Noah Plant

Professor Mann

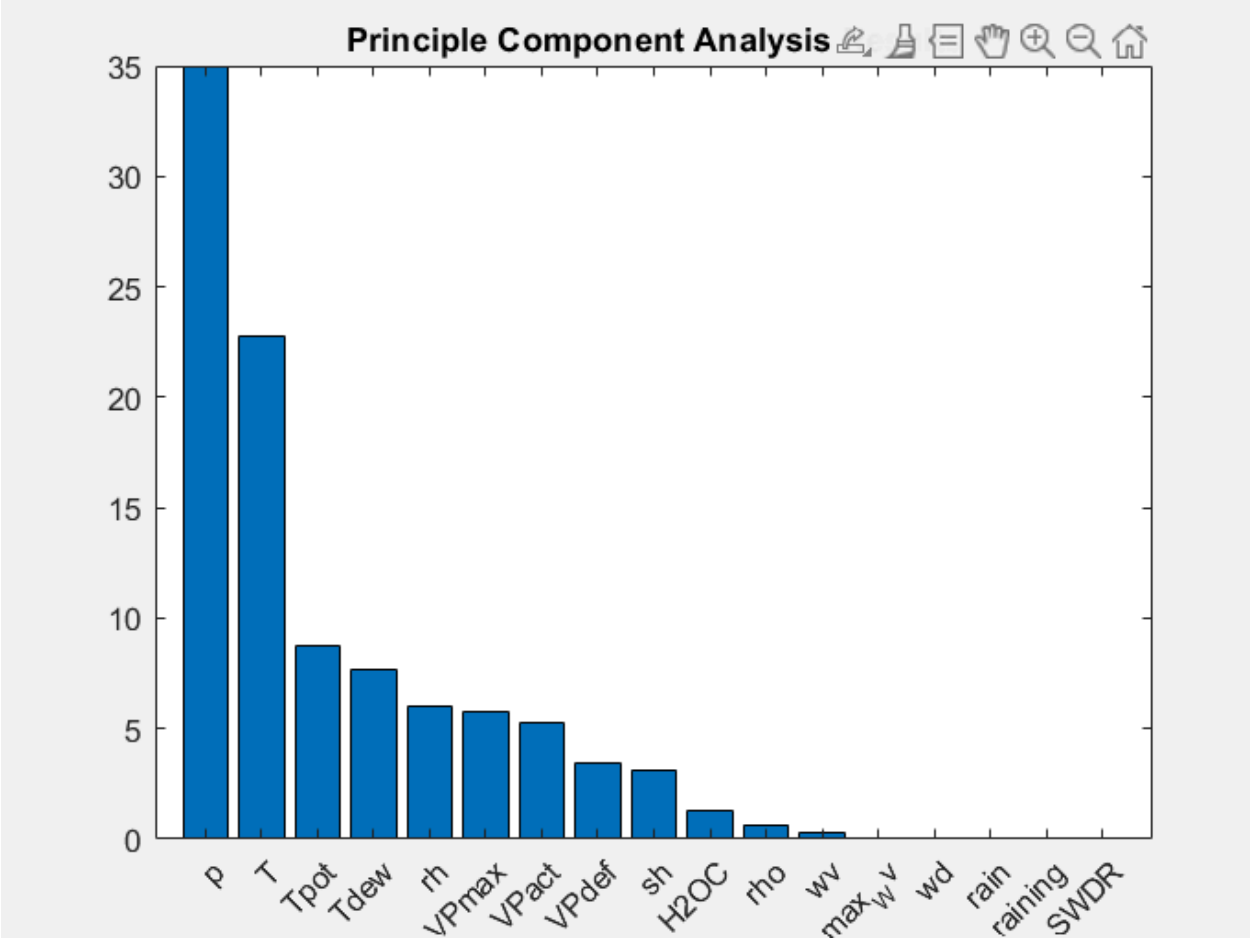
PHY 550

11/15/2024

Assignment 5

Predicting the weather has been a complex problem that humanity has faced since its beginning. Accurately predicting the temperature provides valuable insight on many topics such as knowing what to wear for the day to understanding what plants to grow. For this assignment I analyzed a dataset that took weather measurements every 10 minutes for 1 year. The measurements include Temperature (my target), humidity, rainfall, etc, totally 16 factors. I then created both a shallow neural network and a deep neural network that allowed me to reliably predict future temperatures.

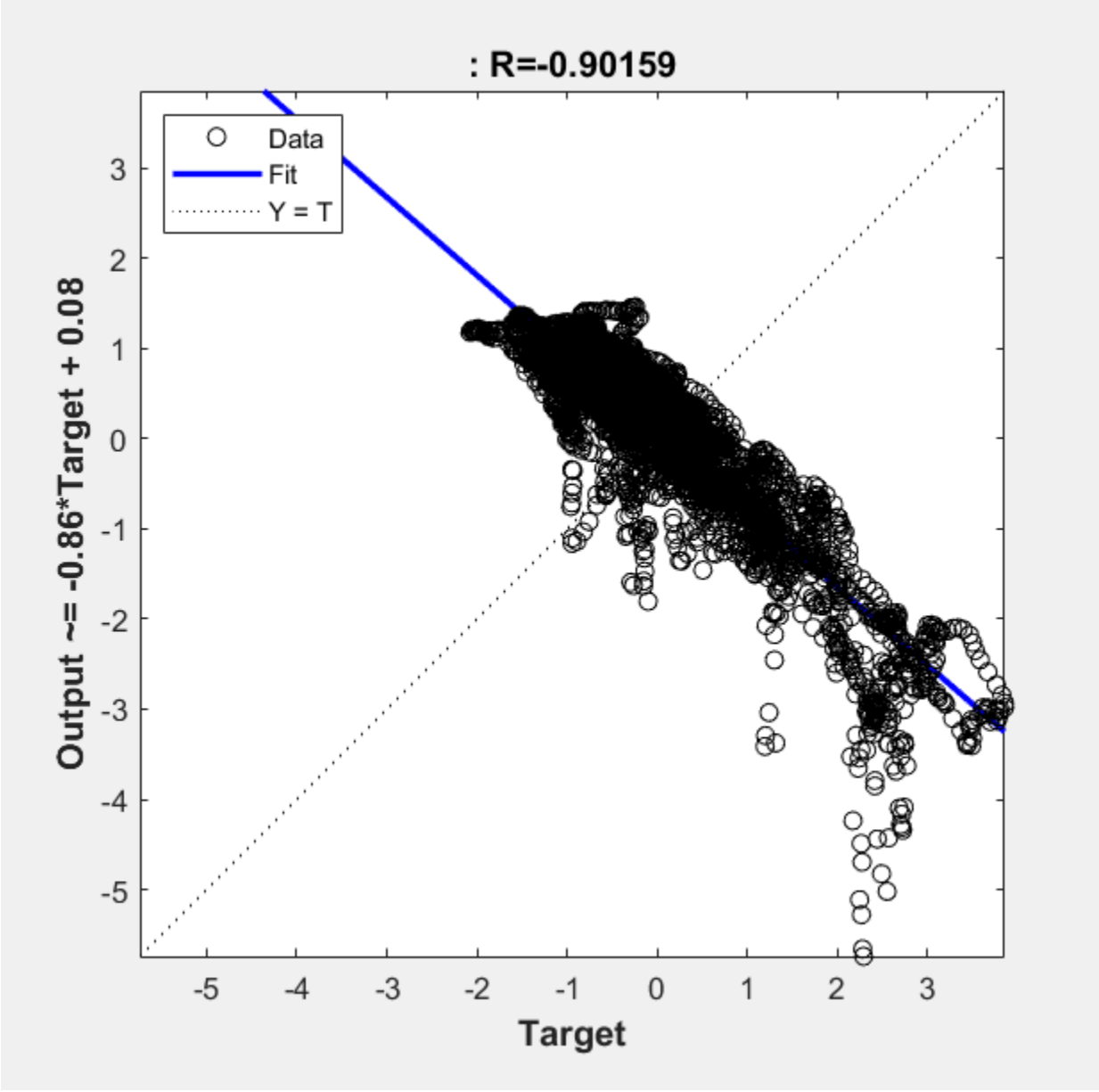
The first challenge I faced while creating this neural network was the size of my dataset. I have access to a fairly limited computer that would not allow me to create a neural network that could analyze all the data in a timely manner. In order to combat this I conducted a principle component analysis, which allowed me to reduce the dimensionality of my data from 16 to 9 while still accounting for 95% of the variability in the data.



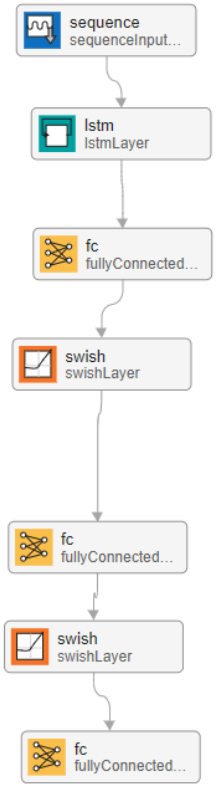
As you can see from my figure, there are only a handful of components that contribute to the variability of the data in a meaningful way. These include the pressure, time, potential temperature (I later removed this from the data as I learned contained look ahead bias.) and humidity. Interestingly enough rainfall had next to no impact on the temperature, I found this to be very strange. From the PCA I removed 8 components, reducing my dataset by more than half, this greatly decreased the time it took me to train the neural network and gave me the computational freedom to find the optimal neuron number.

Once the dimensionality of my dataset was reduced I decided to create a shallow narxnet neural network. I created two nested for loops that trained the neural network with different neuron numbers from 1 to 10, after the loop was completed the optimal neural network number

was found to be 9. Once I had discovered this I was able to fine tune my neural network by changing various options such as epoch number, the training function, as well as the parameter goal. Using various techniques I was able to obtain a regression of -.90159 and an mse of 3.85:



After creating a shallow neural network I decided that a deep neural network could prove to be more accurate. To create my deep neural network I used matlab's deep learning toolbox to create the following network:



Upon training my neural network I was met with yet another computational limit, my computer was unable to train the neural network in a timely fashion. I set the epoch limit to be 500 but after 2 hours the training had only gone to epoch 200, I decided to end the training early and was able to get a regression of .92751, and an MSE of 3.1664, I am confident that with more time and computational power I can get my Neural network to pass the .95 threshold.

For further information on the dataset I used visit the following URL: https://www.kaggle.com/datasets/alistairking/weather-long-term-time-series-forecasting