Loan default prediction
Technical architecture

MySQL Database

account	card
client	desposition
district	loan
order	trans

EDA

amount
duration
payments
district_id
gender

Features creation :
date btw. account open & loan
owner_age

Demographic features:
no_inhabitants
average_salary
average_unemployment_rate
average_crime_rate

Features calculation :
average order_amount
average_trans_amount
average_trans_balance
no_transaction
card_type
same_district

Visualization with Matplotlib

Decision gate

Feature finalization

Data Manipulation
- Correlation
- Profile report
- Transformation
  - Numerical – scaling
  - Categorical – OHE
- Upscaling min. class

Models
- KNN
- Linear regression
- Random forest
- XGBoost

Evaluation
- ROC
- AUC
- Precision, recall, F1 scores
- Confusion matrix

Tuning
- Cross-validation
- Kfold
- GridSearch
- Threshold Tuning

Decision gate

Training on optimized parameters

Saving model
in pickle file

Streamlit webapp
- Train model
  - Interactive parameters
- Interactive data input
- Run on:
  - Tuned model
  - Trained model
- Result/prediction