

LITERATURE SURVEY

TITLE	AUTHOR'S NAME	ABSTRACT	METHODOLOGY
CROP YIELD PREDICTION IN AGRICULTURE USING DATA MINING PREDICTIVE ANALYTIC TECHNIQUES	1 P.Surya, 2 Dr. I.Laurence Aroquiaraj	Data Mining is emerging research field in Agriculture especially in crop yield analysis and prediction. As early into the growing season as possible, a farmer is focused in perceptive how much yield they about to expect. As with many other sectors the amount of agriculture data are increasing on a daily source. In our proposed work, collected agriculture dataset will be used to get crop yield prediction model using various regression techniques. Regression analysis was tested for the effective prediction or forecast of the	Data Mining

		<p>agriculture yield for various crops in Tamilnadu state particularly in North Western zone of Tamilnadu. North western zone of tamilnadu state data consist four districts. The North western zone of Tamilnadu districts are Dharmapuri, Salem, Namakkal, Krishnagiri. By the analysis depends on the results of predictor model, in the north western zone, under the area having more cultivated crops are Tapiaco, Sugar cane, Ragi, Maize, Groundnut.</p>	
<p>CROP YIELD PREDICTION BASED ON INDIAN AGRICULTURE USING MACHINE LEARNING</p>	<p>1 Potnuru Sai Nishant , 2 Pinapa Sai Venkat, 3 Bollu Lakshmi Avinash , 4 B. Jabber</p>	<p>In India, we all know that Agriculture is the backbone of the country. This paper predicts the yield of almost all kinds of crops that are planted in India. This script</p>	<p>Stacked Regression</p>

		<p>makes novel by the usage of simple parameters like State, district, season, area and the user can predict the yield of the crop in which year he or she wants to. The paper uses advanced regression techniques like Kernel Ridge, Lasso and ENet algorithms to predict the yield and uses the concept of Stacking Regression for enhancing the algorithms to give a better prediction.</p>	
<p>FORECASTING MODEL FOR VEGETABLE PRICE USING BACK PROPAGATION NEURAL NETWORK</p>	<p>1 G.M.Nasira, 2 N.Hemageetha</p>	<p>The Agricultural sector needs more support for its development in developing countries like India. Price prediction helps the farmers and also the Government to make effective decision.</p> <p>Based on the complexity of vegetable price</p>	<p>Neural Networks</p>

		<p>prediction, making use of the classification technique like neural networks such as self build up the model of Back-propagation neural network (BPNN) to predict vegetable price. A prediction model was set up by applying the neural network. Taking tomato as an example, the parameters of the model are analyzed through experiment. At the end of the result of Back shows accuracy percentage of the price prediction.</p>	
<p>DISEASE PREDICTION ON SOIL MICRONUTRIENTS ANALYSIS OF BT COTTON BY J48 CLASSIFICATION</p>	<p>1 Youvrajsinh Chauhan, 2 Jignesh Vania</p>	<p>The decision tree is one of the common modelling methods to classify. This is an emerging research field that is experiencing a constant development. In this paper, we present the basic knowledge of 16 soil nutrients of crop and</p>	<p>Classification</p>

		<p>how related with other factors like Soil PH, Cation Exchange, Soil Fertility. We add general concept of decision tree and recent applications that already tested and predict the use in field of agriculture. We make the conceptual model that helps to feature research work in agriculture with use of soil nutrients.</p>	
<p>A STUDY OF CLUSTERING TECHNIQUES FOR CROP PREDICTION - A SURVEY</p>	<p>1 Utkarsha P. Narkhede, 2 K. P. Adhiya</p>	<p>Farming community necessitate for well organized system to predict and improve the crop over the world. The complexity of predicting the best crops is highly due to unavailability of proper knowledge discovery in crop knowledgebase which affects the quality of prediction. However, Clustering is an important step in mining useful information. There are</p>	<p>Clustering</p>

		<p>several clustering methods such as partitioning, hierarchical, model-based, grid-based, constrained-based which make this task complicated due to problems related to optimization and noise. In this review paper there is a comparative study of clustering algorithms. Out of these BeeHive and Improved k-means clustering algorithm are outstanding in solving the optimization problem which led to select for performance evaluation in order to get good quality of clusters.</p>	
<p>KNOWLEDGE DISCOVERY ON AGRICULTURAL DATASET USING</p>	<p>1 Farah Khan, 2 Dr. Divakar Singh</p>	<p>Data mining is a technique of analyzing the dataset such that the final conclusion can be accessed easily and quickly from the dataset. Here in this</p>	<p>Association Rule Mining</p>

ASSOCIATION RULE MINING		<p>paper association rule mining technique is implemented for the analysis of agricultural dataset. The idea is to apply association rule mining technique for generating rules and establishing a relationship between them .So as to enhance crop productivity. Also a performance comparison is done between Apriori and FP Growth algorithm.</p>	
PREDICTION OF THE PRODUCTION OF CROPS WITH RESPECT TO RAINFALL	1 Benny Antony	<p>Agriculture is one of the most important sectors in the Indian context. It is one of the highest employing sectors in the Indian scenario. Unlike other sectors agriculture is highly dependent on the quality and the quantity of both the external factors like rainfall, climate, pH of the soil, fertilizers and insecticides used, and</p>	Regression Method

		<p>internal factors like the quality of seeds. This paper predicts the production of crops as a function of rainfall for four Indian States. This knowledge can be implemented in generating a rough overview of how the production is based on rainfall and how much can a specific crop production for the amount of rainfall it receives. Two crops each belonging to four different states are chosen and the best regression model for the crop of the state is chosen. There is no research done solely on how rainfall affects crops of particular states. The proposed method of evaluation is better than other existing methods of evaluation as it evaluates all the regression techniques (Linear Regression,</p>	
--	--	---	--

		Polynomial Regression, Support Vector Regression, Decision Tree Regression, Random Forest, and XGBRegression) for two crops of four individual states. For balanced evaluation, two states of North India and two states of South India are selected. The regression techniques are evaluated based on their Mean Squared Error	
AN APPLICATION OF MACHINE LEARNING TECHNIQUE IN FORECASTING CROP DISEASE	1 Gianni Fenu, 2 Francesca Maridina Mallocci	In the recent years, Big Data Analytics and Machine Learning techniques are playing an increasingly key role in the agriculture sector in order to tackle the increasing challenges due to the climate changes which are causing serious damage production. The analysis of environmental, climatic	Support Vector Machine Classification

		<p>and cultural factors allows to establish the irrigation and nutritional needs of crops, forecast crop disease, improve crop yield, as well as improve the quantity and the quality of agricultural output while using less input. Potato late blight is considered one of the most devastating disease world over, including Sardinia. Unexpected epidemics can result in significant economic and yield losses. In this paper, we describe the test conducted using the DSS LANDS in order to predict potato late blight disease in Sardinia. The object of the study was to investigate if regional weather variables could be used to predict potato late blight risk in southern Sardinia using a Machine Learning approach. The disease</p>	
--	--	---	--

		severity is predicted using Feed-forward Neural Network and Support Vector Machine Classification based on meteorological parameters provided by ARPAS weather stations. The prediction accuracy for ANN was 96% and for SVM Classification was 98%	
WHEAT YIELD PREDICTION USING MACHINE LEARNING AND ADVANCED SENSING TECHNIQUES	X.E., Moshou, D., Alexandridis, T., Whetton, R.L. and Mouazen	Understanding yield limiting factors requires high resolution multi-layer information about factors affecting crop growth and yield. Therefore, on-line proximal soil sensing for estimation of soil properties is required, due to the ability of these sensors to collect high resolution data (>1500 sample per ha), and subsequently reducing labor and time cost of soil sampling and analysis. The aim	Supervised Kohonen Networks

		<p>of this paper is to predict within field variation in wheat yield, based on on-line multi-layer soil data, and satellite imagery crop growth characteristics.</p> <p>Supervised self-organizing maps capable of handling existent information from different soil and crop sensors by utilizing an unsupervised learning algorithm were used.</p> <p>The performance of counter-propagation artificial neural networks (CP-ANNs), XY-fused Networks (XY-Fs) and Supervised Kohonen Networks (SKNs) for predicting wheat yield in a 22 ha field in Bedfordshire, UK were compared for a single cropping season. The self organizing models consisted of input nodes corresponded to</p>	
--	--	---	--

		<p>feature vectors formed from normalized values of on-line predicted soil parameters and the satellite normalized difference vegetation index (NDVI). The output nodes consisted of yield isofrequency classes, which were predicted from the three trained networks. Results showed that cross validation based yield prediction of the SKN model for the low yield class exceeded 91% which can be considered as highly accurate given the complex relationship between limiting factors and the yield. The medium and high yield class reached 70% and 83% respectively. The average overall accuracy for SKN was 81.65%, for CP-ANN 78.3% and for XY-F 80.92%, showing that the SKN model had the</p>	
--	--	--	--

		best overall performance.	
A SMART FARMING APPROACH IN AUTOMATIC DETECTION OF FAVORABLE CONDITIONS FOR PLANTING AND CROP PRODUCTION IN THE UPPER BASIN OF CAUCA RIVER	1 López, I.D. , 2 Corrales, J.C.	The upper basin of Cauca River is one of the most important agricultural and industrial corridors in Colombia. In addition, Colombian relief has a variety of soils and climatic zones, in which the slightest modification of agro-climatic conditions could imply a great impact on sowing, production, yield, crop quality, and hence, the farmers' income sources. However, these conditions are not the only factors that influence a crop growth; dimensions like biophysical, economic-productive, political-institutional, and socio-cultural establish interrelations that affect the viability of planting and produce	Big Data Analytics

		<p>a crop in a geographic territory. For a more efficient information management, Smart Farming represents the application of information and communication technologies in order to support more intelligent decisions in the agricultural sector. This study presents a theoretical proposal based on Big Data Analytics to detect appropriate conditions for sowing and crop production in the upper basin of Cauca River. Similarly, the preliminary public datasets (corresponding to biophysical and economic-productive) and main proposal components are described in a general view. Finally, a future implementation and evaluation of this research work is proposed.</p>	
--	--	--	--

INTERNET OF THINGS-DRIVEN DATA MINING FOR SMART CROP PRODUCTION PREDICTION IN THE PEASANT FARMING DOMAIN	Colombo-Mendoza, L.O., Paredes-Valverde, M.A., Salas-Zárate, M.D.P. and Valencia-García, R.	Internet of Things (IoT) technologies can greatly benefit from machine-learning techniques and artificial neural networks for data mining and vice versa. In the agricultural field, this convergence could result in the development of smart farming systems suitable for use as decision support systems by peasant farmers. This work presents the design of a smart farming system for crop production, which is based on low-cost IoT sensors and popular data storage services and data analytics services on the cloud. Moreover, a new data-mining method exploiting climate data along with crop-production data is proposed for the	Internet of Things, Data Mining

		<p>prediction of production volume from heterogeneous data sources. This method was initially validated using traditional machine-learning techniques and open historical data of the northeast region of the state of Puebla, Mexico, which were collected from data sources from the National Water Commission and the Agri-food Information Service of the Mexican Government.</p>	
--	--	---	--