**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

Belagavi-590014, Karnataka



**MINI PROJECT REPORT**

**OF**

**“WIND STORM PREDICTION”**

Submitted by

**KUMARI SUPRIYA** (3GN16CS031)

**POOJITA GURAM**  (3GN16CS057)

**NEHA M** (3GN16CS048)

**NISHCHITH KULKARNI** (3GN16CS049)

Under the guidance of

**Prof. Gururaj.S**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**GURU NANAK DEV ENGINEERING COLLEGE**

**BIDAR-585403, KARNATAKA**

**2019-2020**

**VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI**

**GURU NANAK DEV ENGINEERING COLLEGE**

**BIDAR-585401, KARNATAKA**

****

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**CERTIFICATE**

This is to certify that the mini project report entitled “**WIND STORM PREDICTION**” is a bonafied work carried out by KUMARI SUPRIYA (3GN16CS031), POOJITA GURAM (3GN16CS057) ,NEHA M (3GN16CS048) AND NISHCHITH KULKARNI (3GN16CS049) in practical fulfilment for the award of IA marks for **MACHINE LEARNIG (15CS73)** in COMPUTER SCIENCE AND ENGINEERING of the **GURU NANAK DEV ENGINEERING COLLEGE**, **BIDAR** during the year 2019-2020. It is certified that all the corrections/suggestions indicated for the internal assessment have been incorporated in the report. The Mini Project Report has been approved as it satisfies the academic requirements.

Signature of Guide .....................................

**(Prof. GURURAJ.S)**

**MARKS AWARDED**

**KUMARI SUPRIYA** (3GN16CS031) = \_\_\_\_\_\_\_

**POOJITA GURAM** (3GN16CS057) = \_\_\_\_\_\_\_

**NEHA M** (3GN16CSO48) = \_\_\_\_\_\_\_

**NISHCHITH KULKARNI** (3GN16CS049) = \_\_\_\_\_\_

**PROBLEM STATEMENT**

1. **Wind** **storm prediction depend on several variable which are not directly measurable.**
2. **Weather storm prediction is the application of current technology and science by which we can predict the state of wind storm for a given time and for a given location.**
3. **We cannot accurately predict wind storm .**

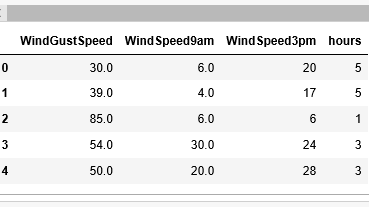
**Description of the dataset**

**We took the data set from kaggle and the they have collected the data from the national weather service .Some information appearing in storm data may be provided by or gathered from source outside the national weather services(NWS),such as media,law enforcement and / or other government agencies, emergency managers ,private companing,individuals,etc.**

**In this project we have used the data and they are WindGustspeed,WindSpeed9am,Windspeed3pm,Hours.**

**WindGustspeed is the increase in speed of wind,usually less than 20 seconds.WindSpeed9am it is the speed of the wind at 9am.WindSpeed3pm it is the speed of the wind at 3pm.Hours is the time at which the wind storm will be visible.**

**Dataset of Wind Storm Prediction**

****

**Methodology**

The ML process can be represented as a series of steps:

•

**Data acquisition and preprocessing**: where possibly different data sets and modalities are integrated, cleaned of outliers, etc.

•

**Feature selection and extraction**: important signals and characteristics are identified and extracted from the data.

•

**Model selection**: an appropriate model is chosen, taking into consideration the task to be solved.

•

**Validation**: a [performance measure](https://www.sciencedirect.com/topics/engineering/performance-measure) is used that is specific to the task, including accuracy (classification) and [mean absolute error](https://www.sciencedirect.com/topics/engineering/mean-absolute-error) (regression), evaluated on a validation set of data.

**Steps in Wind Storm Prediction**

1. **First step in wind strom prediction.**

* **Pandas: pandas is the python package providing fast, flexible and expressive data structures designed to make working with relational or labeled data both easy and intuitive.**
* **Seaborn : It is a python data visualization based matplotlib . It provide high level interface for drawing attractive and informative statistical graphics.**
* **Matplotlib: It is an amazing visualization library in python for 2D plots of arrays it is a multi platform data visualization library built on arrays.**
* **NumPay: It is a general purpose array processing package . It provide high performance. It can also be used as an efficient multidimensional container of generic data.**
* **Sickit- learn: it is afree machine learning library for python.it features various algorithms like vector machine , random forest and k-neighbours and it also support library like Numpy and Scipy.**

1. **Splitting the data set into the training set and test set**

**Training dataset: It is used to make sure the machine recognizes patterns in the data, the cross validation data is used to insure better accuracy.**

**x-train: it is a training data set**

**y-tarin: it is set of labels to all the data in x-train.**

**x-test:**

**y-test:**

1. **Fitting simple linear regression to the training set**

**Linear regression:It is a linear approach to modeling the relationship between a scalar response.**

**The case of one explanatory variable is called simple linear regression.**

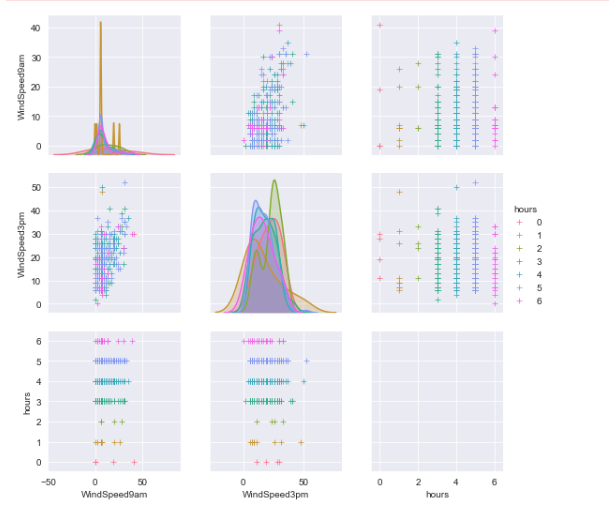
**The case of many explanatory variable is called multiple linear regression.**

**In this we are using simple linear regression.**

1. **Precdicting the test set result**
2. **Visualizing the training set result**

**Experminting result**

****

****

**Code Listing:**

#importing libraries

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

dataset = pd.read\_csv('modified.csv')

X = dataset.iloc[:, 0:3].values

y = dataset.iloc[:,3].values

import seaborn as sns

%matplotlib inline

sns.set\_palette("rainbow")

sns.set\_style('darkgrid')

sns.pairplot(dataset)

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size = 0.2, random\_state = 0)

from sklearn.linear\_model import LinearRegression

regressor = LinearRegression()

regressor.fit(X\_train, y\_train)

# Predicting the Test set results

y\_pred = regressor.predict(X\_test)

from sklearn.metrics import r2\_score

mm=r2\_score(y\_test, y\_pred)

import seaborn as sns

sns.set\_palette('husl')

tmp = dataset.drop('WindGustSpeed', axis=1)

g = sns.pairplot(tmp, hue='hours', markers='+')

plt.show()

**Conclusion**

**Hence we have concluded that we can predict the wind storm for a given period of time and for a given place or area . And how many hours it may be visible.**