Professor: Sylvain Jaume

Name: Spoorthi Kondagadapu

Course: Big Data and Business Analytics[Capstone]

Date: April 22, 2017

**Abstract:**

Apache Zeppelin:

Apache Zeppelin is a multipurpose web based notebook which brings the features like

Data exploration and Data visualization to Hadoop and Spark.

The components of the workflow are:

1.Data discovery

2.Data exploration

3. Data reporting

4. Data visualization

Earlier the dynamic prediction was used, but now dynamic and statistical methods are used for

Weather forecasting.

Smart cities are now being established with smart, resilient infrastructure.

Smart cities are now facing extreme climate change and weather events which

need to be monitored. Public transport is ideal for reducing the overall gas emissions.

Smart cities: Smart city refers to modern urban growth which makes the cities

Competitive and productive. Top leaders and decision makers play a vital role

In the development of smart cities with balanced weather conditions.

Smart cities strive to achieve stable economy, good environment and smart governance.

The population growth in the smart cities will be growing exponentiallyin both developed and developing countries. World wide climate changes are increasing drought, flooding and large storms

which should be tackled effectively by the smart cities

Smart cities are generating 70% of global co2 emissions. Cities are rising to regulate

the environmental conditions. Smart cities have to face economic, environmental and social

challenges. The population growth is increasing tremendously in the smart cities.

Smart cities are using digital technologies to enhance conservation of natural resources

and energy use. By the use of new sensors the study of the environment pollution levels have

been closely monitored.

The city must have required funds to become a smart city.

World wide climate changes are increasing drought, flooding and large storms

Which should be tackled effectively by the smart cities.

The third benchmark forecast is produced by the Earthsat in real time which is highly sophisticated

one. The third type of forecast is achieved by the numerical weather prediction model.

The weather derivatives market depends on weather modeling and forecasting for both the demand and supply sides. To eliminate the weather derivatives we have to see how much weather noise exists.

To forecast daily average temperature in four US cities, the nonstructural time series approach works out well.

**Contribution:**

The Different layers of the Atmosphere based on temperature are: Troposphere, Stratosphere, Mesosphere and Thermosphere.

The Long wave radiation (Green house gases) are primarily absorbed by water vapor and carbon dioxide.

The atmospheric window is the range of wavelengths not absorbed.

The temperature gradient develops as the surface warms. The energy transfer upward through convection is induced by the temperature gradients.

The energy is transferred downward at nights.

There are two processes in fluids that generate convection. They are: Free convection and Forced

Convection. The incoming and outgoing radiation generate the earths radiation equilibrium.

The equation of state describes how pressure, Temperature and density are related to one another.

The air density and pressure are directly proportional to constant temperature.

The increase in temperature causes an increase in pressure under constant density.

The movement of air is provided by the pressure gradients. The large areas are concentrated

With small gradients. The large pressure gradients across small areas are displayed by hurricanes

and tornadoes.

With the altitude, the pressure gradients always decreases. The mass is proportional to gravitational force. The pressure gradients lead to hydrostatic equilibrium for warm air.

The pressure gradients lead to hydrostatic equilibrium for cold air. When the atmosphere is saturated

,it generates saturation vapor pressure.

The vapor pressure on surfaces is generated by the movement of water vapor.

Saturation vapor pressure depends on temperature. The saturation vapor pressure increases rapidly

at higher temperatures.

The addition of water vapor, moist air causes the air to become saturated. The addition or removal of energy is done by the diabatic process. The adiabatic process is the process in which cloud formation

involves temperature changes. The decrease in air temperature with height is referred by the environmental lapse rate.

The formation of clouds and precipitation is the result of adiabatic cooling. The adiabatic cooling is the result of the upward movement of air. Collisions and coalescence occur when the collector drops fall

and overtake small droplets. The adiabatic cooling is a result of the upward movement of air. Frontal lifting occurs when two air masses converge. Clouds are formed when warm air advances towards

the cold air. The air is formed to rise when the air converges to the Earth’s surface. The air is caused to

lift, when differential heating occurs at the surface. This causes cloud development. When the air is lifted and moves in spite of saturation, the air becomes unstable.

Cloud drops occur from the condensation nuclei, when they have been attracted by water. Collisions

Occur when the collector drops overtake the small droplets. At temperature below zero degree, ice crystals are formed. At temperature above zero degree the subfreezing conditions occur.

The cool cloud precipitation depends upon the existence of ice and supercooled water. When the water

freezes into ice crystals, a phenomenon called rimming occurs. When the ice crystals bond with the

Surface water, it results in aggregation. Ridges are formed when the heights extends towards the poles.

The troughs are formed when the heights extend towards the equator.

Uniform temperature and humidity are the characteristics of the air masses. The different air masses

Have boundaries called Fronts. The aerosol forms are taken by the sulfur compounds. The acid fog and acid rain are formed by the sulphate aerosols. The respiratory problems are caused by the Ozone.

The environmental degradation is caused by the Ozone. The pollutants near the earth’s surface are trapped by the inversions. A quarter of the earth’s surface is constituted by the mountains

and plateau areas. Climate change is the critical challenge facing today. Climate control heating is used

In maintaining the room temperature. The settings are chosen so that comfort requirements are

met and the energy costs are minimized. A neural network is used to forecast the minimum ozone in the Atmosphere. In the real time the model called Neurozone is being used.

Forecasting pollution peaks and informing the authorities and public are the responsibilities of the

air quality agencies.

**Discussion:**

The forecasting tools have developed statistical models which correspond to

the level of pollutants. The electric power plays a role in the production of green house gases.

The hydro energy resources and scheduling will be affected by the changes in precipitation.

The catastrophic wild fires in some of the places are also caused by climate change.

Earlier the dynamic prediction was used, but now dynamic and statistical methods are used for

Weather forecasting.

The predicted variables and the predictors are the observed quantities of the Model.

Step wise regression is the technique used in selecting the predictors, used in the meteorology.

Continous as well as binary variables are used in estimating the probabilities in the equations.

We need a separate equation , if the minimum Root , mean square error of the wind is desired.

In the forecasting , the surface weather variables, the model output statistics technique is used.

The atmospheric boundary layer plays an important role in weather forecasting.

The air mass transformation is studied in the Weather forecasting.

The output of the air mass transformation is useful for air pollution and agricultural problems.

The air mass transformation model is used as a forecast tool. There are two types of forecast good forecast and bad forecast. Good forecast is one in which the conditions are similar to observed conditions and bad forecast is one in which the beneficial outcomes are not produced.

There are three types of goodness, which are consistency, quality , value.

Forecasters derive their conclusions from the knowledge base which contain the weather

Conditions.

The current statements about the weather forecast are called Judgments and about the

future are called forecasts.

The inherent element of uncertainty changes from event to event. The requisite forecast is one in

In which the users find the complete data to make good decisions.

The requisite forecast is synonymously called as forecast.

The forecaster’s base their judgment on the information contained in the knowledge base, however the complete information may not be required.

The forecaster’s judgment always corresponds to the requisite forecast.

The ultimate goal of the forecasting is to provide the best judgment to satisfy the users.

A forecast can contain less spatial specificity than the judgment. The forecaster’s judgment

Is supposed to predict whether there is a likelihood of the precipitation, if the value is 0.2, according to the knowledge base and the real value obtained is also 0.2, then we say that the consistency exists.

The forecasters control the type 1 goodness which is consistency. If the forecasts correspond to judgments, type 1 goodness is satisfied. This makes it different from the other two types of

goodness.

It is the totality of the information, which contributes to the quality of the forecast.

The other aspects of quality are referred to as discrimination.

The sharpness of the forecasts correspond to the marginal distribution.

The relative forecasting is known by comparing the two levels of accuracy. The forecasts are

assessed by the information provided to the user.

By achieving the three types of goodness, we serve the users interest as well as forecaster’s interest.

Ignoring the uncertainty element affects the quality and value. In the distributed-oriented approach,

It is possible to reduce the dimensionality and complexity.

No weather forecasting system can acquire hundred percent accuracy. Accuracy and skills are the two factors on which the weather forecasting depends upon.

The conditional mean dynamics reveal the time series modeling. In the weather derivates context,

The weather forecasting methods are useful.

If the prespecified weather events occur, then the weather derivates are a new type of security.

If the prespecified weather events occur, then the weather derivates are a new type of security.

Precipitation [Rainfall, Snow], maximum temperature and humidity are the different kinds of weather

related variables.

Quantity hedging is determined by the weather forecasting, but price hedging is not.

The weather related changes in terms of quantities are protected by the weather derivative products.

Weather is a location specific commodity unlike crude oil. In the weather derivatives market, demand and supply sides are affected by weather forecasting.

Energy companies, utilities and insurance companies are affected by the weather driven poor performance.

The very large weather changes can pose very little risk, if they are predicted well.

Energy companies, utilities and insurance companies are affected by the weather driven poor performance.

The very large weather changes can pose very little risk, if they are predicted well.

The physical models of atmospheric conditions produce the short horizon forecasts.

National climate Data Center collects data from all the measurement stations.

The unconditional temperature densities are shown by the seasonality in daily average temperature.

The very large weather changes can pose very little risk, if they are predicted well.

The physical models of atmospheric conditions produce the short horizon forecasts.

National climate Data Center collects data from all the measurement stations.

The unconditional temperature densities are shown by the seasonality in daily average temperature.

The two benefits of the Fourier series are smooth seasonal pattern and parsimony.

Models like time series trend, Seasonal and cyclical components are greatly studied by the meteorologists to achieve good forecasts.

The trend in the daily average temperature is seen in most of the cities. The large increase in the

trend is due to the air pollution.

Auto regressive forecasts are the type of forecasts which involve seasonal, trend, cycle model.

The no change forecast is the first bench mark forecast which is known as persistence forecast.

The two component seasonal and trend model is the second benchmark model which uses

dummy variables to study the daily seasonal effects.

The third benchmark forecast is produced by the Earth sat in real time which is highly sophisticated

one. The third type of forecast is achieved by the numerical weather prediction model.

The weather derivatives market depends on weather modeling and forecasting for both the demand and supply sides. To eliminate the weather derivatives we have to see how much weather noise exists.

To forecast daily average temperature in four US cities, the nonstructural time series approach works out well.

In the future weather forecasting, multivariate analysis and cross hedging will be used. Weather

Surprises translate into earning and then translate to stock return movements.

The artificial neural networks have advantages over the linear combination. The data about the different attributes are used to train and test different models. Accurate forecasts were produced by the artificial Neural networks.

Weather forecasting is a challenging task since weather is a data intensive and continuous process.

Artificial Neural networks help in solving many nonlinear problems of weather forecasting.

The process of Weather forecasting is done well by the artificial Neural Networks.

In The process of Neural networks, the different neural networks are trained and their predictions are collected. The weighted average of each network is the output of an ensemble.

The final decisions are based on the certainty of each classifier. The different results are produced

by the different networks. From the view point of bias, the neural networks bring about greater accuracy.

The final decision is obtained by the certainty of each classifier. Artificial Neural networks do the training and testing of data in its models.

To minimize the Mean square error, the initial weights are trained.

The lower error and a reasonable learning time are the characteristics of the optimal network.

The different data patterns result in the different values of temperature, humidity and wind speed.

The Neural networks are employed for hourly weather forecasting.

There is a challenge today to study the climate change and minimize the energy savings.

The room temperature is kept within the range to build climate control heating. The different

Actuators keep the room temperature in the comfortable range. The Actuators settings are adjusted to minimize the energy costs.

For simulation and control, a bilinear model is used. The actuators are in each system. They are employ

ed for the following subsystems such as heating, cooling, ventilation.

The non linear problems are satisfied by the sequential linear programming.

Performance bound is used for perfect weather and its predictions.

There is an urgent need for reliable warning systems as the severity of floods increases.

The part of the uncertainties is predicted by the EPS forecasts.

In an urban atmosphere, the real time forecasting is caused by the neural network.

The success index forecasting is 78% with the neural classifier. In the real world a model called Neurozone is being used. The training data is updated every year. The air quality agencies perform tasks like monitor pollutants and inform authorities and public. The first air quality model

Is a combination of chemistry, transport and dispersion.

The second model connects the level of pollutants directly to the meteorological conditions.

For the real time forecasting, the second approach is preferred. The daily ozone concentrations are affected by the daily emissions. The variability of the Ozone levels is best explained by the daily weather variations. The photochemical smog is produced in the summer where high temperature

and low humidity prevail. The air quality agencies use the available variables in the statistical regression function to predict the weather.

The regression models and Ozone modeling are studied in depth.

The autocorrelation of the Ozone variability, the seasonality and trend are modeled by the time series

Analysis.

The cart method is associated with the regression analysis. The ozone prediction is tested by the

non linear techniques. The neural network with a neural classifier and the Ozone modeling are

used to improve the prediction.

In the summer the ozone reaches peak values. The weather forecasts deliver the meteorological data.

**Data and Method:**

The neural model is based on the multilayer perceptron network.

Computers are used to find the best parameters to avoid the time consuming process.

The pruning technique is used to avoid overtraining, which has poor performance on fresh data.

Overtraining occurs when the noisy details of the training data are learnt. To compute the

Confidence interval of the predicted values in the training data, a theory was proposed.

The particular observation on the fitted regression is a measure of leverage.

To optimize the parameters of the regression, the training data set is used. The performance of the regression is tested by the testing data set. The difference between the two sets is not there

at the 95% confidence level from the Anova test.

By using the neural model, there is significant improvement of the scores. The data by the National weather service is used in the meteorological forecasts. By using the neural model, there is significant improvement of the scores. The data by the National weather service was used in the meteorological

Forecasts. At the end of the year, validation data is used as training data and retrained to obtain new regression parameters.

In order to simulate, better non-linear problems in numerical weather predictions,

The non-linear functions are used. The wind speed and air temperature are the two essential parameters by which the numerical weather predictions are done.

The small flow characteristics can be characterized by increasing the model resolution.

The higher resolution increases the forecast skill.

The seasonal changes and the local effects are obtained from the Model output statistics.

The parameters air temperature and the wind speed are used in the study of the different order polynomials.

The state of the weather is uncertain. Daily we face challenges

regarding how to equip ourselves appropriately. The greatest factor

which is used in studying the weather conditions is the solar radiation.

The weather forecasters use supercomputers. Satellite data has contributed

to the weather predictions. If the supercomputers improve in resolution,

the weather predictions will become accurate. Warm temperatures are easy to predict.

All the weather predictions are based on the data from the US National weather service.

Wet bias is when the weather forecasters predict more rain than the actual rain.

Weather predictions are based on sound decisions to help people in their day to day activities.

Super computers are used to study the mysteries of weather and climate.

Weather forecasting has led to safe transportation, productivity and resilience.

There will be more benefits to the individuals if the weather forecasting techniques are improved.

The warnings of weather conditions have increased from 13 percent to 69 percent over time.

Researches have developed a Radar which predicts the weather more accurately.

Accurate forecasts helps in optimizing the staff in the various areas of business and sectors of the industry.

Weather forecasting may undermine people’s ability in the day to day activities.

Predicting the rain is harder. Global warming affects every nation.

Businesses like Walmart, JPMorgan Chase have developed their own statistical models to predict extreme weather conditions.

Weather forecasting is a complex process which needs reliable data to make accurate

predictions. Weather forecasting developed meteorology as Science, where meteorologists are

extensively used in making the weather predictions.

Weather data is collected to study weather forecasting. Weather prediction makes use of numerical data such as temperature, atmospheric pressure and humidity.

The different forecast range types are short range forecast, extended forecast, Medium

forecast, Long range forecast.

Weather forecasts need more improvement even though the modern technology

is being used. The earth’s atmosphere is a complex process where rain or snow cannot always

be predicted.

Temperature warnings are used to protect life and property. The data about the state

of the atmosphere is calculated. The back propagation neural networks which comes

under Neural networks is used to calculate the temperature which supports various algorithms.

A neural network is the efficient data model used to study the input/output relationships.

A neural network represents the human brain in functioning. The neural network collects the

Input data called the testing data and processes it to make classification and prediction.

The several steps in prediction are data collection, data analysis, prediction, output processing

The support vector machines are also used for weather prediction.

The time series data is used to calculate the maximum temperatures of two to three days. The parameters which need to be analyzed are: Atmospheric pressure, Temperature, Humidity, wind velocity and wind direction.

We start with the initial conditions and put them in the statistical models and we get the predictions.

Bayesian probabilistic arguments are used for numerical weather prediction. The supervised learning

techniques are applied to the prediction problems. The weather prediction is achieved by the projection of uncertainties. The green house effect was discovered by Joseph Fourier in 1824.

The increased levels of greenhouse gases warm the earth. The number of high temperature events has been increasing since 1950. Global warming is permanently changing the Earth’s climate.

Forty two percent of the scientists think that the Global warming is happening.

The Global warming results in rising sea levels due to the melting of the polar ice caps.

The amount of the water in the atmosphere affects the temperature. The factors affecting the weather conditions are temperature, pressure, humidity, precipitation and wind.

Meteorology is the study of the weather. Certain weather features indicate the respective

weather conditions. Over the years the study of the weather patterns has influenced the weather

predictions. Weather forecasting is done by the study of the science and technology to give

weather predictions. The meteorological information is produced by the models, which are

computer programs to give the future predictions.

The techniques of weather forecasting are persistence, use of a barometer, looking at the sky, now casting, use of forecast models and Analog technique.

Persistence is the study of today’s weather conditions to predict tomorrow’s weather.

Nowcasting is the forecasting of the weather in the next six hours.

Accurate weather forecasting are required since the Aviation industry is sensitive to the weather.

The agricultural industry depends on the weather forecasting to decide what to do on a particular

day. The forestry department uses weather forecasting to prevent and to control wildfires.

The navy has a special team of forecasters and weather observers for the study of the

weather conditions.

Competitor article:

The author uses the artificial intelligence techniques for forecasting wind power. The meteorological forecasts are used for improving predictions. The method of applying

Forecasting method on power production shows the validity of the method.

The priority of the grid operator is to predict changes in the wind power. Researches focus on

A forecasting tool to predict wind power with good accuracy. The model output statistics is used to reduce the remaining error. The physical models may use the numerical weather predictions. Researchers have developed models for short term prediction based on fuzzy neural networks.

The main contribution of the numerical weather prediction is to make them available based on fuzzy logic rules. The time series of wind power is used as input and the proposed model is based on neural networks. The forecasting models contain information about the wind speed, Wind direction and temperature.

The numerical weather predictions provide information in the long term horizons.

The spatial resolution of the meteorological model is the factor on which accuracy is dependent on. Models with high resolution require more time, but perform better.

The numerical weather predictions contribute to the accuracy having effect on the long time and short time horizons. In order to optimize the use of Numerical weather predictions, artificial neural networks combined with fuzzy logic are used. Inaccurate numerical weather predictions make the wind power prediction difficult.

Weather forecasts are uncertain because they depend on the spatial and temporal resolution

of the Numerical weather predictions. The contribution of the proposed method is that it uses

artificial neural network with a fuzzy logic model. The author outperforms my algorithms in this aspect.

**Reference for Competitor article**:

Sideratos, G., & Hatziargyriou, N. D. (2007). An advanced statistical method for wind power forecasting. *IEEE Transactions on power systems*, *22*(1), 258-265.

My algorithm:

Weather forecasting is a complex process which needs reliable data to make accurate

predictions. Weather forecasting developed meteorology as Science, where meteorologists are

extensively used in making the weather predictions.

weather data is collected to study weather forecasting. Weather prediction makes use of numerical data such as temperature, atmospheric pressure and humidity.

The different forecast range types are short range forecast, extended forecast, Medium

forecast, Long range forecast.

weather forecasts need more improvement even though the modern technology

is being used. The earth’s atmosphere is a complex process where rain or snow cannot always

be predicted.

Temperature warnings are used to protect life and property. The data about the state

of the atmosphere is calculated. The back propagation neural networks which comes

under Neural networks is used to calculate the temperature which supports various algorithms.

A Neural network is the efficient data model used to study the input/output relationships.

A Neural network represents the human brain in functioning. The neural network collects the

Input data called the testing data and processes it to make classification and prediction.

The several steps in prediction are data collection, data analysis, prediction, output processing.

The support vector machines are also used for weather prediction. The time series data is used

to calculate the maximum temperatures of two to three days. The parameters which need to

be analyzed are: Atmospheric pressure, Temperature, Humidity, wind velocity and wind direction.

We start with the initial conditions and put them in the statistical models and we get the predictions.

Bayesian probabilistic arguments are used for numerical weather prediction. The supervised learningtechniques are applied to the prediction problems. The weather prediction is achieved by the projection of uncertainities. The green house effect was discovered by Joseph Fourier in 1824.

The increased levels of greenhouse gases warms the earth. The number of high temperature events has been increasing since 1950. Global warming is permanently changing the Earth’s climate.

Fourty two percent of the scientists think that the Global warming is happening.

The Global warming results in rising sea levels due to the melting of the polar ice caps.

The amount of the water in the atmosphere affects the temperature. The factors affecting the weather conditions are temperature, pressure, humidity, precipitation and wind.

**Conclusion:**

The factors that influence weather are solar distance, Latitudinal location, Air pressure, water presence.

The latitudinal location also affects the weather. The presence of water plays an important

role in weather. The places near the oceans are cooler than the areas away from them.

Climate is the description of the weather over a period of years. Solar energy is received

by the earth by the process known as radiation. Weather is categorized into equatorial, tropical, polar

and arctic. Air pressure is the gravity on air mass. It is measured in millibars. Areas closest to the equator receive most of the sunlight. The heat energy is trapped by the element carbon and carbon dioxide. The collection of ice crystals and water droplets form the clouds.

Higher latitudes get lower solar radiation. The moisture content of air is called humidity. The ocean side facing areas receive more rain. The minimum temperatures occur during the early morning hours. Deforestation, urbanization, green house gases effect the weather conditions.

Weather refers to the atmospheric conditions over a short period of time and climate refers to the conditions over a longer duration of time. The important elements of the weather are temperature

, relative humidity, precipitation, pressure and winds.

The temperature is affected by the solar radiation. The factors that affect solar radiation are latitude, altitude and cloud cover.

The temperature at poles and equator are different because sun’s rays travel the earth’s surface

at different angles. The poles experience low temperatures than the regions closer to the equator.

The sun’s rays are less intense at the poles, so they have lower temperatures.

The high altitudes have lower temperatures, since they have smaller concentration of

gases to trap heat. The amount of water vapor in air is called humidity.

Humidity indicates the happening of precipitation, dew or fog.

The maximum amount of water vapor indicates saturation. Rain is most likely to occur

in the high humidity areas. Relative humidity and temperature are inversely proportional.

The water falling from the atmosphere is called precipitation. The movement of air from high

Pressure region to a low pressure region is called wind.

Weather, Climate affects the tourist industry. The seasonal movement of large air masses

are called Monsoon winds. The cold winds of Asia blow towards the areas of Australia.

The heavy rain to Southwestern India and Bangladesh is bought by the moisture carried by the wind.

The winds bring rain to India and Bangladesh blowing across the Asian continent. The condition

Of the atmosphere is called weather and its study is called meteorology. The experts in meteorology

are called meteorologists. The different layers of the atmosphere are Thermosphere, Mesosphere,

Stratosphere and Troposphere. Wind velocity describes the wind speed and the wind direction.

Wind speed is measured by the device called Anemometer. North wind is the wind coming from the north. Rain gauge is the instrument which measures the amount of rainfall. Barometer is used

to measure Air Pressure. We have cool summers and mild winters in Britain. Both the temperature of the ground and the temperature of the air are recorded by the weather stations. Beaufort scale is used to measure wind strength. In summers the southern areas get warmer as we go further down.

The weather systems called depressions bring about the rain in Britain. The areas where

Cold air and warm air meet are called depressions. The air masses move in two types of fronts. Computers are used for the modern forecasts based on the careful observations from the

weather stations.

Satellite images and radar give better weather predictions.

The six major controls of the climate are latitude, altitude, pressure and wind system, distance from the sea and ocean currents. High mountains pose as barriers for cold or hot wind. When the rainfall

Continues for several days, it is known as the burst of monsoon. The vertical movement of air

Is called air current and the horizontal movement is known as winds.

Weather forecasts are being made more accurate to protect life and property.   
Technological apps are making weather predictions accessible to people immediately.

Super computing is used extensively to study the mysteries of weather forecasting and climate.

During the hurricanes like Katrina and sandy, five day forecasts were given to make

People comfortable to face the hurricanes.