Aplicación de las Ciencias de la computación

Francisco Flores Ibarra1, Alejandro Vázquez Marín2, Dr. Ignacio de Jesús Segovia Domínguez3

1 Lic. En Sistemas Computacionales, Universidad Autónoma de Nayarit, Ciudad de la Cultura "Amado Nervo", Tepic, Nayarit. México. C.P. 63155

2 Lic. En Matemáticas, Universidad Veracruzana,

3 Instituto Tecnológico Superior de Guanajuato.

1 [francisco.ibarra@uan.edu.mx](mailto:francisco.ibarra@uan.edu.mx), 2 [alejandrovazquezmarin@yahoo.com.mx](mailto:alejandrovazquezmarin@yahoo.com.mx) , 3 [jsegovia@itesg.edu.mx](mailto:jsegovia@itesg.edu.mx).

## Resumen

El presente trabajo consiste en la investigación acerca de la calidad con la que se realiza una colposcopia en tres distintas modalidades: Green, Hinselmann y Schiller, a través de la creación de una página web que cuenta con una base de datos y en la que un usuario puede ingresar (ya sea de manera aleatoria o manual) 62 datos correspondientes a una exploración y de esta manera pueda elegir uno de los dos algoritmos de clasificación disponibles (KNN, y Naive Bayes) que le permitirá determinar la pertenencia de la clase (buena calidad o mala calidad).

El resultado es una página web que sirva como herramienta de apoyo en la investigación acerca de la calidad de las diferentes exploraciones de colposcopía y brinde la oportunidad de obtener diversas interpretaciones de los resultados.

## Introducción

“El auténtico genio consiste en la capacidad para evaluar información incierta, aleatoria y contradictoria.” (Winston Churchill, estadista) (Universitat Oberta de Catalunya. Estudios de Ciencias de la Información y de la Comunicación, n.d.)¿Qué habrá querido decirnos Winston con dicha frase?, Quizás fue conclusión de lo que realizó en su vida como estadístico, o de sus observaciones como político en la época de la Segunda Guerra Mundial; la verdad es que esta frase en nuestros tiempos adquiere mayor significancia por ser ahora motivadora en el estudio del análisis de datos y en la contribución de esto a las ciencias y su desarrollo.

Internet puede ser tanto una herramienta de comunicación como de acceso a la información, según las necesidades del usuario, brinda la oportunidad de romper las barreras del tiempo y la geografía, el fácil acceso y la utilización de algunos de sus servicios facilita el proceso de socialización, hace accesible la cultura, la ciencia, salud y el ocio

El objetivo principal de la realización de este proyecto es aportar en la búsqueda de conocimientos y tratamiento de la información desde una perspectiva cómputo-matemática haciendo un análisis de reconocimiento estadístico de patrones, resultando en aplicaciones en ciencias de la computación.

Se pretende construir un instrumento que se encargue de la clasificación de datos recopilados por los relacionados con el tema de colposcopía, con la ayuda de la programación de algoritmos que discriminen en dos clases y fundamentados en conocimientos estadísticos y matemáticos.

## Metodología

### Desarrollo de los algoritmos en Matlab

Sobre nuestro conjunto de datos se dispone de 3 distintas tablas para las modalidades Green, Hinselmann y Schiller, que se modificaron para tener un total de 63 columnas (las conformadas por las variables y el consenso). Así, la tabla de Green consta de un total de 98 x 63 datos numéricos, donde 98 fueron las personas analizadas, y de las cuales 31 se caracterizan por ser de mala calidad y 67 concluyeron ser de buena calidad; para la información de la modalidad Hinselmann se contaba con la información de 97 personas con 15 pertenecientes a la clase de mala calidad y 82 de buena calidad, finalmente, los datos de Schiller corresponden a 92 personas, con 25 correspondientes a los realizados con mala calidad y 67 de buena calidad. El análisis a continuación será realizado sobre los datos de Green, y el proceso será análogo para la demás información; cabe mencionar, que el usuario tendrá la opción de poder seleccionar los datos con los que gustará de trabajar en el documento.

KNN

Para la realización del algoritmo KNN llamaremos a , el vector de prueba que será ingresado por el usuario, donde corresponde al dato relacionado con las variables, , consideremos también dos matrices y las matrices que tienen la información de los correspondientes a realización de colposcopia de mala calidad () y de buena calidad (), respectivamente. Sea y un vector que tiene como entradas cualquiera de las filas de la matriz y respectivamente, , , .

A continuación, se tiene que realizar el cálculo de las distancias de a y , es decir, se debe determinar y para cada , y , luego, se toman en cuenta los k vecinos más cercanos y se cuenta el números de vecinos correspondientes a cada clase y , finalmente es asignado a la clase que haya tenido el mayor número de vecinos cercanos a . Este procedimiento es análogo para las demás distancias.

Para realizar nuestras pruebas, utilizamos la plataforma de Matlab para programar el algoritmo, como se puede ver en la siguiente imagen

Naive Bayes

Tomando las mismas consideraciones que en la sección anterior, buscaremos ahora encontrar el valor mayor entre

y

Para ambos casos

: Probabilidad de ocurrencia de la clase , este se puede calcular como

: Probabilidad de ocurrencia del vector (en la práctica esta es considerada como constante.

.

Para calcular el valor de asumiremos (tomando en cuenta que los datos ingresados son números reales) que cada clase se distribuye según una distribución Normal con media y desviación estándar correspondientes a la columna de la variable con la que esté relacionada, de esta manera

### Desarrollo de la página web

Aunque existen diversos métodos para la realización de un sistema informático, no todos sirven para un determinado sistema sino que se debe optar por el que mejor se adapte al proyecto y ayude al desarrollo del mismo de una manera funcional. El uso de estos métodos debe favorecer al desarrollo del proyecto en distintos aspectos tales como: ahorro de tiempo, costos, control en la planeación y administración. Cada método tiene diferente enfoque, y dichos enfoques se basan en distintos modelos metodológicos como lo son: modelo cascada, modelo de prototipado, modelo Incremental y modelo espiral, sea cual sea el que se elija debe llevar al mismo resultado. Por ejemplo en el modelo espiral, cada bucle o iteración representa un conjunto de actividades, dichas actividades deben reflejar los requerimientos de la página y dejar soporte para futuras aclaraciones, por estas y otras razones, se eligió el modelo espiral (véase imagen 1), para este proyecto ya que se trata de un sistema amplio. Este modelo se eligió después de escuchar las observaciones de los integrantes del equipo.

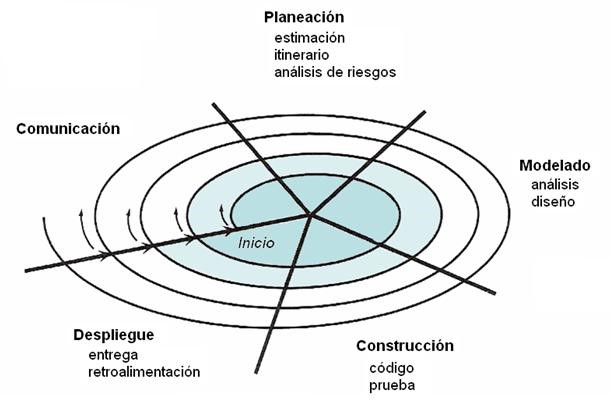


Imagen 1: Modelo de Espiral

Con el fin de mejorar la planeación del desarrollo de la página web, se buscaron alternativas que ayudaran a cumplir con este objetivo.

En la primera fase de nuestro modelo de espiral se definen los objetivos específicos para posteriormente identifica las limitaciones del proceso y del sistema de software, además se diseña una planificación detallada de gestión y se identifican los riesgos.

En la siguiente etapa se realizó el modelando de la página web, precediendo en a la elaboración de estructura de la esta, posteriormente en base a la estructura se crearon bosquejos abstractos para facilitar la construcción de la página.

La etapa de construcción, se enfocó en el desarrollo completo de la página web utilizando los lenguajes de programación correspondientes. Una vez hecho esto continuar con la etapa siguiente, en la ejecución de las pruebas de la página.

Y como el modelo en espiral permite a quien lo desarrolla aplicar el enfoque de construcción de prototipos en cualquier etapa de evolución del producto. Con referencia en la etapa de comunicación, se decidía si se aplicaba algún cambio o se continuaba con la idea principal.

## Resultados

1.-Planeacion

Se obtuvieron los requisitos de la página web, los cuales serían útiles para el desarrollo web como la ejecución de las funciones matemáticas establecidas. Algunos por ejemplo

* Algoritmos a utilizar
* Variables que se utilizan
* Funciones necesarias
* Conjuntos de datos

2. Modelado

Tras haber fijado los puntos principales para comenzar el desarrollo se elaboró un esquema tipo árbol, de la estructura de cómo estaría dividida la página web (véase imagen 6).

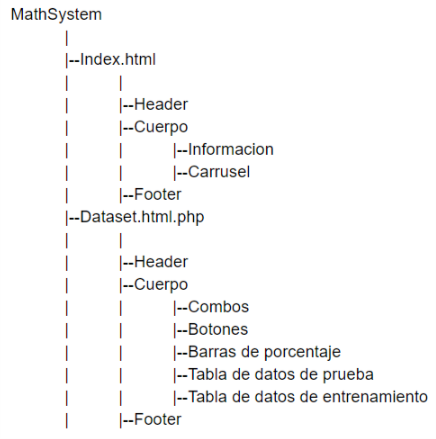


Imagen 6: Estructura de la página

Después se realizaron los bosquejos de las ventanas, con sus respectivos botones, menús, paneles, entre otras cosas. Estos bosquejos fueron fabricados en una herramienta llamada Pencil, esta fue descargada de la página oficial (<https://pencil.evolus.vn>) (véase imagen 7).



Imagen 7: Página oficial de Pencil.

Cuyos bosquejos quedaron de la siguiente manera, cabe mencionar que es una herramienta muy práctica debido a la gran variedad de elementos disponibles (véase imagen 8).

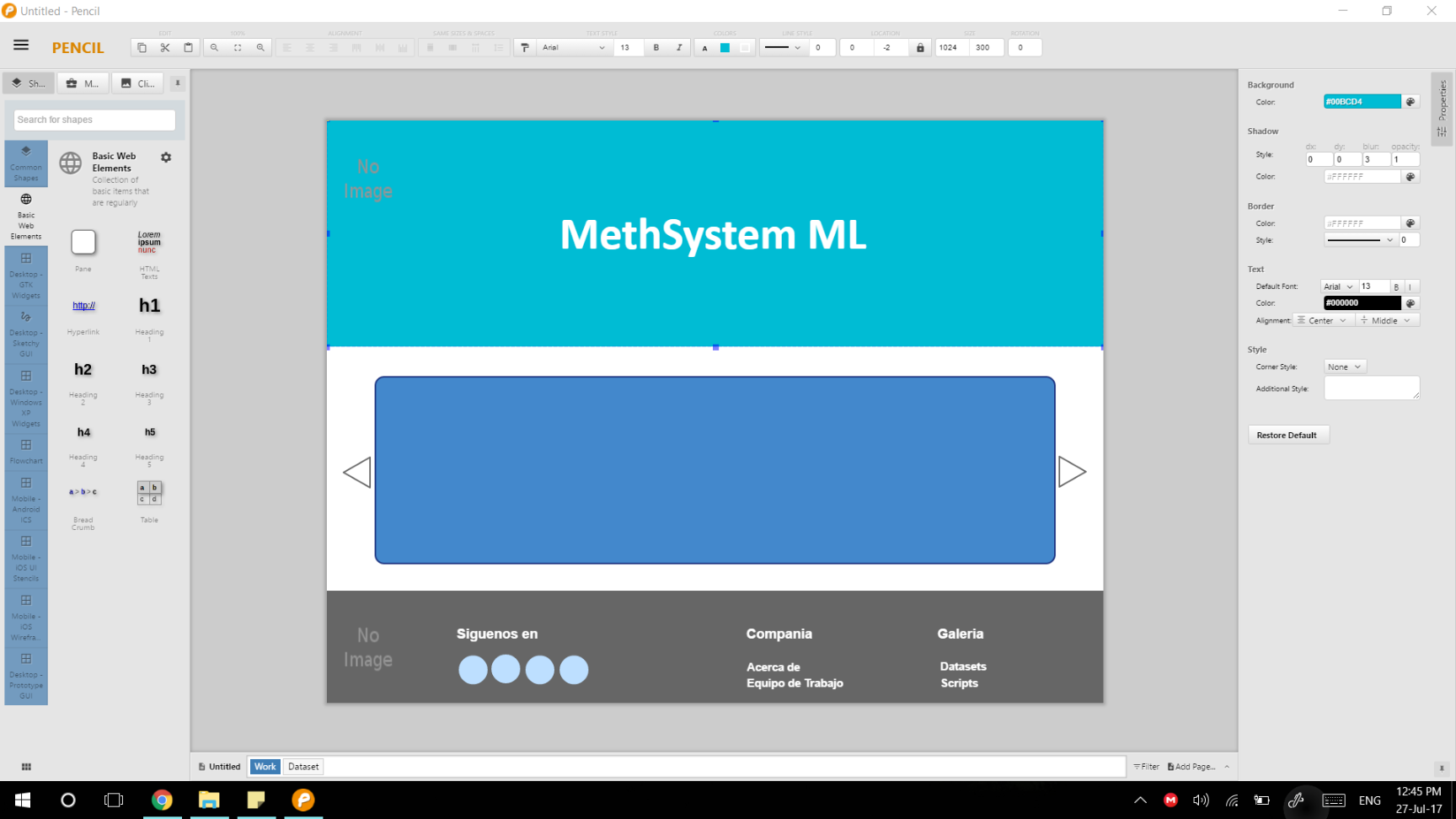


Imagen 8: Bosquejo de la ventana de Index.html

3. Construcción

Para la continuación del proyecto lo siguiente, es utilizar los bosquejos para elaborar las pantallas con HTML5, estilo en cascada (CSS) y JAVASCRIPT (véase imagen 9, 10,11 y 12).

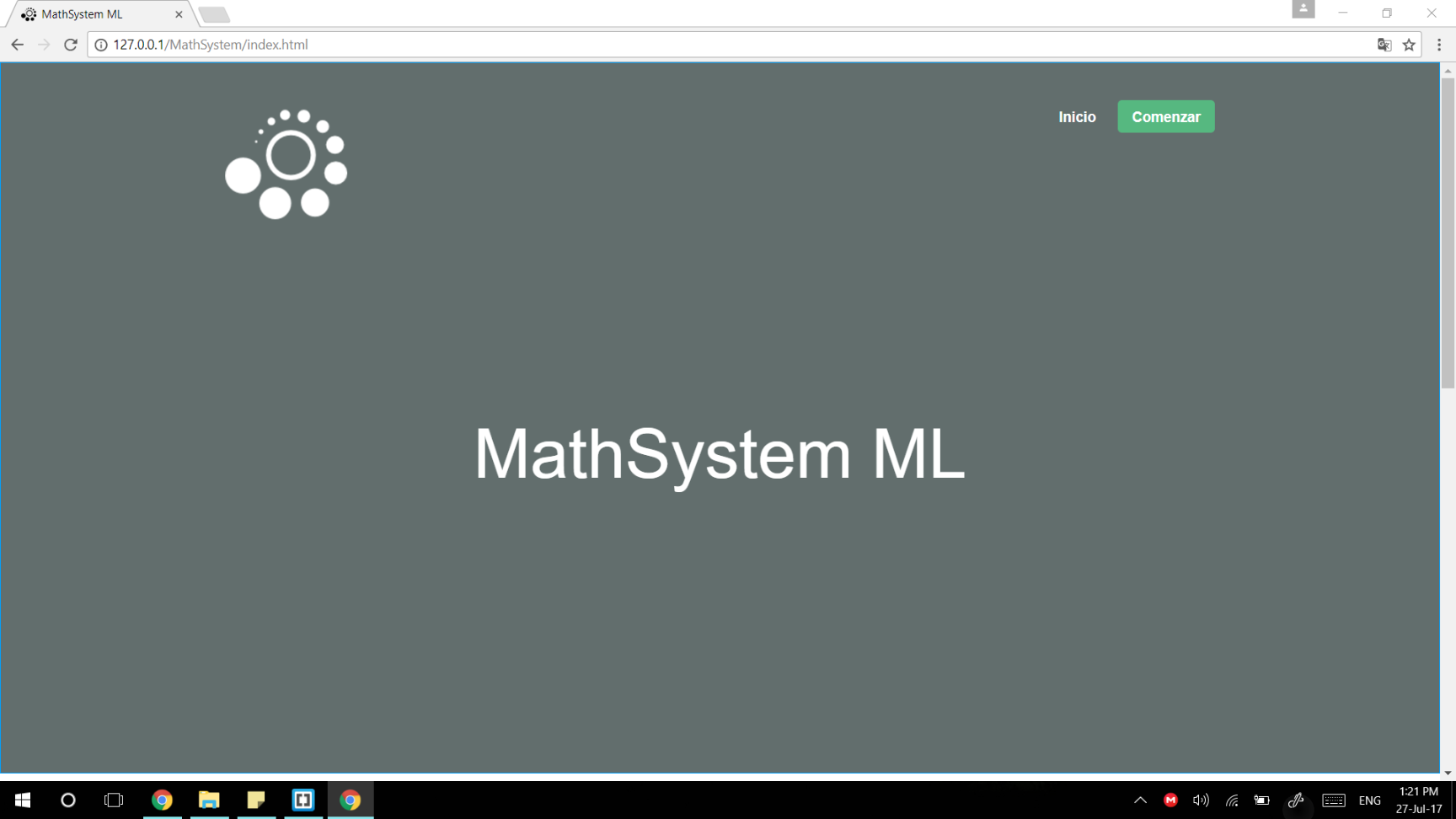


Imagen 9: Header de la ventana principal

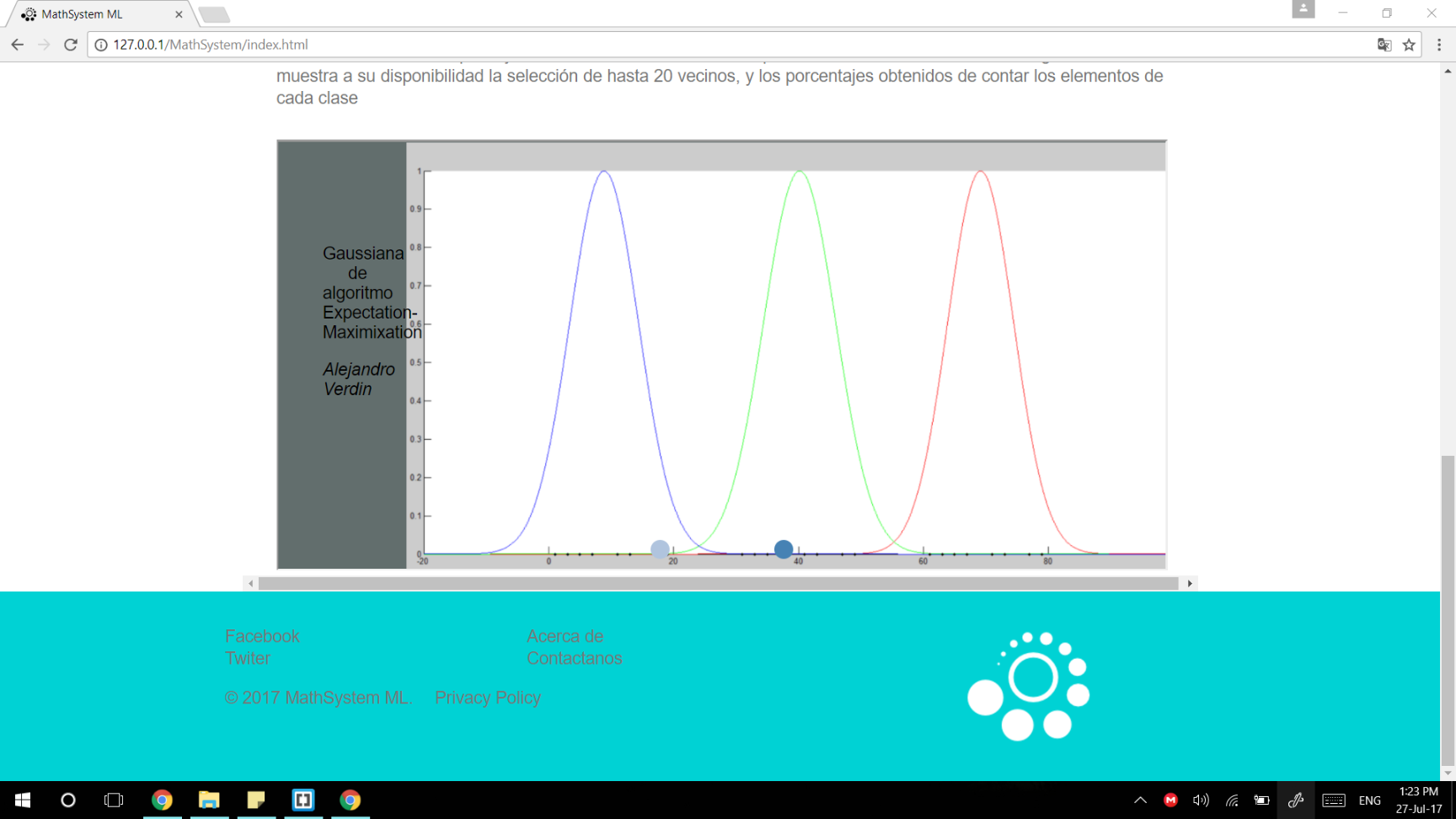


Imagen 11: Carrusel y footer de la página web

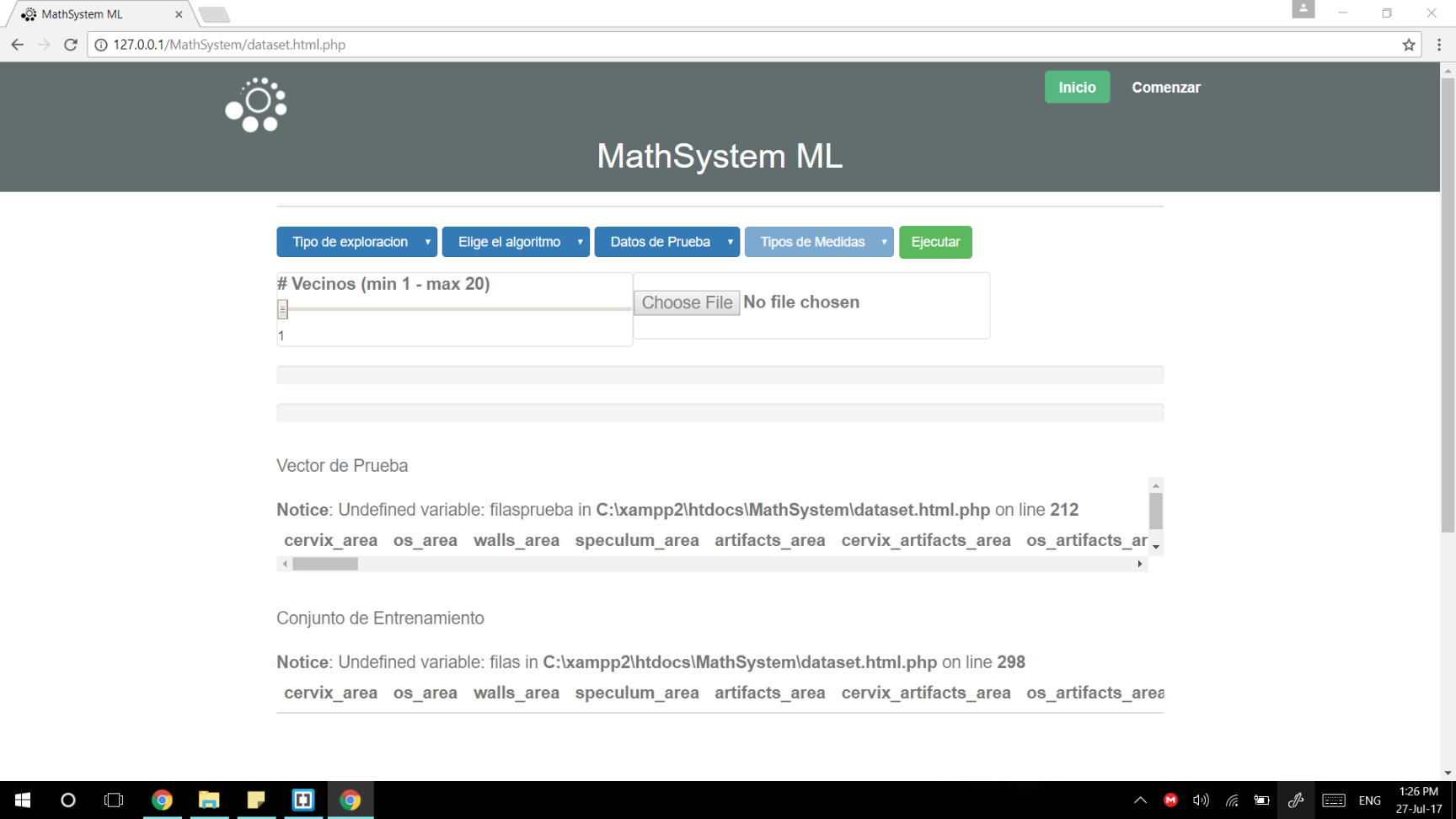


Imagen 12: Pantalla de dataset

Además se realizó una búsqueda de un servidor que sirviera como hosting para el alojamiento de la página web. Se consultaron distintos hosting entre ellos Hostinger y 000webhost (véase imagen 13), de los cuales se eligió 000webhost.

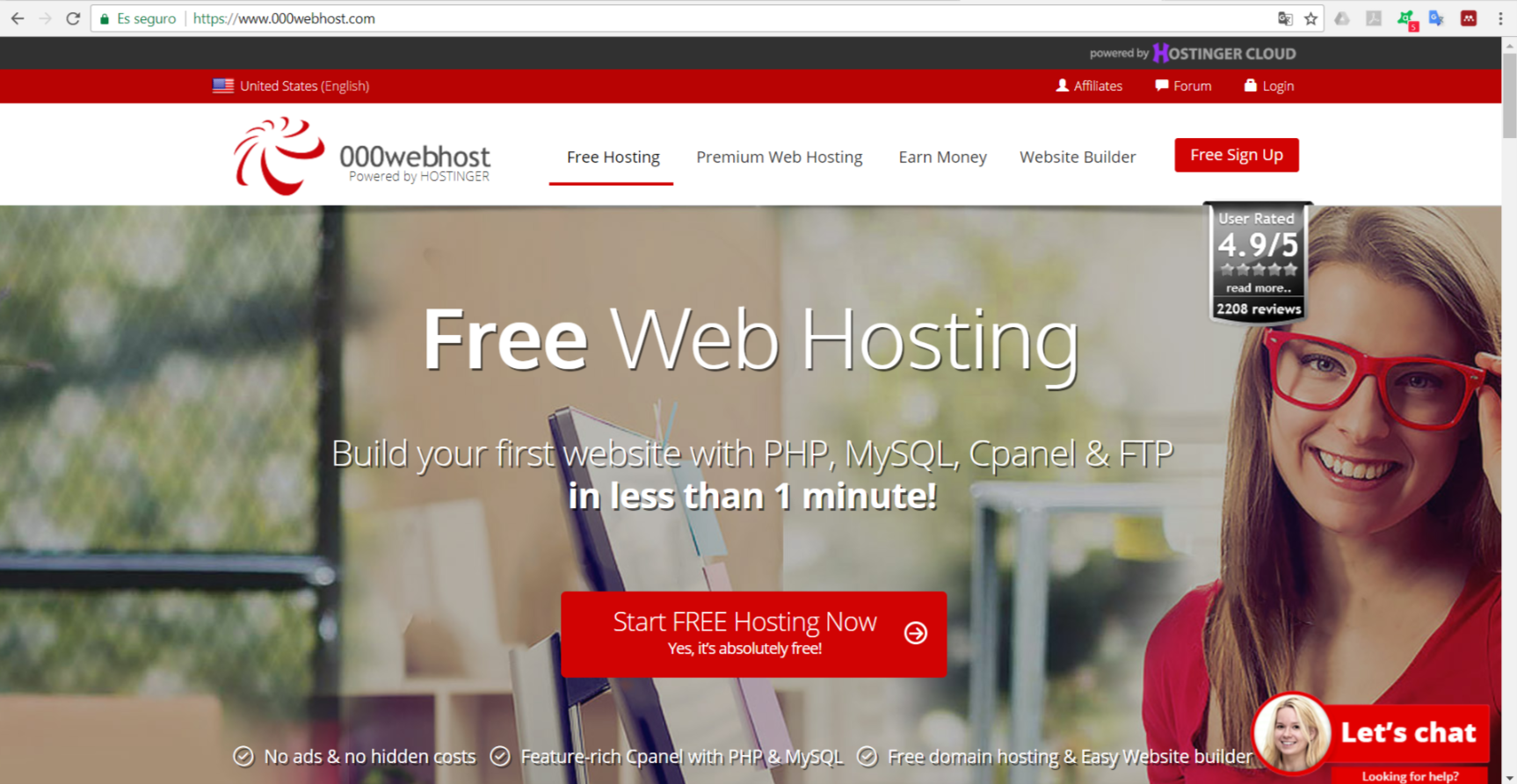


Imagen 13: Pantalla principal de Hostinger

Para el alojamiento de la base de datos y fines prácticos, se instaló un servidor local llamado XAMPP, cuya herramienta cuenta con distintas funciones, entre ellas cuenta con un Gestor de Base de Datos (MySQL), fue aquí donde se montó la base de datos con los conjuntos de datos establecidos.

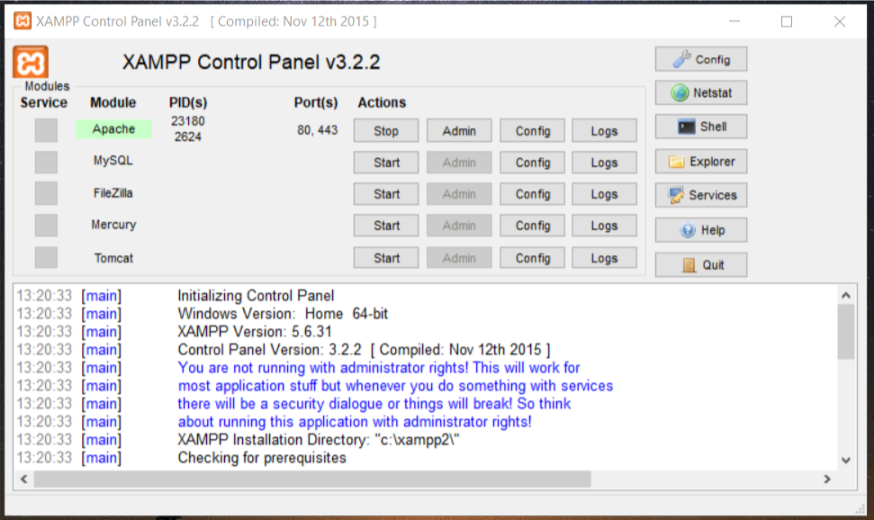


Imagen 14: Interfaz de XAMPP

El Desarrollador tomo todos los elementos gráficos del prototipo, la base de datos y los utilizó para crear un sitio funcional. También se implementarán en esta etapa los elementos interactivos, habilitar y deshabilitar campos. En esta etapa además de usar HTML, CSS y JAVASCRIPT, se utilizó PHP, ya que en este lenguaje de programación se extraen los datos de la base de datos y se realizan las distintas operaciones y algoritmos que maneja la página web.

4. Despliegue

Se probó el funcionamiento completo de la página web, por pruebas segmentadas tomando en cuenta las distintas modalidades, valores, ítems y conjuntos de datos. Dando un satisfactorio resultado en cada una de ellas.

## Discusión

Se desarrolló una página web de calidad, la cual cumple con todas las funciones y requisitos fundamentales para su uso. Con una interfaz agradable al usuario así como también intuitiva.

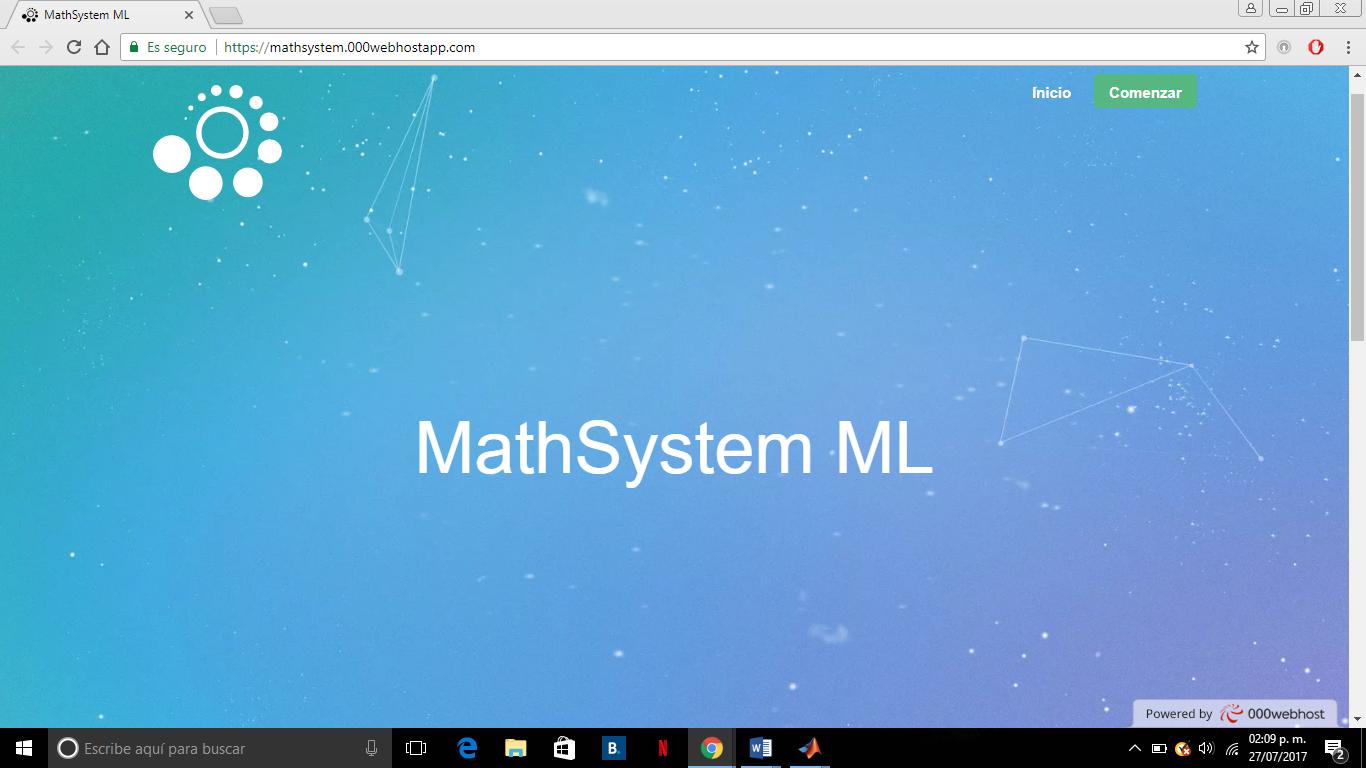


Imagen 15: Pantalla principal de la Página Web

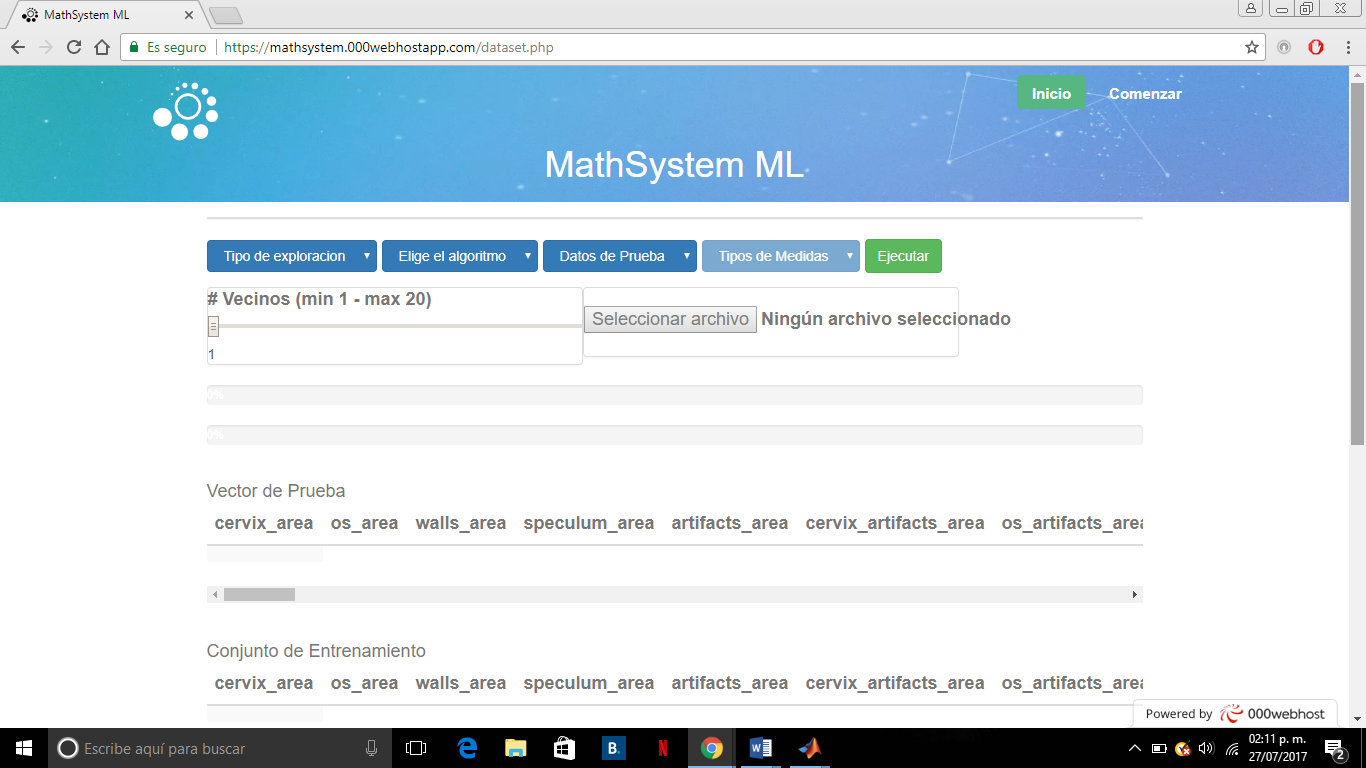


Imagen 16: Pantalla del contenido de la Página Web

<https://mathsystem.000webhostapp.com/>

## Conclusiones

El aprendizaje automático prácticamente cuenta con tantas aplicaciones como imaginemos, pudiéndose adaptar a tantas situaciones como datos con los que contemos.

Motores de búsqueda, diagnósticos médicos, reconocimiento del habla y del lenguaje, robótica… Entre otras.

El Machine Learning resulta especialmente efectivo en problemas de naturaleza compleja en los que la aplicación de algoritmos ayuda a la obtención de soluciones precisas y, claro está, con el consecuente ahorro de tiempo que este método implica.

Por medio del trabajo presentado, se puede concluir que las matemáticas y los sistemas computacionales van tomados de la mano, que cooperan para que el análisis y la aplicación hacia las ciencias lleve consigo el manejo de la inteligencia artificial en los ámbitos conocidos y por saber.

Gracias a la inteligencia artificial, por ejemplo, se ha logrado que una máquina sea capaz de desarrollar áreas de conocimiento muy específicas y complicadas, haciendo que la máquina pueda simular procesos que el hombre realiza, como el autoaprendizaje.

Pero no podemos olvidar que el desarrollo de estas tecnologías no pretende reemplazar al ser humano sino que tratan de mejorar el estilo de vida de la especie dominante en la tierra.

## Literatura Citada

¿Qué es MySQL? - Definición en WhatIs.com. (n.d.). Retrieved July 27, 2017, from http://searchdatacenter.techtarget.com/es/definicion/MySQL

Larrañaga, P. (n.d.). Tema 6. Clasificadores Bayesianos. Retrieved from http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t6bayesianos.pdf

Ludwig Ovalle, D., Rosales Vice ministro Técnico Rafael Haeussler Jefe de Dirección Regulación de Programas de atención las Personas -DRPAP Dra Miriam Bathancourt Coordinadora Programa Nacional de Salud Reproductiva -PNSR Erick Jacobo Alvarez Rodas, P., Jacobo Alvarez -Rodas, E., Equipo técnico OPS, R., & Guatemala, O. (n.d.). Material con propósitos de capacitación. Retrieved from http://www.paho.org/gut/index.php?option=com\_docman&view=download&category\_slug=temas-de-salud&alias=225-manual-de-la-clinica-de-deteccion-temprana-colposcopia&Itemid=518

MathWorks. (n.d.). MATLAB - El lenguaje del cálculo técnico. Retrieved July 27, 2017, from https://es.mathworks.com/products/matlab.html

medical-dictionary. (n.d.). Green-light colposcopy | definition of green-light colposcopy by Medical dictionary. Retrieved July 27, 2017, from http://medical-dictionary.thefreedictionary.com/green-light+colposcopy

Montes, M. C. (n.d.). K-Nearest Neighbors - Graficas, estadistica y minerra de datos con python. Retrieved from http://wwwae.ciemat.es/~cardenas/docs/curso\_MD/knn.pdf

php Group. (n.d.). PHP: ¿Qué es PHP? - Manual. Retrieved July 27, 2017, from http://php.net/manual/es/intro-whatis.php

Transformación de Descubrimientos en Salud. (n.d.). Definición de colposcopia - Diccionario de cáncer - National Cancer Institute. Retrieved July 27, 2017, from https://www.cancer.gov/espanol/publicaciones/diccionario?cdrid=45989

Universitat Oberta de Catalunya. Estudios de Ciencias de la Información y de la Comunicación. (n.d.). *COMEeIN revista dels Estudis de Ciències de la Informació i de la Comunicació*. Estudis de Ciències de la Infromació i de la Comunicació de la UOC i Publicacions a Internet (Àrea de Comunicació). Retrieved from http://comein.uoc.edu/divulgacio/comein/es/numero36/articles/Article-Josep-Cobarsi.html