

Project Design Phase – I

Proposed Solution Template

S.No	Parameter	Description
1	Problem Statement(Problem to be solved)	This comparative study is conducted concentrating on the following aspects: modeling inputs, Visualizing the data, modeling methods, and pre-processing techniques. The results provide a comparison of various evaluation metrics of these machine learning techniques and their reliability to predict rainfall by analyzing the weather data. 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.
2	Idea / Solution description	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.
3	Novelty/Uniqueness	It seems that there are no application or web application which is common for all over

		India for predicting the rain fall data for Indian agriculture , so the proposed application will contain the rain fall data of all the Indian agriculture.
4	Social Impact/Customer Satisfaction	<p>It Deals efficiently with data having missing values.</p> <p>Computation of relation between variables and classification.</p>
5	Business Model(Revenue Mode)	<p>Rainfall forecasting has been around for years using traditional methods that employ statistical techniques to assess the correlation between rainfall, geographic coordinates (such as latitude and longitude), and other atmospheric factors (like pressure, temperature, wind speed, and humidity). However, the complexity of rainfall such as its non-linearity makes it difficult to predict. Consequently, attempts have been made to reduce this non-linearity by using Singular Spectrum Analysis, Empirical Mode Decomposition, Wavelet analysis.</p>
6	Scalability of the Solution	<p>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.</p>