

Predictive Modeling Prototype

On Risk of Home Equity Line of Credit (HELOC)

Presented By Zizhuo Chen, Songyang Ma, Yuqi Su, Yancheng Xiang

Contents

- ♦ Reasoning for Log. Reg.
- ♦ Model design
- ♦ Interface demonstration
- ♦ Limitations on interface
- ♦ Limitations on Model

Reason for Log. Reg.

- ♦ High accuracy
- ♦ High interpretability
- ♦ Longer time but within a reasonable range
- ♦ Easy to understand and flexible
- ♦ Low variance

Model Design - Accuracy

Log. Reg and LDA
Linear SVC and Naïve Bayes
Tree and KNN

Train accuracy

Log. Reg. accuracy: 0.741

Linear SVC accuracy: 0.728

Naive Bayes accuracy: 0.676

LDA accuracy: 0.741

Tree accuracy: 1.000

KNN accuracy: 0.775

Validation accuracy

Log. Reg. accuracy: 0.74406

Linear SVC accuracy: 0.72939

Naive Bayes accuracy: 0.67881

LDA accuracy: 0.74355

Tree accuracy: 0.63784

KNN accuracy: 0.67931

Test accuracy

Log. Reg. accuracy: 0.72211

Linear SVC accuracy: 0.70522

Naive Bayes accuracy: 0.65302

LDA accuracy: 0.72723

Tree accuracy: 0.63153

KNN accuracy: 0.67349

Model Design – Processing Time

Logistic Regression Processing Time: 1.7161731719970703

Linear SVC Processing Time: 0.03702974319458008

Naive Bayes Processing Time: 0.005909442901611328

LDA Processing Time: 0.020890235900878906

Decision Tree Processing Time: 0.03191041946411133

KNN Processing Time: 0.003125429153442383

Model Design - Interpretability

- ♦ Easiest: Log Reg, Naïve Bayes
- ♦ Medium: LDA, Linear SVC
- ♦ Difficult: Tree and KNN

Interface Design

- ♦ Export the prediction model
- ♦ Detailed feature input
- ♦ Low tech proficiency required
- ♦ Generate probability

Credit Risk Prediction System

Disclaimer: The predictive system prototype provided is based on a combination of historical data and statistical modeling techniques, and is intended for informational purposes only. It is not intended to provide any guarantees or predictions with absolute accuracy. The prototype is subject to change.

Please put all applicant's information below

External risk estimate index

0.00

- +

Months since oldest trade open

0.00

- +

Months since most recent trade open

0.00

- +

Average months in file

0.00

- +

Number of satisfactory trades

0.00

- +

Interface Demonstration

- ♦ 32 features
- ♦ 74.33% training accuracy,
- ♦ 74.41 % validation accuracy
- ♦ 72.31 % test accuracy
- ♦ 2.489 seconds running time for test data

Limitations - Model

- ♦ Long processing time
- ♦ Limit to the linear relationship

Logistic Regression Processing Time: 1.7161731719970703

Linear SVC Processing Time: 0.03702974319458008

Naive Bayes Processing Time: 0.005909442901611328

LDA Processing Time: 0.020890235900878906

Decision Tree Processing Time: 0.03191041946411133

KNN Processing Time: 0.003125429153442383

Limitations - Interface

- ♦ Lack of visualization
- ♦ Repetitive
- ♦ Absence of feedback
- ♦ Low Accessibility

Prediction Result

The application is predicted as HIGH credit risk. Additional information may needed from applicant.

Prediction Probabilities

Probability of applicant having high credit risk: 0.9856599049658635

Probability of applicant having low credit risk: 0.014340095034136469

Thank you for using our system, hope you have a nice day!

Lessons

- ♦ Model Functions
- ♦ Choice of Model
- ♦ Limitations and Trade-off
- ♦ Different Aspects

Thank You