

Año I, No. 01 Enero-Junio 2013

ISSN: 2395-9029

PROYECTOS INSTITUCIONALES Y DE VINCULACIÓN



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



FIME

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

CLASIFICACIÓN DE TUMORES CANCERÍGENOS POR MEDIO DE OPTIMIZACIÓN

Juan Salvador Hinojosa, Álvaro, Arriaga García, Mario Josué Saavedra Olague,
Guadalupe Trujillo Sánchez

RESUMEN.

La detección oportuna del cáncer es la mejor forma de combatir dicha enfermedad, desgraciadamente para realizar dicha tarea es necesario que sea ejecutada por un especialista y no todo el mundo tiene acceso a este o no les es posible costearlo. Para reducir el costo y el tiempo de que se invierte en hacer estas pruebas, así como para ponerlo al alcance de un mayor número de personas hemos desarrollado un sencillo programa funcionara como una aplicación capaz de predecir hasta cierto grado el que un tumor sea maligno o benigno. El programa utilizara un método de optimización para, por medio de una gráfica y basándose en los datos de pacientes previamente diagnosticados, determinar qué tanta posibilidad hay de que un tumor se convierta en maligno.

PALABRAS CLAVE: Cáncer, tumor, matlab, grafica, predicción.

ABSTRACT.

The program used a method of logistic regression, by means of a graph, and based on data from previously diagnosed patients; determine how much chance is there that a tumor becomes malignant.

KEY WORDS. Cáncer, tumor, matlab, grafica, predicción.

INTRODUCCIÓN.

El cáncer es una enfermedad de las células que componen las distintas partes de nuestro organismo. En un individuo sano, las células se dividen y después, como todo ser vivo, mueren; este proceso de muerte, o suicidio celular, también se denomina apoptosis. Durante los primeros años de vida, las células se multiplican con rapidez, hasta alcanzar la edad adulta, en ese momento la división celular se produce no para crecer, sino para reparar los tejidos que se van deteriorando o paliar determinadas lesiones. A diferencia de las células sanas, las cancerosas continúan multiplicándose, por un cúmulo de causas aún desconocidas, en las que pueden influir factores genéticos y exógenos (estilo de vida, medioambiente, infecciones víricas, etc.) Esta continua división celular produce acumulaciones de células o bultos anómalos, que son los tumores.

La detección temprana del cáncer aumenta enormemente las posibilidades de un tratamiento exitoso.

Hay dos componentes principales de la detección temprana del cáncer: la educación para promover el diagnóstico precoz y el cribado. Reconociendo los posibles signos de advertencia del cáncer y tomar medidas rápidas conduce a un diagnóstico precoz. Mayor conciencia de posibles señales de advertencia de cáncer, entre los médicos, enfermeras y otros profesionales de la salud, así como entre el público en general, puede tener un gran impacto en la enfermedad. Algunos de los signos tempranos de cáncer incluyen bultos, llagas que no cicatrizan, sangrado anormal, indigestión persistente y ronquera crónica. El diagnóstico temprano es particularmente relevante para los cánceres de mama, cuello uterino, próstata, boca, laringe, colon y recto, y la piel.

El análisis de optimización es una técnica estadística utilizada para estudiar la relación entre variables. Se adapta a una amplia variedad de situaciones. En la investigación social, el análisis de regresión se utiliza para predecir un amplio rango de fenómenos, desde medidas económicas hasta diferentes aspectos del comportamiento humano. Es así pues que, basándonos en este tipo de análisis de precisión, nos hemos dado a la tarea de desarrollar un programa especializado en el diagnóstico médico que sea capaz de predecir qué tan probable es que un tumor pueda convertirse en maligno.

OBJETIVO

Crear una aplicación que tenga una interfaz práctica y entendible que pueda predecir con un alto grado de precisión la posibilidad de que un tumor pase de ser benigno a convertirse en maligno.

JUSTIFICACIÓN

No hay mejor forma de combatir el cáncer que detectándolo a tiempo, desgraciadamente para hacer esto es necesario invertir una cantidad considerable de tiempo y dinero, los cuales no están al alcance de toda la población, por lo que un programa que sea capaz de predecir el comportamiento de un tumor cancerígeno que sea compatible con cualquier computadora y tenga modelo entendible y sencillo sería una gran ventaja no solo para los usuarios del mismo si no para los centros médicos que deseen y necesiten apresurar los procedimientos de chequeo de tumores.

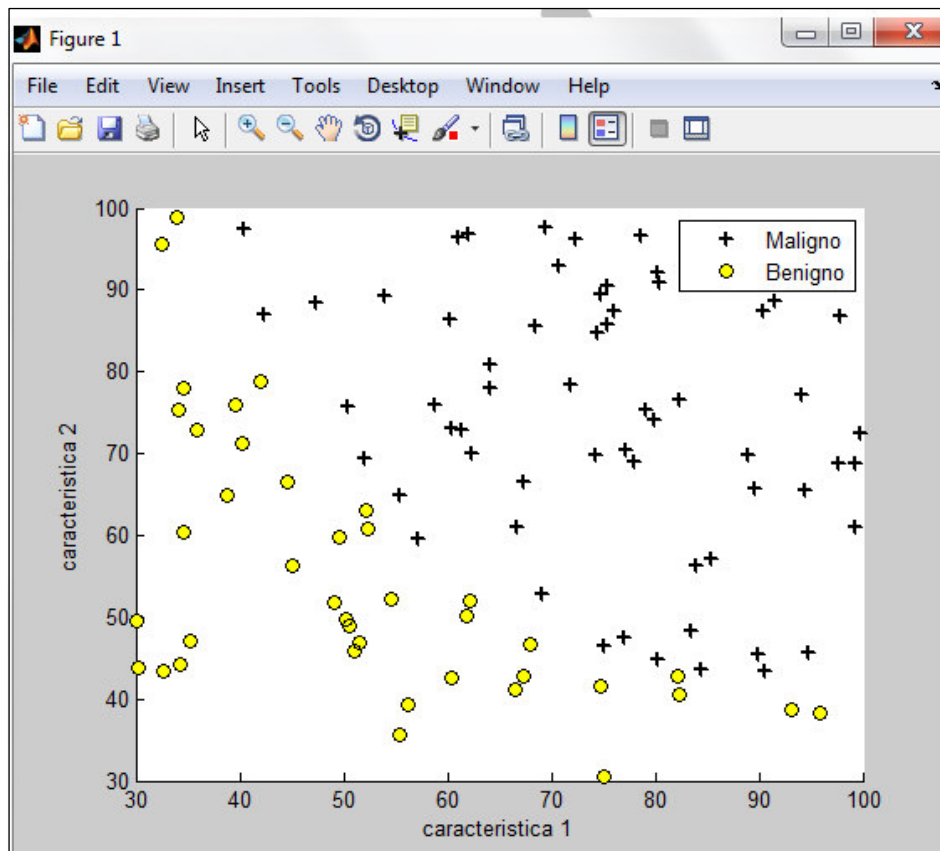
METODOLOGÍA

Obtener un conocimiento profundo acerca de los tumores es una tarea que tomaría años aprender, que una persona sea capaz de decidir si un tumor cancerígeno es maligno o benigno teniendo como único dato las características de éste y solamente utilizando su conocimiento es algo que no se ve muy a menudo, se necesitan gran cantidad de estudios y experiencia previa por parte del médico para determinar si el tumor cancerígeno es maligno o benigno y aun así los resultados no están garantizados.

Lo que se busca, es que este algoritmo de aprendizaje sea capaz de predecir (con la mayor exactitud posible) si el tumor es maligno o benigno. Se utiliza un método de optimización en donde en base de resultados previos de pacientes anteriores, el algoritmo aprende en cuestión de fracciones de segundo lo que a un médico le costaría aprender en años.

El programa empieza por cargar los datos de pacientes anteriores en donde crea un arreglo de m filas y n columnas donde m representa a cada paciente. Por ejemplo, si tenemos $m = 100$, entonces tenemos los datos de 100 pacientes anteriores, y n representa el número de características que se están tomando en cuenta, por ejemplo el tamaño, si es regular o irregular, etc.

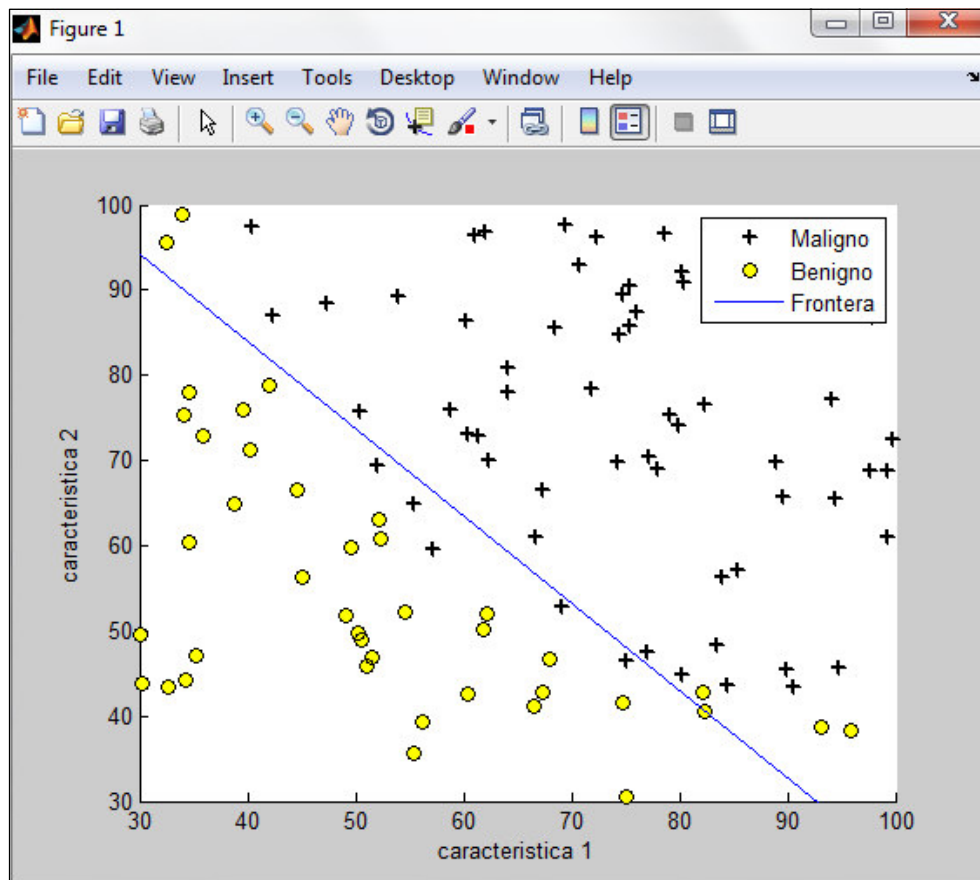
Inmediatamente después se procede a graficar los datos (en este caso solamente tomamos en cuenta 2 características, por lo que la gráfica está representada por dos dimensiones, pero puede haber n cantidad de características) de manera que los signos "+" representan tumores malignos y los signos "o" representan los tumores benignos.



Gráfica de los datos recopilados; véase que los "+" indican los malignos y los "o" indican los benignos.

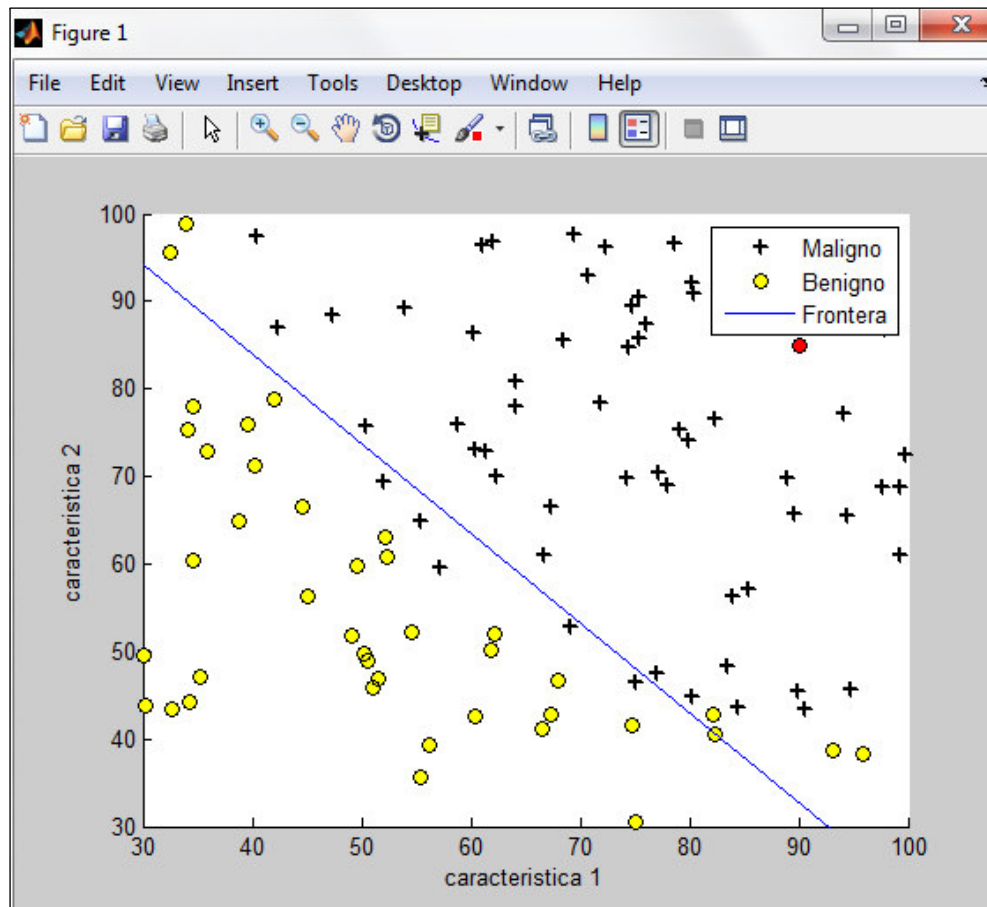
En base a estos datos haremos nuestras predicciones. Ya que los datos están cargados y graficados lo siguiente en hacer es optimizar las funciones que precisamente harán la predicción, para que puedan tomar la mejor decisión posible es necesario optimizarla, en base a los datos cargados el algoritmo es capaz de tomar la mejor decisión posible.

Se ingresan los datos del paciente al que queremos predecir si su tumor es maligno o benigno, la función optimizada colocará la posición del paciente y determinará si está en el área de los malignos o de los benignos y que probabilidades tiene de que sea maligno o benigno.



Gráfica después de optimizar la función véase la decisión de frontera, la cual divide la zona de los malignos y los benignos

Por ejemplo, si el algoritmo nos dice que la posición del tumor del paciente se encuentra justo en la división (en la frontera) entre los malignos y los benignos éste tendrá exactamente un 50% de probabilidad de que este sea maligno o benigno. Aunque claro, es muy poco probable de que esto pase, después de todo existen una infinidad de posiciones en las que el tumor, en base a sus características, tienda a estar más de un lado o del otro, pero no justo en la frontera.



Gráfica para un paciente con características de 90 y 85 para la característica 1 y 2 respectivamente véase que el círculo rojo indica la posición del paciente

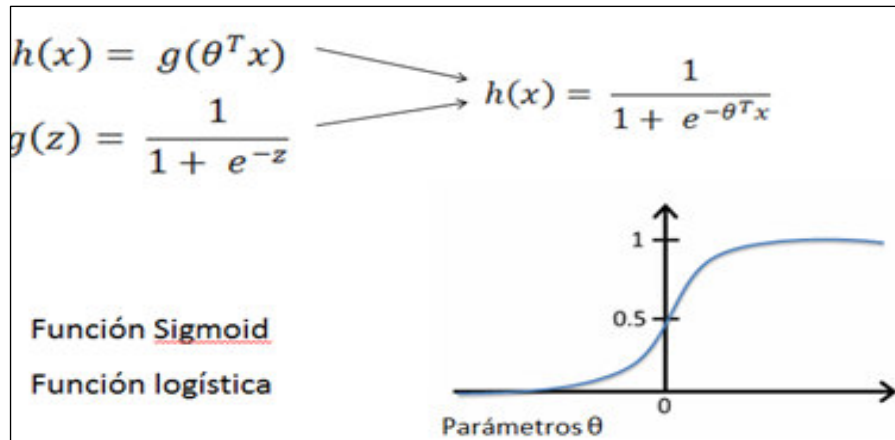
El programa utiliza funciones matemáticas optimizadas para realizar dicha tarea, que de otra manera sería imposible, el algoritmo prácticamente se vuelve un experto en cuestión de segundos en determinar si los tumores cancerígenos son malignos o benignos, lo único que hay que hacer es proporcionarle los datos recopilados de pacientes anteriores y el programa aprenderá por sí mismo en base a estos datos. Con los datos correctos, no existe mejor decisión que la que puede tomar este algoritmo.

RESULTADOS

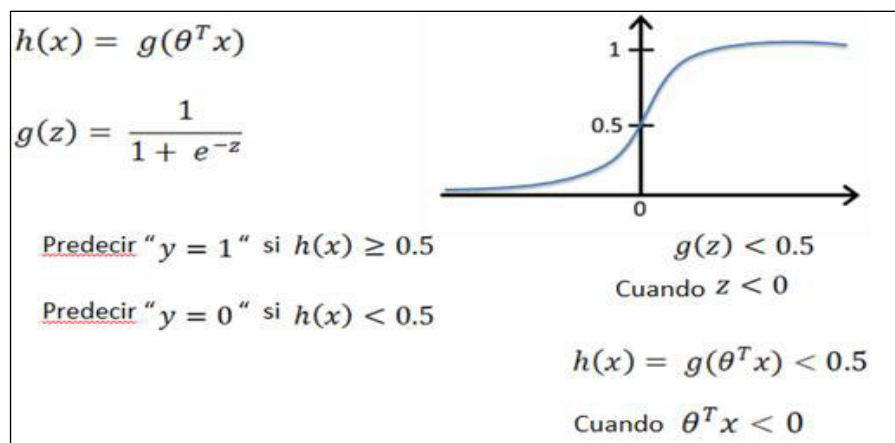
Ya que no nos pudieron proporcionar datos reales, Se tomaron un conjunto de datos de manera aleatoria del 1 al 100.

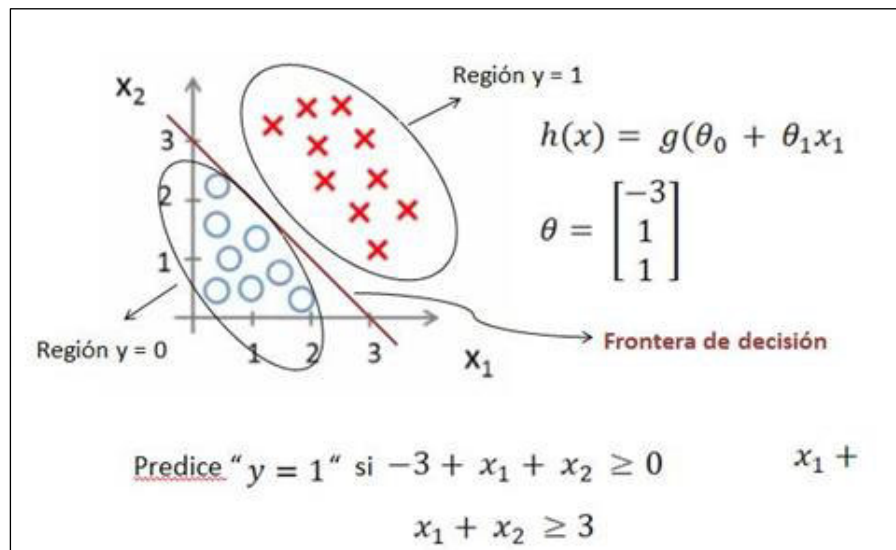
34.6237	78.0247	0
30.2867	43.8950	0
35.8474	72.9022	0
60.1826	86.3086	1
79.0327	75.3444	1
45.0833	56.3164	0
61.1067	96.5114	1
75.0247	46.5540	1
76.0988	87.4206	1
84.4328	43.5334	1
95.8616	38.2253	0
75.0137	30.6033	0
82.3071	76.4820	1
69.3646	97.7187	1
39.5383	76.0368	0
53.9711	89.2074	1
69.0701	52.7405	1
67.9469	46.6786	0
70.6615	92.9271	1
76.9788	47.5760	1
67.3720	42.8384	0
89.6768	65.7994	1
50.5348	48.8558	0
34.2121	44.2095	0
77.9241	68.9724	1
62.2710	69.9545	1
80.1902	44.8216	1
93.1144	38.8007	0
61.8302	50.2561	0
38.7858	64.9957	0
61.3793	72.8079	1
85.4045	57.0520	1
52.1080	63.1276	0
52.0454	69.4329	1
40.2369	71.1677	0
54.6351	52.2139	0
33.9155	98.8694	0
64.1770	80.9081	1
74.7893	41.5734	0
34.1836	75.2377	0
83.9024	56.3080	1

Al aplicar la regresión logística tenemos que $h(x)$ puede ser >1 ó > 0 , entonces queremos que $0 \leq h(x) \leq 1$



Lo cual nos serviría para predecir el comportamiento de un tumor en el futuro.





Dependiendo en que zona del plano quede ubicada dicho tumor, se diagnostica su tipología

CONCLUSIONES

Utilizando modelos matemáticos podemos controlar, predecir y calcular cualquier evento, fenómeno físico o desarrollo biológico siempre y cuando se tomen los parámetros correctos y se mantenga simple dicho modelo matemático. Ciertamente nos hizo falta el apoyo de un experto en cancerología que pudiera proporcionarnos datos relevantes y correctos sobre las tipologías de los tumores y de esta manera darle una validez real a nuestro programa.

BIBLIOGRAFÍA

- Dmedicina.com*. (31 de Marzo de 2009). Obtenido de www.Dmedicina.com: <http://www.dmedicina.com/enfermedades/cancer/actualidad/cancer-tumores>.
- Ng, A. (2013). *Andrew Ng*. Obtenido de [www.cs.stanford.edu](http://cs.stanford.edu): <http://cs.stanford.edu/people/ang/>
- Ng, A. (2013). *coursea.org*. Obtenido de www.coursea.org: <https://www.coursea.org/learn/machine-learning/home/info>