea de detección de emociones en la experimentación realizada.

Con los clasificadores entrenados, los comentarios que se utilicen para hacer la predicción de intensidad emocional ya pueden ser automáticamente clasificados y se pueden generar las intensidades históricas del usuario y sus contactos.

4.5.2 MEDICIÓN DE VARIABLES DE ENTRADA

En nuestro modelo, la medición de las variables de entrada (es decir, intensidades emocionales y entorno) puede hacerse en intervalos de tiempo regulares (minutos, horas, días, semanas) o irregulares (por ejemplo, durante ciertos eventos o cuando las condiciones necesarias para hacerlo se satisfagan). En nuestro caso, utilizamos intervalos regulares para poder comparar contra modelos como ARIMA, pues los intervalos de tiempo irregulares impiden que las series de tiempo puedan obtener la función de predicción f , ya que se basan en la suposición de que las mediciones se hacen a intervalos regulares. Como en nuestro modelo hacemos referencia a lo sucedido dentro del intervalo de tiempo observable en ese momento, lo llamamos ventana de tiempo t . Por ejemplo, $\zeta(t)$ representa el conjunto de comentarios que fueron publicados dentro de un lapso de tiempo concreto; no se puede obtener su valor hasta que dicho intervalo de tiempo haya transcurrido. Esto contrasta con el caso de simplemente tomar el valor del total de comentarios publicados al momento de la observación, ya que dicho valor continuaría creciendo sin un límite. En términos matemáticos, la cantidad de comentarios publicados por el usuario u , denotada por ψ_u , está dada por

$$\psi_u = \zeta_u(t_1) + \dots + \zeta_u(t_2) + \dots + \zeta_u(t_k), \quad (4.10)$$

donde las t_i son las ventanas de tiempo y k es el periodo histórico con $1 \leq i \leq k$. En el caso, por ejemplo, del peso corporal, el valor de las variables en cada ventana de tiempo podría estar dado por alguna función de agregación (promedio, mínimo o máximo) o por el cambio detectado en dicha ventana de tiempo.

En nuestro modelo, asumimos que en las mediciones de estas variables de entrada obtenemos números reales y que su valor se encuentra dentro de un rango cerrado, puesto que buscamos tener un modelo computacional y esto permite el uso de métodos de aprendizaje máquina. Para las redes neuronales, el teorema de aproximación universal requiere que las entradas de la red o variables independientes sean valores reales en rangos cerrados (Csáji, 2001); esto permite que puedan aproximar funciones continuas como la función de predicción f . En el caso de datos categóricos como el día de la semana, es posible utilizar una clasificación numérica como en el caso de la ecuación 4.7 y la matriz \mathbf{E} . Estas mediciones han de almacenarse en alguna base de datos para posteriormente ser pasadas a la función de predicción. Actualmente, no se sabe la granularidad adecuada para las ventanas de tiempo; sin embargo, en algunos trabajos relacionados se ha utilizado el día o la semana. En el capítulo 5, se analizarán ventanas de tiempo con diferentes granularidades para compararlas y observar qué cambios producen en los resultados.

4.5.3 PREDICCIÓN

El objetivo del modelo es predecir la intensidad emocional futura de los usuarios. Para tal propósito, se ha diseñado una función de predicción f que toma como variables de entrada información del usuario, sus contactos y el entorno. En la ecuación 4.3 se hace referencia, mediante p , al número de ventana futura que se desea predecir. En el modelo propuesto, cada valor diferente de p requerirá encontrar una nueva expresión para f . Es decir, f está diseñada para predecir una sola ventana de tiempo futura y no un intervalo. Como veremos más adelante en el capítulo 5, la calidad en la predicción es medida a través del error absoluto medio y la corre-

lación. El primero indica qué tan cercano al valor real se encuentra la predicción y la segunda indica si la función f tiene un comportamiento creciente o decreciente similar.

Una de las principales diferencias de nuestro modelo con respecto a otros, es que nuestro modelo emplea datos históricos de *todas* las emociones del usuario y sus contactos para predecir la intensidad emocional de *una* emoción $e \in \xi$. Por ejemplo, se considera que la alegría está en función del enojo, la tristeza y otras emociones, y así también para el resto de las emociones. Otros modelos, como hemos visto, se centran frecuentemente en predecir una sola emoción (como la alegría o el estrés) usando como variable de entrada solamente la misma emoción (además de otras variables exógenas o ambientales).

4.6 RESUMEN

En este capítulo, describimos el modelo propuesto, sus componentes y las fases que se llevan a cabo en la predicción, las cuales incluyen la recopilación de datos, la detección de emociones básicas, la definición de una ventana de tiempo y la predicción. En la etapa de detección de emociones, se analizan los comentarios con el fin de reconocer la presencia de estas en el texto. En este trabajo, suponemos que un comentario puede tener o no tener una emoción, por lo que nos referimos a una clasificación binaria. Una vez clasificados los comentarios, podemos calcular la intensidad emocional $I_{u,e}(t)$, cuyos valores históricos son utilizados posteriormente. La predicción se lleva a cabo a través de una función de predicción f que toma en cuenta información del usuario, de sus contactos y del entorno. Se puede utilizar cualquier método de aprendizaje máquina como función f , o incluso algún método probabilístico. Hemos optado por el método de regresión lineal debido a su velocidad, lo cual nos permitió realizar experimentos exhaustivos.

Para que los datos puedan ser procesados en la función f , definimos matrices de

intensidad emocional, que incluyen datos históricos de las intensidades emocionales para cada emoción básica: alegría, enojo, miedo y tristeza. Asimismo, definimos una matriz de entorno cuyas componentes reflejan la presencia de eventos externos que tienen influencia en las emociones del usuario (en nuestro caso, el día de la semana). La función f toma en cuenta toda esta información en conjunto para hacer la predicción de una intensidad emocional futura del usuario.

CAPÍTULO 5

EXPERIMENTOS Y RESULTADOS

En este capítulo, se explican los experimentos realizados para comprobar las hipótesis planteadas. Primero se describe el repositorio utilizado y las razones por las cuales fue elegido (sección 5.1); con este repositorio, como se explica más adelante, se generaron las instancias de entrenamiento para validar la detección de emociones y se generaron, asimismo, las instancias de prueba para diversos usuarios. Esto se detalla en la *configuración* experimental (sección 5.3.1), donde además se describen otros datos necesarios para la preparación de los experimentos. Por cada tipo de experimento, se exponen los *resultados* obtenidos — como precisiones y medidas de error — y los principales hallazgos (secciones 5.2.2 y 5.3.2). Finalmente, también por cada tipo de experimento se presentan secciones de *discusión* (secciones 5.2.3 y 5.3.3), donde se da una interpretación de los resultados y un razonamiento sobre lo obtenido, así como su impacto en la comprobación de las hipótesis.

5.1 REPOSITORIO

Utilizamos Twitter como caso de estudio. Esta red social en línea es una de las más visitadas actualmente y cuenta con más de 67 millones de usuarios registrados en Estados Unidos, con lo que se posiciona dentro de los primeros cinco lugares de uso, según un estudio realizado por Pew Internet (Perrin, 2016). En el estado de

Nuevo León, México, Twitter es una de las dos redes más utilizadas junto con Facebook. Dado que existen pocos trabajos en español sobre detección y predicción de emociones, hemos decidido trabajar con este idioma como parte de nuestra contribución. Por tales motivos, la red social en línea seleccionada para los experimentos — como ya lo mencionamos — es Twitter. Podrían extraerse datos de usuarios seleccionados al azar de esta red social en línea con la restricción de que se tratasen de usuarios hispano-hablantes. Sin embargo, esto podría ocasionar algunos problemas: (1) que no se cuente con información suficiente de sus contactos, (2) que los usuarios tengan una ubicación geográfica distante, lo cual incrementa la probabilidad de que utilicen modismos o expresiones distintas que incluso puedan ser incompatibles y (3) que expresen sus emociones de maneras distintas en la red social. Con el fin de minimizar estos problemas, es deseable que los usuarios seleccionados pertenezcan a una sola ciudad.

Por este motivo, se hizo uso de un conjunto de comentarios extraídos de Twitter a partir de usuarios pertenecientes a la ciudad de Monterrey, N.L. Estos comentarios fueron utilizados por Rodríguez Aldape (2013) para evaluar el interés de un usuario en un tema (por ejemplo, en un producto o servicio). El repositorio contiene 40,186,542 comentarios provenientes de 80,954 usuarios. Se decidió utilizar este repositorio, además, por las siguientes razones: (1) el repositorio contiene una cantidad suficientemente grande de usuarios y relaciones de amistad, así como de publicaciones, (2) se trata de un repositorio en español, lo cual permite responder algunas de las hipótesis planteadas sobre este idioma, (3) el repositorio permite que personas quienes radican en la ciudad de Monterrey entiendan el lenguaje coloquial utilizado en los comentarios y estén familiarizadas con las formas de expresar las emociones de los mismos y (4) la relación de amistad de usuarios de una misma ciudad podría ser más personal, por lo que también podría presentarse con mayor fuerza el fenómeno de contagio emocional, el cual deseamos observar.

Como veremos más adelante, es necesario crear herramientas automáticas de detección de emociones en texto, ya que sería imposible etiquetar manualmente todos

los comentarios del repositorio. En comparación con otras herramientas como Amazon's Mechanical Turk¹, se tiene la ventaja de que personas de la localidad realicen el etiquetado, con lo cual se espera una mejor calidad y, por ende, una mejor precisión para los clasificadores automáticos. De igual manera, en general se decidió utilizar el repositorio mencionado debido a que el lenguaje coloquial utilizado en la ciudad es compartido por sus habitantes. Por esta razón, no tienen demasiada confusión en la interpretación de los comentarios — lo que supone una mejor comunicación; gran parte del trabajo con estos comentarios ha sido realizado por personas que viven en dicha ciudad, ya que a estas les ha sido más fácil comprender los comentarios y etiquetarlos adecuadamente.

Dado que la muestra de usuarios en el repositorio no es representativa de la población mundial ni de las personas hispano-hablantes, los patrones encontrados representan solo sugerencias de patrones de comportamiento en los seres humanos. Estos patrones han de ser estudiados posteriormente con mayor profundidad y quizá desde una perspectiva psicológica con el fin de comprobarlos.

5.2 DETECCIÓN DE EMOCIONES

A continuación se presentan los experimentos relacionados con la detección de emociones en los comentarios extraídos del repositorio previamente mencionado. Se describen la configuración experimental y los resultados más sobresalientes. De igual manera, se discute brevemente el impacto de estos resultados en la siguiente fase, que es la predicción de intensidades emocionales.

¹Mechanical Turk de Amazon.com es una herramienta en la que los trabajadores reciben cantidades pequeñas de dinero a cambio de realizar actividades que requieren inteligencia humana y que posiblemente sean repetitivas. Los trabajadores realizan las actividades de forma anónima, por lo que la plataforma se presta para que ingresen datos falsos o contaminados con tal de hacer las actividades en menor tiempo y así recibir un poco más de dinero. Esto conlleva a que al utilizar estas plataformas sea necesario realizar operaciones de limpieza y pruebas de confianza a las evaluaciones y posiblemente rechazar todas las evaluaciones de un usuario o grupos de usuarios.

5.2.1 CONFIGURACIÓN

Para implementar el modelo, es necesario detectar las emociones que contienen los comentarios. Este proceso es indispensable para poder obtener un valor de intensidad emocional según la ecuación 4.1. Considerando lo anterior, se etiquetó manualmente un conjunto pequeño de comentarios para entrenar clasificadores de cada emoción básica.

Para reducir el esfuerzo realizado por los evaluadores, como previamente se mencionó en el capítulo 4, se utilizó el clasificador de sentimiento *Twitter Opinion Mining*, también conocido como TOM (Rodríguez Aldape, 2013). Este clasificador considera tres categorías: positivo, negativo y neutro. Al hacer una evaluación previa con esta herramienta y excluir los comentarios neutros, se esperó obtener aquellos comentarios que con mayor probabilidad tenían una emoción presente. También para facilitar el etiquetado, solo se seleccionaron usuarios del repositorio que cumplieran las siguientes condiciones:

1. Provenir del área metropolitana. Se eligieron aquellos usuarios que en su locación publicada tuvieran alguna ciudad del área metropolitana de Monterrey, N.L. Esta característica ya está diferenciada en el repositorio, lo cual facilitó el proceso de selección.
2. Tener como mínimo 30 contactos (*followees*). Lo anterior fue con el fin de obtener una muestra estadísticamente significativa.
3. Haber publicado, tanto el usuario como sus contactos, 400 comentarios como mínimo. Lo anterior fue con el fin de obtener una muestra estadísticamente significativa de comentarios, lo cual permite que se apliquen las fórmulas para obtener la intensidad emocional.

5.2.1.1 ETIQUETADO DEL CONJUNTO DE COMENTARIOS

Durante un periodo de un año, se les asignó la tarea de etiquetar comentarios a varios grupos de alumnos de licenciatura de la Facultad de Ingeniería Mecánica y Eléctrica de la UANL. Cada comentario fue evaluado para las emociones básicas indicando el nivel de intensidad de cada emoción (*sin emoción, poca, presente y mucha*). Se utilizó el prototipo en línea mostrado en la figura 5.1. Los alumnos recibieron puntos extras en sus calificaciones como recompensa por llevar a cabo la evaluación. En general, se les pidió etiquetar por lo menos 400 comentarios. Cada comentario a etiquetar fue elegido de forma aleatoria y se permitieron hasta doce evaluaciones por comentario. Si un comentario alcanzaba las doce evaluaciones, se daba por concluido y se continuaba con el comentario siguiente. En el cuadro 5.1, se puede observar un resumen de los comentarios etiquetados y la cantidad de evaluadores que participaron; en 2015 y 2016, los grupos de alumnos que participaron etiquetaron 4,090 comentarios.

Participaron 141 evaluadores diferentes. El proceso de selección de comentarios a etiquetar fue el siguiente: si un comentario era detectado por lo menos por un evaluador con una emoción, se le daba prioridad para que otros evaluadores lo etiquetaran, y si en su primera evaluación el comentario resultaba no tener ninguna emoción, se le daba menos prioridad en la selección aleatoria de comentarios siguiente. Este mecanismo permitió ahorrar trabajo para encontrar comentarios con emoción, ya que el primer evaluador actuaba como filtro. Si todos los comentarios del conjunto de datos ya contaban con una evaluación y en ninguno se había detectado emoción, se podía elegir entre hacer una segunda ronda de todos los comentarios (por si otro evaluador detectaba alguna emoción) o terminar el proceso de evaluación.

Para elegir el nivel de emoción presente en cada comentario y posteriormente decidir si una emoción estaba presente en el mismo, utilizamos el *acuerdo de inter-correlación* de Qing *et al.* (2014), visto en la sección 2.6, por lo que solo tomamos en cuenta aquellos comentarios que recibieron por lo menos siete evaluaciones de

Comentario a evaluar:

@Tomasitu @sergioortega Pero cómooo Aysel es naa Maravillaa :P

Emoción	Nada	Poco Observable	Observable	Muy Observable
Alegria	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable
Tristeza	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable
Enojo	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable
Miedo	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable
Repulsión	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable
Sorpresa	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable
Desesperanza	<input type="radio"/> Nada	<input type="radio"/> Poco Observable	<input type="radio"/> Observable	<input type="radio"/> Muy Observable

Otro idioma

No sé

Sarcasmo

Neutral

Figura 5.1: Prototipo utilizado para el etiquetado manual de emociones en comentarios de Twitter.

personas diferentes; estos comentarios fueron en total 1,895. El acuerdo de intercorrelación tiene la ventaja de que no necesita un proceso de limpieza de evaluaciones debido a que asigna automáticamente pesos menores a los evaluadores con una calificación más pobre en las evaluaciones. Es decir, las etiquetas con mayor calidad de los evaluadores tienen más importancia en la selección de la etiqueta final. Para obtener la confiabilidad de cada evaluador (ecuación 2.16, página 33), empleamos primero $|K| = 4$ niveles de presencia (*sin emoción, poca, presente y mucha*). Sin embargo, después de obtener los resultados preliminares mostrados en el cuadro 5.2, se vio un desbalance en comparación con la clase *sin emoción*, por lo que se decidió utilizar solamente las etiquetas de *ausente* o *presente* ($|K| = 2$) para entrenar los clasificadores binarios, tal como se menciona en el capítulo 4. Utilizando estas etiquetas, los comentarios por cada emoción quedan distribuidos como se muestra en el cuadro 5.3.

Cuadro 5.1: Cantidad de evaluaciones y evaluadores que participaron.

Año	Mes	Evaluaciones	Evaluadores
2015	Abril	1349	10
2015	Mayo	8386	49
2015	Junio	449	2
2015	Agosto	12091	68
2015	Noviembre	380	1
2016	Abril	4	1
2016	Mayo	3853	19
Total		27712	150

Cuadro 5.2: Resumen de clasificaciones de comentarios (preliminar).

Emoción	Sin emoción	Poca	Presente	Mucha
Alegría	1167	134	499	95
Enojo	1698	64	107	26
Miedo	1867	14	13	1
Tristeza	1713	104	63	15
Total	6445	316	682	137

Se puede observar que en el caso del miedo, existen solamente 28 casos con esta emoción presente. Esta situación nos obligó a excluir el miedo del análisis, ya que se requiere una mayor cantidad de instancias para entrenar un clasificador. Este resultado indica que las personas tienden a escribir pocos comentarios expresando miedo en esta red social, sin embargo, esto necesitaría una investigación más profunda.

5.2.1.2 CREACIÓN DE LOS CLASIFICADORES

En la creación de los clasificadores se utilizó un enfoque llamado *saco de palabras* (en inglés, *bag of words*), el cual considera a los comentarios como un conjunto

Cuadro 5.3: Resumen de clasificaciones de comentarios (final).

Emoción	Ausente	Presente
Alegría	1167	728
Enojo	1698	197
Miedo	1867	28
Tristeza	1713	182
Total	6445	1135

de palabras sin un orden aparente. Este método es uno de los más simples — aunque comúnmente utilizados — en el área de clasificación de sentimiento o minería de opinión; en ese sentido, existen otros enfoques que utilizan conocimiento (Wu *et al.*, 2015), así como reglas sintácticas y semánticas (Cambria *et al.*, 2014; Ren *et al.*, 2016). Sin embargo, debido a que en nuestro caso solo es necesario tener una precisión aceptable en la clasificación y no es el problema principal a resolver, hemos elegido utilizar este enfoque. La creación de los clasificadores se puede ver como un proceso de tres etapas:

1. Selección de comentarios con sus etiquetas de emoción.
2. Limpieza del texto en los comentarios.
3. Entrenamiento de los clasificadores automáticos.

Para la creación de los clasificadores automáticos, se utilizaron las máquinas de soporte vectorial que se incluyen en WEKA² (Hall *et al.*, 2009), la cual es una herramienta de software gratuita que incorpora implementaciones de diversos algoritmos de minería de datos. Antes de obtener los vectores de características, se hizo limpieza de los comentarios con el fin de evitar agregar ruido (características innecesarias) durante el entrenamiento. En este proceso se incluyó una lista de *palabras vacías* (en inglés, *stopwords*), las cuales son palabras que en sí no tienen relevancia. La

²Disponible en <https://www.cs.waikato.ac.nz/ml/weka/>.

lista de palabras vacías en español fue extraída de la herramienta Freeling (Padró y Stanilovsky, 2012), la cual es una herramienta gratuita para el análisis de texto que permite trabajar con el idioma español. Con esto, evitamos procesar palabras como “la”, “el”, “donde”, “que”, etcétera, dado que no aportan mucha información sobre la emoción. Además, se utilizaron expresiones regulares e interpretación de emoticones para estandarizar y agrupar términos (y de esta manera concentrar el peso en una sola forma del término). Entre las expresiones regulares empleadas, por ejemplo, se encuentran “jajjjajaaa” o “jaaaja”, que se convirtieron a “jajaja”, de manera que se facilitara el aprendizaje de los clasificadores. Por otro lado, los emoticones tales como :), :D, :(y :(((también fueron asociados a formas más simples y constantes como :-)) y :-(. En el proceso de limpieza, de igual manera, se eliminaron direcciones de sitios, correos electrónicos y teléfonos. Asimismo, algunos símbolos tales como las comillas (“”) y la coma (,) fueron cambiados por un espacio en blanco.

Para cada emoción se implementó una máquina de soporte vectorial. Se utilizó un 70 % de las instancias (ejemplos) para el conjunto de entrenamiento y el 30 % restante para el conjunto de pruebas. Para tener la misma cantidad de instancias en ambas clases (es decir, equilibrar las clases *ausente* y *presente*), se añadieron instancias de la clase *ausente* elegidas de forma aleatoria en una cantidad igual a las instancias de la clase *presente*. Cada palabra encontrada en los comentarios correspondía a una característica y la frecuencia de dicha palabra en el comentario correspondía al peso de la característica. Los comentarios se pasaron al formato ARFF utilizado por WEKA, lo que hizo posible su interpretación por parte de la herramienta.

5.2.2 RESULTADOS

Los resultados de cada clasificador (alegría, tristeza y enojo) se evaluaron utilizando las métricas de precisión (ecuación 2.8, página 31), exhaustividad (ecuación 2.9, página 31) y valor-F (ecuación 2.10, página 31) descritas en el capítulo 2. Estas

métricas se muestran en el cuadro 5.4. Podemos ver que las precisiones se encuentran por encima del 70% en la mayoría de los casos, excepto en la emoción del enojo, donde es del 66% para la clase *ausente*. En trabajos como el de Mogadala y Varma (2012), reportan una precisión de aproximadamente 71% con un valor-F de aproximadamente 65% al clasificar emociones en texto utilizando etiquetas de emociones como *alegría* y *felicidad*. Nuestros clasificadores arrojan un mejor valor-F para las emociones, lo cual nos indica que son aptos para ser utilizados en los demás experimentos.

Cuadro 5.4: Evaluación de clasificadores de emoción

Emoción	Clase	Precisión	Exhaustividad	Valor-F
ALEGRÍA	Ausente	0.707	0.815	0.757
	Presente	0.786	0.668	0.722
	Promedio	0.747	0.741	0.74
ENOJO	Ausente	0.667	0.807	0.73
	Presente	0.776	0.623	0.691
	Promedio	0.723	0.712	0.71
TRISTEZA	Ausente	0.771	0.673	0.718
	Presente	0.705	0.796	0.748
	Promedio	0.738	0.734	0.733

5.2.3 DISCUSIÓN

Como se puede apreciar en los resultados, los clasificadores para cada emoción básica se comportan casi igual. Esto resulta interesante, ya que la cantidad de comentarios etiquetados para enojo y tristeza abarcan (respectivamente) alrededor del 20% de los comentarios que fueron etiquetados para alegría (cuadro 5.3) — es decir, los comentarios para esta última emoción representan una cantidad considerablemente mayor. En ese sentido, los clasificadores para enojo y tristeza obtuvieron una

calidad equiparable con una menor cantidad de comentarios para entrenamiento. Razones posibles para este resultado incluyen la limpieza de los comentarios, el uso del acuerdo de intercorrelación y el balanceo de clases.

Ahora bien, algunas implicaciones de este resultado incluyen consideraciones en cuanto al esfuerzo manual que se debe aplicar para el etiquetado de comentarios (realmente qué tan grandes deben ser los conjuntos de entrenamiento), la calidad del preprocesamiento de los datos y la búsqueda de aumento de la calidad en los clasificadores a través de las técnicas de aprendizaje (es decir, diferentes tipos de clasificadores o vectores de características) en lugar de hacerlo mediante el aumento de los conjuntos de entrenamiento. En cuanto a esto último, pensamos que utilizar otro tipo de enfoque para obtener las características de los comentarios previo al entrenamiento podría mejorar los resultados. Por ejemplo, en el trabajo con SenticNET (Cambria *et al.*, 2014), las precisiones reportadas son de 91.4 % y 86.9 % para clasificaciones de comentarios como positivos y negativos respectivamente con una exhaustividad de 81.2 % y 74.3 %. El valor-F para ambas categorías es de 85.9 % y 80.1 %, respectivamente. Utilizar este tipo de métodos podría verse reflejado en una mejor predicción de la intensidad emocional.

Vimos que en general nuestros clasificadores arrojaron mejores valores-F que otros trabajos recientes, lo cual es interesante ya que nosotros hemos utilizado un método simple. Atribuimos este resultado a que los comentarios pertenecen a una misma región geográfica por lo que presentan los mismos modismos y uso coloquial del idioma, a que los evaluadores trabajaron en su idioma natal y al proceso de limpieza de los comentarios. En ese sentido, también consideramos que esta contribución a la detección de emociones en español puede abrir las puertas a otros desarrollos, en especial si se publican los datos utilizados.

En general, y como conclusión principal de estos primeros resultados, observamos un nivel de calidad suficiente para continuar con la siguiente fase: predicción de intensidades emocionales. Desafortunadamente, no pudimos realizar el entrena-

miento en el caso de la emoción del miedo, por lo que incluirla quedará para trabajo futuro.

5.3 PREDICCIÓN DE INTENSIDAD EMOCIONAL

El objetivo del modelo es la predicción de la intensidad emocional (ver ecuación 4.5, página 67). Como se explicó en el capítulo 4, el modelo toma información a partir de (1) el usuario, (2) sus contactos y (3) el ambiente. Sin embargo, existen también un par de parámetros que han de definirse previamente para hacer una predicción. Estos son los parámetros k , que representa el periodo histórico con que se cuenta y el parámetro p , que representa la ventana de tiempo a predecir. Esta puede ser la ventana de tiempo siguiente o alguna otra más a futuro (es decir, cuando existe holgura).

El proceso para llevar a cabo la predicción se detalla en las subsecciones siguientes.

5.3.1 CONFIGURACIÓN

Esta sección se divide en dos: (1) preparación de instancias de prueba y elección de parámetros k y p y (2) método de validación del modelo. En la primer subsección se explica la selección y preprocesamiento de los datos para ser utilizados por el modelo. En la segunda subsección se explica la manera en que será evaluado el modelo. El cuadro 5.5 muestra un resumen breve de estas configuraciones las cuales se explican con más detalle posteriormente.

Cuadro 5.5: Resumen de configuraciones. Muestra a grandes rasgos las diferentes configuraciones utilizadas en la experimentación con relación al modelo propuesto. Se dividen en dos partes principales: 1) parámetros y 2) modelos. La primera parte hace referencia a la variación que se hace sobre los parámetros globales como el tamaño de la ventana de tiempo y el número de ventana futura que se va a predecir. La segunda se refiere a los distintos modelos contra los que será comparado el modelo propuesto.

Tipo	Descripción
Parámetros	
Tamaño de ventana de tiempo	Se especifican diferentes tamaños para la experimentación.
Parámetro p	Es la ventana de tiempo futura que se va a predecir.
Modelos	
Modelos de referencia	Son modelos que incluyen parcialmente propiedades del modelo propuesto completo.
Modelo comparativo	Es un modelo en el estado del arte que se compara con el modelo propuesto.
Modelo propuesto	Es el modelo final evaluado.

5.3.1.1 PREPARACIÓN DE INSTANCIAS DE PRUEBA

A partir del conjunto de datos extraído de Twitter, se obtuvo una muestra de 5,000 usuarios elegidos al azar de entre los usuarios disponibles en el repositorio. Los usuarios tienen la característica de especificar su idioma como español — como se ha descrito anteriormente — así como su ubicación en el área metropolitana del estado de Nuevo León. De cada usuario se obtuvieron sus últimos mil comentarios publicados más recientes. Debido a la distribución de los comentarios, se eligió el periodo de cinco meses que va desde julio de 2011 a noviembre de este mismo año, ya que es el periodo con más actividad presente en el conjunto de datos contando con el 70% de los comentarios.

Una de las primeras variables que definimos es la interpretación de la ventana de tiempo. Hasta donde sabemos, no hay una granularidad o tipo de unidad idónea para predecir las emociones. Algunos trabajos han utilizado una semana, mientras que otros han utilizado días o eventos puntuales. En esta tesis, se utilizaron tres distintas ventanas de tiempo:

- Un día

- Una hora
- Un bloque de 15 minutos

Esto se llevó a cabo con el objetivo de identificar diferencias debido a la ventana de tiempo utilizada y requiere que se lleven a cabo experimentos independientes para cada ventana con diferente granularidad. En cada caso y para cada variable del modelo se obtuvo la intensidad emocional correspondiente de acuerdo con la ecuación 4.1 (página 64). Es decir, se obtuvieron los valores de intensidad emocional necesarios para obtener \mathbf{U} , \mathbf{A}_f , \mathbf{A}_m y \mathbf{E} para cada uno de los usuarios en el repositorio. Antes de obtener las matrices definitivas, fue necesario definir el valor k , es decir, el tamaño del periodo histórico a tomar en cuenta para cada variable. En otros trabajos se ha identificado que existen patrones en Twitter que se repiten semanalmente (Dodds *et al.*, 2011). Por tanto, se escogió $k = 7$.

En cuanto al parámetro p , es importante ya que se refiere a la ventana de tiempo cuyo valor de intensidad emocional será predicho. Utilizamos $p = 1$, es decir, predijimos la ventana de tiempo siguiente al periodo histórico, pues consideramos que debía ser el primer valor a probar. Aunque se podían utilizar otros valores, tales como $p = 2, 3, \dots$, consideramos que — por lo pronto — estos otros valores quedan fuera del alcance de esta tesis.

Considerando nuestra muestra de usuarios, así como k , p y en general la información disponible en el repositorio, se extrajeron instancias para validar el modelo según las diferentes granularidades para las ventanas de tiempo, donde cada instancia consiste propiamente en las matrices \mathbf{U} , \mathbf{A}_f , \mathbf{A}_m y \mathbf{E} para el periodo histórico de tamaño $k + p$. Para la ventana de un día, fue posible extraer 80,541 instancias, mientras que para la de una hora se obtuvieron 35,487 y para el bloque de 15 minutos se obtuvieron 8,980. Cada instancia cumple la condición de tener k periodos previos como información histórica y, por lo tanto, a medida que se hacen más cortos los intervalos de tiempo la cantidad de instancias disminuye. Esto es porque es más difícil que un usuario continúe publicando comentarios de una manera tan frecuente.

5.3.1.2 COMPARACIÓN CON OTROS MODELOS

Para validar el modelo propuesto, se comparó contra los modelos que a continuación se listan. Algunos de estos modelos se pueden considerar *puntos de base* o referencia (*baselines*), otros son métodos clásicos (por ejemplo, ARIMA) y otros consisten en modelos parciales que se basan fuertemente en el propuesto. La comparación de modelos no es una tarea trivial, pues depende de la complejidad computacional del modelo y la calidad. Nos enfocamos en este segundo aspecto al evaluar el parecido de la solución entregada con respecto a la realidad.

Las dos métricas escogidas para hacer la evaluación (mismas que son comunes para predicción y regresión, en general) son el error absoluto medio (MAE, por sus siglas en inglés), descrito en la ecuación 2.11 (página 32) y la correlación entre valores reales y valores predichos, descrita en la ecuación 2.12 (página 32). A medida que el MAE se acerque a cero, tendremos correlaciones más cercanas a 1 y por lo tanto podríamos utilizar solamente el valor MAE como una medida para comparar los modelos; sin embargo, la correlación nos da una visión de la dependencia lineal que existe entre estas variables (Crawford, 2006).

Para saber si uno de los modelos examinados genera valores MAE más bajos que otro, es necesario comparar su media estadística. Para lograr esto, se realizaron 300 corridas, donde cada corrida consistió en tomar el total de instancias y dividirlo en dos: el conjunto de entrenamiento y el conjunto de prueba (cada instancia quedando aleatoriamente en alguno de los dos conjuntos para cada corrida). El conjunto de entrenamiento abarcó el 70 % de las instancias y el de prueba el 30 % restante en cada corrida. Esta cantidad de corridas se escogió para poder obtener una muestra suficientemente grande de valores MAE que permitiera llevar a cabo una prueba *t-student* de validez estadística y así concluir que el modelo propuesto genera (o no genera) valores significativamente menores para MAE en comparación con otros. En general, se trabajó en el nivel de significancia de $\alpha = 0.05$.

MODELOS UTILIZADOS PARA LA COMPARACIÓN

Aleatorio: asume que la intensidad emocional del usuario es una variable aleatoria con una función de densidad uniforme cuyos valores se encuentran en el rango $[0, 1]$.

Promedio global: asume que la intensidad emocional del usuario en una ventana de tiempo futura es igual al promedio de la emoción de todos los usuarios en el conjunto de datos en una ventana de tiempo anterior, donde $|U|$ es la cantidad de usuarios. Es decir,

$$I_{u,e}(t+p) = \frac{\sum_u I_{u,e}(t)}{|U|}. \quad (5.1)$$

Promedio de n ventanas anteriores: asume que la intensidad emocional del usuario en una ventana de tiempo futura es igual al promedio de sus emociones registradas hasta n ventanas de tiempo anteriores. Es decir,

$$I_{u,e}(t+p) = \sum_{t-n+1}^t \frac{I_{u,e}(t)}{n}. \quad (5.2)$$

ARIMA: asume que la intensidad emocional es una serie de tiempo a la cual se puede adaptar un modelo ARIMA para obtener el siguiente valor. Es decir,

$$I_{u,e}(t+p) = \text{ARIMA}(p, d, q), \quad (5.3)$$

donde p , d y q son las componentes autorregresiva, integrada y de media móvil, respectivamente.

Una sola emoción: asume que la intensidad emocional del usuario en una ventana de tiempo futura se puede obtener a partir de aplicar un modelo de regresión lineal a los datos históricos de la intensidad emocional en particular. Es decir, existen unas constantes $\psi_1, \psi_2, \dots, \psi_n$ tales que:

$$I_{u,e}(t+p) = \psi_1 I_{u,e}(t) + \psi_2 I_{u,e}(t-1) + \dots + \psi_n I_{u,e}(t-n+1). \quad (5.4)$$

Alegría y Enojo: asume que la intensidad emocional del usuario en una ventana de tiempo futura se puede obtener a partir de aplicar un modelo de regresión lineal utilizando los datos históricos de las intensidades emocionales de alegría y enojo. Es

decir, existen unas constantes $\psi_1, \psi_2, \dots, \psi_n$ y $\phi_1, \phi_2, \dots, \phi_n$ tales que:

$$\begin{aligned} I_{u,e}(t+p) = & \psi_1 I_{u,A}(t) + \psi_2 I_{u,A}(t-1) + \dots + \psi_n I_{u,A}(t-n+1) + \\ & \phi_1 I_{u,E}(t) + \phi_2 I_{u,E}(t-1) + \dots + \phi_n I_{u,E}(t-n+1), \end{aligned} \quad (5.5)$$

donde A y E denotan las emociones de alegría y enojo respectivamente. Por lo tanto $I_{u,A}(t)$ y $I_{u,E}(t)$ son sus respectivas fórmulas para la intensidad emocional.

Alegría y Tristeza: asume que la intensidad emocional del usuario en una ventana de tiempo futura se puede obtener a partir de aplicar un modelo de regresión lineal utilizando los datos históricos de las intensidades emocionales de alegría y tristeza. Es decir, existen unas constantes $\psi_1, \psi_2, \dots, \psi_n$ y $\phi_1, \phi_2, \dots, \phi_n$ tales que:

$$\begin{aligned} I_{u,e}(t+p) = & \psi_1 I_{u,A}(t) + \psi_2 I_{u,A}(t-1) + \dots + \psi_n I_{u,A}(t-n+1) + \\ & \phi_1 I_{u,T}(t) + \phi_2 I_{u,T}(t-1) + \dots + \phi_n I_{u,T}(t-n+1), \end{aligned} \quad (5.6)$$

donde A y T denotan las emociones de alegría y tristeza respectivamente. Por lo tanto $I_{A,u}(t)$ y $I_{T,u}(t)$ son sus respectivas formulas para la intensidad emocional.

Enojo y Tristeza: asume que la intensidad emocional del usuario en una ventana de tiempo futura se puede obtener a partir de aplicar un modelo de regresión lineal utilizando los datos históricos de las intensidades emocionales de enojo y tristeza. Es decir, existen unas constantes $\psi_1, \psi_2, \dots, \psi_n$ y $\phi_1, \phi_2, \dots, \phi_n$ tales que:

$$\begin{aligned} I_{u,e}(t+p) = & \psi_1 I_{u,E}(t) + \psi_2 I_{u,E}(t-1) + \dots + \psi_n I_{u,E}(t-n+1) + \\ & \phi_1 I_{u,T}(t) + \phi_2 I_{u,T}(t-1) + \dots + \phi_n I_{u,T}(t-n+1), \end{aligned} \quad (5.7)$$

donde E y T denotan las emociones de alegría y tristeza respectivamente. Por lo tanto $I_{u,E}(t)$ y $I_{u,T}(t)$ son sus respectivas fórmulas para la intensidad emocional.

Todas: asume que la intensidad emocional del usuario en una ventana de tiempo futura se puede obtener a partir de aplicar un modelo de regresión lineal utilizando los datos históricos de las intensidades emocionales de todas sus emociones. Es decir, existen unas constantes $\psi_1, \psi_2, \dots, \psi_n, \phi_1, \phi_2, \dots, \phi_n$ y $\rho_1, \rho_2, \dots, \rho_n$ tales que:

$$\begin{aligned} I_{u,e}(t+p) = & \psi_1 I_{u,A}(t) + \psi_2 I_{u,A}(t-1) + \dots + \psi_n I_{u,A}(t-n+1) + \\ & \phi_1 I_{u,E}(t) + \phi_2 I_{u,E}(t-1) + \dots + \phi_n I_{u,E}(t-n+1) + \\ & \rho_1 I_{u,T}(t) + \rho_2 I_{u,T}(t-1) + \dots + \rho_n I_{u,T}(t-n+1), \end{aligned} \quad (5.8)$$

donde A , E y T denotan las emociones de alegría, enojo y tristeza respectivamente. Por lo tanto $I_{u,A}(t)$, $I_{u,E}(t)$ y $I_{u,T}(t)$ son sus respectivas fórmulas para la intensidad emocional. Este modelo en particular comienza a parecerse más al modelo propuesto y puede escribirse de la siguiente manera:

$$I_{u,e}(t+p) = f(\mathbf{U}), \quad (5.9)$$

donde \mathbf{U} es la matriz de intensidades emocionales del usuario definida en la ecuación 4.5 y f es la función de predicción definida en la sección 4.4, que en este caso representa una regresión lineal.

Todas + Día Semana: este modelo es similar al anterior, pero además agrega una matriz de constantes de tamaño 1×7 , la cual es utilizada en una operación de producto punto que multiplica a la matriz \mathbf{E} definida en la ecuación 4.5. Se puede interpretar como una constante que se suma al modelo anterior y cuyo valor depende únicamente del día de la semana al que pertenece la ventana de tiempo cuya intensidad emocional se desea predecir. Asimismo, este modelo podemos escribirlo como

$$I_{u,e}(t+p) = f(\mathbf{U}, \mathbf{E}), \quad (5.10)$$

donde \mathbf{U} es la matriz de intensidades emocionales del usuario y \mathbf{E} es la matriz de entorno; ambas matrices fueron definidas en la ecuación 4.5. Al igual que en el modelo anterior, f es implementada a través de una regresión lineal.

Todas + Contagio: asume que la intensidad emocional del usuario en una ventana de tiempo futura se puede obtener a partir de aplicar un modelo de regresión lineal (denotado por f) utilizando los datos históricos de las intensidades emocionales de sus contactos y de las menciones recibidas. Este modelo es similar al modelo **Todas**, salvo que además agrega constantes para cada una de las componentes de las matrices de intensidad emocional provenientes de los contactos del usuario. En la ecuación 4.5 se definen las matrices \mathbf{A}_f y \mathbf{A}_m . Este modelo particular puede escribirse como

$$I_{u,e}(t+p) = f(\mathbf{U}, \mathbf{A}_f, \mathbf{A}_m), \quad (5.11)$$

donde las matrices \mathbf{U} , \mathbf{A}_f y \mathbf{A}_m son las matrices de intensidades emocionales del usuario, contactos y menciones definidas en la ecuación 4.5.

Modelo Propuesto: este modelo incluye toda la información disponible del usuario, sus contactos y el entorno. Puede verse como una combinación de los modelos **Todas + Día Semana** y **Todas + Contagio**. En este modelo, f puede representar una regresión lineal u otro método. El modelo propuesto podemos expresarlo como

$$I_{u,e}(t+p) = f(\mathbf{U}, \mathbf{A}_f, \mathbf{A}_m, \mathbf{E}). \quad (5.12)$$

Asimismo, este modelo está definido con más detalle en la sección 4.3.

5.3.1.3 COMPARACIÓN CON OTROS TRABAJOS

Como referencia, el modelo propuesto puede ser comparado con el trabajo de Mogadala y Varma (2012), descrito en la sección 3.2 del capítulo 3. En este trabajo, se busca predecir una “transición” en la emoción de los comentarios. La transición representa un cambio positivo o negativo (numéricamente) en la emoción del usuario, es decir, se vuelve más positivo o más negativo que antes. Se utilizan emociones abstractas para etiquetar los comentarios y una escala numérica. A través de máquinas de soporte vectorial y regresión lineal, llevan a cabo una predicción de dicha transición. Para evaluar, se emplean métricas clásicas como la correlación y el error absoluto medio, por lo que con unas pequeñas modificaciones en la implementación de nuestro modelo, podemos comparar los resultados con los de este trabajo. Para hacerlo, en vez de predecir la intensidad emocional del usuario, lo que hacemos es predecir el cambio en la intensidad emocional correspondiente a la ventana de tiempo que se desea predecir y la actual:

$$\Delta(t, t+p) = I_{u,e}(t+p) - I_{u,e}(t). \quad (5.13)$$

Como ya hemos mencionado anteriormente, manejamos el valor $p = 1$, ya que predecimos solo la ventana de tiempo siguiente a la actual. Al igual que en los casos

anteriores, usamos un 70 % del conjunto de instancias para el entrenamiento de los modelos y el 30 % restante para las pruebas y obtención de valores de error. Para obtener el modelo final, al igual que Mogadala y Varma (2012), usamos tanto regresión lineal como máquinas de soporte vectorial. Para estas pruebas, utilizamos conjuntos de 30,000 instancias seleccionadas al azar a partir del conjunto para la ventana de tiempo de un día para cada una de las emociones básicas (exceptuando el miedo). Como medidas para cuantificar el error, consideramos las mismas que el trabajo mencionado: correlación (ecuación 2.12, página 32), error absoluto medio (ecuación 2.11, página 32), raíz del error cuadrático medio (ecuación 2.13, página 32), error absoluto relativo (ecuación 2.14, página 33) y raíz del error relativo cuadrático (ecuación 2.15, página 33).

5.3.2 RESULTADOS

Los cuadros 5.6 y 5.7 muestran los resultados de la comparación del modelo propuesto contra los modelos descritos en la sección 5.3.1.2; se pueden apreciar los valores MAE y COR respectivamente para los modelos analizados. Como hemos mencionado en la sección anterior, se hicieron 300 iteraciones en todos los modelos con el fin de comparar estadísticamente los valores MAE a través de pruebas *t*-student. Todos los valores fueron comparados con el resultado del modelo propuesto en el nivel de significancia $\alpha = 0.05$. Vemos que solamente en algunos casos donde la ventana de tiempo es de una hora o bloques de 15 minutos no se generan valores MAE diferentes estadísticamente con respecto al modelo propuesto, lo cual nos hace sentido ya que la predicción se hace con datos que por lo general ocurren dentro del mismo día. Estos valores que no son estadísticamente diferentes con respecto al modelo propuesto están encerrados entre paréntesis. Los valores MAE mínimos y COR máximos están marcados con texto en negritas para facilitar la lectura.

La figura 5.2 muestra de forma visual una comparativa del error absoluto medio para los distintos modelos probados usando las ventanas con los tres tipos de

granularidad. Notamos que, para la ventana de tiempo de un día, el error absoluto medio alcanza un mínimo cuando se utilizan siete días para obtener el promedio, mientras que para las ventanas de tiempo de una hora y bloques de 15 minutos — a medida que se incrementa el número de ventanas para obtener el promedio — el error absoluto medio aumenta también. Asimismo, vemos que el modelo ARIMA no resultó ser mejor que los modelos base obtenidos mediante el promedio de varias ventanas de tiempo, e incluso resulta ser comparable al modelo promedio de siete ventanas de tiempo.

La figura 5.3 muestra de forma visual una comparativa de la correlación de los distintos modelos probados en los tres tipos ventana (un día, una hora y bloques de 15 minutos). Notamos que, para la ventana de tiempo de un día, la correlación alcanza un máximo cuando se utilizan siete días para obtener el promedio, mientras que para las ventanas de tiempo de una hora y bloques de 15 minutos — a medida que se incrementa el número de ventanas para obtener el promedio — la correlación disminuye. Asimismo, vemos que el modelo ARIMA no resultó ser mejor que los modelos base obtenidos mediante el promedio de varias ventanas de tiempo.

La figura 5.4 resume los resultados de la comparación del modelo propuesto contra el resto de los modelos considerados. Como podemos apreciar, salvo algunos modelos parciales, todos los modelos son superados con diferencias estadísticamente significativas — tanto en correlación como en error absoluto medio. Con respecto a los modelos parciales que incorporan todas las emociones en el periodo histórico (Todas, Todas + Contagio, Todas + Día Semana), aunque las diferencias no son significativas, estos también son superados.

En la sección 5.3.1.3 se mencionó que se compararían los resultados con otros trabajos relacionados. Con respecto a la comparación con el trabajo de Mogadala y Varma (2012), el cuadro 5.8 muestra los resultados obtenidos al predecir la transición de la emoción utilizando el modelo propuesto y desglosado por emoción básica, mientras que el cuadro 5.9 se hace propiamente la comparación, promediando para

ello los resultados que obtuvo nuestro modelo propuesto. Como se mencionó anteriormente, la transición indica un cambio de la emoción en el tiempo cuyo valor está dado por la ecuación 5.13. Vemos que el modelo propuesto puede predecir la transición con un $RAE = 68.82\%$ como valor más bajo, mientras que en el trabajo de Mogadala y Varma (2012) se reporta un $RAE = 98.36\%$, lo cual representa una mejora de más de un 40% a favor del modelo propuesto.

Cuadro 5.6: Comparación de modelos utilizando error absoluto medio. Se comprobó en el nivel de significancia $\alpha = 0.05$ que los valores fueran estadísticamente diferentes con respecto al valor del modelo propuesto. Los mejores valores están en negritas; si además tienen un asterisco, la diferencia es estadísticamente significativa. AL = Aleatorio, GB = Global, P_i = promedio utilizando i ventanas de tiempo, CG = Contagio, DS = Día Semana, ARI = ARIMA, S = Una sola emoción, AE = Alegría y Enojo, AT = Alegría y Tristeza, TE = Tristeza y Enojo, T = Todas, TDS = Todas + Día Semana, TC = Todas + Contagio, MP = Modelo Propuesto.

	Un día			Una hora			15 min.		
	ALEGRÍA	ENOJO	TRISTEZA	ALEGRÍA	ENOJO	TRISTEZA	ALEGRÍA	ENOJO	TRISTEZA
AL	0.333	0.336	0.330	0.391	0.394	0.389	0.409	0.409	0.407
GB	0.224	0.219	0.225	0.288	0.284	0.300	0.309	0.309	0.324
P_1	0.273	0.269	0.271	0.318	0.311	0.327	0.331	0.324	0.345
P_2	0.245	0.242	0.243	0.344	0.335	0.340	0.356	0.346	0.357
P_3	0.236	0.234	0.233	0.366	0.359	0.359	0.378	0.367	0.373
P_4	0.231	0.230	0.229	0.384	0.378	0.374	0.395	0.384	0.386
P_5	0.229	0.228	0.227	0.399	0.394	0.387	0.410	0.398	0.397
P_6	0.228	0.226	0.226	0.411	0.408	0.399	0.422	0.410	0.408
P_7	0.228	0.225	0.225	0.422	0.420	0.409	0.433	0.421	0.417
P_8	0.228	0.226	0.225	0.432	0.431	0.418	0.442	0.430	0.425
P_9	0.230	0.228	0.227	0.441	0.441	0.426	0.452	0.439	0.433
P_{10}	0.231	0.229	0.228	0.447	0.449	0.433	0.459	0.447	0.439
P_{11}	0.233	0.231	0.229	0.453	0.456	0.438	0.467	0.454	0.446
P_{12}	0.235	0.233	0.231	0.458	0.462	0.443	0.473	0.461	0.452
P_{13}	0.236	0.235	0.232	0.462	0.467	0.448	0.479	0.467	0.457
P_{14}	0.238	0.236	0.234	0.466	0.472	0.452	0.484	0.473	0.462
P_{15}	0.239	0.238	0.235	0.470	0.476	0.456	0.489	0.478	0.466
ARI	0.228	0.225	0.226	0.422	0.415	0.412	0.432	0.423	0.426
CG	0.208	0.201	0.214	0.281	0.276	0.295	0.297	0.299	0.317
DS	0.210	0.203	0.214	0.281	0.277	0.295	0.299	0.300	0.318
S	0.198	0.191	0.199	0.273	0.267	0.283	0.291	0.289	0.306
AE	0.195	0.189	0.204	0.272	0.265	0.285	0.289	0.287	0.309
AT	0.197	0.190	0.197	0.273	0.268	0.282	0.290	0.291	0.305
TE	0.203	0.191	0.198	0.278	0.266	0.282	0.296	0.288	0.306
T	<i>(0.195)</i>	<i>(0.188)</i>	<i>(0.197)</i>	<i>(0.272)</i>	<i>(0.265)</i>	<i>(0.281)</i>	<i>(0.289)</i>	<i>(0.287)</i>	<i>(0.305)</i>
TDS	<i>(0.195)</i>	<i>(0.188)</i>	<i>(0.197)</i>	<i>(0.272)</i>	<i>(0.265)</i>	<i>(0.281)</i>	<i>(0.289)</i>	<i>(0.287)</i>	<i>(0.305)</i>
TC	<i>(0.194)</i>	<i>(0.188)</i>	<i>(0.197)</i>	<i>(0.270)</i>	<i>(0.262)</i>	<i>(0.279)</i>	<i>(0.287)</i>	<i>(0.285)</i>	<i>(0.303)</i>
MP	0.194*	0.187*	0.197*	0.270*	0.262*	0.279*	0.287*	0.285*	0.303*

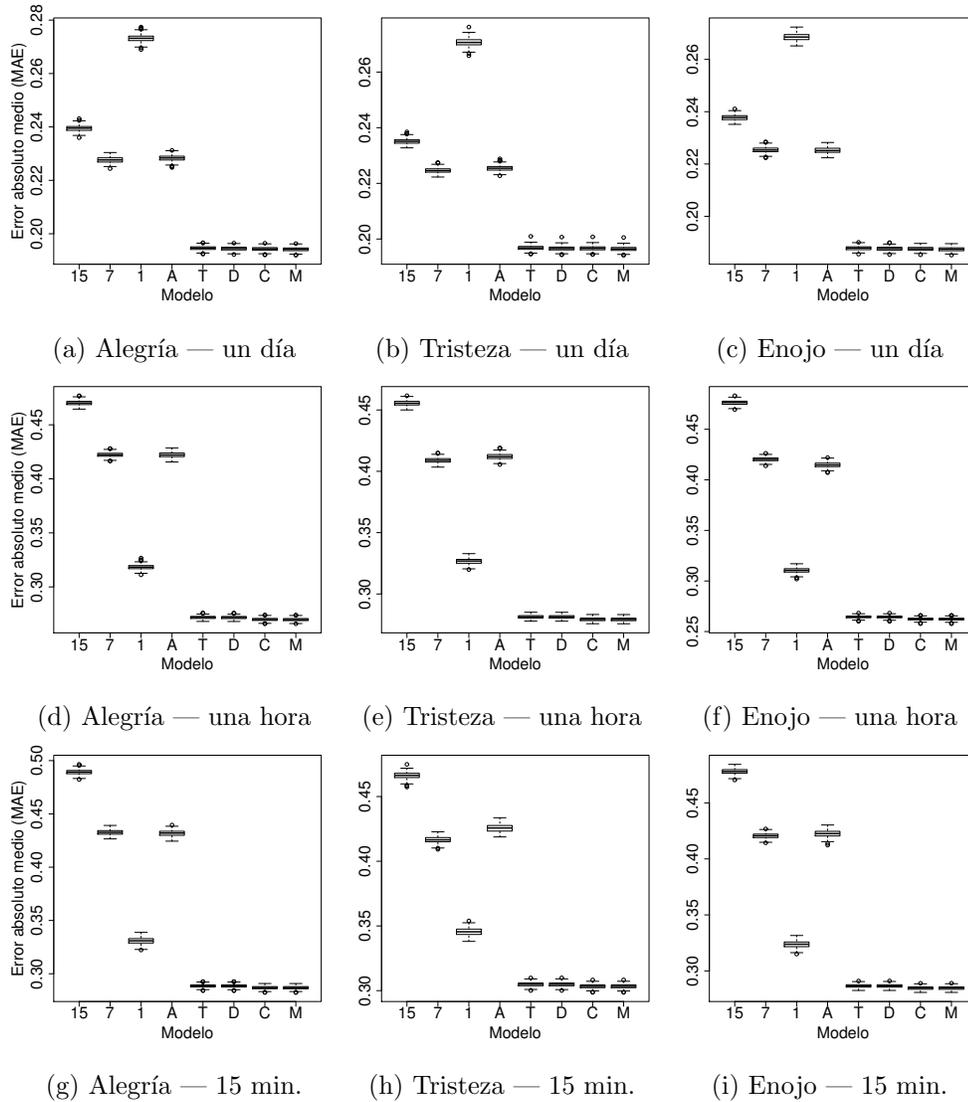


Figura 5.2: Error absoluto medio de los modelos más relevantes. El modelo “1” representa el modelo promedio de 1 día, “7” representa el modelo promedio de 7 días, “15” es el modelo promedio de 15 días, “A” representa al modelo ARIMA, “T” representa al modelo “Todas”, “D” representa al modelo “Todas + Día Semana”, “C” representa al modelo “Todas + Contagio” y “M” representa al modelo propuesto que incluye información del usuario, de los contactos y del entorno.

Cuadro 5.7: Comparación de modelos utilizando correlación. Se comprobó en el nivel de significancia $\alpha = 0.05$ que los valores fueran estadísticamente diferentes con respecto al valor del modelo propuesto. Los mejores valores están en negritas; si además tienen un asterisco, la diferencia es estadísticamente significativa. AL = Aleatorio, GB = Global, P_i = promedio utilizando i ventanas de tiempo, CG = Contagio, DS = Día Semana, ARI = ARIMA, S = Una sola emoción, AE = Alegría y Enojo, AT = Alegría y Tristeza, TE = Tristeza y Enojo, T = Todas, TDS = Todas + Día Semana, TC = Todas + Contagio, MP = Modelo Propuesto.

	Un día			Una hora			15 min.		
	ALEGRÍA	ENOJO	TRISTEZA	ALEGRÍA	ENOJO	TRISTEZA	ALEGRÍA	ENOJO	TRISTEZA
AL	-0.0001	-0.00004	-0.0002	-0.0005	0.0007	0.0003	-0.0003	-0.0005	-0.0001
P_1	0.198	0.197	0.215	0.183	0.193	0.201	0.177	0.197	0.186
P_2	0.239	0.234	0.258	0.162	0.191	0.201	0.160	0.190	0.185
P_3	0.261	0.252	0.281	0.148	0.190	0.199	0.145	0.185	0.181
P_4	0.272	0.261	0.290	0.142	0.186	0.193	0.135	0.177	0.176
P_5	0.281	0.270	0.299	0.138	0.188	0.191	0.128	0.174	0.173
P_6	0.283	0.277	0.305	0.136	0.188	0.189	0.123	0.173	0.169
P_7	0.286	0.282	0.309	0.136	0.186	0.186	0.118	0.172	0.166
P_8	0.286	0.282	0.310	0.134	0.184	0.183	0.113	0.168	0.162
P_9	0.285	0.282	0.310	0.135	0.182	0.181	0.108	0.168	0.161
P_{10}	0.284	0.280	0.309	0.136	0.180	0.181	0.105	0.168	0.159
P_{11}	0.282	0.278	0.307	0.138	0.179	0.179	0.102	0.167	0.158
P_{12}	0.280	0.276	0.306	0.141	0.178	0.178	0.100	0.166	0.157
P_{13}	0.278	0.275	0.304	0.144	0.178	0.178	0.099	0.165	0.154
P_{14}	0.276	0.274	0.304	0.148	0.179	0.179	0.098	0.165	0.154
P_{15}	0.274	0.273	0.302	0.151	0.181	0.182	0.098	0.163	0.153
ARI	0.272	0.266	0.298	0.123	0.158	0.155	0.113	0.145	0.135
CG	0.132	0.123	0.071	0.065	0.078	0.055	0.077	0.076	0.034
DS	0.042	0.034	0.041	0.034	0.021	0.025	0.017	0.027	0.024
S	0.291	0.288	0.315	0.198	0.227	0.232	0.187	0.223	0.212
AE	0.322	0.316	0.253	0.214	0.239	0.206	0.208	0.238	0.184
AT	0.299	0.306	0.330	0.204	0.209	0.240	0.197	0.198	0.221
TE	0.216	0.289	0.320	0.128	0.231	0.240	0.114	0.228	0.217
T	<i>(0.327)</i>	<i>(0.329)</i>	<i>(0.334)</i>	<i>(0.215)</i>	<i>(0.246)</i>	<i>(0.245)</i>	<i>(0.210)</i>	<i>(0.243)</i>	<i>(0.224)</i>
TDS	<i>(0.328)</i>	<i>(0.330)</i>	<i>(0.336)</i>	<i>(0.216)</i>	<i>(0.246)</i>	<i>(0.245)</i>	<i>(0.210)</i>	<i>(0.243)</i>	<i>(0.224)</i>
TC	<i>(0.330)</i>	<i>(0.332)</i>	<i>(0.336)</i>	<i>(0.231)</i>	<i>(0.262)</i>	<i>(0.259)</i>	<i>(0.223)</i>	0.255	0.235
MD	0.332*	0.334*	0.338*	0.233*	0.262*	0.259*	0.223*	0.254	0.234

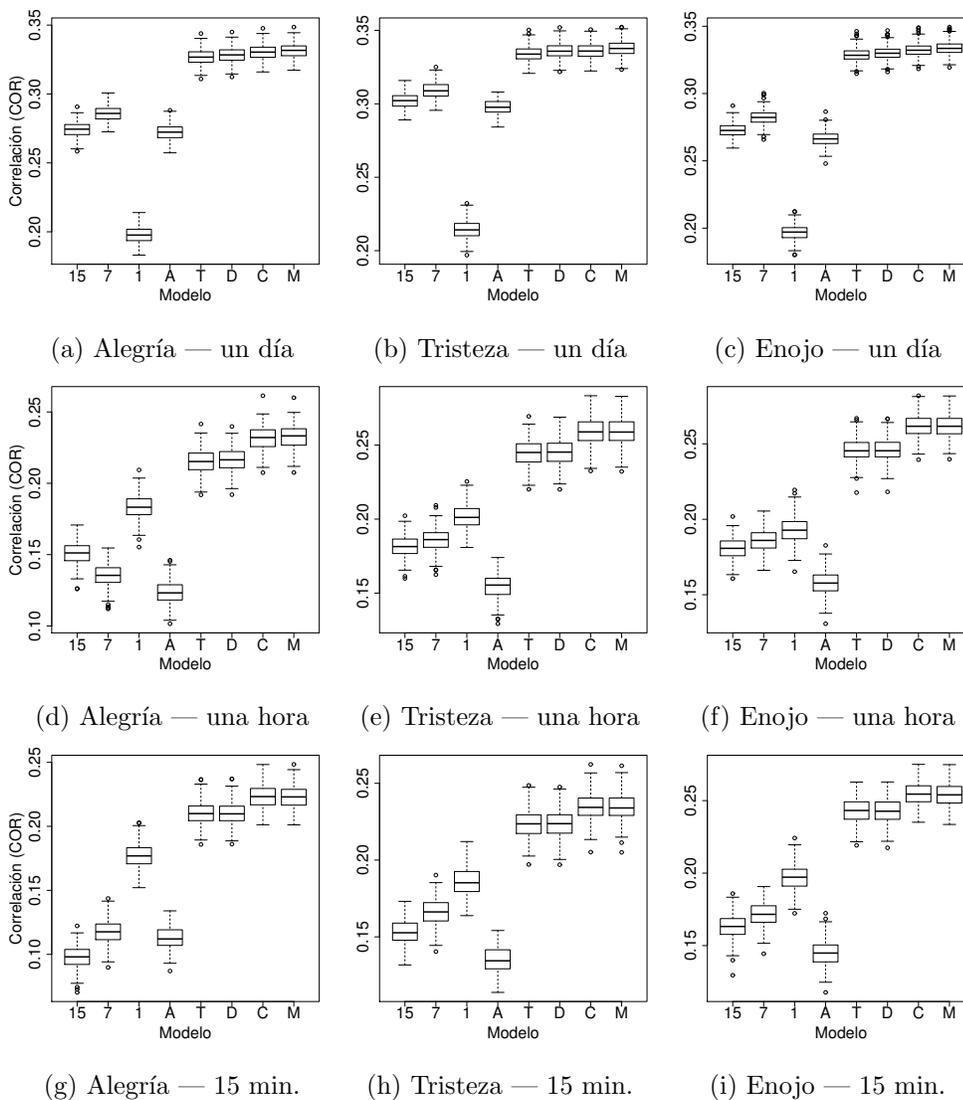


Figura 5.3: Correlación para los modelos más relevantes. El modelo “1” representa el modelo promedio de 1 día, “7” representa el modelo promedio de 7 días, “15” es el modelo promedio de 15 días, “A” representa al modelo ARIMA, “T” representa al modelo “Todas”, “D” representa al modelo “Todas + Día Semana”, “C” representa al modelo “Todas + Contagio”, “M” representa al modelo propuesto que incluye información del usuario, de los contactos y del entorno.

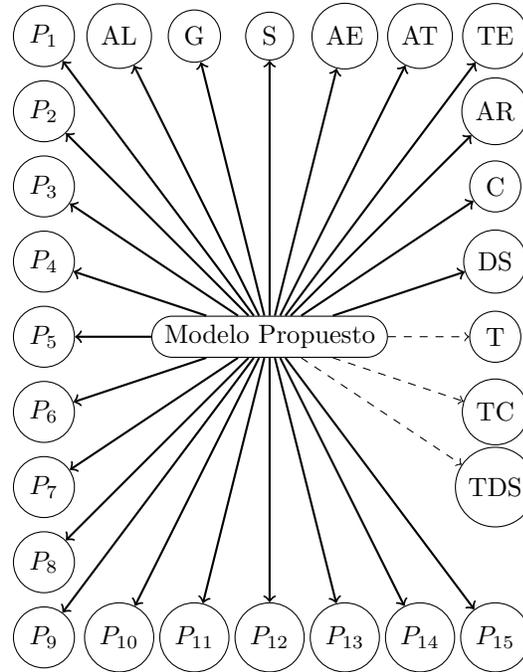


Figura 5.4: Resumen de comparación entre modelos. Una flecha saliente del modelo indica que este obtiene mejores resultados que aquel al cual entra la flecha. Una flecha gruesa indica diferencias estadísticamente significativas, mientras que una flecha punteada indica solamente que los resultados fueron superiores (sin ser estadísticamente significativos).

Cuadro 5.8: Resultados de predicción de transición con el modelo propuesto. Las siglas SVM hacen referencia al método de máquinas de soporte vectorial.

ALEGRÍA		COR	MAE	RMSE	RAE	RRSE
	Regresión lineal	0.7451	0.1940	0.2473	69.66 %	66.71 %
	SVM	0.7434	0.1937	0.2485	69.54 %	67.03 %
ENOJO						
	Regresión lineal	0.7493	0.1874	0.2396	68.86 %	66.23 %
	SVM	0.7474	0.1873	0.2404	68.82 %	66.46 %
TRISTEZA						
	Regresión lineal	0.7345	0.1938	0.2478	70.50 %	67.86 %
	SVM	0.7325	0.1932	0.2488	70.28 %	68.15 %

Cuadro 5.9: Resultados comparativos para predicción de transición emocional. Las siglas RL hacen referencia a regresión lineal, SVM hace referencia a máquinas de soporte vectorial y MV hace referencia al trabajo de Mogadala y Varma (2012).

Método Predicción	Trabajo	COR	MAE	RMSE	RAE	RRSE
RL	MV	-0.2518	1.9645	2.7469	100.00 %	100.00 %
	Propuesto	0.743	0.192	0.245	69.67 %	66.93 %
SVM	MV	-0.0787	1.9323	2.7276	98.36 %	99.29 %
	Propuesto	0.741	0.191	0.246	69.55 %	67.21 %

5.3.3 DISCUSIÓN

En cuanto a los resultados obtenidos, cabe destacar algunos datos sobresalientes. Por ejemplo, las emociones de enojo y tristeza fueron las que mejor correlación y error obtuvieron por parte del modelo propuesto (que fue el mejor); de hecho, las diferencias entre cada emoción son estadísticamente significativas ($\alpha = 0.05$) para ambas métricas en cada ventana de tiempo utilizada, por lo que el enojo es la emoción cuya intensidad obtiene las mejores predicciones. Según Brans y Verduyn (2014), tanto tristeza como enojo se encuentran dentro de las emociones negativas más intensas y duraderas, aunque sitúan a la tristeza por encima del enojo dentro de sus hallazgos. Pensamos que nuestros resultados, por otro lado, favorecen al enojo debido — por una parte — a la naturaleza de las redes sociales, en las cuales los sentimientos negativos tienden a esparcirse con mayor facilidad (Chmiel *et al.*, 2011b), y también creemos que puede deberse a algunos eventos que se dieron en el lapso de recolección del repositorio, pues estos incluyen, por ejemplo, el incendio del Casino Royale en Monterrey a manos del crimen organizado. Este evento generó indignación por parte de las personas, y estos patrones pueden ser apreciados en trabajos como el de Torres y Rodríguez (2013).

Aún cuando el enojo obtiene mejores predicciones en el modelo propuesto, una tendencia interesante se puede observar en los modelos P1-P15, donde el enojo comienza obteniendo los errores más bajos para los modelos que utilizan poca información histórica (P1-P3), pero esta ventaja desaparece cuando se incorpora una mayor cantidad de información (P4-P15), favoreciendo así a la tristeza. Esto concuerda ampliamente con los hallazgos de Verduyn y Lavrijsen (2015), los cuales establecen que la tristeza es una emoción más duradera que el enojo; en ese sentido, puede apreciarse esta disipación del enojo a través de una pérdida en la calidad de predicción de la intensidad emocional. No obstante, en nuestro caso el enojo parece disiparse más rápido (tres horas) que lo reportado por Verduyn y Lavrijsen (2015), que es seis horas; sin embargo, esto concuerda con lo reportado por Nguyen *et al.*

(2012). Consideramos que la coincidencia con estos trabajos valida de manera externa nuestros resultados y, por ende, nuestra metodología — tanto la fase de predicción como la de detección y el preprocesamiento de datos.

Los modelos base cuyo pronóstico se obtiene promediando intensidades emocionales históricas también muestran una tendencia en la que la calidad del modelo disminuye a medida que se agregan más ventanas de tiempo. Al utilizar ventanas de un día, el mejor modelo se obtiene cuando se utilizan siete — es decir, una semana. Lo anterior nos sugiere que pueden existir patrones repetitivos semanalmente en el comportamiento de las emociones dentro de las redes sociales. Sin embargo, cuando el tamaño de la ventana de tiempo es de horas o bloques de 15 minutos, a medida que aumenta el número de ventanas, la calidad del modelo disminuye (aumenta el error absoluto medio y disminuye la correlación).

Se puede observar en algunos de los resultados (no así en la mayoría) que, mientras el error se muestra bajo, la correlación también; parecieran no coincidir del todo. Esto sucede, por ejemplo, cuando se utiliza solo el día de la semana; en ese sentido, podríamos decir que este dato por sí solo resulta más informativo que los promedios de días anteriores (en cuanto a error absoluto medio), pero debido a que la correlación de dicho modelo es muy pobre, podríamos descartarlo para aplicaciones prácticas. De igual manera, podemos ver que la correlación para el modelo propuesto es más baja que el modelo de Todas + Contagio para enojo y tristeza en la ventana de tiempo de 15 minutos; lo atribuimos al hecho de que el modelo propuesto difiere de Todas + Contagio en la inclusión del día de la semana. Probablemente esta variable resulta irrelevante con ventanas de tiempo tan pequeñas, y — sobre todo — que transcurren dentro del mismo día. Más bien parece que esta variable genera ruido durante la regresión, aunque es poco el ruido porque se trata de una milésima.

En cuanto a la calidad del modelo propuesto, vimos que supera los modelos contra los cuales fue comparado. En general, al parecer una característica que contribuyó a lo anterior fue la inclusión de distintas emociones del usuario como parte de

su información personal. En el caso en el que el día de la semana es introducido en el modelo como información del entorno, existe una mejora en la correlación obtenida para los modelos que utilizan la ventana de tiempo de un día, no así para los que utilizan ventanas de tiempo de una hora o bloques de 15 minutos. Esto tiene sentido, ya que en estos últimos modelos el día no varía mucho entre ventana y ventana de tiempo. En todos los casos podemos observar que el modelo de contagio, el cual incluye información de las emociones de los contactos, genera errores más bajos que el modelo que incluye solo las emociones, lo cual nos indica que el efecto del contagio emocional existe en la red social; a pesar de que esta diferencia entre los modelos no es estadísticamente significativa, sí es consistente en los tres tipos de ventana: un día, una hora y bloques de 15 minutos.

Con respecto a la ventaja que se obtuvo sobre ARIMA, se vio que el modelo propuesto fue mejor en todos los casos. Creemos que esto se debe a que el modelo ARIMA tiene la desventaja de que utiliza una única serie de tiempo de la emoción correspondiente para cada usuario. Es decir, utiliza información proveniente de un único usuario y una única emoción para hacer la predicción dada la naturaleza del método; esta misma desventaja ya había sido mencionada por van Breda *et al.* (2016). En cambio, los métodos de regresión lineal y máquinas de soporte vectorial utilizan información de múltiples usuarios y emociones para hacer el entrenamiento y posterior prueba en otro conjunto amplio independiente. Pensamos que esta es una de las razones por las que el modelo ARIMA no pudo superar al modelo propuesto. Existen otras variaciones del modelo ARIMA, como el modelo VARIMA, el cual puede trabajar con varias series de tiempo; sin embargo, no fue implementado como modelo base debido a que la utilización de otras emociones es parte del modelo propuesto y, debido a su eficiencia, se escogió la regresión lineal como principal método. VARIMA aún tendría la desventaja de utilizar varias series de tiempo provenientes de un único usuario en su entrenamiento para hacer la predicción. Su análisis quedará para un trabajo futuro.

Como también vimos, el modelo propuesto se comparó con el modelo de Mo-

gadala y Varma (2012) mostrando una mejoría al predecir la transición en las emociones. Pensamos que una de las principales ventajas que se obtuvo fue durante el etiquetado de los comentarios, ya que en este trabajo fue realizado manualmente por personas que conocen el lenguaje. En el trabajo contra el cual comparamos, por otro lado, utilizaron un método semiautomático para hacer la clasificación. Asimismo, el método de selección final de la emoción para cada comentario discutido en la sección 5.2.1 provocó una mejora en la calidad de los clasificadores automáticos para las emociones.

A pesar de que el modelo propuesto superó los 28 métodos contra los cuales fue comparado, las correlaciones en general siguen siendo débiles. Creemos que esto se debe a que hay factores externos a la red social que no se están considerando y que probablemente están afectando a las personas (televisión, noticias, eventos deportivos, fiestas, periodos escolares). En ese sentido, probablemente estemos llegando a una cota superior en la predicción de emociones basada en una sola red social y se necesite agregar otros mecanismos que afectan a las emociones de las personas y que no se han considerado hasta el momento.

5.4 COMPROBACIÓN DE HIPÓTESIS

El eje rector de los experimentos y resultados es poder aceptar o rechazar las hipótesis planteadas. Debido a que en el capítulo 6 se describen detalladamente las respuestas a las preguntas de investigación, las cuales están estrechamente ligadas a las hipótesis, nos centramos en la hipótesis general: *Es posible crear un modelo para predicción de intensidad emocional en una red social en línea mediante técnicas de procesamiento de texto en español y aprendizaje máquina.*

Para aceptar la hipótesis, consideramos necesario que el modelo propuesto cumpla con lo siguiente:

1. Ser un modelo predictivo.
2. Predecir intensidad emocional en una red social.
3. Utilizar técnicas de procesamiento de texto en español y aprendizaje máquina.

Además de cumplir con lo anterior, la hipótesis en su conjunto denota la calidad (parecido de la predicción a la realidad) de la que se habla en la sección 5.3.1.2. Por tanto, se verá si la evidencia provista es suficiente para aceptar la hipótesis.

En resumen, se tiene la siguiente evidencia:

- El modelo propuesto obtiene valores futuros (en la experimentación, el valor corresponde a la siguiente ventana de tiempo) para la variable objetivo de intensidad emocional. (*Punto 1*)
- El modelo propuesto utiliza como características cuatro tipos de variables: intensidades emocionales históricas del usuario (\mathbf{U}), intensidades emocionales de sus contactos (*followees*) en la red social (\mathbf{A}_f), intensidades emocionales de sus contactos por mención en la red social (\mathbf{A}_m) y el día de la semana correspondiente a la predicción. Las variables relacionadas con contactos se extraen de una red social. En la experimentación, la red social Twitter se utiliza como caso de estudio. Por tanto, se puede considerar que el modelo cumple con predecir intensidad emocional en una red social. (*Punto 2*)
- El modelo predictivo está basado en regresión lineal y máquinas de soporte vectorial para reconocimiento de emociones en texto, el cual a su vez utiliza un método de clasificación de sentimiento en español. Por tanto, se puede considerar que cumple con utilizar técnicas de procesamiento en español y aprendizaje máquina. (*Punto 3*)
- El modelo propuesto supera a otros 28 métodos, incluyendo el modelo ARIMA y un método que se encuentra en el estado del arte. La diferencia entre el modelo y 24 de estos métodos es estadísticamente significativa con $\alpha = 0.05$

(la mayoría de los métodos con los cuales no hay diferencia significativa son modelos parciales basados en el propuesto). Por tanto, se puede considerar que es competitivo. (*Calidad*)

- La fase de reconocimiento (detección) de emociones en texto obtuvo un valor-F superior a 70 %. Por tanto, también se puede considerar que la metodología utilizada para esta fase es competitiva. (*Calidad*)

Considerando la anterior evidencia, la hipótesis general se acepta.

5.5 RESUMEN

En este capítulo expusimos las dos fases por las que tuvimos que pasar para poder atacar el problema de predicción de intensidad emocional: (1) la creación de clasificadores automáticos y (2) la implementación del modelo propuesto. Hablamos sobre el repositorio utilizado, el cual proviene de la red social de Twitter: una de las más populares y utilizadas actualmente a nivel mundial. El repositorio proviene de usuarios de la ciudad de Monterrey, Nuevo León y fue extraído anteriormente por Rodríguez Aldape (2013). Los comentarios seleccionados fueron etiquetados manualmente por alumnos de la Facultad de Ingeniería Mecánica y Eléctrica de la Universidad Autónoma de Nuevo León. Mediante el método del *acuerdo de intercorrelación*, se eligió la etiqueta global más apropiada para el comentario. Para cada emoción básica se creó un clasificador automático — excepto para la emoción del miedo ya que para esta no se logró obtener una cantidad de datos suficiente. Los clasificadores obtuvieron precisiones por encima del 70 %, lo cual resultó aceptable para los experimentos posteriores con el modelo propuesto.

Para evaluar el modelo, se llevó a cabo la predicción de intensidad emocional para una ventana de tiempo siguiente a los datos históricos. El modelo se comparó con otros *modelos base* a través de métricas de uso estándar para métodos de

aprendizaje máquina (correlación y error absoluto medio). Con el fin de comparar estas métricas a través de una prueba de t -student, se realizaron 300 repeticiones de los experimentos y se compararon los resultados con los otros modelos base.

El modelo propuesto resultó ser mejor ($p \leq 0.05$) que los demás modelos base. Los modelos que no incluyen información del día de la semana o de las emociones de los contactos del usuario tienen una ligera desventaja en cuanto a precisión sin ser esta significativa estadísticamente. El modelo ARIMA, el cual es utilizado en problemas donde se asume que se tiene una serie de tiempo, obtuvo resultados por debajo del modelo propuesto. Esto se debió a que dicho método no considera otra información excepto la de la emoción en cuestión y a que los datos históricos son escasos para tal. En los experimentos vimos que a partir de la inclusión de otras emociones y a que todas tienen un efecto en la predicción, se mejora la precisión de forma significativa. Este es uno de los resultados más relevantes, ya que nos invita a explorar más la interacción que tienen nuestras emociones pasadas y su efecto en el futuro.

CAPÍTULO 6

CONCLUSIONES Y TRABAJO FUTURO

En este capítulo, se mencionan y discuten las principales conclusiones acerca del trabajo propuesto. Comenzamos mencionando las conclusiones más importantes y posteriormente describimos las posibles mejoras futuras y otras áreas que se pueden explorar en la predicción de emociones en redes sociales. Separamos las conclusiones en dos categorías: (1) clasificadores automáticos y (2) modelo de predicción de intensidad emocional. Comenzamos primero hablando sobre la etapa de los clasificadores automáticos de emociones básicas y posteriormente del modelo. Finalmente hablamos de las mejoras que se pueden realizar en el modelo para obtener resultados de mayor calidad y así tener más aplicaciones prácticas.

6.1 RESUMEN

Se propuso un modelo para la predicción de emociones en redes sociales. Específicamente, este modelo predice, para un usuario, la proporción de comentarios que publicará con una emoción en particular para una ventana de tiempo futura; a esta proporción la definimos como *intensidad emocional*. Contrario a otros modelos que se enfocan en una sola emoción, el modelo propuesto consideró un esquema básico de emociones y utilizó todas las emociones para la predicción de la intensidad de cada emoción. El modelo se basó tres fuentes distintas de información: (1) pro-

veniente del usuario, (2) proveniente de los contactos del usuario y (3) proveniente del entorno. La predicción finalmente se hizo mediante regresión lineal. Encontramos que nuestro modelo, aplicado sobre un conjunto de datos extraído de Twitter, produjo mejores resultados (hasta en un 35 %) comparado con modelos como ARIMA, y mejores resultados en comparación con otros 27 modelos de base y un método del estado del arte.

Con el fin de detectar emociones básicas en los comentarios, creamos clasificadores automáticos para texto en español extraído de Twitter. El repositorio para la tarea de clasificación contiene 1,895 comentarios con etiquetas de emociones para alegría, enojo y tristeza, donde estas etiquetas fueron asignadas por estudiantes de la Facultad de Ingeniería Mecánica y Eléctrica a nivel de licenciatura. Los clasificadores pudieron detectar emociones con más del 70 % de precisión, lo cual es aceptable para los experimentos que se realizaron.

Las principales contribuciones son: el modelo de predicción de intensidad emocional y su habilidad para considerar la interacción entre varias emociones, el repositorio de datos etiquetado con emociones y el establecimiento de una base comparativa para futuros trabajos.

6.2 COMENTARIOS FINALES

Esta sección se divide en dos partes: (1) clasificadores automáticos y (2) modelo. Cada una contiene los comentarios finales acerca de los resultados que se obtuvieron y las posibles mejoras que pueden realizar en cada parte del proceso de experimentación.

6.2.1 CLASIFICADORES AUTOMÁTICOS

Los clasificadores automáticos obtuvieron una precisión aceptable — alrededor del 70 % — para detectar la presencia de emociones básicas en los comentarios. Esta precisión es comparable a la de otros trabajos que emplean técnicas de sacos de palabras; a través de la incorporación de técnicas más sofisticadas de procesamiento de lenguaje natural se podría mejorar la precisión. La técnica utilizada en la que se calculó la frecuencia de cada palabra en el comentario correspondía a una característica es muy sencilla; sin embargo, no fue necesario agregar más complejidad a este proceso, ya que el objetivo fue mostrar la viabilidad de la predicción de emociones básicas a futuro. Aún así, un análisis más exhaustivo de las características se podría aplicar a los comentarios. Esto se discute con más detalle en la sección 6.6. En el caso del miedo, no se pudo obtener un clasificador debido a que la cantidad de comentarios etiquetados con miedo por los evaluadores fue demasiado baja. Esto puede deberse al método de selección de comentarios a evaluar, pero será necesario investigar más a fondo dicha emoción y cómo se manifiesta en las redes sociales. Con el fin de recaudar más comentarios de este tipo, se podrían aprovechar fechas en las que sucedan acontecimientos como desastres naturales, juegos olímpicos u otro tipo de situaciones en las que se propicie la expresión de emociones a través de las redes sociales.

Se usaron todas las palabras del comentario como características para las máquinas de soporte vectorial, provocando que los vectores se hicieran demasiado grandes; si bien esto no nos generó problemas para obtener una precisión aceptable, quizá esta podría mejorarse a partir de una mejor selección de características. El tiempo invertido en el procesamiento de los comentarios se mejoraría, ya que al evaluar mejor los comentarios obteniendo las características más informativas, el entrenamiento de los métodos de aprendizaje máquina se reduciría.

Hemos utilizado también un solo tipo de herramienta para generar los clasificadores automáticos: máquinas de soporte vectorial. Sin embargo, sería conveniente

investigar otro tipo de mecanismos que pudieran resultar mejores. Estos mecanismos incluyen redes neuronales, redes bayesianas, k -vecinos más cercanos y agrupamiento, entre otras. Optamos por esta técnica debido a su amplio uso en el área de clasificación de texto.

6.2.2 MODELO

Vimos que el modelo propuesto fue mejor que los modelos base y el modelo ARIMA en cuanto a correlación y a medida del error. El modelo ARIMA tuvo la desventaja de utilizar solo datos de la emoción a predecir y no de todas las emociones, así como de la información global disponible del usuario. Esto ocasionó que la calidad fuera más baja en comparación. Notamos también que los patrones semanales que se observan en Twitter influyen en el modelo, por lo que el periodo histórico es de aproximadamente una semana para el caso en el que las ventanas de tiempo representan días. En el caso en el que la ventana de tiempo es de horas o minutos, la cantidad de datos históricos tiene un impacto positivo en la calidad a medida que el periodo crece, pero solo hasta un punto máximo a partir del cual la calidad comienza a bajar. Cada vez que se agrega un dato histórico, se incrementa el ruido, y cuando este comienza a superar la información valiosa, la calidad disminuye.

Uno de los resultados más importantes que podemos observar es que, cuando se incluye información sobre las diferentes emociones en el modelo, la calidad se mejora considerablemente ($p \leq 0.05$). Es decir, encontramos evidencia de que una emoción anterior está relacionada a la manera en que el usuario expresa sus emociones en una ventana de tiempo futura. Notamos que dicha mejora existe desde que se agrega una sola emoción más al modelo como parte de los datos históricos y sigue mejorando cuando se agregan las tres emociones básicas estudiadas. También observamos que al agregar información acerca del contagio emocional o menciones de otros usuarios, existe una mejora en la calidad del modelo. Aunque esta no sea estadísticamente significativa con respecto al modelo que incluye información de varias emociones,

sí es consistente en las tres granularidades para las ventanas de tiempo analizadas. Pensamos que la diferencia corresponde a una mejora real que podría confirmarse estadísticamente al hacer un estudio más enfocado en esa parte o agregando más instancias en un estudio similar a este. De igual manera, agregar información acerca del día de la semana tuvo una mejora en el modelo con una ventana de tiempo de un día, y esta diferencia no se observó en los modelos en donde se utilizaban horas o minutos debido a que es más probable que el día no cambie. Para nosotros, esto nos invita seguir investigando el efecto de la información acerca del ambiente y su efecto en la expresión de emociones por parte de los usuarios.

Al comparar nuestro modelo con otros parecidos, vimos que genera una calidad mejor al momento de predecir una “transición” en las emociones. También es necesario seguir investigando en este sentido para colocar el modelo más precisamente dentro del estado del arte y sugerir mejoras más concretas.

Podemos concluir que el modelo presentado puede ser utilizado para anticipar la intensidad emocional de los usuarios y que la utilización de diversas emociones puede afectar en la predicción. Sin embargo, aún queda mucho por mejorar con el fin de obtener un modelo que sea aplicable de forma práctica. Los factores externos (noticias, lanzamientos televisivos y espectáculos, por mencionar algunos) afectan a los usuarios y, por lo tanto, deben tomarse en cuenta. Solo consideramos un solo factor: el nombre del día de la semana. Aún así, vimos una pequeña mejora, aunque no fue significativa estadísticamente. Es necesario seguir investigando sus efectos en las redes sociales.

6.3 RESPUESTA A LAS PREGUNTAS DE INVESTIGACIÓN

Anteriormente en el capítulo 1, enumeramos las preguntas de investigación que se analizaron, las cuales respondemos a continuación.

- *¿Cómo puede crearse un modelo de predicción de intensidad emocional de los usuarios en una red social en línea utilizando técnicas de procesamiento de texto en español y aprendizaje máquina?* Nuestro modelo toma en cuenta información proveniente del usuario, sus contactos y el entorno. Vimos que es posible utilizar un método de aprendizaje máquina que considere esta información para hacer una predicción de la intensidad emocional en una ventana de tiempo futura. Esta es una forma en la que se pueden crear modelos de este tipo; sin embargo, como mencionamos en la sección 6.6, otros modelos matemáticos también se pueden utilizar.
 - *¿Es posible utilizar minería de texto en español para este propósito?* Sí es posible; lo pudimos comprobar al crear los clasificadores automáticos.
 - *¿Es posible detectar emociones básicas en redes sociales a través del análisis de texto en español?* Sí es posible; los clasificadores automáticos pueden detectar emociones básicas con precisiones por encima del 70% para la alegría, enojo y tristeza. Si bien no pudimos tener suficientes datos para crear un clasificador para la emoción del miedo, es suficiente una de las demás emociones para contestar afirmativamente esta pregunta.
 - *¿Es posible medir la intensidad emocional de los usuarios para cada emoción básica?* Sí es posible; de acuerdo con nuestra definición de intensidad emocional, vimos que es posible calcularla para cualquier usuario con la condición de que este debe haber publicado por lo menos un comentario como emoción. De lo contrario, dicho valor es igual a cero.
 - *¿Es posible crear un modelo de predicción de intensidad emocional de los usuarios utilizando información que procede de ellos, de sus contactos y del medio ambiente?* Sí es posible; vimos que los modelos base predicen, llegando hasta cierto grado de calidad, la intensidad emocional; el modelo propuesto los supera.
 - *¿Es posible utilizar el modelo para predecir la intensidad emocional de los usuarios para cada emoción básica?* Sí es posible; en la experimentación

vimos resultados favorables para la predicción de la intensidad emocional para la siguiente ventana de tiempo.

6.4 CONTRIBUCIONES

En resumen, nuestras aportaciones fueron las siguientes:

- Un modelo de predicción de emociones.
- La introducción del concepto de *intensidad emocional*.
- La utilización de emociones básicas: alegría, enojo, miedo y tristeza.
- El considerar interdependencia de emociones para la predicción de intensidad emocional.
- La incorporación de *contagio emocional* a través del uso de intensidades emocionales de contactos en la red social.
- La creación de clasificadores automáticos de emociones.
- Un repositorio de comentarios en español con etiquetas de emociones básicas.

6.5 POSIBLES APLICACIONES

Predecir emociones en un futuro cercano permite una gran cantidad de aplicaciones prácticas. Por ejemplo, se ha visto que en Twitter existen patrones temporales que se repiten semanalmente (Dodds *et al.*, 2011). Por tanto, los días lunes y martes — en los que la felicidad es baja — se pueden plantear estrategias con el fin de evitar un decaimiento del ánimo laboral dentro de una empresa. Esto se puede hacer a través de una dinámica, de hacer un pequeño convivio por las mañanas,

o de realizar alguna actividad por la tarde que favorezca el ánimo (como un show cómico). En las redes sociales, se podría utilizar esta información para crear sistemas de recomendación que favorezcan contenido que levante el ánimo. Las ventas por Internet se podrían planear mejor al analizar los comportamientos de compra de los usuarios bajo la influencia de ciertas emociones, como la alegría o el enojo. En el área médica, la detección de personas con problemas de bipolaridad, estrés crónico o depresión podría ayudar a brindarles una orientación y darles una asistencia dentro de la red social con el fin de que el problema no siga creciendo. Asimismo, la detección de personas que de forma consistente expresan enojo o tristeza hacia otros usuarios podrían ser analizadas para clasificarse como *bullies* o acosadores y de esta manera recibir alguna penalización o advertencia de parte de la red social con el fin de modular dicho comportamiento ofensivo.

Algunas redes sociales como Facebook permiten la creación de eventos para convocar a una reunión. En este sentido, podrían proponer de forma automática la organización de un evento una vez que tenga un pronóstico para un tipo de emoción futura. Por ejemplo, si se pronostica una racha de felicidad, podría proponerse una reunión para ir a pasear con amigos. En la actualidad, muchas parejas conviven dentro de las redes sociales. A través de un modelo como este, podrían predecirse conflictos futuros y tomar alguna acción correctiva antes de que sucedan. Por ejemplo, si uno de los dos va a estar enojado al día siguiente, podría recomendarse a la pareja enviar un mensaje agradable.

6.6 TRABAJO FUTURO

Aún queda mucho trabajo por hacer. Al igual que las conclusiones separamos el trabajo futuro en dos áreas: (1) clasificadores automáticos y (2) modelo.

6.6.1 CLASIFICADORES AUTOMÁTICOS

Debido a que el objetivo principal es con respecto al modelo, hemos utilizado clasificadores de emociones más sencillos. Sin embargo, sería conveniente en aplicaciones prácticas utilizar clasificadores más sofisticados de emociones que incluyan técnicas de clasificación a partir de análisis semántico o de aplicación de conocimiento. Las técnicas más actuales para analizar texto incluyen la aplicación de ontologías, ya que estas pueden representar el conocimiento humano.

En trabajos futuros podríamos crear una ontología (similar a la *Open Mind Common Sense* del MIT) para la detección de emociones en español. De igual manera, ayudaría investigar estas técnicas desde el punto de vista de emociones básicas, debido a que para cada emoción se podrían obtener redes semánticas u ontologías independientes que ayuden a mejorar la precisión. Este tipo de observaciones aplican también para emociones complejas como la desesperanza. Una mejor selección de características también sería de utilidad, ya que hemos utilizado todas las palabras como características — a excepción de un listado de palabras vacías que fueron eliminadas. Si pudiéramos reducir la cantidad de estas y aumentar su calidad la clasificación de comentarios, también podría ser más rápida y aplicarse en tiempo real durante la publicación de comentarios en las redes sociales.

6.6.2 MODELO

Como primera mejora, debemos investigar más a fondo la emoción del miedo: seguir obteniendo repositorios de información que nos lo permitan, ya que no pudimos hacerlo. El miedo podría ser un gran modificador para las otras emociones básicas y de esta manera se incrementaría la calidad en la predicción. En este modelo, hemos propuesto la regresión lineal como método principal con fines comparativos; sin embargo, sería conveniente investigar otras técnicas de aprendizaje máquina, como máquinas de soporte vectorial y redes neuronales — por mencionar algunas. La

selección adecuada de características también podría mejorar la calidad de la predicción y, por lo tanto, sería conveniente hacer un análisis más exhaustivo para obtener las características más útiles para los métodos de aprendizaje máquina. Asimismo, podríamos hacer una investigación del método de MoodCast y ampliarlo para que se tomen en cuenta diferentes emociones básicas e información del usuario, y revisar si efectivamente encontraríamos un mejor modelo. Este modelo puede extenderse para predecir otras emociones complejas como los celos, la soledad o la desesperanza, e incluso detectar patrones que indiquen alguna tendencia hacia un problema de salud como un trastorno de bipolaridad.

Las emociones negativas juegan un papel importante para encontrar casos de acoso en redes sociales, por lo que la presencia de estas podría anticipar un ataque de este tipo hacia algún usuario — tanto desde el punto de vista de quien lo hace como del que lo recibe. Los sucesos externos afectan las emociones humanas; por ejemplo, en México ocurrió un terremoto el 24 de Septiembre de 2017, el cual provocó varios destrozos en la capital el país. Este suceso ha ocasionado diversas emociones, tanto negativas (por ejemplo, tristeza) como positivas durante varios días. Por lo tanto, incluir en el modelo un listado de los sucesos más importantes históricos sería de utilidad para predecir las emociones.

BIBLIOGRAFÍA

- ACEMOGLU, D., A. OZDAGLAR y A. PARANDEHGHEIBI (2010), «Spread of (mis) information in social networks», *Games and Economic Behavior*, **70**(2), págs. 194–227.
- AGGARWAL, C. C. y C. ZHAI (2012), «A survey of text classification algorithms», en *Mining text data*, Springer, Massachusetts, Estados Unidos, págs. 163–222.
- AKNIN, L. B., G. M. SANDSTROM, E. W. DUNN y M. I. NORTON (2011), «It's the recipient that counts: Spending money on strong social ties leads to greater happiness than spending on weak social ties», *PLOS ONE*, **6**(2), pág. e17018.
- ALTAY, E. y M. H. SATMAN (2005), «Stock market forecasting: artificial neural network and linear regression comparison in an emerging market», *Journal of Financial Management & Analysis*, **18**(2), págs. 18–33.
- BACCIANELLA, S., A. ESULI y F. SEBASTIANI (2010), «SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining», en *4th Language Resources and Evaluation Conference*, tomo 10, European Language Resources Association, Piran, Eslovenia, págs. 2200–2204.
- BALAHUR, A., J. M. HERMIDA, A. MONTOYO y R. MUÑOZ (2011), *EmotiNet: A Knowledge Base for Emotion Detection in Text Built on the Appraisal Theories*, Springer Berlin Heidelberg, Berlín, Alemania, págs. 27–39.
- BALDONI, M., C. BAROGLIO, V. PATTI y P. RENA (2012), «From tags to emo-

- tions: Ontology-driven sentiment analysis in the social semantic web», *Intelligenza Artificiale*, **6**(1), págs. 41–54.
- BAO, S., S. XU, L. ZHANG, R. YAN, Z. SU, D. HAN y Y. YU (2012), «Mining social emotions from affective text», *IEEE Transactions on Knowledge and Data Engineering*, **24**(9), págs. 1658–1670.
- BARSADE, S. G. (2002), «The ripple effect: Emotional contagion and its influence on group behavior», *Administrative Science Quarterly*, **47**(4), págs. 644–675.
- BENNETT, W. L. y A. SEGERBERG (2011), «Digital media and the personalization of collective action: Social technology and the organization of protests against the global economic crisis», *Information, Communication & Society*, **14**(6), págs. 770–799.
- BIANCO, V., O. MANCA y S. NARDINI (2009), «Electricity consumption forecasting in Italy using linear regression models», *Energy*, **34**(9), págs. 1413–1421.
- BIEVER, C. (2010), «Twitter mood maps reveal emotional states of America», *New Scientist*, **207**(2771), pág. 14.
- BISPO, J. y A. PAIVA (2009), «A model for emotional contagion based on the emotional contagion scale», en *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009*, IEEE, California, Estados Unidos, págs. 1–6.
- BLISS, C. A., I. M. KLOUMANN, K. D. HARRIS, C. M. DANFORTH y P. S. DODDS (2012), «Twitter reciprocal reply networks exhibit assortativity with respect to happiness», *Journal of Computational Science*, **3**(5), págs. 388–397.
- BOLLEN, J., H. MAO y A. PEPE (2011), «Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena», en *5th International AAAI Conference on Web and Social Media*, MIT Press, Cambridge, Estados Unidos, págs. 450–453.

- BORTH, D., T. CHEN, R. JI y S.-F. CHANG (2013), «Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content», en *Proceedings of the 21st ACM International Conference on Multimedia*, ACM, Nueva York, Estados Unidos, págs. 459–460.
- BOSCO, C., V. PATTI y A. BOLIOLI (2013), «Developing corpora for sentiment analysis: The case of irony and senti-tut», *IEEE Intelligent Systems*, **28**(2), págs. 55–63.
- BOTELLA, C., G. RIVA, A. GAGGIOLI, B. K. WIEDERHOLD, M. ALCANIZ y R. M. BAÑOS (2012), «The present and future of positive technologies», *CyberPsychology, Behavior, and Social Networking*, **15**(2), págs. 78–84.
- BOVIK, A. C., M. CLARK y W. S. GEISLER (1990), «Multichannel texture analysis using localized spatial filters», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**(1), págs. 55–73.
- BOX, G. E., G. M. JENKINS, G. C. REINSEL y G. M. LJUNG (2015), *Time series analysis: forecasting and control*, quinta edición, John Wiley & Sons, Nueva Jersey, Estados Unidos.
- BRANS, K. y P. VERDUYN (2014), «Intensity and Duration of Negative Emotions: Comparing the Role of Appraisals and Regulation Strategies», *PLOS ONE*, **9**(3), págs. 1–13.
- BROWN, C. M., A. B. DIEKMAN, R. E. TENNIAL y E. D. SOLOMON (2011), «Alone and happy: Personality moderates the effect of happy mood on social approach», *Journal of Research in Personality*, **45**(6), págs. 702–705.
- CAI, Z., B. GOERTZEL, C. ZHOU, D. HUANG, S. KE, G. YU y M. JIANG (2013), «OpenPsi: A novel computational affective model and its application in video games», *Engineering Applications of Artificial Intelligence*, **26**(1), págs. 1–12.
- CAMBRIA, E. (2016), «Affective computing and sentiment analysis», *IEEE Intelligent Systems*, **31**(2), págs. 102–107.

- CAMBRIA, E. y A. HUSSAIN (2012), *Sentic computing: Techniques, tools, and applications*, tomo 2, Springer Science & Business Media, Nueva York, Estados Unidos.
- CAMBRIA, E. y A. HUSSAIN (2015), «Sentic Computing», *Cognitive Computation*, **7**(2), págs. 183–185.
- CAMBRIA, E., A. LIVINGSTONE y A. HUSSAIN (2012), «The hourglass of emotions», en *Cognitive Behavioural Systems*, Springer, Berlín, Alemania, págs. 144–157.
- CAMBRIA, E., D. OLSHER y D. RAJAGOPAL (2014), «SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis», en *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, MIT Press, Massachusetts, Estados Unidos, págs. 1515–1521.
- CAMBRIA, E., S. PORIA, R. BAJPAI y B. W. SCHULLER (2016), «SenticNet 4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives.», en *COLING*, págs. 2666–2677.
- CAMBRIA, E., R. SPEER, C. HAVASI y A. HUSSAIN (2010), «SenticNet: A Publicly Available Semantic Resource for Opinion Mining», en *AAAI Fall Symposium: Commonsense Knowledge*, tomo 10, MIT Press, Massachusetts, Estados Unidos, págs. 14–18.
- CAO, L., G. BALA, M. ZHENG y K. CALDEIRA (2015), «Fast and slow climate responses to CO₂ and solar forcing: A linear multivariate regression model characterizing transient climate change», *Journal of Geophysical Research: Atmospheres*, **120**(23), págs. 37–53.
- CASTI, J. (2010), «Why social mood moves the world», *New Scientist*, **206**(2761), págs. 30–31.
- CAVNAR, W. B., J. M. TRENKLE *et al.* (1994), «N-gram-based text categorization», *Ann Arbor MI*, **48113**(2), págs. 161–175.

- CHAI, T. y R. R. DRAXLER (2014), «Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature», *Geoscientific Model Development*, **7**(3), págs. 1247–1250.
- CHEN, J., B. HU, P. MOORE y X. ZHANG (2016), «Ontology-Based Model for Mining User's Emotions on the Wisdom Web», en *Wisdom Web of Things*, Springer, Zug, Suiza, págs. 121–153.
- CHEN, Y.-Y., T. CHEN, W. H. HSU, H.-Y. M. LIAO y S.-F. CHANG (2014), «Predicting viewer affective comments based on image content in social media», en *Proceedings of International Conference on Multimedia Retrieval*, ACM, Nueva York, Estados Unidos, pág. 233.
- CHMIEL, A., J. SIENKIEWICZ, M. THELWALL, G. PALTOGLOU, K. BUCKLEY, A. KAPPAS y J. A. HOLYST (2011a), «Collective emotions online and their influence on community life», *PLOS ONE*, **6**(7), pág. e22 207.
- CHMIEL, A., P. SOBKOWICZ, J. SIENKIEWICZ, G. PALTOGLOU, K. BUCKLEY, M. THELWALL y J. A. HOLYST (2011b), «Negative emotions boost user activity at BBC forum», *Physica A: Statistical Mechanics and its Applications*, **390**(16), págs. 2936 – 2944.
- COCHRAN, A. L., A. SCHULTZ, M. G. MCINNIS y D. B. FORGER (2017), *Computational Neurology and Psychiatry*, capítulo A Comparison of Mathematical Models of Mood in Bipolar Disorder, Springer International Publishing, Zug, Suiza, págs. 315–341.
- COHEN, I., N. SEBE, A. GARG, L. S. CHEN y T. S. HUANG (2003), «Facial expression recognition from video sequences: temporal and static modeling», *Computer Vision and Image Understanding*, **91**(1), págs. 160–187.
- CRAWFORD, S. L. (2006), «Correlation and regression», *Circulation*, **114**(19), págs. 2083–2088.

- CSÁJI, B. C. (2001), «Approximation with artificial neural networks», *Faculty of Sciences, Etsz Lornd University, Hungary*, **24**, pág. 48.
- DAI, W., D. HAN, Y. DAI y D. XU (2015), «Emotion recognition and affective computing on vocal social media», *Information & Management*, **52**(7), págs. 777–788.
- DESMET, B. y V. HOSTE (2013), «Emotion detection in suicide notes», *Expert Systems with Applications*, **40**(16), págs. 6351 – 6358.
- DEWALL, C. N., R. F. BAUMEISTER, D. S. CHESTER y B. J. BUSHMAN (2016), «How often does currently felt emotion predict social behavior and judgment? A meta-analytic test of two theories», *Emotion Review*, **8**(2), págs. 136–143.
- DHALL, A., R. GOECKE, T. GEDEON y N. SEBE (2016), «Emotion recognition in the wild», *J. Multimodal User Interfaces*, **10**(2), págs. 95–97.
- DHALL, A., R. GOECKE, S. LUCEY y T. GEDEON (2011), «Acted Facial Expressions in the Wild Database», *Informe técnico*, The Australian National University.
- DODDS, P. S., K. D. HARRIS, I. M. KLOUMANN, C. A. BLISS y C. M. DANFORTH (2011), «Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter», *PLOS ONE*, **6**(12), pág. e26752.
- DURBIN, J. y S. J. KOOPMAN (2012), *Time series analysis by state space methods*, tomo 38, Oxford University Press, Oxfordshire, Reino Unido.
- EBRAHIMI KAHOU, S., V. MICHALSKI, K. KONDA, R. MEMISEVIC y C. PAL (2015), «Recurrent neural networks for emotion recognition in video», en *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ACM, Nueva York, Estados Unidos, págs. 467–474.
- EKMAN, P. (1992), «Are there basic emotions?», *Psychological Review*, **99**(3), págs. 550–553.

- EL AYADI, M., M. S. KAMEL y F. KARRAY (2011), «Survey on speech emotion recognition: Features, classification schemes, and databases», *Pattern Recognition*, **44**(3), págs. 572–587.
- FATHY, S., N. EL-HAGGAR y M. H. HAGGAG (2017), «A Hybrid Model for Emotion Detection from Text», *International Journal of Information Retrieval Research*, **7**(1), págs. 32–48.
- FENTON-O’CREEVY, M., E. SOANE, N. NICHOLSON y P. WILLMAN (2011), «Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders», *Journal of Organizational Behavior*, **32**(8), págs. 1044–1061.
- FRIJDA, N. H. (1986), *The emotions*, Cambridge University Press, Nueva York, Estados Unidos.
- FRIJDA, N. H. (1988), «The laws of emotion», *American Psychologist*, **43**(5), pág. 349.
- FUNG, K. Y. (2012), «An intelligent system for supporting affective product design», <http://ira.lib.polyu.edu.hk/handle/10397/6114>.
- GARCÍA-CUMBRERAS, M. Á., A. MONTEJO-RÁEZ y M. C. DÍAZ-GALIANO (2013), «Pessimists and optimists: Improving collaborative filtering through sentiment analysis», *Expert Systems with Applications*, **40**(17), págs. 6758–6765.
- GEUSEBROEK, J.-M. (2006), «Compact Object Descriptors from Local Colour Invariant Histograms», en *British Machine Vision Conference*, Escocia, Reino Unido, págs. 1029–1038.
- GRANGER, C. W. J. y M. HATANAKA (2015), *Spectral Analysis of Economic Time Series (PSME-1)*, Princeton University Press, Nueva Jersey, Estados Unidos.
- GRANGER, C. W. J. y P. NEWBOLD (2014), *Forecasting economic time series: Economic theory, econometrics, and mathematical economics*, segunda edición, Academic Press, Inc., California, Estados Unidos.

- GRASSI, M. (2009), «Developing HEO human emotions ontology», en *European Workshop on Biometrics and Identity Management*, Springer, Berlín, Alemania, págs. 244–251.
- HALL, M., E. FRANK, G. HOLMES, B. PFAHRINGER, P. REUTEMANN y I. H. WITTEN (2009), «The WEKA data mining software: an update», *ACM SIGKDD explorations newsletter*, **11**(1), págs. 10–18.
- HAN, B.-J., S. RHO, S. JUN y E. HWANG (2010), «Music emotion classification and context-based music recommendation», *Multimedia Tools and Applications*, **47**(3), págs. 433–460.
- HANNAK, A., E. ANDERSON, L. F. BARRETT, S. LEHMANN, A. MISLOVE y M. RIEDEWALD (2012), «Tweetin' in the Rain: Exploring Societal-Scale Effects of Weather on Mood», en *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, MIT Press, Massachusetts, Estados Unidos, págs. 479–482.
- HATFIELD, E. y J. T. CACIOPPO (1994), *Emotional contagion*, Cambridge University Press, Nueva York, Estados Unidos.
- HO, C. y C. SPENCE (2013), «Affective multisensory driver interface design», *International Journal of Vehicle Noise and Vibration*, **9**(1), págs. 61–74.
- HO, C.-C. y K. F. MACDORMAN (2017), «Measuring the Uncanny Valley Effect», *International Journal of Social Robotics*, **9**(1), págs. 129–139.
- HOFFMAN, M. L. (2002), «How automatic and representational is empathy, and why», *Behavioral and Brain Sciences*, **25**(01), págs. 38–39.
- HOOGENDOORN, M., J. TREUR, C. N. VAN DER WAL y A. VAN WISSEN (2011), «Agent-based modelling of the emergence of collective states based on contagion of individual states in groups», en *Transactions on computational collective intelligence III*, Springer, Berlín, Alemania, págs. 152–179.

- HU, Q., A. LI, F. HENG, J. LI y T. ZHU (2015), «Predicting Depression of Social Media User on Different Observation Windows», en *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology 2015*, tomo 1, IEEE, California, Estados Unidos, págs. 361–364.
- HU, X. y J. S. DOWNIE (2010), «Improving mood classification in music digital libraries by combining lyrics and audio», en *Proceedings of the 10th annual joint conference on Digital libraries*, ACM, Nueva York, Estados Unidos, págs. 159–168.
- HUANG, G. B., M. RAMESH, T. BERG y E. LEARNED-MILLER (2007), «Labeled faces in the wild: A database for studying face recognition in unconstrained environments», *Informe técnico*, Technical Report 07-49, University of Massachusetts, Amherst.
- IZARD, C. E. (2007), «Basic emotions, natural kinds, emotion schemas, and a new paradigm», *Perspectives on psychological science*, **2**(3), págs. 260–280.
- JACK, R. E., O. G. GARROD y P. G. SCHYNS (2014), «Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time», *Current biology*, **24**(2), págs. 187–192.
- JOHNSON, G. (2009), «Theories of Emotion», <http://www.iep.utm.edu/emotion/>.
- KAGAN, V., E. ROSSINI y D. SAPOUNAS (2013), *Sentiment analysis for PTSD signals*, Springer, Nueva York, Estados Unidos.
- KAHOU, S. E., C. PAL, X. BOUTHILLIER, P. FROUMENTY, Ç. GÜLÇEHRE, R. MEMISEVIC, P. VINCENT, A. COURVILLE, Y. BENGIO, R. C. FERRARI *et al.* (2013), «Combining modality specific deep neural networks for emotion recognition in video», en *Proceedings of the 15th ACM on International conference on multimodal interaction*, ACM, Nueva York, Estados Unidos, págs. 543–550.
- KAJIWARA, Y., M. NAKAMURA, H. KIMURA y T. OYABU (2017), «A Mood Prediction System for Preventing Diseases Using Biological Information and Weather Information», *Electronics and Communications in Japan*, **100**(2), págs. 12–24.

- KENNEDY, A. y D. INKPEN (2006), «Sentiment classification of movie reviews using contextual valence shifters», *Computational Intelligence*, **22**(2), págs. 110–125.
- KHORRAMI, P., T. LE PAINE, K. BRADY, C. DAGLI y T. S. HUANG (2016), «How deep neural networks can improve emotion recognition on video data», en *IEEE International Conference on Image Processing 2016*, IEEE, California, Estados Unidos, págs. 619–623.
- KIM, H.-J., S.-B. PARK y G.-S. JO (2014), «Affective social network-happiness inducing social media platform», *Multimedia Tools and Applications*, **68**(2), págs. 355–374.
- KIM, H.-R. y D.-S. KWON (2010), «Computational model of emotion generation for human–robot interaction based on the cognitive appraisal theory», *Journal of Intelligent & Robotic Systems*, **60**(2), págs. 263–283.
- KIM, Y. E., E. M. SCHMIDT, R. MIGNECO, B. G. MORTON, P. RICHARDSON, J. SCOTT, J. A. SPECK y D. TURNBULL (2010), «Music emotion recognition: A state of the art review», en *Proceedings of The International Society of Music Information Retrieval*, International Society of Music Information Retrieval, Utrecht, Países Bajos, págs. 255–266.
- KIRBY, R., J. FORLIZZI y R. SIMMONS (2010), «Affective social robots», *Robotics and Autonomous Systems*, **58**(3), págs. 322–332.
- KLEINSMITH, A. y N. BIANCHI-BERTHOUBE (2013), «Affective body expression perception and recognition: A survey», *IEEE Transactions on Affective Computing*, **4**(1), págs. 15–33.
- KONTOPOULOS, E., C. BERBERIDIS, T. DERGIADIS y N. BASSILIADES (2013), «Ontology-based sentiment analysis of Twitter posts», *Expert systems with applications*, **40**(10), págs. 4065–4074.
- KOSHAL, R. K., M. KOSHEL y J. D. LINDLEY (2015), «Forecasting the Index of

- Consumer Sentiment», en *Proceedings of the 1985 Academy of Marketing Science Annual Conference*, Springer, Zug, Suiza, págs. 450–451.
- KRAMER, A. D., J. E. GUILLORY y J. T. HANCOCK (2014), «Experimental evidence of massive-scale emotional contagion through social networks», *Proceedings of the National Academy of Sciences*, **111**(24), págs. 8788–8790.
- LANG, P. J., M. M. BRADLEY y B. N. CUTHBERT (1997), «International Affective Picture System (IAPS): Technical manual and affective ratings», *Informe técnico*, NIMH Center for the Study of Emotion and Attention.
- LERNER, J. S., R. M. GONZALEZ, D. A. SMALL y B. FISCHHOFF (2003), «Effects of fear and anger on perceived risks of terrorism: A national field experiment», *Psychological Science*, **14**(2), págs. 144–150.
- LEWIS, M., J. M. HAVILAND-JONES y L. F. BARRETT (2010), *Handbook of emotions*, tercera edición, Guilford Press, Nueva York, Estados Unidos.
- LI, G. y F. LIU (2012), «Application of a clustering method on sentiment analysis», *Journal of Information Science*, **38**(2), págs. 127–139.
- LI, T. y M. OGIHARA (2004), «Content-based music similarity search and emotion detection», en *IEEE International Conference on Acoustics, Speech, and Signal Processing 2004*, tomo 5, IEEE, California, Estados Unidos, págs. V–705–V–708.
- LI, Y., J. HUANG, H. WANG y L. FENG (2015), «Predicting Teenager’s Future Stress Level from Micro-Blog», en *IEEE 28th International Symposium on Computer-Based Medical Systems*, IEEE, California, Estados Unidos, págs. 208–213.
- LIU, B. (2010), «Sentiment Analysis and Subjectivity», *Handbook of natural language processing*, **2**, págs. 627–666.
- LIU, B. (2012), «Sentiment analysis and opinion mining», *Synthesis lectures on human language technologies*, **5**(1), págs. 1–167.

- LIU, D., L. LU y H.-J. ZHANG (2003), «Automatic mood detection from acoustic music data», *Informe técnico*, Microsoft Research Asia.
- LIU, H. y P. SINGH (2004), «ConceptNet: A practical commonsense reasoning toolkit», *BT technology journal*, **22**(4), págs. 211–226.
- LU, L., D. LIU y H.-J. ZHANG (2006), «Automatic mood detection and tracking of music audio signals», *IEEE Transactions on audio, speech, and language processing*, **14**(1), págs. 5–18.
- LYNCH, A. (1996), *Thought contagion: How belief spreads through society*, Basic Books, Nueva York, Estados Unidos.
- MATSUMOTO, D. (2007), «Playing catch with emotions», *Journal of Intercultural Communication*, **10**, págs. 39–49.
- MAXHUNI, A., P. HERNANDEZ-LEAL, E. F. MORALES, L. E. SUCAR, V. OSMANI, A. MUÑOZ-MELÉNDEZ y O. MAYORA (2017), *Applications for Future Internet*, capítulo Using Intermediate Models and Knowledge Learning to Improve Stress Prediction, Springer International Publishing, Zug, Suiza, págs. 140–151.
- MEDHAT, W., A. HASSAN y H. KORASHY (2014), «Sentiment analysis algorithms and applications: A survey», *Ain Shams Engineering Journal*, **5**(4), págs. 1093–1113.
- MEHRABIAN, A. (1996), «Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament», *Current Psychology*, **14**(4), págs. 261–292.
- MICHEL, P. y R. EL KALIOUBY (2003), «Real time facial expression recognition in video using support vector machines», en *Proceedings of the 5th international conference on Multimodal interfaces*, ACM, Nueva York, Estados Unidos, págs. 258–264.
- MILLER, G. A. (1995), «WordNet: a lexical database for English», *Communications of the ACM*, **38**(11), págs. 39–41.

- MOGADALA, A. y V. VARMA (2012), «Twitter User Behavior Understanding with Mood Transition Prediction», en *Proceedings of the 2012 Workshop on Data-driven User Behavioral Modelling and Mining from Social Media*, ACM, Nueva York, Estados Unidos, págs. 31–34.
- MOORE, P. J., M. A. LITTLE, P. E. MCSHARRY, J. R. GEDDES y G. M. GOODWIN (2012), «Forecasting depression in bipolar disorder», *IEEE Transactions on Biomedical Engineering*, **59**(10), págs. 2801–2807.
- MORAES, R., J. F. VALIATI y W. P. G. NETO (2013), «Document-level sentiment classification: An empirical comparison between SVM and ANN», *Expert Systems with Applications*, **40**(2), págs. 621–633.
- MORAL, C., A. DE ANTONIO, R. IMBERT y J. RAMÍREZ (2014), «A survey of stemming algorithms in information retrieval.», *Information Research: An International Electronic Journal*, **19**(1), pág. n1.
- MORENO-ORTIZ, A. y C. P. HERNÁNDEZ (2013), «Lexicon-based sentiment analysis of twitter messages in spanish», *Procesamiento del Lenguaje Natural*, **50**, págs. 93–100.
- MOU, W., O. CELIKTUTAN y H. GUNES (2015), «Group-level arousal and valence recognition in static images: Face, body and context», en *IEEE 11th International Conference and Workshops on Automatic Face and Gesture Recognition*, tomo 5, IEEE, California, Estados Unidos, págs. 1–6.
- MUNEZERO, M. D., C. S. MONTERO, E. SUTINEN y J. PAJUNEN (2014), «Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text», *IEEE Transactions on Affective Computing*, **5**(2), págs. 101–111.
- NG, Y., C. KHONG y H. THWAITES (2012), «A Review of Affective Design towards Video Games», *Procedia-Social and Behavioral Sciences*, **51**, págs. 687–691.
- NGUYEN, L. T., P. WU, W. CHAN, W. PENG y Y. ZHANG (2012), «Predicting Collective Sentiment Dynamics from Time-series Social Media», en *Proceedings of*

- the First International Workshop on Issues of Sentiment Discovery and Opinion Mining*, ACM, Nueva York, Estados Unidos, págs. 6:1–6:8.
- NIELSEN, F. Å. (2011), «A new ANEW: Evaluation of a word list for sentiment analysis in microblogs», abs/1103.2903, arXiv preprint.
- NIU, J., X. ZHAO, L. ZHU y H. LI (2013), «Affvir: An affect-based Internet video recommendation system», *Neurocomputing*, **120**, págs. 422–433.
- NORDHAUS, W. D. y A. MOFFAT (2017), «A Survey of Global Impacts of Climate Change: Replication, Survey Methods, and a Statistical Analysis», *Informe técnico*, National Bureau of Economic Research.
- NOROOZI, F., M. MARJANOVIC, A. NJEGUS, S. ESCALERA y G. ANBARJAFARI (2017), «Audio-visual emotion recognition in video clips», *IEEE Transactions on Affective Computing*, **PP**, págs. 1–1, paper still not assigned to a volume.
- NWE, T. L., S. W. FOO y L. C. DE SILVA (2003), «Speech emotion recognition using hidden Markov models», *Speech Communication*, **41**(4), págs. 603–623.
- O’CONNOR, B., R. BALASUBRAMANYAN, B. R. ROUTLEDGE y N. A. SMITH (2010), «From tweets to polls: Linking text sentiment to public opinion time series.», *ICWSM*, **11**(122-129), págs. 1–2.
- ORTIZ, A. M., F. P. CASTILLO y R. H. GARCÍA (2010a), «Análisis de Valoraciones de Usuario de Hoteles con Sentitext: un sistema de análisis de sentimiento independiente del dominio», *Procesamiento del Lenguaje Natural*, **45**, págs. 31–39.
- ORTIZ, A. M., Á. P. POZO y S. T. SÁNCHEZ (2010b), «Sentitext: sistema de análisis de sentimiento para el español», *Procesamiento del Lenguaje Natural*, **45**, págs. 297–298.
- PADRÓ, L. y E. STANILOVSKY (2012), «Freeling 3.0: Towards wider multilinguality», en *Language Resources and Evaluation*, European Language Resources Association, Cataluña, España, págs. 2473–2479.

- PAK, A. y P. PAROUBEK (2010), «Twitter as a Corpus for Sentiment Analysis and Opinion Mining», en *Language Resources and Evaluation*, European Language Resources Association, La Valeta, Malta, págs. 1320–1326.
- PANG, B. y L. LEE (2005), «Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales», en *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, ACM, Nueva York, Estados Unidos, págs. 115–124.
- PANG, B., L. LEE y S. VAITHYANATHAN (2002), «Thumbs up?: Sentiment classification using machine learning techniques», en *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, ACM, Nueva York, Estados Unidos, págs. 79–86.
- PANG, B., L. LEE *et al.* (2008), «Opinion mining and sentiment analysis», *Foundations and Trends in Information Retrieval*, **2**(1–2), págs. 1–135.
- PARROTT, W. G. (2001), *Emotions in social psychology: Essential readings*, Psychology Press: Taylor & Francis Group, Pennsylvania, Estados Unidos.
- PATWARDHAN, A. y G. KNAPP (2016a), «Aggressive actions and anger detection from multiple modalities using Kinect», abs/1607.01076, arXiv preprint.
- PATWARDHAN, A. S. y G. M. KNAPP (2016b), «Multimodal Affect Recognition using Kinect», abs/1607.02652, arXiv preprint.
- PÉREZ ESPINOSA, H., H. J. ESCALANTE, L. VILLASEÑOR-PINEDA, M. MONTES-Y GÓMEZ, D. PINTO-AVEDAÑO y V. REYEZ-MEZA (2014), «Fusing Affective Dimensions and Audio-Visual Features from Segmented Video for Depression Recognition: INAOE-BUAP's Participation at AVEC'14 Challenge», en *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge*, ACM, Nueva York, Estados Unidos, págs. 49–55.
- PERRIN, A. (2016), «Social Media Usage», URL <http://www.pewinternet.org/2016/11/11/social-media-update-2016/>.

- PICARD, R. W. (1995), «Affective Computing», *Informe técnico*, Massachusetts Institute of Technology.
- PISANI, A. R., P. A. WYMAN, M. PETROVA, K. SCHMEELK-CONE, D. B. GOLDSTON, Y. XIA y M. S. GOULD (2013), «Emotion Regulation Difficulties, Youth–Adult Relationships, and Suicide Attempts Among High School Students in Underserved Communities», *Journal of Youth and Adolescence*, **42**(6), págs. 807–820.
- PLUTCHIK, R. (1984), «Emotions: A general psychoevolutionary theory», *Approaches to Emotion*, **1984**, págs. 197–219.
- PLUTCHIK, R. (2001), «The Nature of Emotions Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice», *American Scientist*, **89**(4), págs. 344–350.
- POMERLEAU, D. A. (1991), «Efficient training of artificial neural networks for autonomous navigation», *Neural Computation*, **3**(1), págs. 88–97.
- POPESCU, S. A. (2012), *SAGE: A simple affective game engine*, Tesis de Maestría, Universitet van Amsterdam.
- QING, C., U. ENDRISS, R. FERNÁNDEZ y J. KRUGER (2014), «Empirical Analysis of Aggregation Methods for Collective Annotation», en *24th International Conference on Computational Linguistics*, Universidad de Dublín y Association for Computational Linguistics, Leinster, Irlanda, págs. 1533–1542.
- QIYAO, W., L. ZHENGMIN, J. YUEHUI, C. SHIDUAN y Y. TAN (2016), «ULM: A user-level model for emotion prediction in social networks», *The Journal of China Universities of Posts and Telecommunications*, **23**(3), págs. 63 – 88.
- REN, Y., Y. ZHANG, M. ZHANG y D. JI (2016), «Context-Sensitive Twitter Sentiment Classification Using Neural Network», en *Proceedings of the 13th AAAI Conference on Artificial Intelligence*, MIT Press, Massachusetts, Estados Unidos, págs. 215–221.

- RINGEVAL, F., S. AMIRIPARIAN, F. EYBEN, K. SCHERER y B. SCHULLER (2014), «Emotion recognition in the wild: Incorporating voice and lip activity in multi-modal decision-level fusion», en *Proceedings of the 16th International Conference on Multimodal Interaction*, ACM, Nueva York, Estados Unidos, págs. 473–480.
- RIVA, G., R. M. BANOS, C. BOTELLA, B. K. WIEDERHOLD y A. GAGGIOLI (2012), «Positive technology: using interactive technologies to promote positive functioning», *Cyberpsychology, Behavior, and Social Networking*, **15**(2), págs. 69–77.
- RODRÍGUEZ ALDAPE, F. M. (2013), *Cuantificación del interés de un usuario en un tema mediante minería de texto y análisis de sentimiento*, Tesis de Maestría, Universidad Autónoma de Nuevo León.
- RUI, T., P. CUI y W. ZHU (2017), «Joint user-interest and social-influence emotion prediction for individuals», *Neurocomputing*, **230**, págs. 66 – 76.
- SARIYANIDI, E., H. GUNES y A. CAVALLARO (2015), «Automatic analysis of facial affect: A survey of registration, representation, and recognition», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**(6), págs. 1113–1133.
- SCHEDL, M., E. GÓMEZ, J. URBANO *et al.* (2014), «Music information retrieval: Recent developments and applications», *Foundations and Trends® in Information Retrieval*, **8**(2-3), págs. 127–261.
- SHEPHERD, L., R. SPEARS y A. S. MANSTEAD (2013), «“This will bring shame on our nation”: The role of anticipated group-based emotions on collective action», *Journal of Experimental Social Psychology*, **49**(1), págs. 42–57.
- SIDOROV, G., S. MIRANDA-JIMÉNEZ, F. VIVEROS-JIMÉNEZ, A. GELBUKH, N. CASTRO-SÁNCHEZ, F. VELÁSQUEZ, I. DÍAZ-RANGEL, S. SUÁREZ-GUERRA, A. TREVIÑO y J. GORDON (2012), «Empirical study of machine learning based approach for opinion mining in tweets», en *Mexican International Conference on Artificial Intelligence*, Springer, Berlín, Alemania, págs. 1–14.

- SINGH, P., T. LIN, E. T. MUELLER, G. LIM, T. PERKINS y W. L. ZHU (2002), «Open Mind Common Sense: Knowledge acquisition from the general public», en *OTM Confederated International Conferences: On the Move to Meaningful Internet Systems*, Springer, Berlín, Alemania, págs. 1223–1237.
- SMALL, D. A. y N. M. VERROCHI (2009), «The face of need: Facial emotion expression on charity advertisements», *Journal of Marketing Research*, **46**(6), págs. 777–787.
- SNOW, R., B. O’CONNOR, D. JURAFSKY y A. Y. NG (2008), «Cheap and fast—but is it good?: Evaluating non-expert annotations for natural language tasks», en *Proceedings of the conference on empirical methods in natural language processing*, ACM, Nueva York, Estados Unidos, págs. 254–263.
- SOLOMON, R. C. (1977), «The logic of emotion», *Nous*, **1**(1), págs. 41–49.
- SPIELBERGER, C. D. y E. C. REHEISER (2009), «Assessment of emotions: Anxiety, anger, depression, and curiosity», *Applied Psychology: Health and Well-Being*, **1**(3), págs. 271–302.
- STATHOPOULOU, I.-O. y G. A. TSIHRINTZIS (2011), «Emotion recognition from body movements and gestures», en *Intelligent Interactive Multimedia Systems and Services*, Springer, Berlín, Alemania, págs. 295–303.
- STRAPPARAVA, C., A. VALITUTTI *et al.* (2004), «WordNet Affect: an Affective Extension of WordNet», en *2004 Language Resources and Evaluation Conference*, tomo 4, European Language Resources Association, Lisboa, Portugal, págs. 1083–1086.
- STURM, B. L. (2014), «The state of the art ten years after a state of the art: Future research in music information retrieval», *Journal of New Music Research*, **43**(2), págs. 147–172.
- SUN, H., H. X. LIU, H. XIAO, R. R. HE y B. RAN (2003), «Short term traffic forecasting using the local linear regression model», en *82nd Annual Meeting of*

- the Transportation Research Board*, Transportation Research Board, Washington, DC, Estados Unidos, págs. 326–334.
- TABOADA, M., J. BROOKE, M. TOFILOSKI, K. VOLL y M. STEDE (2011), «Lexicon-based methods for sentiment analysis», *Computational Linguistics*, **37**(2), págs. 267–307.
- TAN, P.-N., M. STEINBACH y V. KUMAR (2006), *Introduction to Data Mining*, Addison Wesley, Massachusetts, Estados Unidos.
- TEIXEIRA, J. P., C. OLIVEIRA y C. LOPES (2013), «Vocal acoustic analysis–jitter, shimmer and hnr parameters», *Procedia Technology*, **9**, págs. 1112–1122.
- TOMKINS, S. S. y R. MC CARTER (1964), «What and where are the primary affects? Some evidence for a theory», *Perceptual and Motor Skills*, **18**(1), págs. 119–158.
- TORRES, F. y F. RODRÍGUEZ (2013), «Correlación de comentarios en Twitter desde una perspectiva de apoyo social y violencia», *Revista Psicología Científica.com*, **15**(4).
- TSAI, J., E. BOWRING, S. MARSELLA y M. TAMBE (2011a), «Empirical evaluation of computational emotional contagion models», en *Intelligent Virtual Agents*, Springer, Berlín, Alemania, págs. 384–397.
- TSAI, J., N. FRIDMAN, E. BOWRING, M. BROWN, S. EPSTEIN, G. KAMINKA, S. MARSELLA, A. OGDEN, I. RIKI, A. SHEEL *et al.* (2011b), «ESCAPES: evacuation simulation with children, authorities, parents, emotions, and social comparison», en *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, ACM, Nueva York, Estados Unidos, págs. 457–464.
- TURNEY, P. D. (2002), «Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews», en *Proceedings of the 40th annual meeting on association for computational linguistics*, ACM, Nueva York, Estados Unidos, págs. 417–424.

- VALSTAR, M., J. GRATCH, B. SCHULLER, F. RINGEVAL, D. LALANNE, M. TORRES TORRES, S. SCHERER, G. STRATOU, R. COWIE y M. PANTIC (2016), «AVEC 2016: Depression, mood, and emotion recognition workshop and challenge», en *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*, ACM, Nueva York, Estados Unidos, págs. 3–10.
- VAN BREDÁ, W., J. PASTOR, M. HOOGENDOORN, J. RUWAARD, J. ASSELBERGS y H. RIPER (2016), *Innovation in Medicine and Healthcare*, capítulo Exploring and Comparing Machine Learning Approaches for Predicting Mood Over Time, Springer International Publishing, Zug, Suiza, págs. 37–47.
- VAN DEN OORD, A., S. DIELEMAN y B. SCHRAUWEN (2013), «Deep content-based music recommendation», en *Advances in neural information processing systems*, Nips Foundation, California, Estados Unidos, págs. 2643–2651.
- VERDUYN, P. y S. LAVRIJSEN (2015), «Which emotions last longest and why: The role of event importance and rumination», *Motivation and Emotion*, **39**(1), págs. 119–127.
- VERVERIDIS, D. y C. KOTROPOULOS (2006), «Emotional speech recognition: Resources, features, and methods», *Speech Communication*, **48**(9), págs. 1162–1181.
- WALTER, A. (2012), *Designing for emotion*, Vegas Lex, Nueva York, Estados Unidos.
- WAN, J., S. ESCALERA, A. GHOLAMREZA, H. J. ESCALANTE, X. BARÓ, I. GUYON, M. MADADI, A. JURI, G. JELENA, L. CHI *et al.* (2017), «Results and Analysis of ChaLearn LAP Multi-modal Isolated and Continuous Gesture Recognition, and Real versus Fake Expressed Emotions Challenges», en *IEEE International Conference on Computer Vision Workshops*, IEEE, California, Estados Unidos, págs. 1–9.
- WANG, X., J. JIA y L. CAI (2013a), «Affective image adjustment with a single word», *The Visual Computer*, **29**(11), págs. 1121–1133.

- WANG, X., J. JIA, J. YIN y L. CAI (2013b), «Interpretable aesthetic features for affective image classification», en *IEEE 20th International Conference on Image Processing*, IEEE, California, Estados Unidos, págs. 3230–3234.
- WILKS, Y. (1997), *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology*, capítulo Information extraction as a core language technology, Springer Berlin Heidelberg, Berlín, Alemania, págs. 1–9.
- WILLMOTT, C. J. y K. MATSUURA (2005), «Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance», *Climate Research*, **30**(1), págs. 79–82.
- WILSON, T., J. WIEBE y P. HOFFMANN (2005), «Recognizing contextual polarity in phrase-level sentiment analysis», en *Proceedings of the conference on human language technology and empirical methods in natural language processing*, ACM, Nueva York, Estados Unidos, págs. 347–354.
- WU, F., Y. SONG y Y. HUANG (2015), «Microblog Sentiment Classification with Contextual Knowledge Regularization», en *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, MIT Press, Massachusetts, Estados Unidos, págs. 2332–2338.
- XIE, P. y E. P. XING (2013), «Integrating Document Clustering and Topic Modeling», abs/1309.6874, arXiv preprint.
- XU, R. y D. WUNSCH (2005), «Survey of clustering algorithms», *IEEE Transactions on Neural Networks*, **16**(3), págs. 645–678.
- YACOUB, S. M., S. J. SIMSKE, X. LIN y J. BURNS (2003), «Recognition of emotions in interactive voice response systems», en *Interspeech*, Ginebra, Suiza, págs. 1–4.
- YANG, Y.-H. y H. H. CHEN (2012), «Machine recognition of music emotion: A review», *ACM Transactions on Intelligent Systems and Technology*, **3**(3), pág. 40.

- YANG, Y.-H., Y.-C. LIN, Y.-F. SU y H. H. CHEN (2008), «A regression approach to music emotion recognition», *IEEE Transactions on Audio, Speech, and Language Processing*, **16**(2), págs. 448–457.
- YANULEVSKAYA, V., J. C. VAN GEMERT, K. ROTH, A.-K. HERBOLD, N. SEBE y J.-M. GEUSEBROEK (2008), «Emotional valence categorization using holistic image features», en *IEEE 15th International Conference on Image Processing, 2008*, IEEE, California, Estados Unidos, págs. 101–104.
- YE, Q., B. LIN y Y.-J. LI (2005), «Sentiment classification for Chinese reviews: A comparison between SVM and semantic approaches», en *Proceedings of the 2005 International Conference on Machine Learning and Cybernetics*, tomo 4, IEEE, California, Estados Unidos, págs. 2341–2346.
- YIN, D., S. BOND y H. ZHANG (2013), «Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews», *MIS Quarterly*, **38**, págs. 539–560.
- YOU, Q., J. LUO, H. JIN y J. YANG (2015), «Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks.», en *AAAI*, MIT Press, Massachusetts, Estados Unidos, págs. 381–388.
- ZAFEIRIOU, S., C. ZHANG y Z. ZHANG (2015), «A survey on face detection in the wild: past, present and future», *Computer Vision and Image Understanding*, **138**, págs. 1–24.
- ZHALEHPOUR, S., Z. AKHTAR y C. E. ERDEM (2016), «Multimodal emotion recognition based on peak frame selection from video», *Signal, Image and Video Processing*, **10**(5), págs. 827–834.
- ZHANG, X., B. HU, J. CHEN y P. MOORE (2013), «Ontology-based context modeling for emotion recognition in an intelligent web», *World Wide Web*, **16**(4), págs. 497–513.

- ZHANG, Y., J. TANG, J. SUN, Y. CHEN y J. RAO (2010), «Moodcast: Emotion prediction via dynamic continuous factor graph model», en *IEEE 10th International Conference on Data Mining*, IEEE, California, Estados Unidos, págs. 1193–1198.
- ZHAO, L., J. JIA y L. FENG (2015), «Teenagers’ stress detection based on time-sensitive micro-blog comment/response actions», en *IFIP International Conference on Artificial Intelligence in Theory and Practice*, Springer, Zug, Suiza, págs. 26–36.
- ZHAO, S., Y. GAO, X. JIANG, H. YAO, T.-S. CHUA y X. SUN (2014a), «Exploring principles-of-art features for image emotion recognition», en *Proceedings of the 22nd ACM international conference on Multimedia*, ACM, Nueva York, Estados Unidos, págs. 47–56.
- ZHAO, S., H. YAO, Y. GAO, R. JI, W. XIE, X. JIANG y T.-S. CHUA (2016), «Predicting personalized emotion perceptions of social images», en *Proceedings of the 2016 ACM on Multimedia Conference*, ACM, Nueva York, Estados Unidos, págs. 1385–1394.
- ZHAO, S., H. YAO, Y. YANG y Y. ZHANG (2014b), «Affective image retrieval via multi-graph learning», en *Proceedings of the 22nd ACM international conference on Multimedia*, ACM, Nueva York, Estados Unidos, págs. 1025–1028.

RESUMEN AUTOBIOGRÁFICO

Fernando Manuel Rodríguez Aldape

Candidato para obtener el grado de
Doctorado en Ingeniería
con Orientación en Tecnologías de la Información

Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica

Tesis:

PREDICCIÓN DE INTENSIDADES EMOCIONALES EN UNA RED SOCIAL

Nací en la ciudad de Monterrey, Nuevo León, México el año de 1985. Obtuve el grado de Licenciado en Ciencias Computacionales en la Facultad de Ciencias Físico-Matemáticas en el año 2007. En 2009 viajé a la ciudad de Barcelona, España para estudiar un posgrado en Dirección y Gestión de Marketing en la Universitat de Barcelona. Posteriormente, en el año 2010, regresé a mi ciudad y continué mis estudios en la Facultad de Ingeniería Mecánica y Eléctrica en la Maestría en Ingeniería de la Información con Orientación en Inteligencia Artificial graduándome en el año 2013. Posteriormente, continué en la Facultad de Ingeniería Mecánica y Eléctrica estudiando el Doctorado en Ingeniería con Orientación en Computación.