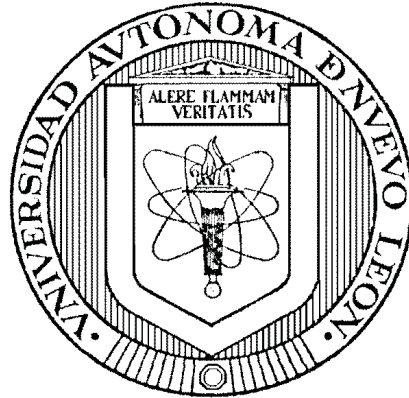


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO



MODELADO DE UN MOLINO DE LAMINACIÓN EN FRIO USANDO
TÉCNICAS DE SISTEMAS INTELIGENTES PARA PREDICCIÓN DE
PLANEZA.

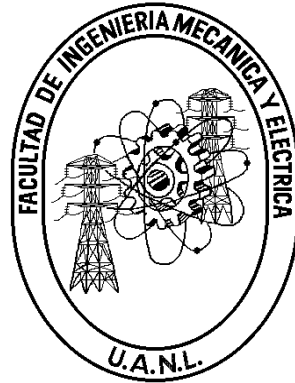
POR
LUIS ALBERTO ALVARADO YAÑEZ

TESIS
EN OPCIÓN AL GRADO DE MAESTRO EN CIENCIAS DE LA
INGENIERÍA ELÉCTRICA

CD. UNIVERSITARIA

SEPTIEMBRE 2019

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



MODELADO DE UN MOLINO DE LAMINACIÓN EN FRIO USANDO
TÉCNICAS DE SISTEMAS INTELIGENTES PARA PREDICCIÓN DE
PLANEZA.

POR
LUIS ALBERTO ALVARADO YAÑEZ

TESIS
EN OPCIÓN AL GRADO DE MAESTRO EN CIENCIAS DE LA
INGENIERÍA ELÉCTRICA

CD. UNIVERSITARIA

SEPTIEMBRE 2019



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

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 “Modelado de un molino de laminación en frío usando técnicas de sistemas inteligentes para la predicción de planeza”, realizada por el alumno Luis Alberto Alvarado Yañez, con número de matrícula 1400693, sea aceptada para su defensa como requisito para obtener el grado de Maestría en Ciencias en Ingeniería Eléctrica.

El Comité de Tesis

Luis Martín Torres Treviño
Director

Dr. Alberto Cavazos González
Revisor

Dr. Mario Alberto Bello Gómez
Revisor

Vo. Bo.

Dr. Simón Martínez Martínez
Subdirector de Estudios de Posgrado



116

San Nicolás de los Garza, Nuevo León, Octubre de 2020



A mi esposa que me apoyo durante este periodo dando consejos y aliento para completar este logro, sus palabras y cariño dieron soporte en los momentos difíciles.

A Sofia por ti y para ti siempre.

A mis padres por el apoyo incondicional durante toda mi formación, dándome los valores que me formaron como persona, mostrándome su trabajo y perseverancia, muchas gracias.

A mis amigos y familia que siempre están cuando los necesito y me dieron palabras de aliento para terminar esta tesis.

Gracias, habría sido imposible lograrlo sin cada uno de ustedes.

*.
. .
.*

Agradezco al Dr. Luis Torres, quien me apoyo para el desarrollo de esta tesis de manera amable y cordial.

Agradezco al CONACYT por el soporte económico que brinda a sus becarios, también a Ternium por su apoyo económico, facilidades y su excelente plan de catedra que permite producir investigaciones en conjunto con las universidades.

Resumen

El proceso de laminación en frío conlleva la interacción de diferentes variables, parámetros y perturbaciones, las cuales interactúan entre sí en una forma altamente no lineal, por lo cual es difícil obtener un modelo matemático.

Este proyecto tiene como finalidad obtener un modelo mediante sistemas inteligentes que permita predecir los valores necesarios para mejorar los atributos de calidad en el producto como lo son el espesor y la planeza. Para el desarrollo del modelo es necesario recabar datos de proceso, los cuales son obtenidos de un molino de laminación de la empresa "Ternium", estos datos son recabados de diferentes bases de datos del proceso, sin embargo debido a la gran cantidad de variables que registran los sistemas de adquisición de datos es necesario adecua los datos y eliminar las variables redundantes para poder alimentar los diferentes sistemas.

Se propone el uso de redes neuronales tales como la red neuronal "backpropagation" y la red neuronal de máxima sensibilidad para el modelado del proceso, además de el uso de regresión simbólica y algoritmos evolutivos para la optimización.

Índice General

Índice General	v
Índice de Tablas	vii
Índice de Figuras	viii
1 Introducción	1
1.1 Laminación en frío.	2
1.2 Delimitación del problema.	3
1.3 Antecedentes.	5
1.4 Hipótesis.	7
1.5 Objetivo General.	7
1.5.1 Objetivos Particulares.	7
2 Marco Teórico.	9
2.1 Procedimiento.	9
2.1.1 Diagrama a bloques.	9
2.2 Análisis de las componentes principales.	9
2.3 Redes Neuronales.	13
2.3.1 Red neuronal tipo perceptrón.	14
2.3.2 Red neuronal de máxima sensibilidad.	17
2.4 Regresión simbólica.	20
2.4.1 "Evonorm".	24
2.5 Estadísticos	26
2.5.1 Error Cuadrático Medio ("MSE")	26
2.5.2 Error Residual Predicho por Suma de Cuadrados ("PRESS")	26
2.5.3 " $R^2 PRED$ " ("R-squared predicted")	27
3 Metodología.	32
3.1 Proceso.	32
3.1.1 Espesor	33
3.1.2 Planeza	36
3.1.3 Medición de Planeza	38
3.2 Parámetros	40
3.2.1 Redes neuronales	40

3.2.2	Arquitectura y parámetros.	41
3.2.3	Perceptrón Multicapas	42
3.2.4	Máxima Sensibilidad	42
3.2.5	Regresión simbólica	43
3.3	Entrenamiento	44
3.3.1	Perceptrón multicapas	44
3.3.2	Máxima sensibilidad	44
3.3.3	Regresión simbólica	45
3.4	Validación	45
4	Experimentación y resultados.	47
4.1	Espesor	47
4.1.1	Desempeño de la red neuronal.	47
4.1.2	Red Neuronal Perceptrón con entrenamiento "Backpropagation"	48
4.1.3	Máxima sensibilidad	49
4.1.4	Regresión simbólica	50
4.1.5	Tablas de comparación	51
4.2	Planeza	52
4.2.1	Desempeño de la red neuronal.	53
5	Conclusiones.	57
5.1	Trabajo Futuro	58
	Bibliografía	60

Indice de Tablas

2.2.1 Información utilizada.	10
2.2.2 Información con la media sustraída.	11
2.3.3 Determinación de máxima sensibilidad.	18
2.3.4 Aprendizaje sin máxima sensibilidad.	20
2.4.5 Operadores α y su función matemática relacionada.	21
3.1.1 Entradas y salidas del proceso	35
4.1.1 Operadores encontrados α	52
4.1.2 Operadores encontrados β	53
4.1.3 Análisis estadístico para cada método considerando la mejor solución encontrada y el tiempo de procesamiento consumido por cada método.	53

Índice de Figuras

2.1	<i>Diagrama Bloques</i>	28
2.2	<i>Datos del ejemplo del ACP.</i>	29
2.3	<i>Grafica de los valores normalizados con los eigenvectores de la matriz de covarianza.</i>	30
2.4	<i>Partes principales de la red neuronal.</i>	30
2.5	<i>Partes principales de la red neurona máxima sensibilidad.</i>	31
3.1	<i>Principales variables y parámetros de entrada y salida del molino de laminación</i>	33
3.2	<i>Esquema del modelo del proceso</i>	34
3.3	<i>Definición de planeza</i>	38
3.4	<i>Medidor de planeza y formas generadas</i>	39
3.5	<i>Proceso de simplificación de base de datos y utilidad en redes neuronales.</i>	41
4.1	<i>Épocas de entrenamiento contra el error cuadrático medio.</i>	48
4.2	<i>Comparación de valores predichos contra los valores de las muestras.</i>	49
4.3	<i>Comparación de valores predichos contra los valores nuevos desconocidos por la red neuronal "Backpropagation".</i>	50
4.4	<i>Comparación de valores predichos contra los valores nuevos desconocidos por la red neuronal MS.</i>	51
4.5	<i>Épocas de entrenamiento contra el error cuadrático medio.</i>	54
4.6	<i>Épocas de entrenamiento contra el error cuadrático medio.</i>	54
4.7	<i>Comparación de valores predichos contra los valores de las muestras.</i>	55
4.8	<i>Comparación de valores predichos contra los valores nuevos desconocidos por la red neuronal "Backpropagation".</i>	55
4.9	<i>Comparación de valores predichos contra los valores nuevos desconocidos por la red neuronal Máxima Sensibilidad.</i>	56
4.10	<i>Comparación de valores predichos contra los valores nuevos desconocidos por la red neuronal Máxima Sensibilidad.</i>	56

Capítulo 1

Introducción

El proceso de laminación consiste en pasar metal entre rodillos los cuales realizan un esfuerzo de compresión entre ellos para reducir el espesor del metal y ampliar su longitud. Cada conjunto de rodillos es llamado un castillo (stand), y para realizar el proceso un castillo es lo mínimo requerido, más sin embargo existen laminadores que cuentan con más de un solo castillo, este tipo de laminadores son conocidos con el nombre de tándem. Cada castillo está formado por un conjunto de rodillos y la cantidad con la cual están formados estos deriva un nombre para la descripción del arreglo de rodillos, estos pueden ser desde dos rodillos los cuales designarían el nombre de two high, four high para cuatro rodillos, six high para seis rodillos y cluster para arreglos especiales de más rodillos. Cabe mencionar que se designa especialmente como "rodillos de trabajo" a los rodillos que realizan contacto directo con el metal a procesar, el resto de los rodillos son conocidos como "rodillos de apoyo". El producto a laminar puede pasar en más de una ocasión a través de un castillo para obtener el espesor necesario, a este tipo de laminador se le conoce como molino laminador reversible. Cada molino laminador está diseñado para diferentes productos dependiendo de las propiedades que se requieran en este. Por otra parte la temperatura de la lámina a la cual se realiza este tratamiento divide el proceso en dos grandes grupos; la laminación en caliente la cual se realiza con el acero a una temperatura mayor a la de recristalización lleva el acero de una forma de planchas a un espesor menor el cual es necesario para trabajar los procesos posteriores. La laminación en frío se realiza a temperatura ambiente,

por lo cual la fuerza necesaria para formar la deformación es mayor, este trabajo esta enfocado completamente a la laminación en frío.

1.1 Laminación en frío.

La laminación en frío es uno de los procesos más importantes en la acería ya que incrementa la dureza del material, mejora la precisión en tamaño de espesor, y produce espesores más delgados con un mejor acabado superficial, utilizados muchas veces para fines estéticos o mejora en sus propiedades mecánicas. El mejorar la calidad del producto es uno de los temas más importantes de la laminación en frío, debido al incremento en las expectativas de los clientes que esperan una mejor precisión y menores tolerancias en los controles de calidad. Esto ayuda a los clientes a mejorar sus operaciones y por lo tanto sus productos, en muchos casos están utilizando espesores más delgados, lo cual implica tener más control para mantener la calidad en el producto. En conjunto el desarrollo de nuevas tecnologías en electrónica, equipo de computación y modelado han hecho un gran avance para la automatización de este proceso. Avances en teorías de control como control adaptativo, control robusto, o sistemas difusos han provisto a investigadores e ingenieros con nuevas herramientas para el desarrollo de sistemas de control que han ayudado a alcanzar las metas de calidad y producción. Un ejemplo de modelado es por principios físicos, por ejemplo, el trabajo considerando un vector de una dimensión alta en el cual la relación entre el sistema dinámico del proceso de laminación y las fallas es encontrado mediante un análisis de búsqueda [1]. Un modelo matemático de un molino de laminación en frío que predice la vibración mecánica es propuesta en [2] este modelo fue usado para el análisis del máximo desplazamiento y amplitud la frecuencia mecánica en dirección horizontal y vertical, la simulación ayuda a disminuir la vibración mecánica. En [3], un modelo similar basado en forma de series de Fourier fue propuesto para modelar las vibraciones mecánicas; en este caso, la malla de engranes y la frecuencia predicha fue similar a la determinada en los experimentos. Sung et al. [4] propuso un control de una tira de coronas de un modelo dinámico del sistema del molino de laminación, este

modelo fue realizado un modelo de deformación del metal y un modelo de rodillos dinámicos. El controlador mejoro el desempeño del proceso en precisión y calidad de la chapa. Un modelo matemático de un molino tipo "tándem" de un molino de laminación en frio basado en un desarrollo de simulaciones para el control fue propuesto usando relaciones empíricas en [5]. Para la implementación de una teoría de control efectiva, un vector de regresión cuadrática (SVR) es propuesto como método de medida. Una detección efectiva de defectos de planeza con rápida y generalizada es reportada como resultado [6] generando una reducción en los esfuerzos de diseño, mantenimiento y tiempo de comisionamiento. En [7] un modelo no lineal de un molino tándem de laminación en frio es propuesto para el control de espesor. Otro trabajo similar fue realizado en [8] donde mediando el análisis numérico, el modelo se usa para la predicción del perfil de la chapa en un molino continuo considerando un paso inicial antes del laminado en frio.

1.2 Delimitación del problema.

Los procesos de manufactura como procesos multivariables complejos están sometidos a diversos fenómenos propios de las relaciones multicausal-efecto, donde se conoce que la relación por su misma naturaleza es no lineal. El proceso de laminación conlleva una interacción entre las siguientes variables:

- Propiedades del lubricante (saponificación, ph, estabilidad, etc).
- Características de los cilindros de laminación (rugosidad, corona, diámetros, refrigeración).
- Características del material procesado (dureza, reducción total, características del proceso anterior, rugosidades).
- Estado de los equipos de medición y control (celdas de cargas, tensiómetros, taco generadores, encoder de posición).

- Referencias de actuadores de velocidad y posición para un material con un atributo determinado (espesor, grado de reducción, etc.).

La variable determinante de la conjugación de todo este proceso multivariable es el coeficiente de fricción que determina la fuerza de laminación y las referencias de los actuadores electromecánicos necesarios para conseguir los atributos dimensionales y superficiales que el cliente requiere, las desviaciones de cualquiera de las variables antes mencionadas puede causar un efecto en el coeficiente de fricción que se manifiesta con los siguientes síntomas:

- Variaciones en la calidad superficial (Patinado, líneas de Luders).
- Variaciones en el espesor de la chapa.
- Variaciones en la forma del material (centro ondulado, borde ondulado).
- Variaciones en la generación de finos de hierro.
- Variaciones en la corriente de laminación limitando la velocidad del tren y por ende la productividad.

La información para la elaboración del modelo es recolectada directamente de la instrumentación de campo en el laminador, donde la información es dirigida a diferentes sistemas, por lo cual uno de los principales pasos para la elaboración del modelo es recolectar la información y concentrarla en una sola base de datos, para lo cual se realizó un análisis de esta información para asegurar que los datos coincidan para cada producto elaborado en el proceso. Gran parte del tiempo se invierte en asegurar que los datos que utilizamos se encuentren en un orden correcto y que estos sean relevantes para el proceso, ya que en una población tan grande de datos puede llegar a ser abrumador y difícil de tratar para obtener información, incluso para sistemas como lo son las redes neuronales y los algoritmos genéticos. Para la depuración de la información se utilizó el método de análisis de las componentes principales, el cual es una

herramienta muy poderosa para este tipo de análisis, ya que puede bajar el orden del problema de manera muy significativa sin la pérdida de información significativa, esta herramienta ortogonaliza las componentes de los vectores de entrada para que estas ya no estén correlacionadas, y en vector resultante contiene las componentes que aportan mayor variación en orden, lo cual permite eliminar las componentes que contribuyen de menor manera al modelo. De esta manera se obtiene una base de datos que simplifica el modelo y se encuentra acondicionada de manera ideal para los problemas de sistemas inteligentes.

1.3 Antecedentes.

Redes Neuronales fueron utilizadas para la optimización de parámetros en controles de lazo abierto en molinos de laminación[9]. Así como también para predicción de atributos como el límite elástico, la fuerza de rolado y el torque de rolado. En el trabajo de Pourcina [10] parámetros dominantes en el proceso de rolado son calculados mediante datos experimentales. Técnicas como computación evolutiva tienen capacidad de manejar modelos complejos no lineales, así como la regresión simbólica vía programación genética es usada para desarrollar modelos matemáticos relativamente simples. En [11] un modelo matemático de la separación entre los rodillos y la chapa en un molino reversible fue propuesto usando algoritmos genéticos.

Un sistema de laminación consiste en una serie de molinos que comprimen la lámina un determinado porcentaje de su espesor, la lámina a la entrada del molino puede encontrarse a temperatura ambiente o suficientemente caliente para facilitar su deformación, esto clasifica este proceso en laminado en frío y laminado en caliente.

En el proceso de realizar la disminución de espesor, la lámina está expuesta a fuerzas de compresión y tensión, las cuales deberán presenciarse de manera uniforme al ancho de la lámina para poder obtener una deformación uniforme a lo largo de esta, para lo cual se trata de controlar

la fuerza de compresión aplicada, así como la fuerza de tensión que sufre el rollo de lámina. La fuerza aplicada depende de diversos factores, los cuales tienen una relación no lineal en la parte elástica con la deformación que se sufre, más estos factores son conocidos, identificables y medibles. En [12] el análisis de las componentes principales es usado para interpretar múltiple información de un proceso de laminación en frío. En [13] redes neuronales son usadas para modelar una efectiva matriz teórica optimizada mediante "particle swarm optimization" la cual es evaluada mediante una función radial optimizada por grupos.

Herramientas inteligentes tales como las Redes Neuronales están siendo utilizadas para tareas de mejoramiento del desempeño en áreas como el establecimiento de parámetros óptimos en los controles de lazo abierto para los molinos de laminación [9], así como para la predicción de atributos específicos en la lámina como el flujo de la tensión, el límite elástico, la fuerza de laminación o el torque de laminación [14, 15, 16, 17, 18].

Para el diseño de los equipos utilizados en laminación se han creado modelos matemáticos empíricos, más estos muestran muchos problemas debido a la complejidad y la no linealidad de estos. Utilizando técnicas de computación evolutiva, tales como regresión simbólica mediante programación genética o algoritmos genéticos, se lograron desarrollar modelos relativamente simples para molinos de laminación en caliente [19, 20].

Para la predicción de textura en el acero en producción se han implementado sistemas de redes neuronales los cuales permiten una rápida predicción en línea, esto lo logran acercando el problema mediante un marco de Bayesiano de la red neuronal que toma en cuenta 23 parámetros [21].

La finalidad de esta investigación es desarrollar herramientas que permitan la obtención de un modelo afín a este proceso mediante la integración de las herramientas inteligentes [21].

Al momento existe una gran variedad de sistemas inteligentes, para esta investigación pondremos en prueba tres sistemas, los cuales son los siguientes:

- Red neuronal perceptrón con entrenamiento backpropagation.
- Red neuronal de máxima sensibilidad.
- Regresión simbólica alpha-beta.

Si bien cada uno de los sistemas inteligentes es sabido que tienen diferentes características y cualidades esta investigación tratara enfocada en la precisión que cada uno tiene para modelar el proceso. Sin embargo, serán mencionadas las diferentes características de cada uno, así como las ventajas de uno sobre otro. Los comparadores utilizados serán los indicadores estadísticos: Error cuadrático medio, R2pred y el PRESS, los cuales seran referidos mas adelante; los cuales tienen gran aceptación en la comunidad científica.

1.4 Hipótesis.

Mediante la selección de un Sistema Inteligente (SI) y el análisis de los datos de proceso, será posible predecir un atributo de calidad en la lámina.

1.5 Objetivo General.

Obtener un modelo capaz de representar el proceso de laminación en frío de un molino tipo tándem, mediante la entrada de las variables del proceso y como salida una variable de calidad de la lámina de acero.

1.5.1 Objetivos Particulares.

Objetivo 1.1: Obtener una base de datos de proceso que se encuentre de manera óptima para su interpretación y depurada, para su uso en los algoritmos a utilizar.

Objetivo 1.2: Representar el modelo del sistema mediante distintos sistemas inteligentes (SI) y comparar su desempeño para elegir el mejor modelo.

Capítulo 2

Marco Teórico.

2.1 Procedimiento.

2.1.1 Diagrama a bloques.

El procedimiento para la implemetación de esta investigación esta descrito de acuerdo a la estructura mostrada en la figura 2.1.

2.2 Análisis de las componentes principales.

Cualquier tema relacionado con la estadística está basado en la idea de que existe un enorme banco de datos, y esta trata de hacer un análisis en términos de relaciones entre puntos individuales en este banco de datos.

En algunas situaciones este vector de entrada es grande pero los componentes de los vectores están altamente relacionados o son redundantes. Para esto es muy útil reducir el tamaño de los vectores de entrada y el análisis de las componentes principales (ACP) es un método efectivo para realizar esta tarea. Esta técnica ortogonaliza las componentes de los vectores de entradas

Tabla 2.2.1: Información utilizada.

x	y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2.0	1.6
1.0	1.1
1.5	1.6
1.1	0.9

para que estas dejen de estar correlacionadas unas con otras, ordena el resultado de las componentes ortogonales para que estas con mayor variación aparezcan primero; y poder eliminar las componentes que contribuyen menos a la variación del banco de datos.

La principal ventaja del análisis de componentes principales es que una vez encontrados los patrones en los datos, y compriendo los datos, i.e. reduciendo el número de dimensiones, sin perder demasiada información. Para procesar la información con análisis de las componentes principales se siguen los siguientes pasos:

1. Obtener información. Para ilustrar los pasos se utiliza un ejemplo con datos con solo dos dimensiones, para poder realizar gráficos que ejemplifiquen lo ocurrido en cada paso. La información que se utilizó se encuentra ilustrada en la tabla 2.2.1 así como en la figura 2.2 de esta información.
2. Extraer la media. Para hacer que el análisis de la componente principal funcione de manera adecuada tenemos que sustraer la media de cada dimensión de los datos mostrados en la tabla 2.2.2. La media sustraída es el promedio en cada dimensión. Por lo cual todos los valores tienen, y esta es sustraída en cada una de las dimensiones de los datos. Como resultado de esto, se genera un banco de datos el cual tiene como media cero.
3. Calcular la matriz de covarianza. La matriz de covarianza para un banco de datos con n

Tabla 2.2.2: Información con la media sustraída.

x	y
0.69	0.49
-1.31	-1.21
0.39	0.99
0.09	0.29
1.29	1.09
0.49	0.79
0.19	-0.31
-0.81	-0.81
-0.31	-0.31
-0.71	-1.01

dimensiones está dada por la ecuacion 2.2.1.

$$C^{n \times n} = (c_{i,j}, c_{i,j} = cov(Dim_i, Dim_j)) \quad (2.2.1)$$

Donde $C_{n \times n}$ es la matriz con n filas y n columnas, Dim_x es la x th dimensión. Esta fórmula indica que existe un banco de datos de dimensión $-n$, por lo que la respuesta a es una matriz cuadrada de dimensión n , y cada entrada a la matriz es el resultado de calcular la covarianza entre dos dimensiones por separado. Para el banco de datos utilizado en el ejemplo el cual es de dos dimensiones genera la siguiente matriz de covarianza:

$$cov = \begin{bmatrix} .616555556 & .616555556 \\ .615444444 & .716555556 \end{bmatrix}$$

Debido a que los elementos que no se encuentran en las posiciones de la diagonal son positivos esperaremos que las variables x y y incrementen conjuntamente.

4. Calcular los eigenvectores y eigenvalores de la matriz de covarianza.

Gracias a que contamos con una matriz de covarianza cuadrada podemos calcular los eigenvectores y eigenvalores para esta matriz. Estos son de suma importancia ya que cuentan con informacion de suma importancia sobre nuestro banco de datos.

$$eigenvalores = \begin{bmatrix} .0490833989 \\ 1.28402771 \end{bmatrix}$$

$$\text{eigenvectores} = \begin{bmatrix} -.735178656 & -.677873399 \\ .677873399 & -.735178656 \end{bmatrix}$$

Es importante recalcar que los eigenvectores se encuentran normalizados, ya que esto es muy importante para el análisis de componente principal.

En la figura 2.3 se observa que la información tiene un patrón bastante definido. Como se menciona desde la matriz de covarianza, las dos variables incrementan conjuntamente. También se observa las líneas punteadas que representan los eigenvectores y nos proveen información acerca del patrón que siguen los datos. Podemos observar como uno de los eigenvectores pasa por en medio de los puntos dibujando lo que sería la línea de mejor ajuste. Este eigenvector muestra como los dos grupos de datos están relacionados a lo largo de esta línea. El segundo eigenvector muestra el otro, menos importante, patrón en la información, que aunque todos los patrones siguen a la línea principal, estos están fuera de la línea por alguna cantidad. Gracias al proceso de tomar los eigenvectores de la matriz de covarianza, se extraen las líneas que caracterizan la información. Los siguientes pasos transforman la información para que sea expresada en términos de estas líneas.

5. Escoger las componentes y formar un vector mejor. Este paso es cuando la compresión de datos y la reducción de dimensión suceden. Si observamos los eigenvectores y eigenvalores de la sección anterior, notaremos que los eigenvalores son valores diferentes. De hecho, el eigenvector con el mayor valor es la componente principal de nuestro banco de datos. Generalmente, una vez que encontramos los eigenvalores de la matriz de covarianza, el siguiente paso es ordenarlos en orden descendente. Esto muestra las componentes en orden de importancia. Con esto se procede a ignorar las componentes que tienen menos significado. A pesar de que se pierde información debido a que los eigenvalores son valores pequeños no se pierde demasiada.

Finalmente para crear el vector definitivo se toman los eigenvectores que se desea utilizar y se construye una matriz con estos eigenvectores en las columnas, como sigue:

$$VectorCaracteristico = (eig_1, eig_2, eig_3...eig_n) \quad (2.2.2)$$

En el ejemplo que solo se cuenta con dos eigenvectores y se elige el eigenvector con mayor información el cual solo contará con una sola columna, dejando de lado el eigenvector con menor información.

$$VectorPrincipal = \begin{bmatrix} -.677873399 \\ -.735178656 \end{bmatrix}$$

2.3 Redes Neuronales.

Las redes neuronales artificiales fueron inspiradas en reconocimiento del hecho que el cerebro humano procesa de manera diferente la información que una computadora.

De acuerdo a Haykin [22], una red neuronal es un procesador masivo distribuido masivamente con una inclinación por conocimiento experimental haciendo esté disponible de usar. Se asemeja al cerebro en dos aspectos:

1. Su conocimiento se adquiere a través de una red por medio de un proceso de aprendizaje.
2. Las conexiones inter-neurona conocidas como pesos sinápticos son usadas para almacenar el conocimiento.

Las redes neuronales son especialmente adecuadas para clasificar y aproximar y/o mapear problemas que son tolerantes a algo de imprecisión y tienen numerosos datos de entrenamiento disponible. Casi cualquier mapeo entre espacios vectoriales puede ser aproximado con una precisión arbitraria por medio de Redes neuronales tipo "feed forward" si hay suficiente información y recursos computacionales. Desde el punto de vista del aprendizaje o fase de codificación, las

redes neuronales pueden clasificarse en sistemas supervisados y no supervisados. Las redes supervisadas tipo "feed forward" son típicamente utilizadas en tareas de aproximación de funciones. Ejemplos específicos de estas:

- Red neuronal "Back Propagation", también llamada perceptrón.
- Red neuronal de Base Radial.
- Red neuronal Máxima Sensibilidad.

En este trabajo se utiliza la red Back propagation y máxima sensibilidad.

2.3.1 Red neuronal tipo perceptrón.

Una red neuronal es usada para modelar el proceso de rolado; sin embargo, otros enfoques pueden ser utilizados, no obstante, las redes neuronales tienen ciertas características que son deseables, como la robustez y la precisión en la aproximación. Como punto en contra sabemos que las redes neuronales son cajas negras y no explican la relación entre variables como lo realiza la regresión lineal, sin embargo, la precisión para predecir de la regresión es muy baja comparada contra la red neuronal. La figura 2.4 nos muestra las partes principales de una red neuronal.

Un perceptrón multicapas con entrenamiento "back propagation" será usado. Esta estructura consiste en un conjunto de capas compuestas de neuronas. Cada neurona ejecuta un conjunto de operaciones que emulan las funciones de una neurona natural de manera simplificada. La integración de las señales de las entradas capturadas por las dendritas, es hecha por el soma de la neurona. La neurona artificial consiste en la sumatoria de las señales de entrada representadas por los elementos de un vector. El peso que regula cada señal de entrada, emula la función de la sinapsis en las neuronas naturales. Cuando la señal integrada es superior al nivel del umbral, una señal de excitación a la salida de la neurona es generada y dispersada a las neuronas siguientes

por el axón. De otra manera la salida es inhibida. Esta operación es representada por una función de activación no lineal (por ejemplo, una función sigmoidal). Los pesos representan la sinopsis y determinan el desempeño de la red neuronal. La regla de aprendizaje es usada para actualizar estos pesos. Muchas reglas de aprendizaje pueden ser consideradas; sin embargo, la regla de aprendizaje "back propagation" con el momentum es comúnmente utilizada. Las siguientes ecuaciones ilustran la implementación de la regla de aprendizaje back propagation considerando un perceptron multicapas con i entradas, m neuronas en la capa oculta y una salida Y_N . Como función de activación es usada la función sigmoidal (2.3.1).

$$f_a(x) = 1/(1 + \exp^{-x}) \quad (2.3.1)$$

Dos operadores son definidos, la función de activación y entrenamiento en la red neuronal. El objetivo es el ajuste de los pesos en cada capa (las matrices W_M y W_O) para minimizar el error entre la respuesta deseada (representada por el vector Y_D) y la respuesta de la neurona (Y_N) considerando el mismo vector de entrada (X). El primer paso es la activación de la capa interna de neuronas que es hecha usando las ecuaciones 2.3.2 y 2.3.3. Asumiendo que un par de patrones (X_p , Y_{Np}) fueron considerados. Y_{Np} es la salida de la red neuronal generada por el patrón de entrada X_p y la salida deseada es el patrón Y_{Dp} .

$$N_M(m) = \sum_{i=1}^I W_M(m, i)X(i) \quad (2.3.2)$$

$$R_M(m) = f_a(N_M(m)) \quad (2.3.3)$$

Solo tenemos una salida Y_N por lo que la función de activación de una sola neurona es necesaria en las ecuaciones 2.3.4 y 2.3.5.

$$N_O = \sum_{m=1}^M W_O(m)R_M(m) \quad (2.3.4)$$

$$Y_N = f_a(N_O) \quad (2.3.5)$$

El entrenamiento de la red neuronal es realizado mediante el ajuste de los pesos de la capa oculta (W_M) y la salida de los pesos (W_O). Primero, el error entre el valor deseado de salida Y_D y la salida de la red neuronal Y_N es calculado 2.3.6. El ajuste de los pesos se realiza usando la ecuación 2.3.7. El cálculo del momento requiere los valores de los pesos antes de ajustar, así que este cálculo es hecho usando la ecuación 2.3.8.

$$\delta_O = (Y_{D_p} - Y_{N_p})f'_a(N_O) \quad (2.3.6)$$

$$W_O(m) \leftarrow W_O(m) + \eta\delta_O N_M(m) + \alpha W_{OA}(m) \quad (2.3.7)$$

$$W_{OA}(m) = \eta\delta_O N_M(m) \quad (2.3.8)$$

Los pesos internos requieren una señal de error calculada usando la ecuación 2.3.9. Este cálculo requiere una función de activación derivada 2.3.10. Los pesos internos son actualizados usando la ecuación 2.3.11 y los pesos después del ajuste son requeridos usando la ecuación 2.3.12.

$$\delta_m(m) = f'_a(N_m(m))\delta_O W_O(m) \quad (2.3.9)$$

$$f'_a(x) = f_a(x)(1 - f_a(x)) \quad (2.3.10)$$

$$W_M(m, i) \leftarrow W_M(m, i) + \eta\delta_m X(i) + \alpha W_{MA}(m, i) \quad (2.3.11)$$

$$W_M(m) = \eta\delta_m(m)X(i) \quad (2.3.12)$$

Siguiendo las ecuaciones 2.3.6 a la 2.3.12 para un par de patrones P, todos los patrones son ajustados hasta que el error sea minimizado (ecuación 2.3.13).

$$E_G = (1/2) \sqrt{\sum_{p=1}^{N_p} (Y_{D_p} - Y_{N_p})^2} \quad (2.3.13)$$

2.3.2 Red neuronal de máxima sensibilidad.

La arquitectura de red neuronal utilizada en este enfoque es llamada red neuronal de máxima sensibilidad (MSNN) [23]. La estructura comúnmente se asemeja a la mostrada en la figura 2.5. Las entradas están agrupadas en X , W contiene las entradas de los pesos y AC es dedicada a los pesos de salida, donde al inicio ambos pesos se encuentran sin datos. Los operadores de la red neuronal dependen de la máxima sensibilidad, este es un estado en el cual una neurona, la cual tiene una activación almacenada en SN , tiene el mayor valor cuando una entrada es presentada a todas las neuronas. Este estado es determinado de la siguiente manera:

$$SN_n = fa(\sqrt{\sum_{i=1}^{TI} (W_{n,i} - X_i)^2}, 0, \lambda) \quad (2.3.14)$$

$$y = fa(x, cm, \lambda) = \exp(-(x - cm)^2/\lambda) \quad (2.3.15)$$

Donde TI consiste de el total de entradas, $i = 1, \dots, TI$ y $n = 1, \dots, CN$ y CN es el contador de neuronas activas. La función de activación fa tiene un parámetro llamado λ que determina la sensibilidad de cada neurona artificial. El centro de masa cm es establecido en cero debido a que la entrada X con un valor cercano al previo W es esperado, y genera un alto valor de sensibilidad almacenado en SN . Podemos definir $maxp(V) \rightarrow (p, v)$ como el procedimiento que nos regresa el máximo valor v y la posición p de todos los elementos de cualquier vector V ; $minp(V) \rightarrow (p, v)$ es similar a $maxp$ pero nos regresa el valor mínimo de todos los elementos y su posición. El estado de máxima sensibilidad es establecido determinando una neurona ganadora w usando los valores de sensibilidad SN de todas las neuronas. La posición de la neurona ganadora es almacenada en w y su valor tendría que ser mayor al margen preestablecido de sensibilidad ms , si este es el caso, máxima sensibilidad estará activo, de otro modo este estado estará inactivo. La tabla 2.3.3 muestra el procedimiento para determinar máxima sensibilidad.

Se define un peso de uso PU que determinará las veces en que cada neurona ha sido un

Tabla 2.3.3: Determinación de máxima sensibilidad.
 Determinación de máxima sensibilidad.

-
- 1) $maxp(SN) \rightarrow (p, v)$
 - 2) if $v > ms$ then
 - 3) $w = p; s = 1$ (maximum sensibility)
 - 4) else
 - 5) $s = 0$
-

ganador; sin embargo, cada momento un el peso es reducido mediante una constante llamada factor de olvido (fo) como se muestra:

$$PU_n = PU_n - fo \quad (2.3.16)$$

Usualmente una pequeña cantidad de parámetros está definida para controlar la generalización y especialización de MSNN. Existen dos modos de operación en MSNN: *activación y aprendizaje*, y los dos dependen del estado de sensibilidad, activación y aprendizaje en cada instante, también la actualización de los pesos se realiza en todo momento.

Activación

En máxima sensibilidad, la salida y es calculada de la siguiente manera:

$$y = SN_w AC_w \quad (2.3.17)$$

$$PU_w = 1 \quad (2.3.18)$$

Donde w representa el peso de cada neurona.

Cuando no existe un estado de máxima sensibilidad, hacemos una aproximación o una contribución usando la respuesta de todas las neuronas SN:

$$y_o = \frac{\sum_{n=1}^{CN} (AC_n SN_n)}{\sum_{n=1}^{CN} SN_n} \quad (2.3.19)$$

Aprendizaje

El aprendizaje también depende de la máxima sensibilidad. Un par de patrones de entrada y salida X y YD es dado a la red neuronal para aprendizaje. La entrada X activa la red neuronal y establece el máximo estado de sensibilidad; si este estado es activado, entonces los pesos de la neurona ganadora w son ajustados. En este estado, un cálculo en línea de la media es realizado usando las siguientes ecuaciones:

$$W_{w,i} = (W_{w,i} + X_i)/2 \quad (2.3.20)$$

$$AC_{w,o} = (AC_{w,o} + YD_o)/2 \quad (2.3.21)$$

Para w_i tal que $i = 1...TI$ y $o = 1...TO$ para el total de salidas de la red neuronal.

$$W_{w,i} = X_i \quad (2.3.22)$$

$$AC_{w,o} = YD_o \quad (2.3.23)$$

Aprendizaje cuando no existe máxima sensibilidad activa diferentes procedimientos, primero es necesario verificar si existen neuronas libres checando el contador CN , después se realiza el almacenaje. La tabla 2.3.4 muestra un pseudo-código del procedimiento. donde TN es el máximo número de neuronas.

Tabla 2.3.4: Aprendizaje sin máxima sensibilidad.

Aprendizaje sin máxima sensibilidad.

- 1) if $CN > TN$ then
- 2) $CN \leftarrow CN + 1$
- 3) $w = CN$
- 4) else
- 5) $(w, p) \leftarrow \text{minp}(PU)$
- 6) end
- 7) $X_w \leftarrow X_i$
- 8) $W_{w,i} \leftarrow X_i$
- 9) $AC_{w,o} = YD_o$

Si no existen neuronas disponibles ó $CN > TN$ tendremos que buscar la neurona menos usada, después de esto, se sustituye la vieja información por la nueva. Tres parámetros pueden ser establecidos: fo, ms y λ . El máximo número de neuronas TN puede ser cambiado en cada puesta en marcha de la MSNN, [$ms \in [0.8, 0.95]$] y [$\lambda \in [0.05, 0.5]$], ambos parámetros controlan el desempeño de la red neuronal para generalización y especialización. El sistema usa MSNN para aprendizaje y considera n datos. Es necesario entrenar la MSNN con datos confiables según lo posible. MSNN está lista para ser integrada al sistema y aprender nuevos valores de los sensores o predecir en caso de falla de estos.

2.4 Regresión simbólica.

En este enfoque, una ecuación matemática es representada por la combinación de operadores α y β . Un operador α es definido como una función que requiere solo un argumento y aplica una operación matemática. Considerando un estudio de diferentes modelos de procesos reales, trece operaciones son consideradas como operadores α (ver Tabla 2.4.5).

Un operador α usa dos parámetros determinados por números reales llamados k_1 y k_2 y un número entero que describe la operación matemática. El operador α es definido de la siguiente manera:

Tabla 2.4.5: Operadores α y su función matemática relacionada.

Operador α	Operación matemática
1	$(k_1x + k_2)$
2	$(k_1x + k_2)^2$
3	$(k_1x + k_2)^3$
4	$(k_1x + k_2)^{-1}$
5	$(k_1x + k_2)^{-2}$
6	$(k_1x + k_2)^{-3}$
7	$(k_1x + k_2)^{1/2}$
8	$(k_1x + k_2)^{1/3}$
9	$\exp(k_1x + k_2)$
10	$\log(k_1x + k_2)$
11	$\sin(k_1x + k_2)$
12	$\cos(k_1x + k_2)$
13	$\tan(k_1x + k_2)$

$$Opr_\alpha(x, k_1, k_2) = \alpha(k_1 * x + k_2) \quad (2.4.1)$$

Donde x es una variable de entrada y α es una operación. Dependiendo del operador α seleccionado, una operación matemática específica es realizada la cual requiere solo un argumento; ejemplo, si $\alpha = 1$ la operación realizada será $(k_1 * x + k_2)$, si $\alpha = 13$ entonces la operación realizada será $\tan(k_1 * x + k_2)$. El operador α puede ser representado como un número entero. Un operador β está definido como una función que requiere dos argumentos y realiza una de las cuatro operaciones básicas, si el operador β es igual a 1 implica que el operador es la suma o $\beta(a, b) = a + b$, y si $\beta = 4$ $\beta(a, b) = a/b$.

Una configuración básica puede ser definida cuando un operador α es asignado para cada variable de entrada, después se asigna un operador β que funcionará para conectar dos operadores α . Usualmente, una configuración simple es suficiente para la mayoría de los casos para realizar la regresión, por ejemplo:

$$y = \beta_{n-1}(\dots \beta_2(\beta_1(\alpha_1(x_1), \alpha_2(x_2))), \dots \alpha_n(x_n)) \quad (2.4.2)$$

La conectividad está controlada por una variable binaria para seleccionar el número de operadores α usados para cada variable de entrada, inclusive puede no requerirse ninguna variable si este fuera el caso. Un máximo de tres operadores α son determinados para cada variable (de α_1 a α_3), entonces, si existen dos variables, una configuración con seis operadores α y cinco operadores β es definida como se muestra en la ecuación 2.4.3. Dependiendo del número de variables de entrada una configuración es usada.

$$\begin{aligned}
y = & \beta_5(\beta_4(\beta_3(\beta_2\beta_1(b_1\alpha_1(x_1k_{11}+ \\
& k_{21}), b_2\alpha_2(x_1k_{12} + k_{22})), b_3\alpha_3(x_1k_{13}+ \\
& k_{23})), b_4\alpha_4(x_2k_{14} + k_{24})), b_5\alpha_5(x_2k_{15}+ \\
& k_{25})), b_6\alpha_6(x_1k_{16} + k_{26}))
\end{aligned} \tag{2.4.3}$$

La representación depende de la configuración utilizada o del número de variables requerida. El número de variables determina la configuración como se mencionó anteriormente. Un vector con números reales normalizados puede ser usado para representar la conectividad o el número de operadores α por variable, los operadores α y β así como los parámetros k de los operadores α . Como ejemplo consideremos una configuración de dos variables, donde los siguientes parámetros pueden ser extraídos: Parámetros de conectividad serian $b_1, b_2, b_3, b_4, b_5, b_6$ y pertenecen a un vector binario; parámetros de los operadores son $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ los cuales son números enteros; y los parámetros reales son $k_{11}, k_{21}, k_{12}, k_{22}, k_{13}, k_{23}, k_{14}, k_{24}, k_{15}, k_{25}, k_{16}, k_{26}$. Solo un número real es usado para conectividad debido a que este valor es convertido a un valor entero y después convertido a un vector binario de nuevo y extraemos los valores b usando su posición correspondiente de los elementos del vector binario. Los parámetros k son asignados directamente al valor real del vector de representación. Cada operador α y β es un número entero, por lo que se requerirá del siguiente procedimiento para tener su valor de representación V :

$$\alpha = \lceil V(i) * 13 \rceil \quad (2.4.4)$$

$$\beta = \lceil V(i) * 4 \rceil \quad (2.4.5)$$

Donde $\lceil \cdot \rceil$ es la función que acota al entero superior. Recordando que tenemos definidos trece operadores α que se encuentran en la tabla 2.4.5 y cuatro operadores β que representan las operaciones aritméticas básicas. Consideremos el siguiente ejemplo de decodificación; el vector de parámetros es $V = [0.482, 0.963, 0.529, 0.043, 0.294, 0.284, 0.786, 0.569, 0.156, 0.561, 0.315, 0.722, 0.518, 0.201, 0.715, 0.017, 0.911, 0.306, 0.275, 0.797, 0.562, 0.898, 0.021, 0.793]$. La conectividad está definida en $V(1)$, donde tenemos $\lceil (0.482 * 2 \exp 2 * 3 - 1) \rceil = \lceil 15.424 \rceil$ dos variables y tres operadores α por variable, cero es considerado como sin conexión. El vector binario es $[0 \ 0 \ 1 \ 0 \ 1 \ 1]$ donde la variable x_1 está conectada por un operador α y la variable x_2 está conectada por dos operadores α . Todo operador α y β es extraído por decodificación en cada elemento del vector V .

$$\alpha_1 = \lceil V(2) * 13 + 0.5 \rceil = \lceil (0.963 * 13) \rceil = 13 \text{ representa una función "tan"}$$

$$\alpha_2 = \lceil V(3) * 13 + 0.5 \rceil = \lceil (0.529 * 13) \rceil = 7 \text{ representa una función "exponencial" de la función -2}$$

$$\alpha_3 = \lceil V(4) * 13 + 0.5 \rceil = \lceil (0.043 * 13) \rceil = 1 \text{ representa una función "lineal"}$$

$$\alpha_4 = \lceil V(5) * 13 + 0.5 \rceil = \lceil (0.294 * 13) \rceil = 4 \text{ representa una función "inversa"}$$

$$\alpha_5 = \lceil V(6) * 13 + 0.5 \rceil = \lceil (0.284 * 13) \rceil = 4 \text{ representa una función "inversa"}$$

$$\alpha_6 = \lceil V(7) * 13 + 0.5 \rceil = \lceil (0.786 * 13) \rceil = 11 \text{ representa una función "sin"}$$

$$\beta_1 = \lceil V(8) * 4 + 0.5 \rceil = \lceil (0.569 * 4) \rceil = 3 \text{ representa una función "multiplicación"}$$

$$\beta_2 = \lceil V(9) * 4 + 0.5 \rceil = \lceil (0.156 * 4) \rceil = 1 \text{ representa una función "adición"}$$

$$\beta_3 = \lceil V(10) * 4 + 0.5 \rceil = \lceil (0.461 * 4) \rceil = 2 \text{ representa una función "sustracción"}$$

$$\beta_4 = \lceil V(11) * 4 + 0.5 \rceil = \lceil (0.315 * 4) \rceil = 2 \text{ representa una función "sustracción"}$$

$$\beta_5 = \lceil V(12) * 4 + 0.5 \rceil = \lceil (0.722 * 4) \rceil = 3 \text{ representa una función "multiplicación"}$$

Finalmente, los parámetros $k_{(1i)} = V(i)k_{(2i)}V(i+1)$ donde $i = 13, 15, 17, 19, 21, 23$ de esta

manera tenemos el siguiente modelo matemático basado en la configuración:

$$y = x_1 - (x_2^{-1} * \sin(x_2)) \quad (2.4.6)$$

Para resolver los problemas de selección de parámetros ($k's$), la conectividad y enteros para definir las operaciones α y β usamos "Evonorm" [24, 25].

2.4.1 "Evonorm".

"Evonorm" es una simple manera de implementar una estimación de para un algoritmo de distribución. Como un algoritmo evolutivo una selección de nuevos individuos y de nuevas generaciones de población es utilizada; sin embargo, el mecanismo de cruce y la mutación es sustituido por una estimación de parámetros de una función de distribución normal. Los siguientes pasos son utilizados en "Evonorm":

1. Evaluación de la población P.
2. Selección determinista de individuos de P a PS.
3. Generación de nueva población utilizando PS.

Una población P es una matriz de tamaño I_p (total de individuos) y D_r (total de variables de decisión). Una solución es un grupo de variables de decisión y este grupo está representado como un vector real. Cada fila de la población P representa un grupo de variables de decisión. El mecanismo de selección es determinístico debido a que el mejor individuo es seleccionado. Usualmente el número de individuos seleccionados es menor que el número de población inicial. Usualmente un veinte por ciento o diez por ciento. Una variable aleatoria con distribución normal es estimada por cada variable de decisión, así una función de distribución marginal es

utilizada. Dos parámetros son estimados, la media y la desviación estándar, que es determinada usando los valores de los individuos seleccionados. La población de individuos seleccionada es una matriz P_s de tamaño I_s (total de individuos seleccionados) y D_r . Las siguientes ecuaciones 2.4.7, 2.4.8 son usadas para calcular la media y la desviación estándar considerando cada vector de la población P_s .

$$\mu_{pr} = \sum_{k=1}^{I_s} (P_{S_{pr,k}}) / I_s \quad (2.4.7)$$

$$\sigma_{pr} = \sqrt{\left(\sum_{k=1}^{I_s} (P_{S_{pr,k}} - \mu_{pr})^2 \right) / I_s} \quad (2.4.8)$$

Donde $pr=1..Dr$ Una nueva población es generada usando una estimación normal de variables aleatorias. Esto es un proceso estocástico; sin embargo, es usada algo de heurística para mantener el equilibrio entre la exploración y la explotación, para encontrar nuevas soluciones que no se encuentren necesariamente cerca de la media calculada. La mejor solución encontrada I_x al momento es llevado a la siguiente generación así que generalmente el 50% de las veces la media es usada en los cálculos y en el otro 50% del tiempo la mejor solución encontrada I_x es usada como la media como se muestra en la siguiente ecuación:

$$P_{i,pr} = \begin{cases} N(\mu_{pr}, \sigma_{pr}) & U() > 0.5 \\ N(I_{x_{pr}}, \sigma_{pr}) & \text{de cualquier otra manera} \end{cases} \quad (2.4.9)$$

La variable aleatoria $U()$ tiene una distribución uniforme, $N()$ es una variable aleatoria con una función de distribución normal.

2.5 Estadísticos

Una forma efectiva de validar un modelo de regresión es coleccionar nueva información experimental para determinar que tan bien los modelos se desempeñan en la práctica[26]. La forma más sencilla de medir es el residuo calculado mediante la diferencia $e(i)$ entre las nuevas observaciones de la respuesta del proceso $y(i)$ y la respuesta generada por el modelo $\hat{y}(i)$ (2.5.1).

$$e(i) = y(i) - \hat{y}(i) \quad (2.5.1)$$

2.5.1 Error Cuadrático Medio ("MSE")

El error cuadrático medio ("MSE") es un estimador que mide el promedio cuadrático de los errores, el hecho de que la diferencia de los errores sea elevada al cuadrado produce un valor positivo en todas las ocasiones. Debido a lo anterior buscamos que este indicador brinde un valor cercano a cero para un buen modelo.

$$MSE = \sum_{i=1}^n (y(i) - \hat{y}(i))^2 / n \quad (2.5.2)$$

2.5.2 Error Residual Predicho por Suma de Cuadrados ("PRESS")

Por sus siglas en inglés "PRESS" (predicted residual sum of squares) es una medida de qué tan bien trabaja un modelo en predecir nueva información. Usualmente es deseable tener un valor pequeño en el indicador PRESS 2.5.3. En este caso, PRESS es obtenido como validación cruzada.

$$PRESS = \sum_{i=1}^n (y(i) - \hat{y}(i))^2 \quad (2.5.3)$$

2.5.3 " $R^2 PRED$ " ("R-squared predicted")

El porcentaje de variabilidad R^2_{pred} es una medida para indicar la efectividad del modelo para predecir nuevas observaciones. Un valor cercano a uno es deseable para este indicador 2.5.4.

$$R^2_{prec} = 1 - \left(\frac{\sum_{i=1}^n (y(i) - \hat{y}(i))}{y'y - \left(\sum_{i=1}^n y(i) \right)^2} \right) \quad (2.5.4)$$

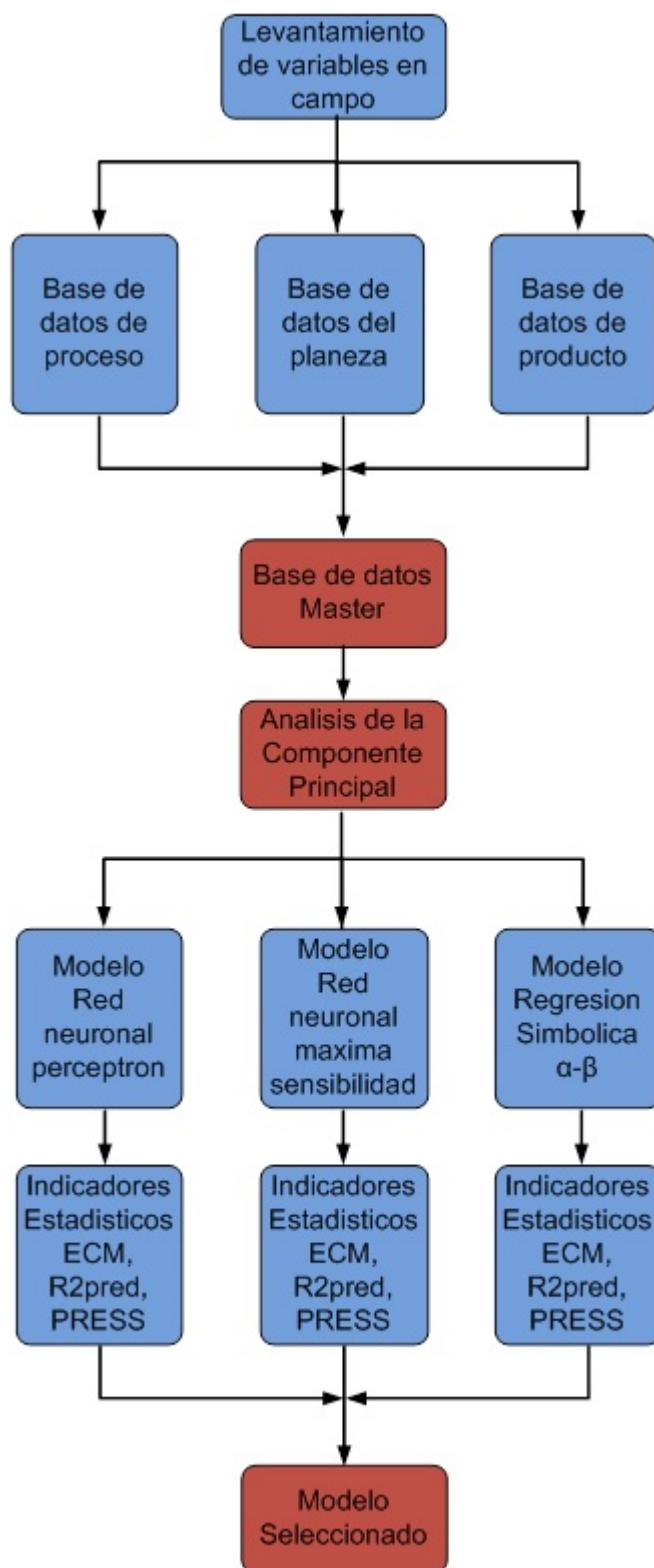


Figura 2.1: *Diagrama Bloques*

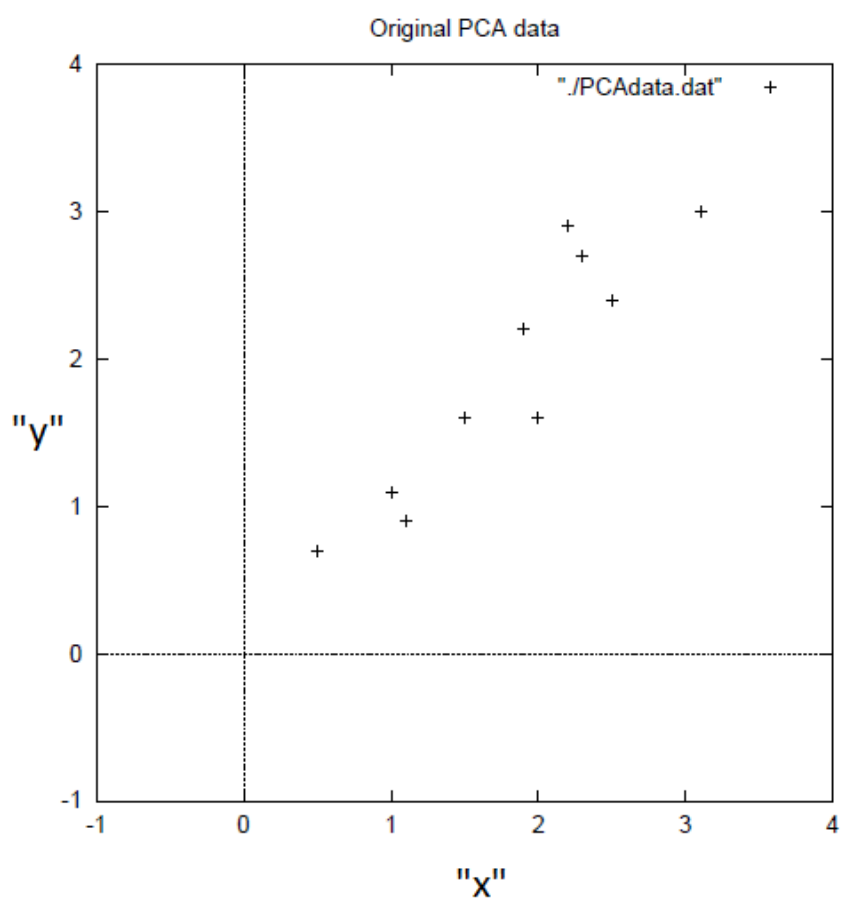


Figura 2.2: Datos del ejemplo del ACP.

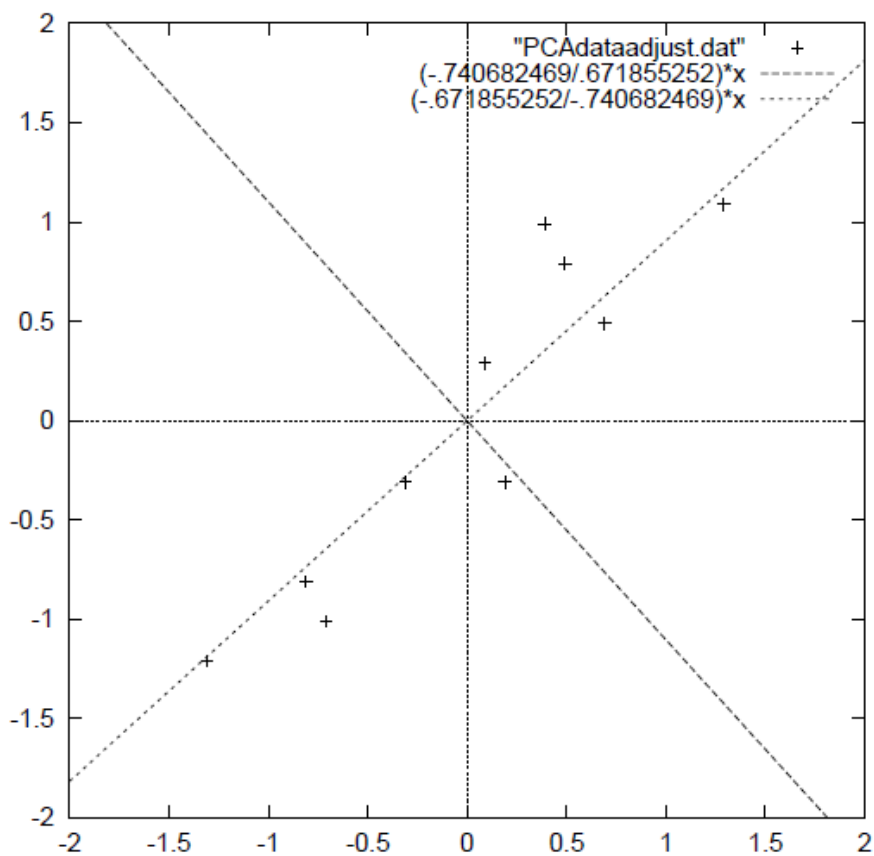


Figura 2.3: Gráfica de los valores normalizados con los eigenvectores de la matriz de covarianza.

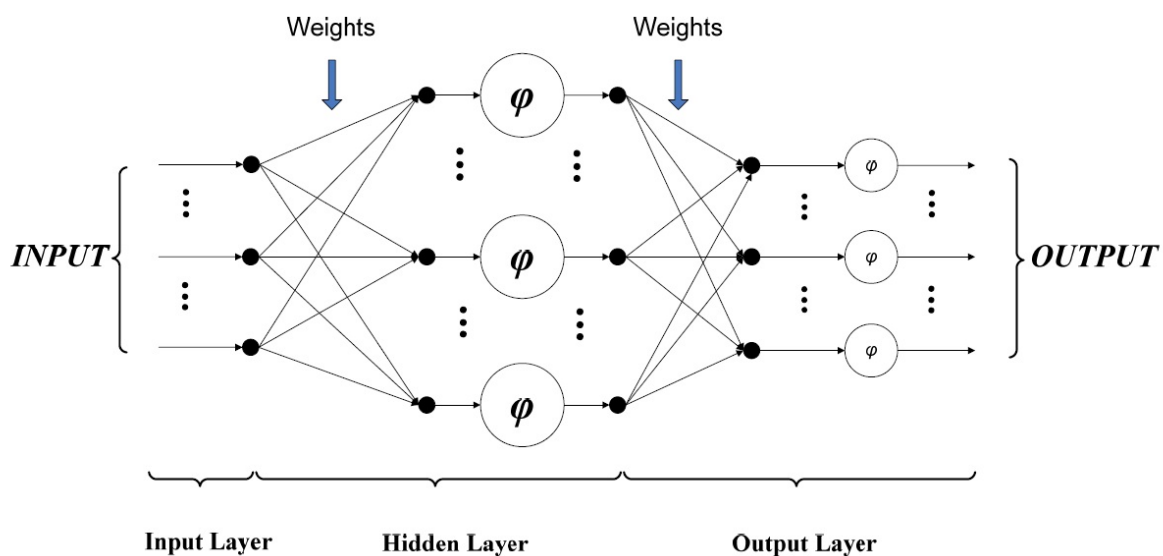


Figura 2.4: Partes principales de la red neuronal.

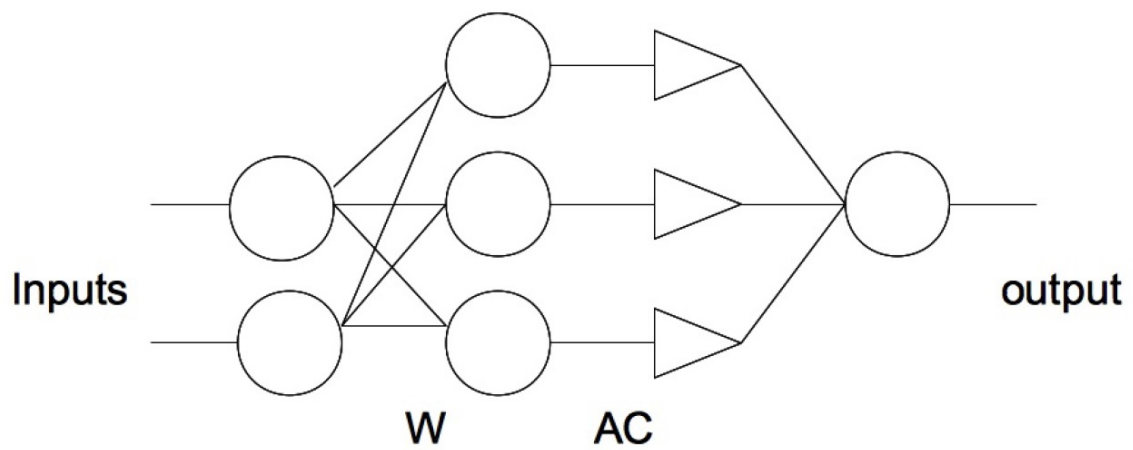


Figura 2.5: *Partes principales de la red neurona máxima sensibilidad.*

Capítulo 3

Metodología.

3.1 Proceso.

Las compañías cuentan con sistemas de recolección de información los cuales permiten conocer el desempeño de cada proceso. El producto del proceso de laminación en frío es afectado por distintos factores, algunos de ellos pueden venir de las condiciones ambientales, la calidad de las entradas y los parámetros impuestos en el proceso. Un conjunto de sensores obtienen mediciones a cada momento y la almacenan en bases de datos, de manera que el número de información disponible es muy alta. La figura 3.1 y la tabla 3.1.1 ilustran la posición de algunas de las variables usadas en el modelo representado del molino de laminación tipo tándem.

Un esquema del proceso incluyendo las variables es mostrado como modelo en la Figura 3.2

El número de variables recolectadas es superior a 100 y de cada una se cuenta con un registro de más de 50000 datos. Toda la información es almacenada por el departamento de procesos, sin embargo, la recolección es un reto en si mismo, ya que este proceso es realizado por diferentes medios y la información no registra una centralización donde se estandarice y se obtenga en un formato único. Mucha de la información debe ser recolectada de manera manual. De la información recolectada se consideran los grupos de productos más relevantes (los cuales son los que tienen mayor producción y mayor demanda). Una base de datos es generada para obtener

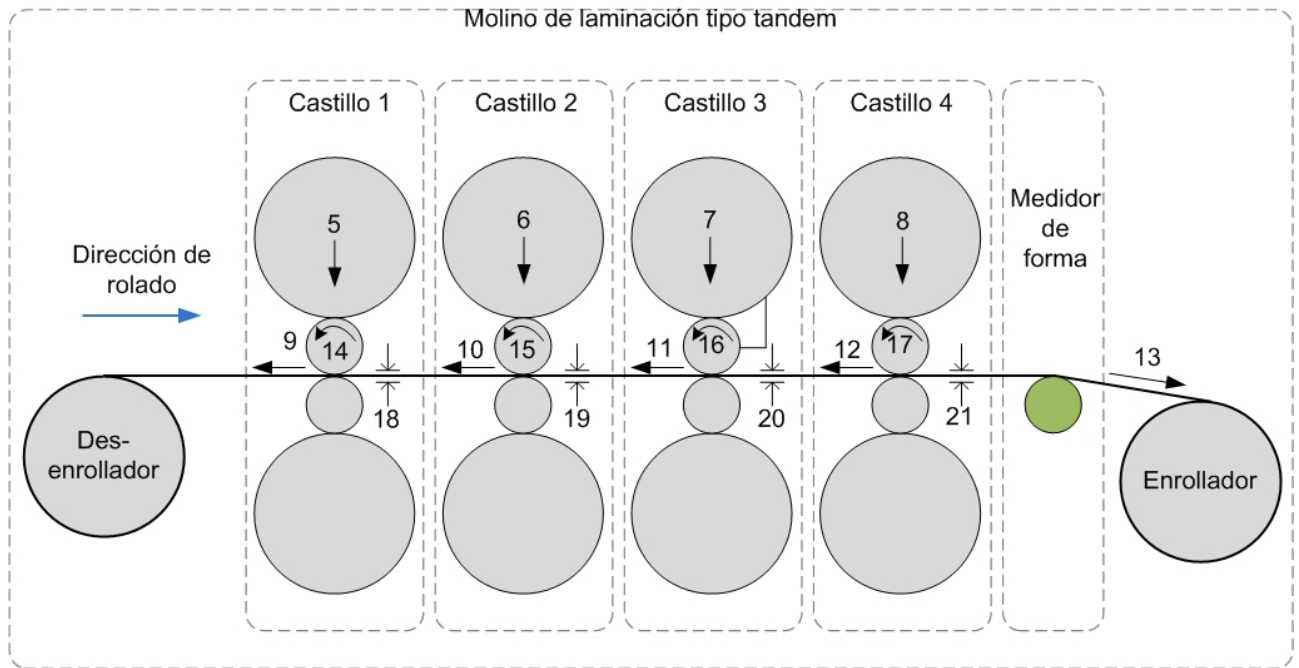


Figura 3.1: *Principales variables y parámetros de entrada y salida del molino de laminación*

datos de entrenamiento y datos para comprobación. Un total de 500 patrones fueron usados para el modelado. El proceso de laminación en frío no es la excepción, las características del producto, del proceso y el desempeño del producto final son almacenados a cada minuto, por cada día. Cada día son almacenados más de 50000 datos. De estos datos la tabla 3.1.1 resume las variables usadas para el modelo del proceso de laminación en frío. .

Las variables son seleccionadas por medio de un análisis de componentes principales para reducir de 50 variables a solo 21 considerando solo las componentes más representativas.

3.1.1 Espesor

Como primera etapa para la realización de la experimentación se optó por utilizar la variable de espesor para predicción, con el fin de obtener resultados confiables se realizaron diferentes experimentos.

El espesor de la lámina en el producto final es uno de los elementos más importantes, ya que

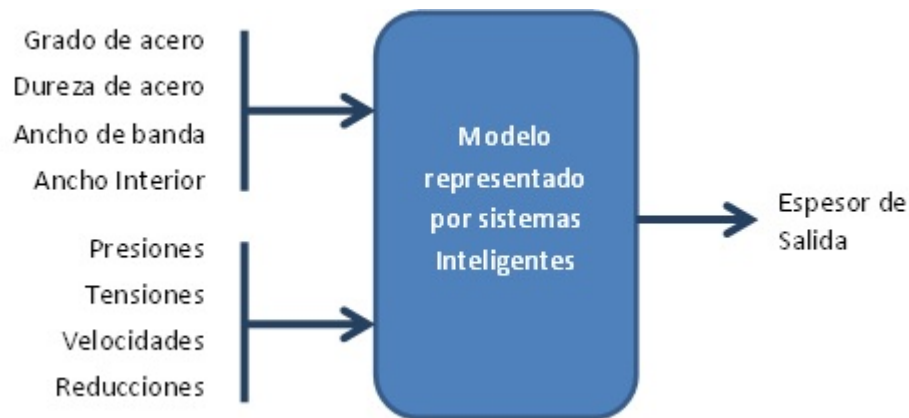


Figura 3.2: *Esquema del modelo del proceso*

este es uno de los principales clasificadores de la gama de productos, y la calidad del producto va de acuerdo a la exactitud que obtengamos en el producto final. Si bien a la salida del proceso contamos con un medidor que nos puede retroalimentar el espesor que estamos obteniendo. La obtención de un rango específico de espesor es sumamente complicada y se requiere de un sistema de control sumamente complejo ya que la interacción entre todas las variables de proceso es sumamente complicada.

Si de antemano se sabe que el proceso de laminación es un proceso donde sus variables interactúan de una manera no lineal, anticipadamente sabremos que la estabilidad y la exactitud en el espesor serán complejas y en el proceso tendremos cierta variabilidad intrínseca la cual inevitablemente afectará en el resultado final. Debido a estos factores consideramos que la predicción en estos sistemas es una buena herramienta para la operación y para el diseño de los sistemas de control.

Específicamente definimos como espesor, el grosor de la lámina en un punto específico del ancho de la lámina. Sabiendo que el perfil de la lámina varía a lo ancho debido a la naturaleza del proceso, sin embargo, idealizando el proceso este espesor se mantendría constante en toda la lámina.

Como parte de las inherencias del proceso, el ancho varía de acuerdo a múltiples causas, entre

Tipo	Número	Nombre
Inputs	1	GradoAcero
	2	Dureza_Clase
	3	AnchoBanda
	4	AnchoCentro
	5	Presión_1
	6	Presión_2
	7	Presión_3
	8	Presión_4
	9	Tensión_final
	10	Tensión_1_2
	11	Tensión_2_3
	12	Tensión_3_4
	13	Tensión_carrete
	14	Velocidad_1
	15	Velocidad_2
	16	Velocidad_3
	17	Velocidad_4
	18	Reducción_1
	19	Reducción_2
	20	Reducción_3
	21	Reducción_4
Salida	1	Salida_Espesor

Tabla 3.1.1: Entradas y salidas del proceso

estas, se espera que el ancho cuente con una deformación llamada corona, la cual tiene raíz en la manera en que el proceso aplica la fuerza sobre la lámina para su reducción en espesor. Esto se origina debido a que la presión está dada por rodillos de acero, los cuales al ser sometidos a elevadas fuerzas de tensión ceden ante la fuerza de reacción que provoca la lámina y presentan deflexión a lo largo de estos. De este fenómeno se obtienen mayores deformaciones cuanto mayor es la distancia a los apoyos de los rodillos, significando que el centro de la lámina generalmente tiene mayor espesor que los extremos de esta.

Existen diferentes maneras de compensar la flexión de los rodillos, una de los principales es la corona en el rodillo de trabajo, la cual consiste en maquinarse un perfil a lo ancho del rodillo que tendrá contacto con la lámina. Este mecanismo tiene el objetivo de compensar la flexión del rodillo con un mayor radio al centro del rodillo que en la orillas, siendo esta una manera efectiva

de lograr esta compensación.

Para lograr ver reflejado el mecanismo de corona, se toman lecturas del perfil de cada rodillo antes y después de salir del ciclo de trabajo. Estas lecturas son almacenadas en una base de datos que lleva historial de cada rodillo utilizado.

El método de corona en el rodillo intenta compensar la flexión de los rodillos, existen otras maneras de atacar este problema. Una de ellas, es el mecanismo de "bending" que consiste en aplicar una fuerza que compense la flexión de los rodillos, esta fuerza aplicada en los extremos del rodillo logra transferir la compensación de la flexión al rodillo, sin embargo este mecanismo no siempre transmite la fuerza al mismo punto donde la flexión es ocurrida, esto generado a partir de que los actuadores que compensan son independientes un lado de otro y se intenta realizar la compensación mediante controladores que aplican la potencia necesaria en cada uno de los lados.

Si bien, aquí se describen algunos de los mecanismos y métodos para intentar tener una superficie completamente plana, se llega a la conclusión de que esto es prácticamente imposible ya que el proceso en sí cuenta con una interacción de muchas variables, las cuales dificultan el obtener una superficie con un espesor específico a lo ancho de la lámina. Por lo anterior, se puede concluir que el producto puede acotarse en cierto límite, donde es considerado como exitoso el hecho de generar esta variación que afectará a otros factores y puntos en la calidad del producto. Uno de estos es la planeza.

3.1.2 Planeza

La planeza puede definirse cualitativamente como una desviación de la forma ideal, hacia una forma latente, evidente o dual. Esta desviación puede ser encontrada al cortar la lamina en cintas estrechas y subsecuentemente medir su longitud.

Los defectos en planeza o variaciones en la forma son resultado de deformaciones plásticas a

través del largo de la lámina y/o variaciones en el espesor de esta, lo cual produce variaciones inherentes en la tensión interna. Como sucede con el espesor, la planeza es afectada por diversos factores que hacen salir a la lámina del ideal en término de planeza.

En un laminado ideal el espesor y el ancho a través de toda la longitud se mantendría constante, lo cual generaría que no existieran tensiones internas en la lámina. Sin embargo, existen numerosos factores que hacen salir a la lámina de este ideal.

Continuando con la experimentación se realizaron pruebas para determinar la planeza de la lámina a partir de los parámetros de entrada del proceso, así como de los parámetros del producto a producir. De esta manera se conoceran de los parámetros para mejorar la calidad y obtener un producto con mejores características.

Sin embargo, para poder conocer la variación en la planeza es necesario recopilar información de los productos que se están generando y comparar el resultado que estos obtuvieron contra las variables establecidas en el momento de generarlo. Si bien este es un problema que podría sonar bastante sencillo, comienza a complicarse debido a la cantidad de variables implicadas en el proceso, como se presenta esta situación en su misma forma para la predicción de espesor.

No obstante, para poder generar un resultado de un solo conjunto de datos se tienen que designar una calificación a la planeza, la cual nos indique el grado de satisfacción en el proceso. De esto se genera un conjunto de problemas ya que la planeza es un atributo que no se considera único para un conjunto de datos, debido a que las mínimas deformidades de la lámina, así como las variaciones intrínsecas del proceso, generan que la lámina sufra variaciones a todo el largo de ésta e inevitablemente a lo ancho también. Estas deformaciones se pueden observar en la medición de planeza, y de esto que la medición forme una matriz con un conjunto de valores que indican la posición respecto al ideal.

Para ejemplificar más a detalle, podemos adentrarnos en lo que se conoce como medición de planeza ó medición de forma.

3.1.3 Medición de Planeza

Cuando la lámina llega al molino de laminación, aunque éste se encuentre en perfectas condiciones, es relativamente difícil evitar que surjan problemas como la deformación en los rodillos debido a las fuerzas aplicadas, lo que a su vez se ve reflejado en una lámina con un centro más grueso y orillas delgadas. Más aun, el calor que se genera durante la reducción causa una expansión térmica en los rodillos. Estos fenómenos sumados a otros factores hacen que la apertura entre rodillos no sea completamente paralela a lo largo de estos, subsecuentemente se genera una reducción no uniforme a lo largo del rodillo. Lo cual finalmente se ve reflejado en fuerzas residuales internas en la lámina que se manifiestan como láminas con ondulaciones al centro u a las orillas. Típicamente la lámina la dividimos en centro, orilla y cuartos para referirnos al lugar donde se encuentra la deformación.

Las fuerzas residuales podemos observarlas como diferencias en elongación, las cuales son cualitativas si dividimos la lámina en tiras como en la figura 3.3. El diferencial de elongación dL por unidad de longitud L , lo cual puede ser expresado como el diferencial de deformación $\delta E = dL/L$. Debido a que el diferencial de deformación es sumamente pequeño, se ha determinado un parámetro normalizado llamado unidades- $I(I - unit)$, el cual es el índice más aceptado en la medición de planeza, en la laminación en frío describe la diferencia en elongación transversal a la dirección de rolado 3.1.1.

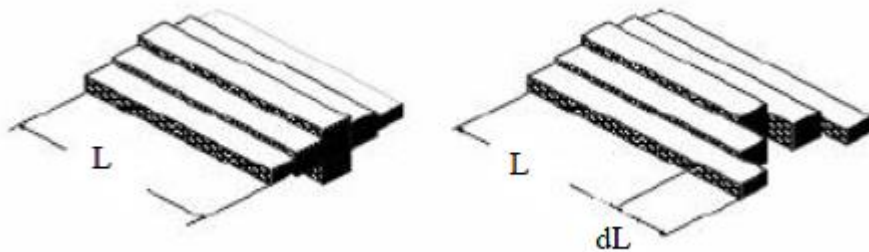


Figura 3.3: Definición de planeza

$$I - unit = \Delta\epsilon * 10^5 = dL/L * 10^5 \quad (3.1.1)$$

Debido a que este índice de planeza en una variables sin dimensiones, una unidad I es igual a 10 *microm/m*.

La figura 3.4 ilustra diferentes ejemplos de los tipos básicos de señales que podemos observar en un medidor de planeza, mostrándonos las unidades I .

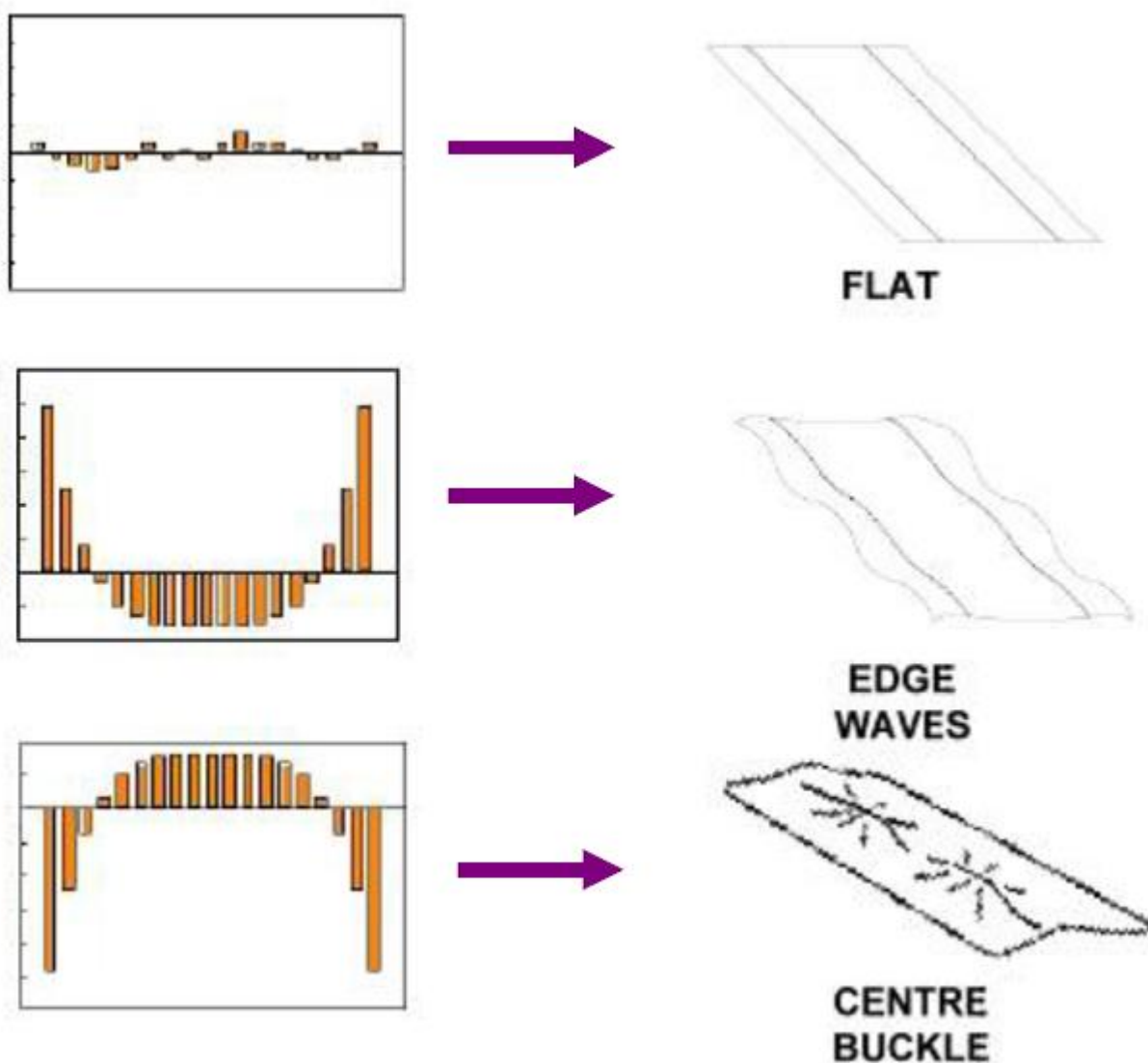


Figura 3.4: *Medidor de planeza y formas generadas*

3.2 Parámetros

Modelar un sistema del mundo real es una tarea sumamente complicada y generalmente caracterizada por ser problemas de una alta dimensión, con complicaciones como la corrupción de las mediciones por ruido y la no linealidad de sus variables. Con certeza se podrá localizar al proceso de laminación en frío en esta categoría.

De lo anterior, se presentan complicaciones para realizar un modelo, actualmente existen modelos matemáticos apoyados en métodos estadísticos que resultan demasiado complicados para un uso práctico. Como alternativa a estos modelos, surge el uso de la inteligencia artificial para manejar la incertinidad, los retrasos de tiempo y la no linealidad de los sistemas.

Un modelo de redes neuronales es presentado, aun no existiendo una representación matemática del proceso, éste puede ser generando por medio de la observación de las entradas y salida del proceso.

3.2.1 Redes neuronales

Las redes neuronales son inspiradas en el Sistema de cognición y reconocimiento que usa el cerebro. De acuerdo a Haykin, una red neuronal es un procesador masivo paralelo propenso a almacenar conocimiento experimental y hacerlo disponible para su uso. También sobresalta es parecido con el cerebro en dos aspectos:

El conocimiento es adquirido a la red por el proceso de aprendizaje. La fuerza de interconexión entre neuronas, conocida como pesos sinápticos, es usada para almacenar el conocimiento.

Las redes neuronales son especialmente útiles para problemas de clasificación, mapeo y aproximación de funciones, ya que son tolerantes a imprecisiones que tengamos en los datos de entrenamiento 3.5.

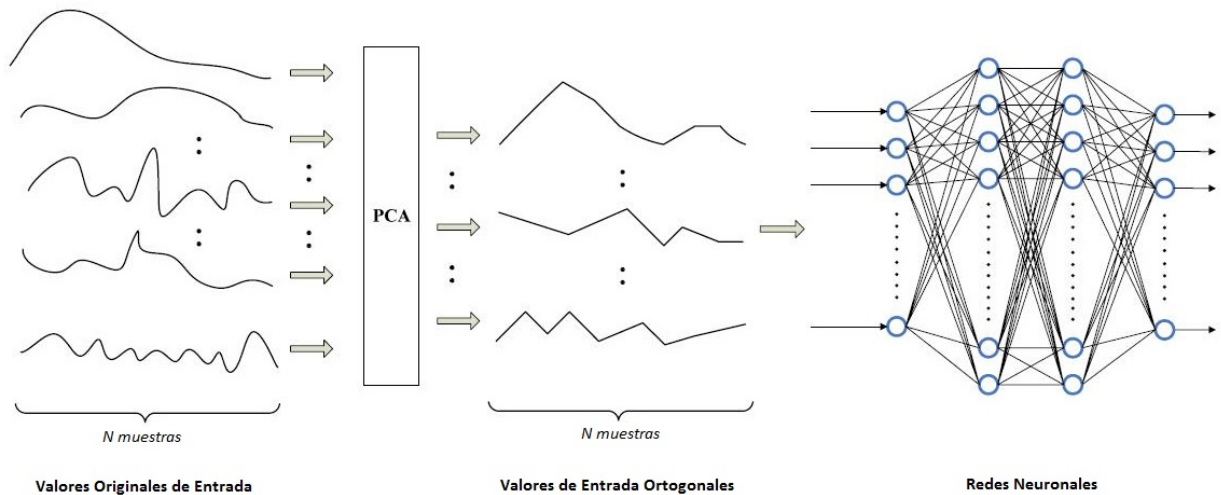


Figura 3.5: *Proceso de simplificación de base de datos y utilidad en redes neuronales.*

Las redes neuronales usan los datos de un proceso para resolver problemas de control, a diferencia de las técnicas tradicionales de control, las cuales se basan en el uso de ecuaciones explícitas. Debido a esto, es importante contar con una cantidad de información acertada para el proceso de entrenamiento. La cantidad requerida dependerá de la complejidad del proceso a ser modelado, el número de variables de entrada involucradas y la calidad de la información disponible. Existen diferentes maneras de obtener datos del proceso, uno de ellos es obtener los datos del proceso bajo las condiciones existentes de operación o diseño experimental y la otra es haciendo un cálculo de optimización de variables para introducir variables específicas de entrada en el proceso. En esta investigación, se trabajó con datos reales de producción, los cuales fueron recolectados del molino de laminación tipo tándem con cuatro castillos de la empresa Ternium México.

3.2.2 Arquitectura y parámetros.

Para la experimentación de esta investigación se seleccionaron dos tipos de redes neuronales, la primera es la red neuronal perceptrón multicapas y como comparativa usaremos la red neuronal

de máxima sensibilidad. Para la selección de parámetros en esta investigación se han hecho un gran número de simulaciones con los datos de proceso, con lo cual hemos tratado de llegar a seleccionar los parámetros óptimos para la modelación. El método utilizado fue mediante un diseño experimental, el cual consistía en iniciar con una proposición de parámetros para la red neuronal y con esto hacer una comprobación de la estimación que producen esos parámetros. Los parámetros los evaluamos con base en los estimadores estadísticos más utilizados y aceptados por la comunidad de investigadores. Seguido de la evaluación, se comenzaba el proceso de nuevo, esta vez con unos parámetros diferentes para proceder con la evaluación y poder tener una medida del desempeño de los nuevos parámetros. Este proceso se realizó de manera recurrente, y sabiendo que parte del proceso es estocástico, se evaluó repetidamente los mismos parámetros y con esto pudimos seleccionar los mejores parámetros.

3.2.3 Perceptrón Multicapas

De acuerdo a lo mencionado anteriormente se tienen que definir ciertos parámetros para la operación correcta de la red neuronal perceptrón. Los parámetros necesarios para el correcto modelado del perceptrón multicapas:

Cantidad de capas de la red neuronal.

Tipo de entrenamiento, estocástico, gradiente, quasi newton, optimización.

Velocidad de Aprendizaje.

Gradiente.

3.2.4 Máxima Sensibilidad

De acuerdo a lo mencionado anteriormente se tienen que definir ciertos parámetros para la operación correcta de la red neuronal de máxima sensibilidad. Los parámetros necesarios para el correcto modelado de la red máxima sensibilidad:

Máximo número de neuronas.

Factor de sensibilidad.

Margen de sensibilidad.

3.2.5 Regresión simbólica

Como variante de los modelos de redes neuronales hemos realizado modelación mediante regresión simbólica, la cual es considerada como la aproximación más cercana a un modelo de control clásico ya que el modelo es completamente representado por operaciones matemáticas en las cuales podemos ver una interacción directa de las variables del modelo. Este tipo de modelo es una variante a los sistemas inteligentes como la red neuronal, ya que las redes son consideradas cajas negras en las cuales nos es imposible observar la interacción de las variables y solo podemos observar las entradas contra la salida que generan.

Actualmente existen diversos métodos y técnicas para aplicar la regresión simbólica, para esta investigación utilizaremos el método de regresión simbólica *alpha - beta*, la cual implica el uso de operadores conocidos como alpha y beta, mismos que nos sirven para representar el modelo mediante funciones matemáticas. Las funciones matemáticas alpha que tendrán que representar al modelo, son propuestas de acuerdo a la observación de distintos modelos y tratando de englobar las operaciones más utilizadas las cuales están ya descritas en el capítulo anterior. Para poder determinar cómo las variables alpha interactúan entre ellas, se proponen los operadores beta los cuales representan a las operaciones algebraicas básicas. Sin embargo, para lograr determinar cuáles funciones y que valores se utilizarán para la representación del modelo, utilizamos un modelo de cómputo evolutivo el cual es conocido como "evonorm" y del cual podemos tener referencia en el capítulo anterior.

Para poder llegar a un resultado en el "evonorm" es necesario introducir el parámetro de número de generaciones, el cual indica la cantidad de generaciones que evolucionará la población.

También es necesario indicar el factor de mutación que existirá en cada generación, este factor proporciona o induce valores aleatorios a cada generación para poder llegar a la generalización del resultado y ayudar a no estancar el resultado en mínimos locales, en el caso de un problema de minimización y viceversa en maximización.

3.3 Entrenamiento

3.3.1 Perceptrón multicapas

El entrenamiento del perceptrón multicapas es conocido por ser un entrenamiento fuera de línea ya que requiere de una enorme cantidad de tiempo computacional para lograr obtener un buen resultado. De los diversos tipos de entrenamientos que existen para esta red neuronal, seleccionamos el método del gradiente, el cual consiste en realizar un acercamiento al resultado en dirección del gradiente obtenido al obtener el error cuadrático medio del valor esperado en el entrenamiento. Este algoritmo al ser repetido durante muchas iteraciones logra acercar el valor lo suficiente para poder determinar los pesos ideales en nuestra red neuronal. Las variables que necesitamos determinar para obtener un buen entrenamiento, son obtenidas mediante un diseño experimental en el cual corremos el algoritmo con diferentes valores en las variables y replicamos esto para determinar cuáles valores son los óptimos.

3.3.2 Máxima sensibilidad

El entrenamiento en la red neuronal de máxima sensibilidad consiste en el empleo de diversas neuronas y sus pesos para representar el modelo que estamos buscando, para el cual idealmente trataremos de utilizar solo las neuronas que sean necesarias para lograrlo. Teniendo como límite el tope de memoria del dispositivo al cual se implementara la red neuronal.

3.3.3 Regresión simbólica

El método de regresión simbólica consiste en representar el modelo mediante una función que contiene unida mediante operadores que relacionan a diferentes argumentos que representan las variables en el proceso. Sin embargo, para alcanzar la mejor representación del proceso se busca el mejor grupo de operadores que pueda acercarse a la función ó modelo al ideal. Para lograr encontrar el mejor grupo de operadores, utilizamos un algoritmo evolutivo, en este caso específicamente el "EVONORM", el cual nos lleva a encontrar el mejor modelo por cada generación que este algoritmo es entrenado. No obstante, las generaciones no son el único factor que influye en encontrar el mejor modelo ya que a este algoritmo hay que adicionarle variables que nos lleven a explorar soluciones que no se encuentren necesariamente cerca del área de la solución encontrada en la generación anterior, para lo cual adicionamos la variable de mutación. Cabe mencionar que para encontrar los valores utilizados en el algoritmo se realizaron múltiples corridas con cada algoritmo que nos llevaron a determinar los mejores valores para este algoritmo.

3.4 Validación

Para cada uno de los métodos que utilizamos es necesario tener un parámetro el cual sea medible y comparable en cada uno para determinar el desempeño de cada algoritmo. Ya que nuestro objetivo es determinar el mejor modelo, utilizamos un método para validar cada modelo.

Para saber la exactitud del modelo, es necesario dividir los valores que fueron adquiridos del trabajo de campo y dividir un porcentaje de estos para poder realizar la validación, el porcentaje restante puede ser utilizado para el entrenamiento. Con esto aseguramos que cada modelo desconoce cada grupo de valores con los cuales evaluaremos el desempeño. La evaluación del desempeño la llevamos a cabo obteniendo un resultado de cada modelo y comparando este con el resultado medido en campo, con estos valores podemos muestrear los resultados mediante los indicadores estadísticos como lo son el "PRESS", "R2PRED" y "MSE". De esta manera,

podemos obtener un resultado comparable entre los distintos modelos obtenidos y partiendo de esto determinar cuál es el mejor modelo para la aplicación.

Capítulo 4

Experimentación y resultados.

Para la generación de resultados consideramos una población total de 500 datos, de los cuales el 80% son usados para entrenamiento y el 20% restante para validación. Considerando 22 entradas y una única salida. Las tablas comparativas consideraran los errores estadísticos y el tiempo de cómputo utilizado para obtener el modelo.

4.1 Espesor

El valor final de espesor de un producto está dado por la interacción de todas las variables que intervienen en el proceso. Esta interacción se da de acuerdo a un modelo de comportamiento, el cual en esta investigación lo definimos mediante distintos paradigmas de los llamados sistemas inteligentes. El espesor es la una sola medición por cada producto producido.

4.1.1 Desempeño de la red neuronal.

Las redes neuronales fueron nuestro principal instrumento en la búsqueda del mejor modelo. De acuerdo al levantamiento y muestreo considerados, encontramos los siguientes resultados.

4.1.2 Red Neuronal Perceptrón con entrenamiento "Backpropagation"

Los resultados obtenidos de nuestra experimentación muestran como indudablemente a lo largo de las épocas con-forme la red neuronal es entrenada logra reducir el error, llegando a un punto en el cual se estabiliza y mantiene un desempeño sólido. Para el entrenamiento utilizamos un η of 0.3, α o momento de 0.15 y 43 neuronas en la capa oculta.

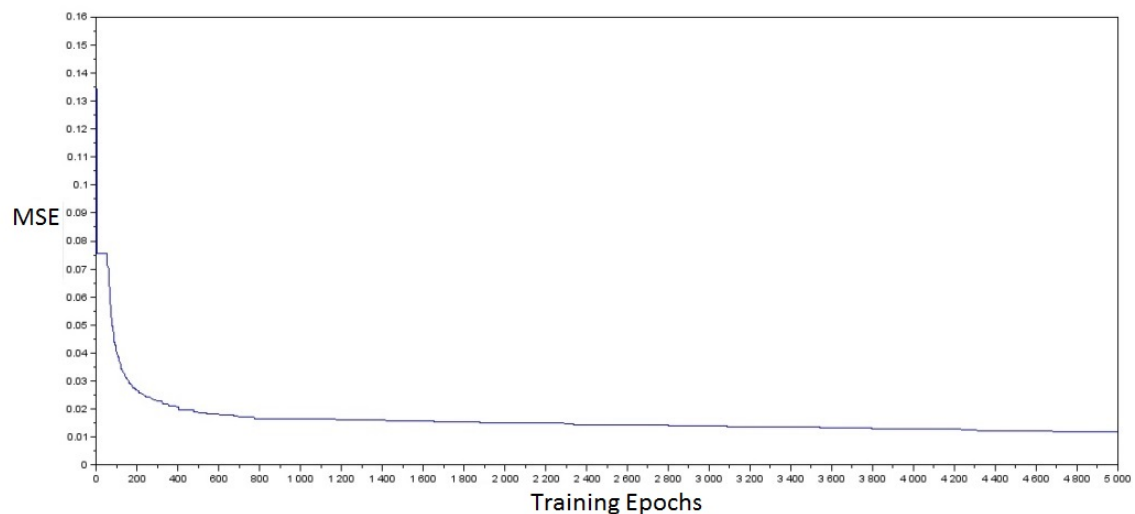


Figura 4.1: *Épocas de entrenamiento contra el error cuadrático medio.*

Los resultados que obtuvimos son directamente medibles en el caso de la red neuronal perceptrón, ya que con cada época de entrenamiento fue evaluada la manera en que la red predecía los valores reales y de acuerdo a la fórmula del error cuadrático medio pudimos determinar el valor de este indicador. Finalizando con un valor de 0.0128007 ver figura 4.1.

La figura 4.2 muestra una comparación en los valores que la red neuronal entrega como resultado para diferentes condiciones de operación contra los valores que fueron muestreados en campo. Con esto podemos observar las diferencias que existen en estos valores.

Finalmente, como se observa en la imagen 4.3 se recolectaron datos para hacer la comparación entre los valores que la red neuronal perceptrón entrega para un conjunto de entradas que la red

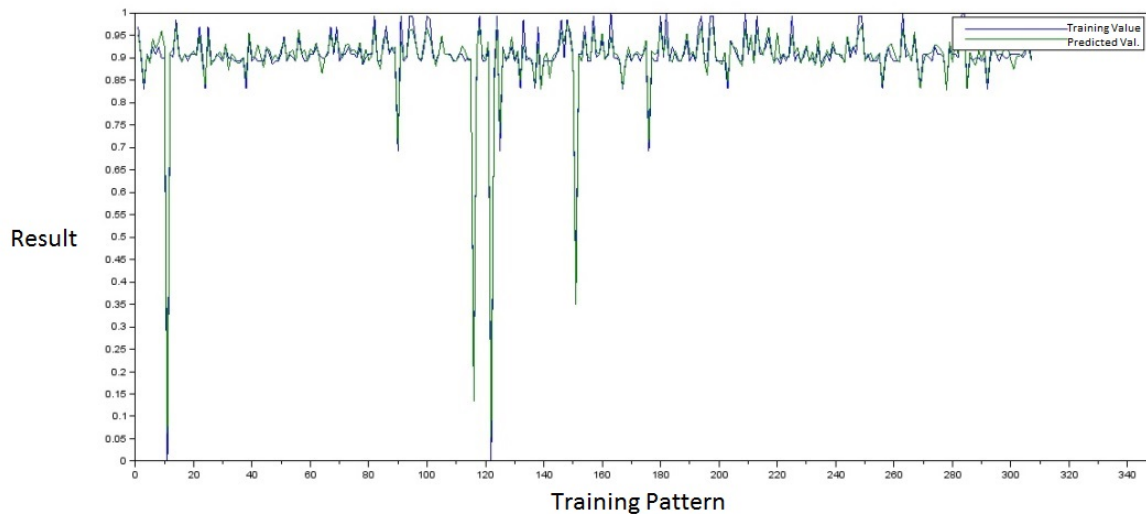


Figura 4.2: *Comparación de valores predichos contra los valores de las muestras.*

desconoce, es decir, valores con los cuales la red neuronal no ha sido entrenada. Esta comparación es sumamente importante ya que nos asegura la generalización de la red neuronal para valores nuevos.

4.1.3 Máxima sensibilidad

Los valores encontrados mediante la red neuronal de máxima sensibilidad son comparados en la figura 4.4 con los valores reales obtenidos en el levantamiento de campo, con lo cual podemos observar la variación que existe entre estos y los resultados de la red neuronal, para este caso se usaron valores nuevos desconocidos por la red neuronal en el modo de entrenamiento, lo cual nos garantiza la generalización del modelo para valores nuevos. Estos resultados fueron logrados usando una arquitectura de 200 neuronas, una lambda de 0.14 y un margen de sensibilidad de 0.85.

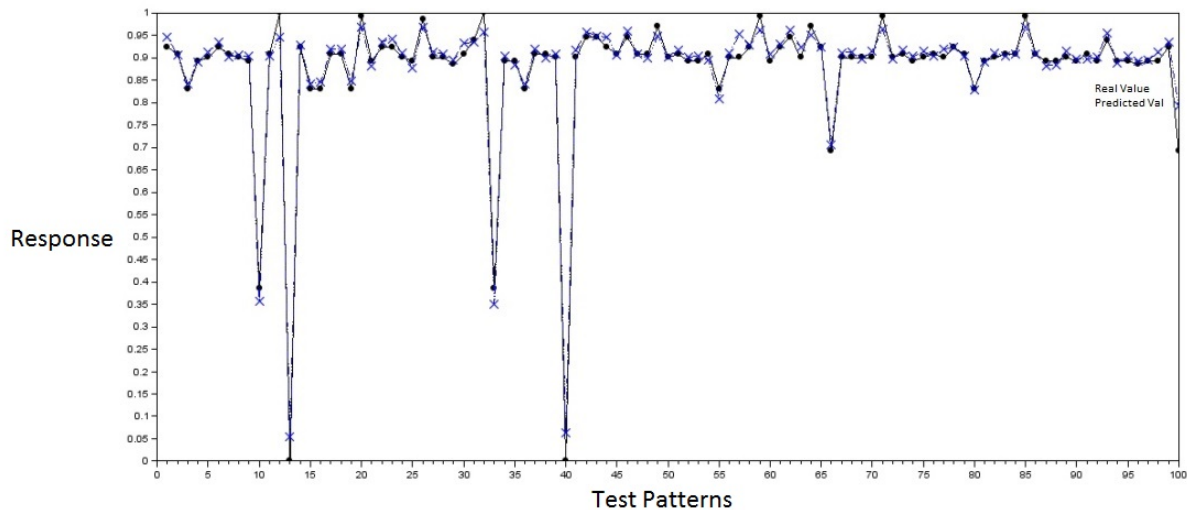


Figura 4.3: Comparación de valores predictos contra los valores nuevos desconocidos por la red neuronal "Backpropagation".

4.1.4 Regresión simbólica

Finalmente en modelo de regresion simbólica fue usado para modelar el proceso. La configuración utilizada considera un operador α por cada variable de entrada. Por lo cual, un total de 84 parametros deben ser calculados, 42 parametros k , 21 α 's, 20 β 's y un entero utilizado para conectividad.

Usando el algoritmo *evonorm* encontramos como mejor resultado el siguiente vector decodificado para conectividad: $B = [0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 1]$

Los resultado de α y β se muestra en las tablas 4.1.1 y 4.1.2.

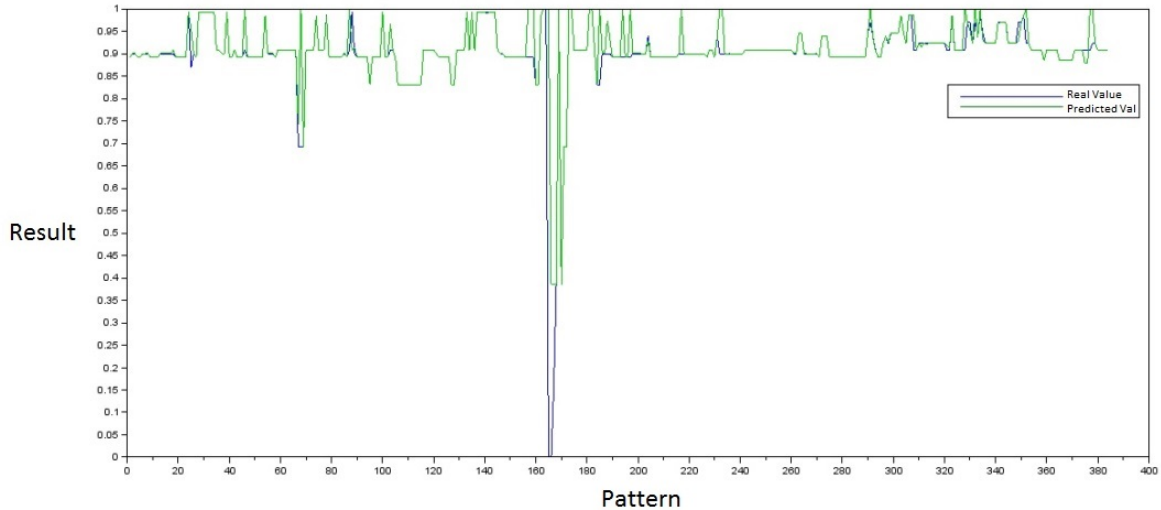


Figura 4.4: Comparación de valores predichos contrá los valores nuevos desconocidos por la red neuronal MS.

Finalmente estos operadores forman la siguiente ecuación 4.1.1.

$$\begin{aligned}
 y(x_3, x_7, x_8, x_9, x_{11}, x_{14}, x_{16}, x_{18}, x_{21}) = & \\
 & ((((((((((((((((((0.5847666 * x_3 + 0.6593538)^3) + \\
 & (0.43712 * x_7 + 0.3346624)^2) + (0.7857256 * x_8 + \\
 & 0.2255187)) * (0.2336883 * x_9 + 0.9180191)))))) + \\
 & (0.9875162 * x_{11} + 0.5309005)))))) + (0.3485697 * x_{14} + \\
 & 0.0419397))) + (0.0234704 * x_{16} + 0.5189331)^2)) + \\
 & (0.6363647 * x_{18} + 0.4944638)^{-1} * (0.8175074)^2)) \\
 & + \cos(0.4328817 * x_{21} + 0.3899872)
 \end{aligned} \tag{4.1.1}$$

4.1.5 Tablas de comparación

La tabla 4.1.3 muestra cuales son las características de cada método, en resumen podemos obtener mejor precisión con la red neuronal entrenada por el metodo "backpropagation" sin embargo debemos disponer y ser pacientes con el tiempo que es necesario para realizar el entrenamiento de esta.

operador	número de operación
α_1	2
α_2	7
α_3	11
α_4	2
α_5	1
α_6	4
α_7	1
α_8	10
α_9	13
α_{10}	2
α_{11}	14
α_{12}	14
α_{13}	5
α_{14}	6
α_{15}	7
α_{16}	3
α_{17}	6
α_{18}	7
α_{19}	7
α_{20}	12
α_{21}	1

Tabla 4.1.1: Operadores encontrados α .

4.2 Planeza

La planeza es considerada como uno de los parámetros definatorios para determinar si el producto producido en el proceso de laminación en frío es satisfactorio o no. Para poder asegurar un valor que satisfaga los estándares de producción, es necesario conocer los valores en los cuales el proceso genera buenos valores, esto lo logramos mediante un modelo que asocie los valores de entrada al proceso y el valor que califica al producto en el ámbito de planeza. Debido a que es complicado obtener una comparativa en tiempo real de los valores contra la matriz que se genera en las mediciones de planeza, se opta por simplificar la medición a un valor que represente el comportamiento general de la planeza en cada rollo producido.

operador	número de operación
β_1	1
β_2	1
β_3	2
β_4	4
β_5	4
β_6	1
β_7	1
β_8	4
β_9	2
β_{10}	2
β_{11}	1
β_{12}	1
β_{13}	1
β_{14}	2
β_{15}	3
β_{16}	2
β_{17}	3
β_{18}	3
β_{19}	3
β_{20}	1

Tabla 4.1.2: Operadores encontrados β

Tabla 4.1.3: Análisis estadístico para cada método considerando la mejor solución encontrada y el tiempo de procesamiento consumido por cada método.

Método	"MSE"	" R_{pred}^2 "	"PRESS"	"CPU time" (seg.)
RN Perpetron	0.0128007	0.9861901	0.0038099	11:51:40
RNMS	0.0151835	0.9199943	0.0885267	0:0:58
RSAB	0.0095267	0.9880336	0.7430802	0:48:20

4.2.1 Desempeño de la red neuronal.

$$R2prec2 = 0.9999996$$

$$PRESS2 = 22.097261$$

$$RMSE2 = 0.1389812$$

$$Errcm = 0.2739009$$

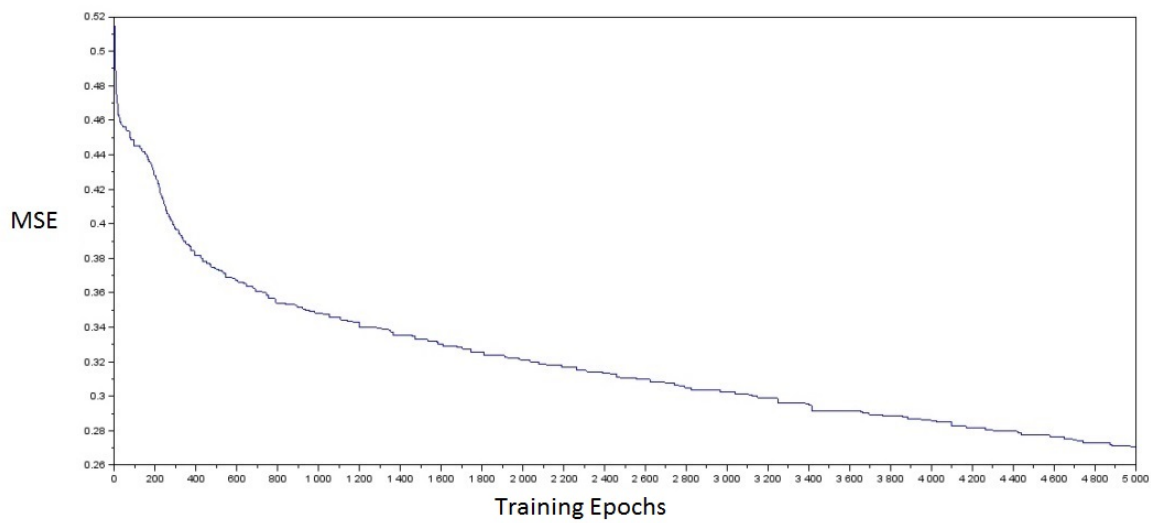


Figura 4.5: *Épocas de entrenamiento contra el error cuadrático medio.*

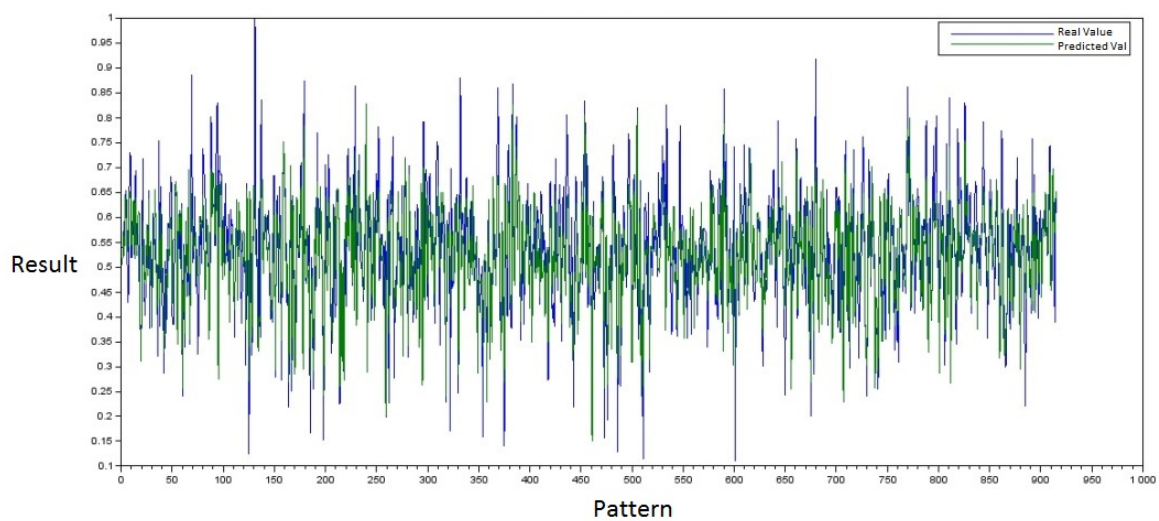


Figura 4.6: *Épocas de entrenamiento contra el error cuadrático medio.*

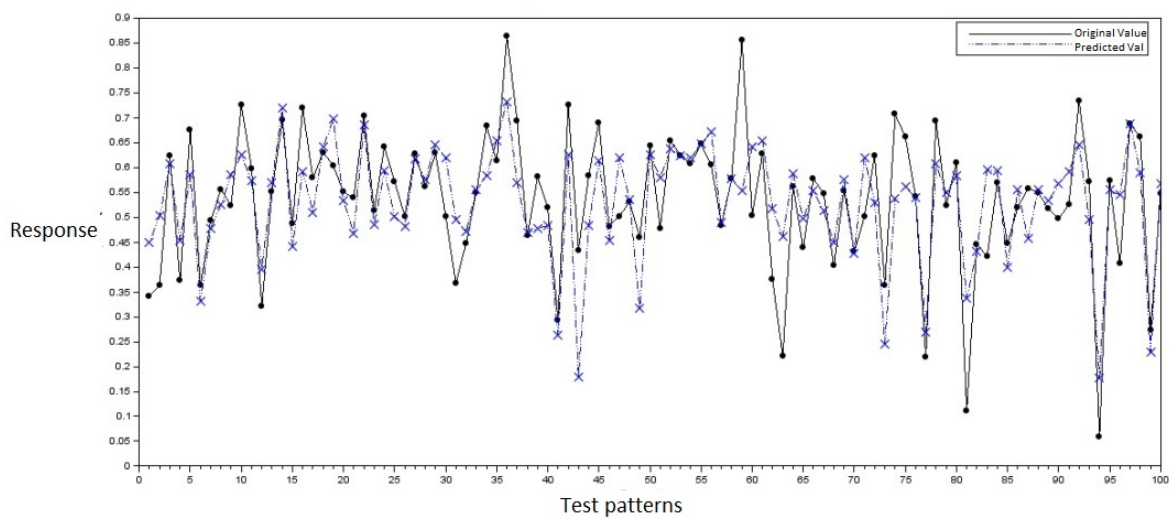


Figura 4.7: Comparación de valores predichos contra los valores de las muestras.

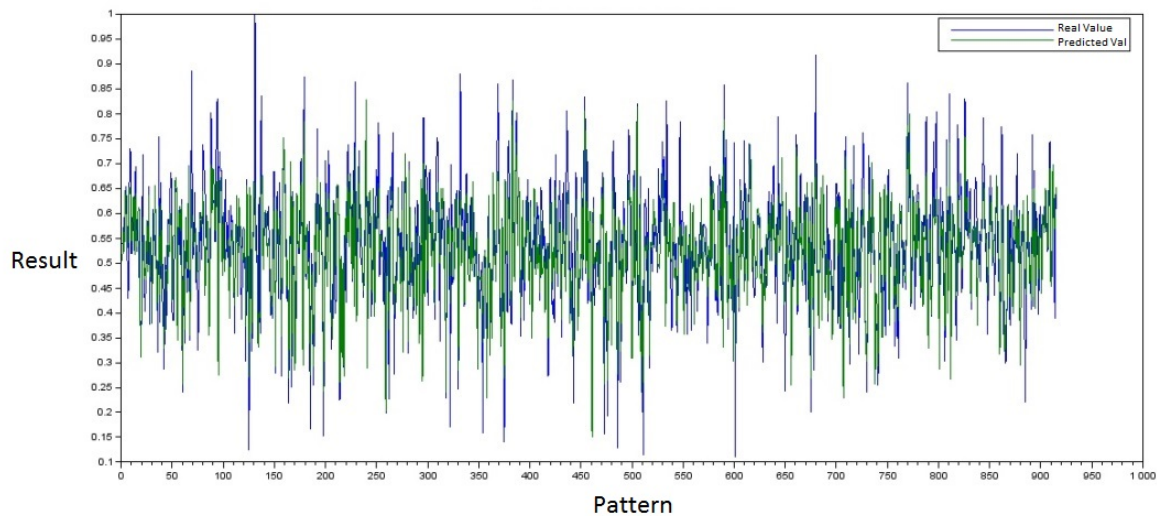


Figura 4.8: Comparación de valores predichos contra los valores nuevos desconocidos por la red neuronal "Backpropagation".

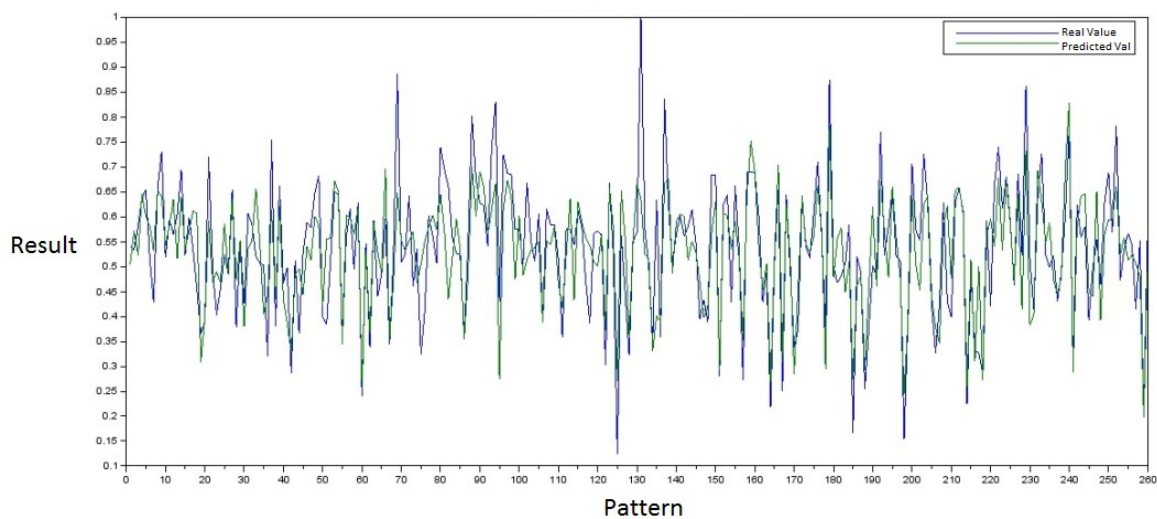


Figura 4.9: Comparación de valores predictos contra los valores nuevos desconocidos por la red neuronal Máxima Sensibilidad.

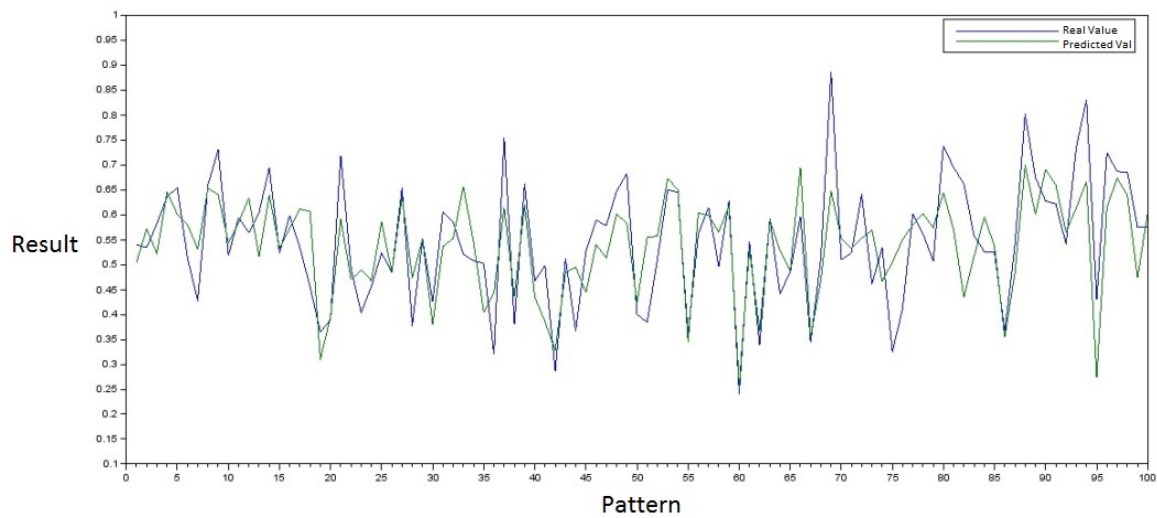


Figura 4.10: Comparación de valores predictos contra los valores nuevos desconocidos por la red neuronal Máxima Sensibilidad.

Capítulo 5

Conclusiones.

El desarrollo de algoritmos inteligentes y las aplicaciones en la industria siderúrgica convergen en soluciones a problemas de alto grado de dificultad. En esta ocasión el resultado es un modelo que puede ayudarnos a solucionar parte de la incertidumbre y evitar trabajar de manera empírica.

Las estructuras desarrolladas en los métodos de redes neuronales, a pesar de claramente simplificar el método de funcionamiento de las neuronas funcionando de manera conjunta, logran emular de una manera precisa el funcionamiento del proceso. Ya que encontrando las opciones adecuadas de entradas a la red neuronal podemos estimar el resultado del atributo indicado, el cual puede ser cualquiera siempre y cuando sea una variable medible, en esta investigación se estudio el espesor y la planeza. De esto podemos concluir que para una variable que puede ser considerada de manera global para un rollo completo de lamina es mucho mas sencillo obtener un resultado del modelo y con esto poder determinar la precision con que este se desempeña. No obstante en esta investigación se intentó buscar un parámetro global que nos indicara la calidad de planeza para cada rollo de lámina, pero al ser una variable compuesta y simplificada no fue posible llegar a obtener el mismo resultado que para el parámetro de espesor.

El proceso de laminación en frio, conlleva la interacción de un gran número de variables y parámetros, los cuales tienen una afectación en la calidad del producto, debido a esto los modelos clásicos de control pueden llegar a ser muy complejos y difíciles de obtener, por lo cual

los modelos de sistemas inteligentes son cada vez mas utilizados en la industria, ya sea realizando modelos con sólo sistemas inteligentes o utilizando sistemas hibridos que mezclan los modelos clásicos con los sistemas inteligentes. Sin embargo, la precisión que estos modelos pueden llegar a conseguir depende enteramente de la precision de la lectura de las variables en el proceso, la correcta asociación de las variables, los parámetros y los resultados en el proceso conllevan a entender el comportamiento de este, esto aun sin ser necesario un modelo matemático que nos explique detalladamente como interactuan las variables, mas utilizando los recursos y técnicas que ofrece esta área de investigación.

La diferenciación que se realiza con lo diferentes métodos de sistemas inteligentes nos ayudan a tener una mejor perspectiva de las cualidades de cada sistema. Además, esta investigación nos ayuda a discriminar las variables del proceso que se consideran mediáticas y redundantes en el proceso de laminación, lo cual servirá de comparativo para futuras investigaciones.

Finalmente concluimos que de acuerdo a los resultados cada metodo mostro atributos unicos, los cuales deberan ser considerados para una selección de un metodo en particular, ya que podemos encontrar precisión en modelos como la red neuronal perceptrón y regresión simbólica o rapidez de entrenamiento en la red neuronal de máxima sensibilidad.

5.1 Trabajo Futuro

Los resultados generados podrán ser comparados contra diferentes vertientes de los sistemas inteligentes y de esto partir para poder tener un panorama más amplio de sus cualidades y deficiencias.

Partiendo de los resultados encontrados se continuará con la investigación del proceso de laminación en frío, conjuntando las diferentes vertientes de investigación, llevádo a conocer mejor el proceso y realizar un modelo que emule el proceso de laminación en frío de una manera precisa.

De manera inmediata, mejorar la medición de planeza para poder indicar resultados que informen las características más importantes para la industria, donde en conjunto con la industria acerera se buscará el mejor indicador que beneficie a su infraestructura.

Bibliografía

- [1] Daniel Pérez, Francisco J. García-Fernández, Ignacio Díaz, Abel A. Cuadrado, Daniel G. Ordonez, Alberto B. Díez, and Manuel Domínguez. Visual analysis of a cold rolling process using a dimensionality reduction approach. *Engineering Applications of Artificial Intelligence*, 26(8):1865 – 1871, 2013.
- [2] Youngdeuk Kim, Chang-Wan Kim, Sungjin Lee, and Hyunchul Park. Dynamic modeling and numerical analysis of a cold rolling mill. *International Journal of Precision Engineering and Manufacturing*, 14(3):407–413, 2013.
- [3] Youngdeuk Kim, Hyunchul Park, Seoung Soo Lee, and Chang-Wan Kim. Development of a mathematical model for the prediction of vibration in a cold rolling mill including the driving system. *ISIJ International*, 52(6):1135–1144, 2012.
- [4] Jianliang Sun, Yan Peng, and Hongmin Liu. Coupled dynamic modeling of rolls model and metal model for four high mill based on strip crown control. *Chinese Journal of Mechanical Engineering*, 26(1):144–150, 2013.
- [5] PÃ Guedes Alves, JosÃ© Adilson de Castro, Luciano Pessanha Moreira, and Elder Moreira Hemerly. Modeling, simulation and identification for control of tandem cold metal rolling. *Materials Research*, 15:928 – 936, 12 2012.
- [6] Xiu ling Zhang, Shao yu ZHANG, Guang zhong Tan, and Wen bao Zhao. A novel method for flatness pattern recognition via least squares support vector regression. *Journal of Iron and Steel Research, International*, 19(3):25 – 30, 2012.
- [7] Kohei Ozaki, Toshiyuki Ohtsuka, Kenji Fujimoto, Akira Kitamura, and Makishi Nakayama. Nonlinear receding horizon control of thickness and tension in a tandem cold mill with a variable rolling speed. *ISIJ International*, 52(1):87–95, 2012.
- [8] Sang-Ho Lee, Gil-Ho Song, Sung-Jin Lee, and Byung-Min Kim. Study on the improved accuracy of strip profile using numerical formula model in continuous cold rolling with 6-high mill. *Journal of Mechanical Science and Technology*, 25(8):2101–2109, 2011.
- [9] José Luis Calvo-Rolle, José Luis Casteleiro-Roca, Héctor Quintín, and María del Carmen Meizoso-Lopez. A hybrid intelligent system for {PID} controller using in a steel rolling process. *Expert Systems with Applications*, 40(13):5188 – 5196, 2013.

- [10] Mehrdad Poursina, Noushin Torabian Dehkordi, Amin Fattahi, and Hadi Mirmohammadi. Application of genetic algorithms to optimization of rolling schedules based on damage mechanics. *Simulation Modelling Practice and Theory*, 22(0):61 – 73, 2012.
- [11] Jian-guo Cao, Xiao-zhao Xu, Jie Zhang, Mu-qing Song, Gui-liang Gong, and Wei Zeng. Pre-set model of bending force for 6-high reversing cold rolling mill based on genetic algorithm. *Journal of Central South University of Technology*, 18(5):1487–1492, 2011.
- [12] Kouros Mousavi Takami, Jafar Mahmoudi, Erik Dahlquist, and Magnus Lindenmo. Multi-variable data analysis of a cold rolling control system to minimise defects. *The International Journal of Advanced Manufacturing Technology*, 54(5-8):553–565, 2011.
- [13] Hong min LIU, Xiu ying SHAN, and Chun yu JIA. Theory-intelligent dynamic matrix model of flatness control for cold rolled strips. *Journal of Iron and Steel Research, International*, 20(8):1 – 7, 2013.
- [14] Suchandan K. Das. Neural network modelling of flow stress and mechanical properties for hot strip rolling of trip steel using efficient learning algorithm. *Ironmaking and Steelmaking*, 40:298–304, 2013.
- [15] Mahdi Bagheripoor and Hosein Bisadi. Application of artificial neural networks for the prediction of roll force and roll torque in hot strip rolling process. *Applied Mathematical Modelling*, 37(7):4593 – 4607, 2013.
- [16] N.S. Reddy, B.B. Panigrahi, and J. Krishnaiah. Modeling mechanical properties of low carbon hot rolled steels. In Jagdish Chand Bansal, Pramod Singh, Kusum Deep, Millie Pant, and Atulya Nagar, editors, *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*, volume 202 of *Advances in Intelligent Systems and Computing*, pages 221–228. Springer India, 2013.
- [17] J. Ghaisari, H. Jannesari, and M. Vatani. Artificial neural network predictors for mechanical properties of cold rolling products. *Advances in Engineering Software*, 45(1):91 – 99, 2012.
- [18] P. Saravanakumar, V. Jothimani, L. Sureshbabu, S. Ayyappan, D. Noorullah, and P.G. Venkatakrishnan. Prediction of mechanical properties of low carbon steel in hot rolling process using artificial neural network model. *Procedia Engineering*, 38(0):3418 – 3425, 2012. {INTERNATIONAL} {CONFERENCE} {ON} {MODELLING} {OPTIMIZATION} {AND} {COMPUTING}.
- [19] Hossam Faris, Alaa Sheta, and Ertan AÜznergiz. Modelling hot rolling manufacturing process using soft computing techniques. *International Journal of Computer Integrated Manufacturing*, 26(8):762–771, 2013.
- [20] M (Arribas Maribel) ; Carrillo A (Carrillo Ana); Arana JL (Luis Arana José) (Marquez, Itziar); Arribas. Optimisation of total roll power using genetic algorithms in a compact strip production plant. *International journal of materials research*, 104:686–696, 2013.
- [21] N. Chakraborti, B. Siva Kumar, V. Satish Babu, S. Moitra, and A. Mukhopadhyay. A new multi-objective genetic algorithm applied to hot-rolling process. *Applied Mathematical Modelling*, 32(9):1781 – 1789, 2008.

- [22] Simon Haykin. *Neural Networks and Learning machines*. Prentice Hall, 2006.
- [23] Gustavo González Sanmiguel, Luis Lauro González, and César Guerra Luis M. Torres-Treviño. On-line learning in an embedded maximum sensibility neural network. In IEEE, editor, *Eleventh Mexican International Conference on Artificial Intelligence 2012. Advances in Artificial Intelligence and Applications. Special Session-Revised papers.*, pages 75–79, November, 2012.
- [24] Luis M. Torres-Trevi no. Evonorm, a new evolutionary algorithm to continuous optimization. In *Workshop on Optimization by Building and Using Probabilistic Models (OBUPM) Genetic and Evolutionary Computation Conference (GECCO)*, 2006.
- [25] Luis Torres. Evonorm: Easy and effective implementation of estimation of distribution algorithms. *Journal of Research in Computing Science*, 23:75–83, 2006.
- [26] Douglas C. Montgomery. *Introduction to Linear Regression Analysis*. John Wiley and Sons, 2007.