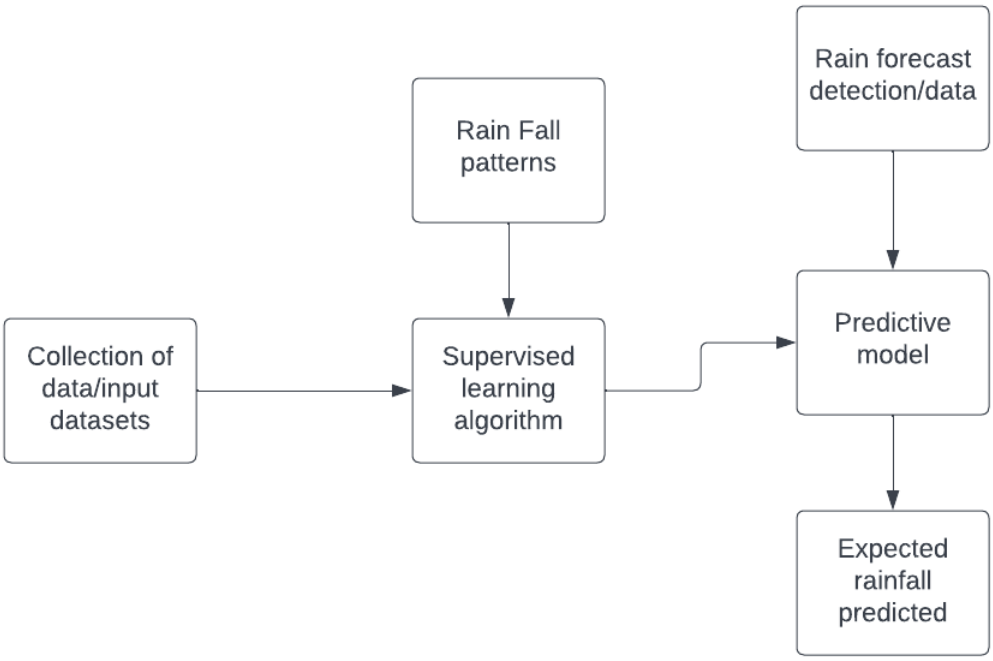
Proposed Solution:

A detailed survey on rainfall predictions using Machine learning algorithm over twenty-five years is done. From the survey it has been found that most of the researchers used different models for rainfall prediction, We will be using classification algorithms such as Decision tree, Random forest, KNN, and xgboost. We will train and test the data with these algorithms. From this best model is selected and saved in pkl format. Once the model is saved, we integrate it with flask application and also deploy the model in IBM.

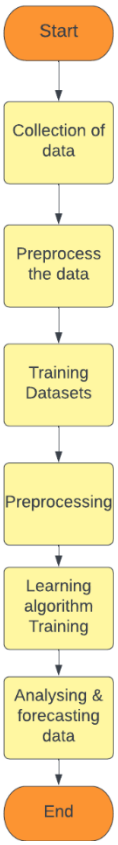
Social Impact:

- High prediction accuracy.
- Hold perfectly good for large scale datasets with large number of variables.
- Integral variable selection based on importance and variable interaction.
- Deals efficiently with data having missing values.
- Computation of relation between variables and classification.
- Proximity calculation between cases.
- Can be used for unsupervised learning and outlier detection.
- Internal unbiased estimation of the generalization error.

Business Model:



Architecture Diagram



Flow Diagram

Scalability of Solution:

Predicted values of rainfall from Machine learning algorithm are available. The mean squared errors of the models are also available. The model with the least mean square error is the best accurate model in predicting the rainfall. Other data analysis graphs for the effects of the rainfall are available for observation. From the observed result we see that classification algorithm has the least mean squared error value hence classification algorithm is predicting the rainfall value accurately.