

Aplicación móvil inteligente para prevención de cáncer de mama

CUAYA-SIMBRO, Germán†*, RUÍZ-HERNÁNDEZ, Elías, HERNÁNDEZ-HERNÁNDEZ, Luis Ángel y LIMA-LUNA, Luis Alberto

Instituto Tecnológico Superior del Oriente del Estado de Hidalgo. Carretera Apan-Tepeapulco Km 3.5, Col Las Peñitas, Apan, Hidalgo, México

Recibido 3 de Julio, 2017; Aceptado 19 de Septiembre, 2017

Resumen

De acuerdo al Instituto Nacional de Estadística y Geografía (INEGI) el total de casos de cáncer diagnosticados en la población mexicana de 20 años es el cáncer de mama. Una de las principales razones por la cual no se detecta a tiempo el cáncer de mama es por no tener herramientas de prevención al alcance de la población afectada. Esta clínicamente comprobado que la detección temprana de cáncer de mama ayuda a tener una mayor probabilidad de supervivencia. Una medida alternativa de detección temprana de cáncer de mama es la autoexploración, por lo que es importante fomentar este hábito para evitar las consecuencias de una detección tardía de la enfermedad. El objetivo de este trabajo es presentar los avances en el desarrollo de una aplicación móvil inteligente la cual apoye en la detección temprana de riesgo de padecer cáncer de mama, la aplicación estará basada en la captura de datos temporales y en el análisis de dicha información mediante técnicas de minería de datos, los datos capturados serán los resultados de una autoexploración de una persona

Prevención, aplicación móvil, minería de datos

Abstract

According to the National Institute of Statistics and Geography (INEGI), the total number of cancer cases diagnosed in the Mexican population of 20 years or more is the breast cancer. One of the main reasons to breast cancer is not detected early is because it doesn't have prevention tools available to the affected population. Clinically has been proved that the early detection of breast cancer helps to have a greater chance of survival. An alternative action to early detection of breast cancer is self-examination, so it is important to encourage this habit to be able to take preventive measures and avoid consequences of a late detection of the disease. Thus, the objective of this work is to present the progress of the development of an intelligent mobile application which supports the early detection of risk of breast cancer, the application will be based on the capture of temporal data and the analysis of such information by data mining techniques, the data will be captured from the results of a self-exploration of a person.

Prevention, mobile application, data mining

Citación: CUAYA-SIMBRO, Germán, RUÍZ-HERNÁNDEZ, Elías, HERNÁNDEZ-HERNÁNDEZ, Luis Ángel y LIMA-LUNA, Luis Alberto. Aplicación móvil inteligente para prevención de cáncer de mama. Revista de Ingeniería Tecnológica 2017. 1-3:9-17

*Correspondencia al Autor (Correo Electrónico: gcuaya@itesa.edu.mx)

† Investigador contribuyendo como primer autor.

Introducción

De acuerdo al Instituto Nacional de Estadística y Geografía (INEGI), en 2014, del total de casos de cáncer diagnosticados en la población mexicana de 20 años, el de mama es el de mayor presencia con un 19.4%. A nivel mundial se detectan 1.38 millones de casos nuevos de cáncer de mama y hay 458,000 muertes siendo este el tipo de cáncer con mayor incidencia entre las mujeres. (INEGI, 2017)

Actualmente se conocen distintos exámenes para la determinación de cáncer de mama cómo lo son, estudios de mastografía, exámenes clínicos de la mama (ECM), imágenes por resonancia magnética (IRM). Estos estudios se realizan de manera esporádica y generalmente bajo las instrucciones de algún médico y tienen la finalidad de confirmar el diagnóstico del especialista, esto es, los estudios no son realizados con fines de prevención. La prevención en este tipo de enfermedades es importante debido a que, si el cáncer es detectado tarde se convertirá en un problema muy grave, además de que puede conllevar problemas mayores y la posibilidad de sobrevivir se reduce, cuando el tejido anormal o el cáncer se detectan a tiempo, puede ser más fácil tratarlo.

Una alternativa para tener un diagnóstico preventivo y así evitar las consecuencias de una detección tardía de cáncer de mama, es la autoexploración periódica, debido a que se ha confirmado que este tipo de estudios personalizados ayudan a detectar de manera temprana la posibilidad de padecer este tipo de cáncer. Desafortunadamente no se cuenta con la cultura suficiente para llevar periódicamente este tipo de exámenes y aún más guardar la información adquirida en cada uno de ellos, para que a partir del análisis de estos datos se pueda determinar el riesgo de padecer esta enfermedad. (American Cancer Society, 2015)

Por lo que el objetivo de este trabajo de investigación es presentar los resultados obtenidos hasta el momento los cuales consisten en el desarrollo de una aplicación móvil en android y el análisis de bases de datos de cáncer de mama mediante algoritmos de minería de datos, lo cual permitiera el desarrollo de una aplicación móvil inteligente que permitirá fomentar la cultura de autoexploración, además, de permitir detectar cambios que impliquen el riesgo de padecer cáncer de mama y mandar una alerta al usuario para que tome acciones de corroboración de dicha enfermedad. Esta aplicación será de fácil adquisición y manejo aprovechando la tendencia en el uso de aplicaciones móviles.

Este artículo consta de las siguientes secciones; Trabajos Relacionados, donde se contrastan los trabajos similares al nuestro; Métodos, donde se describen las técnicas y métodos utilizados para realizar esta investigación; Experimentos, donde se redacta información relacionada al diseño de las pruebas para nuestra investigación; Resultados, en esta sección se presentan y analizan los resultados de los experimentos realizados; Conclusiones y Trabajo Futuro, donde se presentan las conclusiones derivadas de esta investigación, así como las perspectivas a desarrollar.

Trabajos relacionados

Existen trabajos reportados de investigaciones en cuanto a cáncer de mama y su detección temprana, entre los primeros podemos citar a los trabajos con un enfoque médico, los cuales están basados en el análisis de cómo la nutrición, el estilo de vida y la interacción con el ambiente ayuda en la prevención de este tipo de cáncer, como se muestra en (Howell, y otros, 2014), (Key, y otros, 2004).

Por otro lado, desde el área de sistemas computacionales se han realizado diversos estudios en los cuales se aplican técnicas de minería de datos, las cuales permiten realizar la evaluación de diferentes modelos de prevención y diagnóstico de cáncer de mama; estos trabajos reportan el uso de bases de datos con información de pacientes con cáncer de mama y muestran la efectividad de los modelos para detectar o determinar si un paciente tiene riesgo de padecer este tipo de cáncer, como los trabajos (Liou, 2016), (Asri, 2016).

Desafortunadamente, estos estudios solo reportan los resultados crudos de la investigación, pero no se desarrolla algún tipo de sistema el cuál permita hacer uso de dichos resultados, como pudiera ser la construcción de una aplicación móvil.

Por otro lado, en la actualidad los teléfonos móviles o teléfonos inteligentes (Smartphones) son una alternativa viable para poner al alcance nuevas tecnologías o herramientas de apoyo en la detección temprana de cáncer. Se han reportado trabajos en donde se han desarrollado aplicaciones móviles relacionadas al cáncer de mama, por ejemplo: “My cancer coach” (BreastCancer.org, 2017), ayuda a administrar la rutina diaria de gente con cáncer basada en el diagnóstico determinado por el médico, así como el tratamiento a seguir, esta aplicación está basada en grabaciones de audio, calendario, toma de notas y preguntas y respuestas al médico.

Pero esta aplicación, está orientada a la población estadounidense y es completamente desarrollada en inglés, además no permite hacer prevención de cáncer como la que proponemos en nuestro trabajo; “B4BC” (Black Sun Production, 2016), una aplicación orientada principalmente a la comunidad deportiva, la cual permite configurar recordatorios para autoexámenes e incluso puede programar dichos autoexámenes de acuerdo al ciclo menstrual del usuario, esta aplicación ofrece instrucciones sobre cómo realizar un autoexamen, tips médicos para prevención del cáncer, recetas y consejos para llevar una vida saludable para mujeres activas, pero de igual forma esta aplicación está orientada a la población estadounidense y está completamente en inglés, además no realiza un análisis para prevención de acuerdo a cambios registrados en los autoexámenes realizados; “Breast Cancer: Beyond the Shock” (National Breast Cancer Foundation, 2017), esta aplicación proporciona diversa información relacionada al cáncer de mama, puede aprender, preguntar y escuchar, esto es, permite ver video los cuales te enseñan los diferentes tipos de tumores y tratamientos, además de presentar historias de mujeres en la vida real quienes padecen o padecieron este tipo de cáncer, pero básicamente es una aplicación informativa, y de igual forma que las otras dos aplicaciones mencionadas, está desarrollada en inglés para la población estadounidense. Cabe señalar que existen otras aplicaciones similares pero la mayoría están desarrolladas en inglés.

Adicionalmente, en el Centro de Investigación y de Estudios Avanzados (CINVESTAV) creó un software para detectar cáncer de mama. – que principalmente consiste en ayudar en el entrenamiento de radiólogos, así como dar certeza al emitir diagnósticos y evitar las biopsias innecesarias.

El sistema se basa en un software al que previamente “se le entrenó” con un banco de imágenes representativas de cientos de casos de cáncer (Ramírez-Reyes, 2013).

Por otro lado, los estudiantes del Instituto Tecnológico de Colima (ITEC) diseñaron un sostén para detectar este tipo de cáncer de mama de fácil funcionamiento para poder ser utilizado en casa. Sirve para medir la temperatura de los senos y con ello, detectar las condiciones iniciales del cáncer, cuando se va a implantar un tumor, ya que los senos sufren cambio en su patrón termográfico (Ramírez, 2016).

También, existe una aplicación móvil para avisar que se debe hacer la auto exploración una semana después del periodo menstrual de la mujer, en caso de ser mujeres mayores podrán programar el día del recordatorio. Otra de sus características es que enseña paso a paso como realizarse la autoexploración ya sea en imágenes o en un vídeo.

En contraste con la aplicación móvil que se esta desarrollando, ésta tendrá la función de dar recordatorio al usuario sobre el día que debe realizar una autoexploración, al final de esta, el usuario tendrá que llenar un formulario predefinido con datos obtenidos durante la autoexploración, esta acción será realizada mensualmente, la aplicación analizará automáticamente los datos obtenidos para así, determinar si existió algún cambio y con base en ello, dar una alerta al usuario para realizarse algún estudio más detallado propiciando la prevención de la ocurrencia de cáncer.

Métodos

Minería de datos

La definición para ésta puede variar dependiendo su uso y al autor, pero en resumen minería de datos es la extracción de conocimiento a partir de la información almacenada en una base de datos. (Valiente, 2015).

Para realizar minería de datos se han desarrollado diferentes técnicas que permiten a los científicos la clasificación de la información y con los resultados formular alguna hipótesis. Estas técnicas se dividen en dos tipos:

El método descriptivo: Tiene como objetivo la búsqueda de patrones interpretables para describir datos.

El método predictivo: Este utiliza algunas variables para la predicción de valores futuros o desconocidos de otras variables. Estas variables pueden ser de clasificación, regresión y detección de la desviación. Esta técnica a funcionado de manera exitosa en el ámbito de la medicina.

En nuestro caso específico, para el análisis de datos se utilizaron técnicas de clasificación para fines predictivos. Estas técnicas de clasificación consisten en determinación de clases (etiquetas o valores), asociadas a valores continuos, entre las técnicas podemos mencionar, conjunto de reglas, árbol de decisión, redes neuronales, etc. La clasificación consiste en predecir una determinada clase para un objeto.

Los clasificadores utilizados sobre los datos analizados fueron los siguientes:

Bayes Net: Basado en redes bayesianas las cuales modelan un fenómeno mediante un conjunto de variables y las relaciones de dependencia entre ellas. Dado este modelo, se puede hacer inferencia bayesiana; es decir, estimar la probabilidad posterior de las variables no conocidas, en base a las variables conocidas.

J48: Utiliza el algoritmo C4.5 utilizado para generar un árbol de decisión. En cada nodo del árbol, C4.5 elige un atributo de los datos que más eficazmente dividen el conjunto de muestras en subconjuntos enriquecidos en una clase u otra. El atributo con la mayor ganancia de información normalizada se elige como parámetro de decisión.

Bagging: Ofrece un incremento sobre la precisión de cualquiera de los clasificadores individuales utilizados, especialmente si el modelo está sobre ajustado. La idea básica es re-muestrear los datos y calcular las predicciones sobre el conjunto de datos re-muestreados.

SMO: Este algoritmo está basado en redes neuronales (funcionamiento inspirado en el cerebro humano, de ahí su nombre) cuya característica más importante es su capacidad de aprender a partir de ejemplos, lo cual les permite generalizar sin tener que formalizar el conocimiento adquirido. Se caracterizará por tener un aprendizaje no supervisado competitivo y por no tener ningún resultado objetivo al que la red deba tender.

Weka

Para el procesamiento de los datos utilizamos la herramienta Weka. Weka es una herramienta de software para el aprendizaje automático y minería de datos diseñado a base de Java y desarrollado en la universidad de Waikato en Nueva Zelanda en el año 1993.

Permite realizar minería de datos con una interfaz gráfica para facilitar su utilización. Además, permite una comparación con los distintos métodos que se utilizan para el pre procesamiento, clasificación de información, clustering y meta-aprendizaje, contiene una colección de algoritmos para realizar análisis de datos y modelado predictivo, también tiene herramientas para la visualización de estos datos (Valiente, 2015).

Aplicaciones móviles

Una aplicación móvil es un programa que se puede descargar y al que se puede acceder directamente desde el teléfono o desde algún otro aparato móvil como una Tablet. Para poder descargar una aplicación móvil, es necesario tener acceso a internet. Actualmente ya existen aplicaciones móviles que resuelven diferentes problemas cotidianos en casi todos los sectores.

Es por lo anterior que se considera un área de oportunidad el realizar aplicaciones que puedan ser utilizadas por diferentes sectores sociales y que dichas aplicaciones permitan apoyar en el monitoreo, detección y prevención de diferentes enfermedades, en nuestro caso, prevención de cáncer de mama.

Plataformas

Los dispositivos móviles, al igual que los ordenadores, necesitan sistemas operativos que manejen el hardware y ofrezcan funcionalidades básicas para el software de aplicación. A continuación, se mencionan los sistemas operativos más populares en la actualidad.

iOS: Sistema operativo móvil de la empresa Apple Inc. Originalmente desarrollado para el iPhone (iPhone OS), siendo después usado en dispositivos como el iPod Touch, iPad y el Apple TV. Apple, Inc.

Android: Esta basado en el kernel de Linux diseñado principalmente para dispositivos móviles con pantalla táctil, como teléfonos inteligentes o tabletas, inicialmente desarrollado por Android, Inc. Google respaldó económicamente y más tarde compró esta empresa en 2005.

En nuestro caso, seleccionamos la plataforma Android Studio para el desarrollo de nuestra aplicación debido principalmente a que la mayoría de los dispositivos móviles cuentan con este sistema operativo.

Experimentos

Datos

Los experimentos realizados están basados en el análisis de dos bases de datos con información de cancer de mama, extraídas del repositorio de la Universidad de California en Irvine (UCI). La primera base de datos (DB 1) consta de 569 registros con 31 atributos cada uno, de los cuales, se puede observar si el paciente registrado pose un tumor maligno o benigno. La segunda base de datos (DB 2) contiene 198 registros con 34 atributos, en esta base de datos se puede visualizar la recurrencia o no recurrencia del cáncer de mama en los pacientes de la base de datos 1. Adicionalmente, se construyeron dos bases de datos más a partir de las bases de datos DB 1 y DB 2; la primera base construida (DB 3) contiene todos los registros de ambas bases, DB 1 y DB 2, es decir, contiene 767 registros, cada uno con 32 atributos, donde, para unificar la clase se consideró el atributo de recurrencia equivalente a cáncer maligno y el de no recurrencia equivalente a cáncer benigno. La segunda base construida (DB 4) considera aquellos registros que existen en ambas bases de datos, DB 1 y DB 2, por lo que esta base cuenta con 139 registros con 65 atributos, en donde el atributo que determina la clase es el de recurrencia o no recurrencia.

Weka

Como se ha mencionado se utilizó la herramienta de Weka, en específico, los algoritmos de clasificación, los algoritmos utilizados en nuestra investigación fueron: Bayes Net, J48, Bagging y SMO. Para evaluar la efectividad de clasificación de cada uno de ellos, se utilizó validación cruzada de 10 pliegues y para determinar la efectividad de cada uno de ellos se reportó los valores de precisión y recuerdo. Donde precisión es la probabilidad de padecer la enfermedad si el resultado de la prueba diagnóstica es positivo y el recuerdo es la proporción de enfermos correctamente identificados.

Resultados

Aplicación móvil

Como se ha mencionado, se ha desarrollado hasta el momento una primera versión de la aplicación móvil que permitirá enviar recordatorios para realizar una autoexploración periódicamente, ingresar y almacenar los datos obtenidos de dicha exploración y finalmente, enviar alertas de acuerdo al análisis de los datos para tomar medidas de corroboración como puede ser visitar a un médico o realizarse algún tipo de estudio.

A continuación, se muestran imágenes de la aplicación desarrollada hasta el momento.



Figura 1 Pantalla de inicio de sesión

Fuente: Elaboración Propia



Figura 3 Pantalla de configuración de alarma

Fuente: Elaboración Propia



Figura 2 Pantalla de los factores de riesgo

Fuente: Elaboración Propia

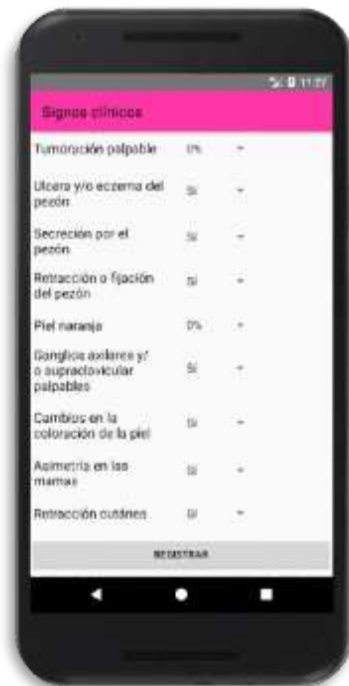


Figura 4 Pantalla de signos clínicos del paciente

Fuente: Elaboración Propia

Minería de datos

Como se ha comentado se han probado diversos clasificadores con el propósito de verificar la efectividad de cada uno de ellos y así, determinar la viabilidad de incorporar como motor de inferencia a alguno de ellos a nuestra aplicación móvil. Además de verificar con cuál de las bases de datos utilizadas hasta el momento se obtienen mejores resultados.

Los resultados obtenidos de precisión y recuerdo se muestran en los gráficos 1 y 2 respectivamente.

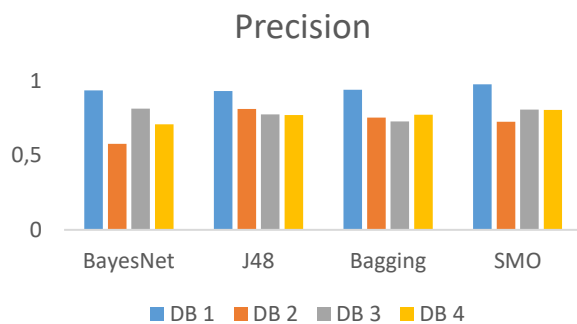


Gráfico 1 Comparación del valor de precisión obtenido por los clasificadores en las diferentes bases de datos

Fuente: *Elaboración Propia*

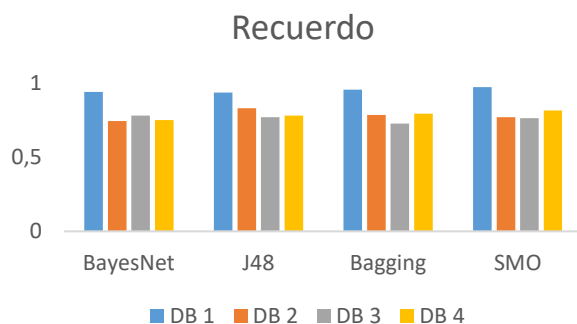


Gráfico 2 Comparación del valor de recuerdo obtenido por los clasificadores en las diferentes bases de datos

Fuente: *Elaboración Propia*

Conclusiones y trabajo futuro

De acuerdo a los resultados obtenidos, se observa que el uso de la segunda base de datos y de las que fueron construidas en combinación de ambas, no producen mejores resultados, debido, probablemente, a que la segunda base de datos contiene información pertinente a la recurrencia de la enfermedad y no tanto al padecerla por primera vez.

Los resultados indican que cualquiera de los clasificadores probados da una efectividad similar por lo que cualquiera de ellos es viable de ser el motor de inferencia en nuestra aplicación.

Como trabajo futuro se tienen las siguientes actividades; la terminación de la aplicación móvil, en específico el módulo de inferencia, el de almacenamiento de datos y el módulo de recordatorio, por otro lado, se realizará la adquisición de datos reales de exámenes de autoexploración, así como su análisis y procesamiento, también se realizarán pruebas de la aplicación ya terminada con diferentes usuarios para evaluar su eficiencia, y finalmente, con base en los resultados obtenidos y la valoración de la aplicación, se realizará un proceso de mejora de la misma.

Referencias

American Cancer Society. (2015). Breast Cancer Facts & Figures 2015-2016. Atlanta: American Cancer Society.

Asri, H. (2016). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Computer Science*.

Black Sun Production. (2016). B4BC Boarding 4 Breast Cancer.

BreastCancer.org. (2017). My Cancer Coach. Genomic Health.

Howell, A., Anderson, A., Clarke, R., Duffy, S., Evans, D., Garcia-Closas, M., . . . Harvie, M. (2014). Risk determination and prevention of breast cancer. *Breast Cancer Res.*

INEGI. (2017). Estadística a propósito del día mundial del cáncer de mama.

Key, T., Schatzkin, A., Willet, W., Allen, N., Spencer, E., & Travis, R. (2004). Diet, nutrition and the prevention of cancer. Oxford: The Authors.

Liou, D. (2016). Applying data mining for the analysis of breast cancer data. Taiwan : Yang Ming University.

National Breast Cancer Foundation. (2017). Breast Cancer: Beyond The Shock.

Ramírez, J. A. (2016). Sistema de monitoreo de temperatura de los senos humanos en la detección temprana del cáncer de seno. Colima: ITEC.

Ramírez-Reyes, A. (2013). Investigación y Desarrollo Tecnológico en Mamografía Digital. CINVESTAV.

Valiente, A. B. (2015). Inteligencia en redes de comunicaciones. Diagnóstico cardiología.