

## Introduction







# 1st Run

# 2nd Run 3rd Run

### **EDA**

Data Pulling Feature Engineering

### **MODELING**

What Models? How it failed

### **ARIMA**

Failed Results

### **EDA**

What was different?

### **MODELING**

What was the results?

# TABLE OF CONTENTS

### **Locations Selected**

Cheung Chau Hong Kong Airport King's Park



How I selected? - Looks nice on Google maps



Maximum Temperature
Minimum Temperature
Mean Temperature
Sunshine Values
Wind Speed
Wind Angles
Dates in float format



# Date Transformation

Date were in float. Changes were made to change to datetime

### Seasons

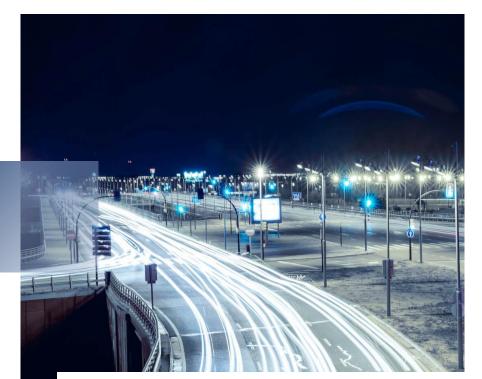
One-hot encoding of season, Spring, Summer, Autumn, Winter

### Scaling

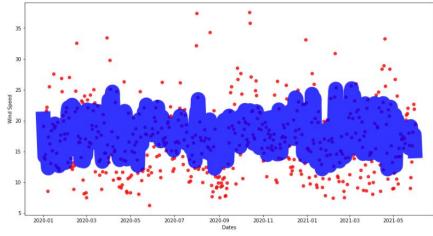
Max-min Standard Scaling

	Year	Month	Day	Temperature	Day_Mm	Day_\$\$	Wind	Temp_Diff	Sun	Wind_log	Spring	Summer	Autumn	Winter	Month_str	Day_str	Date
258	2018	9	16	26.266668	0.517241	0.057767	74.699997	8.166668	0.0	4.313480	0	0	1	0	09	16	2018- 09-16
259	2018	9	17	27.266668	0.551724	0.173301	42.250000	3.866665	1.3	3.743604	0	0	1	0	09	17	2018- 09-17
577	2019	8	1	26.199999	0.000000	-1.677051	43.450001	2.633331	0.0	3.771611	0	1	0	0	08	01	2019- 08-01

# Engineering



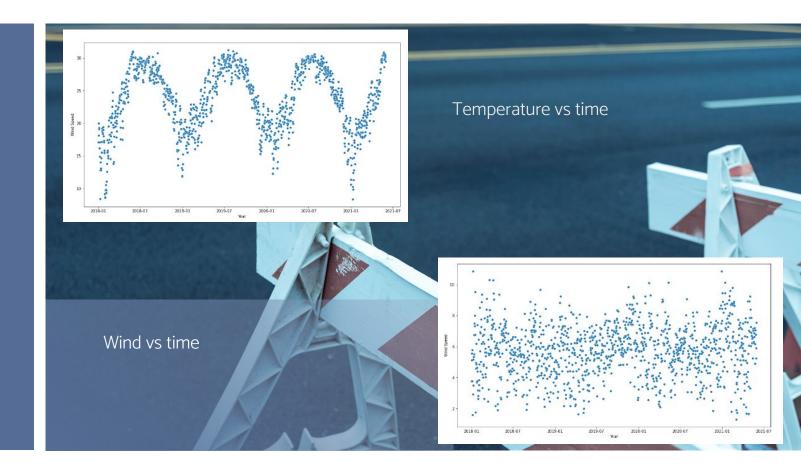
### **Failed Results**

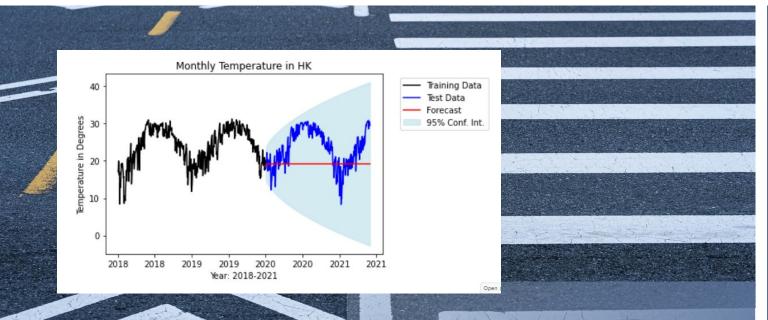


Best score for Ridge Regression on no scaling is - 0.13836111614135033 , and test score is - 0.15358482114740568 , arameters are - {'alpha': 20, 'max\_iter': 10000}

Best score for Random Forest on no scaling is - 0.08997277161300025 , and test score is - 0.13256196614653915 meters are - {'n\_estimators': 20}





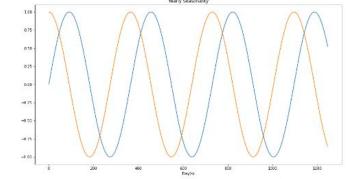


NOPE! Didn't work!

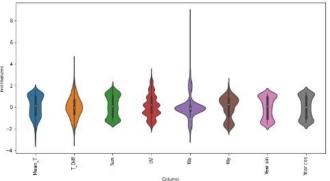


Wind Values can be converted from Speed + Angle to X and Y Vector

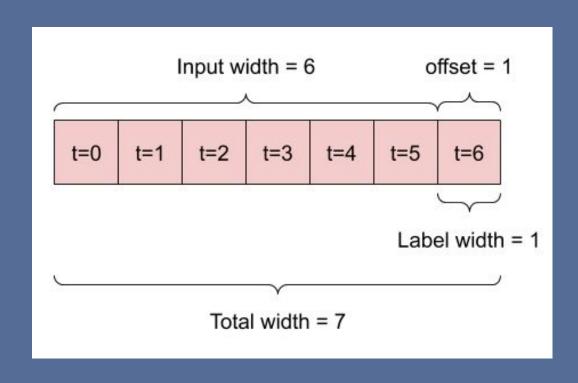
Datetime can be converted into a cosine and sine wave for periodicity



Normalize data with (x-mean)/std

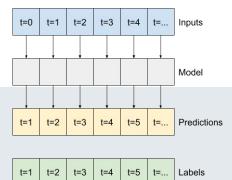


# Windowing for time steps

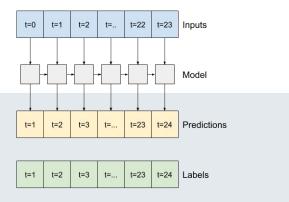


Model

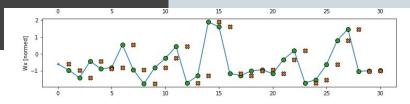
### **Baseline**

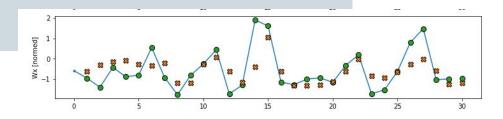


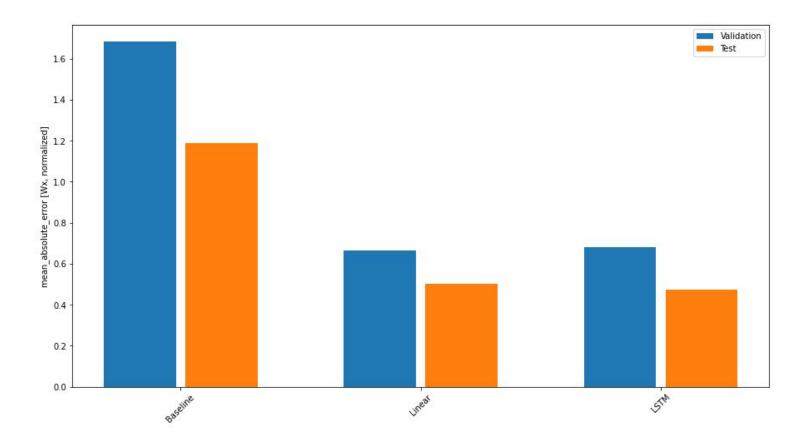
### **LSTM**



### Results

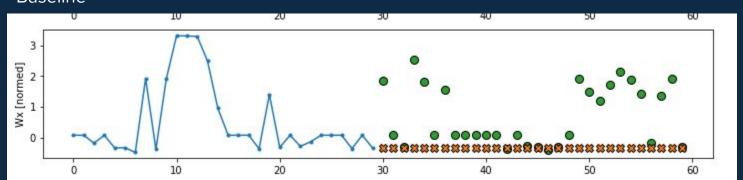




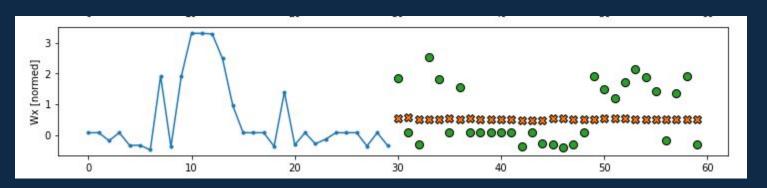


## Multistep Model





### LSTM



### Conclusion and Drawbacks

Due to the lack of feature variables, more could be done to enhance the data model to achieve a better fit.

For the current project, I am using a basic LSTM model, more can be done to include more layers to enhance the parameters further.

Stronger LSTM models can be explored upon to increase better fitting.

Wind measurements in coastal areas might not be a good fit because of the data varies too much on a single time step

