ISBN: 978-628-95207-4-3, ISSN: 2414-6390, Digital Object Identifier: 10.18687/LACCEI2023.1.1.1308

Estratificación para Mejorar el Rendimiento de una ANN en la Detección de Diabetes. Stratification for the Improvement of the Performance of an ANN in Diabetes Detection.

Darwin Patiño-Pérez,Ph.D[1], Felicísimo Iñiguez-Muñoz,MSc[2],Ángel Ochoa-Flores,MSc[1], José Córdova-Aragundi,MATI[3], José Castro-Carrasco,MSc[4], Alex Luque-Letechi,MGPC[3], Celia Munive-Mora,BS[1,5,6],

[1]Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Física, Ecuador,darwin.patinop@ug.edu.ec, angel.ochoaf@ug.edu.ec,celia.munivem@ug.edu.ec,

[2]Bitekso S.A, Guayaquil,Ecuador,bernardo.iniguez@bitekso.com,

[3]Universidad de Guayaquil, Facultad de Ciencias Económicas, Ecuador,jose.cordovaa@ug.edu.ec,alex.luquel@ug.edu.ec,

[4]Universidad de Guayaquil, Facultad de Ciencias Administrativas, Ecuador, jose.castroca@ug.edu.ec,

[5]St Luke's University Hospital Network, PA, Estados Unidos,celia.munive@sluhn.org,

[6]De Sales University,Center Valley, PA, Estados Unidos,cm3877@desales.edu

Resumen. – Uno de los problemas que se han detectado en la generalización de los modelos de machine learning, que se ha implementado para la detección de la diabetes mellitus en el centro integral de diagnóstico de salud privada de la ciudad de Guayaquil; que impiden alcanzar un rendimiento optimo del modelo de aprendizaje supervisado, para su puesta en producción en el centro integral, es el desbalanceo o desequilibrio de los datos en función del tipo de clase. Por tal motivo puesto que la red neuronal artificial obtuvo el rendimiento más aceptable, el presente estudio se centrará en la evaluación de la red neuronal artificial y de la mejora que se ha alcanzado al aplicarse la técnica de estratificación de los datos, que permite seleccionar de forma proporcional a los grupos de datos de una forma equilibrada, permitiendo elevar el rendimiento de la red neuronal artificial en un 3.39% pero con una reducción de perdida muy alta de aproximadamente el 99.98% durante el aprendizaje.

Palabras Claves: Diabetes, Estratificación, Machine Learning, Aprendizaje Profundo, Redes Neuronales.

Abstract. – One of the problems that have been detected in the generalization of machine learning models, which has been implemented for the detection of diabetes mellitus in the comprehensive private health diagnostic center of the city of Guayaquil; What prevents achieving optimal performance of the supervised learning model, for its implementation in the comprehensive center, is the imbalance or imbalance of the data depending on the type of class. For this reason, since the artificial neural network obtained the most acceptable performance, this study will focus on the evaluation of the artificial neural network and the improvement that has been achieved by applying the data stratification technique, which allows selecting from proportionally to the data groups in a balanced way, allowing to increase the performance of the artificial neural network by 3.39% but with a very high loss reduction of approximately 99.98% during learning.

Keywords: Diabetes, Stratification, Machine Learning, Deep Learning, Neural Networks.

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).

I. INTRODUCCION

El *machine learning* (ML) o aprendizaje automático es un subcampo de la inteligencia artificial (IA) que se enfoca en el desarrollo de algoritmos y modelos que permiten a una máquina aprender y mejorar a través de la experiencia[1]. En lugar de ser programada específicamente para realizar una tarea, una máquina de aprendizaje automático utiliza datos y algoritmos para aprender por sí misma y mejorar su rendimiento en una tarea específica[2]. El proceso de aprendizaje en el aprendizaje automático se basa en identificar patrones y relaciones en los datos de entrenamiento, y luego utilizar esos patrones para hacer predicciones o tomar decisiones en nuevas situaciones [3]. El aprendizaje automático se aplica en una amplia variedad de campos, incluyendo reconocimiento de voz y de imágenes, análisis de datos, sistemas de recomendación, detección de fraude, entre otros. La eficacia del aprendizaje automático depende en gran medida de la calidad y cantidad de los datos de entrenamiento, así como de la elección del algoritmo adecuado para el problema que se está resolviendo.

Con el uso del ML se han resuelto muchos problemas del mundo real en diversas áreas de aplicación en el campo del marketing, comercio, PLN, autonomía vehicular, robótica, cambio climático, reconocimiento de voz, video juegos y sobre todo en la medicina llegando a posicionar como uno de los mejores elementos tecnológicos en diversas áreas de la biomedicina. Su aplicación en el diagnostico de enfermedades como covid19[4] así como la diabetes mellitus[5], entre otras ha arrojado resultados muy prometedores. Hay varios tipos de aprendizaje en inteligencia artificial, los cuales se pueden clasificar en tres categorías principales: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo[6]. El aprendizaje supervisado es un enfoque en el que el modelo aprende a partir de un conjunto de datos etiquetados. El conjunto de datos consiste en entradas y salidas esperadas.

El objetivo del modelo es aprender una función que pueda predecir la salida esperada para una entrada dada[7].
Los algoritmos de aprendizaje supervisado incluyen regresión, clasificación y redes neuronales. Sin embargo, el aprendizaje no supervisado es un enfoque en el que el modelo aprende de un conjunto de datos no etiquetados[8]. El modelo busca patrones en los datos y los utiliza para crear grupos o clústeres de datos similares. El objetivo del modelo es encontrar patrones interesantes y estructuras en los datos que puedan ayudar a los humanos a entenderlos mejor. Los algoritmos de aprendizaje no supervisado incluyen agrupamiento, reducción de dimensionalidad y análisis de componentes principales. Por otra parte, el aprendizaje por refuerzo es un enfoque en el que el modelo aprende a través de la interacción con un entorno[9]. El modelo recibe recompensas o castigos por tomar ciertas acciones y su objetivo es aprender una política que maximice la recompensa acumulada. En general, el tipo de aprendizaje que se utiliza depende del problema que se está intentando resolver y de la naturaleza de los datos disponibles. Cada tipo de aprendizaje tiene sus propias fortalezas y debilidades, y es importante seleccionar el enfoque adecuado para obtener los mejores resultados.

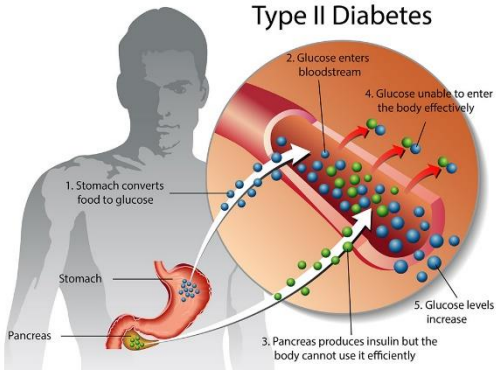


Fig.1 Diabetes Mellitus Tipo-2

Una de las principales enfermedades en el Ecuador es la diabetes que afecta la vida de más de 1.3 millones de ecuatorianos, lo que representa más del 7.5 % de la población[10]. Siendo la mas comun la denominada diabetes mellitus tipo-2 que se debe a una combinación de resistencia a la insulina y una producción insuficiente de insulina según la Fig.1. Esta forma de diabetes generalmente se desarrolla en la edad adulta y se asocia con factores de riesgo como la obesidad, la falta de actividad física y la alimentación poco saludable[12]. El tratamiento puede incluir cambios en el estilo de vida, medicamentos orales y/o inyecciones de insulina[11]. Los principales síntomas de la DM tipo-2 incluyen sed excesiva, micción frecuente[17], fatiga, hambre constante, visión borrosa, infecciones frecuentes y en donde las heridas tardan en sanar. Si no se trata la diabetes puede causar muchas complicaciones, como la cetoacidosis diabética y el coma hiperosmolar no cetósica[13].

Con este tipo de enfermedad es importante tener en cuenta que algunas personas con DM tipo-2 pueden no presentar síntomas en absoluto[14]. Por lo tanto, se recomienda que las personas mayores de 45 años se sometan a pruebas regulares de diabetes mellitus tipo-2 para detectar la enfermedad en sus primeras etapas. De hecho, si detecta los signos de un problema potencial en la etapa de prediabetes de la enfermedad, es posible que pueda detener la progresión antes de que desarrolle diabetes tipo-2[16]. Comience por familiarizarse con los factores de riesgo de la diabetes y los signos que debe observar que podrían indicar la aparición de la enfermedad.

Todavía no existe una cura para la diabetes, pero perder peso, comer alimentos saludables y mantenerse activo realmente puede ayudar, al igual de tomar los medicamentos según sea necesario, obtener educación y apoyo para el autocontrol de la diabetes y asistir a las citas de atención médica también pueden reducir el impacto de la diabetes en su vida[18]. La diabetes es una condición de salud crónica de larga duración que afecta la forma en que su cuerpo convierte los alimentos en energía, la mayor parte de los alimentos que se consumen se descomponen en azúcar (también llamada glucosa) y se liberan en el torrente sanguíneo, cuando el nivel de azúcar en la sangre sube, le indica al páncreas que libere insulina; la insulina actúa como una llave para permitir que el azúcar en la sangre ingrese a las células de su cuerpo para usarla como energía[19].

En un estudio previo[20], se implementaron algunos modelos de aprendizaje automático entre los que destaco el rendimiento de un modelo de predicción basado en una red neuronal artificial o *artificial neural network* (ANN) la misma que obtuvo una exactitud del 93.3% en el mejor de los escenarios, en vista que la puesta en marcha requiere que su precisión sea la más alta se ha aplicado un técnica basada en la estratificación de los datos para garantizar que la precisión a la hora de detectar DM sea la mejor.

II. MATERIALES Y METODOS

A. Dataset

TABLA I
CARACTERÍSTICAS Y ETIQUETA

Variables	Descripción	Unidad
Embarazos	numero de veces de embarazo	-
Sexo	sexo M(1),F(0)	-
Glucosa	concentración de la glucosa	ml/dl
PresionSanguinea	presion arterial diastólica	mm.Hg
PliegueCutaneo	grosor pliegue cutaneo triceps	mm
Insulina	insulina sérica a 2-horas	µU/ml
IndiceDeMasaCorporal	índice de masa corporal	kg/m^2
PedigriDiabetesFuncion	función pedigrí diabetes	-
Edad	edad de la persona en años	-
Etiqueta		
DiabetesResultado	Si(1),No(0)	-

El conjunto de datos para el diagnóstico de pacientes con DM tipo-2 fue recopilado de varios centros de salud privados de la ciudad de Guayaquil en Ecuador. Se conformo un *dataset* que contiene 2768 registros de pacientes cuyas características son: pregnant_times, glucose, blood_pressure , tst, insulin , bmi , dpf , age , is_diabetic(etiqueta). La Tabla I, muestra la descripción de unidades y rangos de atributos de riesgo del conjunto de datos.

Para la implementación de las técnicas de aprendizaje de maquina supervisado se utilizará Python[20] como herramienta de programación que se ejecuta sobre una máquina virtual de Google llamada Colab[21] que está configurada con todas las librerías requeridas para el uso de machine learning y deep learning, además se necesitará una conexión a internet con un amplio ancho de banda para la interacción con colab.

B. *Stratify* - Estratificación

Es un término utilizado en *machine learning* para referirse a una técnica de muestreo que se utiliza para garantizar que las proporciones de las clases en los datos de entrenamiento y prueba sean lo más similares posible. La estratificación se utiliza para abordar problemas de desequilibrio de clases en los datos, donde una o varias clases tienen una representación significativamente menor que las demás. La estratificación se aplica durante la partición de los datos en conjuntos de entrenamiento y prueba. En lugar de simplemente dividir aleatoriamente los datos en un conjunto de entrenamiento y un conjunto de prueba, la estratificación asegura que la proporción de cada clase en los datos originales se mantenga en ambos conjuntos.

Por ejemplo, si un conjunto de datos tiene un 80% de instancias de la clase A y un 20% de instancias de la clase B, la estratificación asegura que el conjunto de entrenamiento y el conjunto de prueba también tengan aproximadamente un 80% de instancias de la clase A y un 20% de instancias de la clase B. La estratificación es importante en situaciones en las que una clase minoritaria es de interés especial y es importante garantizar que el modelo tenga suficientes ejemplos de esta clase para aprender y generalizar correctamente. La estratificación también ayuda a prevenir la sobreestimación de la precisión del modelo en casos de clases desequilibradas.

La estratificación se basa en la definición de una función de densidad de probabilidad conjunta que describe la distribución de las variables de entrada y de la variable de salida. Esta función se puede escribir como:

$$f(x, y) = f(x) * f(y \mid x) \quad (1)$$

donde:

x es el conjunto de variables de entrada
y es la variable de salida
f(x, y) es la función de densidad de probabilidad conjunta de x e y
f(x) es la función de densidad de probabilidad marginal de x
f(y | x) es la función de densidad de probabilidad condicional de y dado x.

La estratificación se puede lograr aplicando esta función de densidad de probabilidad conjunta en cada conjunto de datos de entrenamiento y prueba. Para cada conjunto, se debe asegurar que las proporciones de las clases en la variable de salida sean similares a las proporciones en el conjunto original.

Para lograr esto, se puede utilizar un enfoque de muestreo estratificado, en el cual se divide el conjunto de datos original en diferentes estratos basados en la variable de salida y se extrae una muestra de cada estrato para formar los conjuntos de entrenamiento y prueba. La proporción de cada clase en la variable de salida se mantiene en los dos conjuntos.

C. Marco Teórico

La Neurona Artificial

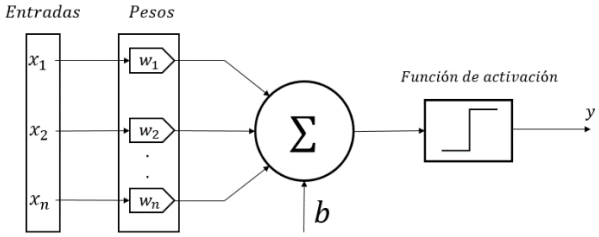


Fig.2 Neurona Artificial

Una neurona artificial como en Fig.2 es una unidad básica de procesamiento en una red neuronal artificial[34]. Se modela matemáticamente como una función que recibe una o más entradas, realiza un cálculo y produce una salida según (1). La entrada a una neurona artificial es ponderada por un conjunto de pesos, que representan la fuerza relativa de cada entrada en el cálculo de la salida. Además, la neurona puede tener un sesgo (bias) que permite ajustar la salida en función de un valor de referencia[35].

$$y = f(w1x1 + w2x2 + ... + wnxn + b) \quad (2)$$

donde:

x1, x2, ..., xn son las entradas a la neurona.
w1, w2, ..., wn son los pesos asignados a cada entrada.
b es el sesgo (bias) de la neurona.
f es la función de activación que se aplica al resultado de la suma ponderada de las entradas y los pesos, más el sesgo.

La salida de una neurona artificial generalmente se procesa por otras neuronas en la red neuronal artificial, y así sucesivamente hasta llegar a la salida final de la red. Este proceso se conoce como propagación hacia adelante (*forward propagation*).

Existen diferentes tipos de funciones de activación que se pueden utilizar en las neuronas artificiales, como la función sigmoide, la función ReLU (Rectified Linear Unit), entre otras. Cada una de estas funciones de activación tiene sus propias ventajas y desventajas en términos de la capacidad de la red para aprender y generalizar a partir de los datos de entrenamiento.

La función de activación se utiliza para introducir no-linealidad en el cálculo de la salida de la neurona como por ejemplo según (2) o (3), lo que permite a la red neuronal aprender relaciones más complejas entre las entradas y las salidas.

La función de activación más comúnmente utilizada es la función sigmoide, que se define como:

$$f(z) = 1 / (1 + e^{(-z)}) \quad (3)$$

donde z es la suma ponderada de las entradas y los pesos más el sesgo.

Otra función de activación popular es la función ReLU (Rectified Linear Unit), que se define como:

$$f(z) = \max(0, z) \quad (4)$$

donde z es la suma ponderada de las entradas y los pesos más el sesgo.

La Red Neuronal Artificial (ANN)

Una red neuronal artificial o ANN es un modelo computacional inspirado en la estructura y funcionamiento del cerebro humano. Consiste en un conjunto de unidades básicas de procesamiento llamadas neuronas artificiales ver Fig.3, conectadas en capas, que trabajan juntas para resolver problemas de aprendizaje y clasificación[36].

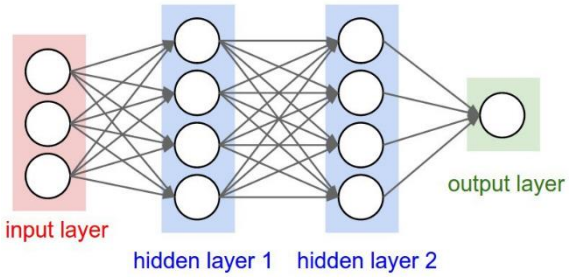


Fig. 3 Red Neuronal Artificial

Cada neurona artificial recibe una o varias entradas, que son ponderadas por un conjunto de pesos y sumadas con un sesgo (bias). La salida de cada neurona se procesa como entrada para las neuronas de la capa siguiente, y así sucesivamente, hasta que se obtiene la salida final de la red. La estructura de la red neuronal se define por el número y la disposición de las capas, así como por el número de neuronas en cada capa y la forma en que se conectan entre sí. Las redes neuronales pueden tener varias capas ocultas, que les permiten aprender características más complejas y abstractas de los datos de entrada[38].

Las redes neuronales se entrenan mediante un proceso de aprendizaje supervisado, en el que se presenta un conjunto de datos de entrenamiento etiquetados a la red, y se ajustan los pesos y sesgos de las neuronas para minimizar la diferencia entre las salidas de la red y las etiquetas de los datos de entrenamiento. Una vez que se entrena la red, se puede utilizar para hacer predicciones sobre datos nuevos y no vistos[39]. Las redes neuronales se han utilizado con éxito en una variedad de aplicaciones, como reconocimiento de voz, reconocimiento de objetos en imágenes, procesamiento de lenguaje natural, y muchas otras áreas de la inteligencia artificial y la ciencia de datos.

Tipos de Redes Neuronales Artificiales (ANN)

Existen varios tipos de redes neuronales artificiales, cada una diseñada para resolver problemas específicos. Aquí se presentan algunos de los tipos más comunes:

Redes neuronales recurrentes: estas redes tienen conexiones bidireccionales que permiten que la información fluya hacia adelante y hacia atrás. Son útiles para procesar secuencias de datos, como en el procesamiento del lenguaje natural, el reconocimiento del habla y la predicción de series temporales.

Redes neuronales convolucionales: son especialmente adecuadas para el procesamiento de datos estructurados, como imágenes y videos. Utilizan capas convolucionales para extraer características importantes de las imágenes y reducir su complejidad.

Redes neuronales de propagación hacia atrás: se utilizan para el aprendizaje supervisado y se basan en la optimización de una función de pérdida mediante el ajuste de los pesos de las neuronas.

Redes neuronales generativas adversarias (GANs): se utilizan para generar datos sintéticos que se parecen a los datos de entrada. Las GANs utilizan dos redes neuronales que compiten entre sí para mejorar la calidad de los datos generados.

Redes neuronales feedforward: son las redes neuronales más simples y comunes. La información fluye en una dirección, desde la entrada a la salida, a través de una o más capas ocultas. Estas redes se utilizan para tareas como la clasificación, la regresión y el reconocimiento de patrones; se puede representar una red neuronal feedforward como una composición de funciones matemáticas.

Sea X un vector de entrada, Y un vector de salida, y f(x) una función que transforma x en una representación interna z:

$$z = f(x) \quad (5)$$

Luego, se pueden aplicar capas adicionales de funciones de transformación lineales y no lineales, denotadas como W y g, respectivamente, para producir la salida:

$$y = g(Wz) \quad (6)$$

donde W es la matriz de pesos que conecta las neuronas de la capa anterior con las de la capa actual, y g es la función de activación que se aplica a la suma ponderada de las entradas.

Estas capas se organizan en una estructura de capas, con la capa de entrada como la primera capa, la capa de salida como la última capa, y las capas ocultas en el medio. Cada capa puede tener un número variable de neuronas, y las conexiones entre las capas están completamente conectadas.

El proceso de entrenamiento de una red neuronal feedforward implica ajustar los pesos de las conexiones para minimizar una función de pérdida, que mide la diferencia entre la salida producida por la red neuronal y la salida deseada. Esto se realiza mediante un algoritmo de optimización, como el descenso del gradiente, que ajusta gradualmente los pesos de las conexiones para mejorar la precisión de la red neuronal.

Redes neuronales autoencoder: son redes feedforward que se utilizan para aprender representaciones de datos. Son útiles para la compresión de datos, la eliminación de ruido y la reconstrucción de datos faltantes.

Cada tipo de red neuronal tiene sus propias fortalezas y debilidades, y es importante seleccionar el tipo correcto para la tarea en cuestión. Además, muchas veces se combinan diferentes tipos de redes neuronales para crear sistemas más complejos y robustos.

Métricas de Evaluación

TABLA II
MATRIZ DE CONFUSIÓN

		VALOR-PREDICCION	
		0 (Negativo)	**1** (Positivo)
VALOR-REAL	**0**	(**TN**) Verdadero Negativo El valor real es negativo y predijo un valor negativo.	(**FP**) Falso Positivo El valor real es negativo y predijo un valor positivo.
	1	(**FN**) Falso Negativo El valor real es positivo y predijo un valor negativo.	(**TP**) Verdadero Positivo El valor real es positivo y predijo un valor positivo.

Matriz de Confusión. – Según la Tabla II, la matriz de confusión[40] es el elemento sobre el cual se basan todas las métricas de clasificación, en ella se agrupan los valores clasificados por un determinado modelo (0) es *Negative* y (1) es *Positive*. Cuando el modelo clasifica adecuadamente se tienen dos valores. Verdaderos Positivos, cuando el modelo ha predicho que SI y en realidad SI. Verdaderos Negativos, aquí el modelo ha predicho que NO y en realidad es un NO. Cuando no ha clasificado adecuadamente se tienen los siguientes valores. Falso Positivo, cuando el modelo ha predicho que SI y en realidad es un NO. Falso Negativo, es cuando el modelo ha predicho que NO pero en realidad es SI.

En este estudio, se emplean dos técnicas de evaluación para determinar el desempeño de cada modelo de aprendizaje predictivo desarrollado basado en varios algoritmos de ML supervisados. Estas técnicas incluyen lo siguiente:
La precisión se utiliza para evaluar los modelos predictivos de ML supervisados. La precisión tiene la siguiente definición:

1)Accuracy.- *Accuracy* o exactitud es la cantidad de predicciones positivas que fueron correctas.

$$\text{Accuracy} = \frac{(TN+TP)}{(FP+TP)+(TN+FN)} \qquad (7)$$

2) La curva ROC o curva Característica Operativa del Receptor es el gráfico, que expone el rendimiento de un clasificador binario en función del umbral de corte, exponiendo la tasa de verdaderos positivos (TPR) contra la tasa de falsos positivos (FPR)[41] .

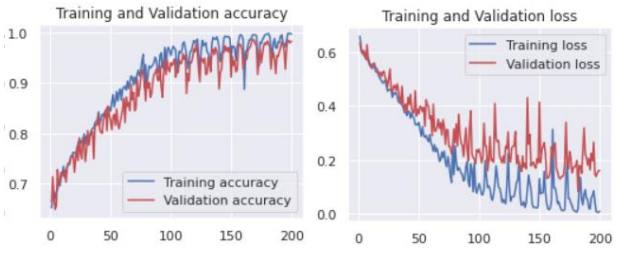
III. Resultado y Discusion

El trabajo se desarrolló bajo un enfoque cuantitativo, en la modalidad experimental de tipo exploratoria. Se ha utilizado una técnica de inteligencia artificial basada en aprendizaje automático o *machine learning* en la que por medio de aprendizaje profundo o *deep learning* se crearon algunos modelos de redes neuronales artificiales para probar la efectividad de la estratificación con un proceso que está enmarcado dentro de los siguientes pasos:

1)Tratamiento de los Datos
2)Creación del Modelo
3)Fase de entrenamiento
4)Fase de Prueba
5)Evaluación del modelo con sus métricas
6)Aplicación del modelo

Fase-I
Se tomo el dataset que contiene 2768 registros de información de pacientes repartidos en clases, hay 1816 pacientes con diabetes(1) y 952 pacientes sin diabetes(0), se procedió a usar un modelo con las nueve columnas de entrada sin ningún tipo de tratamiento y en otros modelos se tomaron las 6 columnas de entrada más significativas o relevantes las cuales fueron estandarizadas y escaladas, en todas las muestras se tomó el 80% para entrenamiento(train) y el 20% para pueba(test), además se tomó una muestra estratificada y otra no estratificada para realizar el comparativo.

Fase II
En esta fase se exponen las etapas que van desde la creación del modelo, entrenamiento, prueba, evaluación y aplicación de estos, por lo que destaca de todos los modelos la red neuronal artificial.
Modelo de ANN-1: Para este modelo, se usó el dataset original que tiene las 9 columnas de características de entrada de donde se tomó el 80% para train y 20% para test, sin aplicarse ningún proceso de estandarización y/o normalización y sin haberse aplicado la estratificación de los datos según se aprecia en la Fig.4.



. Fig.4 Modelo de ANN-1

Se obtuvo una exactitud del 95.85% con una pérdida de 0.2955 y aunque la precisión es muy buena hay mucho ruido según la Fig.7.

Modelo de ANN-2: Se seleccionaron de forma automática, las 6 columnas más relevantes ['Embarazos', 'Glucosa', 'Insulina','IndiceDMC','Pedigri','Edad'] y que son las que aportan de forma significativa al proceso de aprendizaje, las cuales fueron usadas como características de entrada, del total de registros se tomó el 80% para train y el 20% para test, los datos fueron estandarizados, pero no se les aplico el proceso de estratificación de los datos según la Fig.5.

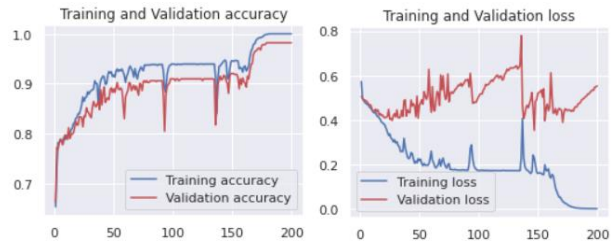


Fig.5 Modelo de ANN-2

Se obtuvo una exactitud del 98.19% con una pérdida de 0.3748 y aunque la precisión es muy buena todavía hay ruido según la Fig.8.

Modelo de ANN-3: Se seleccionaron de forma automática, las 6 columnas más relevantes que aportan significativamente al proceso de aprendizaje, las cuales fueron usadas como características de entrada, del total de registros se tomó el 80% para *train* y el 20% para test, los datos fueron estandarizados, y aquí ya se aplicó el proceso de estratificación de los datos.

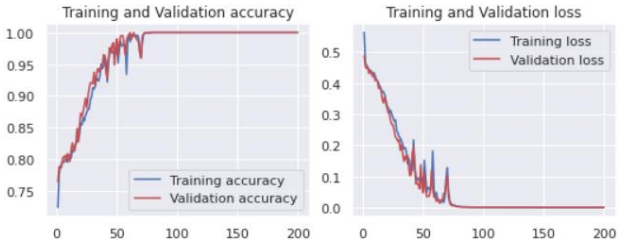


Fig.6 Modelo de ANN-3

Se obtuvo una exactitud del 100% con una pérdida de 0.000002 en la etapa de *train* y una exactitud del 99.1% con una pérdida del 0.114 en la etapa de *test*, tal como se observa en la Fig.6, por lo que el modelo a aprendido adecuamente.

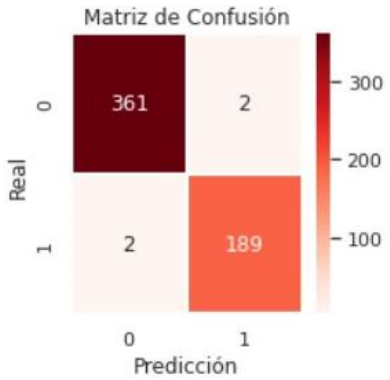


Fig.7 Matriz de Confusión de la ANN

La matriz de confusión según la Fig.7 fue muy importante para poder probar las métricas de clasificación, la matriz resultante está relacionada con el mejor de los modelos ANN-3.

IV. Discusion y Conclusion

La Fig.6 refleja la exactitud o *accuracy* que se ha alcanzado con el modelo de red neuronal artificial o ANN-3 la cual tiene exactitud del 99.1% en su etapa de prueba con una pérdida del 0.114 sin que se vea algún tipo de sobre entrenamiento dado que la etapa de *train* la perdida fue del 0.000002.

Se puede concluir que de los 3 modelos de ANN los modelos el modelo de la Fig.8 ha aprendido a realizar adecuadamente la clasificación.

```
#Creación del Modelo - ANN
model3 = Sequential()
model3.add(Dense(256,input_dim=6,kernel_initializer='uniform',activation='relu'))
model3.add(Dense(512,kernel_initializer='uniform',activation='relu'))
model3.add(Dense(64,kernel_initializer='uniform',activation='relu'))
model3.add(Dense(8,kernel_initializer='uniform',activation='relu'))
model3.add(Dense(1,kernel_initializer='uniform',activation='sigmoid'))

#Compilacion del Modelo
model3.compile(loss='binary_crossentropy', optimizer='adam',metrics=['accuracy'])
#Entrenamiento del Modelo
history3 = model3.fit(X_train3, y_train3, validation_data=(X_val, y_val),
                      epochs=200,batch_size=32, verbose=0)
```

Fig.8 Modelo ANN-3

Al realizarse el comparativo de los 3 modelos de ANN, se observa que el esquema de estatificación de datos se ha reducido la variabilidad dentro de cada clase, y que ha permitido mejorar la precisión de las estimaciones y reducir el margen de error de forma muy significativa.

El incremento de la exactitud entre la ANN-1 que tiene una exactitud del 95.85% y del ultimo modelo ANN-3 que tiene una exactitud del 99.1%, se nota que la exactitud de predicción aumento en un 3.39%.

Con el modelo ANN-3 se puede determinar, con un alto grado de confiabilidad, si un paciente tiene diabetes o no; el uso del predictor será de gran ayuda en al sector de la salud porque puede diagnosticar la presencia o no de la diabetes en un paciente en cualquier momento y a cualquier hora por que se lo recomienda para su implementación.

Referencias

[1] M. Varone, D. Mayer, and A. Melegari, “What is Machine Learning? A definition,” *Expert System*, 2020. .

[2] J. Cabanelas Omil, “Inteligencia artificial ¿Dr. Jekyll o Mr. Hyde?,” *Mercados y Negocios*, no. 40, pp. 5–22, 2019.

[3] A. Panesar, “What Is Machine Learning?,” in *Machine Learning and AI for Healthcare*, 2021.

[4] D. P. Pérez, R. S. Bustillos, C. M. Mora, and M. Botto-Tobar, “Prediction of Covid19 with the use of Random Forests Algorithm and Artificial Neural Networks,” *Ecuadorian Sci. J.*, vol. 4, no. 2, pp. 101–110, Sep. 2020.

[5] M. E. Baldeón *et al.*, “Prevalence of metabolic syndrome and diabetes mellitus type-2 and their association with intake of dairy and legume in Andean communities of Ecuador,” *PLoS One*, vol. 16, no. 7 July, 2021.

[6] J. Luna, “Tipos de aprendizaje automático,” *Medium*, 2018.

[7] V. Roman, “Aprendizaje No Supervisado en Machine Learning: Agrupación | by Victor Roman | Ciencia y Datos | Medium,” *Medium*, 2019.

[8] Mauricio Arango, "Introducción al Aprendizaje por Refuerzo," *Oracle A-Team*, no. August, 2019.
[9] C. González-García, "En qué consiste el aprendizaje automático (machine learning) y qué está aportando a la Neurociencia Cognitiva," *Cienc. Cogn.*, vol. 12, no. 2, 2018.
[10] C. Urrea and A. Mignogna, "Development of an expert system for pre-diagnosis of hypertension, diabetes mellitus type 2 and metabolic syndrome," *Health Informatics J.*, vol. 26, no. 4, 2020.
[11] F. Garmendia-Lorena, "El tratamiento actual de la Diabetes Mellitus Tipo 2," *Diagnóstico*, vol. 59, no. 1, 2020.
[12] R. N. Fatimah, "Diabetes Melitus Tipe 2," *Fak. Kedokt. Univ. Lampung*, vol. 4, 2015.
[13] F. Gómez Rodríguez, P. Ruiz Alcantarilla, L. Rodríguez Félix, A. Martín Santana, and E. Zamora Madaria, "Alteración del receptor Fc(IgG) de los monocitos en la cetoacidosis diabética y el coma hiperosmolar no cetósico.," *Med. Clin. (Barc).*, vol. 84, no. 1, 1985.
[14] J. Vlaški and I. Vorgučin, "Diabetes mellitus type 2 in children and adolescents," *Paediatr. Croat. Suppl.*, vol. 63, 2019.
[15] J. R. Vielma Guevara, J. del C. Villarreal Andrade, and L. V. Gutiérrez Peña, "Pandemia por el SARS-CoV-2: aspectos biológicos, epidemiológicos y clínicos," *Observador del Conocimiento. Revista Especializada de Gestión Social del Conocimiento*, vol. 5, no. 3. 2020.
[16] X. Liu, S. Wu, Q. Song, and X. Wang, "Reversion from pre–diabetes mellitus to normoglycemia and risk of cardiovascular disease and all-cause mortality in a chinese population: A prospective cohort study," *J. Am. Heart Assoc.*, vol. 10, no. 3, 2021.
[17] E. Menéndez, "TRATAMIENTO NO FARMACOLÓGICO EN LA ENFERMEDAD RENAL POR DIABETES," *Rev. la Soc. Argentina Diabetes*, vol. 51, no. 3, 2018.
[18] A. Y. Forero, J. A. Hernández, S. M. Rodríguez, J. J. Romero, G. E. Morales, and G. Á. Ramírez, "La alimentación para pacientes con diabetes mellitus de tipo 2 en tres hospitales públicos de Cundinamarca, Colombia," *Biomédica*, vol. 38, no. 3, 2018.
[19] E. G. Blanco Naranjo, G. F. Chavarría Campos, and Y. M. Garita Fallas, "Insulinización práctica en la diabetes mellitus tipo 2," *Rev. Medica Sinerg.*, vol. 6, no. 1, 2021.
[20] D. Patiño-pérez *et al.*, "Modelos de Machine Learning basados en Aprendizaje Supervisado para la Detección de Diabetes Mellitus en la Ciudad de Guayaquil . Machine Learning Models based in Supervised Learning for the Detection of Diabetes Mellitus in the City of Guayaquil .," *LACCEI*, pp. 1–8, 2022.
[21] G. Bentacourt, "Las Máquinas de Soporte Vectorial (SVMs)," *Sci. Tech.*, vol. 1, no. 27, 2005.
[22] J. Rodrigo, "Máquinas de Vector Soporte (Support Vector Machines, SVMs)," *Cienciadedatos*, 2017.
[23] Z. Zhang, "Introduction to machine learning: K-nearest neighbors," *Ann. Transl. Med.*, vol. 4, no. 11, 2016.
[24] S. Kang, "K-nearest neighbor learning with graph neural networks," *Mathematics*, vol. 9, no. 8, 2021.
[25] N. Ali, D. Neagu, and P. Trundle, "Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets," *SN Appl. Sci.*, vol. 1, no. 12, 2019.
[26] P. Cunningham and S. J. Delany, "K-Nearest Neighbour Classifiers-A Tutorial," *ACM Computing Surveys*, vol. 54, no. 6. 2021.
[27] D. Patiño Pérez, R. Silva Bustillos, C. Munive Mora, and M. Botto-Tobar, "Predicción de Covid19 con el uso del Algoritmo Random Forest y Redes Neuronales Artificiales," *Ecuadorian Sci. J.*, vol. 4, no. 2, 2020.
[28] L. Breiman, "Randon Forests," *Machinelearning202.Pbworks.Com*, 1999.
[29] V. Roman, "Algoritmos Naive Bayes: Fundamentos e Implementación," *Ciencia y Datos*, 2019. .
[30] A. V. Konstantinov and L. V. Utkin, "Interpretable machine learning with an ensemble of gradient boosting machines," *Knowledge-Based Syst.*, vol. 222, 2021.
[31] O. Sprangers, S. Schelter, and M. De Rijke, "Probabilistic Gradient Boosting Machines for Large-Scale Probabilistic Regression," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2021.
[32] J. C. Giraldo Mejía and F. A. Vargas Agudelo, "Aplicación de la técnica regresión logística de la minería de datos en el proceso de descubrimiento de conocimiento (KDD) en bases de datos operativas o transaccionales," *Perspectiv@*, vol. 14, no. 13, 2019.
[33] C. Barrionuevo, J. Ierache, and I. Sattolo, "Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística," *XXVI Congr. Argentino Ciencias la Comput.*, pp. 491–500, 2020.
[34] J. Archila Rodriguez, "La neurona," *LA Neurona*, 2017.
[35] J. A.-T. Barrera, "Redes Neuronales Artificiales," *Univ. Guadalajara*, p. 276, 2016.
[36] B. Martín del Brío and C. Serrano Cinca, "Fundamentos de redes neuronales artificiales: hardware y software," *Scire Represent. y Organ. del Conoc.*, 1995.
[37] A. Novales, "Estimación de modelos No Lineales," *Dep. Econ. Cuantitativa Univ. Complut.*, 2016.
[38] O. Agasi, J. Anderson, A. Cole, M. Berthold, M. Cox, and D. Dimov, "What is an Artificial Neural Network (ANN)? - Definition from Techopedia," *Techopedia*. 2018.
[39] D. Patiño Perez, R. Silva Bustillos, M. Botto-Tobar, and C. Munive Mora, "Análisis de Imágenes de Rayos X por Medio de Redes Neuronales Artificiales," *Ecuadorian Sci. J.*, vol. 5, no. 1, 2021.
[40] P. Recuero de los Santos, "Machine Learning a tu alcance: La matriz de confusión - Think Big Empresas," *Luca Telefonica Data Unit*, 2018. .
[41] A. A. Osi *et al.*, "A classification approach for predicting COVID-19 Patient's survival outcome with machine learning techniques," *medRxiv*, 2020.