How to Predict Type 2 Diabetes Using Machine Learning

Why Diabetes?

- 10% of population by 2015, projected to reach 12% or 5 million people by 2025.
- (https://www.diabetes.ca/health-care-providers/clinical-practice-guidelines/chapter-1#panel-tab_FullText)
- healthcare/aging population,
- Can build preventative treatment models to prevent more people developing diabetes by focusing on other factors can lead to diabetes.

Source of data –

- National Health and Nutrition Examination Survey

- Variables:
  - ID, age, sex, height, weight,
  - BMI, Cholesterol, other health data,
  - Smoker/non-smoker, drinker/non-drinker, family history of diabetes,

Target: Diabetes (Type 2)

Data Exploration

- EDA performed in VS code

- Started renaming columns, replacing values to numeric values, dropping nulls,
-

Data Exploration

- EDA performed in VS code

- Started renaming columns, replacing values to numeric values, dropping nulls,
- Density graphs, counting values, crosstab

Questions we hope to answered:

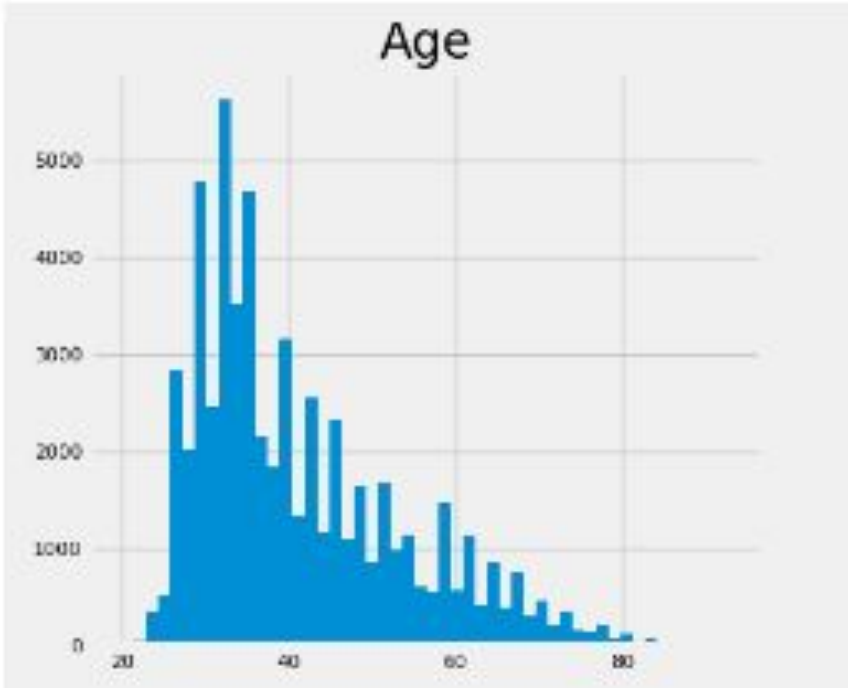
- Has the total number of people with diabetes changed due to the pandemic?
- What are the variables that are key identifiers in determining if someone has diabetes?

Data Analysis

Data was analyzed using pandas and numpy libraries.

- Various statistics were calculated and visualized to showcase variables.
  - Eg. Age

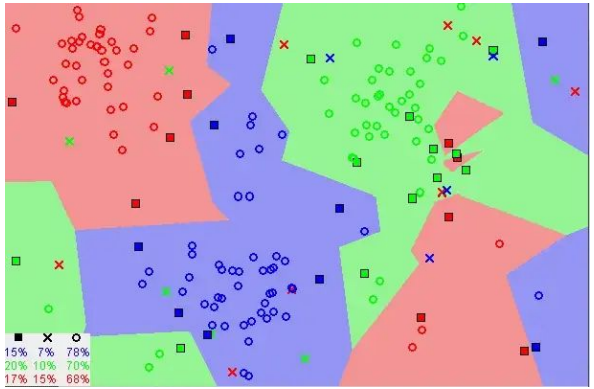
Data Analysis



Model

- KNN model (K- Nearest Neighbors Algorithm)

- Supervised model used for classification.
- Split data
- Train data
- Scale data
- Select number of neighbors



How Was Data Split, features, why this model?

- Data was split using sklearn train_test_split

```
## Split Data

X = df.iloc[:, 1:10]
y = df.iloc[:, 0]
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0, test_size=0.2)
```

- We chose this model because of its use in classification analysis.
- It’s simple to implement.

Storyboard

-Tool to use - Javascript/CSS/HTML

- Interactive elements we wish to have - pages/graphs/toolbar etc (look at examples)

Example Storyboard - like module 11