

Weather Prediction

Using IBM SPSS Modeler

Team Member:

Sidhanta Sahoo

Dheeraj Mahapatra

Shyam Kumar Soni

Aman Kumar Keshri

**Purpose:**

The purpose of this project is to predict weather reports for the public in a simple and accurate way. For this, we collect data from different resources like temperature, humidity, rainfall, and other weather conditions.

Data file: weather\_database.csv

Data folder: D:\SPSS\Project

**Task 1. Start IBM SPSS Modeler and set the working folder.**

1. From the Start menu, click All Programs, click IBM SPSS Modeler 18.0, and then click IBM SPSS Modeler 18.0.

2. When a splash screen appears, click Cancel.

3. From the File menu, click Set Directory.

4. Beside Look in, navigate to D\SPSS\Project, and then click Set.

**Task 2: Import the dataset using the Var. file node.**

In this task you will import the dataset stored in a csv file, weather\_database.csv. The csv file has field names in a first row

1. From a Source palette, place a Var. File node on the stream canvas.
2. Edit the Var. File node.
3. In the Data tab, next to File type, ensure that csv (\*.csv) is selected.
4. Set Import file to D:\SPSS\Project\ weather\_database.csv
5. Ensure that the First row has column names option is enabled, and then click Preview.
6. Click OK, to close the Preview output window.
7. Click OK, to close the Excel dialog box.

**Task 3: Making a Super Node.**

1. Place a Type Node connect with the Source Node and set the measurement level and click on Read Values.
2. Connect a Select Node to the Type Node and give the conditions as shown:

Visibility\_km < 100 and

Temp\_Avg\_C > -50 and Temp\_Avg\_C < 60 and

'Humidity\_%' >= 0 and 'Humidity\_%' <= 100 and

Pressure\_hPa >= 900 and Pressure\_hPa <= 1100 and

WindSpeed\_kmh < 150

Then click on the Include Mode and then OK.

1. Connect the Node to ‘to stream’ to get out of the Super Node.

**Task 4: Making Partition.**

1. Place a Partition Node and connect with a Super Node and make partition (70% training and 30% testing) and the click on OK.

**Task 5: Making Model for Air Quality Index.**

1. Place a Random Trees Node and connect it with the Partition Node and set the **Target field** PM10\_ug\_m3 and the **Prediction Fields**:

* Temp\_Avg\_C
* Humidity\_%
* Pressure\_hPa
* WindSpeed\_kmh
* PM2.5\_ug\_m3
* NO2\_ppb
* SO2\_ppb
* O3\_ppb.

Then Run this model and then you will see a Nugget will be formed .

1. Attach a Table Node to this Nugget to see the result and attach a Analysis Node to check the accuracy of the Model.