**Abdulrhman Abdullah**

**(Vector Wind Forecast System)**

**Project proposal**

**4 Dec 2021**

Abstract:

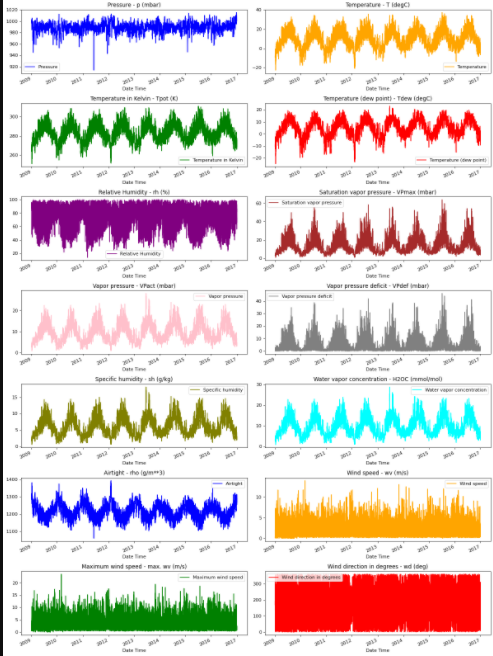
Knowing the direction and speed of the wind is very important in industrial facilities, for example, in the event of a toxic gas leak, residents must be directed to safe areas, the goal of this project is to build deep learning models can predict one-hour wind vectors based one last 48 data to help and improve control of the effects of an environmental accident.

Design:

This project originates from [time series forecasting using TensorFlow](https://www.tensorflow.org/) and the data was recorded by the [Max Planck Institute for Biogeochemistry](https://www.bgc-jena.mpg.de/) are  contains 14 different features such as air temperature, atmospheric pressure, and humidity. These were collected every 10 minutes, beginning in 2003.

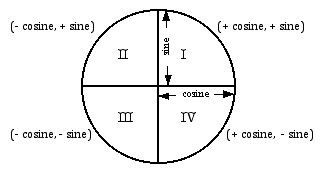
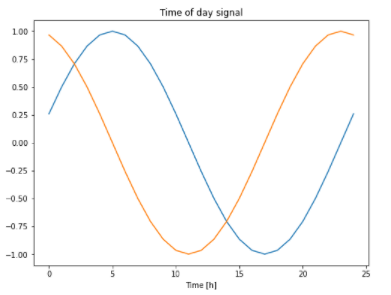
Data:

using Jena Climate dataset recorded by the [Max Planck Institute for Biogeochemistry](https://www.bgc-jena.mpg.de/wetter/). This dataset contains 14 different features such as air temperature, atmospheric pressure, wind speed, wind direction and humidity etc. recorded once per 10 minutes, beginning in 2003 to 2016. First, I just deal with hourly predictions, so start by sub-sampling the data from 10-minute intervals to one-hour intervals, then I process the data by finding any null values or erroneous values and I found some erroneous in the min value of the wind velocity (wv (m/s)) and the maximum value (max. wv (m/s)) columns. This -9999 is likely erroneous. so the velocity should be greater than zero (>=0). I replace it with mean wind velocity and meanMaxWV.



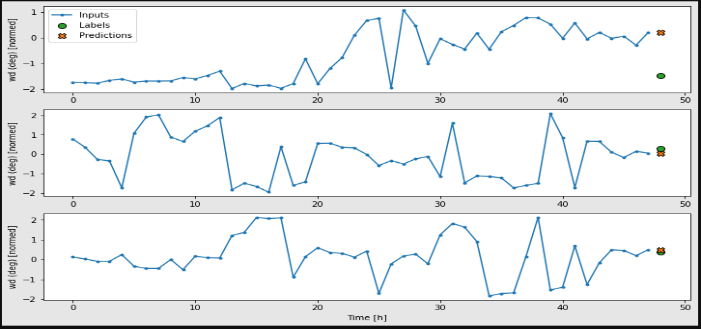
Algorithms:

Then I try to some features engineering to make sure the data appropriately formatted. features engineering for Wind velocity, gives the wind direction in units of degrees. Angles do not make good model inputs: 360° and 0° should be close to each other and wrap around smoothly. Direction shouldn't matter if the wind is not blowing. features engineering for Date and time I do same sine and cosine transforms to clear "Time of day" and "Time of year" signals after I convert it to seconds.



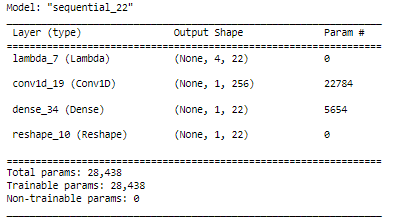
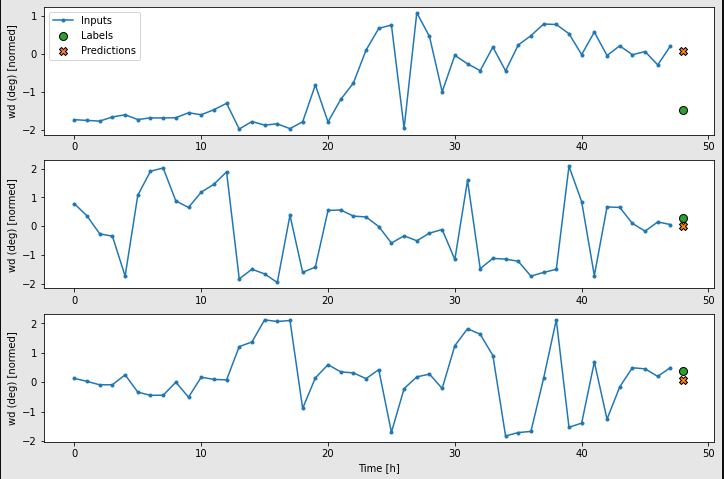
Model:

I split data (70%, 20%, 10%) for the training, validation, and test sets into a training, validation, and test sets. After that I’ll Normalize the data I’ll subtract the mean and divide by the standard deviation of each feature. then build Base Line, deep learning RNN LSTM and CNN models to predict the wind vector and compare the accuracy.

1- Base Line: simple baseline for this task is to repeat the last input time step for the required number of output time steps.

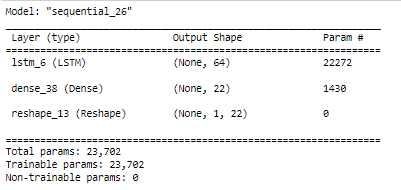
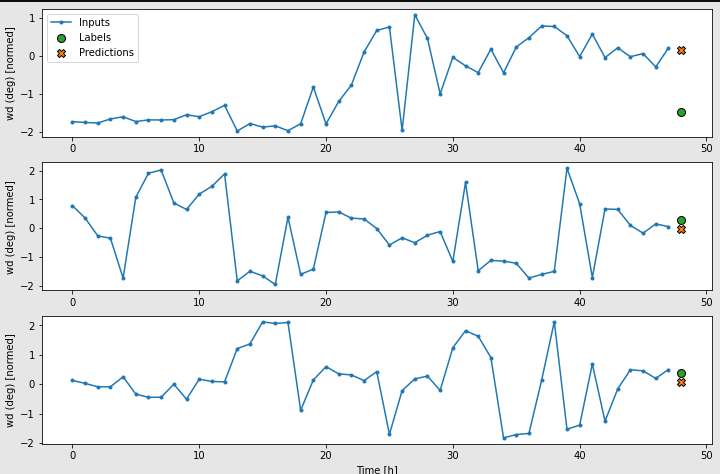
Performance: Base Line : 0.2040

2- CNN: convolutional model makes predictions based on a fixed-width history



Performance: Conv : 0.1696

3- RNN LSTM: recurrent model can learn to use a long history of inputs, if it's relevant to the predictions the model is making. Here the model will accumulate internal state for 48 hours, before making a single prediction for the next 1 hour.



Performance: LSTM: 0.1671

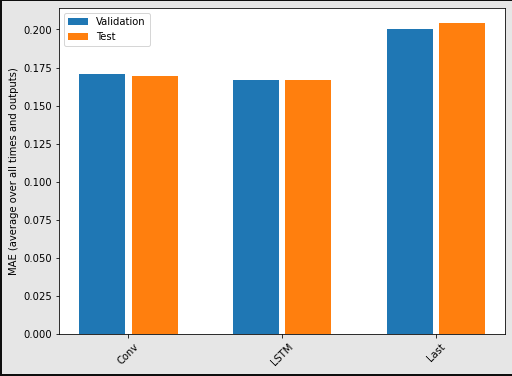
As the following diagram performances are similar but also averaged across output

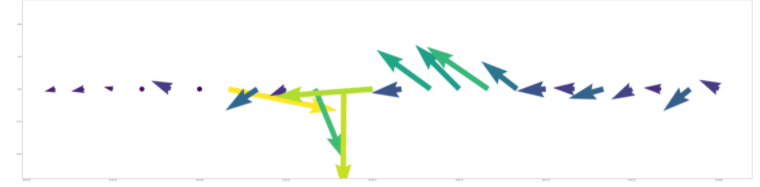
**Tools:**

NumPy, Pandas, sicikit-learn, Matplotlib, Seaborn, Tensorflow, Keras

**Communication:**

In addition to slide I will try to embed this model into Meteorological Data website





Wind vector