Artículo **Revista de Didáctica Práctica**
Junio 2020 Vol.4 No.11 1-7

Selección de un método de aprendizaje automático para clasificar patrones biomarcadores de lesiones precancerosas de las cuerdas vocales

Selecting a machine learning method to classify biomarker patterns of precancerous vocal cords injuries

SIORDIA-VASQUEZ, Xóchitl†*´, VILLAGRAN-VILLEGAS, Luz Yazmin´, PATIÑO-ORTIZ, Miguel´´ y ROJAS-HERNÁNDEZ, Miguel Ángel´

´Universidad Veracruzana, Unidad de Ingeniería y Ciencias Químicas, Prolongación Venustiano Carranza s/n. Col. Revolución, Poza Rica, de Hidalgo, Veracruz, México.
´´Instituto Politécnico Nacional. Unidad Profesional "Adolfo López Mateos"-Zacatenco, Gustavo A. Madero, CDMX, México.

ID 1er Autor: *Xóchitl, Siordia-Vásquez* / **ORC ID:** 0000-0002-8472-8001, **CVU CONACYT ID:** 1036998

ID 1er Coautor: *Luz Yazmin, Villagrán-Villegas* / **ORC ID:** 0000-0003-3860-2923, **CVU CONACYT ID:** 96365

ID 2do Coautor: *Miguel, Patiño-Ortiz* / **ORC ID:** 0000-0002-5630-8077, **CVU CONACYT ID:** 167388

ID 3er Coautor: *Miguel Ángel, Rojas-Hernández* / **ORC ID:** 0000-0001-9294-5842

DOI: 10.35429/JPD.2020.11.4.1.7 Recibido 10 de Enero, 2020; Aceptado 30 de Junio, 2020

Resumen

La Encuesta Nacional de Consumo de Drogas, Alcohol y Tabaco 2016-2017 señala que 15.6 millones de mexicanos son fumadores activos y, en el año 2030, se estima la muerte de 8 millones por cáncer de laringe o de pulmón. Por lo anterior, la Organización Mundial de la Salud (OMS), recomienda detectar lesiones precancerosas de la laringe, lo cual es posible, ya que se caracterizan en un patrón biomarcador manifestado por la alteración del desempeño biomecánico de las cuerdas vocales, sin importar el género y la edad del fumador. El objetivo de este artículo es evaluar tres métodos de aprendizaje automático: redes neuronales, redes bayesianas y árboles de decisión, para seleccionar el que mejor resuelva el problema de detección de biomarcadores de lesiones precancerosas de las cuerdas vocales. Se utiliza la herramienta WEKA y un banco de conocimiento con 250 patrones, proporcionado por el Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra, y avalado por la NOM-012-SSA3-2012. El desempeño de los métodos fue comparado mediante las áreas bajo la curva ROC y matrices de confusión, considerando los criterios establecidos por la norma ISO-5725. Los resultados demostraron que el árbol de decisión resuelve mejor el problema de detección de patrones biomarcadores en un 88 % en la curva ROC y valores mayores al 90% en especificidad y sensibilidad.

Patrones biomarcadores, Lesiones precancerosas de las cuerdas vocales, Métodos de clasificación, Aprendizaje automático

Abstract

The National Survey on Drug, Alcohol and Tobacco, 2016-2017, notes that 15.6 million Mexicans are active smokers and, by 2030, expect the death of 8 million cancers of the larynx or lung. Therefore, the World Health Organization (WHO) recommends detecting precancerous lesions of the larynx. This is possible, as they are characterized by a biomarker pattern manifested by the alteration of the biomechanical interpretation of the vocal cords, regardless of the sex and age of the smoker. The goal of this article is to evaluate three machine learning methods: neural networks, Gaussian networks, and decision tree to determine the method that best solves the problem of detecting patterns of precancerous vocal cord injury biomarkers. It uses the WEKA tool and a knowledge bank, endorsed by NOM-012-SSA3-2012, with 250 patterns, and provided by the Luis Guillermo Ibarra National Institute of Rehabilitation, Ibarra. The performance of the methods was compared by ROC curves and confusion matrices, under the criteria established by ISO-5725. The decision tree the method that best responds to the detection of patterns of biomarkers of precancerous lesions of the vocal cords.

Biomarker patterns, Precancerous vocal cord lesions, Classification methods, Machine learning

Citación: SIORDIA-VASQUEZ, Xóchitl, VILLAGRAN-VILLEGAS, Luz Yazmin, PATIÑO-ORTIZ, Miguel y ROJAS-HERNÁNDEZ, Miguel Ángel. Selección de un método de aprendizaje automático para clasificar patrones biomarcadores de lesiones precancerosas de las cuerdas vocales. Revista de Didáctica Práctica. 2020. 4-11:1-7.

*Correspondencia al Autor (Correo Electrónico: xsiordia@uv.mx)
† Investigador contribuyendo como primer autor.

Introducción

El cáncer de laringe es una enfermedad relacionada con la exposición a diversos agentes tóxicos, contenidos en el humo del cigarro, estas sustancias se mantienen atrapadas en la membrana mucosa, provocando el deterioro celular del tejido epitelial, hasta evolucionar en un tumor maligno que se hace visible sobre las cuerdas vocales. La Encuesta Nacional de Consumo de Drogas, Alcohol y Tabaco, 2016-2017, señala que 15.6 millones de mexicanos son fumadores activos; y el Sistema Nacional de Proyecciones en Salud Publica de México ((SINAIS), 2013) estima, para el año 2030, la muerte de 8 millones de fumadores, jóvenes de edades entre 45 y 55 años, a causa del cáncer laríngeo.

La detección temprana de las lesiones precancerosas de la laringe es importante, dado que el éxito en el tratamiento de este tipo de lesiones disminuye en etapas avanzadas, produciendo, la mayoría de las veces, la muerte. Por ello, la Organización Mundial de la Salud (OMS), recomienda desarrollar estrategias que ayuden a detectar, de forma no invasiva, las lesiones precancerosas de la laringe entre la población de fumadores.

Recientes investigaciones han utilizado las ondas acústicas de las señales de voz para obtener información que ayuda a caracterizar patrones biomarcadores de las lesiones precancerosas de la laringe, a partir de parámetros biológicos que cuantifican una alteración en el desempeño biomecánico que ocurre en las cuerdas vocales de los fumadores. Esta metodología, no invasiva, ofrece la posibilidad de generar herramientas para la detección de lesiones precancerosas de las cuerdas vocales a bajo costo (México Patente nº En trámite, 2015) (Gómez-Vilda, y otros, 2009)

El conjunto de biomarcadores propuesto por Gómez Vilda consiste de 68 parámetros y no existe un único criterio para la categorización de las lesiones precancerosas de las cuerdas vocales, a diferencia, el conjunto de biomarcadores propuesto por Siordia, contiene 8 parámetros específicamente diseñados para medir la alteración biomecánica en lesiones precursoras del cáncer de laringe asociado a hábitos de tabaquismo, sin embargo, no existe evidencia reportada en la literatura que muestre el proceso de unificación de criterios para la categorización de los patrones biomarcadores que Siordia obtiene (Siordia Xóchitl, 2016).

En este artículo se propone la utilización de técnicas de reconocimiento automático de patrones para evaluar tres métodos de aprendizaje automático: redes neuronales, redes bayesianas y árboles de decisión, y seleccionar el que mejor logre aprender las diferencias entre tres clases de patrones biomarcadores propuestos por Siordia et al (Siordia Xóchitl, 2016); correspondientes a lesiones precancerosas de las cuerdas vocales, estrés biomecánico de las cuerdas vocales o laringitis viral u otras afecciones de tipo respiratorio, como una estrategia que ayude a detectar la alteración funcional del desempeño biomecánico que las lesiones precancerosas producen en los fumadores, aun cuando no haya signos visibles de un tumor.

En la sección 2 se explica a detalle la descripción de los materiales y métodos que se utilizaron para el preprocesamiento de los biomarcadores del banco de información, así como la descripción de los métodos de clasificación que se plantea evaluar. En la sección 3 se analizan los resultados de desempeño de cada uno de los métodos de clasificación propuestos y, en la sección 4, se muestra la selección del método que presenta la mejor solución para resolver el problema de detección de las lesiones precancerosas de las cuerdas vocales. Finalmente, se presentan las conclusiones y propuestas de trabajo futuro. (Siordia Xóchitl, 2016)

ISSN: 2523-2444

Materiales y métodos

La colección de los patrones de biomarcadores que forman el banco de conocimiento fue obtenida de pacientes del departamento de foniatría del Instituto Nacional de Rehabilitación Luis Guillermo Ibarra Ibarra de la Ciudad de México, quienes voluntariamente y bajo consentimiento informado participaron en un ensayo clínico, diseñado bajo la norma NOM-012-SSA3-2012.

En la Figura 1 se presenta el esquema del método para extraer la información que modela el patrón biomarcador a partir del banco de voces que proporciono el ensayo clínico.

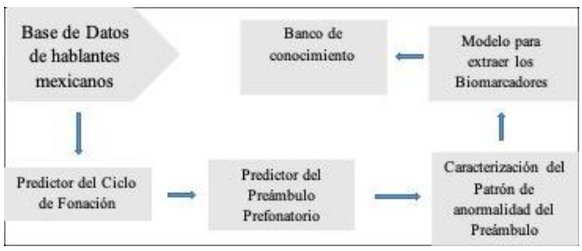

Figura 1 Método para extraer el banco de conocimientos a partir del banco de voces
Fuente: Elaboración Propia

El banco de conocimiento está formado por 250 patrones biomarcadores; se utilizó el software Excel para analizar los registros y reorganizarlos en un nuevo archivo .xls, donde los patrones biomarcadores se clasificaron en tres clases: clase A, que corresponde a las lesiones precancerosas; clase B, relacionada al estrés biomecánico de las cuerdas vocales; y clase C, asociada a laringitis viral o lesiones laríngeas causadas por problemas respiratorios; como se muestra en la figura 2.

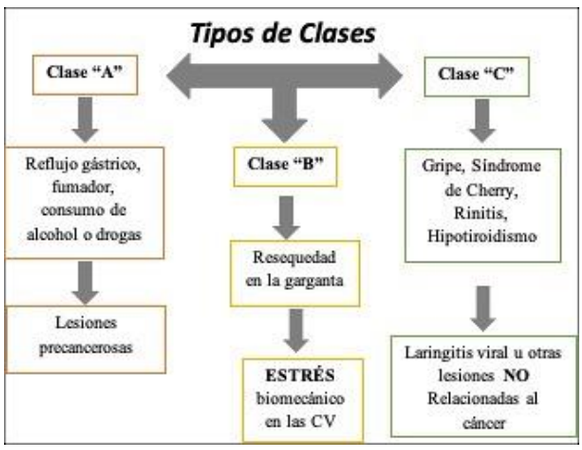

Figura 2 Método para distribuir y organizar los patrones biomarcadores el en banco de información
Fuente: Elaboración propia

Pre-procesamiento de la información

Una vez que se reorganizo el banco de información, éste se trasformó a un archivo de extensión .CVS, para su posterior tratamiento en la herramienta de minería de datos WEKA (Waikato Environment for Knowledge Analysis), que es una aplicación ampliamente utilizada para la experimentación sobre el reconocimiento de patrones y minería de datos, ya que incluye una variedad de algoritmos de procesamiento, clasificación y agrupación de datos e información. Posteriormente, los datos se normalizaron para evitar que la función de decisiones se vea influenciada por las variables de magnitudes considerablemente mayores, en este paso se estandarizan los valores en una misma escala.

Debido a que en el banco de conocimientos existía un desbalance entre las clases A y B, con respecto a la clase C, también fue necesario utilizar la herramienta SMOTE para balancear la distribución de cada clase en el banco de información y evitar sesgos en la clasificación (Mena, 2008).

Una vez balanceadas las clases, se procedió a revisar la distribución de los datos para cada uno de los parámetros que forman el patrón biomarcador, observándose que los valores de los parámetros estaban mezclados entre las tres clases.

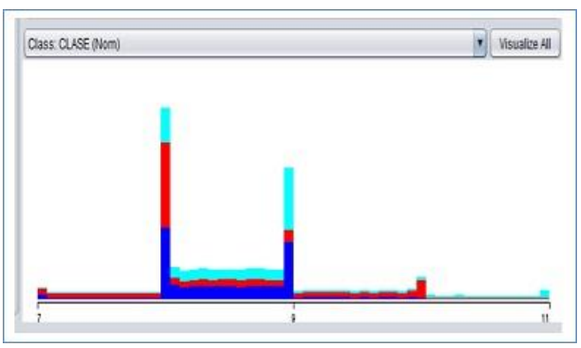

Figura 3 Ejemplo de los datos de un parámetro del patrón biomarcador sin discretizar
Fuente: Elaboración propia

En la Figura 3 se observa el ejemplo de este fenómeno, por lo que fue necesario realizar la discretización de los datos para obtener una mejor distribución de estos; en esta etapa se utilizó la herramienta DISCRETIZE de WEKA. La figura 4 muestra una clara separación entre los valores del parámetro por clase, después de aplicar la discretización.

ISSN: 2523-2444

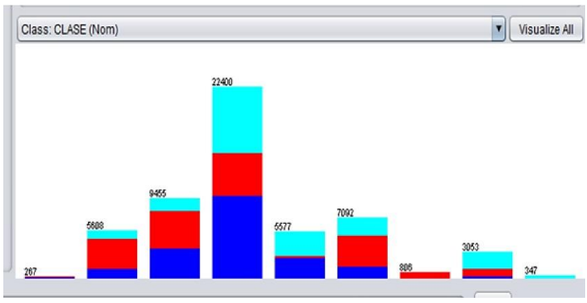

Figura 4 Separación de los valores de los datos de un parámetro del patrón biomarcador mediante la discretización
Fuente: Elaboración propia

El software WEKA posee una herramienta llamada "SMOTE", que permite generar patrones de datos artificiales a partir de los patrones originales, y asegura que los datos numéricos de los nuevos patrones estén contenidos entre los rangos numéricos originales, para asegurar la confiabilidad de los datos.

Debido a que el banco de información es de tamaño reducido, se decidió utilizar SMOTE para ampliar el número de biomarcadores de 250 a 1996, balanceados entre las clases A, B y C, con la finalidad de que el clasificador disponga de una mayor cantidad de información y logre generar un resultado confiable en el reconocimiento de los patrones que forman cada clase.

Selección de los algoritmos de clasificación

En WEKA es posible realizar experimentos de reconocimiento de patrones mediante el uso de algoritmos de clasificación que poseen aplicaciones escritas en leguaje java o usando módulos gráficos (Paramo Lozada, Espitia Betancourt, Menéndez Mora, & Holguín Ontiveros, 2018). En este trabajo se utilizó la segunda alternativa.

En la literatura, Goméz Vilda ha reportado el uso de clasificadores basados en Redes Bayesianas y Redes Neuronales para el reconocimiento de patrones biomarcadores, por lo que se decide probar estos mismos algoritmos por la similitud que guardan con esta investigación (Gorris & Pedro Gómez Vilda, 2020), (.J.I. Godino-Llorente, 2001). Además, se seleccionó el método de árboles de decisión por ser ampliamente utilizados en el ámbito médico para detección de patrones de cáncer (Flores & Dueñas, 2019), (Marante & Marante, 2020).

El método de clasificación basado en las redes Neuronales consiste en emular un sistema de procesamiento basado en el comportamiento biológico de las unidades de procesamiento del cerebro humano, llamadas neuronas, las cuales están organizadas en capas desde donde se recibe, se procesa y se trasmite la información. Los patrones son las entradas que alimentan a la red neuronal y a su vez se trasmiten como señales a la siguiente capa, los enlaces de comunicación entre cada capa se precisan por un peso que biológicamente simboliza el nivel de sinapsis en el enlace. Las salidas de las neuronas se forman en la última capa de la red neuronal y representan la respuesta de la red al estímulo del patrón que se ha introducido en la entrada (Gabriel Mauricio Martínez Toro, Bautista, & Romero-Riaño, 2019), existen una gran variedad de clasificadores basados en redes neuronales, en este artículo se emplea el algoritmo del perceptrón multicapa.

Para evaluar el desempeño de la red bayesiana se utilizó el método de clasificación basado en el algoritmo de Naive Bayes (Castrillón, Eduardo, & F.Castillo, 2018), que ha sido identificado como el más adecuado en el campo médico, por la exactitud que ofrece cuándo se aplica a grandes bases de datos (M. Hariz, 2012). Este algoritmo pertenece a la clase de clasificadores estadísticos que predicen la clase de pertenencia de un ejemplo X, mediante probabilidades, la posibilidad de que un evento dado pertenezca a una clase en particular, esto se basa en el supuesto que los efectos de un atributo son independientes de los valores de los otros atributos de un evento X, como se expresa en la ecuación 1.

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} \quad (1)$$

Otro de los métodos más populares en la comunidad científica es el método de clasificación basado en arboles de decisión C.45, creado por Ross J. Quinlan (Quinlan, 1993). WEKA ofrece este clasificador bajo el nombre de J48. Un árbol de decisión es un clasificador expresado como una partición recursiva del espacio de instancias que se establece como una máquina de aprendizaje y su estructura tiene una similitud a los diagramas de flujo.

ISSN: 2523-2444

En este método se realizan las predicciones basándose en el concepto de entropía minoritaria, expresada en la ecuación (2), usando la idea de proponer dos puntos de separación en el atributo seleccionado, de modo tal que se logre capturar la mayor cantidad de elementos de clase minoritaria en la partición central (Luis A. Caballero-Cruz, 2015)

$$minority - entropy = -\sum_{i=1}^{N} \frac{n_I}{n}\left(log\left(\frac{n_i}{n}\right)\right) \quad (2)$$

Resultados

En esta sección se presentan los resultados obtenidos de la implementación de los tres métodos de clasificación. Los experimentos se realizaron en una Lenovo con un procesador Intel Core i5 y memoria de 8 GB. El banco de información utilizado contiene 1966 patrones biomarcadores con los datos balanceados para las clases A, B y C de acuerdo a la distribución mostrada en la Tabla 1.

Clases	No. Patrones Biomarcadores
A	644
B	629
C	693
TOTAL	1966

Tabla 1 Distribución de los patrones biomarcadores en el banco de información

En las pruebas realizadas se utilizó la validación cruzada con k=10 pliegues, es decir, el conjunto de datos disponibles se divide en diez partes iguales, aprovechando una para la validación del modelo de clasificación y el resto para su entrenamiento, el resultado final se obtiene al promediar las métricas arrojadas en cada una de las pruebas (Enrique, Maya, Giovani, & Gustavo, 2019).

	Naives bayes	Perceptron Multicapa	Arboles de decisión J48
Prevalencia	49.66%	48.66%	61.66%
Sensibilidad	80.66%	74.66%	91.00%
Especificidad	82.00%	72.66%	90.66%
VPP	82.00%	72.33%	91.00%
VPN	81.33%	75.00%	92.00%
Error	17.66%	25.33%	7.66%
Precisión	81.33%	73.66%	91.33%
Exactitud	69.43%	58.69%	84.89%
Correctamente Clasificados	1365	1154	1669
Confusiones	30.56%	41.30%	15.10%
Confundido	601	812	297

Tabla 2 Resultados obtenidos con los métodos de clasificación propuestos

La validez de los tres modelos se efectuó comparando las mediciones del porcentaje de predicciones correctas, confusión, sensibilidad y especificidad. La Tabla 2 presenta los resultados obtenidos después de evaluar los tres métodos de clasificación, observándose que el desempeño del Árbol de decisión, basado en el clasificador J48, obtiene el 84% de aciertos, siendo superior al perceptrón multicapa que distingue el 59.69% y al clasificador bayesiano de Nayves-bayes que logra un 69.43%.

Al hacer el análisis de la sensibilidad y de especificidad el árbol de decisión presenta valores superiores al 90% que no fueron superados por los otros dos métodos.

Una vez obtenidos los resultados anteriormente descritos, se analizaron las áreas bajo la curva ROC que devuelve WEKA para cada clasificador.

En la Figura 5 se observa que el área bajo la curva de que devuelve WEKA al evaluar el clasificador de Nayves Bayes obtiene el 87% de reconocimiento.

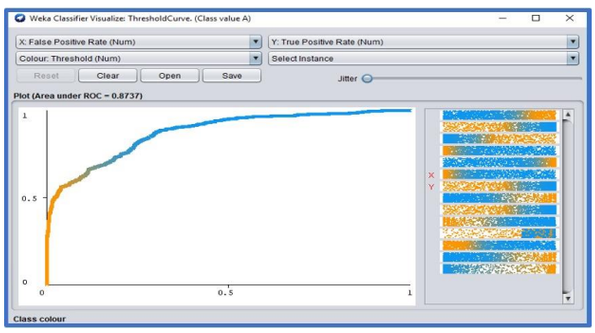

Figura 5 Área bajo la curva ROC del método de clasificación de Nayves Bayes con un 87% de reconocimiento
Fuente: Elaboración propia

La respuesta del clasificador basado en la red neuronal de perceptrón multicapa se muestra en la Figura 6, donde el área bajo la curva ROC devuelve un porcentaje de reconocimiento del 75%, siendo menor que el conseguido por el algoritmo de Bayes.

ISSN: 2523-2444

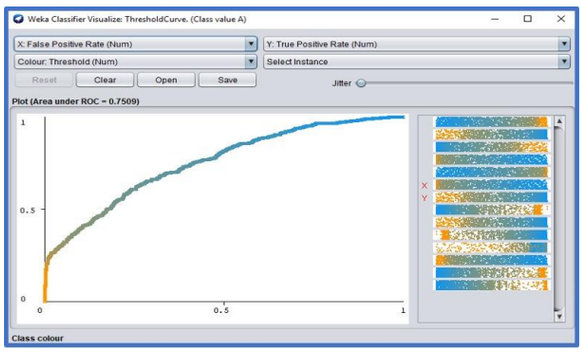

Figura 6 Área bajo la curva ROC del clasificador basado en red neuronal de retropropagación con un 75% de reconocimiento.
Fuente: Elaboración propia

El desempeño del clasificador basado en el algoritmo J48 se presenta en la Figura 7, logrando un área bajo la curva ROC del 88%.

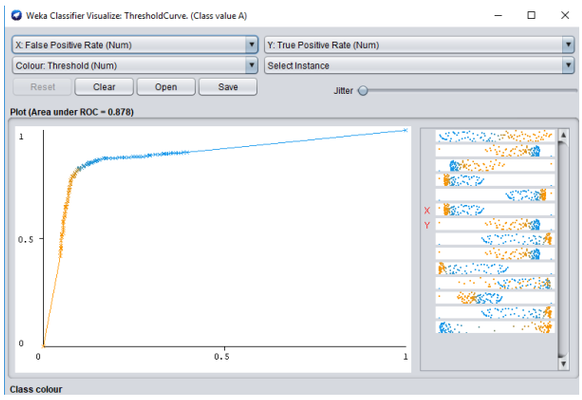

Figura 7 Área bajo la curva ROC del clasificador por árbol de decisión j48 con un 88% de reconocimiento
Fuente: Elaboración propia

Conclusiones

El uso de patrones biomarcadores de la alteración biomecánica de las cuerdas vocales, en conjunto con un sistema de reconocimiento de patrones, entrenado con una amplia base de datos certificada por el ámbito clínico y bajo el criterio de expertos especializados en detección de lesiones precancerosas en fumadores, abre la posibilidad de diseñar una estrategia para detectar de forma automática las lesiones precancerosas, aún y cuando no se presentan signos visibles de un tumor sobre la superficie de las cuerdas vocales, tal acción posibilita el desarrollo de pruebas de tamiz de bajo costo para la detección temprana del cáncer entre la población de fumadores mexicanos.

En cuanto a los resultados obtenidos al evaluar cada uno de los métodos de clasificación elegidos, se concluye que el árbol de decisión basado en el algoritmo J48 presenta una probabilidad del 84% de aciertos, el 88% de que el modelo clasifique como positivas las instancias positivas, además el alto porcentaje en sensibilidad y especificidad mayor al 90%, indican que el modelo posee una mejor capacidad para distinguir entre los patrones biomarcadores de cada una de las tres clases, no distingue de forma binaria entre un patrón biomarcador asociado a una lesión maligna o benigna.

De forma general, todos los algoritmos presentados ofrecen buen desempeño, sin embargo, es importante resaltar la necesidad de extender la experimentación usando una mayor cantidad de patrones biomarcadores, debido a que no se logra alcanzar el 97% establecido por la norma ISO-5725.

Dada la complejidad de ampliar la cantidad de muestras del banco de ondas acústicas originales, se implementan técnicas para la generación de patrones biomarcadores artificiales creados a partir del banco de conocimiento, los cuales ya se tienen a disposición, y en este tema se encamina el trabajo futuro de esta investigación.

Agradecimiento

Se agradece al Instituto Nacional de Rehabilitación todas las facilidades para el uso del banco de hablantes mexicanos.

Referencias

(SINAIS), S. N. (2013). *Proyecciones de la Población en México.* México,d.F: Dirección general de Información en Salud (DGIS). Base de datos de defunciones 1979-2008.

.J.I. Godino-Llorente, S. A.-N.-V. (2001). Automatic detection of voice impairments due to vocal misuse by means of Gaussian mixture models,". *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2*, págs. 1723-1726. Estambul, Turquia: IEEE. doi:10.1109/IEMBS.2001.1020549.

Arango V., S. S. (2011). Biomarcadores para la evaluación de riesgo en la salud humana. *Facultad Nacional de Salud Pública, 30*(1), 75-82.

Cambrigde University, P. (2016). *Cambrigde Dictionary*. Obtenido de http://dictionary.cambrigde.org/dictionary/english/navigation

Castrillón, O. D., E. C., & F.Castillo, L. (2018). Sistema Predictivo Bayesiano para Detección del Cáncer de Mama. *Información Tecnológica, 29*(3), 257-270. doi:http://dx.doi.org/10.4067/S0718-07642018000300257

Enriquez, I. C. (2017). *Diseño y navegacion de un vehiculo terrestre no tripulado con sistema de navegacion autonomo.* Cd. de México: Universidad Autonoma de México.

Flores, L. A., & Dueñas, A. M. (2019). Sistema Experto Probabilístico basado en Redes Bayesianas para la predicción del cáncer de cuello uterino. *Revista peruana de Computación y Sistemas, 2*(1). doi:https://doi.org/10.15381/rpcs.v2i1.16360

Gabriel Mauricio Martínez Toro, Bautista, D. R., & Romero-Riaño, E. (2019). Análisis comparativo de predicción dentro de bases de datos de cáncer: una aplicación de aprendizaje automático. *Revista Ibérica de Sistemas e Tecnologias de InformaçãoIberian Journal of Information Systems and Technologies,* 113-122.

Gorris, J. M., & Pedro Gómez Vilda, &. o. (Octubre de 2020). *Artificial intelligence within the interplay between natural and artificial computation: Advances in data science, trends and applications* (Vol. 210). Elseiver. doi:https://doi.org/10.1016/j.neucom.2020.05.078

Luis A. Caballero-Cruz, A.´.-C.-L. (2015). Arbol´ de decisi´on C4.5 basado en entrop´ıaminoritaria para clasificaci´on de conjuntos dedatos no balanceados. *Research in Computing Science, 92*, 23-34.

M. Hariz, M. A. (2012). Hybrid Approaches Using Decision Tree, Naive Bayes, Means and Euclidean Distances for Childhood Obesity Prediction. *6*(3), 99-106.

Marante, Y. T., & Marante, Y. T. (2020). Evaluación y clasificación de las imagenes de células de cervix usando rassgos morfológicos. *Universidad & Ciencia, 9*(1), 58-68.

Mena, C. (2008). *Aprendizaje automático a partir de un conjunto de datos no balanceados y su aplicación en el diagnóstico y pronóstico médico.* Tozantzintla, Puebla: Instituto Nacional de Astofisica optica y Electrónica.

Paramo Lozada, J. P., Espitia Betancourt, C. A., Menéndez Mora, R. E., & Holguín Ontiveros, E. P. (2018). *Aplicación del aprendizaje automático en la clasificación de textos cortos. Un caso de estudio en el conflicto armado colombiano.* Universidad Católica de Colombia, Facultad de Ingeniería, departamento de Ingeniería de sistemas. Bógota, Colombia: Universidad Católica de Colombia. Recuperado el 20 de Agosto de 2019

Quinlan, J. (1993). *C4.5: Programs forMachine Learning.* San Francisco, CA, USA : Morgan Kaufmann Publis-hers Inc., .

Schworer, I. (2005). *Navigation and control of an autonomous vehicle.* Virginia: Virginia Polythecnic Institute.

Sinder, J. (2009). *Automatic Steering Methods for Autonomous.* Pitsburg: Caenegie Mellon University.

Siordia Xóchitl, C. J. (2016). Parámetros biomecánicos de la membrana epitelial: un biomarcador para el pronóstico de riesgo del cáncer de laringe. *Revista Tecnología e Innovación, 1*(1), 1-15.

Society, M. (2013). The New England Journal of Medicine. Massachusetts.

.