



slington college
(इस्लिङ्टन कलेज)

Module Code & Module Title

CU6051NA - Artificial Intelligence

Assessment Weightage & Type

80% Individual Coursework

Year and Semester

2018-19 Autumn

Student Name: Prayash Bikram Shah

London Met ID:16033180

College ID: SITY1C216131

Assignment Due Date: 11th February, 2019

Assignment Submission Date: 11th February, 2019

I confirm that I understand my coursework needs to be submitted online via Google Classroom under the relevant module page before the deadline in order for my assignment to be accepted and marked. I am fully aware that late submissions will be treated as non-submission and a mark of zero will be awarded.

Table of Contents

1. INTRODUCTION	1
1.1 Aims:	1
1.2 Objectives:.....	1
2. BACKGROUND	2
2.1 Research Conducted:	2
2.2 Research on similar system or works:	3
3. SOLUTION	5
3.1 AI Algorithm:.....	5
3.2 Pseudocode of the solution	7
3.3 Diagrammatical Representation of Algorithms:	9
3.4 Development Process.....	10
3.5 Achieved results	15
4. CONCLUSION	16
5. REFERENCES	17
6. BIBLIOGRAPHY	19

Table of Figures

Figure 1 Artificial Intelligence (IBM , 2018)	2
Figure 2 Machine Learning (The MathWorks, Inc., 2019)	3
Figure 3 Weather Forecast System (KMA, 2009).....	4
Figure 4 AI for Weather Forecast (Joshi, 2018)	5
Figure 5 Detailed Machine Learning algorithms (The MathWorks Inc, 2019).....	6
Figure 6 Clustering (Kurama, 2018)	7
Figure 7 Agglomerative Hierarchical Clustering (Kassambara, 2018)	9
Figure 8 Flowchart for agglomerative clustering (Silva, 2017).....	9
Figure 9 Importing Libraries	10
Figure 10 Reading datasets	10
Figure 11 Dataset contents	11
Figure 12 Data Filtering.....	12
Figure 13 Data Filtering 1	12
Figure 14 Data Filtering 2.....	12
Figure 15 Conversion and Description of columns.....	13
Figure 16 Preprocessing, Training and agglomerative	13
Figure 17 Clustering 1	14
Figure 18 Clustering 2.....	14
Figure 19 Classification and accuracy.....	14
Figure 20 Comparison of predicted and original.....	15
Figure 21 Accuracy of train set.....	15
Figure 22 Accuracy of test set.....	15

1. INTRODUCTION

Prediction has been one of the most important field that the humans have been trying to excel on it and figure out a way to give out the most accurate result. The more the accurate result the more dependent and guarantee we can have and brace the future comings of ours. A prediction or forecast is not something like a fortune teller, people in this modern era believe in facts and figures. Facts and figures are delivered when we pass some statistical data through a certain algorithm and the output will be more reliable and believable by the people of this era rather than depending upon some fortune teller prediction.

Weather forecast maybe a minor topic in Artificial Intelligence field but if we look into it then weather forecast is something that almost everyone looks up on weather widget on their phone, tablet or watches. I am also one of the people who want to know the weather for the next day and what's is going to be like so that I can prepare myself or plan accordingly. Not just for the next day, even when I wake up, I feel it's important to know what will be the weather like the whole day to brace myself for uncertainty of rain or thunderstorm that may hit my schedule for the day. Weather forecast has been playing and important role in our day to day life. Not just an individual need this or uses this, even business and other entities relies on an accurate weather forecast.

1.1 Aims:

- Research about weather forecast.
- Why is it important in today's world?
- Learn and gain proper idea of the concept that is used to solve and gain accuracy on weather forecast.

1.2 Objectives:

- Research about AI.
- Research about the concepts that i can use to solve the problem.
- Select a concept/algorithm.
- Explain how will I use the concept to solve the problem.

2. BACKGROUND

2.1 Research Conducted:

Artificial Intelligence has been in the hype in recent era though this has been a part of the technology for a long time and is responsible for the development and drastic upgrade of the technologies till date. AI can be defined as some algorithm or design and study of an intelligent algorithm. In 1956, it was defined as the science and engineering of making intelligent machines by John McCarthy. (ScienceDaily, 2018)

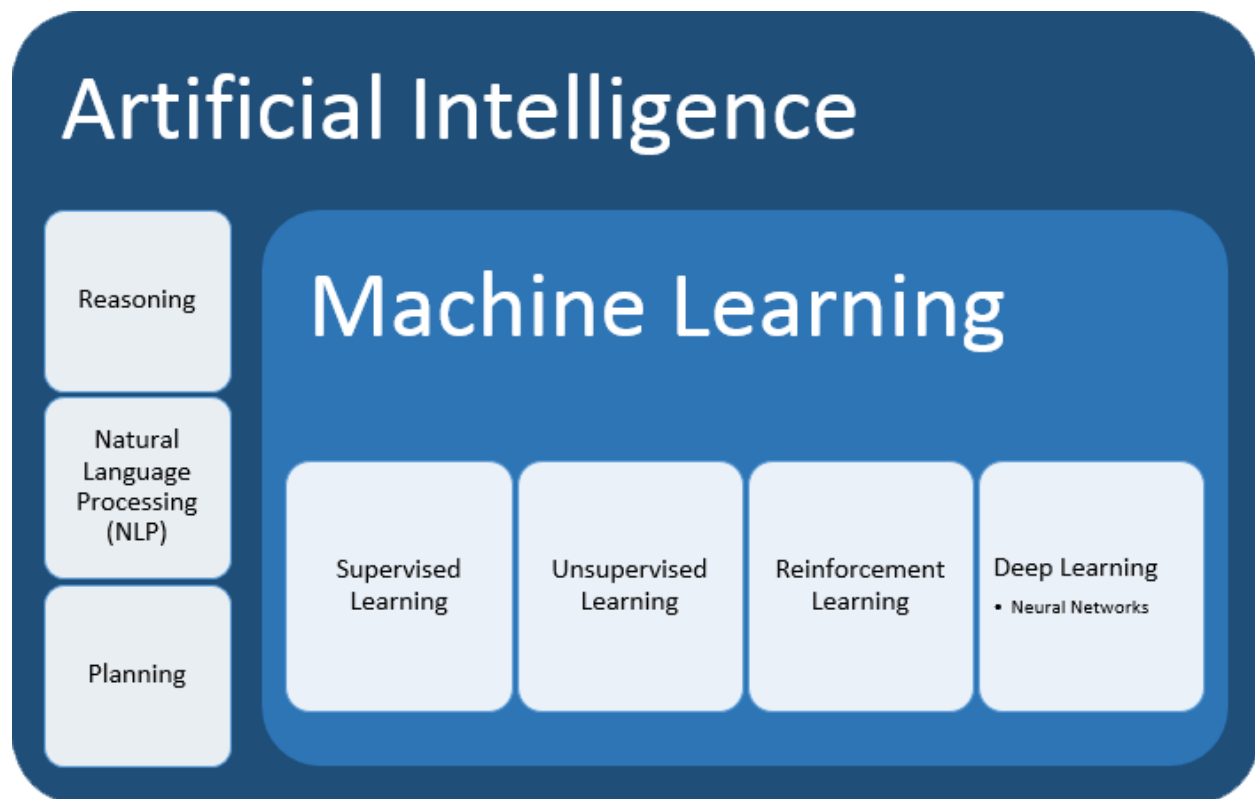


Figure 1 Artificial Intelligence (IBM , 2018)

Machine Learning is a part of Artificial Intelligence which uses data and analyze them to give an output. In Machine Learning data is trained with certain algorithms that give out a model and then users can give an input to that model to get the results. Machine Learning has four categories Supervised and Unsupervised Learning, Reinforcement Learning and Deep Learning. (IBM , 2018)

In this project I have chosen to research about Weather forecast which is a product of AI and due to which we have been cautious about the weather we will be facing from ahead

of time we are on. A small gain of accuracy of weather forecast can give us a huge advantage for an organization or an individual.

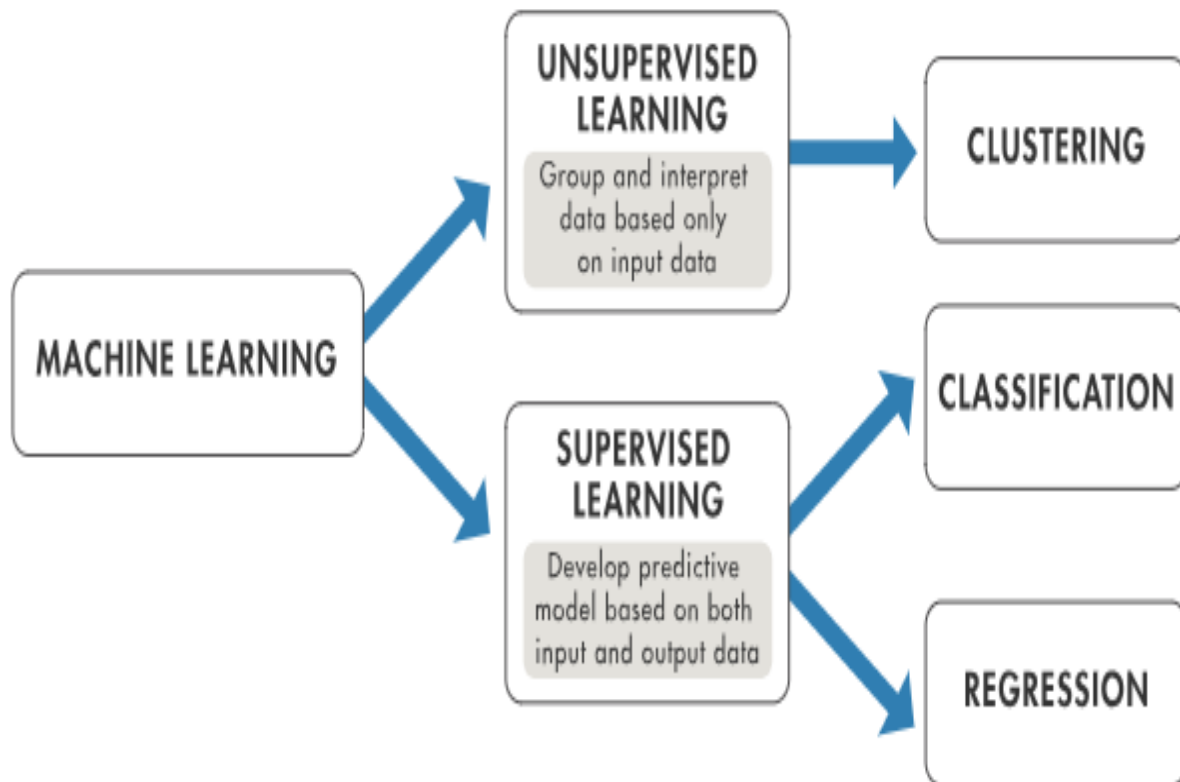


Figure 2 Machine Learning (The MathWorks, Inc., 2019)

2.2 Research on similar system or works:

Weather Forecast is not something that I thought of it first, this has been done by many agencies in the world. People have been trying to develop a model with high accuracy and a reliable system to predict. There are already many models for weather forecast such. Korea Meteorological Administration has issued different types of weather forecast like short, medium and long-range weather forecast. In 1999 Korea Meteorological Administration began to run numerical weather prediction (NWP). Such Models are already out there giving out some accurate result to us. Below figure 3 shows a general idea how weather forecast is done but there is no state of algorithms used in it though some certain algorithms are being used to make models and predict the data fetched to tis such kind of systems. (KMA, 2009)

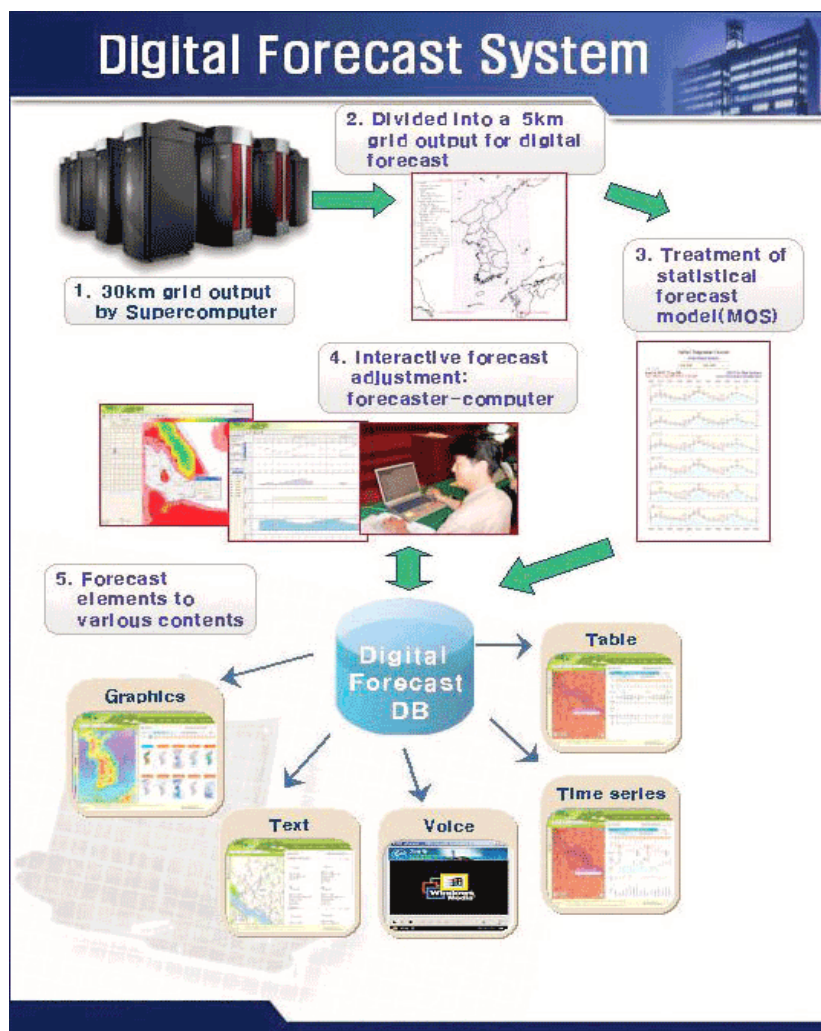


Figure 3 Weather Forecast System (KMA, 2009)

3. SOLUTION

AI has been a game changer in the field of technology bringing the capability of finding solutions for the day to day problems. Being able to generate some facts through a certain algorithm and finding a pattern to predict things makes AI a perfect medium to forecast weather for us. An algorithm is nothing without giving it some data. A set of data of weather should be fetched into an algorithm of AI to let it find a certain pattern. That pattern will help us with the prediction of weather. Though the output will not be 100% accurate but even slightest accuracy can help to believe in the prediction and change the future and fate of the human kind. (Joshi, 2018)

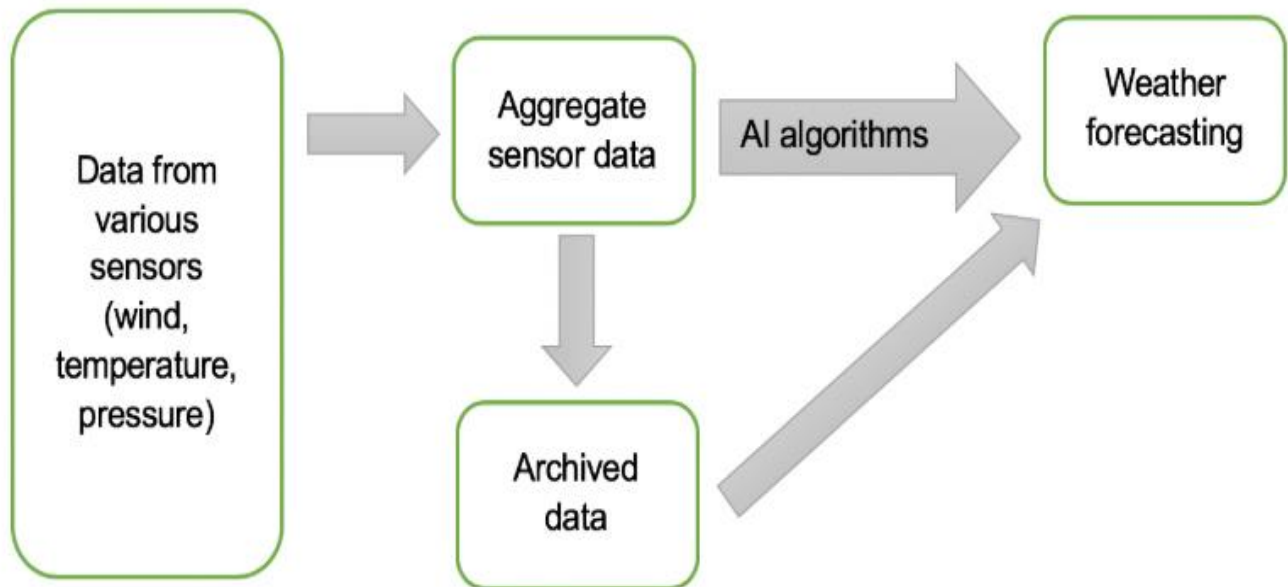


Figure 4 AI for Weather Forecast (Joshi, 2018)

Weather forecast can be done using multiple algorithms from Machine learning such as Decision Tree, Linear Regression, Neural Networks, Clustering. I have chosen Clustering which is a part of Unsupervised Learning. (Expert System, 2018)

3.1 AI Algorithm:

The approach that I will be taking is Unsupervised Learning which is a category of Machine Learning. Unsupervised Learning is mostly used when analyzing a sequence,

data mining for sequence and pattern mining. One of the frequent and common method of this is cluster analysis. (The MathWorks Inc, 2019)

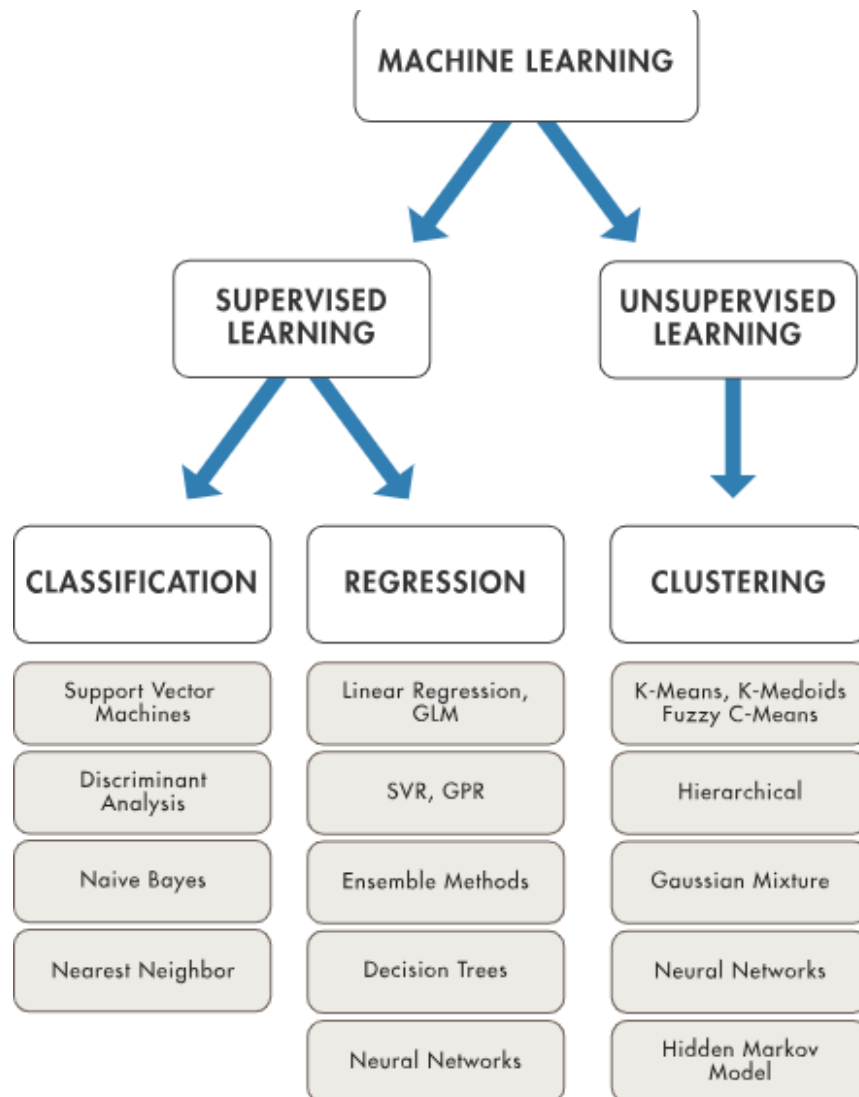


Figure 5 Detailed Machine Learning algorithms (The MathWorks Inc, 2019)

A common method of unsupervised learning, Clustering will be my approach to the solution for making a weather forecast system using certain data sets. In this clustering algorithm the data are manipulated in such a way that there are groups of data and this grouping is done with the similarities and these are assigned into clusters for the further process. (Kurama, 2018)

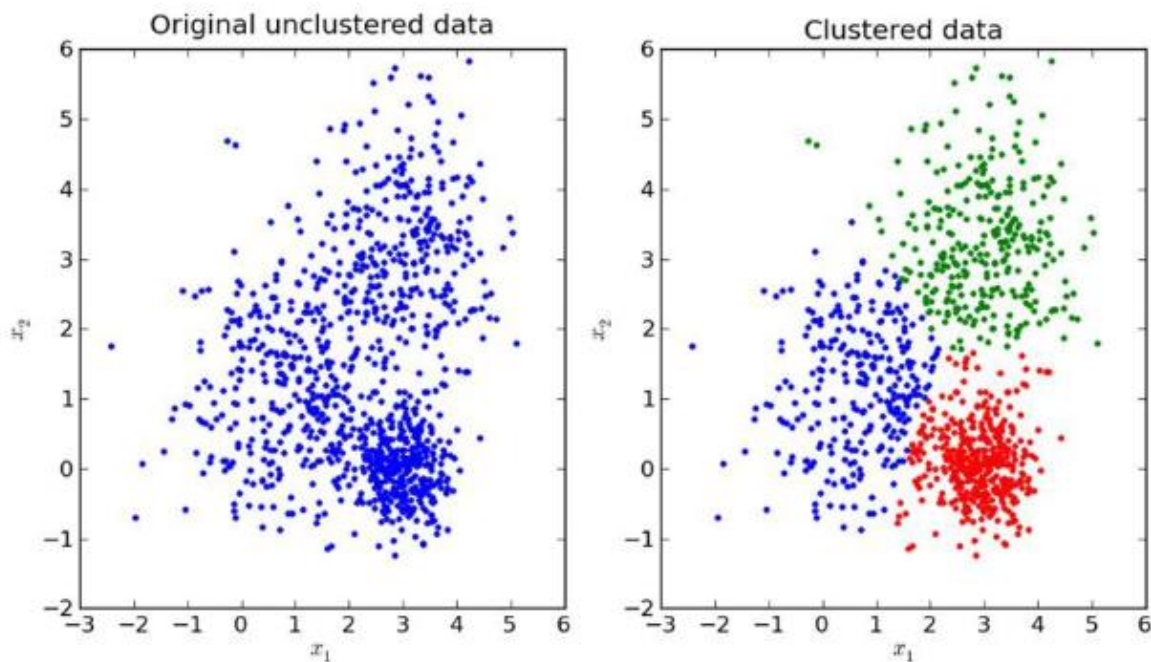


Figure 6 Clustering (Kurama, 2018)

Clustering also has different types of algorithm such as K means and hierarchical clustering being the most common among all the types. Choosing one of these clustering techniques I will be solving the solution and creating a system to forecast weather. The Data sets for the algorithm will be extracted from Kaggle. Below are the diagram explaining the K means clustering and agglomerative clustering.

3.2 Pseudocode of the solution

```
import numpy library as np

import pandas library as pd

import matplotlib.pyplot library as plt

using sklearn.model_selection Scikit-learn library import train_test_split

using sklearn.cluster Scikit-learn library import Agglomerative Clustering

reading dataset weatherdata.csv

choosing specific columns from dataset
```

assigning chosen columns to data variable

function to find non-numeric data

function to find row containing non-numeric data

assigning values to empty rows variable.

Function to replace values where T is there

Classifying the events into different category

Dropping empty rows.

Converting the all the columns to numeric and describing it

Using Scikit-learn library import preprocessing

Normalize the input data

Splitting the data to train and test sets

Agglomerative Clustering and plotting the data

Mapping the cluster numbers to the categories.

Clusters center mapping

Function to calculate the distance between centroids and data points

Adding tolerance so that we can classify the

Function to calculate accuracy

Displaying Accuracy of trained data

Predicting events using agglomerative algorithm

Displaying final accuracy of test data.

3.3 Diagrammatical Representation of Algorithms:

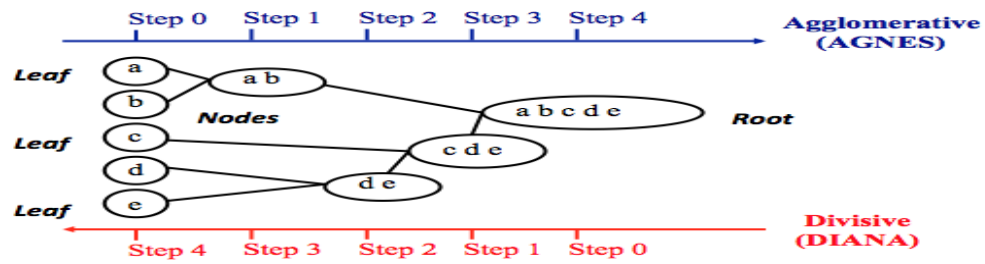


Figure 7 Agglomerative Hierarchical Clustering (Kassambara, 2018)

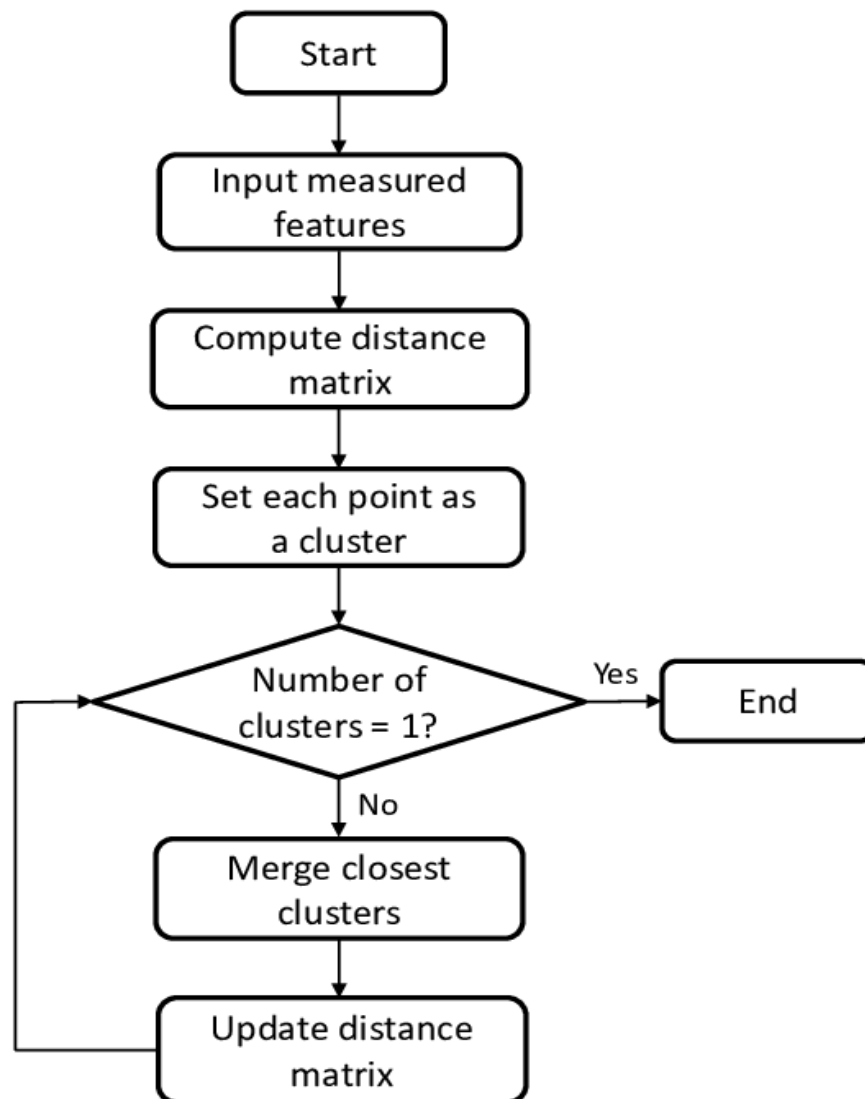


Figure 8 Flowchart for agglomerative clustering (Silva, 2017)

3.4 Development Process

This problem was solved using python 3 on Jupyter Notebook. The main problem while going through this project was to find a dataset that will help me demonstrate the possibility of weather forecast through unsupervised learning. To find a dataset, I used Kaggle where I found a dataset of the Austin KATT station from 2013-02-21 to 2017-07-31.

After finding the dataset, first of all the libraries needed were imported. Libraries such as pandas was used which was needed to manipulate the dataset in the jupyter notebook using python 3. With using pandas, numpy is also one of the libraries that I will be importing which is a fundamental package for scientific computing with python, libraries for training data, agglomerative clustering and data representation were also imported. The below figure shows the code importing the libraries.

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.cluster import AgglomerativeClustering
```

Figure 9 Importing Libraries

After importing the libraries, the dataset was read and the below figure shows the code for it.

```
In [2]: df = pd.read_csv('weatherdata.csv')
```

Figure 10 Reading datasets

The CSV file was first reviewed and looked into it and there were some minor details that could bring complication while manipulating this dataset through our algorithm. As per research conducted data cleaning was the most important part while doing weather forecast using unsupervised learning. When it comes down to Machine Learning the phrase “garbage in, garbage out” is suitable as it can be to it. So, data cleaning was one of the most important process to be done.

The CSV file contained a lot of data and only few of them were necessary for us or useful to us. Below figure shows the data that were on the dataset.

Content

Contains the:

Date (YYYY-MM-DD)

TempHighF (High temperature, in Fahrenheit)

TempAvgF (Average temperature, in Fahrenheit)

TempLowF (Low temperature, in Fahrenheit)

DewPointHighF (High dew point, in Fahrenheit)

DewPointAvgF (Average dew point, in Fahrenheit)

DewPointLowF (Low dew point, in Fahrenheit)

HumidityHighPercent (High humidity, as a percentage)

HumidityAvgPercent (Average humidity, as a percentage)

HumidityLowPercent (Low humidity, as a percentage)

SeaLevelPressureHighInches (High sea level pressure, in inches)

SeaLevelPressureAvgInches (Average sea level pressure, in inches)

SeaLevelPressureLowInches (Low sea level pressure, in inches)

VisibilityHighMiles (High visibility, in miles)

VisibilityAvgMiles (Average visibility, in miles)

VisibilityLowMiles (Low visibility, in miles)

WindHighMPH (High wind speed, in miles per hour)

WindAvgMPH (Average wind speed, in miles per hour)

WindGustMPH (Highest wind speed gust, in miles per hour)

PrecipitationSumInches (Total precipitation, in inches) ('T' if Trace)

Events (Adverse weather events. ' ' if None)

This dataset contains data for every date from 2013-12-21 to 2017-07-31.

Figure 11 Dataset contents

I just needed few of the data so I had to get rid of the unnecessary data for me from the dataframe. The below figure shows code that is just selecting the data that I used from the dataframe.

```
In [3]: usefultdata = ['TempAvgF', 'DewPointAvgF', 'HumidityAvgPercent', 'SeaLevelPressureAvgInches', 'VisibilityAvgMiles', 'WindAvgMPH', 'PrecipitationSumInches', 'PrecipitationTrace']
data = df[usefultdata]
```

```
In [ ]:
```

Figure 12 Data Filtering

After getting rid of unnecessary columns, the columns that I was using had empty data. One of the problems on the dataset was that a column had values 'T' which stands for trace and this could be very disturbing while manipulating the data for the prediction. After a lot of research, I could not find any solution to it so instead I took the help from the dataset source where the manipulation of this dataset was done. The below figure shows the code that I took reference from the dataset source.

```
def nonnumeric(usefultdata, data):
    r = np.zeros(data.shape[0], dtype=bool)
    for column in usefultdata:
        r = r | pd.to_numeric(data[column], errors='coerce').isnull()
    return r
def dfrowsnonnumeric(dataframe):
    return data[nonnumeric(usefultdata, data)]
emptyrows = dfrowsnonnumeric(data).index.values
def numberorzero(value):
    try:
        parsed = float(value)
        return parsed
    except:
        return 0
data['PrecipitationSumInches'] = df['PrecipitationSumInches']
trace = nonnumeric(['PrecipitationSumInches'], data).astype(int)
data = data.assign(PrecipitationTrace=trace.values)
data['PrecipitationSumInches'] = data['PrecipitationSumInches'].apply(numberorzero)
data.iloc[0:10,:]
```

Figure 13 Data Filtering 1

Distinguished Category was needed so the events were separated and classified.

```
In [6]: events = df[['Events']].replace(' ', 'None')
differentevents = set()
for value in events.Events.value_counts().index:
    splitted = [x.strip() for x in value.split(',') ]
    differentevents.update(splitted)
differentevents = set(differentevents)
individualevents = pd.DataFrame()
for typeofevent in differentevents:
    allevents = events.Events.str.contains(typeofevent)
    individualevents = pd.concat([individualevents , pd.DataFrame(data={typeofevent: allevents.values})], join='outer', axis=1)
```

Figure 14 Data Filtering 2

Numerical data was needed so all columns were converted to numerical data types which was needed. And using info (), I viewed the data types to check.

```
In [7]: finaldata = data.drop(emptyrows)
        finalevents = individualevents.drop(emptyrows)

In [8]: finaldata = finaldata.apply(pd.to_numeric)
        finaldata.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1307 entries, 0 to 1318
Data columns (total 8 columns):
TempAvgF          1307 non-null int64
DewPointAvgF      1307 non-null int64
HumidityAvgPercent 1307 non-null int64
SeaLevelPressureAvgInches 1307 non-null float64
VisibilityAvgMiles 1307 non-null int64
WindAvgMPH        1307 non-null int64
PrecipitationSumInches 1307 non-null float64
PrecipitationTrace 1307 non-null int32
dtypes: float64(2), int32(1), int64(5)
memory usage: 86.8 KB
```

Figure 15 Conversion and Description of columns

The data needed to be normalized using sklearn library of preprocessing. The training of the model was done which means the data was distributed into train and test sets. After that I looked if it was possible to figure out events using unsupervised learning which is Agglomerative Clustering.

Preprocessing data

```
In [9]: from sklearn import preprocessing
        dvalues = finaldata.values
        min_max_scaler = preprocessing.MinMaxScaler()
        finaldata = pd.DataFrame(min_max_scaler.fit_transform(finaldata), columns=finaldata.columns, index=finaldata.index)
```

Training the Model

```
In [10]: random_state = 42 #gives out the same output as the first time making the split.
        X_train, X_test = train_test_split(finaldata, test_size=0.2, random_state=random_state)
        y_train, y_test = train_test_split(finalevents, test_size=0.2, random_state=random_state)
        numberofclusters = len(differentevents)
```

Agglomerative Clustering

```
In [11]: clustering = AgglomerativeClustering(n_clusters=numberofclusters, linkage="average").fit(X_train)
        resultac = pd.DataFrame(clustering.labels_)
        resultac.iloc[:,0].value_counts().plot.line()
```

Figure 16 Preprocessing, Training and agglomerative

After this the data was plotted and checked if it could be distinguished and then clustering was done such as centroids, calculation of distance of centroid and datapoints was done.

Multi-Lable Clustering

```
In [12]: eventorder = finalevents.sum().sort_values(ascending=False).index
acorder = resultac.iloc[:,0].value_counts().index
clusterdistribution = {}
for i in range(numberofclusters):
    clusterdistribution.update({acorder[i]:eventorder[i]})
clusterdistribution
```

```
Out[12]: {0: 'None', 4: 'Rain', 1: 'Thunderstorm', 3: 'Fog', 2: 'Snow'}
```

Calculating Clusters Centroids

```
In [13]: clustercenter = {}
for key in clusterdistribution:
    cluster_indices = resultac.loc[resultac[0] == key].index
    cluster_data = X_train.iloc[cluster_indices]
    mean = cluster_data.mean(axis=0).values
    clustercenter.update({key:mean})
clustercenter
```

Figure 17 Clustering 1

```
In [14]: def distance_from_centroid_to_datapoints(dataframe):
    clusterdist = np.zeros((dataframe.shape[0], numberofclusters))
    for i in range(dataframe.shape[0]):
        for key in clusterdistribution:
            distance = np.linalg.norm(dataframe.iloc[[i]].values[0]-clustercenter[key])
            clusterdist[i,key] = distance
    column_names = [clusterdistribution[k] for k in clusterdistribution]
    return pd.DataFrame(clusterdist, index=dataframe.index, columns=column_names)
distances = distance_from_centroid_to_datapoints(X_train)
distances.head()
```

Figure 18 Clustering 2

My classification is the columns with minimum distance and to able to classify a particular data to the clusters I added tolerance. A function to check accuracy was developed.

```
def eventgroup(distances_dataframe):
    return distances_dataframe.apply(lambda x: x<x.min()*1.02, axis=1)
classification = eventgroup(distances)
X_train = classification.reindex(sorted(classification.columns), axis=1)
y_train = y_train.reindex(sorted(y_train.columns), axis=1)
def accuracy(X, y):
    comparison = X == y
    val_counts = comparison.all(axis=1).value_counts()
    percentageCorrect = val_counts.at[True] / X.shape[0] * 100
    return percentageCorrect
```

Figure 19 Classification and accuracy

Finally, the predicted data derived from unsupervised learning with agglomerative clustering was compared with original events and accuracy was checked.

```

distancesDf = distance_from_centroid_to_datapoints(X_test)
classification = eventgroup(distancesDf)
X_test = classification.reindex(sorted(classification.columns), axis=1)
y_test = y_test.reindex(sorted(y_train.columns), axis=1)
fig, ax = plt.subplots(1, 2, figsize=(15, 5))
y_test.sum().plot.line(ax=ax[0], title="Original")
X_test.sum().plot.line(ax=ax[1], title="Forecast")

```

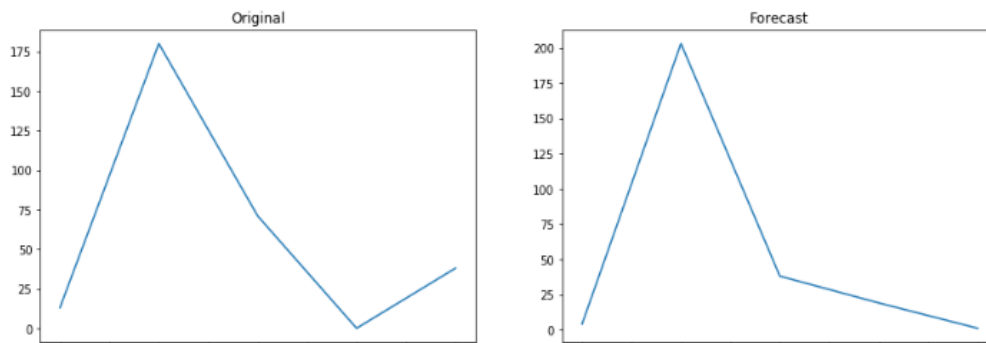
Figure 20 Comparison of predicted and original

3.5 Achieved results

```
In [16]: accuracy(X_train, y_train)
```

```
Out[16]: 65.07177033492823
```

Figure 21 Accuracy of train set



```
In [18]: accuracy(X_test, y_test)
```

```
Out[18]: 69.46564885496184
```

Figure 22 Accuracy of test set

4. CONCLUSION

Weather forecast has been playing a vital role directly or indirectly in each and every one of our life. Knowing what we will be facing in ahead of time is something big to achieve to the human race. Whether its about planning a vacation, trip or setting a routine and plans for the day knowing the weather can be very helpful for the individual.

AI has opened the door for us to explore and do the unimaginable things and discover and reach to the new heights of the possibilities in our life. With given sets of data collected from the past AI algorithms can mine these data and find patterns and predict the future. A weather forecast system will be developed similarly, a set of data collected can be used and fetched into algorithms to find some distinct pattern and gives out the prediction accordingly to that pattern. AI has many branches such as Natural Processing Language, Machine Learning, problem solving and searches. Machine Learning is the category that was decided to be implemented to solve this problem. Machine learning is a huge sector with supervised and unsupervised learning. The algorithm to solve the problem was cluster analysis which is a common method of unsupervised learning. The further work can be achieving more accurate result. Manipulation of the data being fetched to it can be done with more improved way and predict with higher accuracy. Agglomerative clustering was implemented to develop a model for the weather forecast. Prediction can be done with unsupervised learning too as the result was quite unexpected and with higher accuracy than of that expected. I think with more proper research and data filtering and proper datasets the predicted events could be more accurate.

5. REFERENCES

Alkawaz, M. H., 2014. *Flowchart of k-means clustering algorithm* | *Download Scientific Diagram*. [Online]

Available at: https://www.researchgate.net/figure/Flowchart-of-k-means-clustering-algorithm_fig2_271915066

[Accessed 4 February 2019].

Expert System, 2018. *What is Machine Learning? A definition - Expert System*. [Online]

Available at: <https://www.expertsystem.com/machine-learning-definition/>

[Accessed 25 January 2019].

IBM , 2018. *Machine Learning | IBM Analytics | IBM*. [Online]

Available at: <https://www.ibm.com/analytics/machine-learning>

[Accessed 25 January 2019].

Joshi, N., 2018. *Transforming weather forecast with AI | Artificial Intelligence |*. [Online]

Available at: <https://www.allerin.com/blog/transforming-weather-forecast-with-ai>

[Accessed 3 February 2019].

Kassambara, A., 2018. *Agglomerative Hierarchical Clustering - Datanovia*. [Online]

Available at: <https://www.datanovia.com/en/lessons/agglomerative-hierarchical-clustering/>

[Accessed 4 February 2019].

KMA, 2009. *Digital Weather Forecast > Weather Forecast*. [Online]

Available at: https://www.google.com.np/search?hl=en-NP&authuser=0&ei=3zpZXK7rCsyVwgPdpxzACA&q=february&oq=feb&gs_l=psy-ab.1.0.35i39j0i67j0i131i4j0i131j0i2.524.901676..902708...19.0..0.182.3007.0j20.....0..1..gws-wiz.....6..0i131i67.TbH3ZzFn67Q

[Accessed 27 February 2019].

KMA, 2009. *Weather Forecast > Weather Forecast*. [Online]

Available at: https://web.kma.go.kr/eng/biz/forecast_01.jsp

[Accessed 26 February 2019].

Kurama, V., 2018. *Unsupervised Learning with Python – Towards Data Science*. [Online]

Available at: <https://towardsdatascience.com/unsupervised-learning-with-python-173c51dc7f03>

[Accessed 4 February 2019].

ScienceDaily, 2018. *Artificial intelligence*. [Online]

Available at: https://www.sciencedaily.com/terms/artificial_intelligence.htm

[Accessed 25 January 2019].

Silva, M., 2017. *Flow chart of agglomerative hierarchical clustering*. | Download Scientific Diagram. [Online]

Available at: https://www.researchgate.net/figure/Flow-chart-of-agglomerative-hierarchical-clustering_fig1_313238175

[Accessed 8 February 2019].

The MathWorks Inc, 2019. *Unsupervised Learning - MATLAB & Simulink*. [Online]

Available at: <https://www.mathworks.com/discovery/unsupervised-learning.html>

[Accessed 4 February 2019].

The MathWorks, Inc., 2019. *What Is Machine Learning? | How It Works, Techniques & Applications - MATLAB & Simulink*. [Online]

Available at: <https://www.mathworks.com/discovery/machine-learning.html>

[Accessed 25 January 2019].

6. BIBLIOGRAPHY

Alkawaz, M. H., 2014. *Flowchart of k-means clustering algorithm* | *Download Scientific Diagram*. [Online]

Available at: https://www.researchgate.net/figure/Flowchart-of-k-means-clustering-algorithm_fig2_271915066

[Accessed 4 February 2019].

AMERICAN METEOROLOGICAL SOCIETY, 2017. *bamsD160123 1..19*. [Online]

Available at: <https://journals.ametsoc.org/doi/pdf/10.1175/BAMS-D-16-0123.1>

[Accessed 2 February 2019].

Expert System, 2018. *What is Machine Learning? A definition - Expert System*. [Online]

Available at: <https://www.expertsystem.com/machine-learning-definition/>

[Accessed 25 January 2019].

IBM , 2018. *Machine Learning | IBM Analytics | IBM*. [Online]

Available at: <https://www.ibm.com/analytics/machine-learning>

[Accessed 25 January 2019].

Joshi, N., 2018. *Transforming weather forecast with AI | Artificial Intelligence |*. [Online]

Available at: <https://www.allerin.com/blog/transforming-weather-forecast-with-ai>

[Accessed 3 February 2019].

Kaggle Inc, 2019. *Austin weather forecast with unsupervised learning | Kaggle*. [Online]

Available at: <https://www.kaggle.com/sikora507/austin-weather-forecast-with-unsupervised-learning>

[Accessed 3 January 2019].

Kassambara, A., 2018. *Agglomerative Hierarchical Clustering - Datanovia*. [Online]

Available at: <https://www.datanovia.com/en/lessons/agglomerative-hierarchical-clustering/>

[Accessed 4 February 2019].

KMA, 2009. *Digital Weather Forecast > Weather Forecast*. [Online]

Available at: <https://www.google.com.np/search?hl=en->

[NP&authuser=0&ei=3zpZXK7rCsyVwgPdpxzACA&q=february&oeq=feb&gs_l=psy-ab.1.0.35i39j0i67j0i131i4j0i131j0l2.524.901676..902708...19.0..0.182.3007.0j20.....0..1..gws-wiz.....6..0i131i67.TbH3ZzFn67Q](https://www.google.com/search?q=february&oeq=feb&gs_l=psy-ab.1.0.35i39j0i67j0i131i4j0i131j0l2.524.901676..902708...19.0..0.182.3007.0j20.....0..1..gws-wiz.....6..0i131i67.TbH3ZzFn67Q)

[Accessed 27 February 2019].

KMA, 2009. *Weather Forecast > Weather Forecast*. [Online]

Available at: https://web.kma.go.kr/eng/biz/forecast_01.jsp

[Accessed 26 February 2019].

Kurama, V., 2018. *Unsupervised Learning with Python – Towards Data Science*.

[Online]

Available at: <https://towardsdatascience.com/unsupervised-learning-with-python-173c51dc7f03>

[Accessed 4 February 2019].

Pedregosa, F. a. V. G. a. G. A. a. M. V., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(1), pp. 2825-2830.

R.S. Michalski, J. C. T. M., 2013. An Overview of Machine Learning. In: J. C. T. M. R.S. Michalski, ed. *Machine Learning: An Artificial Intelligence Approach*. New York: Springer-Verlag Berlin Heidelberg, pp. 3-16.

ScienceDaily, 2018. *Artificial intelligence*. [Online]

Available at: https://www.sciencedaily.com/terms/artificial_intelligence.htm

[Accessed 25 January 2019].

Silva, M., 2017. *Flow chart of agglomerative hierarchical clustering*. | Download Scientific Diagram. [Online]

Available at: https://www.researchgate.net/figure/Flow-chart-of-agglomerative-hierarchical-clustering_fig1_313238175

[Accessed 8 February 2019].

The MathWorks Inc, 2019. *Unsupervised Learning - MATLAB & Simulink*. [Online]

Available at: <https://www.mathworks.com/discovery/unsupervised-learning.html>

[Accessed 4 February 2019].

The MathWorks, Inc., 2019. *What Is Machine Learning? | How It Works, Techniques & Applications - MATLAB & Simulink*. [Online]

Available at: <https://www.mathworks.com/discovery/machine-learning.html>

[Accessed 25 January 2019].