**SYSTEM ANALYSIS**

**EXISTING SYSTEM:**

* Hegde, M. et al. (2015, June) used multiple linear regression and neuro fuzzy systems for predicting crop yield by taking biomass, soil water, radiation and rainfall as input parameters for the research and their majorly concentrated crop was wheat.
* Sujatha, R., & Isakki, P. (2016, January) used classification techniques like ANN, j48, Naïve Bayes, Random Forest and Support vector Machines. They have also included both climatic parameters and soil parameters as features in their modelling.
* Ramalatha, M. et al. (2018, October) used a hybrid approach of combining Kmeans clustering and classification based on modified K-NN approach. The data was collected from Tamil Nadu, India where the majorly concentrated crops were rice, maize, Ragi, Sugarcane, and Tapioca.

**DISADVANTAGES OF EXISTING SYSTEM:**

* Existing approaches are low accuracy and high computation time and these might be due the use of irrelevant features in dataset. In order to tackle these problems new methods are needed to predict the crop yield correctly. The improvement in prediction accuracy is a big challenge and research gap.
* Accuracy is very low
* Computationally complex
* More execution time required to generate results

**PROPOSED SYSTEM:**

* In India, there are more than a hundred crops planted around the whole country. These crops are categorized for better understanding and visualization. The data for this research has been acquired from the Indian Government Repository.
* The data consists of attributes – State, District, Crop, Season, Year, Area and Production with around 2.5 Lakh observations. The system depicts the states and territories of India which visualize that which category of crops are famous in which season. We used advanced regression techniques – Lasso, ENet and Kernel Ridge and further we used stacking of these models to minimize the error and to obtain better predictions.

**ADVANTAGES OF PROPOSED SYSTEM:**

* The performance of the proposed method is high in terms of prediction accuracy.
* Low computation time.