Predicting mortgage loan delinquency status – CS230, Spring 2018

Ksenia Ponomareva, kp260@Stanford.edu

Problem Statement

Monitoring of loan performance and early identification of high risk consumers aids prevention of loan defaults and is of interest to many banks and investors. Four multi-classification models have been built that take as input characteristics of a residential mortgage loan at inception as well as information about the first twelve monthly mortgage payments and predict the status of these payments over the next twelve-month period.

Dataset

The dataset used is publicly available Fannie Mae [1] single family loan performance data, providing information about 30 year fixed-rate mortgages issued between 2000 and 2016 and their performance. Each row represents an individual mortgage loan and columns contain 25 variables. Original mortgage rate for each example has been scaled by the average 30year federal funds rate for the corresponding quarter. This enables comparison of mortgage loans issued over the last 16.5 years.

Only entries with complete information were used and loans that were terminated prior to month 24 were also discarded. Finally, 730112 data entries were split into 90% used for train set, 5% for dev set and 5% for test set.

Features

The following eight features have been selected: original rate, original amount, original LTV, number of borrowers, debt-to-income ratio, credit score of the borrower, first three digits of the zip code and the mortgage insurance percentage. Values for these features have been normalized. Additionally, loan monthly payment performance information over the first twelve months of the loan term is utilized.

The ground truth output has seven classes with values ranging from zero to six and is based on the actual loan performance over the month 13 to 24.

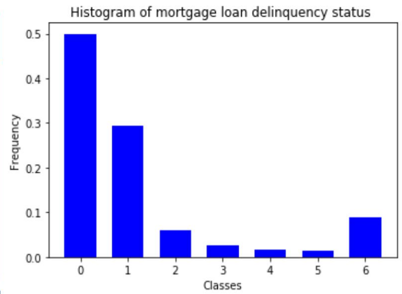

Fig 1. Frequency distribution of each class in the train data set

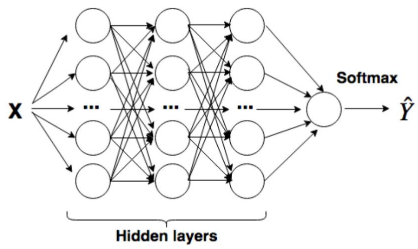

Fig 2. DeepNN model architecture

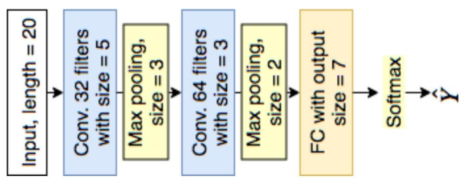

Fig 3. CNN model architecture

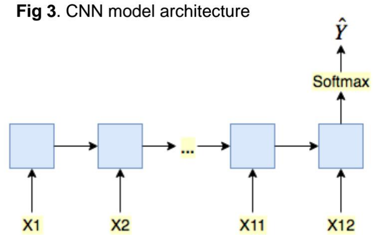

Fig 4. Many-to-one RNN model architecture

Models

The definition of the problem and dataset available enables application of four different multi-classification model architectures (Fig 2-5), which have been implemented using TensorFlow.

Baseline model has a single node with a softmax activation function. This is the simplest model out of the four.
Deep neural network (DeepNN) model has three hidden layers with 100, 50 and 10 nodes respectively. All activations are ReLU apart from the final one which is softmax.
CNN model uses a 1D structure inspired by [2] with two blocks of convolutional and max pooling layers, followed by a fully connected layer and a softmax activation function.
RNN model is a many-to-one one directional LSTM with a softmax activation applied to the last output.

Due to class imbalances in the training data as shown in Fig 1, weights, calculated following median frequency method in [4], were applied to the softmax losses to apply higher penalties for errors for labels 1-6.

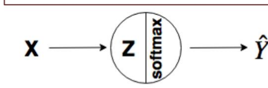

Fig 5. Baseline model architecture

Results

Models have been iterated over various hyperparameters and network structures to provide the best test accuracy. Results are provided in Table 1. DeepNN, CNN and RNN model show an improvement over the baseline, with a small variation of results between them.

	Baseline	DeepNN	CNN	RNN
Training Accuracy	53.3%	60.8%	60.5%	60.8%
Test Accuracy	53.8%	61.1%	61.0%	61.3%

Table 1. Summary of accuracy results achieved by each model

Discussion

Although DeepNN, CNN and RNN models perform better than the baseline, they can be further improved in the future by further fine tuning the hyperparameters (number of layers, hidden units, filter number and size, learning rate) to find the best combination. Additionally, considering more than the first 12 payments as features or supplementing with current/savings/credit account data as in [2] might help improve models' performance. Finally, using a more balanced sample from the data available might be worth consideration.

References

1. https://loanperformancedata.fanniemae.com/lppub/index.html
2. Kvamme, H., Sellereite, N., Aas, K., Sjursen, S. (2018) Predicting mortgage default using convolutional neural networks, Elsevier Expert Systems with Applications, 102, 207-217.
3. Yeh, I., Lien, C., (2009) The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients, Elsevier Expert Systems with Applications, 36, 2473–2480.
4. Eigen, D., Fergus, R., (2015) Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture, arXiv:1411.4734v4.
5. Poster video link: https://youtu.be/vKy1h00gFnk