Project Making Week Presentation

Insulin Dosage Predictor App

Ivan Gan Ee Zern, 7/1/2026, [cohort id]

Problem Statement

Why?

Who?

What if?

Data Overview

10000 rows

	gender	age	family_history	glucose_level	physical_activity	food_intake	previous_medications	BMI	HbA1c	weight	insulin_sensitivity	sleep_hours	creatinine	Insulin
4896	male	59	yes	166.73	0.27	high	insulin	30.04	6.1	44.27	1.01	5.34	0.8	steady
5982	female	53	yes	171.82	0.95	high	none	39.57	11.46	72.27	0.91	6.2	1.32	up
9798	male	41	yes	156.78	1.17	medium	none	24.54	4.3	45.6	1.16	8.64	1.32	no

Source: https://www.kaggle.com/datasets/robyburns/insulin-dosage

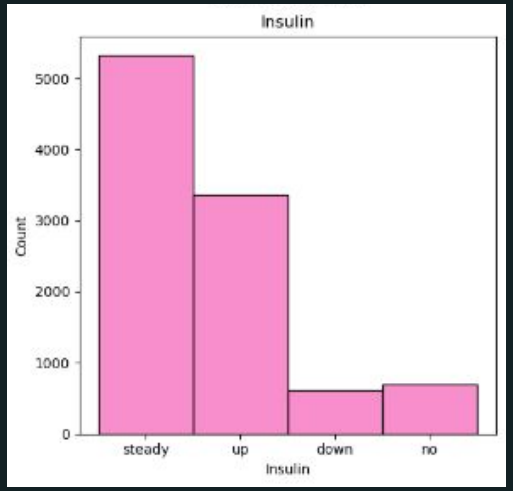
Objectives

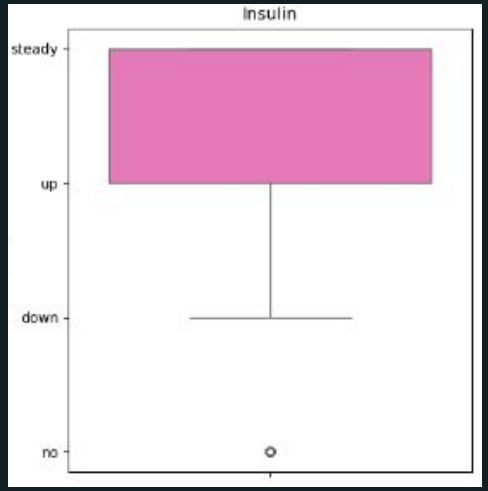
- Predict required insulin dosage
- Visualise the data provided

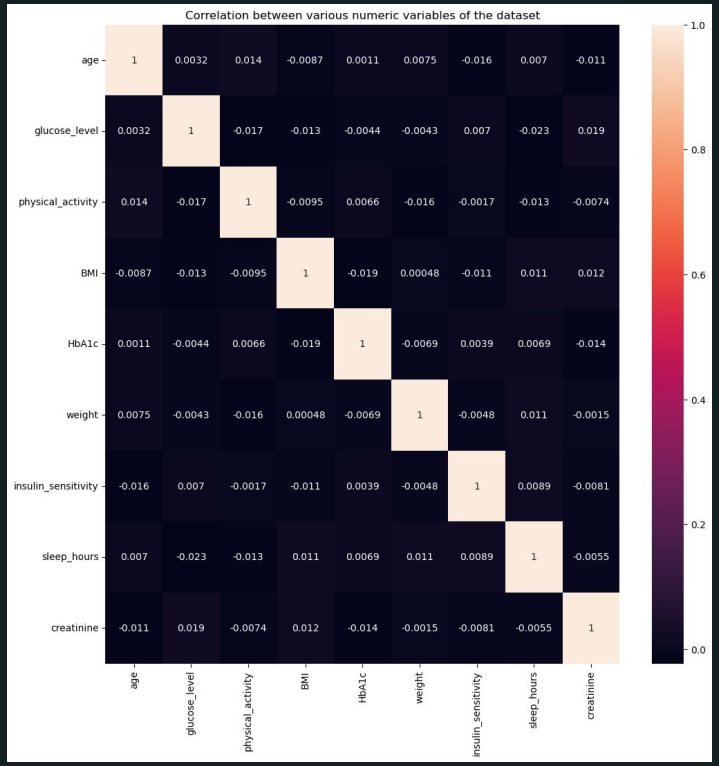
Methodology

- EDA
- Training and testing the model
- Building the web app

EDA


Insulin
Count
5000
4000
3000
2000
1000
0
steady
up
down
no
Insulin


Insulin
steady
up
down
no


Correlation between various numeric variables of the dataset
age
glucose_level
physical_activity
BMI
HbA1c
weight
insulin_sensitivity
sleep_hours
creatinine
1 0.0032 0.014 -0.0087 0.0011 0.0075 -0.016 0.007 -0.011
0.0032 1 -0.017 -0.013 -0.0044 -0.0043 0.007 -0.023 0.019
0.014 -0.017 1 -0.0095 0.0066 -0.016 -0.0017 -0.013 -0.0074
-0.0087 -0.013 -0.0095 1 -0.019 0.00048 -0.011 0.011 0.012
0.0011 -0.0044 0.0066 -0.019 1 -0.0069 0.0039 0.0069 -0.014
0.0075 -0.0043 -0.016 0.00048 -0.0069 1 -0.0048 0.011 -0.0015
-0.016 0.007 -0.0017 -0.011 0.0039 -0.0048 1 0.0089 -0.0081
0.007 -0.023 -0.013 0.011 0.0069 0.011 0.0089 1 -0.0055
-0.011 0.019 -0.0074 0.012 -0.014 -0.0015 -0.0081 -0.0055 1
1.0
0.8
0.6
0.4
0.2
0.0

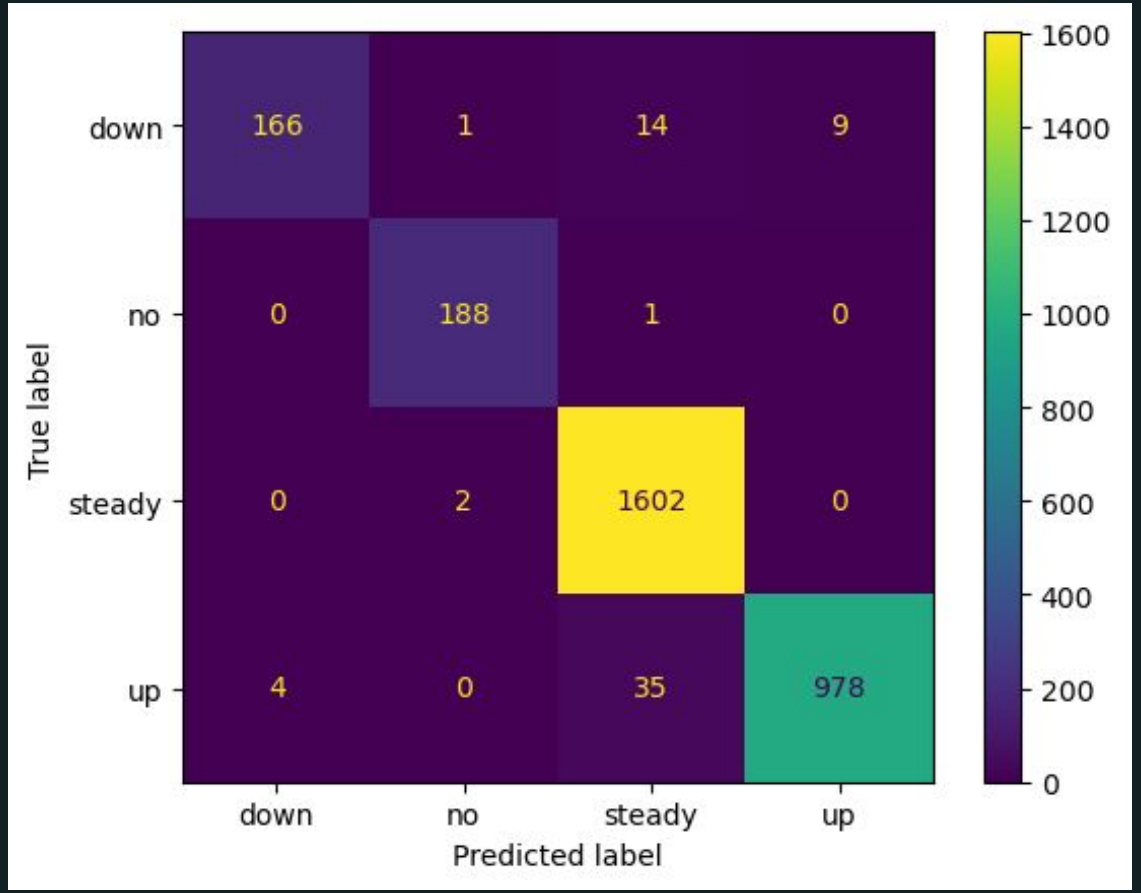
Modeling Approach

RandomForest

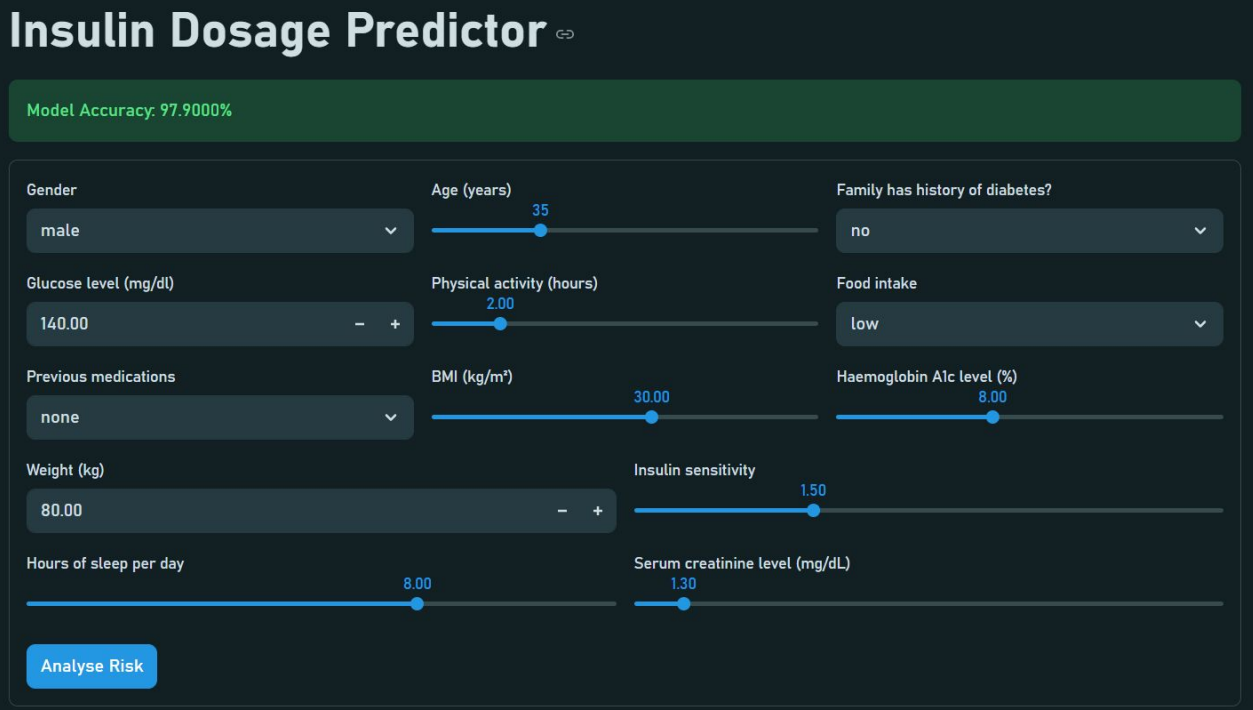
StandardScaler
OneHotEncoder
OrdinalEncoder

Results

Accuracy:
~90%
Goal:
>80%
(achieved!)

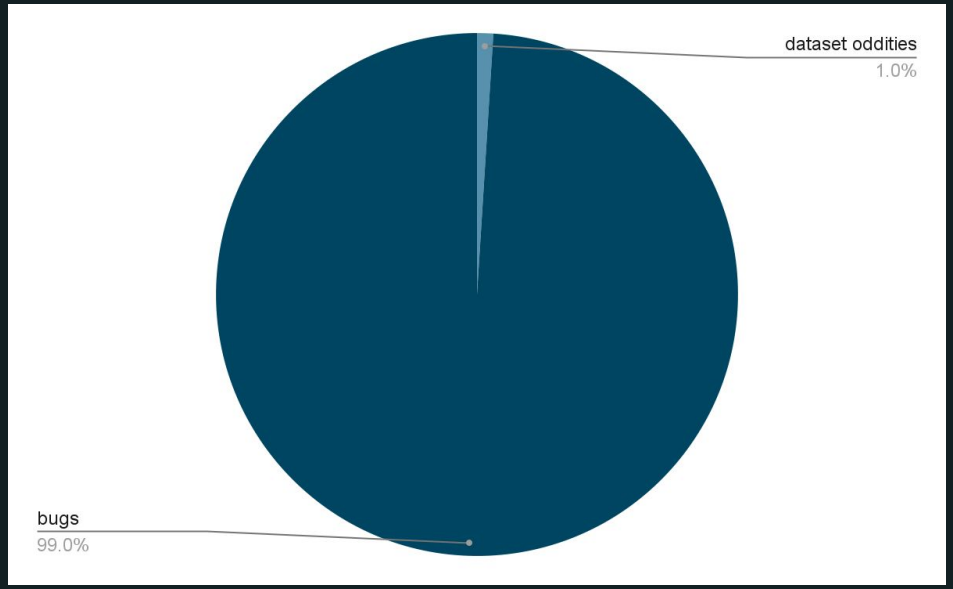


Demo



https://insulin-dosage-predictor-1vyn2r3g4noseeyou5.streamlit.app/

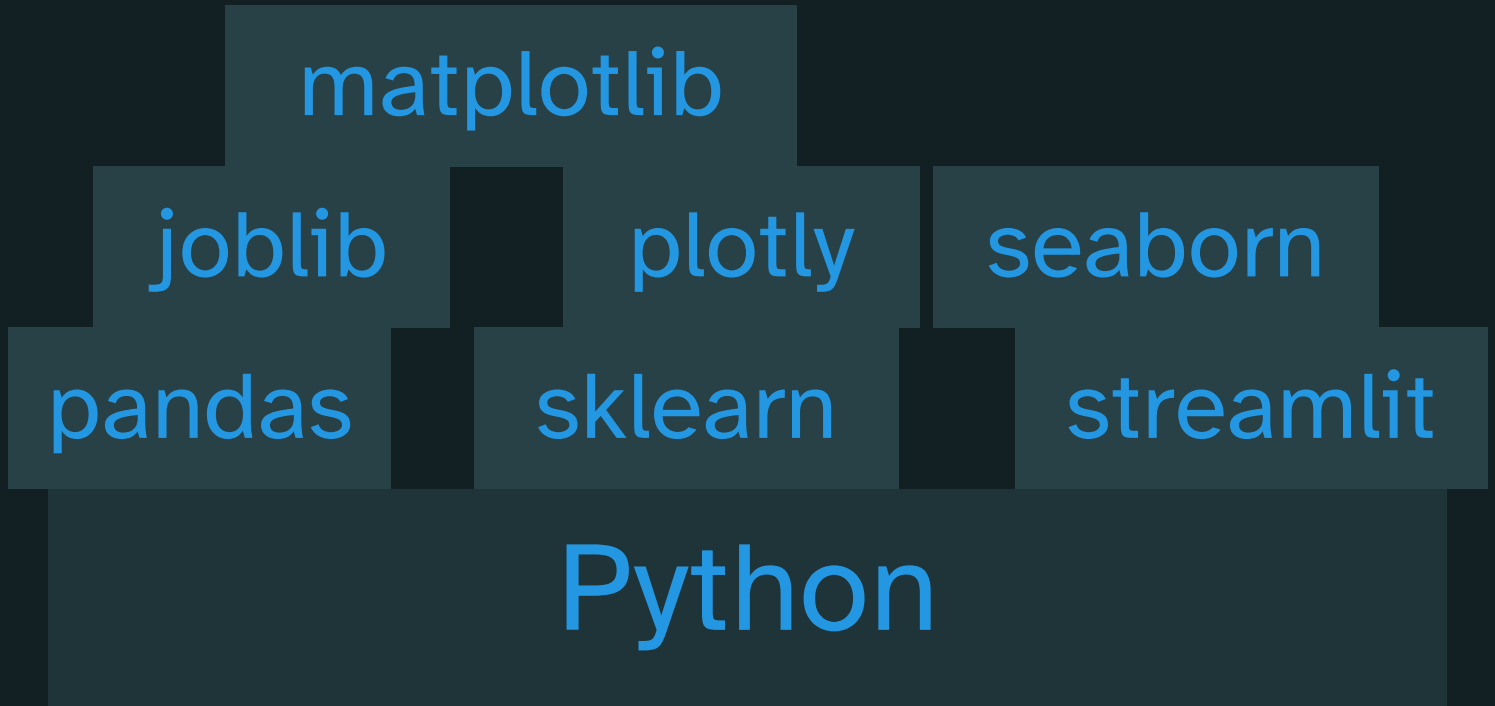
Challenges



Future Improvements

- Optimise code
- Add more insights to Data page
- Possibly hyperparameter tuning?

Tech Stack


matplotlib
joblib
plotly
seaborn
pandas
sklearn
streamlit
Python

Q&A session

Thank you.