

**Tribhuvan University**

**Institute of Science and Technology**

**A Final Year Project Report**

**on**

**Weather Analysis for Crop Prediction**

**Submitted to:**

**Department of Computer Science and Information Technology**

**Ambition College**

**Baneshwor, Kathmandu, Nepal**

**In partial fulfilment of the requirements**

**For the Bachelors of Science in Computer Science and Information Technology**

**Submitted by:**

Rakesh Niraula (10612)

Ranjan Khatiwada (10614)

Subodh Shah (10624)

**Under the Supervision of**

**Mr.**

**February,2020**

# ABSTRACT

Nepal is richly endowed with agro-biodiversity. Despite being an agricultural country, the agricultural yield is found to be very low and the dependency to the nature is very high. The only solution to this is to introduce technology assisted farming. The production of crops mainly depends upon two major factors. One is weather condition and second being soil attributes. The analysis of weather using old dataset can help us in predicting the weather and soil of a place and the crop suitable for that particular location. The accuracy can be increased taking the soil type into account. The analysis of weather performs a key role in planning better farming structure. The use of technology can also help in minimizing the risks involved in agriculture to which the early farmers were awfully exposed.

***Keywords:*** *prediction model, dataset, Euclidian distance*, *test data*

**Table of Contents**

Abstract

[1. Introduction 1](#_Toc33300268)

[2. Problem Statement 1](#_Toc33300269)

[3. Objectives 1](#_Toc33300270)

[4. Scopes and Limitation 2](#_Toc33300271)

[4.1 Scope 2](#_Toc33300276)

[4.2 Limitation 2](#_Toc33300277)

[5. Research Methodology 2](#_Toc33300278)

[5.1 Literature Review 2](#_Toc33300280)

[6. Requirement Analysis 3](#_Toc33300281)

[6.1 Functional Requirement 3](#_Toc33300283)

[7. Feasibility Analysis 4](#_Toc33300284)

[7.1 Economic Feasibility 4](#_Toc33300286)

[7.2 Operational Feasibility 4](#_Toc33300287)

[7.3 Technical Feasibility 4](#_Toc33300288)

[7.4 Schedule Feasibility 4](#_Toc33300289)

[8. Implementation Tools 4](#_Toc33300290)

[9. Implementation Algorithm 5](#_Toc33300291)

[9.1 K-Nearest Neighbors 5](#_Toc33300294)

[10. Proposed Work 7](#_Toc33300295)

[11. Expected Output 7](#_Toc33300296)

[12. Working schedule 8](#_Toc33300297)

References

# Chapter 1: Introduction

## Introduction

Agriculture can be seen as the oldest occupation in the world and it has been evolving ever since along with the human race. Nepal is also an agricultural country and 66% of people are directly involved in farming [1]. Till this date, there’re so many research and developments in this field. There are almost all kinds of technologies which makes work easier in this field. Despite these developments and advancements, agriculture is facing so many challenges. As of now greatest challenges are climate change, degrading soil condition, and production should be increased making lesser investment and time.

Weather Analysis for Crop Prediction is a system combining of machine learning and web field to provide reliable and accurate measure for predicting crops which is suitable under the soil condition and weather scenario. The users will be suggested with the best crop suited by the weather of the place and soil attributes. The system uses open street map API in order to get weather attributes of different places within Nepal. These obtained weather attributes along with the soil attributes are fed to machine learning algorithm in order to compute the most suitable crop for the selected region.

This project classifies the weather conditions and maps the crops suitable under those classified weather condition. Then, soil attributes are taken as input and prediction model which is trained to predict suitable crops for that soil type and weather predicts the crops for that data. This enables the farmers to minimize the loss of crops due to adverse weather conditions and geographical factors too. This application focuses on analyzing the past trends of weather of a particular place to predict the future weather and assist farmers in choosing the right crop with the maximum productivity.

## Problem Statement

It is being difficult to harvest suitable crops in right time due to the recent emerging problems like climate change and global warming. And as soil condition also keeps changing, farmers usually don’t know about that and keeps planting same crops in same field for years which also reduce the soil fertility and crops production. So, it’ll be beneficial in every aspect to predict crops for coming weather based on soil types.

The climate of world has changed a lot due to several factors. Therefore, the yield of crops isn’t found to be as expected. Weather has adverse effect on the yield of crops. Given the weather condition and type of soil, there will be at least one most profitable crop. The farmers can be assisted by suggesting the best crop for a particular area along with the forecast of next x days.

## Objectives

The crucial part of this project is to solve underlying problems in agriculture sector and aid farmers. To overcome the problems faced by farmers in present days, certain objectives were drawn before the start of the project.

Major objectives of this project are:

* To predict upcoming weather for crop prediction.
* To classify weather and find suitable crops based on weather and soil analysis.

## Scopes and Limitations

### Scopes

This project will provide a reliable and accurate model for predicting suitable crops based on weather and soil type. So, this’ll partially solve the problem of crop losses due to climate change.

### Limitations

Despite of its significant benefits and scopes, it has some limitations which are:

* There exist some exceptional cases where crops predicted by this system may not be suitable for that condition as other factors may also affects the crops.
* As this system only analyze the past weather data, this can’t accurately predict upcoming weather.

# Chapter 2: REQUIREMENT ANALYSIS AND FEASIBILITY STUDY



## Literature Review

This review article written by Dr. Bharath Mishra focuses mostly on various attributes to be taken in consideration while applying data mining techniques in the field of agriculture and also concludes that the multidisciplinary approach of integrating computer science with agriculture will help in forecasting/ managing agricultural crops effectively [2].

Sally Jo Cunningham emphasizes on the usage of data mining techniques and its process model to derive innovative applications in the field of agriculture. He also visualizes the applications of data mining, the goal might be to use a model predictively, to provide automated classification of new instances. He finally concludes that the usage of Weka tool and its comprehensive suite of facilities for applying data mining techniques to large data sets [3].

Jayanta *Basak*, the author of this paper discusses about the central problem in weather and climate modelling, which is to predict the future states of the atmospheric system. Also, the atmospheric correlation is being discussed in this paper. The author has provided techniques for determining the strongest independent components in the multidimensional data set [4].

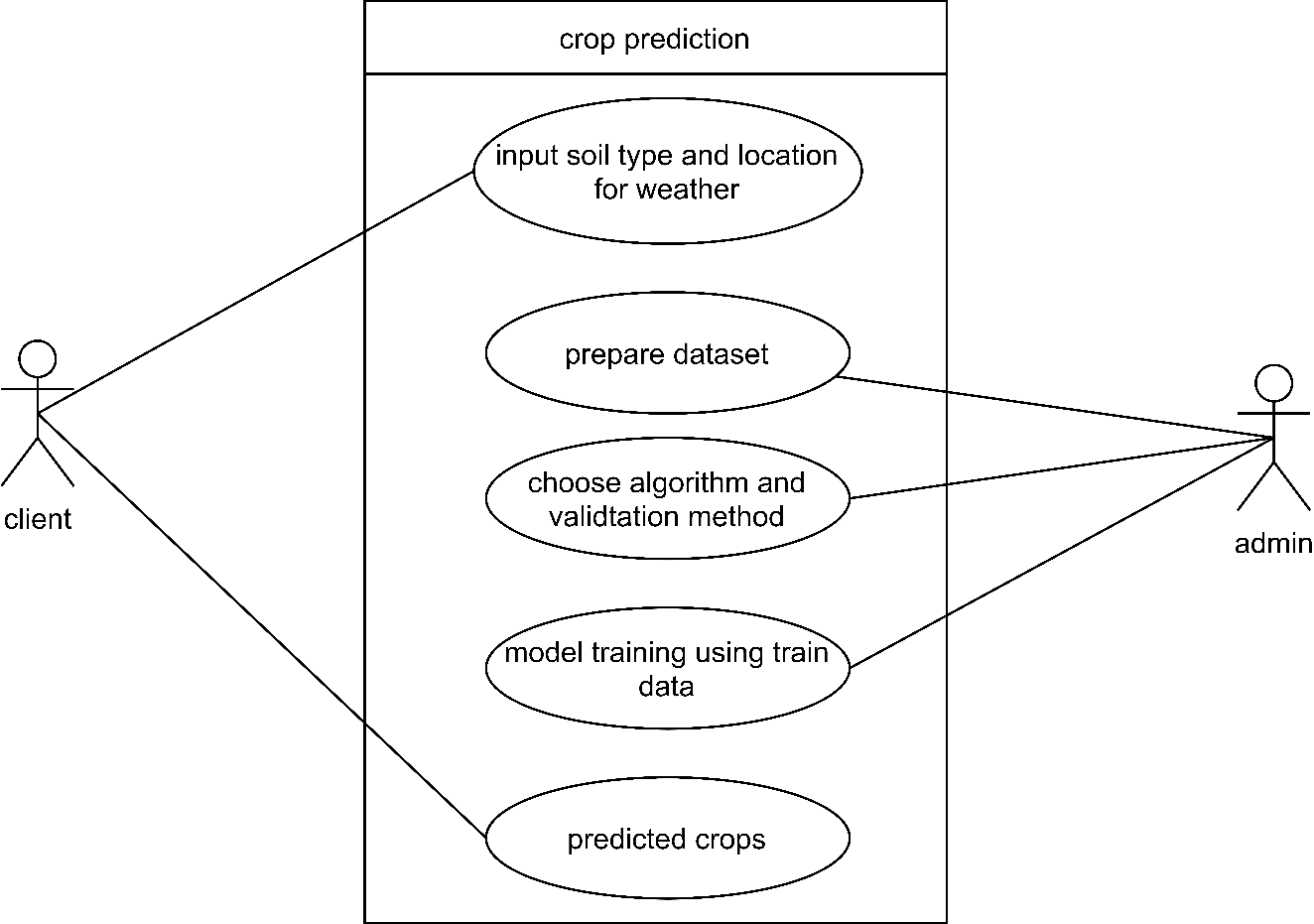
## Requirement Analysis

In software development process, it is must to list out the requirements in any system and those requirements must be analyzed properly.

### Functional Requirement

A functional requirement describes what a system should do when it is given input. Some of the functional requirements are given below:

* User should enter soil types and location to get predicted crops.
* System should classify weather and also classify crops suitable on those weather.
* System should predict upcoming weather and predict suitable crops for that weather and provided soil types.



**Figure 6.1:** use case diagram for proposed work

### Non-functional Requirements

## Feasibility Analysis

This part of system analysis is carried to confirm that whether the system being developed is actually feasible or not. This is the phase where any system designers are able to know whether to start the project or not.

### Economic Feasibility

This system can be developed on low end computers and can be deployed on server with lower specs as only classification and prediction algorithms are used and won’t work on neural networks or deep learning so no any high-end computers with higher GPU’s, TPU’s and huge computation power is required. Thus, this system is economically feasible. Also, client doesn’t need to make any sort of payments in order to find the suitable crops using this system. Thus, the system is economically feasible.

### Operational Feasibility

This system provides a simple web interface, which can be easily used by any type of users having basic idea of interacting with websites. This system will tend to provide accurate results according to the way the system needs to do. The system is operable in any browser in any operating system. Hence, this system is operationally feasible.

### Technical Feasibility

Though this system will have predication and classification algorithms, this system will be built using simple algorithms so, there won’t be any technical complexities. Also, from the user’s perspective, user with basic computing knowledge can use this system. User can simply fill a small form with required information and find the most suitable crops under the required conditions.

### Schedule Feasibility

As we have created a proper schedule for this system, we can say that this system is feasible and can be completed on time

# Implementation Tools

Till this stage, development of planned system will be done so it could be brought into real world.

Different tools and technologies that may be used are:

**Python:** Python 3.8 will be used for machine learning model creation. KNN algorithm is implemented for classification of weather types using python.

**Flask:** Used for creating API. API will be consumed later in the web application.

**.Net Core:** .NET core 3.0 will be used for creating web Interface. The web interface includes interactive forms and pages.

**MS SQL:** Used for storing datasets and other data required for the weather analysis. It uses TSQL as SQL query engine. The version that will be used is MS SQL server 15.

# Implementation Algorithm

Different algorithms and papers were studied and KNN algorithm was found to be most suitable for crops prediction using soil and weather analysis.



## K-Nearest Neighbors

KNN is a type of instance-based learning or the lazy learning where the function is only approximated locally and all computation is deferred until classification. In this particular project, the KNN algorithm is specifically used in order to classify the weather data into different classifications. This is also responsible for recommending the right crops as per the provided weather conditions and soil types.

The case is classified by a majority of votes of its neighbor and the case is assigned to the class common amongst its K nearest neighbors measured by a distance function. If k=1, then the case is simply assigned to the class of its nearest neighbor. In this case we use Euclidian distance as a distance function.

**Euclidean Distance:** Euclidian distance between two points measures the length of segments connecting the two points. It can be given as:

Where,

Xi and Yi are the end points of the line segments,

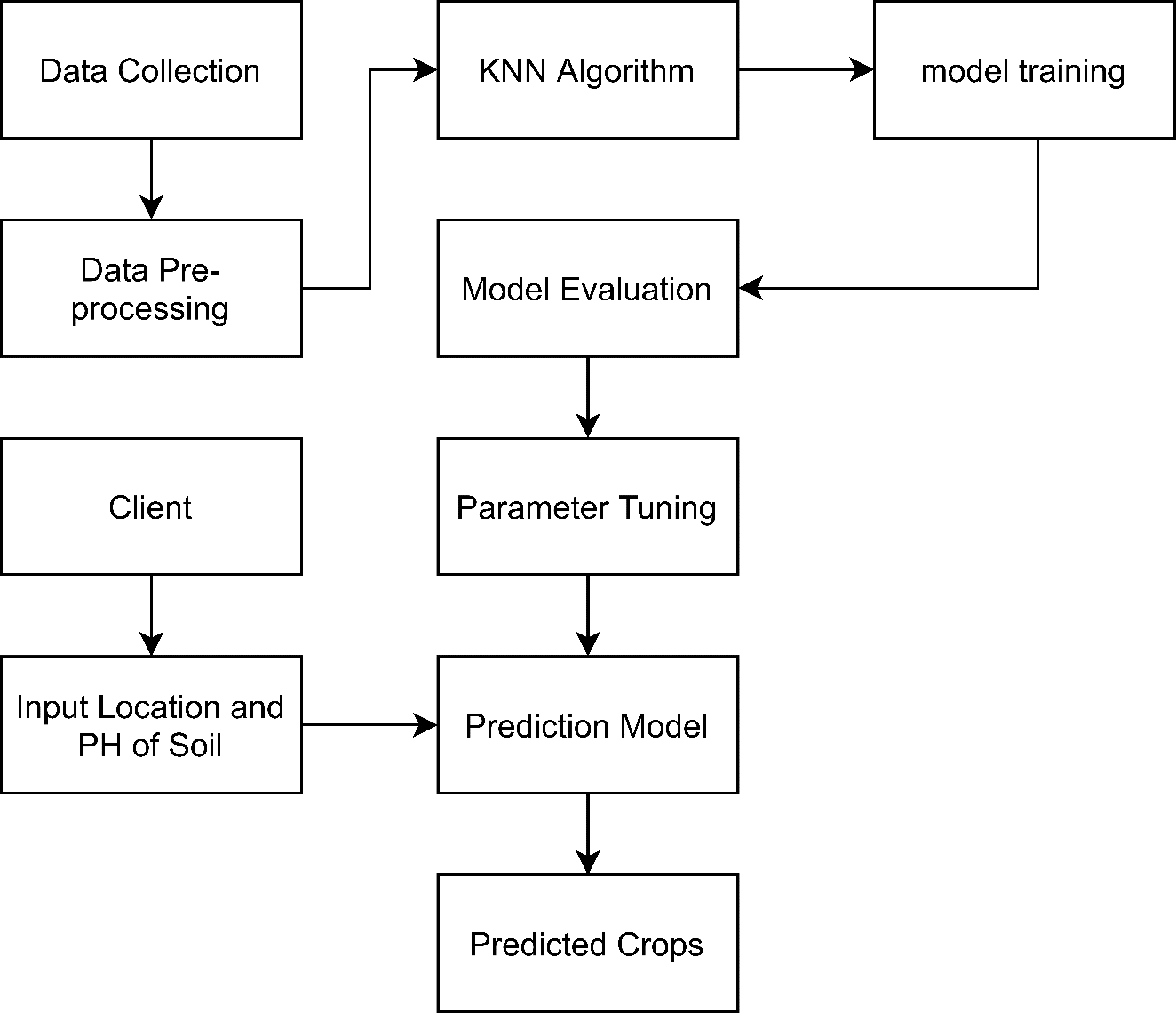
K is the limit of iterative summation

Once the Euclidian distance is evaluated, the training set can be used to classify the weather. The major steps involved in KNN algorithm are,

1. Load the data
2. Choose and initialize appropriate value of K
3. For getting the classification, iterate from 1 to total number of training data
4. Calculate distance between test data and each row of training data using Euclidian distance.
5. Sort the calculated distances in ascending order based on distance values
6. Get top K rows from the sorted array
7. Get the most frequent class of these rows
8. Return the predicted class

To select appropriate value of K for the available data, the KNN algorithm is run several times with different value of K. The value of K is chosen in such a way that it reduces the number of errors encountered while maintaining the algorithm’s ability to accurately make predictions when it is given data it hasn’t seen before.

# Proposed Work



**Figure 10.1:** Process diagram for proposed work

# Expected Output

After the completion of this system, it can be expected that this system provides a platform for crops prediction given that the weather and soil types of a region. It may not provide 100% accurate and reliable result. However as per the researches and available resources, high accuracy (almost 80%) is expected within the end of the project.

# Working schedule



**Figure 12.1:** Gantt chart for proposed work

# References

[1] Fao.org. (2020). Nepal at a glance | FAO in Nepal | Food and Agriculture Organization of the United Nations. [online] Available at: <http://www.fao.org/nepal/fao-in-nepal/nepal> at-a-glance/en/?fbclid=IwAR0AzcFwLprolGMZ6h2ioQ6DVT2N3VXjMbrU5OHZz3O1ZbMvUWb0RjO6QkM [Accessed 10 Feb. 2020].

[2] Mishra, D. and Singh, D. “Data mining Techniques for Predicting Crop Productivity – A review article”. [online] Ijcst.com. Available at: http://www.ijcst.com/vol21/veenadhari.pdf [Accessed 12 Feb. 2020].

[3] Cunningham, S. and Holmes, G. (1999). Developing innovative applications in agriculture using data mining. [online] cs.waikato.ac.nz. Available at: <https://www.cs.waikato.ac.nz/~ml/publications/1999/99SJC-GH-Innovative-apps.pdf> [Accessed 12 Feb. 2020].

[4] Jayanta Basak, Anant Sudarshan, Deepak Trivedi, MS Santhanam, “Weather data mining using independent component analysis”, Journal of Machine Learning Research, Volume 5, 2004