Predicting Flight Delays with Deep Learning

Ethan Chi, Jillian Tang
{ethanchi, jilingt}@stanford.edu

Problem

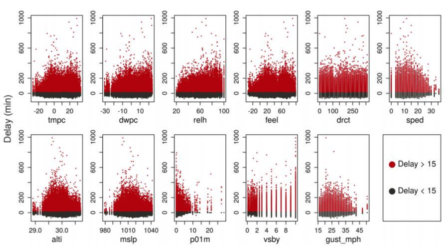
Given a flight and associated weather data, classify it as *delayed* (>15 min) or *non-delayed.*

Background

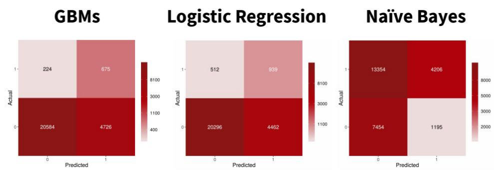
- With 24% of flights in the US delayed, delays are extremely costly both to consumers and airlines.
- Due to the sequential nature of flight delays, previous machine learning approaches have been mostly unsuccessful [1].
- Sequential DL models can better predict delays.

Data

- 314,677 flights departing from Chicago O'Hare International Airport in 2009 [split 80%-10%-10%]
- Time-matched weather data from the NOAA
- Total: 325 features, including:
  - **Weather**: temperature, dew point, cloud coverage, snow, rain, wind speed, wind direction, etc.
  - **Categorical**: carrier, airport (arrival + destination)
  - **Other:** Time, date, arrival delay (= ground truth)



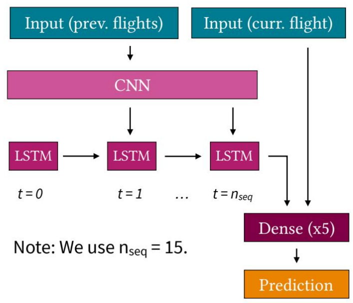
Baseline Models



Methods

Model 1: CNN + LSTM (based on LSTNet [2])

- Previous flight data fed through CNN into LSTM
- Output combined with weather data, fed into dense layers



Note: We use $n_{seq} = 15$.

Model 2: same as above, but with a dense neural network instead of a CNN [3].

We trained 500 epochs with `tf.train.AdamOptimizer`.

Batch size = 5000. L1 regularization applied with λ=0.005

Class imbalance: weighted $\mathcal{L} = -2.6y\log\hat{y} - (1-y)\log(1-\hat{y})$

Results

	GBM	Log. Reg.	Naïve Bayes	LSTM	CNN+LSTM
Accuracy	81.11%	81.02%	44.49%	78.89%	**81.90%**
Precision	**75.08%**	64.71%	23.95%	44.03%	52.32%
Recall	12.50%	17.39%	**77.87%**	43.93%	46.52%
F1	21.43%	27.41%	36.64%	43.98%	**49.25%**

Discussion

- As expected, our model greatly outperformed baseline statistical models.
- We were able to replicate the performance of the state-of-the-art paper in the field. [3]
- Using CNNs for preprocessing led to better performance, and due to the cheaper nature of convolutions, they were much faster than preprocessing with several dense layers.

Future Work

- Inclusion of data from more airports and more years to increase training set diversity
- More complex architectures (e.g. multiple layers of LSTMs)
- Integration of more features (national threat level, economy)

References

[1] A. Sternberg, J. Soares, D. Carvalho, and E. Ogasawara, "A review on flight delay prediction," *arXiv preprint arXiv:1703.06118*, 2017.
[2] Y. J. Kim, S. Choi, S. Briceno, and D. Mavris, "A deep learning approach to flight delay prediction," in *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE, 2016, pp. 1–6.
[3] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling long-and short-term temporal patterns with deep neural networks," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 2018, pp. 95–104.