Predicting air quality in Delhi due to pollution from bordering states

1MS17IS105 Sarthak Jain
1MS17IS124 Tilak Singh
1MS17IS133 Vinay Biradar
1MS18IS416 Umer Faruk

Guided by- S R Mani Sekhar
(Asst. Professor)

Abstract

- **Air Pollution** in delhi is one of the biggest problems. This problem goes to its peak the the month of October and November due to pollution in surrounding areas, north west to Delhi. Air pollution consist of various harmful gases like NO2, SO2, CO, PM2.5 etc. In this project we are focusing on the PM2.5. We want to predict the value of PM2.5 in delhi based on Haryana and Punjab air pollution.
- **PM2. 5** refers to atmospheric particulate matter (PM) that have a diameter of less than 2.5 micrometers, and it is one of the major gases which majorly effacts human health.
- This will help us to take necessary steps to avoid the repercussions of Air pollution on the health of delhi.

Introduction

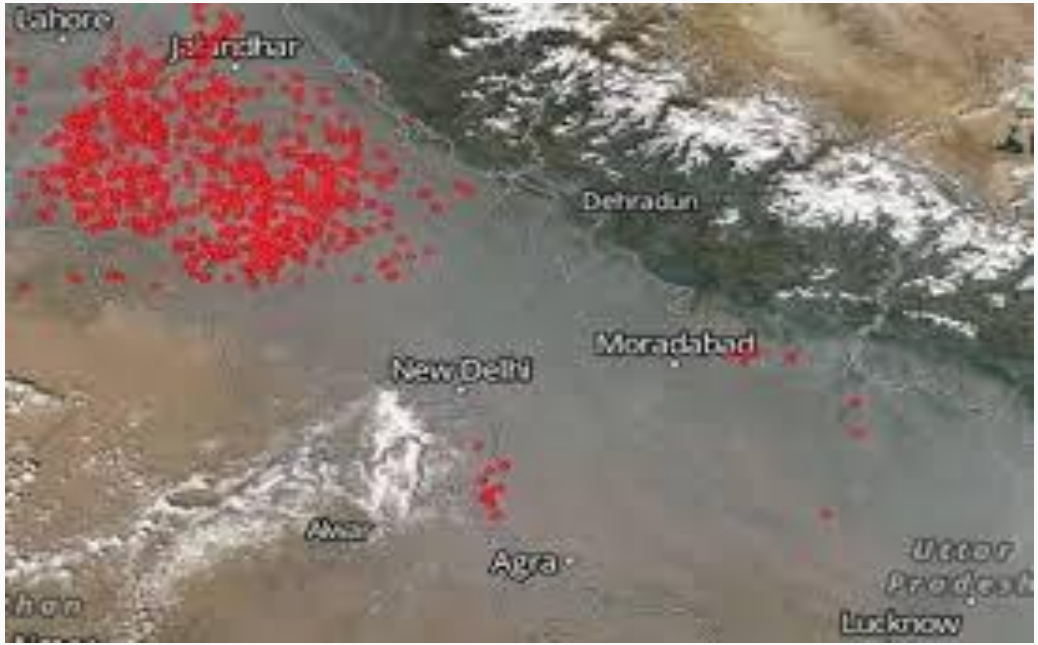
- **History.**

- **Stubble burning** is intentionally setting fire to the straw stubble that remains after grains, like paddy, wheat, etc., have been harvested. The practice was widespread until the 1990s, when governments increasingly restricted its use.

- Why **Stubble burning** ?

- Thick **Smog** in Delhi, How?

Geographical Aspects



Effects Of Stubble Burning

- Helpful Effects
  - Kills slugs and other pests.
  - Can reduce Nitrogen

- Harmful Effects
  - Loss of nutrients.
  - Pollution from smoke.
  - Damage to electrical and electronic equipment from floating threads of conducting waste.
  - Risk of fire spreading out of control.

Literature Review

Earlier Researches in this fields:

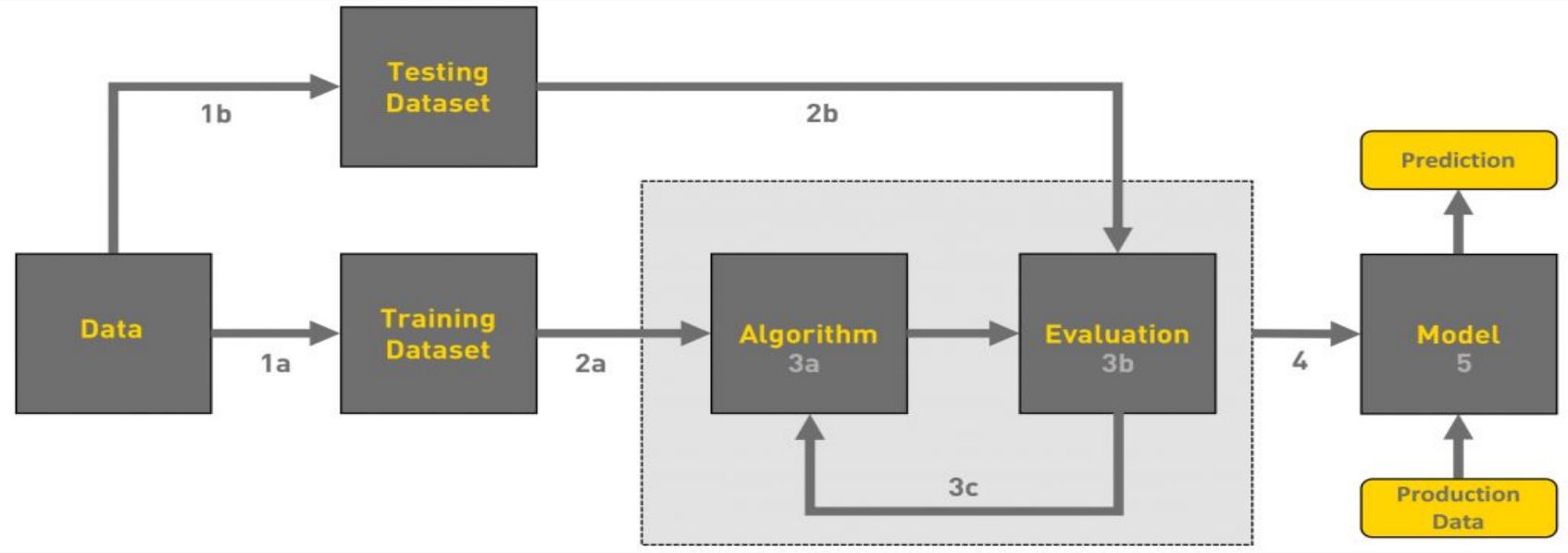
- Assessment of contribution of agricultural residue burning on air quality of Delhi using remote sensing and modelling tools. **Author** – Moorthy Nair, Hemant Bherwani, Suman Kumar, Sunil Gulia, Sanjeev Goyal, Rakesh Kumar.
    - Authors used remote sensing for the observation to find the particulates transmission over Delhi during burning period.
    - In their observation there was no transmission, but when the study was led by AERONET it showcased that some carbon aerosol subtype were identified over Delhi
    - In the same research HYSLIT model was proposed in which it showed that northwesterly winds intersect the agricultural residue burning region which might transport stubble burning particles to Delhi

- Ambient air quality changes after stubble burning in rice-wheat system in an agricultural state of India.[2009] **Author** – Dipti Grover, Smita Chaudhary.
  - They undertook three regions of Haryana and analysed their PM2.5 , SOx, and NOx values and its effect on Delhi.
  - During the research it was founded that the concentration of the three parameters taken into consideration exceeded the NAAQS values by 78%, 71% and 53% respectively. There was a striking increase in PM2.5 value by 3.5 times.
  - It was concluded that it is clearly evident that there is a significant increase in levels of PM2.5 , SOx, and NOx during crop residue burning period as compared to non-burning period.

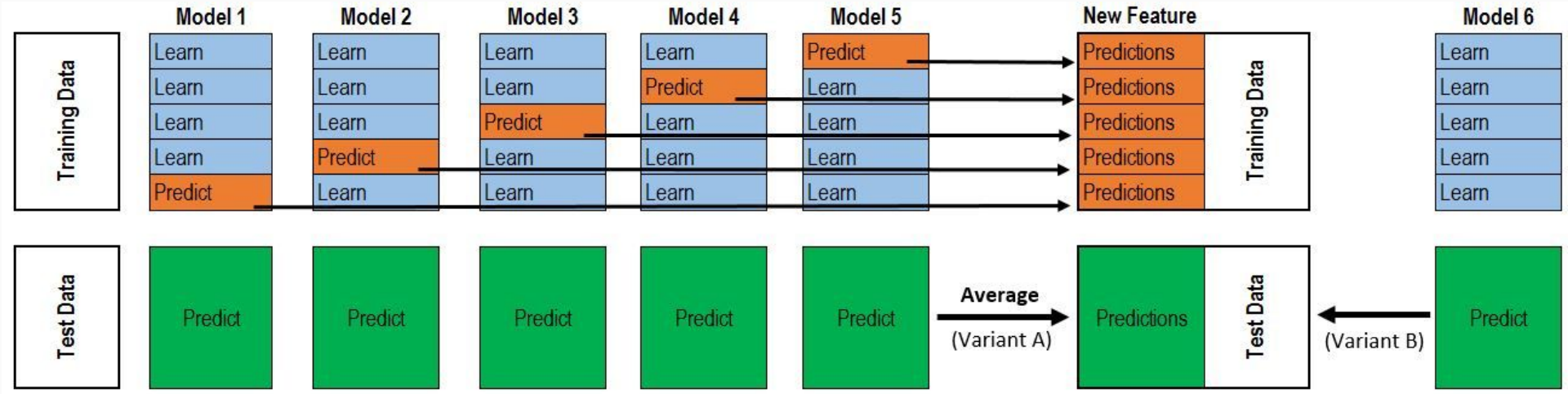
- Ambient air Quality during wheat and rice crop stubble burning episodes in Patiala.[2016] **Author** – Susheel K Mittal, Nirankar Singh, Ravinder Agarwal, Amit Awasthi, Prabhat K. Gupta.
  - The researchers studied the stubble burning practices in five different regions of Patiala to find the level of change in aerosol, SO2 , NO2 .
  - Aerosols which were collected on GMF/A and QMF/A sheets for 24 hrs along with separate collection of SO2 and NO2 throughout the year and the results obtained during stubble burning periods were compared with non-stubble burning period.
  - Results clearly pointed out that there were an increase in the levels of concentration of aerosol, SO2 ,NO2 during stubble burning period.

System Analysis & Design

Workflow



Methods Used



Stacking regression adding a meta model

The procedure, for the training part, may be described as follows:

1. Split the total training set into two disjoint sets (here **train** and .**holdout**)
2. Train several base models on the first part (**train**)
3. Test these base models on the second part (**holdout**)
4. Use the predictions from 3) (called out-of-folds predictions) as the inputs, and the correct responses (target variable) as the outputs to train a higher level learner called **meta-model**.

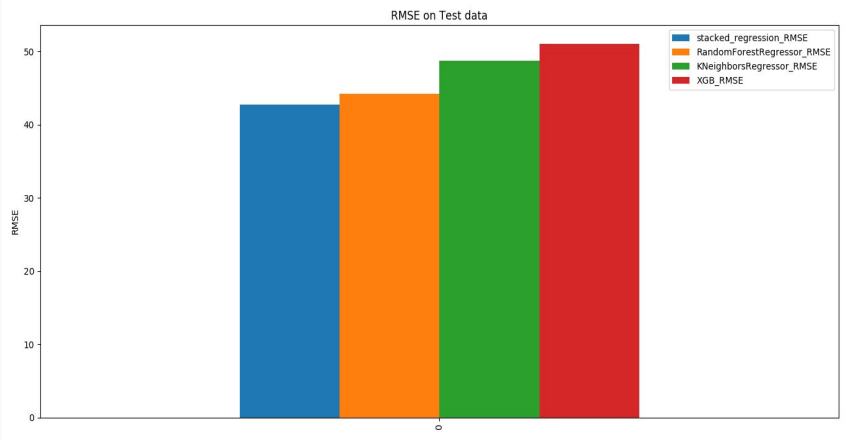
Testing & Results

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

RMSE on testing data

1. Stacked Regression
2. RF
3. KNN
4. XGB

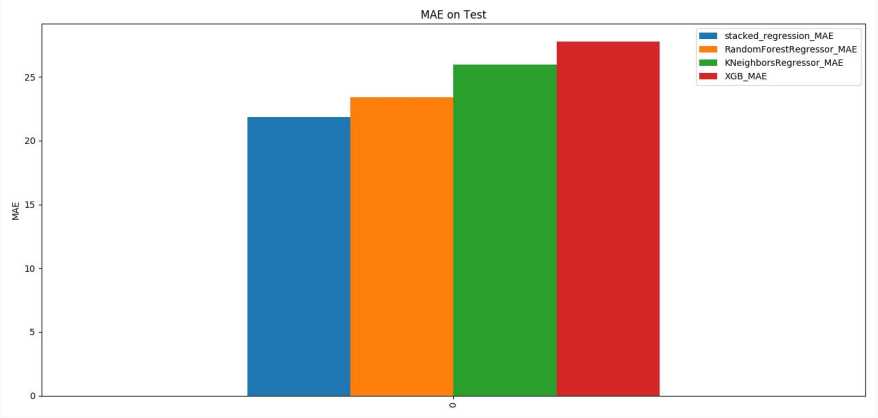


Mean absolute error (MAE)

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j|$$

MAE on testing data

1. Stacked Regression
2. RF
3. KNN
4. XGB

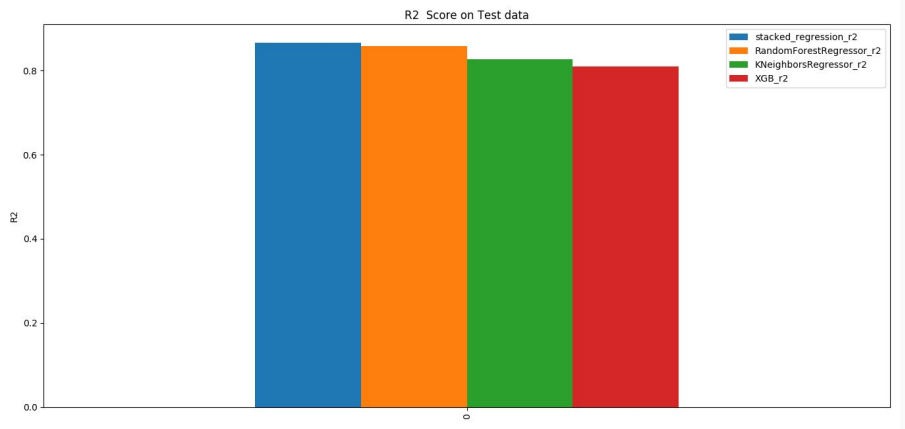


R-squared (R 2)

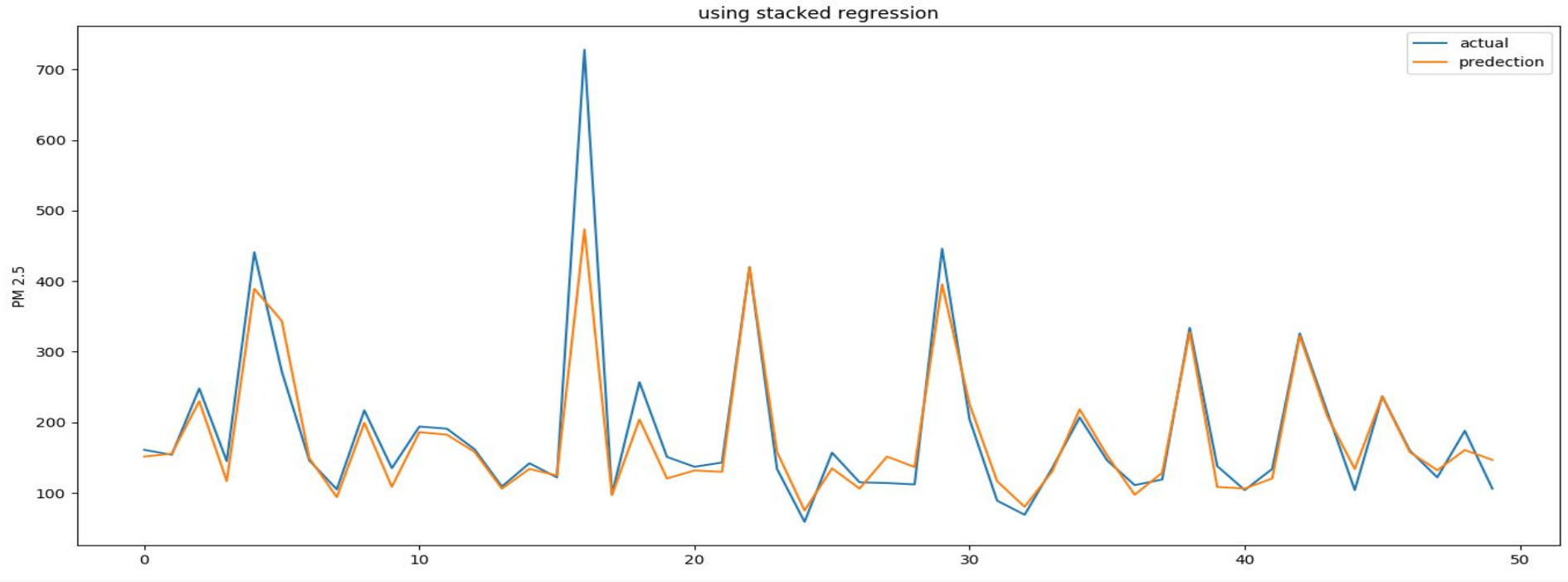
$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

MAE on testing data

1. Stacked Regression
2. RF
3. KNN
4. XGB



Stack Regression Actual vs Predicted



Conclusion

- We have used ensemble methodology called stacking in which improvements are done through boosting.
- Models give best results based on specific parameters, so most of them have limitations in different scenarios.
- We chose models based on their performances, in later work we can consider characteristics and the principles of the model instead of the
- weak accuracy gap between models

Future Works

- Other cities affecting the different states.
- In future other parameters can be taken into consideration like, NO_2, SO_2, SO.
- Limited dataset in current project.
- Other latest models can be used for obtaining higher accuracy
-

THANK YOU