ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 12, December 2017

Big Data Analytical tools for Agriculture

[1] A.Muthu Rathi, [2] B.Jeni Therese, [3]Murugeswari
[1][2] III B.Sc., [2] I B.Sc.,
[1][2][3] Department of Computer Science, Holy Cross Home Science College,Thoothukudi

Abstract:-- **The traditional methods used by the farmers in India, are very slow, undependable and large amount of crops are damaged in fields due to bacterial attacks and lack of information resources. Annually, such loss exceeds 40% in total. The main goal is to provide better solution for the finding problems and to enhance the productivity of the agriculture sector. The main awareness of this work is focused on Indian farmers as it addresses the key problems of getting the market status of different products, weather alerts and also provides multiple language support. This will effectively help farmers to sell their products in global market and earn remarkable profit. There is scope to maintain the information of all these and analyse properly and communicate with farmers. This kind of analysis can be done with latest technologies such as big data analytical tools. A real monitor system needed to communicate with farmers, time to time with support of mobile based application.**

Keywords:-- **Big Data, Agriculture, Data Management, Data Model, Data Accuracy.**

1. INTRODUCTION

Big data in agriculture and its analysis is the next wave in technology to support all aspect of our lives. In agriculture, data management has long been recognised as is an opportunity to drive farm efficiency. It is only through the development new tools in data capture and data sharing that such opportunities can be realised. New technology development in agriculture has shifted from sensors and machines to methods in big data analysis. Sensors on fields and crops are starting to provide literally granular data points on soil conditions, as well as detailed info on wind, fertilizer requirements, water availability and pest infestations. GPS units on tractors, combines and trucks can help determine optimal usage of heavy equipment. Data analytics can help prevent spoilage by moving products faster and more efficiently. Unmanned aerial vehicles, or drones, can patrol fields and alert farmers to crop ripeness or potential problems. RFID-based traceability systems can provide a constant data stream on farm products as they move through the supply chain, from the farm to the compost or recycle bin. Individual plants can be monitored for nutrients and growth rates. Analytics looking forward and back assist in determining the best crops to plant, considering both sustainability and profitability.

II. BIG DATA ANALYTICS FOR AGRICULTURE

Big data analytics has made such a widespread impact in the agriculture industry that it's hard to pinpoint all its effects, and harder still to predict what changes it might bring. So, to clear things up a little, here's a list of four ways big data technology is disrupting the agriculture industry.

1. Accurate crop predictions

Waiting to see how the crops turn out has been driving men to drink since…well, since we've had alcohol to drink. We've tried everything from the rain dance to solemn prayer, but crop predictions throughout the years have been anything but accurate, until now.

By using sophisticated computer algorithms to analyse decades and sometime centuries of weather and crop data, today's farmers can predict crop yields with shocking accuracy, before planting a single seed. The insight provided by data analytics allows farmers to start and harvest their crops at the optimum time, which maximizes crop yields and minimizes stress.

2. Stronger seeds and less hunger

An increasing global population combined with rising temperatures has led to a massive famine in Africa which has left 20 million people at risk of starvation. Humanitarian groups across the globe have geared up to offer assistance however they can, but the solution may lie in big data.
Chemists and agricultural scientists have been analysing plant data for years in the hopes of developing crops that can grow in any environment. We can grow plants faster, taller, and heartier than ever before. Chemically engineered seeds designed using big data may sound like a bad thing on the surface and the news usually portrays it that way. However, seeds created using data analytics could put an end to world hunger.

3. Automated agriculture

Automated farming or precision agriculture is nothing new. For decades we've been using different systems to automate and keep track of as many agricultural processes as possible.

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 12, December 2017

In many ways, data analytics was what separated commercial farmers from the pack long before big data was even a term. Farmers are using drones with advanced sensors to survey their crops, update their data, and notify them of areas that need improvement. As the technology continues to progress you can expect drones to move from surveying to planting and harvesting themselves.

Data analytics is all about finding the minute flaws in a system and correcting them, unfortunately, the human element is all too often that tiny flaw. Big data has us well on our way to farmer less farms.

4. Environmental awareness

Big Data is the reason we can say with certainty that humans have had a negative impact on the environment, and it's the tool we're using to fix it.

Saving the environment sounds great from a human perspective, but the agriculture industry is just that, an industry, a business, and all business decisions live and die by the bottom line. Big data is showing companies in the agriculture industry that not only can you protect the environment without increasing costs, you can reduce them. It's not just the manufacturing industry that's making changes to reduce their environmental impact, farmers and agriculture companies have been leading the way.

III. FROM TRADITIONAL AGRICULTURE TO FARMING WITH BIG DATA

A properly designed big data solution would provide answers to a vast amount of interesting enquiries using historical data and more importantly it would lay the ground for predicting future trends.

The Data Model

Raw data from farm management systems is extracted and fitted into a data domain model that has been meticulously designed to facilitate its manipulation and analysis. This process is called ETL: Extract, Transform, and Load. The data model contains different entities, which individually encapsulate a farming concept (products, fields, farms, planting zones, crop types, field operations, etc.) and collectively integrate to describe agricultural activities. The model revolves around field operations, such as drilling (planting), crop protection (pesticides application), nutrition (fertiliser application), and crop yield (harvest operations).

Enrichments

Inbound data is not only loaded, it is also enriched in various ways, for example:

- Data canonicalization: Farm management systems often allow the farmer to type free text, which generally creates multiple representations of the same entity. For example, 'Abc,' 'ABC +,' and 'ABC, 10' refer to the same pesticide. All these entries will be rolled up into a single canonical form by using pattern matching with regular expressions.
- Component breakdown: Fertilisers' nutrients can be inferred from the name provided by the farmer. For example, '0-0-26-6' can be broken down into Nitrogen (N) 0%, P2O5 (Phosphorus pent oxide) 0%, K2O (Potassium oxide) 26%, SO3 (Sulphur trioxide) 6%.
- Organic soil composition: Based on a particular field's geographical location, soil properties can be obtained from external sources.
- Weather conditions: Similarly, past weather conditions as well as forecasts for a particular location are obtained from an external Application Programming Interface (API).

Technology Stack

Processing and Storage

HPCC

HPCC (High-Performance Computer Cluster), a data-intensive computing platform, has been chosen as the big data platform for this solution. This technology is used to obtain, process, analyse, and massage the data in order to build the desired data model. HPCC has its own programming language called ECL (Enterprise Control Language) which was designed specifically for big data endeavours. ECL is used within THOR (The Data Refinery Cluster) - HPCC's processing component and distributed file system (and a very powerful one). THOR does the heavy lifting of big data very efficiently.

Querying and Searching

Searching and querying technical requirements are driven by emerging use cases as new customers are interested in consuming the data stored in HPCC. Amongst actual and prospective customers, there are product suppliers, growers, cooperatives, regulation authorities, distributors, media outlets, etc., each with their own search needs, which need to be analysed on a case by case basis. However, there are two general strategies currently implemented:

1. Elasticsearch

In order to search, browse, and slice and dice the data, Elasticsearch comes into play. Subsets of the data lake are incrementally pushed from HPCC into Elasticsearch indices

which can vary in content and structure, depending on the search needs of each case. Scalability, speed, stability, and the Elastic Stack (Logstash, Kibana, X-Pack, etc.) make Elasticsearch a perfect fit for the solution architecture, providing search power and facilitating data analytics exercises.

2. ROXIE (Rapid Online XML Inquiry Engine)

ROXIE is HPCC's own data delivery engine. It is used to expose structured and punctual results. Although designing and implementing the queries can be time-consuming, it serves results very quickly.

Architecture Overview

IV. CONCLUSION

The importance of data management is recognised and many growers are providing their farm data to service providers so that the data forms part of a central databases. The hope is that such databases become available as an open source of data to help grow innovations in big data analysis and management. However, it is clear from the recent development in patents covering data management that companies have recognised the importance securing the novel techniques for the use of farm data. If you are a developer of new software tools on big data analysis, it will be important to assess the methods you use against existing patents. Thus we have a tendency to conclude that, there's a growing range of applications and the role of Big data analytics techniques in agriculture and a growing quantity of information that area unit presently obtainable from several resources. This is often comparatively a completely unique analysis field and it's expected to grow within the future.

V.REFERENCES

[1] https://en.wikipedia.org/wiki/Agriculture_in_India

[2] https://en.wikipedia.org/wiki/Agriculture

[3] https://en.wikipedia.org/wiki/Crop

[4] https://en.wikipedia.org/wiki/Kharif_crop

[5] https://en.wikipedia.org/wiki/Rabi_crop

[6] https://en.wikipedia.org/wiki/Zaid_crops

[7] Rupika Yadav, Jhalak Rathod,Vaishnavi Nair, "Big Data Meets Small Sensors in Precision Agriculture", International Journal of Computer Applications (0975 – 8887) Applications of Computers and Electronics for the Welfare of Rural Masses (ACEWRM) 2015.

[8] Min Chen, Shiwen Mao, Yunhao Liu, "Big Data: A Survey", Springer Science Business Media New York 2014.

[9] Rohit Pitre, Vijay Kolekar, "A Survey Paper on Data Mining With Big Data", International Journal of Innovative Research in Advanced Engineering (IJIRAE), Volume 1 Issue 1 (April 2014).