Big Data Analysis

Practical 1

Objective :

Study and explore various applications of big data in different domains. Choose one of it and study in detail, Also write down the report on different types of digital data generated in selected applications.

Roll No. & Name :

18bce173 Utpal Patel

18bce183 Prince Prajapati

Submitted to: Prof. Jaiprakash Verma

Task 1:

Introduction

BIG DATA IN AGRICULTURE :

Big data can be used in agriculture for

- increasing the yield from existing farmlands
- predicting what crops to plant for better profitability
- regulation of use of pesticides
- optimizing farm equipments
- managing supply chain issues

- providing agricultural information and market intelligence
- providing weather information
- tracking progress of crop
- reducing wastage
- reducing cost of production
- ensuring global food security measures
- analyzing plant genome information
- government policy making for supply chain management
- useful in field zoning to easily identify most productive areas
- helps in analyzing vehicle's route, engine speed, amount of fuel
- controlling driverless vehicles for farming
- smart greenhouses

Volume,Velocity and Variety of Data

Volume:

For agriculture, collected data to be used when we have a large volume of data available for analytics.

A field that produces 185 bushels of corn is a point, but knowing thousands of points between low yields and high yields gives us more information to understand that report card and how we can influence it throughout the season.

A single weather station on a farm can send 34 MB of data (10 mp3 songs) per year, while a 2,500-acre plant can set a 5GB record during the planting season.

As we understand the size of this data on multiple passes, IoT and FMS to understand it - the volume and size for farms alone becomes too large.

For large amounts of data, this creates an opportunity for further understanding, but it remains more complex.

At the forefront of precision agriculture, yield maps gave an understanding of what changed across regions, and today we can actually measure and manage dozens of other conditions that affect that map, much more so. form, which can help us make connections, causes, and identities. influencers but we must have ways of sorting, filtering, analyzing and accessing this amount of information for it to be useful.

Velocity:

The timing of data we receive is another key part. During the whole year we can collect data on the progress of crops. So we continuously capture the data.Now a days images are used as data. So the velocity has increased which can give us access to quicker decisions, but also can make for a larger volume and more complexity. With the use of some modern tools we can increase the velocity of data, so that we can get information faster and make decisions on time. There are many government apps available which are taking care of farmers' data and collecting information about various tools that farmers are using and various things they are using to protect crops.

Variety:

Data collected for big data analytics :

Traditional tools are being replaced by sensor-equipped machines that can collect data from their environments to control their behavior – such as thermostats for temperature regulation or algorithms for implementing crop protection strategies. Technology, combined with external big data sources like weather data, market data, or standards with other farms, is contributing to the rapid development of smart farming.

1) **Structured Data :-**

- rainfall patterns
- fertilizer requirements
- historical weather information

- seeds and fertilizers information
- market prices
- plant nutrient requirements
- finance and loan
- diseases and pests
- government schemes and subsidies
- irrigation schedule and its frequency
- risk assessment data
- Government historical databases about food production and agriculture
- Tractors, combines, seeders and other agricultural vehicles
- soil types
- Transport, cold storage availability and location
- weather disasters management

2) **Unstructured Data :-**

- satellite and drone images
- drone and aerial photography
- air temperature and humidity sensors
- water evaporation data
- wind direction and strength data
- Geo-spatial data
- remote sensing data
- daily growth of plants and crop health
- Harvesting related data

- spraying related data
- variability in soil
- soil salinity data
- soil acidity
- soil moisture and nutrients
- water cycles
- sunlight intensity

3) **Semi-structured Data :-**

- chemical composition of soil
- Genotype information reciprocal data

The quality of the data -Veracity

For data to provide us answers and value, farmers want to analyze data that gives the real answers, not fancy maps or information that is incorrect that I make next year's decisions on that faulty information, creating that cascading snowball. This is a critical crossroads for us in agriculture, and having more automation and transparency in the processes from seed to harvest will help manage quality and veracity into the future.

The veracity of data, which is also known as its quality, depends on both sensors and human error. For example, the veracity of yield data involves both whether the combined yield monitor is properly calibrated and whether planted data has hybrids correctly tagged to fields. Both sensors and human error influence the quality of data.

Actionable Insight -Value

Big data provides farmers **granular data on rainfall patterns, water cycles, fertilizer requirements, and more**. This enables them to make smart decisions, such as what crops to plant for better profitability and when to harvest. The right decisions ultimately improve farm yields.

Task 2:

List Of Database Queries:

Here different types of Queries are possible based on requirement.

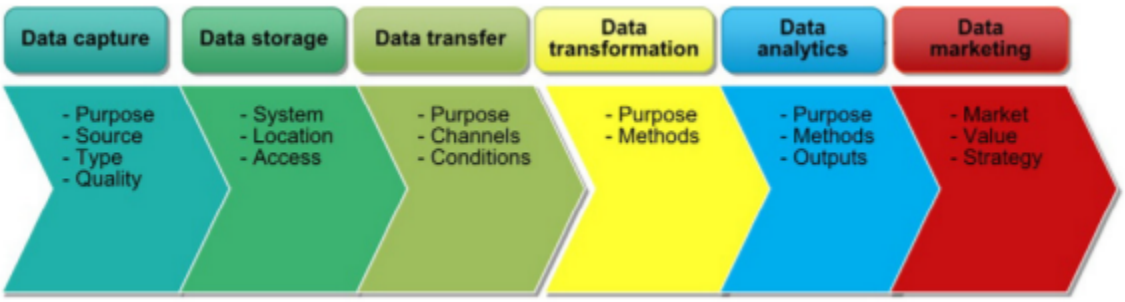
Like,

- Current weather or pricing information .
- Past crop pricing data
- Current selling price of various crops
- Tools and technology used widely
- Electricity timings
- Past drone video recordings
- Current new technologies in market

List of Data Mining and warehousing Queries:

- K-nearest neighbour to weather and forecasts
- K-nearest neighbour to check sound recognition
- K-means to evaluate soil fertility
- K-nearest neighbour for Classification of forest data with remotely sensed images
- Predicting metabolizable energy of poultry feed using group method of data handling-type neural network
- Prediction of problematic wine fermentations
- Relationship between sprays and fruit defects
- Optimizing pesticide use by data mining
- Explaining pesticide abuse

List of Big Data analytical queries:



Task 3:

Methodology for Smart farming using Big Data:

In the first phase, IoT devices collect the data. Sensors plugged in trucks and tractors and fields, plants, and soil aid the real-time data in the collection directly from the ground.

In the second phase, data analysts integrate these large amounts of data collected and other information available in the cloud, which includes weather data and pricing models to determine patterns.

Finally, based on the revealed patterns and insights, it helps in controlling the problem. With big data analytics, we can pinpoint existing issues, like soil quality, operational inefficiencies, etc., and formulate predictive algorithms. These are finally implemented as an alert to prevent future problems.

AgriTech companies working on Big Data in India:

- SMAG
- FASAL
- DeHaat
- Clover
- CropIn
- Intello labs

Big Data Analytics Tools and Technology :

- **Satellites :** Using satellites we can capture the images
- **Sensors :** Weather stations can measure a whole range of features: air temperature and humidity, wind direction and strength, precipitation rate, soil density and acidity
- **Agricultural vehicles :** Tractors, combines, and seeders equipped with onboard computers are a valuable source of information.
- **Drones and Aerial photography :** Drone photography can be utilized to view areas captured from satellite imagery more clearly.

- **Laboratory Analysis :** Agrochemical analysis of soil is one of the traditional estimation methods to determine necessary fertilizer rates.
- **Multipurpose government apps :** We can collect the information from various parts and then use it to analyze.

Conclusion: After implementing this practical we have understood how big data is useful in agriculture and application areas and how five v's are important to understand big data for specific fields.