RESEARCH PROPOSAL

1. Proposed Title:

"Big Data based Optimization and Simulation using Machine Learning to improve crop yield in Madurai District"

2. Staff Details

Mrs.D.Sakthipriya MCA.,M.Phil.,
Department of Computer Science
Mangayarkarasi College of arts & science for Women
Madurai.

3. Problem definition

Improving crop yield is a critical and necessary component of achieving food security and protecting natural resources and environmental quality for future generations. Although significant progress has been made in agricultural science developing seed varieties with genetic traits desirable in different planting environments and in advancing farming technologies and practices, tremendous opportunities exist in exploring how individual farmers, with the limited available resource, can make the best use of what agricultural science offers.

As a small part of the research effort in that direction, it proposes "**an analytics framework for seed variety selection decisions"** – one of the most important decisions a farmer has to make that has significant implications for the yield of the farm. Agribusinesses offer many seed varieties to farmers based on the yield performance of the seed varieties observed over several years at various farm locations at Madurai district. Farmers face the decision of selecting a small set of seed varieties offered and allocating farmland to the selected seed varieties. An informed decision requires **accurate predictions of yield performances of seed varieties on the targeted farmland** and balancing tradeoffs between expected yield and risk associated with the varieties selected.

4. Focus Objectives

i) To develop a methodological framework to help farmers in Madurai District make paddy seed selection decisions for betterment of the Crop Yield.

ii) To apply machine-learning models as the predictive models for the paddy yield with the data set from Syngenta and comparing with the local data set available at agricultural departments to the farmers in Madurai District.

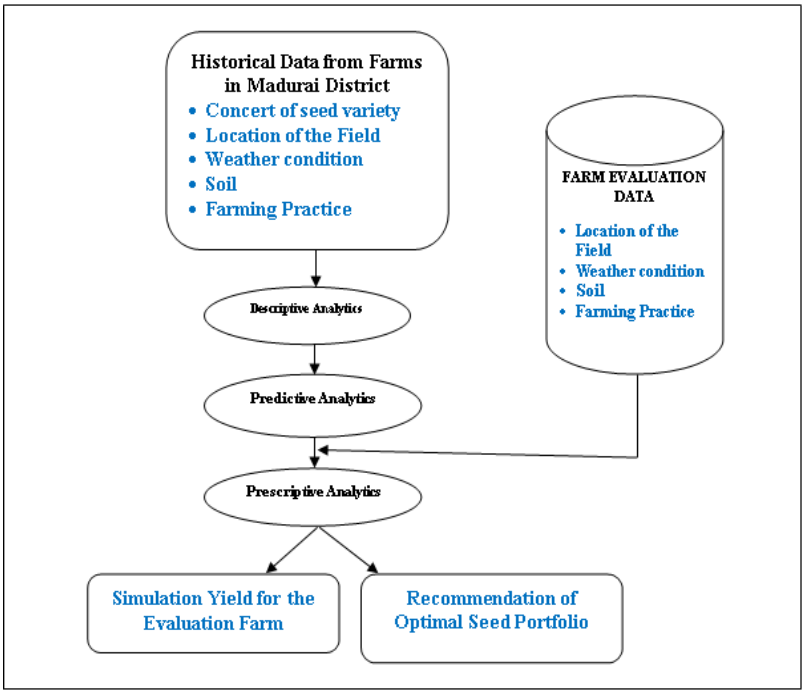
iii) To collect the dataset of Paddy yield and other related data in Madurai district from 2010 and 2019.

iv) Evaluation of machine-learning models to update simulation and optimization of paddy yield under different weather conditions.

v) Formulation of a simulation-based optimization problem to determine the optimal paddy-mix to minimize the risk associated with the yield.

5) Methodology

The methodology framework developed in this research can be applied to seed selection decisions of paddy crops and influence the farming practice positively.



6) National and international Status

Literature source	Contributions
Xu, J., Cai, H., Wang, X., Ma, C., Lu, Y., Ding, Y., & Saddique, Q. (2020). Exploring optimal irrigation and nitrogen fertilization in a winter wheat- summer maize rotation system for improving crop yield and reducing water and nitrogen leaching. Agricultural Water Management, 228, 105904.	**The author found optimal irrigation could be adjusted in variable precipitation years in Guanzhong Plain. Maize seedling and wheat jointing stages were the sensitive phases to water deficit.**
Shirley, R., Pope, E., Bartlett, M., Oliver, S., Quadrianto, N., Hurley, P., & acon, J. (2020). An empirical, Bayesian approach to modelling crop yield: Maize in USA. Environmental Research Communications.	**The author applied an empirical, data-driven approach for describing crop yield as a function of monthly temperature and precipitation by employing generative probabilistic models with parameters determined through Bayesian inference.**
Cui, J., Shao, G., Lu, J., Keabetswe, L., & Hoogenboom, G. (2020). Yield, quality and drought sensitivity of tomato to water deficit during different growth stages. Scientia Agricola, 77(2).	**The author investigated the drought sensitivity of tomato (Lycopersicon esculentum Mill.) yield and quality during different growth stages, field and pot experiments were conducted in a high tunnel in southern China.**
Mendez, J. M., & Dasig, D. D. (2020). Frost Prediction in Highland crops Management Using IoT-Enabled System and Multiple Regression. In Internet of Things and Analytics for Agriculture, Volume 2 (pp. 261-288). Springer, Singapore.	**This article discusses the concepts of Precision Agriculture and applications of the Internet of things (IoT) in agriculture, the design and challenges of an IoT-enabled system for highland crops management, and utilized multiple regression as a frost prediction techniqu**

Dong, Y., Fu, Z., Peng, Y., Zheng, Y., Yan, H., & Li, X. (2020). Precision fertilization method of field crops based on the Wavelet-BP neural network in China. Journal of Cleaner Production, 246, 118735.	**The author proposes a method for the precision fertilization of maize based on the wavelet-BP neural network which increases maize production, while reducing production cost and agricultural pollution.**
Meng, X., Liu, M., & Wu, Q. (2020). Prediction of Rice Yield via Stacked LSTM. International Journal of Agricultural and Environmental Information Systems (IJAEIS), 11(1), 86-95.	**This article proposes Long Short- Term Memory recurrent neural network (LSTM) methods of time series prediction of Rice Yield.**
Maldaner, L. F., & Molin, J. P. (2020). Data processing within rows for sugarcane yield mapping. Scientia Agricola, 77(5).	**The author tested a method for sugarcane yield data cleaning, investigating if the data recording frequency influences the characterization of yield variations in mapping high-resolution spatial data within a single row.**

7)Novelty in Research Proposal:

- **This research suggests a way to better ground their recommendations for seed selection decision on variety of paddy by the farmers by using the application to be built.**
- **The big data analytics framework proposed in this research, which integrates methodologies from machine learning, optimization, and simulation has great potential to improve the quality of seed selection decision, and ultimately improve farmers' crop yield.**
- **Further agricultural department community actively embraces machine-learning and data-driven decision-making methodologies to solve important problems in the agriculture domain.**
- **The method proposed for assessing the environmental impact of agriculture on the scale of a farming region cover a wide range of objectives, users and concepts.**
- **The following six main types of method used for analysis: environmental risk mapping, life cycle analysis, environmental impact assessment, multi-agent system, linear programming and agro-environmental indicators**

8) References:

- **Deep transfer learning for crop yield prediction with remote sensing data ,AX Wang, C Tran, N Desai, D Lobell - Proceedings of the 1st 2018 - dl.acm.org**
- **Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review,A Chlingaryan, S Sukkarieh, B Whelan - Computers and electronics in 2018.**
- **Deep gaussian process for crop yield prediction based on remote sensing data. J You, X Li, M Low, D Lobell… - Thirty-First AAAI …, 2017**
- **Performance Evaluation of Machine Learning Techniques for Mustard Crop Yield Prediction from Soil Analysis,V Pandith, H Kour, S Singh, JManhas… - Journal of Scientific …, 2020.**
- **An improved machine learning algorithm for predictingblast disease in paddy crop Sreevallabhadev Radhakrishnan July 2020.**
- **Deep learning approach for recognition and classification of yield affecting paddy crop stresses using field imagesArtificial Intelligence in Agriculture2020.Basavaraj S. Anami Naveen N. Malvade Surendra Palaiah**
- **Machine-Learning-Based Simulation for Estimating Parameters in Portfolio Optimization: Empirical Application to Soybean Variety Selection Durai Sundaramoorthi, Lingxiu Dong Olin Business School, Washington University in St. Louis.**

9) Additional Information:

The research proposed is to be carried out in Madurai District

- **The main aim of the project proposal is to improve the economic status of the farmer by doubling the production and tripling the income.**
- **For Second Green Revolution Agriculture Department has formulated various policies and strategies to achieve equitable, competitive and sustainable growth in Agriculture Crops. Madurai District is mainly agrarian with an average rainfall of 874.5 mm. Paddy, Millet, Pulses, Cotton are the major crops cultivated in Madurai District. Periyar – Vaigai canal is the main source of irrigation.**
- **The average yield of paddy is high in Madurai regions where the seed replacement rate is high.**
- **The Data collected for research is as follows:**

1. Farming Population (as per 2011 census)
Small Farmers – 58,716
Marginal Farmers – 2, 05,771
Agricultural Laborers- 4, 01, 867
2. No. of Block Assistant Director of Agriculture- 13
3. No. of Agricultural Extension Centre to later farmers needs of Agriculture inputs and technical advices Main AEC – 13, Sub AEC – 18
4. Seed Processing Unit – 2
State Seed Farm – 1
5. Soil Testing Laboratory – 2
Fertilizer Testing Laboratory – 1
Perticide Testing Laboratory – 1
Liquid Bio Fertilizer Production unit – 1
Bio Control Laboratory – 1
IPM Control Laboratory – 1
Sugarcane Parasite Breeding Station – 1
6. Agricultural College & Research Institute – 1
7. Krishi Vigyan Kendra – 1
8. Total No. of Fertilizer Shops – 413
9. Total No. of PACCS – 182
10. Cropping Intensity – 1.16
11. Irrigation Intensity – 1.18