

Prototipo de redes neuronales artificiales de clasificación binaria con una sola etiqueta para la detección del trastorno depresivo mayor e ideación suicida.

Prototype of artificial neural networks of a binary classification with a single label for the detection of major depressive disorder and suicidal thoughts.

Ricardo de Jesús González Morales (1).
Estudiante, Tecnológico Nacional de México/Instituto Tecnológico de Tuxtla Gutiérrez.
116270775@tuxtla.tecnm.mx.

Rafael Antonio Bolaños Camaras (2), estudiante, TecNM/ITTG, 116270740@tuxtla.tecnm.mx.

Aída Guillermina Cossío Martínez* (3), TecNM/ITTG, aida.cm@tuxtla.tecnm.mx.

Héctor Guerra Crespo (4), TecNM/ITTG, hector.gc@tuxtla.tecnm.mx.

Germán Ríos Toledo (5), TecNM/ITTG, german.rt@tuxtla.tecnm.mx.

Galdino Belizario Nango Solís (6), TecNM/ITTG, galdino.ns@tuxtla.tecnm.mx.

*corresponding author.

Artículo recibido en noviembre 24, 2020; aceptado en diciembre 17, 2020.

Resumen.

Este trabajo presenta el diseño de un sistema de 2 redes neuronales artificiales conectadas a una aplicación web para la detección del trastorno depresivo mayor y suicidio, mediante salidas probabilísticas derivadas de cada una de las redes neuronales. El modelo implementa redes neuronales tipo feedforward de clasificación binaria con una sola etiqueta. En el proceso de aprendizaje, se emplean valores obtenidos de exámenes realizados por psicólogos y psiquiatras, para diagnosticar los trastornos en un individuo mediante el Inventario Depresivo de Beck. Dicho diagnóstico será realizado mediante el llenado de un formulario en la aplicación web, el cual pasará los datos a las redes neuronales para realizar la predicción y obtener el resultado del diagnóstico. La aplicación, así como las redes neuronales, son desarrolladas en lenguaje Python. Los resultados indican que después del entrenamiento y afinación, las redes neuronales detectan con un 96% de precisión patrones depresivos y suicidas en datos de pacientes que no fueron usados durante la fase de validación, por lo que se concluye que el sistema es capaz de predecir estos trastornos eficazmente en cualquier individuo.

Palabras clave: Trastorno depresivo mayor, suicidio, Django, Python, redes neuronales.

Abstract.

This work presents the design of a system of 2 artificial neural networks connected to a web application for the detection of major depressive disorder and suicide, through probabilistic outputs derived from each of the neural

networks. The model implements feedforward type neural networks of binary classification with a single label. In the learning process, values obtained from tests carried out by psychologists and psychiatrists are used to diagnose disorders in an individual using the Beck Depressive Inventory. This diagnosis will be made by filling in a form in the web application, which will pass the data to the neural networks to make the prediction and obtain the result of the diagnosis. The application, as well as the neural networks, they are developed in the Python language. The results indicate that after training and tuning, neural networks detect depressive and suicidal patterns with 96% accuracy in patient data that were not used during the validation phase, so it is concluded that the system is capable of predicting these disorders effectively in any individual.

Keywords: Major depressive disorder, suicide, Django, Python, neural networks.

1. Introducción.

Según la Organización Mundial de la Salud (2019), el Trastorno Depresivo Mayor es un trastorno mental bastante frecuente que tiene como sintomatología, la tristeza, la pérdida de interés o placer, sentimiento de culpa, falta de concentración y baja autoestima. Puede llegar a hacerse crónico y recurrente, dificultando el desempeño y calidad de vida de la persona en todos los ámbitos; en la escuela, con la familia y en las relaciones interpersonales. En México, la Secretaría de Salud nos indica que el Trastorno Depresivo Mayor se manifiesta a través de varios síntomas, como una respuesta a largos periodos de estrés y angustia. Existen multitud de factores que están asociados a la depresión y ansiedad. Es difícil que un paciente pueda detectar tempranamente la depresión, regularmente puede llegar a ser confundido con la tristeza, sin embargo, es de vital importancia poder detectarla a tiempo, para que ésta no se convierta en un trastorno permanente. La diferencia entre la tristeza y la depresión es que los síntomas suelen ser más graves, persistentes, e intensos. Este trastorno tiene como tratamiento, la psicoterapia, la cual consiste en ejercicios de relajación, que de manera conjunta ayudan a disminuir los síntomas de tristeza, ansiedad y melancolía, además de que aportan calma y mayor control en situaciones de estrés (IMMS, 2019).

Son los jóvenes de edades comprendidas entre 15 a 29 años y en su mayoría pertenecientes a instituciones escolares los que suelen desarrollar este tipo de problemas, y uno de los mayores retos a los que Chiapas se enfrenta es que no todas las instituciones poseen un psicólogo laborando dentro de ellas, además, la cantidad de especialistas en psiquiatría y psicología en nuestro estado es pequeña comparada con la población que se enfrenta a estos trastornos. Aunado a esto, cuando se habla del Trastorno Depresivo Mayor, en una etapa crónica o catatónica, el paciente necesita de medicamentos antidepressivos, e inclusive se puede llegar hasta el grado de hospitalizarlo en algún centro de salud mental por su seguridad, debido a que algunos presentan pensamientos suicidas que, las personas a su alrededor no pueden detectar (OMS, 2019).

Se reconoce al suicidio como una de las prioridades de salud pública. La Organización Mundial de la Salud (OMS, 2019) asegura que, alrededor de 800 000 personas en todo el mundo, cometen suicidio cada año, además que, por cada caso que se registra, existe más personas que intentan suicidarse y fallan. Es importante recalcar que el suicidio es la segunda causa principal de muerte en personas de entre 15 a 29 años de edad. Está comprobado que el 79% de los suicidios son producidos en países con una economía baja o mediana, en los cuales, mayormente, las personas se suicidan a través de la ingesta de plaguicidas, ahorcamiento o armas de fuego. El suicidio es considerado un grave problema de salud pública, debido a que no existe un grupo específico de personas a las que se les atribuya esta acción, puede ser realizado impulsivamente, en algún momento de crisis que nubla el juicio de las personas, así como su capacidad de afrontar sus problemas. Sin embargo, las tasas de suicidio son elevadas, en aquellos grupos que sufren algún tipo de discriminación, ya sea migrantes, refugiados, comunidad LGBT, personas con discapacidades, comunidades indígenas, jóvenes en edad escolar que sufre de bullying, e incluso la presión de las redes sociales y sus “challenges”. Aún quedan muchos retos para que las personas, reconozcan al suicidio como un trastorno tratable o del que se necesite ayuda profesional. Existen mucho tabúes alrededor de este, que impide que las personas con ideas suicidas busquen ayuda a tiempo, debido a que no se cree que sea un problema de prioridad.

Además, la Secretaría de Salud en México, ve al suicidio como un tema prioritario en las agendas locales de salud pública, y debido a esto, realizan distintos seminarios que tienen como principal objetivo el propiciar un espacio de reflexión en el que aquellos participantes, puedan reconocer factores sobre este tema. No solamente existen estos seminarios, de igual forma para poder superar los retos más importantes de este trastorno (como el estigma o la discriminación) se estableció diseñar y desarrollar modelos para promover la implementación de los servicios psicológicos y psiquiátricos, en aquellos centros que no cuenten con esta ayuda (IMMS, 2019).

El objetivo principal de este sistema de redes neuronales es que, a partir de diagnósticos realizados previamente por psicólogos, detectar si el paciente presenta el trastorno depresivo mayor y/o pensamientos suicidas, con el fin de poder canalizarlo a una clínica especializada, además de detectar los factores de mayor peso que atribuyen a estos trastornos. Estos factores pueden ser de diferente índole, ya sean económicos, sociales, o inclusive religiosos, y que no todos los pacientes presentan en conjunto.

El sistema ayuda a:

1. Psicólogos que estén en instituciones educativas y de servicio que no cuentan con una especialidad en el ámbito de Trastorno Mayor Depresivo, les proporciona un diagnóstico, donde pueda determinar si el paciente padece del trastorno y de ideación suicida, para asignar terapias o indicar el apoyo médico de un psiquiatra.
2. Psiquiatras, como un segundo diagnóstico y que pueda determinar un tratamiento, en su defecto para asegurar la respuesta a todas las necesidades psiquiátricas, del paciente.

2. Métodos.

Redes Neuronales.

Matich (2001) nos explica:

Las redes neuronales son otra forma de emular ciertas características propias de los humanos, tales como la capacidad de memorizar y asociar hechos. Si se examinaran con atención todos los problemas que no son capaz de expresar su contenido en un algoritmo hay una característica en común: la experiencia. En sí, las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano. (pág. 4)

Esto mismo hace que sea una herramienta útil para nuestro sistema, ya que un diagnóstico se basa en la experiencia previa del psicólogo o psiquiatra para dar un resultado con el cual se puede determinar si una persona tiene o no depresión y la terapia a seguir. Para el funcionamiento correcto de una red neuronal, se necesita de los conjuntos de datos guardados en bases de hechos, que son archivos que contienen toda información con la que será entrenada la red.

Las redes neuronales pueden ser clasificadas según el tipo de aprendizaje (supervisado, no supervisado), el tipo de aplicación y la arquitectura de la conexión (monocapa, multicapa). Una de las principales características de las redes neuronales, es su capacidad de aprendizaje; a la hora de entrenar la red neuronal, se busca conseguir que una aplicación determinada para un conjunto de entradas, produzca el conjunto de salidas deseadas o mínimamente consistentes. Este proceso de entrenamiento es llevado a cabo mediante la aplicación secuencial de diferentes conjuntos o vectores de entrada para que se ajusten los pesos de las interconexiones según un procedimiento predeterminado.

En las redes neuronales artificiales existen dos capas con conexiones, con el mundo exterior. Una capa de entrada, donde se presentan los datos a la red, y una capa de salida que mantiene la respuesta de la red a una entrada. El resto de las capas reciben el nombre de capas ocultas. La Figura 1, muestra el aspecto de una red neuronal artificial multicapa.

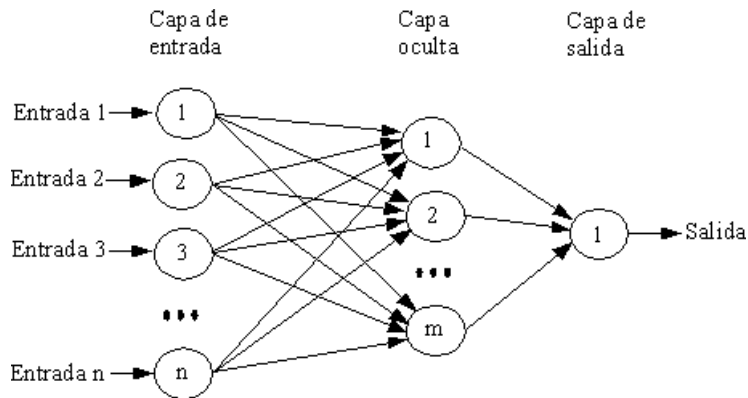


Figura 1. Arquitectura de una Red Neuronal Artificial Multicapa.

Red neuronal prealimentada (Feed-forward).

Una red neuronal prealimentada es una red neuronal artificial donde las conexiones entre las unidades no forman un ciclo.

La red neuronal prealimentada fue la primera y más sencilla forma de red neuronal artificial ideada. En esta red, la información se mueve en una única dirección: adelante. De las neuronas de entrada, a través de las neuronas escondidas (si los hay) hacia las neuronas de salida, no hay ningún ciclo o bucle en estas redes, un ejemplo es la que podemos observar en la figura 1.

Base de Hechos.

Una base de hechos incorpora una representación de hechos referidos al dominio de aplicación del sistema. En este caso, es la espina dorsal de todo el sistema ya que aquí se guardarán las experiencias que la red neuronal deberá analizar para poder aprender de ellas. Martínez (2006) comenta: “La base de hechos, es relevante para determinar la potencia del sistema con el que se está trabajando ya que este depende de la amplitud de los conocimientos contenidos en ella” (p. 47).

Esto nos dice que la elección de la forma en la que se pueda representar un conocimiento es fundamental para obtener un buen rendimiento del sistema. Esta relación debe permitir identificar conocimientos nuevos siendo lo flexiblemente necesario como para decidir incorporarlo a la base de hechos, eliminar un conocimiento contradictorio u obsoleto o sustituir conocimientos por otros más actualizados o precisos. (Martínez, 2006, p. 48).

Aplicación Web.

La World Wide Web (Red Informática Mundial) es un sistema de distribución de hipertexto o hipermedia interconectados y accesibles a través de Internet mediante el uso de los navegadores web. Un usuario de este tipo de sistemas puede visualizar páginas web compuestas de archivos que regularmente tienen extensión HTML que pueden contener multitud de contenidos multimedia. Estos sistemas web pueden ser accedidos navegando a través de hiperenlaces.

Esta manera de difusión de información ha ido en constante actualización comenzando desde las primeras páginas con formato de un boletín de periódico con el único fin de difundir noticias hasta las aplicaciones más complejas como lo son las redes sociales o aplicaciones que permiten una interacción más compleja entre un usuario y el aplicación.

Gracias al avance informático que se ha presentado en los últimos años, mediante los diferentes lenguajes y frameworks de desarrollo web que se han ido creando, existe un mundo de posibilidades en cuanto a los procesos que pueden ser automatizados. Con el fin de una mejor administración y seguimiento del sistema, hemos concluido que esta tecnología

es la indicada para el análisis de información que se desea realizar. Como las redes neuronales en cuestión se encuentran programadas en el lenguaje Python, se necesitó de una aplicación en un lenguaje o framework de desarrollo web del mismo lenguaje, que pueda ser accedido desde cualquier dispositivo por un profesional en la salud mental para usar la red y realizar diagnósticos. Esta aplicación web usa Django el cual es un framework compatible con librerías de Keras y Tensorflow que son usados en la red.

Django.

Django es un framework de aplicaciones web gratuito y de código abierto (open source) escrito en Python. Un framework web es un conjunto de componentes que facilitan y agilizan el desarrollo de sitios web. Cuando se construye un sitio web, se necesita de un conjunto de componentes similares: una manera de manejar la autenticación de usuarios (registrarse, iniciar sesión, cerrar sesión), un panel de administración para el sitio web, formularios, una forma de subir archivos, etc. Los frameworks sirven para que no se tenga que reinventar la rueda cada vez y se avance de una manera más rápida al construir un nuevo sitio.

Se usa el framework previamente mencionado el cual provee una estructura ideal con el patrón MVT (figura 2). Para Django el Modelo-Vista-Template es similar al Modelo-Vista-Controlador utilizado en otros frameworks como Laravel o Spring Boot.

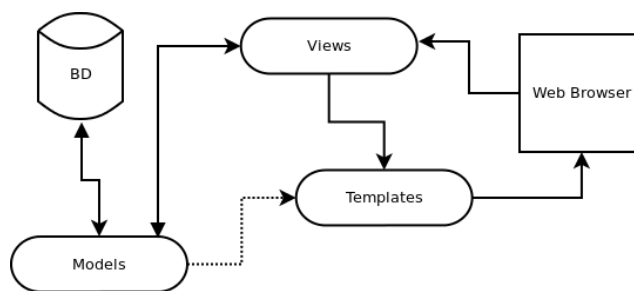


Figura 2. Modelo MVT.

El modelo es la capa de acceso a la base de datos. En esta capa contiene toda la información sobre los datos: accesos, validaciones, comportamientos y relaciones. La plantilla (template) es la capa de la presentación. Esta capa contiene los archivos HTML con la estructura de cómo se mostrará el contenido de la aplicación. La vista (view) es la capa de la lógica de negocios. Dentro de esta capa se contiene la lógica que accede al modelo y la delega a la plantilla correspondiente, este es el puente entre los modelos y las plantillas.

Se utilizó SQLite como el gestor de las bases de datos utilizadas para la aplicación web. Dentro de la base de datos para la aplicación, las tablas principales para que el sistema funcione correctamente deben de incluir los diagnósticos generados, el historial de usuarios y los usuarios registrados con sus respectivas relaciones para la gestión de los resultados de la red.

Inventario Depresivo de Beck.

El Inventario de Depresión de Beck-II (también conocido como BDI-II) es un autoinforme de lápiz y papel compuesto por 21 ítems de tipo Likert. El inventario inicialmente propuesto por Beck y sus versiones posteriores han sido los instrumentos más utilizados para detectar y evaluar la gravedad de la depresión. (Consejo General de Consejos Oficiales de Psicólogos, 2013, pág. 3). Podemos apreciar el encabezado y los primeros 4 ítems del inventario en la figura 3.

Inventario de Depresión de Beck (BDI-2)

Nombre _____ Estado Civil _____ Edad _____ Sexo _____
 Ocupación _____ Educación _____ Fecha _____

Instrucciones: Este cuestionario consta de 21 grupos de afirmaciones. Por favor, lea con atención cada uno de ellos cuidadosamente. Luego elija uno de cada grupo, el que mejor describa el modo como se ha sentido las últimas dos semanas, incluyendo el día de hoy. Marque con un círculo el número correspondiente al enunciado elegido. Si varios enunciados de un mismo grupo le parecen igualmente apropiados, marque el número más alto. Verifique que no haya elegido más de uno por grupo, incluyendo el ítem 16 (cambios en los hábitos de Sueño) y el ítem 18 (cambios en el apetito)

1. Tristeza
0 No me siento triste. 1 Me siento triste gran parte del tiempo.
2 Me siento triste todo el tiempo. 3 Me siento tan triste o soy tan infeliz que no puedo soportarlo.

2. Pesimismo
0 No estoy desalentado respecto de mi futuro. 1 Me siento más desalentado respecto de mi futuro que lo que solía estarlo.
2 No espero que las cosas funcionen para mí. 3 Siento que no hay esperanza para mi futuro y que sólo puede empeorar.

3. Fracaso
0 No me siento como un fracasado. 1 He fracasado más de lo que hubiera debido.
2 Cuando miro hacia atrás, veo muchos fracasos. 3 Siento que como persona soy un fracaso total.

4. Pérdida de Placer
0 Obtengo tanto placer como siempre por las cosas de las que disfruto. 1 No disfruto tanto de las cosas como solía hacerlo.
2 Obtengo muy poco placer de las cosas que solía disfrutar. 3 No puedo obtener ningún placer de las cosas de las que solía disfrutar.

Figura 3. Inventario de Depresión de Beck-II.

3. Desarrollo.

Redes Neuronales.

Se llevó a cabo el siguiente proceso de desarrollo para la implementación de cada una de las redes neuronales:

1. Recolección de los datos.
2. Preparación de los datos.
3. Elección de los hiperparámetros de las redes neuronales.
4. Diseño de la arquitectura de las redes neuronales.
5. Pruebas y validación.
6. Afinación.

Recolección de los datos.

Para la etapa de recolección, se emplean datos de 150 pacientes diagnosticados por profesionales de la salud mental, con el objetivo principal de entrenar las redes neuronales para su posterior aprendizaje. Principalmente, los datos que se buscaron en esta etapa son:

- a) Datos sociodemográficos del paciente diagnosticado, tales como la edad, sexo, ocupación, estado civil, etc. Estos datos se encuentran en la ficha de identificación psicológica del paciente.
- b) Inventario de depresión de Beck, el cual contiene todos los ítems/métricas relacionados con el Trastorno Depresivo Mayor, que indican si el paciente presenta el trastorno o ideación suicida.
- c) Otros factores que los pacientes consideren, fue el detonante de su trastorno.
- d) El diagnóstico final de los profesionales, que indica si el paciente padeció/padece de depresión, además de la ideación suicida.

Preparación de los datos.

Para la codificación de las características (columnas del dataset) utilizamos el método One Hot Encoding. Este es un formato ampliamente usado para datos categóricos, también llamado “codificación categórica”. La estrategia que implementa es crear una columna para cada valor distinto que exista en la característica que estamos codificando y,

para cada registro, marcar con un 1 la columna a la que pertenezca dicho registro y dejar las demás con 0 (Chollet, 2018).

En nuestro caso, el dataset es modificado usando esta estrategia, para aquellas columnas que pueden pertenecer a más de una categoría (Profesión, Escolaridad, Nivel socioeconómico, Religión), de manera que el dataset sea representado con una notación binaria.

Una vez determinados los datos que se requieren, se construye el dataset a partir del análisis de características del problema, en nuestro caso, tendrá la siguiente arquitectura (figura 4):

- Las columnas marcadas por el color amarillo, corresponden al sexo del paciente (hombre o mujer).
- Las columnas color azul corresponden a la escolaridad del paciente (sin estudios, preescolar, primaria, secundaria, preparatoria, universidad, posgrado).
- Las columnas color gris, indican la ocupación del paciente, la cual se dividió en 7 clases (estudiante, ama de casa, empleado, nada, oficio, emprendedor, y profesión) las cuales fueron las más frecuentes, y se englobaron las profesiones en una sola clase.
- Las columnas verdes representan el estado civil del paciente (soltero, casado, unión libre, divorciado, viudo, separado).
- Las columnas cafés corresponden a la religión del paciente (católico, cristiano, ateo, adventista, testigo de Jehová, otra, ninguna).
- La columna “Socioeconómico” indica el nivel socioeconómico del paciente, clasificado en 0 = bajo, 1 = medio bajo, 2 = medio, 3 = medio alto, 4 = alto.

Figura 4. Arquitectura del dataset.

Con esto tenemos listos los datos de tipo sociodemográficos, ya que, con esta representación binaria, la red neuronal podrá aprender de una manera más eficiente y precisa. De igual manera se tienen que preparar los datos depresivos, que son representaciones de los 21 ítems del Inventario Depresivo de Beck. Como podemos ver en la figura 5, el inventario depresivo de Beck puntúa para cada uno de los ítems, en una escala de 0 a 3, la variable que representa el grado de malestar del paciente. Esto significa que cada punto de los diferentes ítems pertenece a una misma clase, por lo que no es necesario utilizar One Hot Encoding en este caso.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16a	16b	17	18a	18b	19	20	21
1	0	2	1	1	0	2	0	1	3	0	3	1	1	0	1	0	1	0	1	1	1	0	0
2	2	2	2	2	2	1	3	2	2	1	0	3	3	3	2	3	0	2	2	0	1	2	1
3	1	1	1	1	1	1	1	2	1	2	1	1	2	1	1	0	1	0	1	0	2	1	1
4	3	3	2	2	3	3	2	1	3	1	1	2	2	2	2	0	3	1	2	0	1	1	0
5	1	1	2	2	1	1	2	2	3	3	1	2	1	1	0	3	2	0	2	0	2	1	3
6	3	3	3	3	3	3	3	3	3	2	1	0	3	3	3	3	0	3	0	3	3	3	3
7	0	1	2	1	1	0	1	0	2	3	0	1	1	0	0	1	1	0	0	1	0	0	1
8	2	3	1	1	1	3	2	1	3	1	0	1	1	1	1	1	0	0	0	1	0	1	1
9	2	1	2	2	1	0	2	1	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2

Figura 5. Inventario de Beck codificado.

Las etiquetas de salida son los pronósticos o resultados a los que queremos llegar, y por supuesto corresponden a los datos que la red neuronal va a predecir, en esta instancia corresponden a las columnas de trastorno (depresión, suicidio o nada).

Como podemos observar en la figura 6, estos campos de igual manera están codificados en binario, en donde 0 significa la ausencia del trastorno y 1 la existencia.

TRASTORNO		
DEPRESIÓN	SUICIDIO	NADA
1	0	0
1	0	0
1	0	0
1	0	0
1	0	0

Figura 6. Datos de salida del dataset.

Sería problemático alimentar la red neuronal con datos en diferentes rangos (la edad no puede ser codificada categóricamente), ya que con el tiempo a pesar de que la red se adaptará a datos heterogéneos, será más difícil para esta aprender.

Chollet nos dice que una práctica muy empleada en el mundo del Deep Learning para tratar con datos de este tipo, es la normalización o estandarización según las características, para esto se sustrae la media de cada columna y se divide entre la desviación estándar, esto para que la característica representada en la columna este centrada en 0 y tenga una desviación estándar en común con las otras características. (Chollet, 2018, pág. 86).

Debido a esto, realizamos el proceso de normalización en nuestro dataset, usando el método **zscore** de **scipy**, dando como resultado la siguiente tabla (figura 7).

	0	1	2	3	4	5	6	7	8	9	10	11	12
0	-0.342149	0.986754	-0.986754	-0.0819232	-0.142857	-0.613308	0.728431	-0.237356	1.12815	-0.267261	-0.380839	-0.281312	-0.204124
1	-0.0253444	0.986754	-0.986754	-0.0819232	-0.142857	1.6305	-1.37281	-0.237356	1.12815	-0.267261	-0.380839	-0.281312	-0.204124
2	0.29146	0.986754	-0.986754	-0.0819232	-0.142857	-0.613308	0.728431	-0.237356	-0.886405	-0.267261	2.62578	-0.281312	-0.204124
3	-0.658954	-1.01342	1.01342	-0.0819232	-0.142857	1.6305	-1.37281	-0.237356	1.12815	-0.267261	-0.380839	-0.281312	-0.204124
4	-0.817357	0.986754	-0.986754	-0.0819232	-0.142857	1.6305	-1.37281	-0.237356	1.12815	-0.267261	-0.380839	-0.281312	-0.204124
5	-1.45097	-1.01342	1.01342	-0.0819232	-0.142857	1.6305	-1.37281	-0.237356	1.12815	-0.267261	-0.380839	-0.281312	-0.204124

Figura 7. Dataset codificado y estandarizado.

Una vez que los datos están estandarizados, el dataset está listo para alimentar a la red neuronal. La preparación y limpieza de los datos es un paso primordial para el análisis de datos, debido a que muchos algoritmos no funcionan de una forma correcta si los alimentamos con “raw data” que no pueden entender. Antes de construir nuestras redes neuronales, separamos el dataset en 3 secciones, las cuales utilizamos para el entrenamiento (los datos que se utilizaron para entrenar a la red neuronal), la validación (la sección que usó la red en el momento de observar cómo se comporta con datos que nunca ha visto), y el test (una última prueba realizada después de haber afinado las redes neuronales, para observar que no existe overfitting).

Elección de los hiperparámetros de las redes neuronales.

Ya tenemos listos los datos, por lo que a partir de los 2 problemas de clasificación que se quieren resolver con las redes, elegimos los siguientes hiperparámetros para cada red:

- 1) Para la primera red neuronal, encargada de detectar si el paciente sufre de ideación suicida o no, usamos la función de pérdida **“binary_crossentropy”** debido a que es la función determinada por Keras para los problemas de clasificación binaria (esto debido a que se clasifica si el paciente refiere ideación suicida o no, 0 ó 1), la capa de salida implementa la función de activación **“sigmoid”** (figura 9) ya que para todos los valores de x, solamente puede arrojar datos entre 0 y 1, los cuales corresponden a los valores binarios que se buscan, esto a diferencia de otras funciones de activación como por ejemplo **“softmax”** (figura 21) que sirve

cuando las clases son mayores a dos. (Chollet, 2018, págs. 68-77). El optimizador es “**ADAM**”, ya que es un optimizador que funciona excelente en los problemas de clasificación.

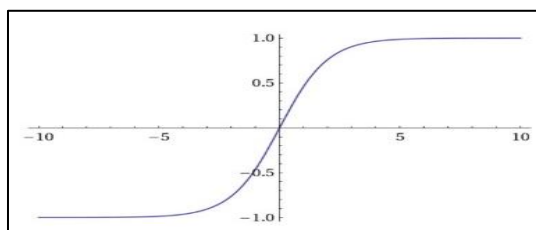


Figura 8. Función Softmax.

- 2) En el caso de la segunda red encargada de predecir si un individuo padece del Trastorno Depresivo Mayor o no, se nos presenta un problema de clasificación binaria o de “dos clases” (un sí o no), por lo que la función de pérdida que usamos es “**binary_crossentropy**”, siendo los valores cercanos a 0 la probabilidad de que el paciente no refiera el trastorno y cercanos a 1 que sí. Nuevamente el optimizador es “**ADAM**” mientras que la capa de salida tiene como función de activación “**sigmoid**” (figura 9).

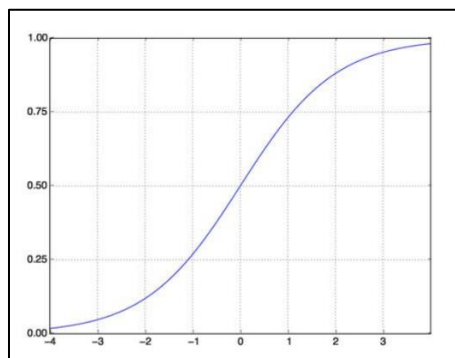


Figura 9. Función Sigmoid.

Diseño de la arquitectura de las redes neuronales.

Para la elección de la arquitectura de cada una de las redes neuronales, analizamos las características del dataset que se construyó (figura 10):

1. La primera red neuronal, es de tipo “clasificación binaria con una sola etiqueta”, debido a que nos permite clasificar si el paciente padece de ideación suicida o no (el paciente solamente puede caer en una de estas situaciones, lo cual hace al problema excluyente). Según el número de columnas del dataset, la capa de entrada tiene 61 entradas para 128 neuronas, mientras que la capa de salida está compuesta de 1 neurona, la cual arroja la probabilidad de que sea positivo el diagnóstico (1 o valores cercanos a 1) o no (0 o valores cercanos a 0).
2. La segunda red neuronal, se encarga de predecir si el paciente sufre del Trastorno Depresivo Mayor o no. La capa de entrada implementa las mismas 61 entradas, mientras que su salida tiene una sola neurona ya que la salida es binaria.
3. Las 2 redes neuronales comienzan con 1 sola capa oculta, esto debido a que con una capa oculta es suficiente en la mayoría de problemas donde se aplican estos algoritmos (Aravena, 2007, pág. 41). Esta capa oculta tiene 16 neuronas.

Layer (type)	Output Shape	Param #
dense_497 (Dense)	(None, 128)	7936
dense_498 (Dense)	(None, 16)	2064
dense_499 (Dense)	(None, 1)	17

=====
Total params: 10,017
Trainable params: 10,017
Non-trainable params: 0

Figura 10. Arquitectura de las redes neuronales.

Pruebas y validación.

Una vez construidas las redes neuronales, podemos comenzar a entrenarlas con nuestros datos de entrenamiento y validación (figura 11).

```
Instructions for updating:
Use tf.cast instead.
Train on 100 samples, validate on 50 samples
Epoch 1/5
100/100 [=====] - 3s 31ms/step - loss: 0.5843 - accuracy: 0.8500 - val_loss: 0.3728 - val_accuracy: 0.8400
Epoch 2/5
100/100 [=====] - 1s 6ms/step - loss: 0.2786 - accuracy: 0.8500 - val_loss: 0.3094 - val_accuracy: 0.8400
Epoch 3/5
100/100 [=====] - 0s 4ms/step - loss: 0.2065 - accuracy: 0.8600 - val_loss: 0.3198 - val_accuracy: 0.8600
Epoch 4/5
100/100 [=====] - 1s 6ms/step - loss: 0.1467 - accuracy: 0.9700 - val_loss: 0.3725 - val_accuracy: 0.8600
Epoch 5/5
100/100 [=====] - 1s 7ms/step - loss: 0.1073 - accuracy: 0.9800 - val_loss: 0.3996 - val_accuracy: 0.8400
```

Figura 11. Redes siendo entrenadas en 5 épocas.

Se usaron 100 registros para el entrenamiento y 50 para la validación. Podemos observar en la figura 11 que mientras aumenta el número de épocas, se reduce la pérdida (loss: 0.1073) e incrementa la precisión (accuracy: 98%), lo cual es una señal de que el modelo está aprendiendo. La métrica **val_loss** indica la pérdida en el conjunto de validación la cual de igual manera va disminuyendo, aunque de manera más lenta que la pérdida de entrenamiento, y la métrica **val_accuracy**, nos indica cómo se comportan las redes con datos que nunca han visto, como podemos observar la precisión real en este entrenamiento es de 84%, mientras que en la cuarta época llega al 86%. (figura 12)

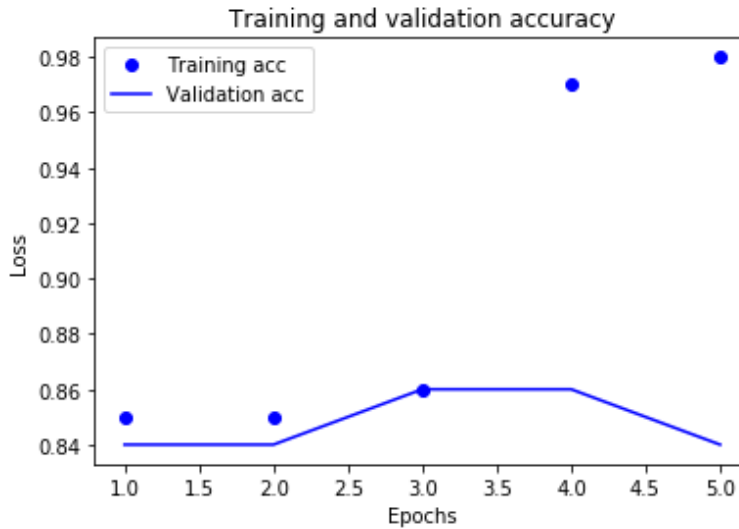


Figura 12. Gráfica de la precisión de la red durante el entrenamiento y la validación.

Afinación.

En el caso del sistema de redes neuronales para el trastorno depresivo y suicidio, al no poseer demasiados registros en el dataset (150 son pocos datos), es necesario llevar a cabo procesos de validación que permitan que la red pueda aprender y alcanzar una buena precisión, con un mínimo de datos (algunos datasets pueden tener miles de registros), por lo que la afinación consistió en llevar a cabo el proceso conocido como K-FOLD Cross Validation (figura 13), el cual es un algoritmo iterativo en donde se divide el dataset y se entrena la red neuronal K veces, cada una de estas partes será usada como parte del conjunto de entrenamiento y será usada solo una vez como el conjunto de validación, de manera que la red neuronal pueda aprender y ser testeada con absolutamente todos los registros del dataset. Preferentemente, dejaremos la arquitectura de las 2 redes con 1 sola capa oculta, ya que, aunque no exista una fórmula que nos asegure el número de capas o neuronas ocultas, Aravena nos dice que “Un aumento del número de capas de neuronas ocultas se traduce en un cambio en la estructura de la red, pudiéndose obtener resultados diferentes. Generalmente con una capa oculta es suficiente.” (Aravena, 2007, pág. 41).

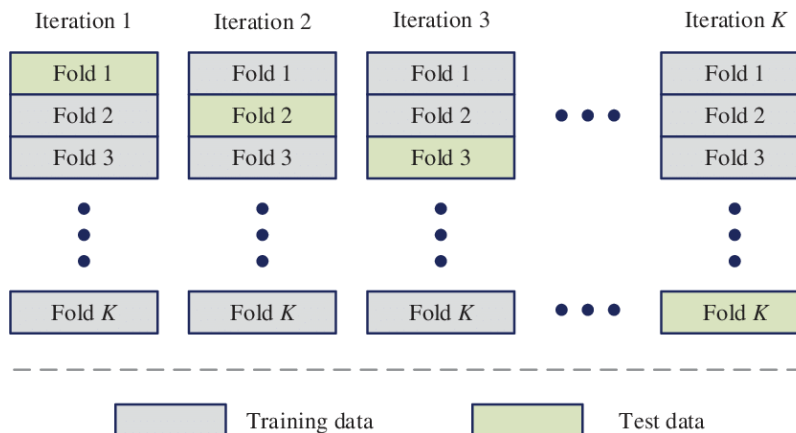


Figura 13. Algoritmo K-fold Cross Validation.

Al realizar este paso, la red aprende utilizando todos los datos para entrenarse y validarse, por lo que la precisión aumenta considerablemente. Podemos observar en la figura 14 que en la iteración número 7, alcanzamos el 100% de

precisión en validación y una pérdida de 0.124, lo cual quiere decir que, al momento de validar la red, acertó en todos los diagnósticos de prueba.

```

> Fold 1 - Loss: 0.14749717712402344 - Accuracy: 93.75%
-----
> Fold 2 - Loss: 0.8180436491966248 - Accuracy: 75.0%
-----
> Fold 3 - Loss: 0.1744883507490158 - Accuracy: 93.75%
-----
> Fold 4 - Loss: 0.23311275243759155 - Accuracy: 87.5%
-----
> Fold 5 - Loss: 0.2512471079826355 - Accuracy: 93.75%
-----
> Fold 6 - Loss: 0.29416775703430176 - Accuracy: 81.25%
-----
> Fold 7 - Loss: 0.12467191368341446 - Accuracy: 100.0%
-----
> Fold 8 - Loss: 0.9844509959220886 - Accuracy: 80.0000011920929%
-----
> Fold 9 - Loss: 0.21479323506355286 - Accuracy: 86.66666746139526%
-----
> Fold 10 - Loss: 1.4714584350585938 - Accuracy: 73.33333492279053%
-----
    
```

Figura 14. Resultados del entrenamiento de las redes usando KFold Cross Validation.

Conexión entre la aplicación web y las redes neuronales.

Para que un doctor tenga acceso a la aplicación, primero debe registrarse en el sistema y posteriormente iniciar sesión, a través de las pantallas que se muestran en las figuras 15 y 16. Las rutas están protegidas para que únicamente los usuarios registrados puedan utilizarlo, implementando un nivel de seguridad óptimo.

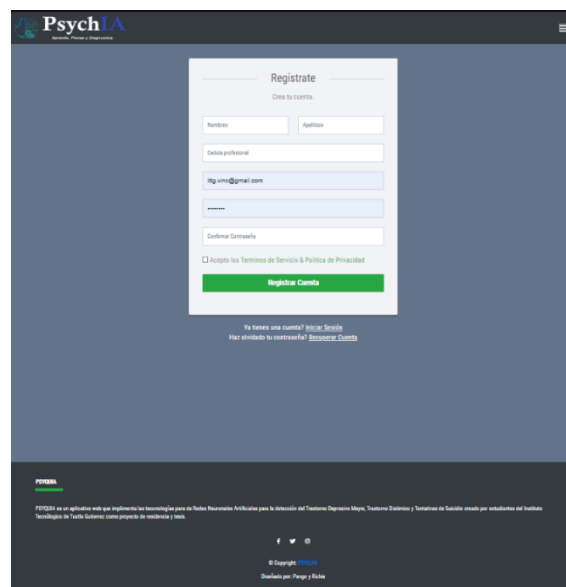


Figura 15. Formulario de registro.

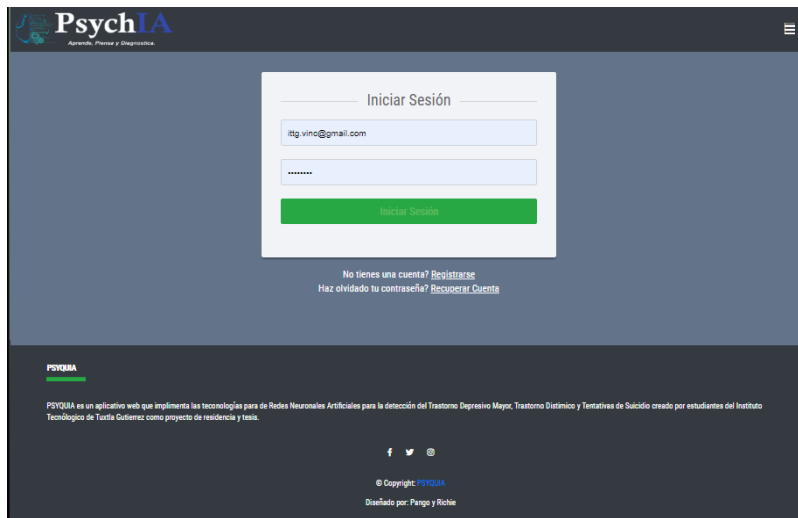


Figura 16. Formulario de inicio de sesión.

Lo primero que se muestra es el dashboard del doctor. Este está dividido en tres secciones:

1. Datos generales de los diagnósticos realizados y pacientes registrados (figura 17).

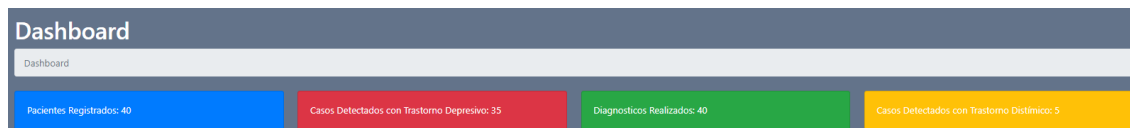


Figura 17. Dashboard: Datos generales de los diagnósticos.

Estos rectángulos muestran el total de los pacientes que ha registrado el doctor para realizar diagnósticos. El total de los diagnósticos que ha realizado el doctor durante todo el tiempo usando la aplicación y los totales de casos en donde el resultado del diagnóstico haya detectado un trastorno depresivo mayor o un trastorno distímico.

2. Tabla de pacientes registrados.

Nombre	Grado Académico Alcanzado	Municipio de Residencia	Edad	Primer Diagnóstico Realizado
Tiger Nixon	System Architect	Edinburgh	61	2020/09/25
Último Diagnóstico Realizado 2020/10/28				
Sonya Frost	Software Engineer	Edinburgh	23	2008/12/13
Rhona Davidson	Integration Specialist	Tokyo	55	2010/10/14
Herrod Chandler	Sales Assistant	San Francisco	59	2012/08/06
Garrett Winters	Accountant	Tokyo	63	2011/07/25
Colleen Hurst	Javascript Developer	San Francisco	39	2009/09/15
Cedric Kelly	Senior Javascript Developer	Edinburgh	22	2012/03/29
Brielle Williamson	Integration Specialist	New York	61	2012/12/02
Ashton Cox	Junior Technical Author	San Francisco	66	2009/01/12
Airi Satou	Accountant	Tokyo	33	2008/11/28

Figura 18. Tabla de pacientes registrados.

La tabla muestra todos los pacientes registrados con sus respectivos datos generales y las fechas del primer y del último diagnóstico registrado.

Para poder navegar en las demás ventanas que maneja la aplicación, se tiene una barra a la izquierda la cual muestra el nombre del usuario que ha iniciado sesión y las opciones de navegación (figura 19).

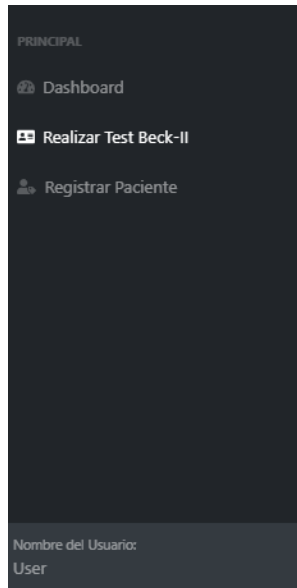


Figura 19. Barra lateral izquierda.

Para que un doctor pueda realizar un diagnóstico, primero debe de realizar un registro del paciente en cuestión. Esto se hace dando clic en la opción “*Registrar Paciente*” en la barra lateral izquierda. Esto abre el siguiente formulario para realizar el registro.

Figura 20. Formulario para registrar paciente.

Con los pacientes registrados, se tiene la posibilidad de realizar el test de Inventario Depresivo de Beck dando clic en su opción dentro de la barra lateral izquierda. Esto se realiza llenado el formulario que se muestra en la figura 21, este

tiene un botón llamado **“Registrar Test Depresivo de Beck”** el cual envía los datos ingresados en los campos hacia la red neuronal para su análisis y posteriormente llenar las tablas de la aplicación donde se consultan los diagnósticos depresivos realizados, casos depresivos encontrados y casos distímicos encontrados.

Figura 21. Formulario para diagnosticar.

Hemos probado el sistema con datos que no ha visto anteriormente, desde la interfaz web. Utilizamos 5 casos que no fueron vistos por la red en ningún momento de su entrenamiento ni validación. Los resultados obtenidos fueron de 100% de aciertos (en la derecha de la figura 22 están los datos extraídos del dataset, y a la izquierda las predicciones hechas por la red, siendo 1 la existencia de los trastornos), sin embargo, al momento del entrenamiento, la precisión alcanzó 96% por lo que esta es la precisión real de las redes.

0	0	150	0
1	0	151	0
2	1	152	1
3	0	153	0
4	0	154	0
5	0	155	0

Figura 22. Datos predichos vs datos reales.

Conclusiones.

El uso de un modelo que implementa redes neuronales en este tipo de problemas, donde se debe tener cierto conocimiento especializado, es cada vez más frecuente e incluso en los últimos años se ha estado investigando como aprovechar el gran poder de estos algoritmos de inteligencia artificial, en tareas que antes se pensaba que solamente los humanos podían realizar. Las redes neuronales son ampliamente utilizadas en el ámbito médico, para la detección de múltiples enfermedades, tales como el cáncer o tumores (por medio de redes neuronales convolucionales en este

último), debido a que han demostrado tener un grado de precisión superior a otros algoritmos, por supuesto las métricas y el alcance de estas dependerán del proceso de entrenamiento.

Como se observa en el desarrollo del proyecto, el proceso de construcción de las redes neuronales depende de un análisis de los datos a profundidad, ya que a partir de la arquitectura que tenga nuestro dataset, será el tipo de la red y los hiperparámetros que usaremos para ella. No solo debemos preocuparnos por la red neuronal, si no que la base de todo está dentro de los datos, no podemos alimentar una red con estos sin antes procesarlos, es por ello importante llevar a cabo la ingeniería de características y limpieza de datos, de manera que puedan ser una representación útil del problema. En la actualidad no existen reglas como tal para adivinar el número de capas ni de unidades ocultas, por lo que el entrenamiento es un proceso iterativo, en donde se tiene que probar diferentes arquitecturas para tratar que el modelo muestre resultados óptimos.

En este proyecto, se logró alcanzar una precisión del 96% durante el entrenamiento y del 100% en la fase de validación, por lo que se espera coadyuvar en el proceso de diagnóstico de depresión, y detectar con mayor facilidad si una persona puede sufrir de estos trastornos y de ser así, canalizarlo con un psicólogo que ayude al paciente a tratar con estos sentimientos negativos. La red neuronal aprendió correctamente y saco conclusiones a partir de los datos con los que fue entrenada, asignándole un peso a cada variable, el cual indica que tan importante es esta para el diagnóstico depresivo.

La implementación de la aplicación web en este proyecto, dio múltiples ventajas que muchos usuarios buscan actualmente, tal es el caso de la portabilidad que se les brinda en cuanto al uso de interfaces responsivas, la comodidad de tener el sistema disponible las 24 horas, sin necesidad de tener que encender equipos cada vez que se quiera acceder a este, y la seguridad de que los datos de los diagnósticos no se perderán si algún accidente llegara a suceder. La programación web acerca a los usuarios cada vez más a los sistemas de todo tipo, no solo inteligentes, por lo que es de esperarse que, en un futuro, la mayoría de servicios sean ofrecidos a través de páginas electrónicas, al igual que lo servicios que ya han sido dominados por este tipo de innovación tecnológica (e-commerce principalmente).

Agradecimientos.

Le extendemos nuestro principal agradecimiento a la M.C. Aida Guillermina Cossio Martínez por el apoyo que nos han brindado durante el desarrollo del proyecto, así como a la psicóloga Sofía Amairani González Morales por facilitarnos el Inventario Depresivo de Beck, además del conocimiento brindado en el ámbito de los trastornos psicológicos.

Referencias bibliográficas.

- Aravena, R. (Noviembre de 2007).** Comparación de MLP-R y MLP-C para aproximación de funciones no lineales.
- Chollet, F. (2018).** Deep Learning with Python. Shelter Island, NY 11964: Manning Publications.
- Consejo General de Consejos Oficiales de Psicólogos. (2013).** Evaluación del Inventario. Madrid España: EFPA.
- IMMS. (2019).** Obtenido de Atención a Conducta Suicida: www.gob.mx/salud/sap/documentos/guia-practica-para-la-atencion-del-pacientecon-conducta-suicida-en-hospitales-generales.
- Matich, D. (2001).** www.frro.utn.edu.ar. Obtenido de Informática Aplicada a la Ingeniería de Procesos- Orientación I.
- Martínez, M. (2006).** repositorio.unican.es. Obtenido de conocimiento y bases de datos: una propuesta de integración inteligente: repositorio.unican.es/xmlui/bitstream/handle/10902/1172/1de3.MAMcap2.pdf?s equence=2.

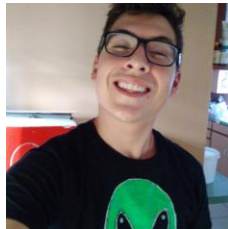
OMS. (2019). Organización Mundial de la Salud. Obtenido de Suicidio: www.who.int/es/news-room/fact-sheets/detail/suicide.

OMS. (2019). Obtenido de Organización Mundial de la Salud: www.who.int/topics/depression/es/.

Información de los autores.



Ricardo de Jesús González Morales es estudiante del Instituto Tecnológico de Tuxtla Gutiérrez de la carrera de Ingeniería en Sistemas Computacionales, se especializa en el área de inteligencia artificial, desarrollo de proyectos web y sistemas embebidos.



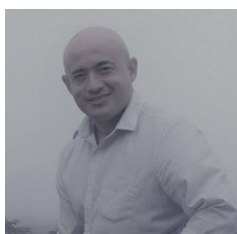
Rafael Antonio Bolaños Camaras es estudiante del Instituto Tecnológico de Tuxtla Gutiérrez de la carrera de Ingeniería en Sistemas Computacionales, se especializa en el área de Desarrollo Web, enfocado en frameworks web entre ellos Laravel, Django y Spring Boot.



Aída Guillermina Cossío Martínez es Maestra en Ciencias en Administración por el Instituto Tecnológico de Tuxtla Gutiérrez en 2002. Es profesora de tiempo completo del área de Ingeniería en Sistemas Computacionales del Instituto Tecnológico de Tuxtla Gutiérrez, desde 1994. Se especializa en la formulación y evaluación de proyectos, así como el emprendimiento y desarrollo de planes de negocio, actualmente es perfil deseable y trabaja en la línea de investigación Tecnología de Información y Base de Datos.



Héctor Guerra Crespo es egresado del I. T. de Mérida (Yucatán, México) de la carrera de Ingeniería en Sistemas Computacionales en 1994, es Doctor en Sistemas Computacionales por la Universidad del Sur (Chiapas, México) en 2011. Es profesor en el área de Ingeniería en Sistemas Computacionales del Instituto Tecnológico de Tuxtla Gutiérrez y en el área de Licenciatura en Sistemas Computacionales de la Universidad Autónoma de Chiapas, en ambas desde 1995. Es miembro del Claustro Doctoral "Doctorado en Ciencias de la Ingeniería" perteneciente al Programa Nacional de Posgrados de Calidad, I.T. de Tuxtla Gutiérrez, desde 2016. Miembro del Sistema Nacional de Investigadores, nivel C (2024-2021). www.hectorguerracrespo.com.



Germán Ríos Toledo obtuvo el grado de Doctor en Ciencias de la Computación en 2019 por el Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) en Cuernavaca, Morelos, México. Actualmente, es profesor de tiempo completo en el Departamento de Computación del Tecnológico Nacional de México (campus Tuxtla Gutiérrez, Chiapas) en la Ingeniería en Sistemas Computacionales y en la Maestría en Ciencias en Ingeniería Mecatrónica. Su área de especialización es el Procesamiento del Lenguaje Natural, particularmente en el uso de información sintáctica como una característica para el análisis de estilo de escritura. Otras áreas de su interés incluyen el procesamiento y análisis de imágenes, audio y video por medio de Algoritmos de Aprendizaje Automático y Aprendizaje Profundo. Miembro del Sistema Nacional de Investigadores, nivel C (2024-2021).



Galdino Belizario Nango Solís obtuvo el grado de Doctor en Desarrollo Tecnológico por la Universidad de Ciencia y Tecnología Descartes, maestro en Ciencias de la Computación por CIC-IPN, ha sido jefe del Centro de Cómputo y también del Departamento de Recursos Humanos del Instituto Tecnológico de Tuxtla Gutiérrez. Actualmente es docente en la carrera de Ingeniería en Sistemas Computacionales del Tecnológico Nacional de México campus I.T. de Tuxtla Gutiérrez, profesor de la maestría en Tecnologías de la Información en la Universidad Pablo Guardado Chávez. Con campos de interés en lenguajes de programación, sistemas de información y bases de datos.