Predicting Conference Paper Acceptance

William Jen, Shichang Zhang, and Muyun Chen
{wjen, shichang, muyunc}@stanford.edu

Background and Motivation

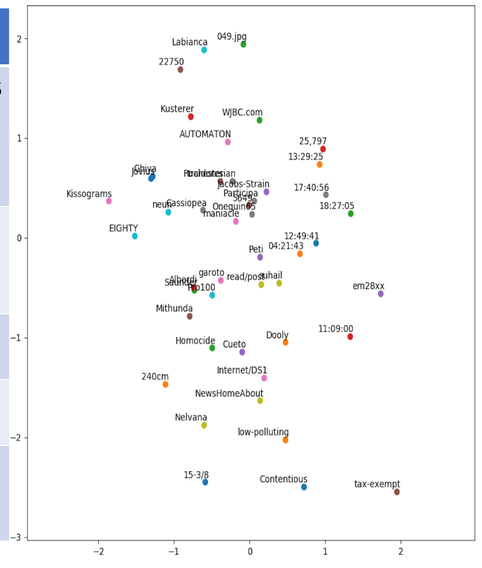
The explosion of scientific research in machine learning has led to the rapid growth of paper submissions to top conferences. Can we predict whether a machine learning paper will be accepted using machine learning methods?

Kang, et. al. published initial work on this topic in April 2018 with the public release of PeerRead, a dataset that collects research papers from AI/ML conferences.

Datasets and Features

We took all the 427 papers submitted to ICLR 2017 as our dataset. There are 172 accepted and 255 rejected papers. For each paper, we extracted 18 features of numerical and Boolean values.

Important Features
Whether abstract contains deep, neural, embedding, outperform, outperform, novel, or state_of_the_art
Number of figures, tables, sections, equations, theorems
Number of references
Bag-of-words in abstract
Average of GloVe word embeddings in abstract



Models

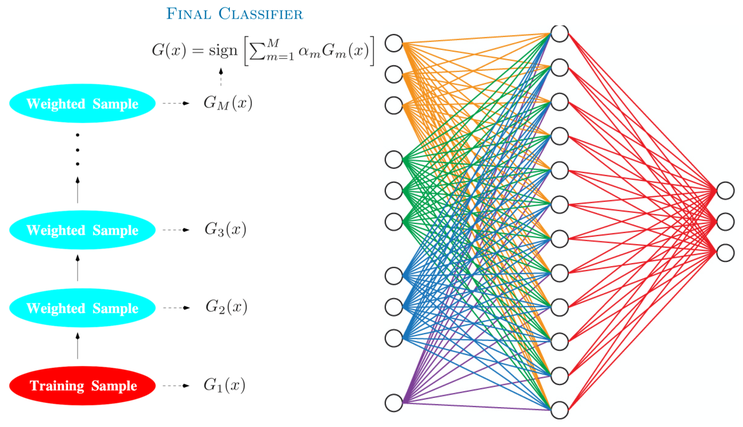
We reproduced the models from Kang's paper together with some other models.

- Logistic regression with L2/L1 regularization.
- Random Forest
- SVM with RBF kernel

$$\min_{w,b} \frac{1}{2}||w||^2$$
$$s.t. y^{(i)}(w^T x^{(i)} + b) \geq 1, i = 1, ..., m$$

$$K(x, \tilde{x}) = \exp\left(-\frac{||x - \tilde{x}||^2}{\sigma^2}\right)$$

- AdaBoost
  - We used 50 weak classifiers.



- Neural Network
  - We used ReLU activation function and 20 different network structures.

Result

The first row shows the accuracy of the baseline model of predicting by majority, which in this case is reject all the papers. Our best models outperform Kang's best model, which has **65.3**% test accuracy on average.

Model	Train Accuracy(%)	Test Accuracy(%)
Majority	60.17	60.53
Logistic L2	42.41	42.10
Logistic L1	68.48	68.42
SVM RBF	**72.49**	**71.05**
Random Forest	99.43	63.16
AdaBoost	96.56	50.00
Neural Network	63.04	60.53

Future Work

Our work focused on the ICLR dataset, which has limited examples. Similar studies can be done on other conferences with more submissions, like NIPS, or for the same conference but with submissions across years. The challenging part is parsing and featurizing the large number of papers, which is computationally expensive.

Reference

[1] D.Kang, W.Ammar, B.Dalvi, M.vanZuylen, S.Kohlmeier, E.Hovy, and R.Schwartz, "A data set of peer reviews (peerread): Collection, insights and nlp applications," in *Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL)*, New Orleans, USA, June 2018.

[2] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.