Diabe-Detect

A ML ANALYSIS ON DIAGNOSING DIABETES

By: JV, Maribel, Nikko, Stephen

Table of Contents

In 2021, 11.6% of the U.S. population, or 38.4 million people, had diabetes.[1]

WHAT IS DIABETES?

Diabetes is a **chronic condition** that occurs when the body doesn't produce enough insulin or can't use insulin properly. Insulin is a hormone that helps regulate blood sugar levels.[2]

THE OBJECTIVE:
IDENTIFY PREDICTORS OF DIABETES

The object was to develop machine learning models to predict the risk of developing diabetes based on **symptoms**, **health habits**, **SDOH**, and **other clinical indicators**. These models will help healthcare professionals and individuals identify at-risk populations and take preventive steps, improving diabetes prevention, early detection, and management.

What factors are strong predictors of diabetes?

1. Symptoms

- Polyuria
- Polydipsia
- Visual blurring
- Partial Paresis
- Delayed healing

2. Lifestyle and Health Habits

- Alcohol
- Physical Activity
- Sleeping
- Smoking
- Med CheckUps

3. Health Indicators/SDOH

- Age
- BMI
- Location
- Blood Sugar
- HbA1C

4. Clinical Indicators

- AIDS
- Cirrhosis
- Hepatic Failure
- Immuno-suppression

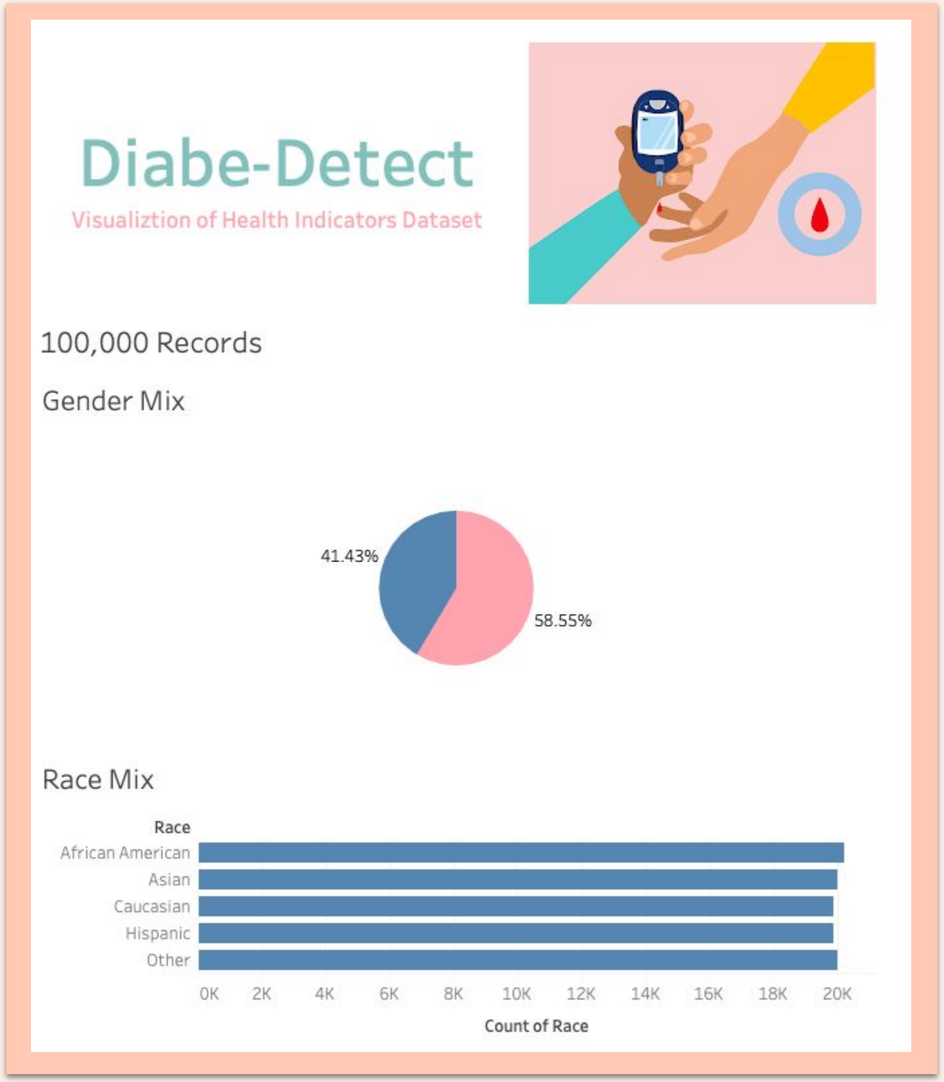
SDOH = Social Determinants of Health

THE HEALTH INDICATORS / SDOH HAD THE STRONGEST PREDICTIONS

	Accuracy	Loss
Symptoms	0.915	0.223
Health Habits	0.814	0.499
Health Indicators/SDOH	0.949	0.202
Clinical Indicators	0.795	0.452

Dataset Visualization

Dataset Visualization



Import because...

1. Data Structure & Quality
2. Pattern Recognition
3. Model Performance

[Tableau Link]

Neural Network Analysis

Demo Overview

NEURAL NETWORK	Sequential Keras Model + KerasClassifier (Weak/Strong Features) + KerasTuner (Auto Optimization)
PHASE 1 Build & Save Model	Tensorflow (keras.models) -3 layers -100 epochs Hyperparameters (keras-tuner) -# Neurons -Activation Function (ReLu & Sigmoid) Save model (*diabetes_mode.h5*)
PHASE 2 Run the model	Tensorflow (load_model) Transform Dataset (match formatting of original dataset) Pass new data into the model Analyze accuracy

Get your check up

(Have your Blood Sugar and A1c measured)

APPENDIX

Project Requirements

- Find a problem worth solving, analyzing, or visualizing.
  - The team focused on exploring a medical use case about finding what factors predict a diabetes diagnosis. JV, Maribel, Nikko, and Stephen each performed a ML analysis on unique datasets.
- Use machine learning (ML) with the technologies we've learned.
  - Several models were tested, however, we all performed a Sequential Keras Model on our own dataset.
- You must use Scikit-learn and/or another machine learning library.
  - As part of our ML analysis, we used Scikit-learn.
- You must you as least two concepts learned in the class.
  - We used:
    i. Python Pandas
    ii. Python Matplotlib
    iii. Tableau