

**UNIVERSIDAD POLITÉCNICA SALESIANA  
SEDE QUITO**

**CARRERA:  
INGENIERÍA DE SISTEMAS**

**Trabajo de titulación previo a la obtención del título de:  
INGENIERA DE SISTEMAS**

**TEMA:  
DESARROLLO DE UN APLICATIVO PARA RECOLECCIÓN DE DATOS  
PARA DIAGNÓSTICO AUTOMÁTICO DE DESÓRDENES MENTALES  
(SIGUIENDO LA CLASIFICACIÓN CIE-10)**

**AUTORA:  
DIANA CAROLINA CHAUCA CHANGOLUISA**

**TUTORA:  
PAULINA ADRIANA MORILLO ALCÍVAR**

**Quito, febrero de 2019**

## **CESIÓN DE DERECHOS DE AUTOR**

Yo, DIANA CAROLINA CHAUCA CHANGOLUISA, con documento de identificación N° 1725395196 manifiesto mi voluntad y cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autora del trabajo de titulación con el tema: “DESARROLLO DE UN APLICATIVO PARA RECOLECCIÓN DE DATOS PARA DIAGNÓSTICO AUTOMÁTICO DE DESÓRDENES MENTALES (SIGUIENDO LA CLASIFICACIÓN CIE-10)”, mismo que ha sido desarrollado para optar por el título de: INGENIERA DE SISTEMAS, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente. En aplicación a lo determinado en la Ley de Propiedad Intelectual, en mi condición de autora me reservo los derechos morales de la obra antes citada. En concordancia, suscribo este documento en el momento que hago la entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Politécnica Salesiana.

Quito, febrero de 2019.



DIANA CAROLINA CHAUCA CHANGOLUISA

CI: 1725395196

## DECLARATORIA DE COAUTORÍA DE LA TUTORA

Yo, PAULINA ADRIANA MORILLO ALCÍVAR, declaro que bajo mi dirección y asesoría fue desarrollado el artículo académico, con el tema: “DESARROLLO DE UN APLICATIVO PARA RECOLECCIÓN DE DATOS PARA DIAGNÓSTICO AUTOMÁTICO DE DESÓRDENES MENTALES (SIGUIENDO LA CLASIFICACIÓN CIE-10)”, realizado por la estudiante DIANA CAROLINA CHAUCA CHANGOLUISA, obteniendo un producto que cumple con todos los requisitos estipulados por la Universidad Politécnica Salesiana, para ser considerados como trabajo final de titulación.

Quito, febrero de 2019.



PAULINA ADRIANA MORILLO ALCÍVAR

CI: 1715646574

## **DEDICATORIA**

### **A Dios.**

Por haberme dado la capacidad de seguir adelante para lograr mis objetivos, por mantenerme con salud y vida, además de su infinito amor y bondad.

### **A mis Padres**

Por nunca dejarme de apoyar a lo largo de mi vida y darme ánimo en todos los aspectos de mi vida, por sus consejos, sus valores, por toda la motivación constante que me ha permitido ser una persona de bien y salir adelante, pero más que nada, por su amor infinito.

### **A mi Familia**

Por su apoyo constante y el ánimo que siempre me han brindado para seguir adelante y conseguir mis objetivos, por sus consejos que siempre han estado presentes, por su amor.

**Diana Carolina Chauca Changoluisa**

# DESARROLLO DE UN APLICATIVO PARA RECOLECCIÓN DE DATOS PARA DIAGNÓSTICO AUTOMÁTICO DE DESÓRDENES MENTALES (SIGUIENDO LA CLASIFICACIÓN CIE-10)

## DEVELOPMENT OF AN APPLICATION FOR DATA COLLECTION FOR AUTOMATIC DIAGNOSIS OF MENTAL DISORDERS (FOLLOWING THE CIE-10 CLASSIFICATION)

Diana Chauca<sup>1</sup>, Paulina Morillo<sup>2</sup>, Holger Ortega<sup>3</sup>

### Resumen

En este artículo se propone el desarrollo de una aplicación web utilizando la metodología *UML-Based Web Engineering* (UWE) y la aplicación del algoritmo de aprendizaje automático kNN, para asistir a los profesionales del área de la salud mental, en el diagnóstico de desórdenes mentales. Esta herramienta, pretende facilitar la tarea de diagnóstico, para superar las limitaciones causadas, principalmente, por la extensa variedad de trastornos mentales según la CIE-10, la ambigüedad de las relaciones entre síntomas y enfermedades, la falta de registro de casos reales que sirvan de referencia frente a casos nuevos, etc. El aplicativo web servirá como plataforma para la digitalización y la adquisición de datos de pacientes que sufren algún tipo de trastorno mental, estos datos servirán a su vez para entrenar un primer modelo de aprendizaje automático, cuyo fin será conseguir el diagnóstico automático de

### Abstract

This paper aims to the development of a web application using the UML-based web engineering methodology (UWE) and the application of a machine learning algorithm kNN to assist professionals in the area of Mental health, for the diagnosis of mental disorders. This tool pretends to facilitate the diagnosis of a patient, and overcome the limitations mainly caused for the extensive variety of mental disorders, according to the ICD-10; the ambiguity of the relationships between symptoms and diseases, the lack of registers of real cases as a reference for new cases, etc. The web application will be used as a platform for the digitalization and acquisition of data from patients that suffer some type of mental disorder, these data will be used for training a first Machine-learning model, whose purpose will be to get the automatic diagnosis of new cases, through the symptomatology of the patient.

---

<sup>1</sup> Carrera de Ingeniería de Sistemas, Universidad Politécnica Salesiana, Quito, Ecuador, dchauca@est.ups.edu.ec.

<sup>2</sup> Universidad Politécnica Salesiana, Grupo de Investigación IDEIAGEOCIA, Quito, Ecuador, pmorillo@ups.edu.ec.

<sup>3</sup> Universidad Politécnica Salesiana, Grupo de Investigación IDEIAGEOCIA, Quito, Ecuador, hortega@ups.edu.ec.

nuevos casos, a través de la sintomatología del paciente. El algoritmo empleado alcanzó un porcentaje de acierto del 44%, debido a la baja cardinalidad del conjunto de entrenamiento y la dispersión de las características de cada instancia.

***Palabras Clave:*** sistema automático de diagnóstico, clasificación multivariante, desórdenes mentales, kNN, aprendizaje automático

The algorithm reached 44% success, due to the low cardinality and the dispersion of the characteristics of each instance.

***keywords:*** automatic diagnostic system, multivariate classification, mental disorders, kNN, Machine Learning

# 1. Introducción

Los trastornos mentales representan uno de los mayores problemas en el área de la Salud Mental Pública, por lo que es indispensable que los profesionales del área realicen un diagnóstico acertado, para poder suministrar tratamientos adecuados. Aunque, la Clasificación Internacional de Enfermedades décima versión (CIE-10) capítulo V rango de códigos F00-F99 brinda una taxonomía detallada de los trastornos mentales y del comportamiento [1], que sirve de guía para lograr un diagnóstico apropiado, esta tarea no siempre resulta sencilla debido, principalmente, a: la gran diversidad de desórdenes mentales [2], la ambigüedad de las relaciones entre síntomas y enfermedades, ya que en muchos casos existen síntomas compartidos con más de un trastorno [3-4] y la falta de documentación de casos previos que sirvan de referencia para nuevos diagnósticos [5].

El aprendizaje de máquina o *Machine Learning* (ML) es una disciplina de la Inteligencia Artificial (IA) que a través del uso de algoritmos ha hecho posible la automatización de diversos procesos y tareas en el campo de los negocios, la ingeniería y la medicina [6]. En el caso puntual de la psicología, el diagnóstico en pacientes con algún tipo de trastorno mental representa un gran desafío, ya que, existen numerosos trastornos y síntomas que dificultan el reconocimiento de un padecimiento específico. Por tanto, se puede aprovechar la capacidad de cálculo y memoria de los computadores, para implementar algoritmos que asistan al médico tratante a dar un diagnóstico confiable, reduciendo el riesgo de prescripciones erradas. [2-4].

Actualmente existen varias aplicaciones que hacen uso de algoritmos de *Machine Learning*, para la descripción, tratamiento y diagnóstico automático de enfermedades

mentales. Sin embargo, estas herramientas se enfocan en un trastorno específico y no toman en cuenta la totalidad de los trastornos descritos en la CIE-10.

En este artículo se propone el desarrollo del aplicativo denominado Psico-Web, que servirá para recolectar datos de casos diagnosticados con algún desorden mental. El aplicativo incluirá la utilización de un algoritmo de clasificación, para identificar de forma automática el tipo de desorden mental que padece un paciente con cierta sintomatología, al ser evaluado por un profesional de la salud mental. La decisión final del diagnóstico seguirá siendo del médico tratante, pero el diagnóstico devuelto por el sistema le brindará una guía que evitará errores y sobrecarga de trabajo mental. El aplicativo cuenta con una interfaz gráfica amigable que permitirá la interacción ergonómica entre el médico tratante y el sistema.

En los siguientes apartados de este artículo se describirán los siguientes puntos: en la Sección 2, se presenta un breve estudio de los trabajos relacionados, incluyendo la descripción de aplicaciones que emplean algoritmos de *Machine Learning* para el diagnóstico de trastornos mentales específicos como la depresión, el alzheimer, la esquizofrenia y los trastornos de personalidad. La Sección 3 muestra la metodología empleada en el desarrollo de la aplicación y se describe el esquema y diseño del aplicativo. También se muestran los antecedentes para la selección del algoritmo de *Machine Learning*, la construcción del *dataset* de entrenamiento a partir de la digitalización de casos reales y la implementación del algoritmo kNN aplicado al conjunto de entrenamiento, para realizar el diagnóstico automático de nuevos casos. En la Sección 4 se describen los experimentos realizados y los

resultados. Finalmente, se exponen las conclusiones y perspectivas de este trabajo.

## 2. Trabajos relacionados

El diagnóstico automático de enfermedades mentales, ha tenido un gran desarrollo en los últimos años debido, principalmente, al surgimiento de nuevas tecnologías y equipos de neuroimagen, que permiten observar el cerebro para intentar explicar su funcionamiento y la relación entre los trastornos mentales y sus síntomas, el escaner fMRI (*funcional Magnetic Resonance Imaging Scan*), se utiliza en [7] para obtener datos de cerebros en reposo de pacientes que sufren depresión, también se utilizan algoritmos de ajuste de curvas para modelar el espectro de frecuencias de cada paciente, los resultados permitieron clasificar la depresión bajo cuatro subtipos, ligados a factores como: la ansiedad, el estrés y la falta de placer [7].

Otra herramienta muy útil para el diagnóstico es la EEG - electroencefalografía, esta permite obtener un registro de la actividad eléctrica del cerebro, los resultados de este examen junto a algoritmos de clasificación se usaron en [8], para lograr el reconocimiento temprano de la enfermedad de Alzheimer.

En el trabajo de [9] se utiliza un avanzado algoritmo de visión artificial para analizar imágenes de resonancia magnética del cerebro y detectar la incidencia de esquizofrenia. También realiza la predicción de la efectividad de un tratamiento farmacológico.

La herramienta de clasificación LAMDA, basada en lógica difusa, realiza la clasificación de la personalidad y sus trastornos, determina índices de depresión y si, paralelamente, las patologías en los rasgos de personalidad del paciente dificultan su proceso de adaptación [9].

Como se observa en los trabajos citados, la mayoría de las aplicaciones usan algoritmos de Inteligencia Artificial, para lograr un diagnóstico temprano y acertado. Sin embargo, aunque son de gran ayuda, no consideran la taxonomía completa de los trastornos que forman parte de la clasificación CIE-10, la principal limitante es la gran variedad de trastornos, sumada al costo y las dificultades propias para adquirir imágenes y/o datos de pacientes diagnosticados que permitan la construcción de un conjunto de datos de entrenamiento [10], requisito indispensable para la aplicación de cualquier algoritmo de IA.

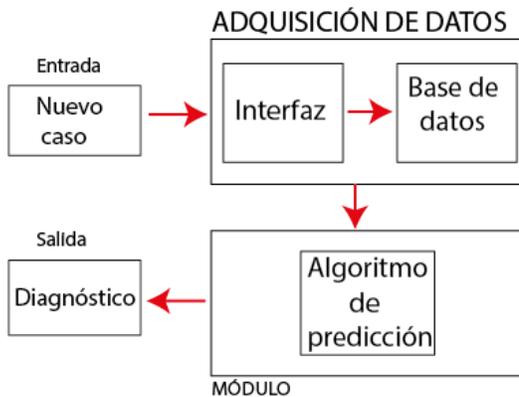
Este trabajo presenta el desarrollo del aplicativo Psico-Web, que servirá como una herramienta de apoyo en el diagnóstico de enfermedades mentales tomando en cuenta la taxonomía completa de la CIE-10. La determinación del trastorno se realizará utilizando la información de la sintomatología del paciente.

## 3. Metodología

### 3.1 Plataforma Psico-Web

#### I. Esquema

La plataforma Psico-Web incorpora una interfaz gráfica encargada de la interacción del sistema con el médico tratante, esta interfaz también permitirá la gestión de usuarios y la edición de la base de datos, por parte del administrador del sistema. La parte nuclear del aplicativo está constituida por dos módulos que son: el módulo de Adquisición de datos y el módulo de Predicción del diagnóstico (véase Figura 1).



**Figura 1.** Esquema del aplicativo Psico-Web [11-12].

El módulo de Adquisición de datos, está diseñado para que el médico tratante ingrese cada uno de los casos diagnosticados, el proceso empieza con una encuesta inicial, que servirá de filtro para la elección posterior de los síntomas que presenta el paciente. Cada nuevo bloque de datos (síntomas, trastorno) formará parte del conjunto de entrenamiento del algoritmo de *Machine Learning* usado para el diagnóstico de casos futuros.

El módulo de Predicción del diagnóstico, por su parte, incorpora un algoritmo de clasificación, este algoritmo discriminará el tipo de trastorno que presenta el paciente, de acuerdo a los síntomas que se han ingresado al sistema.

Una vez que concluye el ingreso de síntomas de un determinado paciente, el aplicativo ejecuta el algoritmo de ML y arroja el posible trastorno. El médico tratante, basado en su conocimiento y experiencia, tendrá que validar esta información, para ello podrá aceptar o rechazar la predicción del aplicativo. En caso de no estar conforme con el diagnóstico mostrado, también podrá seleccionar el nombre del trastorno que considere pertinente. Ambos diagnósticos, el prescrito por el médico tratante y el resuelto por el sistema son almacenados en la base de datos, para poder comprobar

el rendimiento del algoritmo de clasificación.

## II. Implementación

La metodología seleccionada para el desarrollo del aplicativo Psico-Web se denomina Ingeniería Web basada en UML (UWE). Esta metodología provee una guía del proceso de creación de la aplicación web y diagramas estándares para el diseño, ya que es una extensión del Modelo de Lenguaje Unificado (UML) [15], los modelos de UWE como: análisis de requerimientos, modelo de contenidos, modelos de navegación, modelo de presentación y modelo de procesos se desarrollaron bajo la tutela y supervisión de profesionales de la salud mental, con el fin de incluir aspectos importantes que faciliten el uso de la plataforma.

La metodología UWE permitió definir las funcionalidades básicas del sistema y el esquema general del aplicativo (Figura 1). Además, se consiguió agrupar los síntomas en cuatro categorías: Cognitivo, Conductual, Emocional y Fisiológico/Anatómico, cada uno relacionado con un trastorno específico. De esta forma se pueden incorporar filtros, para desplegar únicamente la información necesaria para el diagnóstico.

La plataforma cuenta con una página de inicio, que contiene información básica del proyecto. Para acceder a las funcionalidades es necesario que el usuario realice un registro, facilitando sus datos de contacto e información básica que se validará antes de darle el acceso.

Aunque el perfil de usuario principal de la plataforma será el de médico tratante, también existen otros dos roles que son: Administrador e Invitado, este último solo cuenta con una vista de información básica y de contacto del aplicativo web y es el rol previo, antes de la activación del rol

médico-tratante. El perfil Administrador es el que permite al usuario: la parametrización del sistema, el entrenamiento del algoritmo y la definición de roles para cada usuario registrado. El rol de médico-tratante habilita las opciones para el registro de casos de pacientes diagnosticados, la validación del diagnóstico del sistema y la realización de consultas de casos previos. (Véase Figura 2).

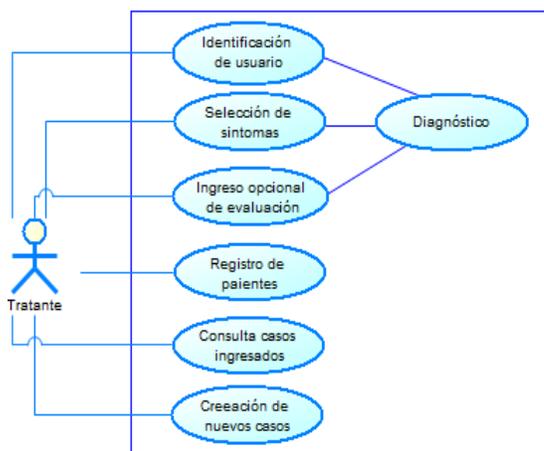


Figura 2. Caso de uso - Usuario tratante.

El proceso general que seguiría el médico-tratante, para ingresar un caso y lograr el diagnóstico, se muestra en la Figura 3.



Figura 3. Modelo de proceso del usuario tratante.

Todos los datos (síntomas-trastorno) seleccionados en el proceso de diagnóstico, se almacenarán en una base de datos relacional, de forma confidencial y sin revelar la identidad del paciente. El conjunto de datos de casos anteriores

servirá para entrenar el modelo de clasificación, que se encargará de hacer la predicción.

La aplicación desarrollada se basa en la arquitectura MVC (Modelo, Vista, Controlador) detallada en la Figura 4. Este modelo permite la separación entre los datos y la aplicación [14]. Así, la aplicación Psico-Web contiene separadamente: la clase de conexión con el gestor de base de datos, la vista con las páginas que el usuario podrá visualizar para interactuar con la aplicación y el controlador que contiene eventos para la realización de peticiones desde el cliente a la vista.

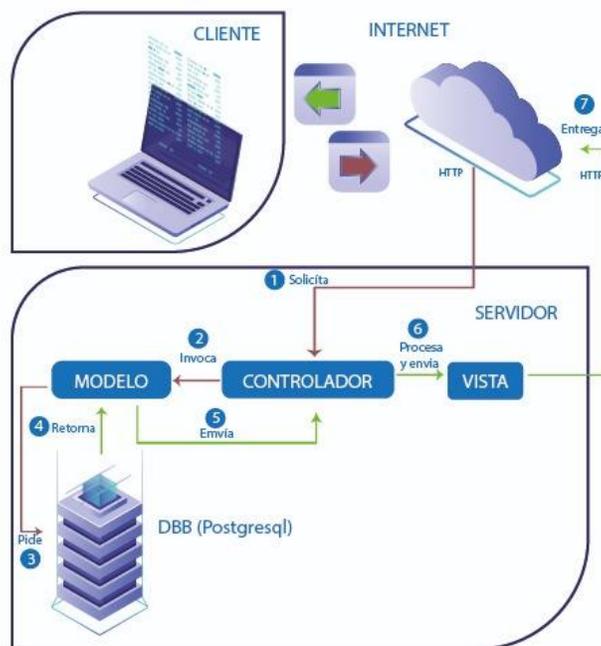


Figura 2. Arquitectura MVC de la aplicación Psico-Web

### 3.2 Algoritmo de clasificación

El propósito general que persiguen los algoritmos de *Machine Learning* es la búsqueda de patrones en grupos de datos específicos que se ajustan a características que pertenecen a un mismo concepto. La premisa básica consiste en la construcción de algoritmos que puedan recibir datos de entrada y a través de un análisis predecir un resultado [15].

Una tarea muy común, abordada por los algoritmos de ML, es la clasificación automática de un objeto, es decir, la determinación de la clase a la que pertenece un objeto en función de sus atributos (características), este proceso consta de tres fases que son: fase de entrenamiento, fase de evaluación y la fase de prueba.

La fase de entrenamiento consiste en generar un modelo a partir de un conjunto de datos. Dicho conjunto de datos está formado por un número de instancias (ejemplos), donde cada instancia contiene un conjunto de atributos y su clase correspondiente. De esta forma, el algoritmo busca reconocer patrones e identificar las relaciones entre los atributos y las clases. En la fase de evaluación, se prueba el modelo, utilizando como entrada un conjunto de datos (con instancias de atributos y clase conocidos), del cual se toman como entrada únicamente los atributos de cada objeto a clasificar. La clasificación realizada por el algoritmo se compara con la clasificación real, para determinar el rendimiento del modelo [16]. Una vez confirmado que la precisión del algoritmo es adecuada, se pasa a la fase de prueba en la que el modelo se pone en producción, es decir, se usa el modelo para clasificar nuevos objetos con clase desconocida.

En el caso del diagnóstico automático de desórdenes mentales, el algoritmo de ML debe ser capaz de determinar el tipo de trastorno que padece un paciente (clase) en función de los síntomas (atributos) que presenta. Para lograr este objetivo, el algoritmo seleccionado se entrenará con datos de casos de pacientes previamente diagnosticados y con los datos de la taxonomía “oficial” de la CIE-10. Cada caso estará constituido por un conjunto de síntomas (atributos) asociados a un solo trastorno (clase).

El algoritmo de ML utilizado en este trabajo es un caso particular (cuando  $k=1$ ) del algoritmo denominado  $k$ -vecinos más cercanos o  $k$ NN por sus siglas en inglés (*k-nearest neighbors*), este algoritmo fue diseñado y desarrollado por Fix y Hodges en 1951, su funcionamiento se basa en el cálculo de distancias entre el objeto que se va a clasificar y los objetos del conjunto de entrenamiento. De este modo, la clase a predecir se escoge entre las clases más frecuentes de los  $k$ -objetos más cercanos al objeto a clasificar [18].

Una de las características más relevantes de este algoritmo es que en la fase de entrenamiento se guardan las instancias, más no se construye un modelo, por lo que la clasificación se hace cuando llega la instancia de prueba. Así, en el aplicativo Psico-Web, cada vez que se ingresa un nuevo caso, se recalculan las distancias que permiten la elección del trastorno diagnosticado. Adicionalmente, se debe considerar que el  $k$ NN no hace suposiciones sobre la distribución que siguen los datos y asume que la clase de un dato depende sólo de los  $k$  vecinos más cercanos sin tomar en cuenta el modelo global [19].

La definición formal del algoritmo aplicado al diagnóstico automático de desórdenes mentales, toma en cuenta el conjunto de datos de entrenamiento, formado por  $m$  instancias. Cada instancia se representa por un vector  $O_j$  ( $j = 1, 2, \dots, m$ ) de coordenadas  $S_i$  ( $i = 1, 2, \dots, N$ ). Las coordenadas representan los síntomas y pueden tomar valores 1 o 0, dependiendo de si el paciente presenta el síntoma o no, respectivamente. De forma general un nuevo caso  $O_D$  está representado por un vector de características  $N$ -dimensional, de la forma:

$$O_D = [S_{1D}, S_{2D}, \dots, S_{ND}] \quad (1)$$

Los sub-índices de las coordenadas de la Ecuación 1, definen el síntoma y objeto, respectivamente. Cuando se ingresa un nuevo caso  $O_D$ , se calcula cada una de las distancias euclidianas  $d$  entre ese objeto ( $O_D$ ) y los demás objetos ( $O_j$ ) del conjunto de entrenamiento. Las distancias se calculan mediante la expresión:

$$d_{O_D, O_j} = \sqrt{\sum_{i=1}^N (S_{iD} - S_{ij})^2} \quad (2)$$

En nuestro caso, como  $k=1$  el diagnóstico estará dado únicamente por la clase del objeto  $O_j$  con la menor distancia al objeto  $O_D$ , (Ecuación 3).

$$\text{Diagnóstico} = \text{clase}(O_j | \min(d_{O_D, O_j})) \quad (3)$$

En la Figura 5, se presenta un pseudocódigo del algoritmo 1-NN utilizado para el diagnóstico automático de enfermedades mentales.

---

**Algoritmo 1 1NN**

---

**Entrada:**  $O_D = [S_1 \ S_2 \ \dots \ S_N]$   
**Salida:**  $\text{clase}(O_D)$

- 1: **para**  $j = 1, 2, \dots, m$  **hacer**
- 2:   **para**  $i = 1, 2, \dots, N$  **hacer**
- 3:     Calcular:  $dis = \sum_{i=0}^N (S_{iD} - S_{ij})^2$
- 4:   **fin para**
- 5:   Calcular:  $d_{O_D, O_j} = \sqrt{dis}$
- 6: **fin para**
- 7: **devolver**  $\text{clase}(O_j | \min(d_{O_D, O_j}))$

---

**Figura 5.** Seudocódigo para el clasificador 1NN [20].

## 4. Experimentos

En la fase de experimentación, se llevó a cabo la construcción del *dataset* de entrenamiento inicial, para ello se utilizaron los casos especificados en la clasificación CIE-10 capítulo VII, rango de códigos F00-F99 perteneciente a Trastornos Mentales y

del Comportamiento. De este modo, se determinaron 185 síntomas diferentes (185 coordenadas) y un total de 77 instancias, cada una asociada a un trastorno diferente.

Los síntomas se dividieron en cuatro grupos: Cognitivo, Conductual, Emocional y Fisiológico/Anatómico, esta agrupación redujo los 77 casos iniciales en 64, ya que los 13 trastornos restantes no presentaban sintomatología que pueda ser utilizada para el entrenamiento del algoritmo, por tanto, el modelo contó con un *dataset* inicial constituido por 64 ejemplos.

Para la evaluación del algoritmo, se utilizó un conjunto de 50 nuevos casos, que fueron ingresados uno a uno al sistema. A medida que se realizaba el diagnóstico automático del caso, un médico tratante validaba el resultado del sistema, de modo que se actualizaba el *dataset* de entrenamiento con la nueva instancia, el diagnóstico predicho y el diagnóstico acertado.

Al realizar la comparación entre el diagnóstico del sistema y el diagnóstico real dado por el profesional de la salud, se obtuvo que 22 de los 50 casos fueron clasificados correctamente (véase Figura 6), los 28 casos restantes representan el error (56%). Aunque puede parecer que los resultados preliminares fueron pocos satisfactorios, se debe resaltar que el conjunto de entrenamiento era disperso y contaba con muy pocos ejemplos de cada trastorno, por lo que, el algoritmo de clasificación se utilizó como una prueba de concepto, que cumplió con el porcentaje de acierto esperado (44%). Es importante recalcar, que la efectividad de los algoritmos de *Machine Learning* es muy sensible al tamaño del *dataset*, por tanto, es altamente probable que los resultados del algoritmo mejoren si la cardinalidad del conjunto de entrenamiento aumenta.

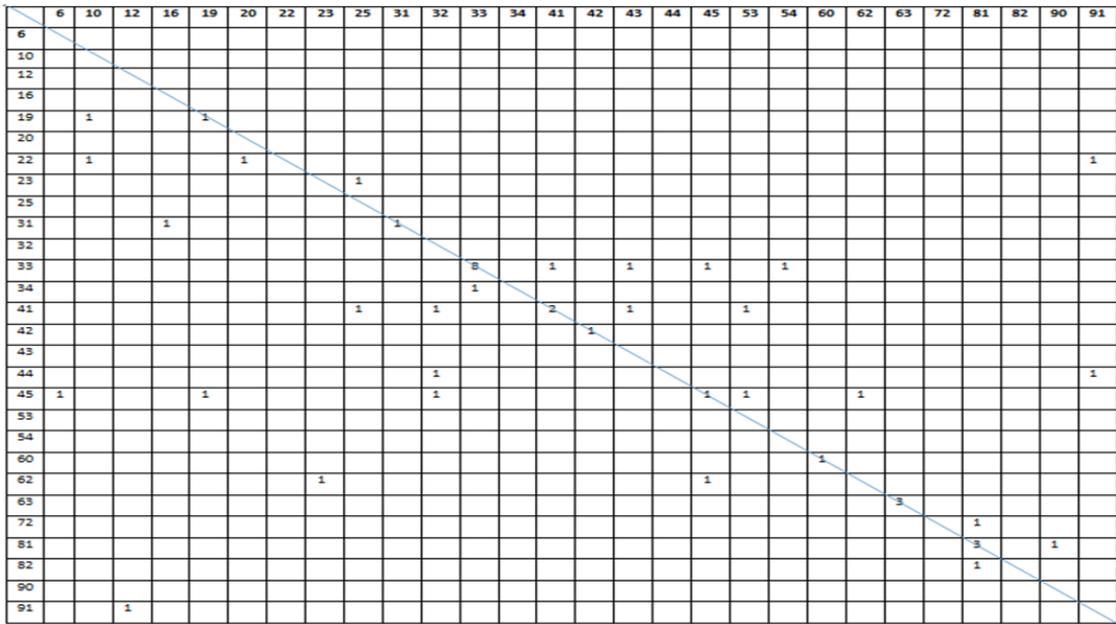


Figura 6. Matriz de confusión.

## 5. Conclusiones

La ambigüedad de los síntomas que presentan los pacientes que padecen algún desorden mental es un hecho frecuente que no puede ser caracterizado por métodos estadísticos tradicionales, por tanto, las técnicas de aprendizaje de máquina son primordiales para conseguir el diagnóstico automático.

El desarrollo de aplicativos Web requiere metodologías afines a las exigencias de los usuarios, por tal razón, Psico-Web fue desarrollado utilizando la metodología UWE basada en UML. Con esta metodología se consiguió que las especificaciones de los componentes de la aplicación quedaran claras para el usuario y el desarrollador. Además, se atendieron los requerimientos y elementos de interés del usuario.

El algoritmo de aprendizaje supervisado kNN (con  $k=1$ ) brinda una primera aproximación del diagnóstico automático. Sin embargo, el porcentaje de acierto es bajo (44%), esto se debe principalmente a que cada instancia solo

contiene unos pocos síntomas (1 a 23) de los 185 que se consideran en la representación de cada caso. Además, el algoritmo empleado es altamente dependiente del tamaño del conjunto de entrenamiento.

Para mejorar el desempeño en el diagnóstico dado por la aplicación es necesario buscar mecanismos para reducir la dimensionalidad de la sintomatología y la dispersión de los datos, es decir, disminuir el número de síntomas asociados a cada trastorno. Adicionalmente, se debería aumentar el tamaño del conjunto de entrenamiento, lo que implica, tener un mayor número de ejemplos por cada trastorno a diagnosticar.

Dado que, el conjunto de entrenamiento inicial solo contaba con un ejemplo de cada trastorno, resultó evidente que el diagnóstico fue mejorando a medida que ingresaban nuevos casos al conjunto de entrenamiento, siendo los últimos ingresados los casos acertados.

El aplicativo Psico-Web permitió la adquisición de datos de pacientes diagnosticados con algún trastorno mental. Estos datos sirvieron para la creación de un primer modelo de ML, que demostró la

factibilidad de un diagnóstico automático en este contexto. Gracias a la interfaz gráfica de la herramienta se puede extender el uso de la aplicación a un mayor número de profesionales de la salud mental, esto contribuirá a la construcción de un conjunto de datos de entrenamiento más sólido y robusto, y por tanto de un modelo más fiable.

El objetivo de trabajos futuros sería enriquecer el conjunto de datos de entrenamiento, con casos clínicos adicionales, y probar nuevos algoritmos de clasificación, que mejoren el rendimiento del diagnóstico. Ergo, la plataforma se diseñó de manera escalable.

## Referencias

- [1] O.M.S. “CIE-10 Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud Décima Edición” [online] 2003. Disponible en: <http://ais.paho.org/classifications/Chapters/pdf/Volume2.pdf>
- [2] J. Vallejo. “Introducción a la psicopatología y la psiquiatría”, 7ma ed., Barcelona: Elsevier Masson, 2015, pp 4-17.
- [3] C. Raheb, “Clasificación en psiquiatría. Conceptos y enfoques. Sistemas de clasificación: DSM-IV, ICD-0” Tesis master, Universidad Autónoma de Barcelona, 2008.
- [4] P. Pichot, J. López, I. Aliño, M.Valdés. “Manual diagnóstico y estadístico de los trastornos mentales”, Barcelona: Elsevier Masson S.A.,2006, pp 279-322.
- [5] E. Bermejo, A. Martínez. “Machine Learning: el smart digital workplace como ventaja competitiva” Raona, 2017.
- [6] Inteligencia Artificial, ¿Qué es machine learning? [Guía completa para principiantes], Adext, 2017.
- [7] B. Resnick. “Treating depression is guesswork. Psychiatrists are beginning to crack the code” [online] Vox, 2017. Disponible en:<https://www.vox.com/science-and-health/2017/4/4/15073652/precision-psychiatry-depression>.
- [8] Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R. E., Wahlund, L. O., and Dierks, T. “Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)”. *Journal of neuroscience methods*, 161(2), pp. 342-350, 2007.
- [9] Bo Cao, Raymond Y. Cho, Dachun Chen, Meihong Xiu, Li Wang, Jair C. Soares & Xiang Yang Zhang. “Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity”. *Molecular Psychiatry*, 2018.
- [10] M.G. Gutiérrez, L. Peña, M. Santiuste, D. García, M. Ochotorena, F. San Eustaquio, M. Cánovas. “Comparación de los sistemas de clasificación de los trastornos mentales: CIE-10 y DSM-IV.” *Dep. Psicología Clínica y Psiquiatría. Hospitales Varios de Murcia*.
- [11] A. Nieves del Valle, “Metodologías de diseño usadas en ingeniería Web, su vinculación con las NTICs”. *Universidad Nacional de la Plata, Madrid*, 2010.
- [12] J. Molina, M. Zea, M. Contento, F. García “Comparación de metodologías En aplicaciones web”, *Revista 3C Tecnología*, vol.7, nº1, pp. 6-15, 2018.
- [13] C. Nieves, J. Ucán, V. Menéndez, “UWE en Sistema de Recomendación

- de Objetos de Aprendizaje. Aplicando Ingeniería Web: Un Método en Caso de Estudio”, Revista Latinoamericana de Ingeniería de Software, vol. I, n° 1, pp. 137-143, 2014.
- [14] J. Pavón “Estructura de las aplicaciones orientadas a objetos, patrón Modelo-Vista-Controlador (MVC)”. Dep. Ingeniería de software e inteligencia artificial. Universidad Computense Madrid. [online] 2008. Disponible en <https://www.fdi.ucm.es/profesor/jpavon/poo/2.14.mvc.pdf>
- [15] D.L. Medin. “Concepts and Conceptual Structure”. American Psychologist, 44(12), pp 1469-1481, 2016.
- [16] J.I. Bagnato, “Aprende Machine Learning” [online] MINI BIO, 2017. Disponible en: <http://www.aprendemachinelarning.com/7-pasos-machine-learning-construir-maquina/>
- [17] C.M. Bishop. “Pattern Recognition and Machine Learning”. New York: Springer, 2006, pp.179-220.
- [18] E. Fix, J.L.Hodges. Discriminatory analysis, nonparametric discrimination. Project 21-49-004, Rept. 4, USAF School of Aviation Medicine, Randolf Field, 1951.
- [19] M. X Uriz, M. Galar. “Aprendizaje de distancias basadas en disimilitudes para el algoritmo de clasificación kNN”. Trabajo fin de grado, Universidad Pública de Navarra, Pamplona, 2015.
- [20] R. Barandela, N. Cortés, A. Palacios. “The Nearest Neighbor Rule and the Reduction of the Training Sample Size” in 9th Symposium on Pattern Recognition and Image Analysis. España, 2001.