



# Congreso Internacional sobre la Enseñanza y Aplicación de las Matemáticas

## diAbassist

### Autores:

Oscar Alberto Chávez Sandoval  
Yamill Adrián López Hernández  
Octavio Muñoz Sacristán

### Asesora:

Judith Mayte Flores Pérez

octavio.sacristan@gmail.com

Artículo incluido en la publicación electrónica Memorias del Congreso ISSN 2448-7945 | Mayo 2024



Departamento de  
Matemáticas





## Objetivo

El objetivo de el desarrollo de la siguiente inteligencia artificial es identificar si una persona puede tener diabetes y con ello la persona que ejecute dicha IA, pueda ir a la brevedad con un médico para que este le otorgue la atención y tratamiento para poder ayudarle con dicho padecimiento médico y con ello evitar complicaciones de dicha enfermedad

## Introducción

En México, la diabetes es una de las principales enfermedades crónicas que afecta a una gran parte de la población, representando un problema de salud pública de gran magnitud. La detección temprana de esta enfermedad es esencial para reducir su impacto y mejorar el pronóstico de los pacientes. En este escenario, la inteligencia artificial (IA) emerge como una herramienta prometedora que puede revolucionar el diagnóstico y tratamiento de la diabetes en el país.



## Metodología y/o desarrollo

### Metodología

Se cargaron los datos mediante un archivo .csv, que contiene los datos de diabetes, recopilados previamente, a través de internet sobre algunos de los síntomas más comunes de pacientes de diabetes.



## Metodología y/o desarrollo

### Desarrollo

Se recopilaron datos clínicos de pacientes que incluyen los síntomas y padecimientos más comunes, como son: pérdida de peso, debilidad, fatiga, poliuria, curación retrasada, etc.

Mediante estos datos se realizó el preprocesamiento de datos realizando una limpieza de estos, como la eliminación de valores atípicos, imputación de valores faltantes y codificación de variables categóricas

Se separaron las variables predictoras (X) de las variables objetivo (Y)



## Metodología y/o desarrollo

Se creó un modelo de árbol de decisión

Se dividió el conjunto de datos en en conjuntos de entrenamiento y prueba, con este modelo se entrenan los datos de entrenamiento y se realizan predicciones con los datos de prueba calculando la precisión del modelo y se muestra la matriz de confusión

Por último, se creó una interfaz gráfica que captura los datos que ingrese el usuario y al dar clic en el botón “Predecir Diabetes” se capturan los datos ingresados, se realiza una predicción con el modelo entrenado y se muestra el resultado mediante un cuadro de mensaje



## Metodología y/o desarrollo

```
1 # 0. importar las librerías necesarias
2 import pandas as pd
3 from sklearn.model_selection import train_test_split
4 from sklearn.tree import DecisionTreeClassifier
5 from sklearn.metrics import accuracy_score, confusion_matrix
6 from sklearn.tree import export_graphviz
7 from tkinter import *
8 from tkinter import messagebox
9
10 # 1. Leer el archivo diabetes.csv
11 data = pd.read_csv("C:/Users/octa_/Downloads/diabetes.csv")
12
13 # 2. Leer las primeras 5 líneas
14 print("Primeras 5 líneas del archivo:")
15 print(data.head())
16
17 # 3. Revisar las columnas
18 print("\nColumnas del archivo:")
19 print(data.columns)
20
21 # 4. Renombrar las columnas
22 column_names = ['Age', 'Gender', 'Polyuria', 'Polydipsia', 'sudden weight loss',
23                'weakness', 'Polyphagia', 'Genital thrush', 'visual blurring',
24                'Itching', 'Irritability', 'delayed healing', 'partial paresis',
25                'muscle stiffness', 'Alopecia', 'Obesity', 'class']
26 data.columns = column_names
27
28 # 5. Utilizar pd.factorize para transformar las etiquetas a numéricas
29 for column in data.columns:
30     data[column], _ = pd.factorize(data[column])
31
32 # 6. Realizar la descripción estadística
33 print("\nDescripción estadística de los datos:")
34 print(data.describe())
35
```



## Metodología y/o desarrollo

```
36 # 7. Mostrar información de los datos
37 print("\nInformación de los datos:")
38 print(data.info())
39
40 # 8. Separar las variables X y Y
41 X = data.drop('class', axis=1)
42 y = data['class']
43
44 # 9. Crear el modelo de árbol de decisión
45 model = DecisionTreeClassifier()
46
47 # 10. Dividir los datos en entrenamiento y prueba
48 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
49
50 # 11. Entrenar el modelo y realizar la predicción para los datos de prueba
51 model.fit(X_train, y_train)
52 y_pred = model.predict(X_test)
53
54 # 12. Evaluar el modelo con accuracy y matriz de confusión
55 accuracy = accuracy_score(y_test, y_pred)
56 conf_matrix = confusion_matrix(y_test, y_pred)
57
58 print("\nAccuracy del modelo:", accuracy)
59 print("\nMatriz de confusión:")
60 print(conf_matrix)
61
62 # 14. Crear una interfaz gráfica para predecir la diabetes
63 def predecir_diabetes():
64     datos = []
65     for entry in entries:
66         datos.append(entry.get())
67
68     # Convertir los datos a números enteros
69     datos = [int(d) for d in datos]
70
```



## Metodología y/o desarrollo

```
70
71     # Realizar la predicción
72     prediction = model.predict([datos])
73
74     # Mostrar el resultado
75     if prediction[0] == 1:
76         messagebox.showinfo("Resultado", "El diagnóstico es positivo para diabetes.")
77     else:
78         messagebox.showinfo("Resultado", "El diagnóstico es negativo para diabetes.")
79
80     # Crear la ventana
81     ventana = Tk()
82     ventana.title("Predicción de diabetes")
83
84     # Crear etiquetas y campos de entrada para los datos del paciente
85     # Texto informativo
86     texto_informativo = "0 para sí, 1 para no"
87     Label(ventana, text=texto_informativo).grid(row=0, column=0, columnspan=2)
88
89     labels = ['Edad', 'Genero(1 Masculino, 2 Femenino)', 'Poliuria', 'Polidipsia', 'Perdida de pe
90             'Debilidad', 'Polifagia', 'Aftas genitales', 'Visión borrosa',
91             'Picor', 'Irritabilidad', 'Curación retrasada', 'Paresia parcial',
92             'Rigidez muscular', 'Alopecia', 'Obesidad']
93
94     entries = []
95     for i, label in enumerate(labels):
96         Label(ventana, text=label).grid(row=i+1, column=0)
97         entry = Entry(ventana)
98         entry.grid(row=i+1, column=1)
99         entries.append(entry)
100
101     # Botón para realizar la predicción
102     predict_button = Button(ventana, text="Predecir Diabetes", command=predecir_diabetes)
103     predict_button.grid(row=len(labels)+1, columnspan=2)
104
105     # Iniciar la ventana
106     ventana.mainloop()
```



## Resultados

Se realizó un proceso de preprocesamiento de datos en el que las etiquetas categóricas se transformaron a valores numéricos para simplificar el análisis y el modelado. Se empleó un árbol de decisión como algoritmo de aprendizaje automático para predecir la presencia de diabetes. El modelo se entrenó con un conjunto de datos dividido en un 80% para entrenamiento y un 20% para prueba. Se evaluó el modelo obteniendo una precisión que muestra qué tan bien puede predecir si una persona tiene diabetes. Además, se generó una matriz de confusión que detalla los verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos. Finalmente, se desarrolló una interfaz gráfica de usuario intuitiva con Tkinter, donde los usuarios pueden ingresar sus datos personales para obtener una predicción sobre la presencia de diabetes.



## Conclusiones

- El modelo de árbol de decisión se muestra efectivo para predecir la presencia de diabetes utilizando las características proporcionadas. Aunque la precisión del modelo ofrece una visión general de su rendimiento, es importante tener en cuenta otras métricas y realizar una validación más completa para evaluar su robustez. Además, la interfaz gráfica desarrollada ofrece una herramienta amigable que permite a los usuarios acceder al modelo sin requerir conocimientos de programación.



## Bibliografía

Admin, & Admin. (2023, 14 septiembre). Diabetes, la otra gran pandemia del siglo XXI - Gaceta UNAM. Gaceta UNAM. <https://www.gaceta.unam.mx/diabetes-la-otra-gran-pandemia-del-siglo-xxi/#:~:text=En%20nuestro%20pa%C3%ADs%2C%20seg%C3%BAn%20la,%2022%20%25%20a%20nivel%20nacional>

¿Qué es la diabetes y por qué es una de las principales causas de muerte en México? (s. f.). IBERO. <https://ibero.mx/prensa/que-es-la-diabetes-y-por-que-es-una-de-las-principales-causas-de-muerte-en-mexico>

Gracias por su atención



Departamento de  
Matemáticas

