USING BIG DATA FOR RAPID DEVELOPMENT OF IOT TECHNOLOGY

T.Sujilatha[1], B.Rajani[2], V.Mamatha[3]

Assistant Professor[1],Assistant Professor[2],,Assistant Professor[3]

Computer Science and Engineering

NBKRIST, Vidyanagar, Kota mandal, Tirupati Dist-524413,Andhra Pradesh, India

Abstract: Now a days usage of Big data has become more effectiveness with the growth of hardware devices forupcoming technology is more helpful to work on for usage and provide more client service for developing and customized marketing business intelligentpurpose and take other features to explore profit and gains for different organizations. In thepresent real world of fast growing technologyfeatures of updates are always changing and formerly popular in demands come outdated. This happens because of excellence of big data technology along growth of IoT technologies. With the support of colorful IoT tools, including tackle and software platforms, network analyzers, and IoT-specific platforms, empower inventors to make, connect, dissect, and cover IoT results efficiently. These tools accelerate development, insure data security, and optimize IoT operation performance.

Index Terms: Big data Technologies, Business Analysis, Machine Learning(ML) ways, IoT Tools, IOT-Hadoop and spark- technologies.

I. INTRODUCTION: Big Datais a collection of information that's large in quantity, rapidly increasing in exponentially with time interval. Data with huge size are difficult for processing among of conventional information that can be operated with help of tools store it or reuse it efficientlyBig data is also a data but with huge size.

II. What Are Big Data Technologies?

It is a term that refers to the massively volume of data or information that associations induce day to day history, this information was too large and complex for traditional data processing tools to handle. still, advances in technology have now made it possible to store, process, and disc There are a variety of data processing technologies available, including Apache Hadoop, Apache technologies has its own strengths and sins, but all of them can be used to gain perceptivity from large dataSpark, andbig data snappily and effectively MongoDB.sets. As associations continue to induce further and further these technologies will come decreasingly essential Big data storehouse technologies is a cipher and storehouse armature that collects and manages large data sets while also allowing real-time data analytics.
 Let's have awareness about these technologies available for big data world.

a. Types of Big Data Technologies: The term" big data" refers to the growing volume of much data that associations are floundering to manage effectively. Whilethe conception of big data isn't new, the technology geography is constantly evolving, making it delicate to keep up withthe rearmost trends. Big data technology results help with this problem. Let's us see how information can be explore with the technologies for managing andassaying big data. Here with a brief overview of some of the most popular big data technologies.

b. what Hadoop big data technology is?
Hadoop is an open- source frame that enables the information that is distributed for processing of large data

sets across clusters for the segregationof commodity waiters. It provides a system(HDFS) that's designed for versatility and trust ability resource director(YARN) which enables effective scheduling and processing of job prosecution.Spark is the fastest and general purpose clustering computing system which provides an interactive shell which takes commands and work on it for ad-adhoc data analysis as APIs for programming in Java which can be used more verbose than Python or Scala. Spark supportsSQL queries andmachinelearning(ML) algorithms. No-SQL databases are designed and developed for dilatability and inflexibility designing them with a well- suited forstoring and retrieving of data from the database.The most useful and popular No-SQL database systems include MongoDB(MDB), Cassandra-C, and H-Base Data storagesvolume of data as atraditional relational database operation systems(RDBMS) which has an advanced feature with new implemented architectural design changes with added operations and functionalities to follow big data analytics process. The two very popular data storehouse systems which are used as a feature system was Teradata and OracleExtradata. Big data technologies can be segregated into four categories as i. Batch process system, Stream process. NoSQL databases system, and iv. Data storage system.Each process has its own features for important to select the right tool for the processing ofjob at hand. In general, Hadoop and Spark are good selection for batch process mean while Kafka and Storm are more suitablefor processingof streaming operationsystems. NoSQL databases is similar as MongoDB and Cassandra are very good choices for processing of data. Scalability is veryimportant to deals or usage while data storages similar as Teradata or Oracle Extra data which are more suitable for process of operationswhich bear complex queries or analytics.

III. **Factors of Big Data Technology:** Big Data technology has mainly divided into four factors Dataprisoner, Data storehouse, Data processing, and Data visualization.
1. **Data prisoner:**It refers to the process of collection of datafrom a variety of sources which includes from social media posts, data files etc., to detect or reading information.
2. Data storehouse: It isthe process of storing the data in a way that makes it accessible for further analysis.
3.Data processing: It is where thealgorithms are used to sort out the data and process for excerptperceptivity.
 4. Data visualization: It is the process of representation of data in a way which caneasily understood for human.

IV. Fields of Top Big Data Technologies: There are four main fields of big data technology Prophetic Analytics (PA), MachineLiteracy (ML), Natural Language Processing (NLP), and Computer Vision (CV).
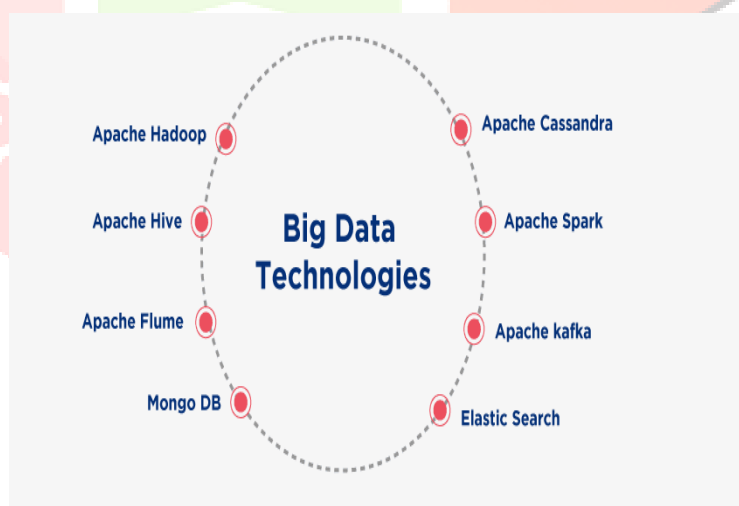

Fig: Big Data Technologies

1. Prophetic Analytics: It is mainly used to identify patterns and trends in data in order to make tools effectively for business process.
2. Machine literacy: It is a type of artificial intelligence that relies uses the application of algorithms and mathematical models for better improvement of performance on specific task.
3. Natural language processing: It is used to dissect textual data in order to analyses and synthesis of speech based on input data.
4. Computer vision: It is area of artificial intelligence which deals with the interpretation of digital images
Mainly four fields are used in van of big data technology and analysis that are essential for understanding and managing large datasets over the internet.

V. Big Data Technologies (BDT): As data or information becomes increasingly and become complexity in our daily lives, the need to effectively collecting, storing, and dissect has noway been lesser. BDT isconstantly evolving to overcome these challenges, and the geography wills more complex in the times to practice. Among them some are the top big data technologies that should be apprehensive of in 2023. Let us check the big data technologies list.

 1. Apache Hadoop: It is one of the very popular big data technologies in 2023. Hadoop is an open- source frame work process that enables the distributed process of large amount of data sets across a cluster of commodity for use many organizations adopt these technologies due to its scalability, flexibility, and cost-effectiveness performances.

2. Apache Spark: It is an open source which is effort for data process for a computerthat can perform batch or real time analytics on large data sets over the internet. It's frequently used in confluence with Hadoop for bettered performance of operations.

3. Apache Flink: Flink is an open source that process data into frame that can perform high- speed analysis on live data aqueducts. Due to its performance it has API and scalable armature which has a growing community of inventors.

4. Presto: Presto is an open source SQL query machine level language that supports interactive of data over databases and huge data sets stored in multiple systems(e.g., HDFS, Cassandra and Hive, etc.,). Due to its distributed data query processing armature offers strong performance in processing of operations.

 5. Druid: Druid is an open source logical data store which can be designed for OLAP& OLTP queries on event grounded data in data processing.These technologies should be apprehensive of in the coming days for accessing data over multiple business logics. The data continues to grow for better results.So,the need for innovative results to effectively collect, store, and dissect.

VI. Big Data Emerging Technologies: with the number of arising technologies are being used to process of data to collect, store, and dissect. In this scenario big data adopts Hadoop, NoSQL databases for large amount of data. While each of these technologies has its own unique benefits they are all capability to handle largequantities or huge amount of data and efficiently. Here some of the big data tools which can be used and categorized as below.

a. Big Data Tool: Big Data tools are software tools which helps to associations collect, store, and dissect large quantities of data among different organizations. Big Data has come increasingly important in recent times for further businesses intelligences in induce large volumes of data. While traditional styles of data can be analysis in frequently time consuming and precious usage. Big Data tools make it possible to reuse large quantities of data for processing into small and less interval of time consuming and precious. Big Data tools are a wide variety of different tools available on the request of processing of data and choosing the right tool for a particular association can be a complex less task. The very most important Big Data tools are Hadoop, Spark, and Flink. Big Data Tools are designed to help associations to deal with the huge amount of information. By using Big Data technologies businesses can gain perceptivity results from all that data and make better opinions that lead to betteroperations as results. There are a number of different Big Data tools and ways available, including Apache Hadoop, NoSQL databases, and MapReduce. Each has its strengths and is more suited for business tasks than others. Choosing the right Big Data tool is important to consider for specific requirements and pretensions.

b. Data Tools: The need for big data tools is more apparent as data becomes increasingly and become more complex. These are some challenges issues to provide remedies.

1. Apache Hadoop: It is an open source to access big data platform that helps to reuse and manage large data sets for the processing.

2. Apache Spark: It is a high- performance big data processing machine that can be used for a variety of task and a high- performance big data processing machine that can be used for a variety of tasks, results, offering a comprehensive platform that includes everything from data storehouse to analysis and machine literacy ways.

3.Hortonworks: It is another important source of IoT and it plays a major role in the big data space, offering a robust platform that helps associations effectively process and dissect large data sets.

4. IBM BigInsights: It is most important for big data platform which helpsassociations to gain perceptivity data from other sources. It includes features similar as textbook analytics andsocial media analytics.

5. MapR: It is a big data platform that helps associations process and analyzes large data sets at high levels. It includes some features which are similar as real time streaming and in memory processing.

6. Oracle Big Data Appliance: It is a turnkey result which helps associations to snappily and fluently emplace a big data structure. It includes Oracle's Extradata database system and other important Oracle software

products.

 1. Vital HDB: vital HDB is a pall grounded big data platform that helps associations process and analyzes large data sets over the internet. It includes features similar as real time processing and in memory computing.

2. Platfora: Platfora is a base grounded big data platform that helps to associations snappily and fluently emplace a big data structure. It includes features like similar as tone service analytics and visualizations.

3. Teradata Aster: One of the important big data platforms thathelps associations to gain perceptivity from their data through advanced analytics ways similar as social network analysisand prophetic modeling.

VII. What Is the Future of Big Data?

Big data is one of the most important resolutions in the business world moment. But what's big data, exactly? And what does it mean for the future of business? Big data generally refers to datasets that are too large and huge complex for traditional data processing systems. As businesses techniques increasingly induce andcollect large quantities of data turning to big data results. To help them make sense of it come from a variety of sources similar as social media, client relations, detectors, and transactional data and other source of systems data. While the volume and variety of big data can be reached, it also provides a wealth of occasion for businesses to gain more profits. By understanding client gets relating trends, and perfecting functional effectiveness, businesses can use big data to gain a competitive advantage. In future businesses intelligent using big data drive decision making and produce value for their guests.

VIII. Conclusion:

Big data technologies can anticipate to see further artificial intelligence and machine learning being used to make sense of all the data out there as well as block chain technology getting more efficient in big data technologies operation and security. However ensure familiar with these big data technologies and it can enroll in the Knowledge Hunt.

References:

1. Dirkks,S., Gurdgievc,C., & Keeling,M.(2021). Smarter metropolises for Smarter Growth Optimize Their Systems for the
   gift- Grounded Frugality. Somers, NYIBM Global Business Services.

2) Chrourabi,H., Naam,T., Wallker,S., Gil- Garcia,J.R., Mellouli,S., Nahon,K.,. &Scholl,H.J.(2019,January). Understanding
   smart metropolises An integrative frame. In System Science(HICSS), 2019 45th Hawaii International Conference

3) V. Raada, Solid Waste Management Policy and Planning for a Sustainable Society, Apple Academic Press, 2016.

4) US Department of Transportion, Pollution Prevention and Waste Management, 2017.(Online).

5) Bonmomi,F., Millito,R., Zhu,J. and Addepalli,S., 2017, August. Fog computing and its part in the internet of effects. In
   Proceedings of the first edition of the MCC factory on Mobile pall computing(pp. 13- 16).

6) Microsoft Azure Iot on SereverlessHackathon, https//azurehacks.devpost.com/(last penetrated Nov 15, 2018