Mini Project 4

Mynhardt Burger

Goals

1. Create model to predict loan status

2. Create public API endpoint on AWS for new predictions

Our Data

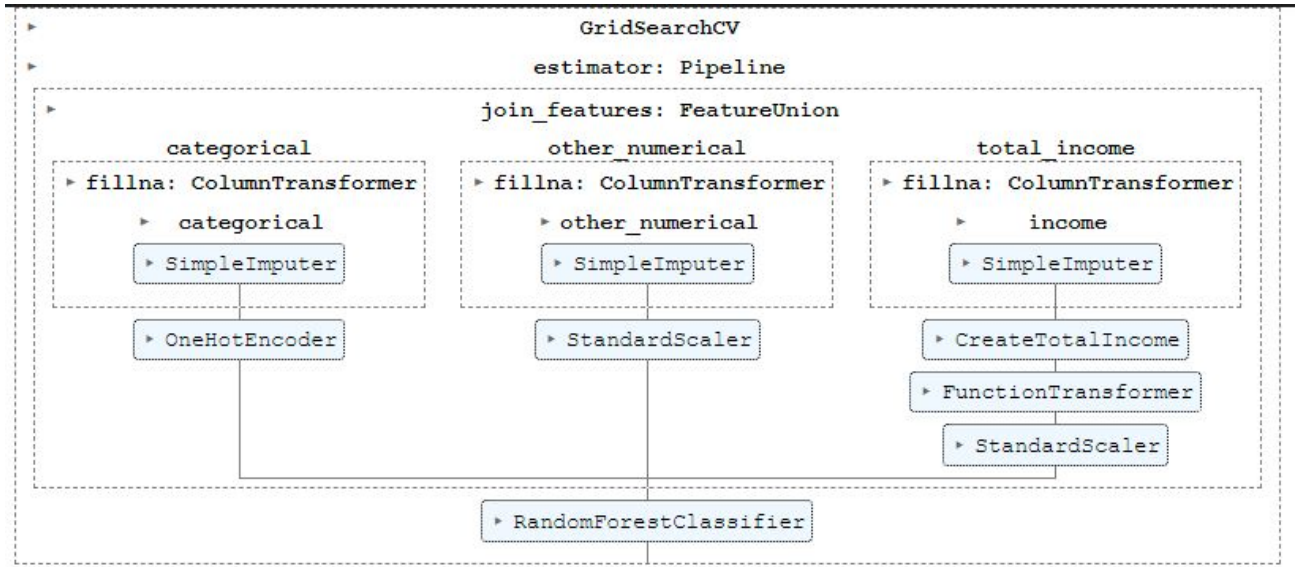
Customer details obtained via online loan application forms.

Numerical	Categorical
ApplicantIncome (Positive skew)	Dependents (missing values)
CoapplicantIncome (Positive skew)	Education
LoanAmount (missing values)	Self_Employed (missing values)
Loan_Amount_Term (missing values)	Credit_History (missing values)
	Property_Area

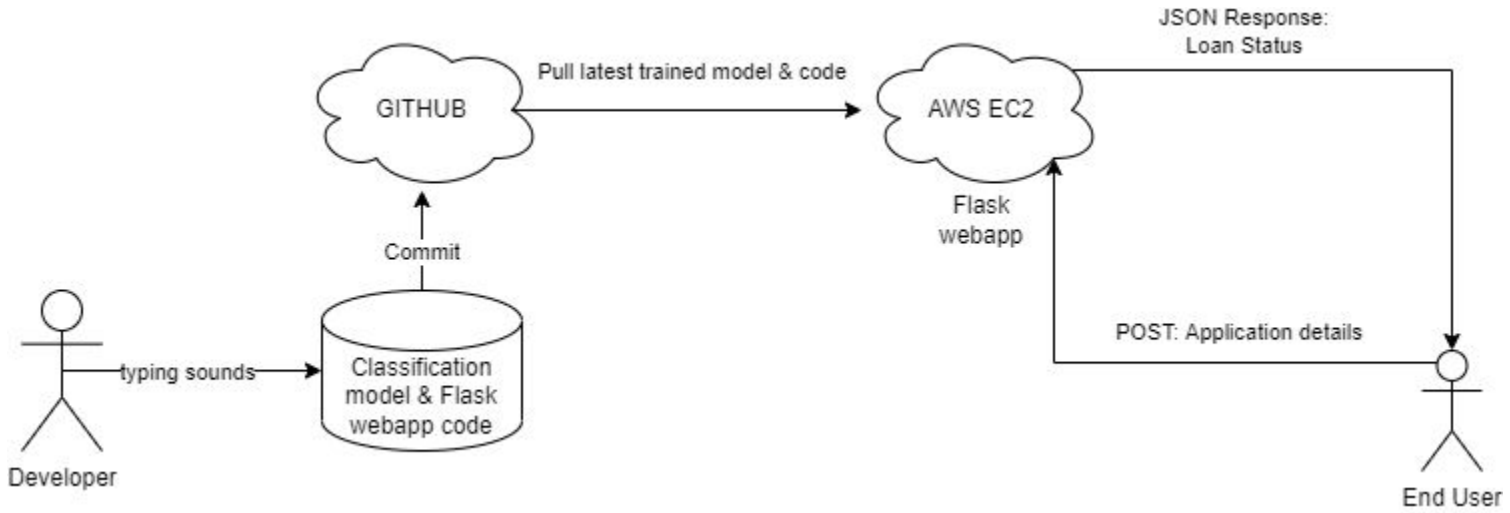
!! *Gender* and *Married* fields were ignored to avoid bias !!

Results

- Best of 7 classifiers using the ROC_AUC scoring metric was a Random Forest with a test score of 0.7845

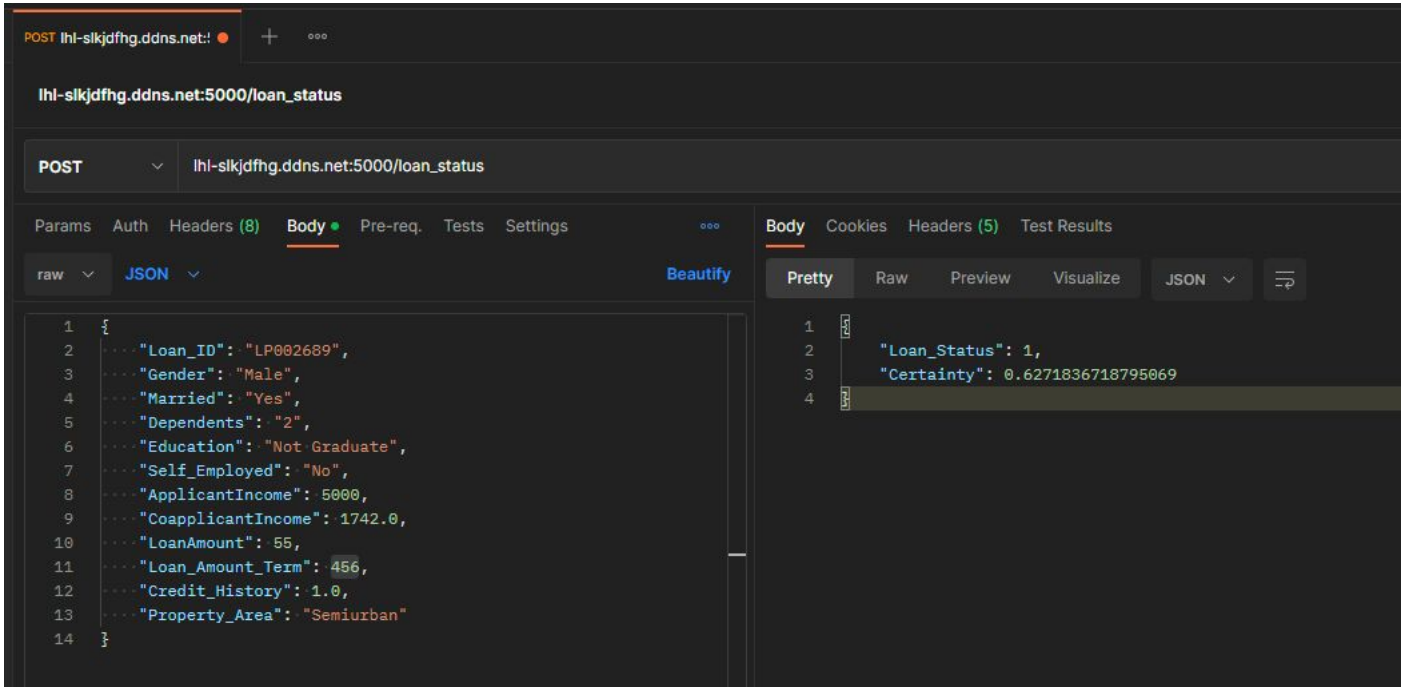


Architecture



API endpoint

Post request to http://lhl-slkjdfhg.ddns.net:5000/loan_status with a JSON object in the body



Challenges & future work

- Custom pipeline transformers
- Hyper parameter tuning
- Error handling to make API more robust
- Infrastructure as code and CI/CD using Terraform and GIT Actions