VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belgaum-590018, Karnataka

A project phase-I report on

"CROP YIELD PREDICTION USING DEEP LEARNING"

Submitted in fulfillment for the requirements of VII semester degree of

BACHELOR OF ENGINEERING

IN

INFORMATION SCIENCE AND ENGINEERING

By

Shantideepa Samanta (1DB16IS035)
Kundan Prasad Kumar (1DB17IS018)

Under the Guidance of

Mrs. Manjula K.

Assistant Professor

Dept of Information Science and Engineering

Don Bosco Institute of Technology

Department of Information Science and Engineering

DON BOSCO INSTITUE OF TECHNOLOGY

Kumbalagodu, Mysore road, Bangalore-560074

2020-2021

Don Bosco Institute of Technology

Kumbalgodu Mysore Road
Bangalore-560074

DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING

CERTIFICATE

Certified that the Project on topic **Crop Yield Prediction using Deep Learning** has been successfully presented at **Don Bosco Institute of Technology** by **Shantideepa Samanta (1DB17IS035) and Kundan Kumar Prasad (1DB17IS018),** in partial fulfillment of the requirements for the *VII Semester degree of* ***Bachelor of Engineering in Information Science and Engineering of*** *Visvesvaraya Technological University, Belagavi* during academic year 2020-2021. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The Project report has been approved as it satisfies the academic requirements in respect of Project work for the said degree

Mrs. Manjula K. **Assist Professor** **Dept. of ISE**	**Mr. Shankara Gowda S R** **Assist Professor** **Dept. of ISE**	**Assoc Prof. Gowramma G S** **HOD of ISE Dept.** **DBIT, Bangalore**

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crowned the efforts with success.

I would like to profoundly thank **Management** of **Don Bosco Institute of Technology** for providing such a healthy environment for the successful completion of Project work.

I would like to express my thanks to the Principal **Dr. Hemadri Naidu T** for their encouragement that motivated me for the successful completion of Project work.

It gives me immense pleasure to thank **Professor Gowramma G S**, Head of Department for his constant support and encouragement.

Also, I would like to express my deepest sense of gratitude to **Mrs. Manjula K**, Assistant Professor and **Mr. Shankara Gowda S R,** Assistant Professor, Department of Information Science & Engineering for their constant support and guidance throughout the Project work.

I would also like to thank the Laboratory System Administrator **Mr. Rangaswamy**, Department of Information Science & Engineering and all other teaching and non-teaching staff of Information Science Department who has directly or indirectly helped me in the completion of the Project work.

- Shantideepa Samanta (1DB17IS035)

-Kundan Kumar Prasad (1DB17IS1018)

ABSTRACT

In India agriculture contributes approximately 23% of GDP and employed workforce percentage is 59%. India is the second-largest producer of agriculture crops. The technological contribution may help the farmer to get more yields. The prediction of the yield of different crops may help the farmer regarding taking the decision about which crop to grow. The research focuses on the prediction of different crops Yield using neural network regression modeling. The data of crop cycle for summer, Kharif, rabi, autumn and whole year is used.

The dataset is filtered using Python Pandas and Pandas Profiling tools to retrieve data. The model is developed using a Multilayer perceptron neural network. Initially the result obtained considering optimizer RMS prop with accuracy 45 %, later it will be enhanced to 90% by increasing layers, adjusting weight, bias and changing optimizer to Adam. This research describes the development of a different crop yield prediction model with ANN, with 3 Layer Neural Network. The ANN model develops a formula to ascertain the relationship using a large number of input and output examples, to establish model for yield predictions an Activation function: Rectified Linear activation unit (Relu) is used. The backward and forward propagation techniques are used

CONTENTS

Chapter 1

INTRODUCTION

The importance of Crop Yield is reflected in recent food security assessment survey conducted by the Department for Environment, Food and Rural Affairs (DEFRA) which concluded that crops are the key indicators in determining global food availability.

Achieving high crop yields is the principle aim of agricultural production. The recognition and management of factors that influence crop yield assist farmers in decision making. The farmer can make a decision about crop choice and can contribute more to its profit. There are a large number of crop yield prediction models available which may use weather real parameters or static parameters.

There are a number of crop yield prediction models which use either statistical or crop simulation models. Machine learning is found to be a very appealing field that can contribute to the agriculture field. The different models built using machine learning can take different crisp inputs to give some concrete output. Over the last decade it has been observed that Artificial Intelligence (AI) techniques provide a more effective approach to predicting crop yield under different cropping scenarios. The use of artificial neural networks can make models with complex inputs easier to interpret.

1.1 Problem Statement

Machine learning is also being used in agriculture for several years. Crop yield prediction is one of the challenging problems in precision agriculture, and many models have been proposed and validated so far. This problem requires the use of several datasets since crop yield depends on many different factors such as climate, weather, soil, use of fertilizer, and seed variety. This indicates that crop yield prediction is not a trivial task; instead, it consists of several complicated steps. Nowadays, crop yield prediction models can estimate the actual yield reasonably, but a better performance in yield prediction is still desirable.

The common problem in existing crop yield prediction methods are given below,

- The most important problem of existing crop yield prediction method is accuracy and time consuming problem.
- In existing time series crop yield prediction method does not react to variations that occur for cycles and seasonal effects.
- Needs extensive information to develop and test the model and also available information in agriculture is sparse and incomplete in existing simulation model.
- Limited studies have been made in crop yield prediction using existing decision tree technique.
- Prediction error value also important problem in crop yield prediction or estimation methods.
- These are the main drawbacks of various existing works, which motivate us to do this research on crop yield prediction.

1.2 Aims and Objective of the Project:

The influence of climate change and its unpredictability has caused majority of the agricultural crops to be affected in terms of their production and maintenance. Therefore, in order to make a prior estimation of the crop yield, considering the previous crop yield records and data, we draw the main objective of this project as follows:

- Forecasting or predicting the crop yield well ahead of its harvest time to assist the strategists and farmers for taking suitable measures for harvest and storage.
- Enhancing the accuracy of prediction, or success rate by using artificial neural networks algorithm, which will further help to resolve the strategizing issues for agronomic production purposes**?**

1.3 Existing System:

Crop yield prediction is a very active area of current research interest and has been so since the 1980s. However, in the early days the work was mainly concerned with the study of linear

systems models and hence was only concerned with the linear relationships among the various agricultural parameters. Therefore, most of the conventional or traditional models are not able perform well because they were not able to effectively deal with the complexity and non-linear nature of the data.

Basically, crop prediction models can be divided into two classes; statistical model and crop simulation model. The early stage of the modeling usually involves statistical methods. This is where the systems use various regression techniques that compute crop yield empirically. On the other hand, simulation models involve the physiologically – based systems of either crops or plant which can affect growth either internally or externally; and is normally involve mathematical analysis in order to predict yield. As an example, let us consider one of the wheat yield prediction models as in the case of the Crop Estimation through Resource and Environment Synthesis – Wheat (CERES) model. This type of simulation model uses a set of data which includes the weather, soil attributes and the detailed management practice which specific farm uses. The next model is representative of the more complex models where a very complex set of data is used to predict wheat yield. ECOSYS and SIRIUS are also categorized as complex models that use lots of different types of data and rely heavily on computer design to simulate the growth of wheat as in the CERES.

1.3.1 Drawbacks in the existing System:

1. Although simulation and statistical models have improved to become better crop prediction models, they still not able effectively with a complex data set.
2. Limited studies have been made in crop yield prediction using existing decision tree technique.

1.4 Proposed System:

Models based on Intelligent Systems (ISs) techniques are able to overcome this limitation. This type of technique can produce good results by manipulating raw and simple or complex data which they perform in competitively with the more complex models. The most popular ISs

technique which has been used for crop or wheat prediction models is Artificial Neural Networks (ANNs).

The project focuses on the prediction of different crops yield using **neural network regression modeling**. The data of crop cycle for summer, Kharif, Rabi, autumn and whole year will be used. The experimental parameters that will be considered for the study are cultivation area, crop, state, district, season, year and production or yield etc. The dataset will be filtered using Python Pandas and Pandas Profiling tools to retrieve data.

The model will be developed using a Multilayer perceptron neural network. This project will describe the development of a different crop yield prediction model with artificial neural network (ANN). Also, the proposed machine learning model built using, the artificial neural network (ANN) will be evaluated to produce higher prediction accuracy over other models.

Artificial Neural network is the system designed to work like a human brain. The computational system inspired by but not identical to the human brain. The ANN systems or model learn from the facts and experiences feed to them and will react to the input situation depending on training given to it. As we add more training it will work more efficiently. The artificial neural network for our study comprised of three-layer. The circle represents nodes and arrow represents the relationship between the input layer and the output layer.

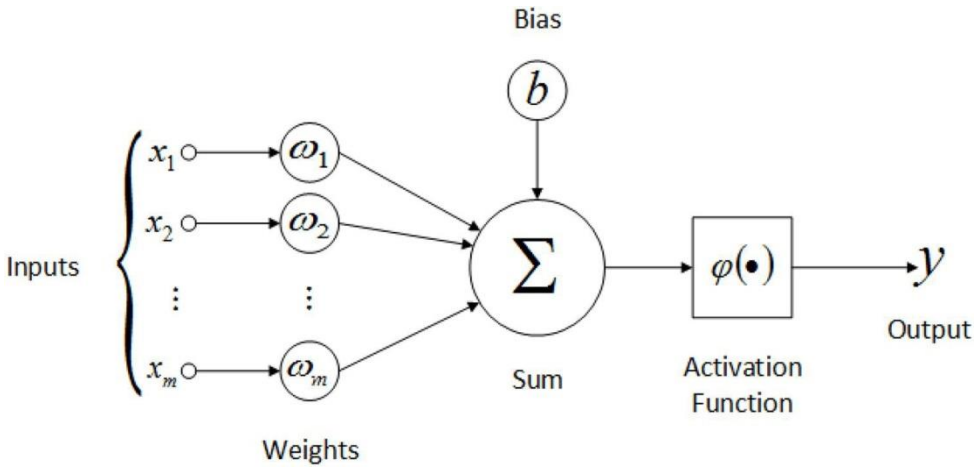

FIG.1 The neural network structure on how data are being processed through each receptor.

The ANN model develops a formula to ascertain the relationship using a large number of input and output examples, to establish model for yield predictions. Also the error rate with the actual will be shown with the assist of Mean Square Error (MSE).

1.4.1 Advantages of Proposed system:

- This type of technique can produce good results by manipulating raw and simple or complex data which they perform in competitively with the more complex models.
- Information such as in traditional programming is stored on the entire network, not on a database. The disappearance of a few pieces of information in one place does not restrict the network from functioning.
- Corruption of one or more cells of ANN does not prevent it from generating output. This feature makes the networks fault-tolerance

Chapter 2

LITERATURE SURVEY

1. It has been observed that with algorithms relating to supervised learning OPF, k-NN and SVM showed better responses compared to unsupervised learning ANN and CNN. ANN had better accuracy when sufficient images are provided to classify the changes within. Different approaches in prediction are also being considered, such as [2] predicted in two steps first of which dedicating ANN to deduce the average rainfall and then using SVR (Support Vector Regression) to predict the changes in the crop yield.

2. There are other source of information that can help providing better prediction of the result i.e. soil parameters from the satellite normalized difference vegetation index (NDVI), the objective of the following paper [3] (used as a reference for this project) is to analyze the main soil properties such as organic matter, essential plant nutrients, micronutrient that affects the growth of crops and find out the suitable relationship percentage among those properties using Supervised Learning, Back Propagation Neural Network.

3. In 2018, Liakos et al. [4], has provided a review on how the standard measure varies from using of different structured algorithm in deducing the predicted output from a sample of related dataset. It has been also said that fusion of different source of information may not be needed or less effective in the change in outcome.

4. Also according to the additional analysis, it was found that apart from ANN, Convolution Neural Networks (CNN) is the most widely used deep learning algorithm in these studies, and the other widely used deep learning algorithms are Long-Short Term Memory (LSTM) and Deep Neural Networks (DNN), described in [5].

Chapter 3

REQUIREMENTS AND SOFTWARE USED

3.1 Software Requirement Specification

A Software Requirements Specification (SRS) -a requirements specification for a software system – is a complete description of the behavior of a system to be developed. In addition to a description of the software functions, the SRS also contains non-functional requirements. Software requirements are a sub-field of software engineering that deals with the elicitation, analysis, specification, and validation of requirements for software.

3.2 Requirements:

3.2.1 Hardware Requirements:

- System : Pentium IV 2.4 GHz or more
- Hard Disk : 40 GB.
- Monitor : 15 VGA Color.
- Ram : 512 Mb
- Processor : Dual Core

3.2.2 Software Requirements:

- TensorFlow
- Pandas
- Numpy
- Matplotlib/Seaborn
- Sklearn
- Flask

3.3 Description about modules:

3.3.1 TensorFlow:

TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. TensorFlow is a symbolic math library based on dataflow and differentiable programming. It is used for both research and production at Google. TensorFlow is available on 64-bit Linux, macOS, Windows, and mobile computing platforms including Android and iOS. Its flexible architecture allows for the easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices

Fig 2. TensorFlow

3.3.2 Pandas:

Pandas are a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language. In computer programming, pandas are a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license. The name is derived from the term "panel data", an econometrics term for data sets that include observations over multiple time periods for the same individuals. Its name is a play on the phrase "Python data analysis" itself. Wes McKinney started building what would become pandas at AQR Capital while he was a researcher there from 2007 to 2010.

Fig 3. Pandas

3.3.3 NumPy:

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Holguin with contributions from several other developers. In 2005, Travis Oliphant created NumPy by incorporating features of the competing Numarray into Numeric, with extensive modifications. NumPy is open-source software and has many contributors.

Fig 4. Numpy

3.3.4 MatplotLib:

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+. There is also a procedural "pylab" interface based on a state machine (like OpenGL), designed to closely resemble that of MATLAB, though its use is discouraged. SciPy makes use of Matplotlib.

Fig. 5 MatplotLib

3.3.5 Sklearn:

Scikit-learn (formerly scikits.learn and also known as sklearn) are a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

Fig. 6 Sklearn

3.3.6 Flask:

Flask is a micro web framework written in Python. It is classified as a micro framework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions. However, Flask supports extensions that can add application features as if they were implemented in Flask itself. Extensions exist for object-relational mappers, form validation, and upload handling, various open authentication technologies and several common framework related tools

Chapter 4

PROJECT DETAILS

4.1 System Design:

System design is the process of defining the architecture, components, modules, interfaces and data for a system to satisfy specified requirements. System design could see it as the application of systems theory to product development. Theory is some overlap with the disciplines of system analysis, systems architecture and systems engineering.

If the broader topic development "blends the perspective of marketing, design, and manufacturing into a single approach to product development," then design the act of talking the marketing information and creating the design of the product to be manufactured. Systems design is therefore the process of defining and developing systems to satisfy specified requirements of the user.

Until the 1990s systems design had crucial and respected role in the data processing industry. In the 1990s standardization of hardware and software resulted in the ability to build modular systems. The increasing importance of software running on generic platforms has enhanced the discipline of software engineering.

Object-oriented analysis and design methods are becoming the most widely used methods for computer systems design. The UML has become the standard language in object-oriented analysis and design. It is widely used for modeling software systems and is increasingly used for high designing non- software systems and organizations.

System design is one of the most important phases of software development process. The purpose of the design is to plan the solution of a problem specified by the requirement documentation. In other words the first step in solution is the design of the project.

The design of the system is perhaps the most critical factor affecting the quality of the software. The objective of the design phase is to produce overall design of the software. It aims to figure out the modules that should be in the system to fulfill all the system requirements in efficient manner.

The design will contain the specification of all the modules, their interaction with other modules and the desired output from each module.

4.2 High level design:

4.2.1 Data flow diagram

A data flow diagram (DFD) is a graphical representation of the flow of the visualization of data processing. On a DFD, data items flow from an external data source or internal data source to internal data source or external data sink via an internal process. DFD provides no information about the timing of process or about whether process will operate in sequence or in parallel.

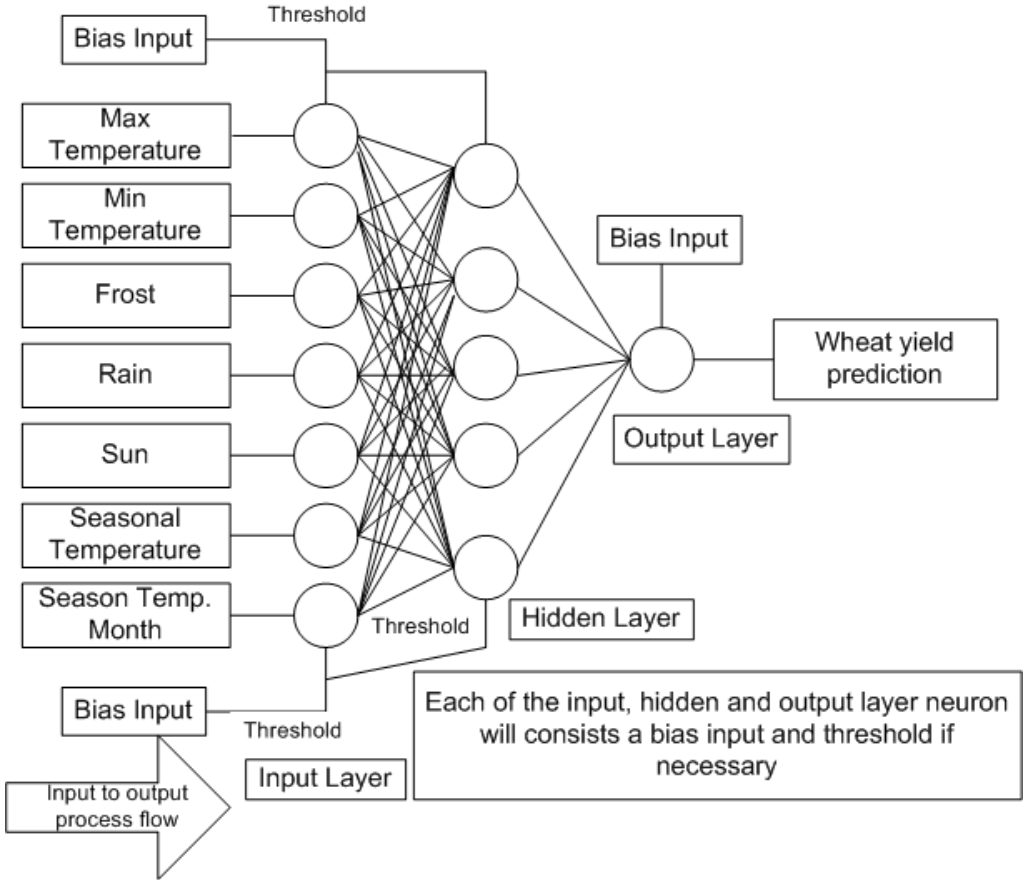

Fig. 7 Design and Network Architecture of the model

Chapter 5

SYSTEM IMPLEMENTATION AND METHODOLOGY

Implementation is the realization of an application, or execution of a plan, idea, model, design, specification, standard, algorithm, or policy. In other words, an implementation is a realization of a technical specification or algorithm as a program, software component, or other computer system through programming and deployment. Many implementations may exist for a given specification or standard.

Implementation is one of the most important phases of the Software Development Life Cycle (SDLC). It encompasses all the processes involved in getting new software or hardware operating properly in its environment, including installation, configuration, and running, testing, and making necessary changes. Specifically, it involves coding the system using a particular programming language and transferring the design into an actual working system. This phase of the system is conducted with the idea that whatever is designed should be implemented; keeping in mind that it fulfils user requirements, objective and scope of the system. The implementation phase produces the solution to the user problem.

5.1 Module Description:

The Project is divided into 2 different modules:

1. Crop yield prediction:

   "Prediction" refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome, such as whether or not a customer will churn in 30 days. The algorithm will generate probable values for an unknown variable for each record in the new data, allowing the model builder to identify what that value will most likely be.

   This module will be responsible for predicting the yield of a particular crop based on factors like geographical data, climatic data, soil data and fertilizer data of the area where it is to be cultivated. Based on the prediction result the system will also make crop

recommendations in concern with the next module described below.

2. Crop Recommendation:

   Recommender systems are an important class of machine learning algorithms that offer "relevant" suggestions to users. Practically, recommender systems encompass a class of techniques and algorithms which are able to suggest “relevant” items to users. Ideally, the suggested items are as relevant to the user as possible, so that the user can engage with those items: YouTube videos, news articles, online products, and so on.

Based on the result of the Prediction system, if the result indicates poor yield on a crop then this module provides a recommendation of crops that could produce a better yield thereby enhancing the profit and cultivation ease of the farmer.

5.2 Model Architecture:

In order to train the ANN to predict wheat yield using the wheat data, we use the Levenberg-Marquadt (LM) technique because it performs better than other approaches such as Gauss-Newton (GN) method and gradient descent algorithm. It also typically performs better in terms of, for example, fastest convergence and also it produces better results; for example the mean square error is lower for this yield prediction model. In other case, the main problem in this model is over fitting which may happen when the training set error becomes relatively small.

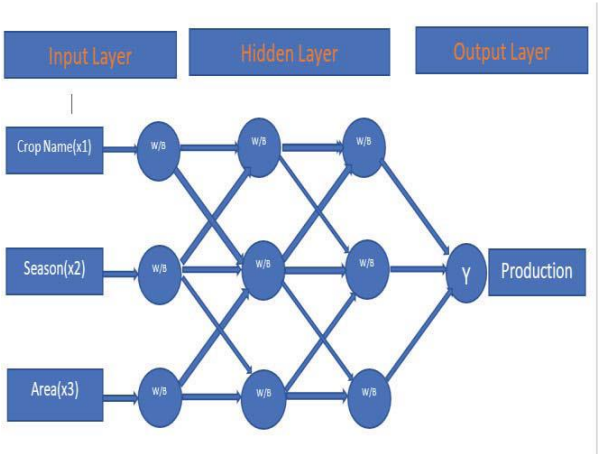

Fig. 5. ANN Model of the Proposed Project

5.3 Overview of ANN Algorithm:

A neural network consists of a set of highly interconnected entities, called nodes or units. Each unit is designed to mimic its biological counterpart, the neuron. Each neuron accepts a weighted set of inputs and responds with an output. Artificial Neural Network (ANN) technique is based upon the model of biological nervous system. The key element of this technique is the novel structure of the information processing system. An ANN is configured for a specific application such as pattern recognition or data classification, through a learning process.

The most important types of Neural Networks for real world problem solving are Multilayer Perceptron, Radial Basis Function Networks and Kohonen Self Organizing Feature Maps.

For present research the multilayer perceptron was used as artificial neural network for the prediction of rice crop yield for Karnataka state, India.

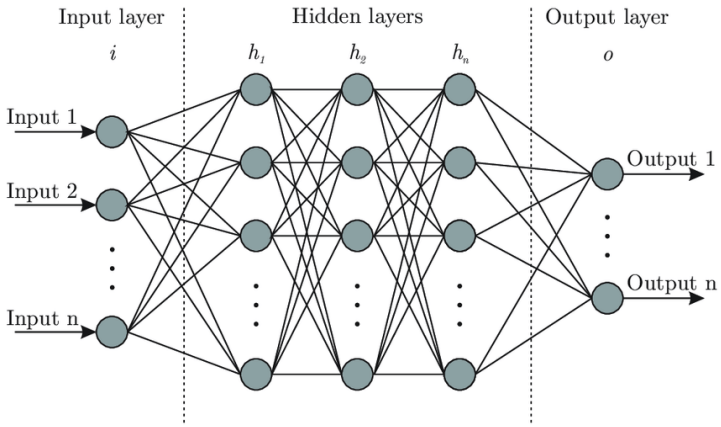

Fig. Artificial Neural Network Architecture

5.4 Dataset Description

The parameters considered for the present project are described below:

Precipitation (mm): Precipitation is a noteworthy segment of the water cycle, and is responsible for saving the fresh water on the planet. The aggregate precipitation for Kharif season (June to November) for each year of each district was computed from the monthly mean precipitation for every year.

Minimum, Average, Maximum Temperature (degree Celsius): Temperature is an essential variable influencing the rate of plant development. Temperatures along with environmental change and the potential for more extreme temperature events will affect plant productivity i.e. crop production. The average, minimum and maximum temperature for the Kharif season (June to November) were calculated for each district on the basis of monthly mean temperatures.

Area (Hectares): The crop cultivated area of the study area in Kharif season (June to November) will be considered for the present project.

Production (Tonnes): The crop production for the study area for Kharif season (June to November) will be considered for the present project.

Yield (Tonnes/Hectare): An amount of crop yield and the area cultivated for a specific crop in Kharif season, of each of the selected district from the Karnataka state, the yield will be calculated and considered for the present project.

5.5 Pseudo Code:

Pseudo code is an informal high-level description of the operating principle of a computer program or other algorithm. It uses the structural conventions of a programming language, but is intended for human reading rather than machine reading. Pseudo code typically omits details that are not essential for human understanding of the algorithm, such as variable declarations, system-specific code and some subroutines. The programming language is augmented with natural language description details, where convenient, or with compact mathematical notations. The purpose of using pseudo code is that is easier for people to understand than conventional programming language code and that it is an efficient and environment independent description of the key principles of an algorithm. It is commonly used in textbooks and scientific publications that are documenting various algorithms, and also in planning of computer program development, for sketching out the structure of the program before the actual coding takes place. No standard for pseudo code syntax exists, as a program in pseudo code is not an executable program. Pseudo code resembles, but should not be confused with skeleton programs, including dummy code,

which can be compiled without errors. Flowcharts and Unified Modeling Language (UML) charts can be thought of as a graphical alternative to pseudo code, but are more spacious on paper.

Both the above mentioned Modules will follow the basic pseudo code briefly discussed below:

The Proposed study will be conducted using Python **matplotlib** and **Seaborn** which is used for data visualization. Data Pre-processing and Data cleaning processes are performed by **Pandas** library of python. Basically broad five steps are used for experiment that are

1. Data collection

2. Data Wrangling

3. Data Pre-processing

4. Data Visualization (Different Visualization Library Used)

5. Explotary Data Analysis (EDA)

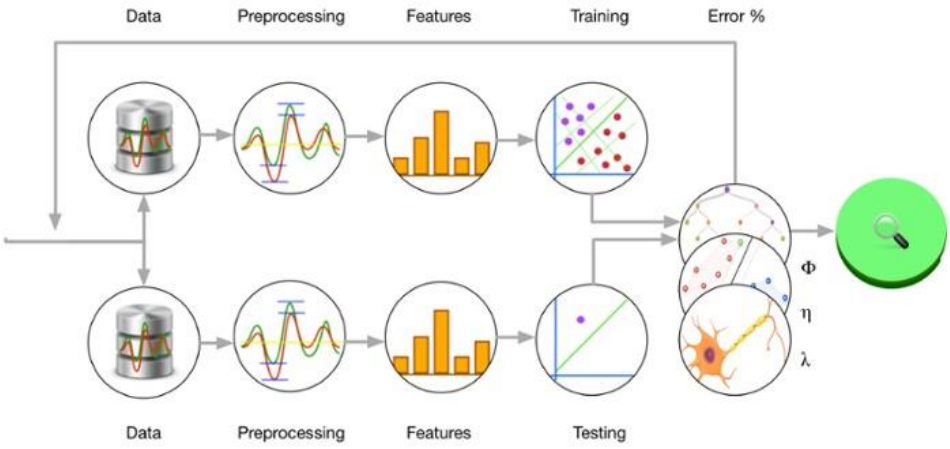

Fig. Steps for model building

The above steps are further explained in detail as follows which are followed for processing and preparing the data for applying the multilayer perceptron technique.

Step 1: **Acquiring each parameter (Area, Crop, State, District, Season, Year and Production**) monthly mean records.

Step 2: Selecting a subset of the **dataset.**

Step 3: Calculating the **Number of records Based on Year with Different categorical Variables**

Step 4: Visualizing a dataset with **matplotlib** and **Seaborn python library**

Step 5: Then row data will be passed through different data Pre-processing steps.

Step 6: Using statistical methods the relation between different variables like are analyzed. When Variables are Categorical used – Frequency, Count.

When variable is Numerical used – mean, median, mode, standard deviation. Exploration of variables or attributes can be done using univariate analysis. Type of variable could be either categorical or numerical. Investigation of each type of variable can be carried by different statistical and visualization techniques. Descretization or Binning process is used for transforming numerical variables to categorical variables. On the other hand, transformation of categorical Crops analysis for number of records for every crop so as to decide parameter priority for accuracy improvement.

Step 7: For preparing the data set for applying multilayer perceptron technique, unwanted columns were removed. They were sr. no, name of the district and year.

Step 8: The data set was then sorted on the basis of area. Outliers will be detected and omitted using various Data cleaning methods available.

Step 9: the data which is present in **label from converted to encoding using sklearn.**

Step 10: The dataset was then sorted on the basis production.

Step 11: we considered production as output parameter and **features like: crop, area, district, and season.**

Step 12: This data set was then saved in .csv format for further application of the multilayer perceptron technique in **Python TensorFlow**.

Step 13: Model is trained **Using linear Regression with Neural Network** with Adam optimizer and 3 layers.

Chapter 6

CONCLUSION AND FUTURE SCOPE

6.1 Conclusion:

The proposed model with back propagation is trying to reduce MSE by using RELU activation function and gradient descent. The Learning rate for each layer is kept constant i.e. 0.001.As we increase number of epochs error will get reduced. This result is used as input for deciding success rate of the crop over another crop. The best crop will be suggested to the farmer depending on the district and weather.

A non-linear technique is requisite to understand the association to find out the interactions between different parameters which are directly or indirectly affecting the crop yield. Due to complexity and severity of crop parameters a linear methodology is insufficient to conclude relationship between factors and crop yield.

6.2 Future Scope:

Traditional linear regression can be replaced with Artificial Neural Network methods to give better accuracy for crop prediction. The ANN algorithm can be improved further to give more accurate prediction by adding layers, and adding some more parameters. This yield prediction surely going to help farmers for better decision making regarding crop hasrvesting. The model also suggest success rate for the crops as per input given by the farmer. So model suggest best possible crop with highest success rate.

BIBLIOGRAPHY

[1]. (Base Paper) Shivani S. Kale, Preeti S. Patil, “A Machine Learning Approach to Predict Crop Yield and Success Rate” in 2019 IEEE Pune Section International Conference (PuneCon) MIT World Peace University, Pune, India. Dec 18-20, 2019

[2]. E. Khosla, R. Dharavath, and R. Priya, “Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression,” in Environment, Development and Sustainability, pp. 1–22, 2019.

[3]. Shivnath Ghosh, Santanu Koley, "Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks ", in International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 2 ISSN: 2321-8169.

[4]. K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine learning in agriculture: A review,” Sensors, vol. 18, no. 8, p. 2674, 2018.

[5]. (Thomas van Klompenburg, Ayalew Kassahun, Cagatay Catal, "Crop yield prediction using machine learning: A systematic literature review", in Computers and Electronics in Agriculture 177 (2020) 105709)