Smart Agriculture Solutions using Machine Learning on Sensor Network Data

Introduction

Agriculture contributes a large percentage towards the global economy and social well being. The needs of an ever expanding population has put an immense pressure on the traditional agricultural system. This pushed us towards the need to incorporate technology in agriculture for enhancing traditional farming techniques. With emergence of the Internet of Things(IOT) which consists of a network of sensors that can gather huge amounts of data (e.g.: light, humidity, temperature, soil moisture, etc). The data gathered can be analysed for suggesting smart farming techniques. Machine Learning algorithms can be developed using the data for more accurate and faster decision making in various phases of farming such as irrigation, soil quality management, crop yield prediction, crop protection etc.
In this project, we aim to exploit several machine learning techniques to aid in some stages of farming. For irrigation, we aim to predict irrigation patterns based on crops and weather scenarios. In Soil Quality Management, we aim to classify soil and tell its quality based on various soil sensor data. The Soil Quality and Climate parameters will further be used for crop yield prediction

Objectives of the Project

The objective of our project is to support smart farming by suggesting simple and innovative solutions for various phases in agriculture. We aim to exploit Machine Learning Techniques which will cover the following aspects of agriculture in this project:

- There is a lot of wastage of water and electricity for irrigation even when using techniques like sprinkler and drip irrigation. Also, the quality of water used for irrigation can severely affect the crop yield and soil quality. Underwater sensors data such as pH, TDS, Electrical Conductivity will be used to determine the quality of water and data such as soil moisture and humidity can be used for predicting efficient irrigation patterns.
- Soil is the most important factor in agriculture. The quality, fertility, type of soil determines which crop is suitable and how much yield can be expected from it. Analysis Soil Sensor data such as pH, moisture, temperature will be used for predicting soil Quality.
- Climatic conditions such as temperature, humidity, light and soil Quality affects the yield of crops and the type of crop that can be grown in an area. Data from Climate sensors and soil sensors will be used to predict the yield for crops and which crop is suitable for a given area.
- Some data gathered from various sensors can have abnormal and missing values. Various prediction techniques such as regression will be used to predict the missing data and methods like localization of neighbourhood sensors' data can be used to correct the abnormal sensor. We will make our model robust so that it doesn't collapse in the time of need and is able to generate necessary warning messages.

Feasibility study

Agriculture is one of the major sectors of Indian economy in which above 40% of people are engaged. The resources used in farming such as water, fertilizers, electricity etc go to waste. To save these resources, we intend to collect and analyze data from various climate, underground and aquatic sensors. This needs the sensor network to be accurate ,durable and fault tolerant in a farming environment. Our system also needs to adapt to the huge amount of data which will be generated from these sensors. We intend to use Machine learning methods on the sensor data to gain insights and suggest possible actions to the user. Simple

algorithms wont work. We need to exploit various state of the art methods such as Deep Learning to improve the performance of our solution.

Methodology

Gathering Sensor Data

We will gather data from three types of sensors: Climate, Underwater and soil. The following are the parameters that will be measured from various sensors.
Climate: Temperature, Humidity
Underwater: pH, Electrical Conductivity, TDS
Soil: pH, Moisture, Temperature

Prediction of missing data from monitoring sensors

Various combinations of different attributes(considering the missing data attribute) will be considered to retrieve their respective contributions.Two methods will be used to further predict the missing data:

a) Artificial neural network followed by Regression(Linear or Polynomial) model
b) Decision tree alongwith K-Nearest Neighbours and polynomial Regression model

Each method will be trained to reduce the prediction loss of the missing data and the model giving the best result will be chosen for a given scenario.

Detection and Correction of Abnormal Data

In case of multiple sensors detecting the same parameter(e.g. Temperature), the sensor measuring abnormal value can be identified by comparing similarity of each sensor's value measured within a geographical cluster(formed with help of latitude and longitude values of sensors). The sensors having a significant difference will be selected based on a criterion.

Water Quality Prediction

For proper irrigation of a farm, we require complete knowledge of water quality to predict the nutrition received by the soil from that water and if its pH and TDS level is optimum for the soil or not. We use K-means for this purpose and the results generated informs the user what changes need to be made to improve water quality for his farm.

Irrigation Pattern Prediction

This pattern can be predicted based on soil moisture data and weather data collected by the sensor which will be used to calculate actual relative humidity. We will process this data using SVR (Support vector regression) and random forest regression.

Soil: Soil Quality, Fertility, Type

We will use the soil data gathered from sensors as the input parameters to the model to detect soil type and Quality. Training an ANN(Artificial Neural Network) on the data will help in predicting the soil quality and type.

Effect of Climate and Soil on crop yield

Climate and Soil data such as temperature, Humidity, soil pH, moisture etc will be used for this. Various combinations of the parameters will be used to train an Artificial neural network that will be used to predict

and improve the crop yield. Appropriate warnings will be generated using decision trees in case there is a drastic change and actions can be taken accordingly.

Facilities required

Software:
Cloud Computing Platform such as google colab
Python v3
Libraries: torch, TensorFlow, pandas, NumPy, OpenCV, scikit-learn, networkx, matplotlib
Datasets: Data from Climate, Underwater and soil sensors
Climate: Temperature, Humidity
Underwater: pH, Electrical Conductivity, TDS
Soil: pH, Moisture, Temperature

Bibliography

- Artificial Neural Networks for Soil Quality and Crop Yield Prediction using Machine Learning
- IOT FOR AUTOMATION OF FARM IRRIGATION
- Precision Agriculture Design Method Using a Distributed Computing Architecture on Internet of Things Context
- Microsoft FarmBeats

By:
Adarsh Jha (2K17/CO/22)
Aditya Choudhary (2K17/CO/24)

Mentor's Signature

Dr Rajesh Kumar Yadav