Enhanced Recurrent Neural Network for Combining Static and Dynamic Features for Credit Card Default Prediction

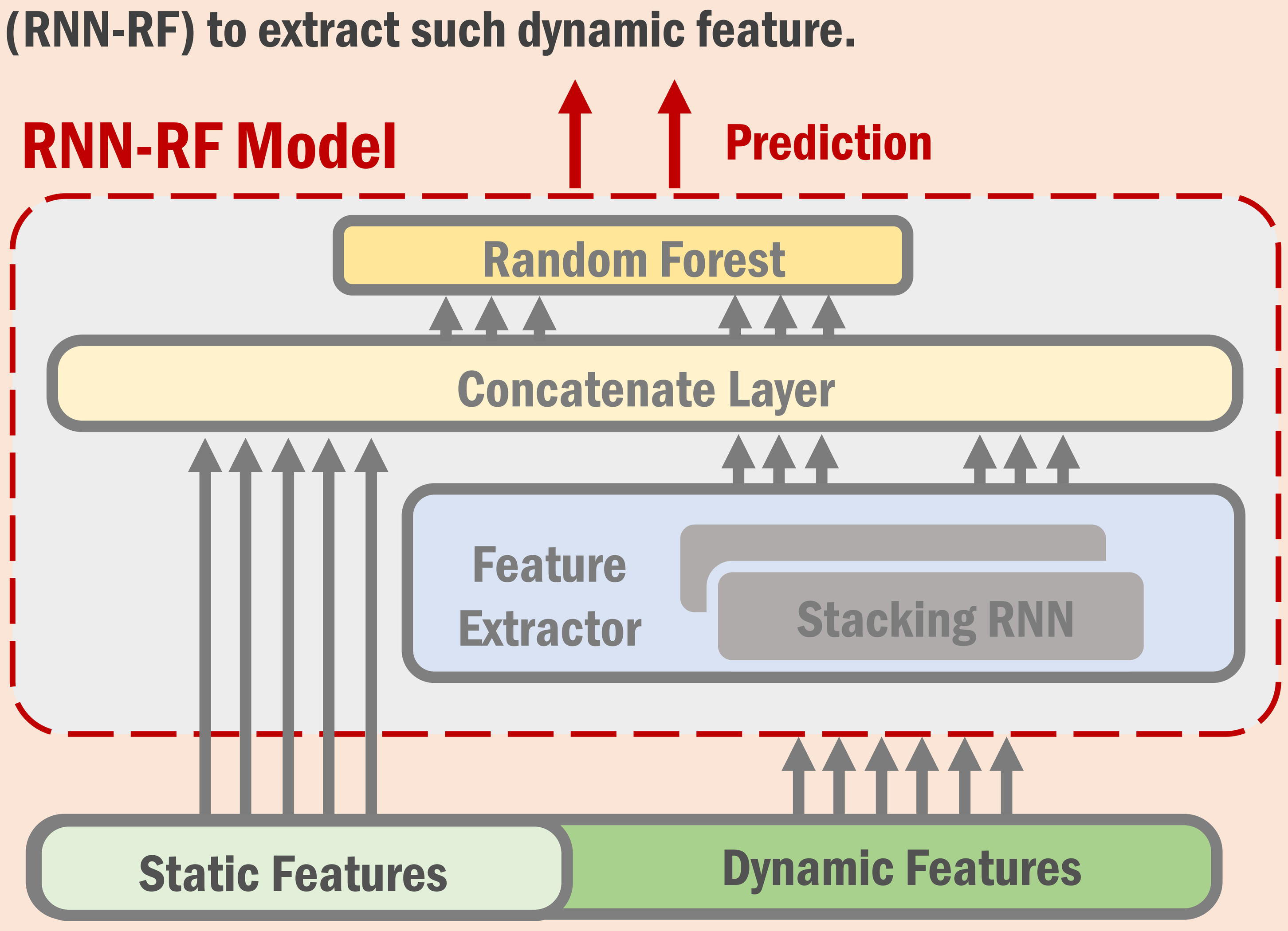
Te-Cheng Hsu, Shing-Tzuo Liou, Yun-Ping Wang, Yung-Shun Huang, Che-Lin
Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan
{andy810436, nick860617, tommt8541, vito31733}@gmail.com, che.lin@gmail.com

Motivation & Contribution

- In financial data, dynamic features that evolve with time are commonly observed. However, such time dependencies are often ignored in classical classification models. In this study, we proposed to learn a Recurrent Neural Network (RNN) feature extractor with GRU on credit card payment history to leverage the time dependencies embedded in these dynamic features.
- Lift was introduced as the evaluation metric for ranking (instead of classification) ability of models. Lift has been widely adopted in various financial applications such as modeling customer churn, subscription renewal, and promotion targeting.
- Numerical experiments showed that enhanced RNN predictor indeed provides the best performance in terms of both lift index (0.659) and AUC (0.782) compared to other benchmark models.

System Model

- To better model time sequences and leverage time dependencies, we proposed to use a RNN feature extractor with random forest (RNN-RF) to extract such dynamic feature.



Evaluation Metric

- Lift and corresponding lift index are defined as a function of n, the number of customers in the top-scoring group, as

$$lift(n) = \frac{\#True\ buyers\ in\ the\ top - n\ scoring\ customers}{n}$$

$$lift\ index = \sum_{i=1}^{10} w_i \cdot S_i$$

where S_i's are the decile lift in decile i and w_i's are prespecified weights. Lift index can be viewed as weighted average of areas divided by deciles under lift curve.

Simulation Result

- Default of credit card clients dataset from UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients) was adopted, consisting of 30,000 samples with 23 features (past 6-month credit card behavior and demographic information).
- Synthetic Minority Oversampling TEchnique (SMOTE) was adopted on the training set to overcome class imbalance issue.
- Model performance

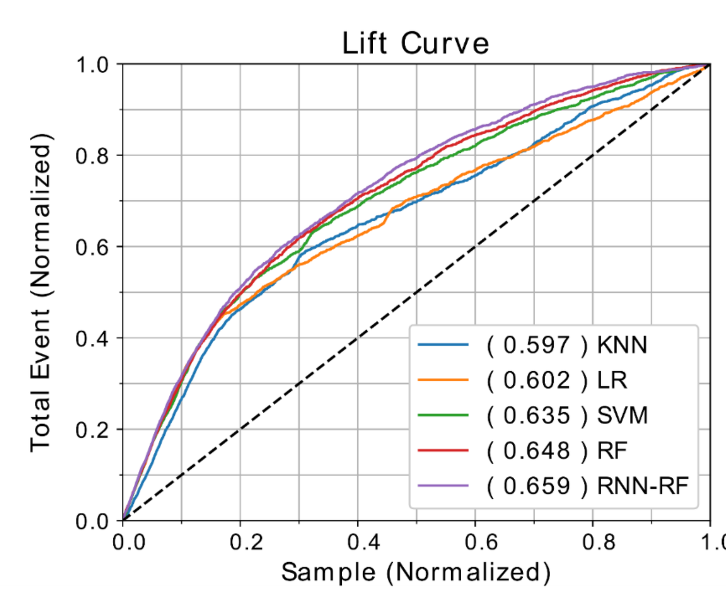


Fig. 1(a)
Lift curves and lift indices for RNN-RF and benchmark models.

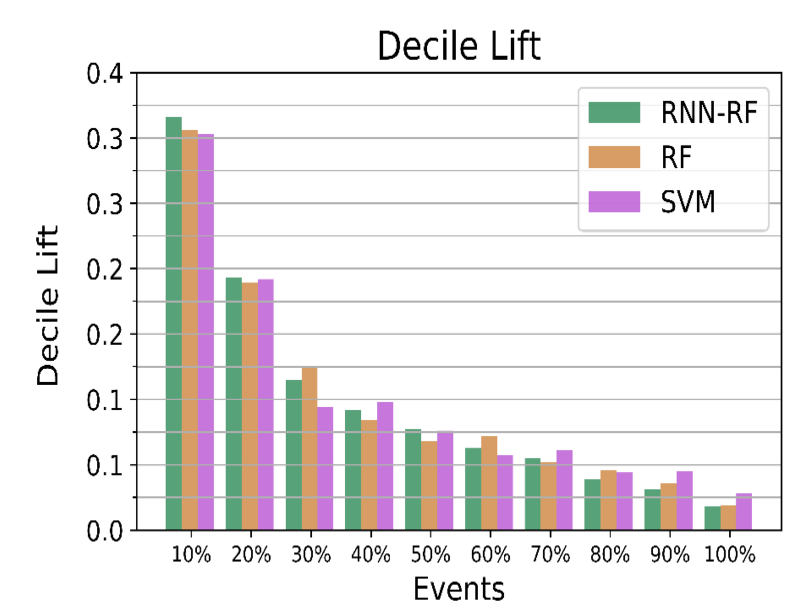


Fig. 1(b)
Decile lifts for RNN-RF, RF, and SVM.

Fig. 1 presents the comparison of our enhanced RNN-RF model with other benchmark models. It can be noted that RNN-RF model achieved the maximum lift index that is higher than KNN, LR, SVM, RF by 0.062, 0.057, 0.024, and 0.011, respectively.

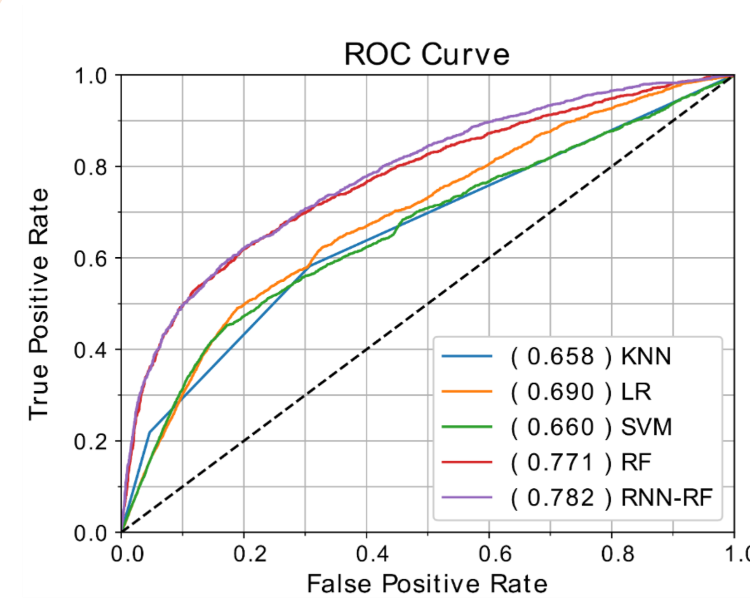


Fig. 2(a)
ROC curves and AUCs for RNN-RF and benchmark models.

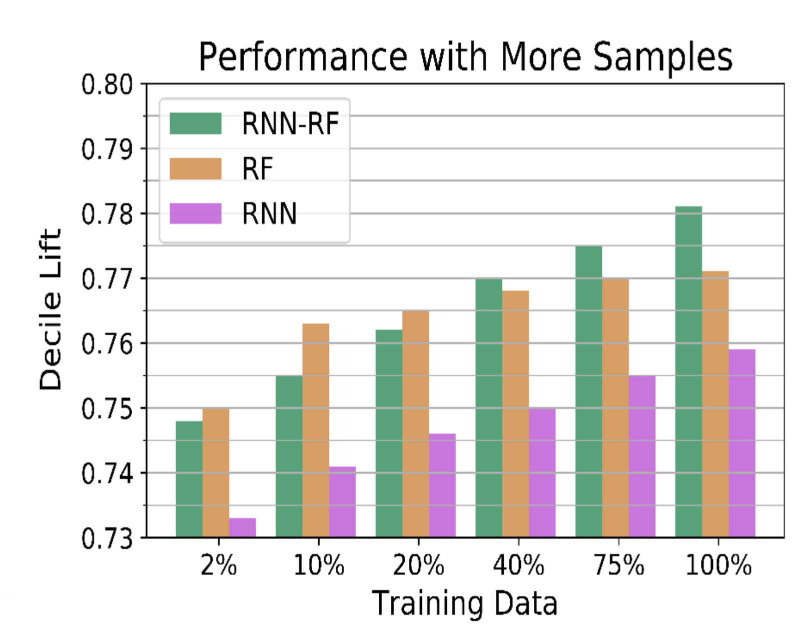


Fig. 2(b)
Decile lifts for RNN-RF, RF, and RNN-GRU with growing amounts of training data

Fig. 2 shows advantage of our proposed model when more training data are available, we varied the amounts of training data available for RNN-RF, RF, and RNN, where GRU cells were fed with both dynamic and static features.

- Model comparison

	KNN	LR	SVM	RF	RNN-RF
Lift Index	0.597	0.602	0.635	0.648	**0.659**
AUC Score	0.658	0.690	0.660	0.771	**0.782**

Table 1 Performance evaluation for RNN-RF and benchmark models.