Deep Learning Implementation of a Recommendation System for Restaurants

Nasreddine El Dehaibi
CS230, Stanford University

Motivation

Restaurant recommendation systems enable customers to find their ideal dining experience while helping restaurants drive traffic growth to their establishments. The goal of this project is to build a recommendation system that leverages the power of deep learning to accurately recommend restaurants to users.

Background

Traditional recommendation systems are based on either collaborative filtering or content-based filtering that group together similar users or items based on user ratings to recommend different items to users [1]. These approaches are popular for applications like movie recommendations but are not as effective for applications like restaurant recommendations. A user's preference can fluctuate much more on any given time or day. This makes it more challenging to build an effective recommendation system for this application.

Data

The dataset consists of 1161 restaurant ratings with corresponding restaurant and user information 9 different data files [2].

Table 1 Dataset

Customer-Restaurant-Ratings	Customers	Restaurants
• 1161 customer ratings of restaurants	• Preferred cuisine for 330 customers • Preferred payment method for 177 customers • User profile for 138 customers	• Methods of payments accepted at 1314 restaurants • Types of cuisine for 916 restaurants • Business hours and days for 339 restaurants • Parking availability for 702 restaurants • Location information for 130 restaurants

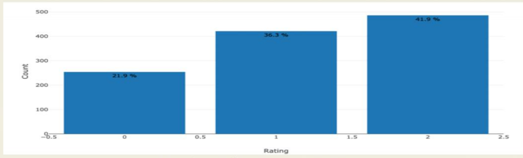

Fig. 1 Distribution of restaurant customer ratings

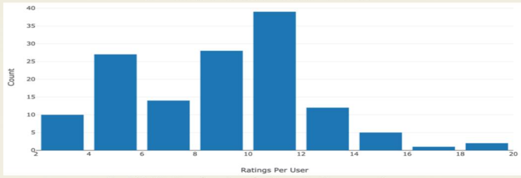

Fig. 2 Distribution of number of restaurant ratings per customer

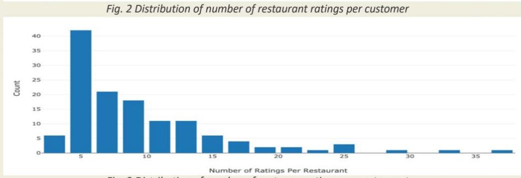

Fig. 3 Distribution of number of customer ratings per restaurant

Method

I used the Surprise Scikit package in Python to build a number of baseline models [3]. Then built a SoftMax classifier in Python using a three-layer neural network with TensorFlow.

Results

Table 2 Baseline RMSE's

Algorithm	test_rmse	fit_time	test_time
SVDpp	0.668294	0.109525	0.009100
SVD	0.679372	0.042878	0.004645
BaselineOnly	0.696977	0.000994	0.001838
KNNWithMeans	0.698772	0.002014	0.006794
KNNWithZScore	0.702664	0.006269	0.007673
CoClustering	0.720291	0.031317	0.002127
SlopeOne	0.736563	0.003677	0.004313
NMF	0.742113	0.047225	0.002896
KNNBaseline	0.759941	0.003650	0.006853
KNNBasic	0.849223	0.000602	0.004808
NormalPredictor	1.003145	0.000735	0.001739

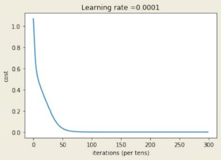

Figure 4 Cost function versus number of iterations

Table 3 Deep learning accuracy RMSE's for validation and test sets

Validation Accuracy	Test RMSE	Baseline RMSE
0.59	0.73	0.66

Conclusion

The lowest deep learning model RMSE was 0.73, which is not as good as the chosen baseline of 0.66. The likely cause of the limited performance is due to overfitting of the training data. For future work, the results from this study encourage building deep learning networks using larger datasets.

References

[1] https://medium.com/recombee-blog/machine-learning-for-recommender-systems-part-1-algorithms-evaluation-and-cold-start-6f696683d0ed
[2] https://www.kaggle.com/uciml/restaurant-data-with-consumer-ratings/version/1
[3] https://surprise.readthedocs.io/en/stable/index.html

Stanford