

# Prediagnóstico Médico de la Diabetes Mellitus tipo 2 mediante Machine Learning

Herrería, Julio <sup>1</sup>; Bonilla, Vladimir <sup>1,2</sup>; Mosquera, Guillermo <sup>1(\*)</sup>; Sánchez, Miguel <sup>1</sup>; Andrade, Cristina <sup>1</sup>

<sup>1</sup>Universidad UTE, Facultad de Ciencias de la Ingeniería e Industria, Quito, Ecuador

<sup>2</sup>Universidad Yachay Tech, Escuela de Ciencias Biológicas E ingeniería, Urcuquí, Ecuador

**Resumen:** En el artículo se presenta el desarrollo un sistema para prediagnosticar la diabetes mellitus tipo 2. El proceso de prediagnóstico consta de tres etapas medición; anamnesis, examen físico y un examen complementario. Como resultado se da a conocer la probabilidad de padecer diabetes mellitus tipo 2 mediante redes neuronales que utilizan como entradas: edad, género, altura, peso, circunferencia abdominal, antecedentes familiares relacionados a la diabetes, patología y estado de gestación de ser el caso. Se utilizaron perceptrones para realizar la clasificación de los patrones y los resultados se validaron con la cuadrícula de análisis consensuado de errores de Clarke, permitiendo obtener un sistema de prediagnóstico no invasivo con una probabilidad de acierto no menor de 90%.

**Palabras clave:** Machine Learning. Prediagnóstico. Diabetes mellitus tipo 2. Redes neuronales artificiales. Cuadrícula de errores de Clarke.

Recibido: 28 de noviembre de 2022. Aceptado: 20 de febrero de 2023

Received: November 28th, 2022. Accepted: February 20th, 2023

## Medical Prediagnosis of Type 2 Diabetes Mellitus using Machine Learning

**Abstract:** The article presents the development of a system to pre-diagnose type 2 diabetes mellitus. The pre-diagnosis process consists of three measurement stages; anamnesis, physical examination and a complementary examination. As a result, the probability of suffering from type 2 diabetes mellitus is disclosed through neural networks that use as inputs: age, gender, height, weight, abdominal circumference, family history related to diabetes, pathology and state of pregnancy if applicable. Perceptrons were used to classify the patterns and the results were validated with Clarke's consensus error analysis grid, making it possible to obtain a non-invasive pre-diagnosis system with a probability of success of no less than 90%.

**Keywords:** Machine Learning. Prediagnosis. Diabetes mellitus type 2. Artificial neural networks. Clarke's error grid

---

(\*) guillermo.mosquera@ute.edu.ec

## 1. INTRODUCCIÓN

La diabetes mellitus tipo 2 es una enfermedad que consiste en niveles altos de glucosa en sangre, y en la actualidad considerada como enfermedad crónica. Dicho nivel alto de glucosa se debe a dos razones; las células beta del páncreas no producen la insulina (hormona que regula la glucosa en sangre) suficiente o las células del organismo se resisten a asimilar la insulina y por consecuencia no pueden absorber la glucosa en sangre (Maitra, 2015).

Según la Federación Internacional de Diabetes, en su informe en el 2019 hubo alrededor de 463 millones de personas diabéticas, más del 50% de casos eran diabéticos tipo 2. Se estiman que dichas cifras vayan en aumento, por el estilo de vida sedentario de las personas (International Diabetes Federation, 2019).

Datos del Instituto Nacional de Estadísticas y Censos del Ecuador indican que la diabetes es una de las principales causas de muerte general del país, y según la Organización Mundial de la Salud en 2012 la diabetes fue la causante de 3,7 millones de muerte en el mundo (Organización Mundial de la Salud, 2019).

Cuando se padece diabetes mellitus tipo 2, se tiene complicaciones agudas y crónicas. Además, que el padecimiento de la enfermedad es considerado como factor de riesgo o agravante para otras enfermedades. Algunas de estas complicaciones se muestran en la Tabla 1:

**Tabla 1.** Complicaciones de la diabetes mellitus tipo 2

Complicaciones Agudas	Complicaciones Crónicas
Cetoacidosis diabética	Microangiopatías
	Complicaciones macrovasculares
Hipoglucemia	Deterioro visual
Síndrome Hiperosmótico	Polineuropatía simétrica distal
Hiperosmolar	Infecciones cutáneas

## 2. MARCO TEÓRICO

Debido al riesgo que se presenta el padecer diabetes mellitus tipo 2, es importante su diagnóstico prematuro con la finalidad de evitar o minimizar el desarrollo de esta. La enfermedad posee distintos factores de riesgo, algunos de los cuales se nombra a continuación:

- Análisis de glucosa en ayunas (mayor a 126 mg/dl indica diabetes, entre 100-125 mg/dl indica prediabetes e inferior a 100 mg/dl, no presenta alteraciones de la glucosa).
- Edad (edad avanzada incrementa probabilidad de padecer diabetes).
- Género (sexo masculino propenso a desarrollo de diabetes).
- Antecedentes familiares (el aspecto genético incrementa el riesgo).
- Actividad física (estilo de vida influye en el desarrollo de la enfermedad).
- Índice de masa corporal (malos hábitos alimenticios incrementan el riesgo).

- Perímetro de cintura (indicador de distribución de grasa abdominal)

Si la persona padece diabetes mellitus tipo 2, el tratamiento más común se basa en suministrar insulina mediante inyecciones por vía subcutánea o por bomba de infusión continua. La implementación de estos tratamientos puede ocasionar una hipoglucemia, que consiste en niveles bajo de glucosa en sangre (Gobierno de México secretaria de Salud, 2014).

Otra alternativa es el reemplazo de células  $\beta$  por células madre, con la finalidad de promover la producción de insulina para el organismo. Algunos proyectos prometedores se basan en la creación de tejidos vivos, para la producción de un páncreas artificial.

Los pacientes diabéticos con el fin poder controlar la enfermedad, practican un estilo de vida con una dieta equilibrada y actividad física. Siendo esta una de las mejores maneras de evitar complicaciones a futuro. Todas las complicaciones y futuros tratamientos pueden ser detectados en una etapa temprana mediante la implementación de sistemas inteligentes de análisis de datos basados en redes neuronales. (Santander, 2019)

Algunos autores (Meza, 2018) sugieren el uso de lógica difusa para el prediagnóstico, obteniendo resultados menores al 80% de exactitud, mientras tanto las propuestas (Diagnóstico Asistido Por Inteligencia Artificial De Aprendizaje Profundo Para El Tamizaje Automático De Anormalidades De Fondo De Ojo De Múltiples Patologías, 2019) para el uso de redes neuronales y redes neuronales profundas evidencian resultados mayores.

## 3. METODOLOGÍA

El objetivo del presente proyecto es el diseño de un sistema de prediagnóstico médico no invasivo que permita determinar si el paciente padece o no de diabetes mellitus tipo 2. Para cumplir con este objetivo se utilizarán redes neuronales como herramienta de machine learning.

Las etapas para el diseño del sistema se describen a continuación.

### 3.1 Registro de datos de los pacientes y protocolo de la medición.

Se estableció como unidad de análisis un centro médico de la ciudad de Santo Domingo de los Tsáchilas (Ecuador), se creó un registro de datos con 50 casos clínicos variados (diabéticos, prediabéticos y no diabéticos), esto debido a las restricciones de la pandemia Covid-19. Además, se estableció un protocolo para las mediciones médicas, con la finalidad que se realicen bajo las mismas condiciones (Ministerio de Salud Pública del Ecuador, 2012). Entre los principales puntos a considerar en el protocolo se encuentran: solicitar al paciente estar en ayuno al menos de 8 horas; medir la altura de la persona manteniendo la espalda recta; con una balanza

digital obtener la masa corporal verificando que el paciente lleve ropa ligera; realizar la medición de glucosa con un glucómetro certificado y calibrado, realizar la medición del perímetro abdominal con una cinta métrica verificando que el paciente no comprima el abdomen, y anamnesis con preguntas claras. Adicional, con el fin de estandarizar las mediciones no invasivas sobre los pacientes, se diseñó un estand que permite obtener la altura, peso y diámetro abdominal

### 3.2 Diseño y entrenamiento de las redes neuronales artificiales.

Se establecieron tres redes neuronales con aprendizaje supervisado, los conjuntos de entrenamiento fueron generados por un médico experto en diabetes mellitus tipo 2 a partir de los registros de 50 casos clínicos. El 80% (40 registros) de los datos registrados se utilizaron para el entrenamiento de las redes neuronales, y el 20% restante se aplicó para pruebas de funcionamiento y ajuste de las redes neuronales. Los conjuntos de entrenamiento tienen diferentes casos que permiten evitar el “overfitting” en sistema (Baradó, 2013).

Las redes neuronales seleccionadas para el presente proyecto son perceptrones simples con una sola capa de entrada y una neurona en la salida. Las funciones de activación se seleccionaron de acuerdo a la naturales y tendencia de los datos del conjunto objetivo. Su arquitectura se muestra a continuación:

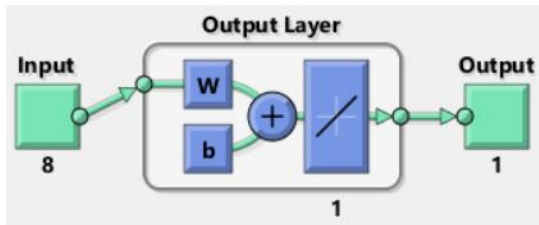


Figura 1. Arquitectura de las redes neuronales

La primera red neuronal estima la circunferencia abdominal del paciente, teniendo como entrada los parámetros de altura y masa corporal. Se seleccionó un perceptrón con el algoritmo Levenberg – Marquardt [1] con propagación inversa del error y función de activación lineal. En la figura 2 se muestra el ajuste de la red neuronal durante el entrenamiento. Similar comportamiento se puede observar para las otras dos redes neuronales.

$$x_i = x_{i-1} - \alpha (\epsilon I + H_f)^{-1} * \nabla f \quad (1)$$

- $\alpha$  y  $\epsilon$  son positivos
- $I$  es la matriz de identidad
- $\epsilon I - H_f$  sean positiva

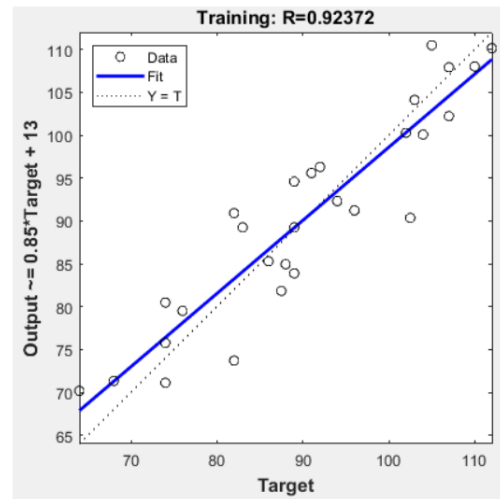


Figura 2. Ajuste de la red neuronal durante la etapa de entrenamiento

El ajuste de esta red neuronal es de 0.92372.

La segunda red estima los niveles de glucosa en sangre, teniendo como entrada los parámetros de edad, género, embarazo, antecedentes clínicos, actividad física, altura, masa corporal, y mediciones del sensor no invasivo de glucosa. Para la activación de la red se utilizó la función tansig.

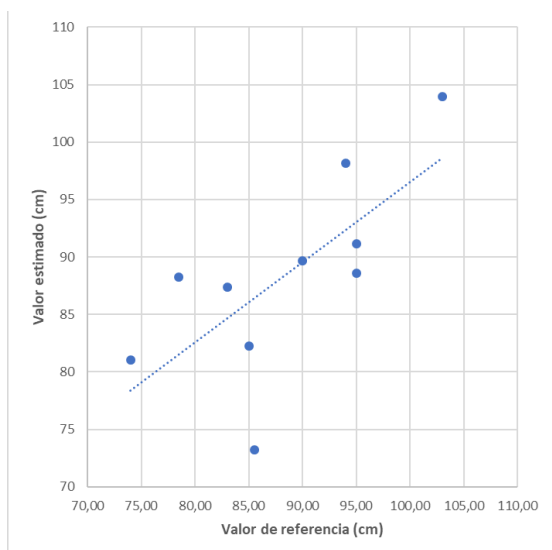
La última red se encarga de proporcionar el prediagnóstico por medio de una probabilidad obtenida de los datos de entrada de edad, género, embarazo, antecedentes familiares, actividad física, IMC, nivel de glucosa y circunferencia abdominal. Los dos últimos datos son proporcionados de las dos redes neuronales anteriores conectadas en cascada. En esta red se consideró el escalamiento y estandarización de los parámetros de entrada debido a la diferencia en sus unidades (American Diabetes Association, 2019).

## 4. RESULTADOS Y/O DISCUSIÓN

Los resultados mostrados a continuación se obtuvieron con los registros y diagnósticos de 10 pacientes (20% de la muestra). Estos datos no formaron parte del conjunto de entrenamiento de las redes neuronales, pero si del conjunto de validación.

### 4.1 Resultados de la estimación de circunferencia abdominal.

En la figura 2 se muestran los resultados de estimación de la circunferencia abdominal mediante el uso de un perceptrón con función de activación *purelin*.

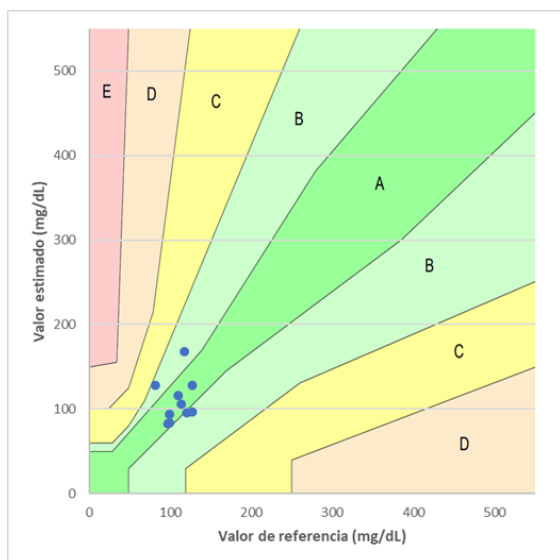


**Figura 2.** Coeficiente de correlación entre los datos de objetivo y de salida de la circunferencia abdominal - Prueba

En el experimento se obtuvo una exactitud del 93,87%, con un error promedio absoluto de 5,19 cm.

#### 4.2 Resultados de estimación de glucosa.

El resultado de la estimación de glucosa con la segunda red neuronal se validó mediante la cuadrícula de errores consensuada según la norma ISO 15197 que indica los requisitos para los sistemas de monitorización de glucosa en sangre para autodiagnóstico en la gestión de la diabetes mellitus, y se muestra en la figura 3.

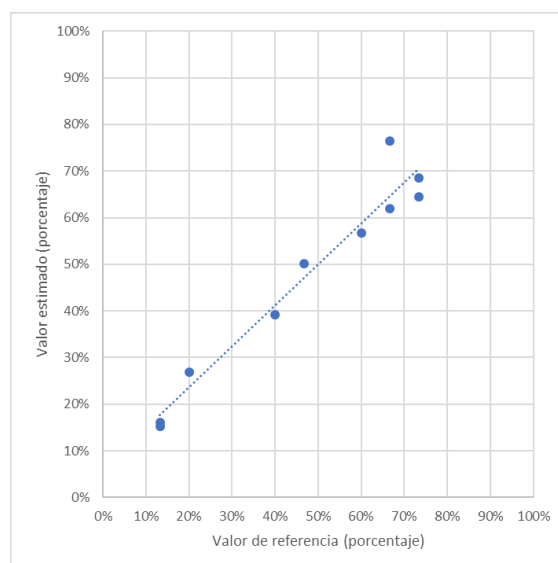


**Figura 3.** Cuadrícula de errores consensuada – 10 puntos dentro de la zona A y B.

Para que las estimaciones cumplan con la norma, es necesario que el 99% de las mediciones se encuentren dentro de las zonas A y B. En el caso de la red neuronal utilizada, el 100% de las estimaciones se encuentran dentro de los parámetros establecidos.

#### 4.3 Resultados de prediagnóstico

Los resultados del prediagnóstico mediante el uso de redes neuronales se muestran en la figura 4.



**Figura 4.** Correlación entre los datos objetivo y de salida del prediagnóstico o probabilidad de riesgo – Prueba

De los 10 casos analizados, se obtuvieron 9 aciertos con la condición o patología del paciente y 1 prueba no acertada, sin embargo, la red neuronal indica la probabilidad de que el paciente padece la enfermedad por su edad y antecedentes familiares, este resultado coincide con el registro de la historia clínica del paciente.

Los 50 casos clínicos fueron la muestra de los pacientes que permitieron el acceso a los datos y aceptaron participar del experimento en la casa de salud mencionada. Sin embargo, con la misma metodología se podrá ajustar las redes neuronales cuando se logre aumentar la muestra.

### 5. CONCLUSIONES

- Las redes neuronales utilizadas en el presente trabajo permitieron realizar un prediagnóstico que coincidió con el diagnóstico establecido para el 90% de los casos analizados.
- La estimación de la glucosa cumple con la norma consensuada ISO 15197 debido a que todas las estimaciones realizadas por la red neuronal se encuentran dentro de las zonas A y B.
- El uso de redes neuronales permitió la creación de un sistema experto para el prediagnóstico médico de la diabetes Mellitus tipo 2.
- El sistema propuesto, tiene la finalidad de actuar como un instrumento o agente coadyuvante para el diagnóstico final de la enfermedad. Dando a conocer al paciente su estado de salud y el riesgo de padecer la misma.

## REFERENCIAS

American Diabetes Association. (2019). Standards of Medical Care in Diabetes-2019. Obtenido de <https://care.diabetesjournals.org/>.

Baradó, S., Ibañez, L.J. y Agüero, M.J. (2013). *Sistemas Expertos: Fundamentos, Metodologías y Aplicaciones*. Palermo, Argentina: Universidad de Palermo.

Campos, E. (2019). Diagnóstico Asistido Por Inteligencia Artificial De Aprendizaje Profundo Para El Tamizaje Automático De Anormalidades De Fondo De Ojo De Múltiples Patologías. Universidad Nacional Autónoma de México, Ciudad de México, México.

Gobierno de México Secretaría de Salud. (2014). Tratamiento de la Diabetes Mellitus Tipo 2 en el primer nivel de atención. Obtenido de <http://educads.salud.gob.mx/sitio/vRes/recursos.php>.

Instituto Nacional de Estadísticas y Censos del Ecuador. (2017). Diabetes, segunda causa de muerte después de las enfermedades isquémicas del corazón. Obtenido de <https://www.ecuadorencifras.gob.ec/diabetes-segunda-causa-de-muerte-despues-de-las-enfermedades-isquemicas-del-corazon/>.

International Diabetes Federation. (2019). Atlas de la Diabetes de la FID Novena Edición. Obtenido de <https://idf.org/>.

Organización Mundial de la Salud. (2019). Informe Mundial sobre la Diabetes (16.3). Obtenido de [https://apps.who.int/iris/bitstream/handle/10665/204877/WHO\\_NMh\\_NVI\\_16.3\\_spa.pdf;jsessionid=4477B049211B933F8DA5BECA2B39C290?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/204877/WHO_NMh_NVI_16.3_spa.pdf;jsessionid=4477B049211B933F8DA5BECA2B39C290?sequence=1).

Maitra, A. (2015). Sistema endocrino. En Kumar, V., Abbas, A.K. Y Aster, J.C. (ELSEVIER), Robbins y Cotran. *Patología estructural y funcional* (pp. 1105-1122), Barcelona, España: ELSEVIER.

Meza, R., Aguilar, A., Ureña, E., Vázquez, C., Posada, R. y González, M. (2018). Sistema experto difuso para el control metabólico en pacientes con diabetes mellitus tipo 2. *SciELO*, 28(2), doi: 1015174/au.2018.1561.

Ministerio de Salud Pública del Ecuador. (2012). Manual de procedimiento de Antropometría y determinación de la presión arterial. Obtenido de <https://bibliotecapromocion.msp.gob.ec/greenstone/collect/promocin/index/assoc/HASH9d40.dir/doc.pdf>.

Santander, N. (2019). Inteligencia artificial en la detección precoz de retinopatía diabética en África. Hospital Universitario de Burgos. Burgos, España.