Machine Learning Applied to Weather Forecasting

Mark Holmstrom, Dylan Liu, Christopher Vo
markholm@stanford.edu, dliu3@stanford.edu, cvo9@stanford.edu

Predicting

We seek to make accurate weather predictions using machine learning techniques. Traditional forecasting techniques rely on physical simulations using fluid dynamics and thermodynamics. However, the system of ordinary differential equations that govern this physical model is unstable under perturbations, and perhaps a data-driven approach could be more robust and attain higher accuracy. The high level idea behind the project is that we will use what we know about past trends to make a 7-day forecast of the daily high and low temperatures in an area. To narrow the scope of the project, we are only looking at the weather at Stanford.

Data

We obtained weather data for Stanford from Weather Underground for the years 2011-2015. For each day, the features obtained were the maximum temperature, the minimum temperature, the mean humidity, the mean atmospheric pressure, and the weather classification. There were originally nine classifications, but we condensed them into four: clear, moderately cloudy, very cloudy, rain/storm.

Features:

We used 5 features for the complicated model: for the previous days before our prediction, we had the weather label, high temperature, low temperature, mean humidity, and mean pressure. We represent these features as function and use the metric

$$d(f_1, f_2) = \sum_{j=1}^{m} \sum_{k=2}^{5} w_k \left(f_1(j)_k - f_2(j)_k\right)^2 + w_1 \mathbf{1}[f_1(j)_1 \neq f_2(j)_1]$$

to represent the distance between two such functions. For the linear regression model, we used 8 variables: high temperature, low temperature, humidity, and pressure for the two days before the first prediction in the time series.

Models

We first implemented a linear regression model for predicting temperatures. We then implemented a new algorithm that predicted the high and low temperatures based upon how similar the weather of the past two days is to historical weather sequences of two days. More formally, we consider weather sequences of nine days, and we define a metric to measure how dissimilar two sequences are in the first two days. We then find the historical weather sequences whose first two days are the most similar to the weather of the past two days, and we finally forecast the high and low temperatures of the next seven days as a weighted average of the high and low temperatures of the last seven days of the historical weather sequences.

Results

	Linear Regression Model	Sequence Comparison Model
Temp Deviation on Train Set	5.013	5.959
Temp Deviation on Test Set	5.366	6.122
Day 1 Temp Deviation	5.039	5.572
Day 2 Temp Deviation	5.157	5.922
Day 3 Temp Deviation	5.300	6.106
Day 4 Temp Deviation	5.379	6.291
Day 5 Temp Deviation	5.446	6.343
Day 6 Temp Deviation	5.566	6.291
Day 7 Temp Deviation	5.642	6.328

Fig. 1. The average temperature deviations in degrees Fahrenheit of predicted temperature values vs actual temperature values for the two models used.

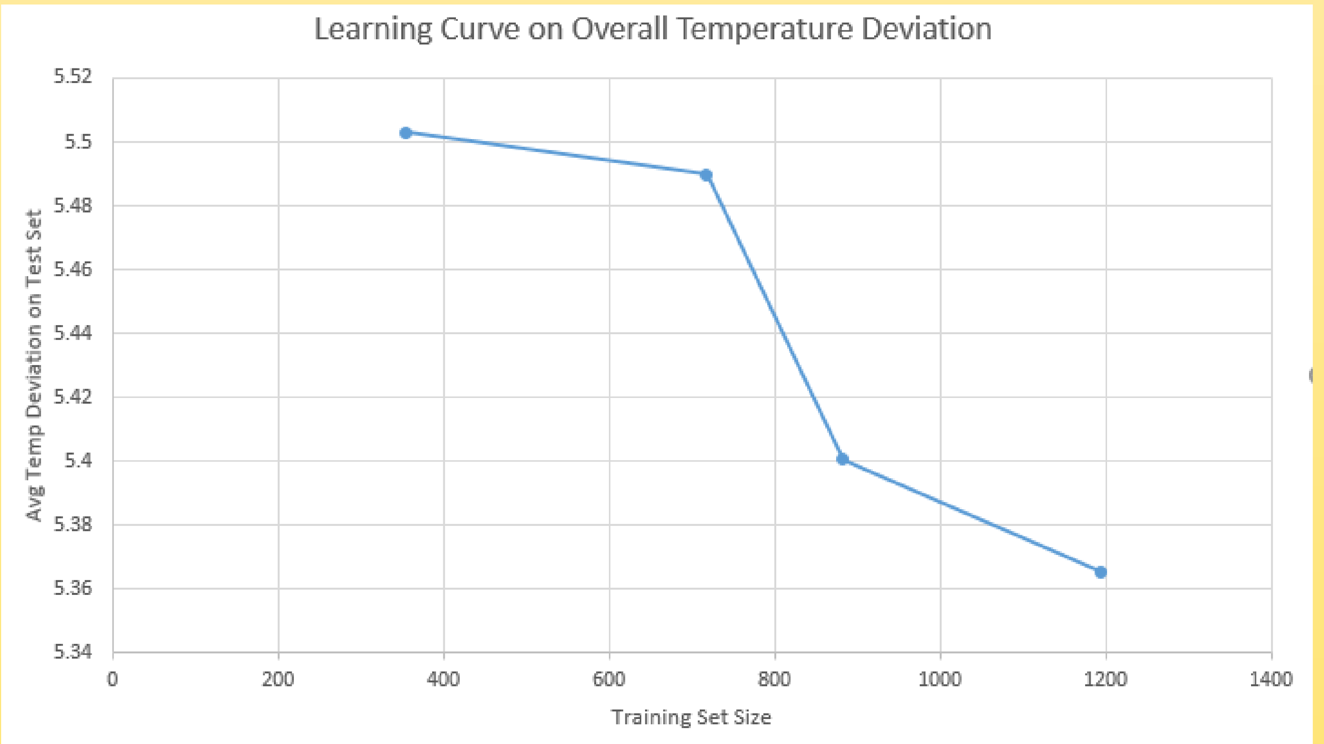

Fig. 2. The learning curve for linear regression, using the rms error averaged over the seven day sequences.

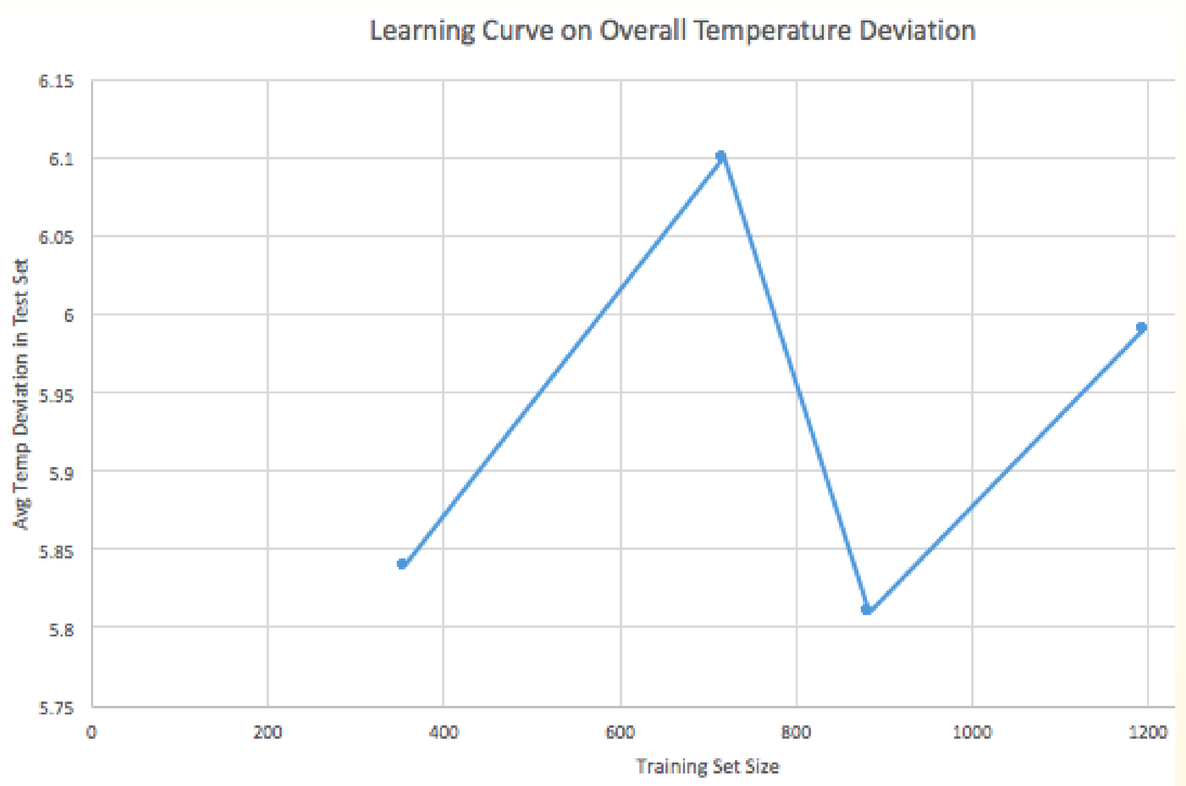

Fig. 3. The learning curve for the sequence comparison model, using the rms error averaged over the seven day sequence.

Discussion

Results showed that for both models, there was a small gap between training and testing set accuracy which indicates slight overfitting of the data. Additionally, predicted temperatures for later days were less accurate than earlier days which is expected as earlier days correlate to the first two days of features more than later days do. Additionally, the linear regression model seems to do better overall than the sequence comparison algorithm. This may be because the sequence comparison model depends solely on similarity between previous days where regression takes into account dependencies on previous days and predicted days. Overall, both models perform generally well given the average temperature deviation for many professional forecasting services is about 8 degrees Fahrenheit for ten day sequences.

Future

If given more time, we could extend the project by trying new feature combinations for both our regression and sequence comparison models to see which work best. Additionally, we could examine different models and perform feature engineering given our different deviations based on the days.

References

[1] A. Ng, “CS229 Lecture Notes Supervised Learning
[2] “Stanford, CA” in Weather Underground, The Weather Company, 2016. [Online]. Available: https://www.wunderground.com/us/ca/palo-alto/zmw:94305.1.99999. Accessed: Nov 20, 2016.