Big Data Mining: An Overview Of Current Practices And Future Innovations

Dhruvitkumar Patel[1*], Priyam Vaghasia[2]

[1*,2]Staten Island Performing Provider System, Mondrian collection, pateldhruvit2407@gmail.com, priyamvaghasia57@gmail.com

ARTICLE INFO	ABSTRACT
	Within the later a long time, we seeing a colossal increment in how much information we are able collect. There's information coming from sensors, gadgets, and other places in numerous distinctive groups. But presently, we got so much information, it's exceptionally difficult to oversee all of it. For illustration, Google had almost a million web pages in 1998. But by 2000, that number already went up to 1 billion. And in 2008, it coming to 1 trillion! This is often developing indeed more with social media like Facebook and Twitter, where individuals can make substance effectively, making the information indeed greater. Presently, with Web of Things (IoT), it's getting more insane. Everything, from coffee machines to cars, is connected to internet. All these things are making information all the time. Take for illustration, your morning commute. To figure out perfect way, the most perfect way to urge to work, a framework requiring to check activity, climate, street works, and indeed your calendar. All these information ought to be prepare quick to induce you to work on time. But indeed we have all this information, knowing what to do with it is huge issue. We requiring superior frameworks and calculations to go through the information and discover the valuable things. Also, with so much information around, we must make beyond any doubt it is safe and not being utilized wrongly. In this paper, we are looking at where Enormous Information mining is right presently and what might happen in future. We'll see the issues we confronting, like framework capabilities, making calculations, and keeping information secure. **Keywords:** Electrical Engineering, Construction Sites, Artificial Intelligence, HVAC Systems, Predictive Maintenance

1. INTRODUCTION

In today's world, the way we assemble information from diverse sources have changed past acknowledgment. We're presently able to gather information from sensors, gadgets, and incalculable applications, and this blast of information has been nothing brief of mind blowing. In any case, even as our capacity to gather information has skyrocketed, we are still battling with the enormous errand of handling, analyzing, and understanding all this data. Fair think around it: in 1998, Google had around one million web pages recorded. But by 2000, that number hopped to one billion, and by 2008, it as of now surpassed one trillion, and presently, the information is developing speedier than our conventional strategies can handle. Social media stages like Facebook, Twitter, and Weibo have included to the complexity. These stages let users unreservedly make and share substance, causing the net to grow at an even more fast pace. Keeping track of all this information feels like attempting to drink from a fire hose—there's fair as well much information coming in as well rapidly. [1]

Besides, the approach of versatile phones has significantly expanded the sum of information accessible to us. Versatile phones presently collect a wide extend of data, from our area to our browsing propensities and app utilization. It's not around call records any longer. This tremendous sum of data can be utilized to move forward administrations and make our lives easier, but processing it effectively may be a significant challenge. The issue is developing bigger as increasingly information is collected each day. With the rise of the Web of Things (IoT), things are planning to get indeed more complicated. IoT interfaces ordinary devices—think coffee machines, cars, and indeed whole buildings—to the web, permitting them to communicate and share data. This implies we're looking at a enormous increment in information from trillions of gadgets. For illustration, envision waking up within the morning and your associated gadgets planning to optimize your commute by checking

activity, climate, and development updates. For this to happen, we require to handle a gigantic sum of data rapidly and precisely, which may be a colossal challenge given the scale of information included.

Enormous Information has gotten to be a major center for the information mining community. Occasions just like the KDD conference, which is all almost 'Mining the Huge Data,' have gotten to be exceptionally prevalent. At these occasions, analysts and industry specialists accumulate to share their most recent discoveries and talk about ways to handle the challenges displayed by Huge Information. As Huge Information proceeds to develop, it is seen as one of the foremost energizing ranges with bounty of openings, but it too comes with noteworthy challenges. We are managing with various issues related to Enormous Information, such as framework capabilities, calculation plan, and commerce models. Conventional frameworks are basically not built to handle the scale and complexity of Huge Information. We got to create modern innovations and strategies to handle, analyze, and store all this information viably. Businesses ought to figure out how to monetize Huge Information and utilize it to make esteem whereas remaining competitive in a world where information is getting to be the foremost profitable resource. [2]

Security and privacy are too major concerns when managing with such endless sums of information. With so much information being collected, there's a genuine hazard of abuse or unauthorized get to. Individual information is especially touchy, and the more information that's collected, the more prominent the chance of security breaches. We need to discover ways to secure this information and guarantee it is used responsibly. Privacy could be a principal right that should be secured, indeed as we proceed to gather and analyze increasingly information. The sheer assortment of information being collected includes another layer of complexity. Information comes in all sorts of formats—text, pictures, recordings, and more. Making sense of this different information isn't simple. We got to guarantee that the information is dependable, reliable, and total. With so numerous distinctive sources and types of information, guaranteeing quality is pivotal, particularly when it's being utilized to form imperative choices or forecasts. Imperfect information can lead to imperfect comes about. [3]

Interactive media information postures its claim one-of-a-kind challenges. This sort of information can uncover exceptionally touchy data, and the speed at which it is produced makes it indeed harder to oversee. The speed of data—how rapidly it is made and transmitted—adds another layer of trouble. We ought to prepare this data in real-time, but doing so safely and precisely could be a huge challenge. The quicker the data moves, the harder it is to guarantee its quality and astuteness. As we proceed to gather more information, the related dangers moreover increment. The danger scene is advancing quickly, with modern sorts of cyber dangers rising all the time. Cybercriminals are getting to be more sophisticated, using progressed instruments and procedures to misuse Huge Information for noxious purposes. This makes modern security challenges that have to be be addressed. On the off chance that we do not address these challenges, we might confront large-scale security occurrences that we're not arranged to handle. We are at a basic crossroads within the improvement of Enormous Information. The openings it presents are gigantic, but so are the challenges. We ought to create modern frameworks, calculations, and trade models to keep pace with the developing scale of information. At the same time, we must guarantee that information is utilized mindfully and safely. Long-term of Enormous Information is both promising and dubious. We have to be be arranged for the challenges ahead and proceed to improve in arrange to fully realize the potential of Huge Information. [4]

2. BIG DATA

Huge Information alludes to the endless volumes of information that are created and collected each day from a large number of sources. The term was to begin with presented in 1998 by John Mashey in a introduction titled "Enormous Information and the Following Wave of InfraStress." From its initiation, Huge Information mining was recognized as pivotal since indeed the primary book specifying the term was a information mining book distributed the same year by Weiss and Indrukya. The term picked up assist footing when Diebold used "Huge Information" within the title of a paper in 2000.

So, why do we call it "Big Data"? It's since the sum of information we create is enormous—indeed, enormous. For occurrence, Usama Fayyad highlighted in a conversation at the KDD BigMine'12 Workshop that each day, Google forms over 1 billion look questions, Twitter produces more than 250 million tweets, Facebook sees over 800 million overhauls, and YouTube garners more than 4 billion views. The information being delivered is presently assessed within the run of zettabytes, which are trillions of gigabytes, and this information volume develops at an surprising rate of around 40% per year. [5]

A critical portion of this information comes from portable gadgets. Major tech companies like Google, Apple, Facebook, Yahoo, and Twitter are closely analyzing this information to reveal designs that can improve client encounters. For case, Alex 'Sandy' Pentland at MIT's Human Elements Research facility is inquiring about designs in portable information to get it users' behaviors better—focusing on what individuals really do, instead of what they say they do.

To oversee and make sense of such tremendous amounts of information, we require unused calculations and devices. Doug Laney was one of the primary to characterize the "Three V's" of Huge Information administration, which are:

1. Volume: This alludes to the sheer sum of information. We are managing with more information than ever some time recently, and its measure keeps developing. Tragically, our existing devices can handle as it were a little division of this information.

2. Assortment: Information comes in numerous diverse forms—text, sensor information, sound, video, and graphs, to title some . Dealing with and analyzing these assorted sorts requires specialized strategies.

3. Velocity: Information streams in ceaselessly, regularly in real-time. We got to extricate profitable bits of knowledge from this spilling information because it arrives.

In expansion to these, two more V's have been included to the Huge Information system:

4. Changeability: This includes changes in information structure and how clients translate the information over time. Information may alter in arrange or setting, influencing its ease of use.

5. Esteem: This speaks to the trade advantage picked up from analyzing Huge Information. Organizations can make educated choices based on bits of knowledge from the information, which can lead to a noteworthy competitive edge.

Concurring to Gartner's 2012 definition, Enormous Information includes high-volume, high-velocity, and high-variety data resources that require imaginative and cost-effective shapes of preparing for improved experiences and decision-making. [6]

Huge Information has various applications over different areas:

• **Trade:** Companies utilize Enormous Information for client personalization and churn location. By analyzing client behavior, businesses can tailor their items and administrations to way better meet person needs.

• **Innovation:** Huge Information makes a difference in lessening prepare times drastically, from hours to seconds. This may lead to speedier and more productive operations.

• **Wellbeing:** In healthcare, Enormous Information is utilized to analyze DNA and other health-related information to monitor and make strides person wellbeing results. This may lead to early location and avoidance of ailments.

• **Shrewd Cities:** Huge Information underpins the advancement of shrewd cities that point for economical financial development and a tall quality of life through cleverly administration of assets and administrations.

These applications are revolutionizing how we connected with information, advertising more personalized administrations, making strides client encounters, and empowering superior wellbeing administration. The control of Enormous Information lies in its capacity to supply bits of knowledge that were already past our reach, in this way driving innovation and effectiveness over different spaces. [7]

3. GLOBAL PULSE: "BIG DATA FOR DEVELOPMENT"

Let's dive into an invigorating illustration of how Enormous Information isn't fair a buzzword but a transformative drive for great. Enter Worldwide Beat, a groundbreaking activity initiated by the Joined together Countries, which kicked off in 2009. This activity isn't fair fiddling in Enormous Information; it's utilizing it as a signal of trust to progress lives in creating countries. Think of Worldwide Beat as an avant-garde lab where the enchantment of Enormous Information is saddled to tackle some of the foremost squeezing issues confronted by these nations. Here's how Worldwide Beat is shaking things up: To begin with, they're not fair skimming the surface. They're plunging profound into real-time advanced information to uncover early signs of developing vulnerabilities. This isn't fair information collection; it's approximately creating imaginative methods and procedures to analyze this information because it streams in. Moment, they're making a treasure trove of free, open-source instruments outlined to analyze this real-time information. These apparatuses are not fair for appear; they're implied to share bits of knowledge and speculations that can drive real change. And third, they're rolling out an yearning arrange to set up a worldwide arrange of Beat Labs, each acting as a pilot ground to test and refine these approaches at a nation level. In their eye-opening White Paper titled "Big Information for Advancement: Challenges & Opportunities," Worldwide Beat traces the mind-blowing openings Huge Information presents for creating nations. Let's break it down:

Early Caution: Envision the capacity to spot emergencies some time recently they winding out of control. By identifying peculiarities in advanced media utilization, Worldwide Beat points to create quick reaction components that can handle emergencies head-on. [8]

Real-Time Mindfulness: It's about crafting programs and arrangements that are not fair based on obsolete data but on a nuanced, up-to-the-minute representation of reality. This fine-grained understanding permits for more exact and viable policymaking.

Real-Time Criticism: Here's where the enchantment of Huge Information truly comes into play. By checking the adequacy of arrangements and programs in real time, Worldwide Beat can quickly recognize what's falling flat and make the vital alterations on the fly. This iterative approach guarantees that intercessions are persistently optimized. The transformation of Huge Information isn't restricted to the created world; it's making waves in creating nations as well. With over five billion mobile phones in circulation and a amazing 80% of them being utilized in creating nations, the potential for Huge Information mining is enormous.
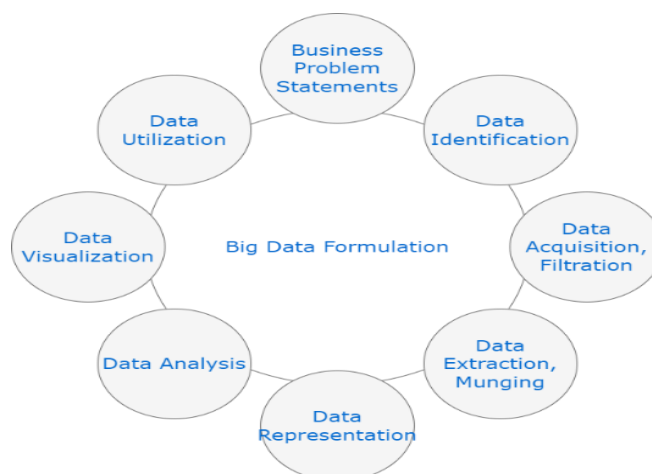
Fig.1: Big Data analysing stages

Versatile innovation is democratizing information get to and examination, engaging indeed the foremost inaccessible locales to take an interest in this data-driven change. Worldwide Beat represents how Huge Information can be a game-changer, advertising a help to creating nations and turning information into a effective instrument for social great. Figure 1 describes the various stages of big-data formulation analysis. [9]

4. COMPARISON OF ADVANTAGES
We've curated four standout commitments that collectively paint a distinctive picture of cutting-edge inquire about in Huge Information Mining and offer a clearing see of the field's direction into long term. For those hungry for indeed more groundbreaking work, the most conferences—such as KDD, ICDM, ECMLPKDD—or diaries like "Information Mining and Information Disclosure" and "Machine Learning" offer a treasure trove of critical thinks about.

Scaling Enormous Information Mining Framework:
The Twitter Involvement by Jimmy Lin and Dmitriy Ryaboy (Twitter, Inc.). Jump into the heart of Enormous Information mining framework through the focal point of Twitter's involvement. This paper pulls back the shade on the abrasive reality of analytics at a gigantic scale, uncovering that in spite of propels, the journey from crude information to significant experiences is anything but direct. Much of the exertion is wrapped up within the preliminary crush, with the genuine challenge lying in changing preparatory models into strong, operational arrangements. [10]

Mining Heterogeneous Data Systems:
A Auxiliary Examination Approach by Yizhou Sun (Northeastern College) and Jiawei Han (College of Illinois at Urbana-Champaign). Investigate the promising wilderness of heterogeneous data arrange mining. This paper dives into the profundities of interconnected, multi-typed data—reframing conventional social databases as dynamic, heterogeneous data systems. These semi-structured organize models use the complex semantics of written hubs and joins, uncovering layers of shocking and wealthy information buried within interconnected information. [11]

Enormous Chart Mining:
Algorithms and Disclosures by U Kang and Christos Faloutsos (Carnegie Mellon College). Take a travel through the domain of enormous chart mining, enlightened by the Pegasus device. This paper offers a all encompassing see of mining enormous charts, spotlighting interesting revelations inside the Internet Chart and Twitter's social arrange. It not as it were presents groundbreaking discoveries but moreover seeds future inquire about bearings with motivation for the another jumps in enormous chart mining.
Mining Huge Streams of Client Information for Personalized Proposals by Xavier Amatriain (Netflix). Inundate yourself within the world of personalized suggestions with experiences gathered from the Netflix Prize. This paper reflects on the lessons learned and the recommender and personalization strategies that Netflix utilizes. It addresses key challenges, later headways, and the continuous talk about of whether more information or way better models are the key to refining our learning strategies. [12]
These papers are not just showcasing the cutting edge of Big Data mining research, but highlighting the dynamic and evolving landscape of this field, pushing boundaries and charting new domains.

5. COMPARISON OF DRAWBACKS
Using As Enormous Information takes center arrange within the tech universe, it hasn't gotten away the whirlpool of discussion and talk about. Undoubtedly, the talk around it is as energetic as the information it looks for to analyze. Here's a synthesized outline of the winning contentions and viewpoints:

Is Huge Information Really Diverse from Information Analytics? A few contend that the refinement between Enormous Information analytics and traditional data analytics could be a unsettled point. Information is on an exponential development direction, and the concept of "little information" might fair be a antique of the past. In quintessence, in case data's persistent development proceeds, at that point what we consider "Huge Information" nowadays may simply ended up the standard of tomorrow.

Hadoop:

The Buildup vs. Reality There's a discernable skepticism almost whether Enormous Data's buildup is simply a ploy to thrust Hadoop-based computing frameworks. Pundits contend that whereas Hadoop is frequently touted as the go-to arrangement, it's not always the leading fit. MapReduce, for illustration, may not be the foremost proficient programming stage for medium-sized endeavors. It shows up that a few information administration merchants are intensely advancing Hadoop-based frameworks, possibly overselling its capabilities [13].

The Significance of Information Recency When jumping into real-time analytics, the focus may move from the sheer volume of information to its convenience. In scenarios where information is in flux, its recency gets to be fundamental. The esteem of information can be less approximately how much there's and more around how current it is, particularly in quickly advancing settings.

Exactness and Correlation Claims Be attentive of exactness claims. As Nassim Nicholas Taleb talks about in his most recent book, as the number of factors increments, so does the probability of finding spurious relationships. A amusing however telling illustration is Leinweber's finding that the S&P 500 record was once connected with butter generation in Bangladesh—an outline of how a few relationships can be both questionable and engaging. [14]

Measure Does Not Rise to Quality. "Greater is superior" doesn't continuously apply to information. Quality trumps amount, and boisterous or unrepresentative information can be more problematic than useful. For occurrence, accepting that Twitter users represent the worldwide populace can be deluding, as their behavior may not continuously reflect broader patterns.

Moral Situations of Information Openness One of the central moral wrangles about spins around whether it's right for people to be analyzed without their unequivocal assent. The address of protection in the age of Huge Information is basic, because it challenges our ideas of assent and reconnaissance.

The Rise of Computerized Isolates The uneven get to to Enormous Information is giving rise to modern computerized partitions. There's a developing dissimilarity between those who can tackle the control of Enormous Information and those who cannot. Organizations with Enormous Information get to can extricate bits of knowledge that may be out of reach for others, possibly making a chasm between Huge Information "haves" and "have-nots."

In rundown, whereas Enormous Information guarantees progressive bits of knowledge and capabilities, it moreover brings with it a complex cluster of challenges and moral concerns. The wrangle about proceeds, as we explore this quickly advancing scene, looking for to adjust development with duty. [14]

6. TOOLS: OPEN-SOURCE REVOLUTION

Exploring The blast of Enormous Information is inseparably connected with the open-source computer program revolution—a advantageous relationship where both domains bolster off each other's progressions and commitments. Major tech monsters like Facebook, Yahoo!, Twitter, and LinkedIn not as it were use but moreover effectively contribute to the dynamic biological system of open source ventures that support their Huge Information frameworks. Among these, Hadoop stands out as a foundation, interwoven with a plenty of related computer program devices that collectively shape the scene of information preparing and investigation.

Let's plunge into the domain of Apache Hadoop, an effective system outlined for taking care of data-intensive conveyed applications. It's in a general sense built upon the MapReduce programming show, combined with the Hadoop Disseminated Filesystem (HDFS). Hadoop encourages the composing of applications able of handling colossal volumes of information in parallel over broad clusters of computing hubs. The method kicks off with a MapReduce work those allotments the input dataset into discrete subsets, which are at that point prepared at the same time by outline assignments. Taking after this, a diminishment stage solidifies the comes about from the map tasks to determine the ultimate yield. [16]

Digging more profound into the Hadoop environment, we experience an cluster of complementary ventures that improve its capabilities. Apache Pig, Hive, HBase, ZooKeeper, Cassandra, Cascading, and Recorder, among others, contribute specialized functionalities, from high-level scripting dialects to disseminated databases and real-time coodination.

At that point, there's Apache S4, a stage made for handling nonstop information streams with artfulness. Not at all like conventional batch-processing systems, S4 exceeds expectations in overseeing and analyzing information streams in genuine time, permitting for the creation of advanced applications that coordinated stream and preparing components consistently.

Storm, another player within the gushing information space, created by Nathan Marz at Twitter, offers a strong elective to S4. It's designed to handle data-intensive dispersed applications, amplifying the capabilities of stream preparing with its possess set of highlights. [17]

Turning our look towards the domain of open-source activities particularly custom fitted for Huge Information mining, a few conspicuous apparatuses develop:

Apache Mahout: This adaptable computer program library gives a plenty of machine learning and information mining calculations, all built on beat of Hadoop. It underpins clustering, classification, collaborative sifting, and visit design mining, advertising a wide range of explanatory apparatuses.

R: Born out of the College of Auckland in 1993, R is an open-source programming dialect and computer program environment committed to factual computing and visualization. It's eminent for its capacity to perform complex measurable examinations on enormous datasets, making it irreplaceable for information researchers. [18]

MOA: The Gigantic Online Examination system, starting from the College of Waikato, specializes in real-time stream information mining. It gives executions for classification, relapse, clustering, and visit thing set mining, all equipped towards preparing information because it streams. The streams system inside MOA bolsters the definition and execution of stream forms through XML-based definitions and coordinating with stages like Android and Storm. The rising SAMOA venture points to bind together S4 and Storm with MOA, moving dispersed stream mining into unused wildernesses. [19]

Vowpal Wabbit: At first created at Yahoo! Investigate and presently proceeded at Microsoft Investigate, Vowpal Wabbit is an open source extend planned for high-speed, adaptable learning. It exceeds expectations in taking care of terafeature datasets and offers surprising throughput through parallel learning strategies. When we contract our center to Huge Chart mining, a unmistakable set of open source devices comes into play:

Pegasus: Built on the MapReduce system, Pegasus stands as a vigorous huge chart mining framework able of revealing designs and peculiarities inside enormous, real-world charts. It's an fundamental device for those investigating the complex networks of associations inside huge datasets.

GraphLab: This high-level graph-parallel framework takes a diverse approach by working freely of MapReduce. It forms subordinate records spoken to as vertices in a dispersed data-graph, executing vertex-programs in parallel and permitting vertices to interact with their neighbors for upgraded expository profundity.

In short, the open-source development has catalyzed a transformation in Huge Information mining, advertising a assorted toolkit that engages analysts and professionals to handle the complexities of cutting edge information scenes with phenomenal dexterity and development. [20]

7. CONCLUSION

Within the domain of Huge Information administration and analytics, various imposing challenges linger on the skyline, driven by the inalienable nature of information: its immensity, differences, and steady advancement. One of the vital problems is the engineering of analytics systems—how to ideally plan a framework competent of concurrently overseeing both chronicled and real-time information remains an puzzler. Enter the Lambda Design, proposed by Nathan Marz, which handles this issue by breaking down the issue into three particular layers: the bunch layer, the serving layer, and the speed layer. This engineering consistently coordinating Hadoop for clump preparing and Storm for speed, gloating properties such as vigor, blame resistance, versatility, and negligible upkeep. Another pivotal jump is accomplishing factual noteworthiness in the midst of the endless ocean of information; as Efron highlights, exploring through gigantic datasets with a large number of questions can effortlessly lead one off track in the event that not drawn closer with thorough measurable strategies. Dispersed mining presents however another challenge; numerous traditional data mining strategies don't loan themselves effectively to parallelization, requiring broad investigate to create both commonsense and hypothetical systems for modern strategies. In addition, the energetic nature of data—constantly advancing over time—requires Enormous Information mining strategies to not as it were adjust but also proactively distinguish changes. Procedures in information stream mining offer promising arrangements for such errands. Compression versus inspecting postures a basic problem in overseeing capacity space: compression keeps up information keenness but requires more time and space, whereas inspecting decreases data but offers critical space investment funds. Coresets, as utilized by Feldman and colleagues, outline a strategy of diminishing Huge Information complexity by approximating the initial dataset with littler, sensible sets. Visualization of Huge Information comes about remains a overwhelming errand; with information being so monstrous, making user-friendly and smart visual representations requests inventive strategies and systems, associated to the reminiscent photos and infographics found in "The Human Confront of Huge Data." In conclusion, a critical parcel of important information remains concealed and undiscovered, as untagged, file-based, and unstructured information multiplies. The IDC ponder on Huge Information illustrates that a stunning rate of the advanced universe's potential remains torpid due to lacking labeling and investigation, taking off endless supplies of potentially noteworthy bits of knowledge mulling in lack of clarity.

REFERENCES

1. Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... & Zaniolo, C. (2007). Top 10 algorithms in data mining. Knowledge and information systems, 14(1), 1-37.
2. Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Morgan kaufmann.

3. McSherry, F., & Isard, M. (2013). Dremel: Interactive analysis of web-scale datasets. Communications of the ACM, 56(1), 90-99.
4. Chen, M., & Mao, K. (2014). Big data: A survey. Mobile networks and applications, 19(2), 171-179.
5. Dean, J., & Ghemawat, S. (2004). MapReduce: Simplified data processing on large clusters. Communications of the ACM, 51(1), 107-113.
6. Fan, W., Li, C., & Liu, J. (2014). Challenges of big data analysis. National science review, 1(2), 253-262.
7. Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live and work. Houghton Mifflin Harcourt.
8. Mahajan, Lavish, Rizwan Ahmed, Raj Kumar Gupta, Anil Kumar Jakkani, and Sitaram Longani. "DESIGN OF WIRELESS DATA ACQUISITION AND CONTROL SYSTEM USING LEGO TECHNIQUE." International Journal of Advance Research in Engineering, Science & Technology 2, no. 5 (2015): 352-356.
9. Barbará, D., & Pigot, S. (2013). Challenges in big data management. Computing surveys, 45(4), 41.
10. Sarawagi, S., & Dhulipala, S. (2014). Data quality in big data. In Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1-10).
11. Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large datasets. In Proceedings of the IEEE symposium on security and privacy (pp. 111-125).
12. Racharla, Mr Sathya Prakash, Mr Kontham Sridhar Babu, and Anil Kumar Jakkani. "An Iterative approach for the Restoration of Motion Blurred Images."
13. Lee, Y., & Chen, A. S. (2015). Big data applications in healthcare. Business informatics, 57(2), 383-404.
14. Sharda, K., Delen, D., & Sharda, R. (2016). Big data applications in marketing: A review and future directions. Journal of marketing analytics, 4(1), 35-53.
15. Mehrotra, S., Sanketi, P., & Upadhyaya, S. (2016). Big data analytics in finance: A review. International journal of computer applications, 141(11), 1-8.
16. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.
17. Srivastava DP. Prof. Anil Kumar Jakkani,"Android Controlled Smart Notice Board using IOT". International Journal of Pure and Applied Mathematics.;120(6).
18. McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. Harvard business review, 90(6), 60-68.
19. Jakkani, Anil Kumar. "Enhancing Urban Sustainability through AI-Driven Energy Efficiency Strategies in Cloud-Enabled Smart Cities." (2024).
20. Nalla, Akash, and Anil Kumar Jakkani. "A Review on Recent Advances in Chatbot Design." integration 3.3 (2023).
21. White, T. (2012). Hadoop: The definitive guide. O'Reilly Media.
22. Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., & Stoica, I. (2012). Spark: Cluster computing with working set management. In Proceedings of the 2nd USENIX conference on hot topics in cloud computing (pp. 10-10).
23. Srivastava, Pankaj Kumar, and Anil Kumar Jakkani. "FPGA Implementation of Pipelined 8× 8 2-D DCT and IDCT Structure for H. 264 Protocol." 2018 3rd International Conference for Convergence in Technology (I2CT). IEEE, 2018.
24. Stonebraker, M., & Cetintemel, U. (2011). Big data: Challenges and opportunities. ACM SIGMOD Record, 40(4), 4-11.
25. Srivastava, P. K., and Anil Kumar Jakkani. "Non-linear Modified Energy Detector (NMED) for Random Signals in Gaussian Noise of Cognitive Radio." International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy. Singapore: Springer Nature Singapore, 2020.
26. Ghodsi, A., Zaharia, M., Shenker, S., & Stoica, I. (2011). I/O-centric cluster computing. In Proceedings of the 8th USENIX symposium on operating systems design and implementation (pp. 265-278).
27. Agbonyin, Adeola, Premkumar Reddy, and Anil Kumar Jakkani. "UTILIZING INTERNET OF THINGS (IOT), ARTIFICIAL INTELLIGENCE, AND VEHICLE TELEMATICS FOR SUSTAINABLE GROWTH IN SMALL, AND MEDIUM FIRMS (SMES)." (2024).
28. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
29. Jakkani, Anil Kumar, Premkumar Reddy, and Jayesh Jhurani. "Design of a Novel Deep Learning Methodology for IOT Botnet based Attack Detection." International Journal on Recent and Innovation Trends in Computing and Communication Design 11 (2023): 4922-4927.