**Modelación de Problemas Reales, Reconocimiento Estadístico de Patrones e Inteligencia Artificial**

Félix Covarrubias Héctor Alionzo1, Limas Pérez Luis2

1,2Universidad Autónoma de Nayarit

1hector.felix@uan.edu.mx, 2 luislimasperez@hotmail.com

**Asesores:**

Mtra. Sandra Marcela Munguía Gutiérrez.

Dr. Ignacio de Jesús Segovia Domínguez

**Resumen**

El desarrollo tecnológico actual conduce a cualquier ciencia a realizar uso de esta, ya que mediante herramientas desarrolladas se produce de manera eficaz y eficiente el trabajo en el que se emplee, siendo así que en el área de la medicina se presentan comúnmente un sin fin de problemas, dentro de los que destaca el “Parkinson”, la cual es un trastorno caracterizado por síntomas motores como temblor en reposo, bradicinesia, rigidez y alteraciones en la marcha. Gracias a la tecnología es que se pueden y deben desarrollar aplicaciones que funcionen como herramienta médica, es por ello que la siguiente investigación indagó acerca del padecimiento del Parkinson, obteniendo como producto final un clasificador web, el cual permite por medio de cálculos matemáticos y algoritmos de clasificación de inteligencia artificial realizar una aproximación al padecimiento, por lo que el médico especialista podrá diagnosticar el padecimiento y de esta manera contar con un respaldo de los datos del paciente y así ser accedidos por este mismo, gracias a ello se estará involucrando más con los datos acerca de su mismo padecimiento

**Palabras Clave:** Clasificador, Parkinson, KNN, Perceptrón Multicapa.

**Introducción.**

“*Ninguna investigación humana puede ser llamada ciencia real si no puede demostrarse matemáticamente” (Leonardo da Vinci, año).*

Se conoce como enfermedad degenerativa a un desequilibrio en los mecanismos de regeneración, que en realidad no se debe fundamentalmente a factores psicosomáticos o bien físicos externos que ocasionan una falta de regeneración (aplasia) o un exceso descontrolado de regeneración (neoplasia), si no que se originan por la alteración anatómica y funcional de los tejidos de cualquier órgano, aparato o sistema del organismo (Secretaria de Salud, 2008). Como ejemplos de este tipo de enfermedades es el infarto del miocardio, que se debe a una falta de riego y por tanto del oxígeno en una porción del corazón por obstrucción de una arteria coronaria, también el mal de Parkinson y el Alzheimer, que afectan al sistema nervioso central. Los síntomas son variados, como vértigo y la fibra muscular que va adelgazando mientras que el cuerpo anatómicamente pierde ciertas estructuras óseas: como aumento de campo craneal lo que afectaría el sistema neurológico, hasta invalidez (Secretaria de Salud, 2008).

Según el Centro Nacional de Excelencia Tecnológica en Salud (CENETEC) define la enfermedad de Parkinson como una degenerativa del sistema nervioso central caracterizada por pérdida neuronal, que ocasiona una disminución en la disponibilidad cerebral del neurotransmisor denominado dopamina; esto, a su vez, se manifiesta como una desregulación en el control de movimiento (CENETEC, 2010). Se han realizado diferentes estudios con respecto a esta enfermedad los cuales arrojan que el Parkinson es el padecimiento más común de inicio en el adulto; siendo la segunda más frecuente después de la demencia y Alzheimer.

Se estima que en México se presentan entre 40 a 50 casos por cada 100,000 habitantes/año. De igual modo se calculó que, en el mundo, debido al aumento de la tasa de sobrevida, la enfermedad de Parkinson afecta de 4.1 a 4.6 millones de personas mayores a 50 años. Para el año 2030 se espera que esta cifra sea duplicada y esté superada con una facilidad, provocando un problema de salud inverosímil, debido a que la enfermedad es progresiva con una posibilidad de duración media de 10 a 13 años (Salud, 2015).

Los autores Hughes, Daniel, Ben‐Shlomo, & Lees (2002) afirman que al realizar un diagnóstico presenta una tasa de error del 24% al 56% cuando es realizado por los médicos generales, con lo antes mencionado, se observa que sin una medida de control ante este padecimiento podrá volcarse de manera abrumadora para los años siguientes, es por esto, el interés de contar con un clasificador web como herramienta para el médico especialista, en el cual se pueda apoyar de manera confiable, claro está basándose en las medidas tomadas con anterioridad al paciente y de esa manera tener un mayor respaldo para tomar un diagnostico más acertado y con menos incertidumbre, a su vez conllevaría una mejoría y evolución de los pacientes, y no solo a ellos sino que de esta manera tendrá un control de datos y mantener o mejorar su estado de salud.

**Marco teórico**

La asociación de Parkinson en Madrid (2016), define el Parkinson como una enfermedad degenerativa producida por la muerte de neuronas de la sustancia negra, que produce dopamina, la cual es un neurotransmisor importante en el circuito de los ganglios basales, cuya función primordial es el correcto control de los movimientos; cuando hay una marcada reducción del nivel de dopamina, se altera la información de movimientos e inestabilidad postural, entre otros síntomas. El desarrollo tecnológico actual conduce a cualquier ciencia a realizar uso de esta, ya que mediante herramientas desarrolladas se produce de manera eficaz y eficiente el trabajo en el que se emplee, siendo así que en el área de la medicina se presentan comúnmente un sin fin de problemas, dentro de los que destaca el “Parkinson”.

Es necesario desarrollar una aplicación que funcione como herramienta médica, por ello que la siguiente investigación indagó acerca del padecimiento, obteniendo como producto final un clasificador web, el cual permite por medio de cálculos matemáticos y algoritmos de clasificación de inteligencia artificial realizar una aproximación al padecimiento, por lo que el médico contará con ella como medio de ayuda para diagnosticar el padecimiento, y de esta manera podrá contar con un respaldo de los datos del paciente y así ser accedidos por este mismo, gracias a ello se estará involucrando más con los datos acerca de su mismo padecimiento.

Las causas principales causantes de la enfermedad de Parkinson tienen un comienzo insidioso, y sus características esenciales son; temblor, rigidez, bradicinesia y alteración en los reflejos posturales. Existe un par de síntomas secundarios los cuales no afectan a todos los enfermos, sin embargo, provocan trastornos importantes, empeorando síntomas principales y agravando las condiciones físicas y psicológicas del paciente: problemas urinarios, constipación, trastornos del sueño, disminución del olfato, entre otros (Secretaría de Salud, 2008).

Síntomas no menos importantes como la depresión y ansiedad son presentados de manera frecuente en la enfermedad del Parkinson, estos afectan al 40% de los pacientes. Como características principales que presentan la depresión son: estados bajos de ánimo y pérdida de interés (anhedonia) entre una larga lista. Cabe destacar que al comenzar el tratamiento la enfermedad no disminuye y mucho menos desaparecerá, ya que el objetivo de los fármacos es simplemente aliviar los síntomas, y efectos secundarios derivados de los mismos fármacos utilizados para combatirla. Una vez que realizas el tratamiento se debe manejar adecuadamente ya que en personas avanzadas una modificación en la dosis sería garrafal al punto que se presenta un empeoramiento del paciente.

Mediante la realización de búsquedas y el resultado de algunas otras investigaciones anteriores, se encontraron más de 125 aplicaciones, de las cuales 56 se clasificaron con potencialidad, considerando 69 con diseños específicos para la enfermedad de Parkinson, siendo 23 apps de contenido solo informativo, 29 de valoración, 13 de acerca de cómo llevar a cabo el tratamiento y solo 4 que abordan una valoración general acerca del tratamiento (Universidad Politécnica de Madrid, 2017). De esta manera se percató que existen un gran número de aplicaciones móviles con potencial utilidad y diseño; sin embargo; puntualizando algunas de estas herramientas cabe mencionar un monitorizador de pacientes mediante sensores vestibles y una plataforma web realizada por la Universidad Politécnica de Madrid, la cual cuenta con datos recabados por medio de pruebas a más de 100 pacientes con el padecimiento (Linares-del Rey, Vela-Desojo, & Cano-de la Cuerda, 2017)⁠, se plantea poder realizar algún acercamiento con ellos para poder intercambiar datos y de manera conjunta proponer una herramienta potente y confiable a la comunidad de salud.

**Objetivo**

El presente proyecto, tiene un enfoque de aprendizaje general, de manera que involucra diversidad en áreas del conocimiento, tales como: Sistemas Computacionales y Matemáticas, planteando realizar un clasificador web, en el cual de forma general se pueda clasificar el padecimiento de Parkinson, teniendo como objetivo principal: *Desarrollar un clasificador utilizando datos de ciencias de la salud, mediante la teoría del reconocimiento estadístico de patrones en torno al padecimiento Parkinson.*

Se establecieron los siguientes objetivos específicos:

1. Elaborar ejemplos-script reconocimiento de patrones en Matlab.
2. Construir un clasificador mediante Matlab.
3. Diseñar y crear la Base de Datos con el conjunto de datos a utilizar.
4. Elaborar un diseño de la aplicación.
5. Realizar modelado, codificación y pruebas de la aplicación.

Para lograrlo se utilizó un conjunto de datos, los cuales fueron obtenidos de UCI (Machine Learning Repository)(“UCI Machine Learning Repository: About,” 2007), el cual contiene una diversidad de archivos, y ellos servirán de gran ayuda en el desarrollo del proyecto. En el contenido se adjuntan dos Dataset, uno de características test data, incluyendo un total de 1050 row, con un total de 28 variables, conteniendo información de vital importancia para la clasificación de Parkinson, el otro Dataset, train data, comprende información que fue obtenida sobre pacientes que contaban con este padecimiento con un total de 520, y el resto 520 que no cuentan con el padecimiento. En general se plantea que, en el conjunto de datos train data su última columna clasification, contiene el resultado de padecimientos 0, no se cuenta con el padecimiento y 1, se cuenta con el padecimiento de Parkinson. El proyecto tendrá como enfoque realizar un clasificador web donde se pueda catalogar mediante cálculos matemáticos si enfrenta o no la enfermedad.

Todo lo anterior se realizará con la ayuda de las herramientas que se aprovisionan por el área de conocimiento Matemático en conjunto con el apoyo de los Sistemas de Computación, siendo este un soporte en el desarrollo de interfaces y procesamiento matemático como la base para hacer la conjunción final y obtener el resultado esperado.

**Metodología y resultados**

Para lograr el objetivo del proyecto, se plantearon y desarrollo 4 etapas, que son:

1. Indagación de los requerimientos.
2. Modelo del análisis, Modelo del diseño, Descripción de Sistema (Modelo de requerimientos).
3. Codificación del clasificador web.
4. Pruebas y regulación del clasificador web.

En la *indagación de requerimientos* se realizó la recolección de datos acerca del padecimiento Parkinson, recabados de la plataforma UCI, la cual cabe destacar que es un repositorio de Machine Learning, un Centro de Aprendizaje Automático y Sistemas Inteligentes el cual consta de una colección de datos: estos pueden ser utilizados para diversas finalidades, en esta ocasión se utilizaron como base para poder realizar un análisis detallado de la enfermedad del Parkinson y poder determinar si un individuo tiene o no dicha enfermedad.

De igual manera se utilizó el algoritmo K-NN como primer clasificador, su funcionamiento se basa en la obtención de las distancias que existen entre el dato de prueba con los de entrenamiento como se mencionó anteriormente; entre las cuales se puede elegir la distancia euclidiana, la Manhattan y Chebyshev permitiendo así observar si los resultados se veían afectados al cambiar de distancia o se mantenían igual.

A partir de la base de datos especificada se crearon 2 nuevas bases donde se dividieron a las personas que padecían la enfermedad y otra base donde estaban las personas que no tenían la enfermedad, fue así como se obtuvieron 2 conjuntos de entrenamiento. Después de contar con los datos de entrenamiento se crearon datos de prueba que también fueron obtenidos de la plataforma UCI con la intención de poder comparar ambos datos ya que nuestra finalidad fue crear un algoritmo de clasificación dentro de nuestra página web que nos permitiera predecir si el paciente de prueba que se ingresara en la página web era una persona que padecía de Parkinson o no.

Como parte de la segunda etapa de la metodología y para realizar una indagación más específica y concreta en cuanto a requerimientos, se eligió el *Modelo de requerimientos*, el cual permitió hacer la recolección de manera más organizada y eficiente, ya que como producto final se tendría un clasificador web por lo tanto existe un puente entre la descripción del sistema y el modelo del diseño, es decir sin una buena exploración de requerimientos lo posterior se generaría de manera errónea causando un conjunto de problema, retrasos y mayor trabajo para realizar la próxima etapa. Una vez que se tuvo la Descripción del sistema se procedió a realizar una *Modelo del análisis*, muy básico pero funcional, en el cual se pudieron identificar anomalías en el proceso anterior y así corregirlas antes de tiempo, finalmente se realizó un modelo del diseño en el cual se exploró hasta no encontrar error alguno que pudiese interferir en el proceso de las posteriores actividades.

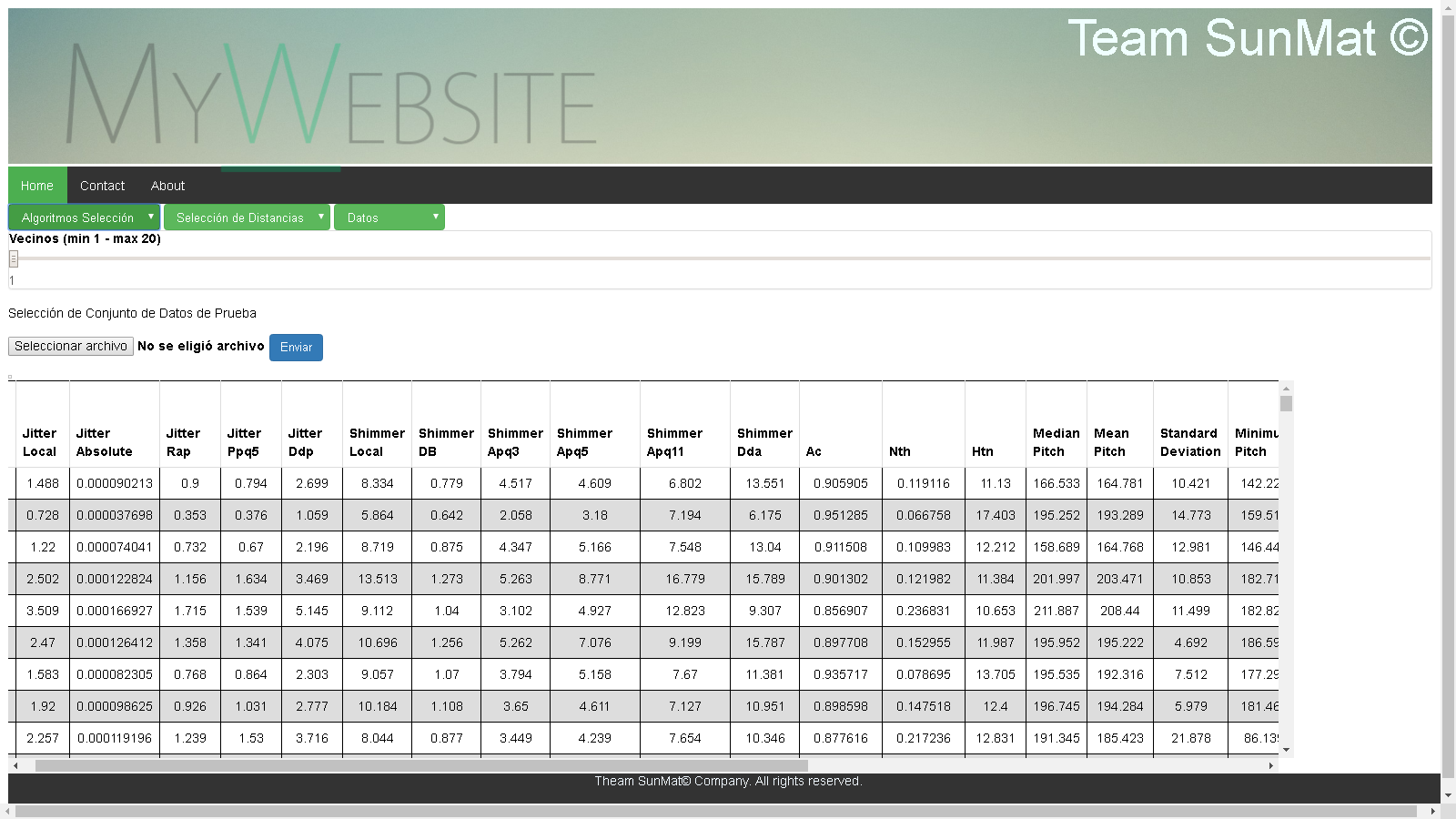
Comprendiendo la tercera etapa *Codificación del clasificador web* se comenzó por elegir un modelo de desarrollo que permitiera realizar el producto final desde principio hasta la última actualización de esta, es por ello que se elige el **modelo incremental**, ya que consiste en la creación de incrementos o módulos, mediante el primer incremento se ejecuta una serie de avances, y a su vez se generan nuevos incrementos, que en forma progresiva dan más funcionalidad al proyecto conforme se entrega cada módulo. En general el modelo incremental combina elementos de los flujos de proceso lineal y paralelo. Es así como en este modelo se aplica secuencias lineales en forma escalonada a medida que avanza el calendario de actividades, cada secuencia lineal produce incrementos de software susceptibles de entregarse de manera parecida a los incrementos producidos en un flujo de proceso evolutivo.

Elegir el modelo incremental, permite crear módulos o incrementos en cada una de las etapas, lo que permitió que los primeros incrementos simplemente estuvieran basados en los requerimientos básicos, es decir el primero solo se obtuvo un producto fundamental o medular para la conectividad del proyecto.

Como cuarta etapa: Pruebas y regulación del clasificador web está comprendida en la etapa tres, ya que gracias al modelo incremental se permitió ir mejorando cada entregable que se desarrolla del proyecto, es decir en cada nuevo entregable se realizaban pruebas al sistema en general para encontrar fallos y estos mejorarlos en el próximo incremento, gracias a esto al realizar la cuarta etapa se presentó de manera asíncrona y paralela, lo que permitió una optimización en los tiempos de trabajo y a su vez en el esfuerzo por parte de las personas involucradas en el proyecto.

Realizar un proyecto de cualquier índole, incluye indagación de requerimientos básicos para poder subsistir y funcionar de manera correcta, posteriormente se continúa con el proceso de efectuar reuniones, elaborar listas de funciones y clases que serán utilizadas en todo el proyecto, así mismo construyendo listas de actividades relacionadas entre una y otra, así se opta en crear prioridades formales, dándole importancia de manera formal a los requerimientos. Es así como realizar una indagación de requerimientos pertinente ayuda en el ahorro de tiempo y esfuerzo del personal que lo están desarrollando, todo esto gracias a los modelos del proceso de software que ayudan a minimizar todos estos detalles.

Al contar principalmente con un conjunto de requerimientos tan extensos, completos y fidedignos, se obtuvo como base medular de la investigación una base de datos (BD) en la cual se incluyen los datos descargados de la plataforma UCI, la cual permite poder realizar cálculos mediante PHP, fórmulas matemáticas y algoritmos de inteligencia artificial. De esta manera que se concluyó con la elaboración de la BD, posteriormente se obtuvo como producto final y objetivo general, el clasificador web (Figura 1), el cual permitirá funcionar de manera paralela con el médico especialista y así generar un diagnóstico confiable y pertinente, mediante los datos que se le introduzcan al sistema. Cabe destacar que el clasificador aún no está libre en la red, para realizar pruebas solo se estableció en un servidor local en el equipo del programador.



**Figura 1**. Home, página donde se realizan las operaciones del Clasificador web TEAMSUNMAT.

**Conclusión**

Cabe destacar que el proyecto fue terminado con éxito, gracias al conocimiento creado en conjunto por las áreas de estudio Matemáticas y Sistemas, las cuales en unión pudieron obtener el potencial y obtener un resultado satisfactorio, todo este potencial no hubiese podido lograr sin la ayuda del Dr. y la Mtra. Sandra Munguía, los cuales brindaron ese apoyo y conocimiento acerca de lo que se quería lograr y las dudas que se presentaban a lo largo del proyecto.

En general se pueden utilizar otros algoritmos de clasificación, pero el KNN en específico está dando resultados favorables ya que las pruebas que se realizaron tuvieron una efectividad en la clasificación de un 90%, es decir solo cuenta con un 10% de error, el cual en el transcurso de mediciones y pruebas no presento inconveniente. Se implementaron diversas maneras de trabajar en Matlab una de ellas fue practicar los algoritmos primeramente en esta plataforma para después hacer lo mismo en la página web.

La elaboración de una clasificación del área de estudio en una serie de clases relativas a paisajes, tipos de suelo, texturas, etc.; es uno de los objetivos fundamentales de un clasificador de imágenes. La clasificación es, por tanto, un caso particular del problema general de clasificar N individuos en un conjunto de K clases en función de una serie de variables cuantitativas.

Para poder resolver este problema se necesita una medida de la semejanza o diferencia entre los diferentes individuos y entre los individuos y las clases. Dos individuos muy parecidos pertenecerán probablemente a la misma clase, mientras que dos individuos distintos pertenecerán a diferentes clases. La medida más utilizada es la distancia euclidiana, aunque pueden utilizarse otro tipo de distancias. se puede potencializar qué es una herramienta que no suple en ningún momento la observación y diagnóstico de algún médico, sino que todo lo contrario viene a pasar como una herramienta más que puede ser utilizada por especialistas para poder realizar de una manera eficaz su trabajo, ya que se podrá introducir el conjunto de datos y mediciones que ha realizado al paciente.

Es por ello que los resultados fueron favorables, ya que se concluyó de manera satisfactoria con el proyecto presentado un producto final funcional, ya que en las diferentes ejecuciones del proyecto se pudo constatar que en todos los casos realizó una clasificación acertada, es decir que al menos en todos los casos probados acertó un 100% a todos ellos.

**Referencias**

CENETEC. (2010). GUÍA DE PRÁCTICA CLÍNICA. <http://www.cenetec.salud.gob.mx/descargas/gpc/CatalogoMaestro/305_SSA_10_PARKINSON_3ER_NVL/GRR_Parkinson.pdf>.

Hughes, A. J., Daniel, S. E., Ben‐Shlomo, Y., & Lees, A. J. (2002). The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain*, *125*(4), 861–870. <https://doi.org/10.1093/brain/awf080>.

Linares-del Rey, M., Vela-Desojo, L., & Cano-de la Cuerda, R. (2017). Aplicaciones móviles en la enfermedad de Parkinson: una revisión sistemática. *Neurología*. <https://doi.org/10.1016/j.nrl.2017.03.006>.

Salud, S. de S. (2015). GUÍA DE PRÁCTICA CLÍNICA, 96. [www.cenetec.salud.gob.mx](http://www.cenetec.salud.gob.mx/).

Secretaria de Salud. (2008). Diagnóstico y tratamiento de la enfermedad de Parkinson inicial en el primer nivel de atención. [www.cenetec.salud.gob.mx](http://www.cenetec.salud.gob.mx/).

UCI Machine Learning Repository: About. (2007). Retrieved September 5, 2017, <https://archive.ics.uci.edu/ml/about.html>

Universidad Politécnica de Madrid. (2017). PARKINSON-MONITOR. Monitorización y seguimiento de pacientes con la Enfermedad de Parkinson y Parkinsonismos. Retrieved September 5, 2017, http://www.upm.es/observatorio/vi/index.jsp?pageac=innovacion/articulo.jsp&id\_articulo=151&id\_tipo\_articulo=3.