International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

A SURVEY ON TOWARD EFFECTIVE RESPONSE TO NATURAL DISASTERS: A DATA SCIENCE APPROACH

K.U.Ranjith[1], Dr. S. Karuppusamy M.E., Ph.D[2]
[1]PG Scholar, Department of Computer Science and Engineering, Nandha Engineering College (Autonomous), Erode, Tamilnadu, India
[2]Associate Professor, Department of Computer Science and Engineering, Nandha Engineering College (Autonomous), Erode, Tamilnadu, India[2]

ABSTRACT
Natural catastrophes have the potential to destroy large portions of infrastructure and kill thousands of people. Both the populace and the government find it challenging to deal with these situations. Particular attention must be given to the following two difficult problems: find a workable solution first evacuating people, then rebuilding homes and other infrastructure. Then, a successful recovery plan that prioritises the reconstruction of damaged areas and the evacuation of people can be a game-changer for overcoming those horrible circumstances. In this light, we introduce DiReCT, a method based on I a dynamic optimization model created to quickly develop an evacuation plan of an earthquake-stricken area, and ii) a double deep Q network-based decision support system capable of effectively guiding the rebuilding of the affected areas. The latter operates by taking into account the needs of the many stakeholders (such as citizens' social benefits and political priorities) as well as the resources available. The foundation for both of the aforementioned solutions is a specialized geographic data extraction Method called "GisToGraph," which was created expressly for this use. We used extensive GIS data, information on the vulnerability of urban land structures, and the historical city centre of L'Aquila (Italy) to test the applicability of the entire strategy.
KEYWORDS: Data Science,Decision-Support System, Deep Reinforcement Learning, Evacuation Plan,Flow Model, Geographic Information, Network

1. INTRODUCTION

1.1 DATA SCIENCE

Data science is the study of data with the goal of gaining important business insights. It is a multidisciplinary method for analyzing massive volumes of data that integrates ideas and techniques from the domains of mathematics, statistics, artificial intelligence, and computer engineering. Data scientists can ask and receive answers to questions like what occurred, why it occurred, what will occur, and what can be done with the outcomes thanks to this study. Because it integrates tools, techniques, and technologies to derive meaning from data, data science is significant. A profusion of gadgets that can automatically gather and store data has flooded modern enterprises with data. In the areas of e-commerce, healthcare, banking, and every other facet of human existence, online systems and payment portals collect more data.

1.2DECISION-SUPPORT SYSTEM

An interactive information system called decision support system (DSS) analyses enormous amounts of data to help guide business decisions. By evaluating the relevance of uncertainties and the tradeoffs involved in making one choice over another, a DSS assists management, operations, and planning levels of an organization in making better decisions.To assist users in making decisions, a DSS uses a variety of raw data, papers, personal knowledge, and/or business models. Relational data sources, cubes, data warehouses, electronic health records (EHRs), income estimates, sales projections, and other sources may all be utilized by a DSS.Business intelligence (BI) and DSS are frequently confused. Some professionals view BI as DSS's successor.

1.3DEEP REINFORCEMENT LEARNING

Reinforcement learning has gained a lot of popularity in recent years as a result of its success in solving difficult sequential decision-making problems. To solve difficult sequential decision-making

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

problems, deep reinforcement learning combines deep learning techniques with reinforcement learning. Deep learning is most helpful when solving issues with high-dimensional state spaces. As a result of its capacity to learn various levels of abstraction from data, deep learning enables reinforcement learning to solve more challenging problems with less prior knowledge. To use reinforcement learning successfully in situations approaching real-world complexity, however, agents are confronted with a difficult task: they must derive efficient representations of the environment from high-dimensional sensory inputs, and use these to generalize past experience to new situations. As a result, machines can now imitate some aspects of human problem-solving abilities, even in high-dimensional space, which was previously unthinkable.

1.4 EVACUATION PLAN

Not many people are aware, but evacuations happen frequently. They are typically brought on by fires and floods. Large-scale evacuations are frequently the result of severe storms like hurricanes. Additionally, hundreds of industrial and transportation incidents each year result in the release of hazardous materials, forcing many people to abandon their homes and places of employment. The hazard will determine how long you have to leave. If it's a weather-related calamity, like a hurricane, you might have a day or two to prepare. Planning ahead is crucial because many calamities don't give individuals enough time to acquire even the most basic supplies. Plan how you'll gather your family (or your colleagues if you're planning an evacuation from the job) and your supplies, and consider where you'll go in various scenarios. Choose a few locations in various directions so you'll have options in an emergency and are aware of the evacuation routes to get there.

1.5 FLOW MODEL

At its core, the flow model is a straightforward graphical depiction of how data and artifacts move through the system as it is used. When conducting usage study, it's critical to pinpoint the fundamental system flow as soon as possible. An overview of the flow of information, artifacts, and work products between user work roles and various components of the system or product as a result of user activities is provided by a flow model. What happens, for instance, when a song or other piece of music is bought, downloaded from the Internet, and then loaded or synchronized to a personal device? A flow model is a top-down representation of the work domain, its elements, and linkages between them. It provides a high-level overview of how people in various work roles interact with one another and with other system entities to complete tasks.

1.6 GEOGRAPHIC INFORMATION

A computer system that evaluates and presents information with a geographic context is known as a Geographic Information System (GIS). It employs information linked to a certain place. The majority of the knowledge we have about the world includes a place reference: In what location do USGS stream gages

Exist? From whence was a rock sample taken? Where are all the fire hydrants in a city located? If a rare plant is discovered in three distinct locations, for instance, GIS analysis may reveal that all of the plants are located on slopes with a northerly aspect, are located above 1,000 feet in elevation, and get more than ten inches of precipitation annually. GIS maps can then show all areas in the region that have comparable circumstances, enabling researchers to find more of the uncommon plants. A GIS analysis of farm sites, stream locations, altitudes, and rainfall will reveal which streams are likely to carry that fertilizer downstream provided one knows the precise position of the farms employing a given fertilizer. These are only a few instances of the diverse applications of GIS in the domains of biology, earth sciences, and resource management.

1.7 NETWORK

Two or more computers connected together to share resources (like printers and CDs), exchange files, or enable electronic communications make up a network. A network's connections to its computers can be made by cables, phone lines, radio waves, satellites, or infrared laser beams.

Networks can be divided into two main categories:

Local Area Network (LAN)

Wide Area Network (WAN)

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

The terms Metropolitan Area Networks (MAN), Wireless LAN (WLAN), and Wireless WAN may also be used (WWAN). A network that is contained inside a very limited region is known as a local area network (LAN). It is typically constrained to a certain location, such a writing lab, building, or school. Wide Area Networks (WANs) link networks over wider geographic regions, like Florida, the US, or the entire planet. The connections between this kind of global networks may be made using specialized transoceanic cabling or satellite uplinks. Schools in Florida can quickly communicate with locations like Tokyo using a WAN without incurring astronomical phone bills. A real-time teleconference between two users who are located half a world apart and have computers with microphones and webcams is possible. A WAN is challenging. Local and metropolitan networks are linked to international communications networks like the Internet using multiplexers, bridges, and routers. However, a WAN won't seem all that different to users from a LAN.

2.LITERATURE REVIEW

2.1 IMPLICATIONS FOR SMART CITIES FROM SOCIAL SCIENCE ANALYSIS OF OBJECTIVE AND SUBJECTIVE DATA

The proposal in this paper by Lauraethane et al.We now have the ability to conduct extensive social studies and gather enormous volumes of data from our cities thanks to the ease of deployment of digital technologies and the Internet of Things. With thisIn this research, we investigate a novel approach to analyze social science study data using machine learning and data science methods. By combining objective (sensor information) and subjective data, this helps us to optimise the knowledge from these types of investigations (direct input from the users).A deeper understanding of how people engage with urban green spaces is the goal of the pilot project. In Sheffield, England, 1870 people participated in a field experiment over the course of two different time periods (7 and 30 days). Both objective and arbitrary data were gathered with the aid of aSmartphone app. People entering any of the publicly accessible green places were tracked according to their location.Users could supplement this by adding textual and visual data on their own or in response to prompts (when entering a green space). We find the key qualities noticed by the citizens in both text and photos by utilizing data science and machine learning approaches. Additionally, we examine how much time people spend in parks and the prime places for social interaction. This essay demonstrates the feasibility of integrating technology into extensive sociological studies while giving us a broad picture of specific patterns and the behavior of the individuals within their surroundings.

2.2 THE FOLLOWING IS FROM THE ERUDITE OF THE BD2K TRAINING COORDINATING CENTER:

Jose luis ambite et al., have proposed in this paper the Educational Resource Discovery Index for Data Science. The area of data science has grown to enable the effective integration and analysis ofever-growing data sets in a variety of fields. Big data in especially in genomics, neuroimaging, and mobile healthand other branches of biomedical science, while presenting difficulties, also promise new insights. In order to do thisthe Big Data to Knowledge (BD2K) initiative, which includes a Training Coordinating Center (TCC) tasked with creating a resource for tailored data science training for biomedical researchers, was introduced by the National Institutes of Health. The Erudite, or Educational Resource Discovery Index, which powers the BD2K TCC website compiles training resources for data science, such as online courses, tutorial and research talk videos, textbooks, and other web-based resources. While the sheer number of available learning resources is amazing, their extreme diversity in terms of topic, format, quality, and difficulty makes the area intimidating and challenging to traverse. Additionally, fresh information and ideas are constantly emerging because data science is continually developing. By utilizing data extraction, data integration, machine learning, information retrieval, and natural language processing, we use data science techniques to build Erudite itself. These techniques automatically gather, integrate, describe, and arrange existing online resources for learning data science. When viewed as a standalone data science project, Erudite has advanced significantly in terms of data gathering, integration, exploration, and analysis. In the process of creating Erudite, we have so far created and implemented a configurable scraping

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

framework, a unified schema, a tagging ontology, a resource exploration visualization method, and a collection of automatic tagging algorithms.

2.3 CONNECTING PLAYER BEHAVIORS TO LEARNING OBJECTIVES THROUGH CONTEXTUAL MARKUP AND MINING IN DIGITAL GAMES FOR SCIENCE LEARNING

In this paper, John S et al. make the case that digital games have the potential to significantly improve K–12 science instruction. Unrealized. Pre- and post-test data continue to be the primary source of research used to evaluate learning games, which limits the insights that can be gained. Interactions between game play, game design elements, and formal evaluation that are more complex. Making rich representations for studying game play data is therefore a crucial step forward. Using a metadata markup language that links game actions to ideas pertinent to particular game contexts, this paper uses data mining techniques to model learning and performance. We discuss the findings of a classroom study and point out possible connections between students' planning and prediction behaviors seen across game levels and advancement on formal assessments. The findings have implications for the scaffolding of particular tasks, such as effect prediction, solution planning, and physics learning while playing video games. The strategy underlines the importance of our contextualized method for marking up game play to aid in data mining and discovery overall. Commercial game designs offer strong opportunities for engagement and learning in science. It's important to keep in mind, though, that these affordances developed in response to various pressures and objectives that might not have anything to do with learning science. In order to support logging and analysis of game play behavior with respect to both the learning context and the gaming context, these conventions must be rethought and redesigned.

2.4 TWENTY YEARS LATER, CRISP-DM: FROM DATA MINING METHODS TO DATA SCIENCE PATHWAYS

The CRISP-DM (CRoss-Industry Standard Process for Data Mining) method, which Fernando Martnez-Plumed et al. proposed in this publication, has been around for about 20 years. It continues to be the de facto standard for creating data mining and knowledge, according to numerous studies and user polls. discovery initiatives. The field has unquestionably advanced much in the last twenty years, with data science currently preferred to data mining as the dominant phrase. In this study, we examine whether and under what circumstances CRISP-DM is still appropriate for use with data. scientific endeavors We contend that the process model view still substantially holds if the project is goal-directed and process-driven. On the other hand, as data science initiatives get more exploratory, the potential directions they can go down become more varied, necessitating the need for a more adaptable model. We define the general structure of such a trajectory-based model and discuss how it might be applied to classify data science projects (goal-directed, exploratory or data management).

2.5 A STEP TOWARDS DATA-INTENSIVE SCIENCE: DATA PROSPECTING

In this article, Rahul Ramachandran et al. claimed that data-intensive science is a method of scientific discovery that is fueled by knowledge gleaned from vast amounts of data. instead of the usual hypothesis-driven research methodology. The creation of supporting technologies that enable researchers to efficiently use these massive amounts of data is one of the main issues in data-intensive science. To solve the difficulties of data-intensive science, this study introduces the idea of "data prospecting." In order to characterise the preliminary stage of data exploration used to identify interesting regions for more in-depth research, we expand the widely used metaphor of data mining to the concept of data prospecting. Data prospecting uses interactive discovery engines to improve data selection. A researcher can filter the data using interactive exploration based on "first look" analytics, find intriguing and previously undiscovered patterns to launch new scientific investigations, confirm the accuracy of the data, and confirm whether patterns in the data match current scientific theories or mental models. This study outlines our preliminary assessment of the benefits of "data prospecting" for Earth Science researchers conducting research. The report explains the current limitations of our discovery engine prototype, which supports data prospecting for particular data products. Additionally, three distinct researchers' example science projects that used

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

our prototype discovery engine to examine the Special Sensor Microwave/Imager and Sounder (SSM/I, SSMIS) data products are provided.

2.6 APPLICATION OF DATA SCIENCE ANALYSIS AND PROFILE REPRESENTATION TO SECONDARY ACUTE CORONARY SYNDROME PREVENTION

In this study, Antonio Garcia-Garcia and colleagues suggest that the analysis of massive volumes of data from electronic medical records (EMRs) and routine clinical In recent years, practice data sources have drawn more and more attention. To facilitate the extraction of the amount and diversity of information from these data sources, yet, few systematic methodologies have been put forth. Because ACS exhibits higher morbidity and mortality, data on ACS are specifically available in many hospitals and healthcare facilities. In order to examine and utilize the scientific information content in small ACS samples in a univariate manner, this work suggests a technique called Data Science Analysis and Representation (DSAR). For reliable, cross-sectional, and non-parametric statistical tests on categorical and metric variables, DSAR employs Bootstrap Resembling. Additionally, it creates a useful graphical depiction of the database variables that aids in understanding the results and locating pertinent variables. When searching for the most pertinent variables in the secondary prevention of ACS, our goals were to validate DSAR by comparing it to traditional statistical methods and to ascertain the degree to which these variables were correlated with the Exitus event. In order to accomplish this goal, we used DSAR on a sample of 270 characteristics collected anonymously from 2377 individuals who had been diagnosed with ACS.

2.7 GROUNDING DATA SCIENCE IN A POLITICS OF JUSTICE: DATA SCIENCE AS POLITICAL ACTION

In this study, Ben Green et al. suggest The area of data science has incorporated ethics in reaction to criticism of data-driven algorithms from the public. principles and instruction. Ethics can aid data scientists in considering some normative aspects of their work, but such attempts fall short of producing data science that is socially responsible and supports social justice. In this piece, I contend that data science needs to adopt a political perspective. Data scientists must acknowledge that they are political actors involved in the normative formation of society and judge their own work based on how it will affect people's lives in the long run. I begin by explaining why data scientists need to understand that they are political actors. In this part, I address three objections that data scientists frequently raise when asked to express political opinions on their work. In response to these claims, I explain why trying to be apolitical is in and of itself a political position—a fundamentally conservative one—and why attempts by data science to advance "social good" dangerously rely on unarticulated and increment list political assumptions. After that, I put forth a framework for how data science may develop into a rigorous and deliberative politics of social justice. I view the development of a politically engaged data science as taking place over a period of four stages. By pursuing these novel ideas, data scientists will gain new tools for thinking and methodically advancing social justice.

2.8 A BLOCK-BASED ENVIRONMENT'S DESIGN AND EVALUATION IN A DATA SCIENCE CONTEXT

The proposal made in this work by AUSTIN CORY BART et al. Introduction to computing programmes need new tools to inspire and educate the flood of students with little prior knowledge and varied ambitions as computers becomes more widespread across sectors. To find to enhance courses by adding rich, authentic environments and effective scaffolding that can direct learners toward success using automated technologies, relieving the burden on scarce human instructional resources. To solve these problems, we developed the web-based, open-source, open-access Blocky programming environment for beginning computer science students (https://www.blockpy.com). Through engaging tasks, learners can relate the educational material to real-world situations using the embedded data science framework in Blocky. By facilitating bidirectional, seamless transitions between block and text programming, the block-based system not only guides learners as they finish issues but also facilitates migration to more advanced programming environments.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

2.9 LARGE-SCALE DATA ANALYSIS SUPPORTED BY DISTRIBUTED DATA STRATEGIES ACROSS GEO-DISTRIBUTED DATA CENTERS

Big data storage in a single data centre is no longer practical when the volume of data increases quickly, according to a proposal made in this research by TAMER Z. EMARA et al. As a result, businesses have created two scenarios for storing their large data across several data centers. The large data of the organization are dispersed over several data centers in the first case without data replication. The second scenario involves the storage of data in different data centers, but it also includes the replication of critical data to increase data availability and safety. However, in these situations, it becomes difficult to analyze huge data that is dispersed over numerous data centers. We provide two data distribution mechanisms in this research to assist big data analysis across geographically dispersed data centers. In these solutions, we partition massive data into sets of random sample data blocks and distribute those data blocks across several data centers, either without replication or with replication, using the recently developed Random Sample Partition data model. We choose random samples of data blocks from several data centers and download them to one data centre for analysis while examining huge data across numerous data centers without replication. By randomly choosing a sample of data blocks replicated from other data centers, we can examine large data on any data centre using the second technique with replication of data blocks.

2.10 A DESCRIPTION OF THE ARCHITECTURE AND CAPABILITIES OF THE EOS MLS SCIENCE DATA PROCESSING SYSTEM

The Earth Observing System (EOS) Microwave Limb Sounder (MLS) is an atmospheric remote sensing experiment led by David T et al., and it was proposed in this study. by the California Institute of Technology's Jet Propulsion Laboratory. The goals of the EOS MLS are to increase our knowledge of stratospheric chemistry, mechanisms influencing climate variability, and upper troposphere pollution. The National Aeronautics and Space Administration's (NASA) EOS Aura spacecraft was launched on July 15, 2004, and it carries four instruments, the longest of which, the EOS MLS, has an operating lifespan of at least five years. The Science Data Processing System (SDPS) for the EOS MLS is described in this study along with its capabilities and architectural layout. The Science Computing Facility and the Science Investigator-led Processing System are the two main parts of the SDPS. The EOS MLS Science Team may design scientific algorithms, create processing software, control the quality of data products, and conduct scientific analyses with the help of the Science Computing Facility.

2.11 CREATING A SMART DATA INTEGRATION PLATFORM TO IMPROVE URBAN MOBILITY

The proposal made in this paper by PALOMA CCERES et al. Mobility defines a collection of flows and linkages that limit those citizens' individual and communal well-being in the urban environment, making it one of the primary elements used to describe that well-being. behavior. However, the complexity of this activity on a city-scale renders this a computationally challenging issue. One of the main causes of this is the information asymmetry: many players only have access to incomplete or outdated information, and many pertinent data are just unavailable. In this article, we suggest an architecture and platform for data integration that can be used to combine pertinent data from numerous sources and deliver the results in a number of formats. By utilizing semantic technologies, this integration makes sure that the connections between the data are understood and reflect their true meaning. The resulting platform combines open data, which is accessible from public sources, extracted data, obtained from public sites using scraping techniques, pre-processed data, kept in public databases, aggregated data, obtained from pervasive devices using crowd sourcing, and smart data, provided by mobile applications and enhanced with contextual information, or data concerning specific incidents, frequently provided by the users themselves. This data's semantic integration enables the coordinated computation of a wide range of outputs, from recognizable events to accessible transportation routes. Following that, the general public is given access to these results through particular software, either online or through mobile applications. We believe that by using this knowledge collectively, urban welfare could be enhanced.

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

2.12 SMART DATA ENGINEERING FOR THE TRANSITION TO SECURE NOSQL ENVIRONMENTS

In this study, SHABANA RAMZAN et al. make a proposal. Data is growing quickly in the age of super computers, demanding greater skill. nfrom the data processing, storage, and analysis technologies that are now available. The term "Big data" refers to this ongoing, huge growth of structured and unstructured data. the handling and keeping of largeData collection via traditional methods is not feasible. Developers started using big data databases like Apache Cassandra, Oracle, and NoSQL in the preceding decade as a result of big data solutions' increased proficiency in handling data, such as NoSQL. NoSQL is a cutting-edge database technology that offers scalability to support massive amounts of data, which has helped it become the most practical database option. These contemporary databases strive to get around relational databases' restrictions in terms of continuous availability, high performance, data modeling, and unrestricted scaling. Due to their more adaptable structures, NoSQL databases must now be switched by larger businesses. Given the variability and complexity of relational data, it is a significant barrier for businesses to convert their current databases to NoSQL databases.

2.13 A COMPREHENSIVE REVIEW OF DATA SCIENCE'S USE IN COMBATING COVID-19

The proposal made in this work by Siddique Latif et al. The World Health Organization designated COVID-19, a virus-caused illness, to be a pandemic. By March 2020, World Health Organization (WHO). More than 21 million people had tested positive globally by the middle of August 2020. Infections have been expanding quickly, and great efforts are being undertaken to combat the illness. In this paper, we attempt to systematize the various COVID-19 research activities utilizing data science. We define data science broadly to include all techniques and tools, including those from artificial intelligence (AI), machine learning (ML), statistics, modeling, simulation, and data visualization, that can be used to store, process, and extract knowledge from data. Along with evaluating the rapidly expanding body of recent research, we look over open datasets and repositories that can be used for future research into the COVID-19 outbreak and prevention measures. We also give a bibliometric analysis of the papers published over this brief period as part of this. Finally, based on these observations, we outline frequent difficulties and problems seen in the works surveyed.

2.14 INTERSECTING OR PARALLEL LINES? INTELLIGENT BIBLIOMETRICS FOR EXAMINING DATA SCIENCE'S ROLE IN POLICY ANALYSIS

In this study, Yi Zhang et al. suggest making efforts to include data science into policy analysis. can be traced back many years, yet turning analytical results into judgments is still a difficult undertaking. It is intriguing to explore if data science and policy analysis are developing independently or whether their paths have crossed since data-driven decision-making necessitates a grasp of methodologies, best practices, and research findings from numerous fields. We have developed an intelligent bibliometric framework that combines a number of conventional bibliometric approaches with a novel technique for mapping the evolutionary pathways of scientific innovation, which is used to identify predecessor-descendant relationships in technological topics. From a bibliometric perspective, our investigation is motivated by a comprehensive set of research questions. Our research shows that policy analysis and data science have crossing lines, and it can be predicted that both communities are moving in a cross-disciplinary manner where policy analysis and data science are interacting.

2.15 DATA SCIENCE WORK AND WORKERS: PASSING THE DATA BATON: A RETROSPECTIVE ANALYSIS

In this research, Anamaria Crisan et al. make the suggestion that data science is a rapidly expanding field and that businesses are relying more and more on data science work. The uncertainty, though It might be challenging for visualization researchers to pinpoint productive research trajectories because of the confusion around data science, what it is, and who data scientists are. We've done a retrospective examination of the data science work and the people who did it. data science literature, human computer interface, and visualization. We have created a thorough model from this analysis,

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

which divides data scientists into nine different jobs and defines the effort that goes into data science. We review and discuss the significance of visualization in data science work as well as the diverse tool support requirements of data scientists themselves. Our research aims to provide visualization researchers with a more tangible understanding of data science in the hopes that this would enable them to identify novel potential for influencing data science work. As we've said repeatedly, data scientists frequently visualize data, but the visualization research community is mostly unaware of the visualization artifacts they produce, how they're produced, and how they're used.

2.16 A DATA SCIENCE PERSPECTIVE ON PROACTIVE SCHEDULING AND RESOURCE MANAGEMENT FOR CONNECTED AUTONOMOUS VEHICLES

In this research, Sayyam Malik et al. suggest that ride-sharing and carpooling are currently providing solutions to a number of problems that face contemporary civilizations. The difficulties with excessive oil consumption, transportation congestion, and ineffective time consumption, traffic pollution brought on by excessive vehicle use, and health issues. Because autonomous cars are unmanned and fully autonomous, it is also anticipated that ride-sharing and carpooling will be more effective for them.Many concerns relating to booking rides, location sharing, money handling, and privacy issues need to be improved when unmanned cars take on the task of carpooling, ride-sharing, or car-hailing. We need efficient scheduling strategies to handle all types of emotional difficulties and offer a pollution-free and accident-free environment on the roads for autonomous vehicles in order to deal with these concerns, which largely affect the scheduling of resources. We believe that, among other ways, data science offers a perfect chance to use machine learning models to categories and determine what factors can influence customers to choose a move toward linked autonomous vehicles. We explore autonomous vehicles, vehicle-as-a-service, and their contribution to CO2 emissions reduction in this essay.

2.17 CONNECTING DATA SCIENCE AND PROCESS SCIENCE WHEN PROCESSES MEET BIG DATA

In this study, Wil van der Aalst et al. make a proposal. It has become clear that connecting enormous amounts of event data to extremely dynamic processes is the biggest difficulty as more and more businesses use big data. Events must be properly managed in order to release the potential of event data. related to the administration and control of operational processes. But right now, the main emphasis of big data technologies is on simple analytical tasks like storing and processing data. Rarely do big data efforts priorities streamlining complete processes. We propose improved data science, data technology, and process science integration to address this gap. Process science approaches are model-driven without taking into account the "evidence" hidden in the data, whereas data science approaches are typically process antagonistic. The goal of process mining is to close this gap. This editorial addresses how process mining is related to Big Data technologies, service orientation, and cloud computing as well as the interaction between data science and process science.

2.18 SNP DATA SCIENCE FOR BIPOLAR DISORDER I AND BIPOLAR DISORDER CLASSIFICATION

In this research, Chia-Yen Lee et al. suggest that bipolar disorder I (BD-I) and bipolar disorder II (BD-II) have distinct features and distinct diagnostic criteria, but very different treatment recommendations. BD-II is frequently misdiagnosed in clinical settings as a minor variant of BD-I. In order to improve the diagnosis process, this study employs data science techniques to find the relevant Single Nucleotide Polymorphisms (SNPs) that have a substantial impact on the classifications of BD-I and BD-II. 316 Han Chinese were subjected to screening evaluations and SNP genotyping using the Affymetrix Axiom Genome-Wide TWB Array Plate. According to the data, the classifier created using 23 SNPs had an area under the ROC curve (AUC) level of 0.939, while the classifier created using 42 SNPs had an AUC level of 0.9574, which is only an increase of 1.84 percent. The categorization accuracy rate increased by 3.46 percent.

2.19 THEORY-GUIDED DATA SCIENCE: A NEW APPROACH TO DATA-DRIVEN SCIENTIFIC DISCOVERY

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

In this study, Anuj Karpatne et al. suggest Data science models have had limited success in solving scientific issues involving complicated physical phenomena, despite being successful in a variety of commercial fields. An emerging approach called "theory-guided data science" (TGDS) seeks to take use of the abundance of scientific knowledge to increase how well data science models facilitate scientific discovery. The main goal of TGDS is to establish scientific consistency as a necessary prerequisite for mastering generalizable models. Additionally, by creating models that can be understood by science, TGDS hopes to expand knowledge by gaining fresh domain insights. In fact, the paradigm of TGDS has begun to acquire popularity in a number of scientific fields, including hydrology, turbulence modeling, material discovery, quantum chemistry, bio-medical science, and the study of biomarkers. We officially construct the TGDS paradigm in this study, and we also give taxonomy of TGDS research themes.

2.20 A DATA SCIENCE APPROACH TO EFFECTIVE RESPONSE TO NATURAL DISASTERS

It has been suggested in this study by GHULAM MUDASSIR et al. that natural disasters can kill thousands of people and severely destroy infrastructure and buildings. These occurrences are challenging for the populace as well as for government agencies. Finding a reliable method of evacuation for people first, then rebuilding homes and other infrastructure, are two difficult concerns that need to be handled in particular. It can then be a game changer to effectively overcome those terrible circumstances with an adequate recovery strategy to evacuate people and begin reconstructing damaged areas on a priority basis. In this light, we present DiReCT, a method based on a dynamic optimization model intended to quickly formulate an evacuation plan for a region hit by an earthquake and a decision support system based on a double deep Q network capable of effectively guiding the reconstruction of the affected areas. The latter operates by taking into account the resources available as well as the requirements of the many stakeholders (such as citizens' social benefits and political priorities) involved.

2.21 KERBEROS PROTOCOL WITH IMPROVED KEY AGREEMENT FOR M-HEALTH SECURITY

In this research, P. Thirumoorthy et al. proposes the creation of a wireless sensor network using Internet of Things. (IoT) forecasts a variety of uses in healthcare and cloud computing. This has the potential to yield good outcomes in mobile health care (M-health) and Telerate. Information systems for medicine Internet of Things-based m-health system (IoT) via wireless sensor network (WSN) are a growing study area. the need of contemporary civilization Sensors placed to the patients' bodies that Being linked to a mobile device can make medical services more convenient. The first concern is security. crucial link for efficient operation of the m-health system that shares data of patients in wireless networks in order to protect their privacy This research provided a method for securely transmitting M-health data on wireless networks. Using the planned Kerberos protocol based on key agreements. The patients who were processed Doctors and caregivers can access data saved on a cloud server. The information the suggested method of communication between patients, servers, and doctors is used. Procedure to ensure the secrecy and integrity of authentication The suggested algorithm's efficiency is compared to that of current protocols. The calculation time for 100 devices is only 91 milliseconds.

2.22 IMPROVED ENERGY-USE MULTI-SENSOR OBJECT DETECTION IN WIRELESS SENSOR NETWORKS DANIYAL

Alghazzawi et al. suggest that independent sensor networks capable of sensing physical parameters like temperature, pressure, and humidity are distributed geographically within Wireless Sensor Networks (WSNs). Examples include energy, pressure, and sound. WSNs are resilient and have a secure connection to the physical environment. Data aggregation (DA) is an important part of WSN. helps cut down on energy use (EC). Existing research efforts have discovered DA with a high aggregation rate for WSNs in order to have reliable data. centered on DRINA (In-Network

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

Aggregation Data Routing). Nevertheless, there is none. achieving an effective balance between routing and overhead; however, the EC DA requirements remained unmet. The Bayes Node's detection of things places the same event in specific locations. nodes of sensors (SNs). For effective DA at the sink in a heterogeneous environment, the Scheduling Multi-Sensor Data Synchronization (MSDSS) framework is proposed. Secure and energy-efficient In-Network Aggregation Sensor Data Routing (SEE-INASDR) is developed using a sensor network based on dynamic routing. (DR) structure in WSNs to ensure the security of data transfers. The Polynomial Distribution (BNEPD) method decreased the Energy Drain Rate (EDR), and the poly distribution technique decreased the Communication Overhead (CO) by 39%, as demonstrated by our experimental results. The Network was also improved by the MSDSS structure that was planned. Lifetime (NL) is cut 15% shorter. Additionally, Data Aggregation Routing was improved by 10.5% with this framework. DAR). Last but not least, the SEE-INASDR architecture saved a lot of money. Utilizing a secure and energy-efficient routing protocol (SEERP) results in a 51 percent reduction in EC.

2.23 TASK SCHEDULING IN THE CLOUD BASED ON TWO STAGES DYNAMIC ALGORITHM STRATEGY

M.Deepika et al. presented in this study to improve task allocating performance and reduce illogical task allocation. Allocations in a cloud environment, this research proposes a two-stage technique. Strategy. Initially, a job classifier is driven by the design of a Naive Bayes classifier. The approach is used to classify occupations based on past scheduling data. Certain A number of virtual machines (VMs) of different sorts are built in response. This saves time. the time it takes to generate virtual machines during task allocation Jobs will be available in the next step. are dynamically coordinated with solid virtual machines Several dynamic algorithms are recommended for work allocation, accordingly. The exploratory finding demonstrates that effectively boosts the cloud's task allocation performance and complete the load In compared to conventional solutions, the splints of cloud resources.

2.24 AN INVESTIGATION OF A ROUTING APPROACH FOR IN-NETWORK AGGREGATION IN WIRELESS SENSOR NETWORKS

In this paper, S.SUDHA et al. propose that we may construct data aggregation utilising the Data aggregation and routing techniques can help to lower the cost of Wireless sensor network communication Traffic congestion occurs when one or more of the many sensor nodes detects events. The network should inform the occurrence to save electricity. Only when an event occurs, appropriately. Overhead happens in Because of its poor scalability, InFRA. According to the projected The DRINA algorithm (Data Routing for In-Network Aggregation) decreases communication costs and conserves energy By constructing the routing tree, we optimised the reducing the amount of duplicate routes and removing the superfluous data. The DRINA's performance has been compared to three others. additional protocols known: the Information Fusion-based Role Algorithms for Assignment (InFRA), Shortest Path Tree (SPT), and The algorithm of cantered-at-nearest-source (CNS).

2.25 POLYNOMIAL DISTRIBUTION OF BAYES NODE ENERGY TO IMPROVE WIRELESS SENSOR ROUTING NETWORK

In this paper, Karthikeyan et al. propose a Wireless Sensor Network to monitor and manage the physical environment using a huge number of tiny sensors. low-cost sensor nodes Existing Wireless Sensor Network (WSN) technique given increased latency due to sensed data transfer via continuous data gathering as well as energy consumption To solve the routing problem and decrease energy consumption, The Bayes Node Energy and Polynomial Distribution (BNEPD) approach is presented. In a wireless sensor network, energy-aware routing is used. Energy Distribution at the Bayes Node first distributes sensor nodes that detect a same event (i.e., temperature, etc.) The Bayes rule is used to direct pressure and flow into specified locations. The detection of objects comparable events is completed and delivered to the sink based on the Bayes probability As a consequence, energy usage is reduced. The Polynomial Regression follows. The function is applied to the target object of comparable events evaluated for various sensors. Combined. They are calculated using the lowest

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

and maximum values of object events. Shifted to the sink node finally, the Poly Distribute method distributes the data properly. Nodes of sensing

3. COMPARATIVE ANALYSIS

Title	Techniques & Mechanisms	Parameter Analysis	Future Work
Analyzing Objective and Subjective Data in Social Sciences: Implications for Smart Cities	The aim of this work was to present how data science and machine learning techniques can be used in social science studies in order to maximize the insight gained	It represents a scenario where technology, IoT and artificial intelligence can be used in order to improve current conditions in cities and to implement and monitor large-scale	in the future the app may actively stimulate the improvement of well-being based on known causes of well-being variation; work in this direction is only preliminary at the moment.
BD2K Training Coordinating Center's ERuDIte: The Educational Resource Discovery Index for Data Science	ERuDIte as its own data science project, we have made significant progress on the data collection, data integration, data exploration, and data analysis steps.	In the development of ERuDIte so far, we have designed and implemented a flexible scraping framework, a unified schema, a tagging ontology, a visualization approach for resource exploration, and a collection of automated tagging algorithms.	In future work, we plan to explore active learning techniques to optimize curation and classifier advancement by prioritizing resources that would address key areas where our classifiers need to improve.
Contextual Markup and Mining in Digital Games for Science Learning: Connecting Player Behaviors to Learning Goals	Finally, incorporating the metadata coding into analyses of learning gains provided further insights about specific physics concepts and learning in SURGE.	Based on these findings, we are working to expand and refine the approach in terms of the grain-size of the focal metadata tagging.	In the original SURGE (SURGE Classic), the lack of contextual metadata limited our analyses to pre-post test gains and high-level analyses of game play data
CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories	First, we define trajectories over a well-defined collection of activities, which can be encapsulated and documented, similar to the original sub stages in CRISP-DM. DST	Software development, like many other engineering problems, has a structure that resembles CRISP-DM in many ways (starting with business needs and ending up in deployment.	In future , CRISP-DM still plays an important role as a common framework for setting up and managing data mining projects.
Data Prospecting–A Step Towards Data Intensive Science	Providing discovery engines such as Polaris to support data prospecting can serve as	These primitive features can be subsequently stored as indices independent of any	As part of our future work, we plan to evaluate these technologies as possible

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

	relatively low cost technology enablers as compared to full analytics.	semantic labels and facilitate object-based retrieval within science data	components for our Polaris prototype.
Data Science Analysis and Profile Representation Applied to Secondary Prevention of Acute Coronary Syndrome	Despite having an unbalanced sample size, this method, compared to other conventional analysis methods, demonstrated the ability to obtain relevant knowledge from univariate analysis	Although this does not necessarily imply causality, the review of some of the scientific literature results has agreed with some DSAR results.	Then, we could review the most relevant variable in each group. This interesting path is beyond the scope of the present work but it will deserve particular effort and attention in the near future.
Data Science as Political Action: Grounding Data Science in a Politics of Justice	As a form of political action, data science can no longer be separated from broader analyses of social structures, public policies, and social movements.	By deliberating about political goals and strategies and by developing new methods and norms, data scientists can more rigorously contribute to social justice.	one necessary direction for future research is to develop interdisciplinary frameworks that will help data scientists consider the downstream impacts of their interventions.
Design and Evaluation of a Block-based Environment with a Data Science Context	a comprehensive description and motivation for key features of BlockPy. This paper also continues our evaluation of BlockPy's design and features.	Reflection on design issues involved in BlockPy for developers in the block-based community who would wish to build systems similar to BlockPy.	We now outline future work and directions for BlockPy. Some of this work is technical, some is design decisions that must be revisited in light of evidence collected in its evaluation.
Distributed Data Strategies to Support Large-Scale Data Analysis Across Geo-Distributed Data Centers	The main advantage of this strategy is to separate the storage level from the analysis level. In the second strategy, we consider data replication among different data centers.	We have proposed two strategies to support the approximate big analysis of distributed data across multiple data centers.	. In the future, we will continue to study the effect of the BDMS system to solve streaming data problems.
EOS MLS Science Data Processing System: A Description of Architecture and Capabilities	The SDPS for EOS MLS met all science data processing requirements by assuring the effective cooperation of its components widely dispersed in location and under the responsibility of different institutions	Finally, any problems that may occur are easily localized, diagnosed, and corrected.	In future Fortran standards. MLS restricts the use of Fortran-provided input and output statements in production code; instead MLS relies on appropriate procedures provided in libraries

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

Improving Urban Mobility by Defining a Smart Data Integration Platform	The concept of the smart city includes the use of software solutions to meet the challenges involved in improving citizens' daily lives	The first sources are those provided as open data, many using a Linked Open Data structure, ranging from automatic data streams to more elaborate and even pre-processed data sources.	With regard to future work in this area, we intend to define a systematic and generalized web scraping method that will be based on an underlying reference data model, currently under development, which is directly related to the data found on websites.
Intelligent Data Engineering for Migration to NoSQL Based Secure Environments	The data transformation module can be enhanced as a future work to support other RDBs and NoSQL graph databases.	Transformation rules are generated by this mapping and Sitar engine is used to execute these rules to perform the automatic transformation of SQL Server to Oracle NoSQL.	The data transformation module can be enhanced as a future work to support other RDBs and NoSQL graph databases.
Leveraging Data Science to Combat COVID-19: A Comprehensive Review	Data scientists have been active in addressing the emerging challenges related to COVID-19.	We first summarized publicly available datasets for use by researchers.	in the future as the situation changes. Other difficult questions include the issue of allocation of scarce resources and the trade-offs involved therein.
Parallel or Intersecting Lines? Intelligent Bibliometric for Investigating the Involvement of Data Science in Policy Analysis	However, Anglo/English-speaking countries have the benefit of a similar and valued cultural mindset to the top journals in their field and distinct advantages in terms of language and communication requirements.	We propose an intelligent bibliometric framework for investigating the involvement of data science in policy analysis.	Intelligent bibliometric could and should be expanded, involving intelligent information technologies and bibliometric from many diverse corners
Passing the Data Baton : A Retrospective Analysis on Data Science Work and Workers	Our modeling of data science work and workers is intended to arm visualization researchers with the means to educate, converse, and collaborate with data scientists and others in	Data science and visualization share a common goal of helping people understand their data, offering complementary approaches toward this aim.	Finally, prescriptive modeling identifies a specific intervention that can be taken to modify future outcomes; for example, if the sales manager hires more people, her sales will increase.

	the larger data science community of practice.		
Proactive Scheduling and Resource Management for Connected Autonomous Vehicles: A Data Science Perspective	We have introduced basics of autonomous vehicles, connected autonomous vehicles and car pooling.	By availing ride-sharing in autonomous cars will also reduce stress, and increase productivity while you have no attention on the road and are free to do your work while autonomous cars will drive you toward your destination.	In our future work we plan to work on different aspects that restrict users to opt carpooling mainly the privacy and resource management.
Processes Meet Big Data: Connecting Data Science with Process Science	The special issue "Process Analysis meets Big Data" of the IEEE Transactions on Services	Conventional process mining tools are often deployed on the process owners' premises.	In the future, we hope to see tools for defining Map and Reduce functions on the basis of the business process model data types, of the relations among them and of other semantics-rich context information
SNP Data Science for Classification of Bipolar Disorder I and Bipolar Disorder	GWAS analysis is a powerful tool to explain the degree of contribution of each SNP to diseases; however, for the interactions among multiple SNPs, it is difficult to build a nonlinear model to clarify the complicated effects.	We propose an ensemble method integrating four state-of-the-art algorithms considering the linear and nonlinear structures among SNPs to provide the robust results of identifying the important SNPs and interaction effect.	In future , clinicians have to consider many factors, of which some will be redundant, inconsistent, or noisy, yet some environmental or non-gene-related risk factors, in order to evaluate BD classification and value of information
Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data	We anticipate the deep integration of theory-based and data science to become a quintessential tool for scientific discovery in future research.	While most of the discussion in this paper focuses on supervised learning problems, similar TGDS research themes can be explored for other traditional tasks of data mining, machine learning, and statistics.	We anticipate the deep integration of theory-based and data science to become a quintessential tool for scientific discovery in future research.
Toward Effective Response to Natural Disasters: A Data Science Approach	An integrated framework that, based on data science, can help decision makers to face natural disasters.	Different from other similar algorithms, we are able to manage additional information, needed for evacuation	In the future are further optimization models, exact or approximate, to be employed in order to reduce the

	As first realization, we embed automatic support to evacuation and reconstruction planning	planning and reconstruction, added as attributes to network nodes and arcs.	computational effort presently required by simulations.

4. CONCLUSION

For planning evacuation and reconstruction in the event of natural disasters, we have suggested a comprehensive data science framework termed DiReCT in this research. Contributions made by thisA comprehensive framework that, when based on data science, can assist decision-makers in dealing with natural disasters. As a first realization, we incorporate automatic support for planning evacuation and reconstruction. The description of the GisToGraph algorithm, which creates an enriched underpinning network of any site with information ideally suited for disaster management, particularly in the phases of readiness, response, and reconstruction. The application and validation of the optimization model created by to a genuine outdoor case study, namely the historic city centre of L'Aquila in Italy, for emergency evacuation purposes. Using the double deep Q-network (DDQN) learning method to plan the reconstruction of destroyed structures and their physical
Dependencies in the aftermath of a disaster.Demonstrating the viability and applicability of our approach in a genuine case study, namely the historic district of L'Aquila city.In terms of buildings, intersections, and streets, the network produced by the GisToGraph algorithm depicts the city map. We are able to manage extra data added as characteristics to network nodes and arcs, which makes us superior to other similar algorithms in that it may be used for reconstruction and evacuation planning. We modified the linear optimization model that Arbib et al. Originally created for the evacuation of a building's interior with regard to the evacuation planning model. The model needed to be adjusted for a number of parameters and rescaled to the network at several orders of magnitude. We took into account all important factors, including cost, duration, physical dependencies, social benefits to the affected population, and political priority to take political input into account, when designing the rebuilding. Ultimately, we were able to validate our framework on one of the four sections of the "L'Aquila" city centre.

5. REFERENCES

[1] (Jul. 2020). (Jul. 2020). Social science analysis of objective and subjective data: implications for the smart citizen Retrieved on May 30, 2021. com/news/world-australia-53549936 [online].
[2] (Oct. 2020). ERuDite: The Educational Resource Discovery Index for Data Science, a service of the BD2K Training Coordinating Center. Retrieved on May 30, 2021.[Online]. Available: \shttps://www.bbc.com/news/world-europe-54749509 \s167842 Version 9, 2021Effective Response to Natural Disasters: A Data Science Approach, G. Mudassir et al.
[3] (Aug. 2020). (Aug. 2020). Contextual Markup and Mining for Science Learning in Digital Games: Linking Player Behaviors to Learning Objectives. Retrieved on May 30, 2021.[Online]. The following link is available: https://www.reuters.com/article/usafghanistan-floods-idUSKBN25M148
[4] (Nov. 2020). (Nov. 2020). CRISP-DM Twenty Years Later: Data Science Trajectories from Data Mining Processes. Retrieved on May 30, 2021.[Online].The following link is available: https://www.cbsnews.com/news/hurricane-eta-150-deadmissing-guatemala/
[5] (Nov. 2020). (Nov. 2020). 30 May 2021: Data Prospecting-A Step Towards Data Intensive Science [Online]. Available: \shttps://www.bbc.com/news/world-latin-america-54864963
[6] "Data Science Analysis and Profile Representation Applied to Secondary Prevention of Acute Coronary Syndrome," by R. Abdalla, is found at [6] SpringerPlus, 2016, pp. 1–10, vol. 5, no. 1.
[7] "Data Science as Political Action: Grounding Data Science in a Politics of Justice," Transp. Res. Rec., vol. 2198, no. 1, pp. 152-160, 2010. A. Abdelghany, K. Abdelghany, H. Mahmassani, H. Al-Ahmadi, and W. Alhalabi.

DOI:10.46647/ijetms.2023.v07i02.072 ISSN: 2581-4621

[8] A. Abdelghany, K. Abdelghany, H. Mahmassani, and W. Alhalabi, "Design and Evaluation of a Block-based Environment with a Data Science Context," Eur.
[9] "Distributed Data Strategies to Support Large-Scale Data Analysis Across Geo-Distributed Data Centers," Transp. Res. Rec. J. Transp. Res. Board, vol. 1939, no. 1, pp. 123–132, 2006. A. Abdelghany, K. Abdelghany, H. S. Mahmassani, and S. A. Al-Gadhi.
[10] "IEOS MLS Science Data Processing System: A Description of Architecture and Capabilities," in A View of Operations Research Applications in Italy, vol. 2, by C. Arbib, M. T. Moghaddam, and H. Muccini Springer, Cham, Switzerland, 2018, pp. 115–131.
[11] "Improving Urban Mobility by Defining a Smart Data Integration Platform," Math. Models Methods Appl. Sci., vol. 22, no. 2, August 2012, Art.no. 1230004; N. Bellomo, B. Piccoli, and A. Tosin.
[12] Intelligent Data Engineering for Migration to NoSQL Based Secure Environments, J. D. Brooks, K. Kar, and D. Mendonça, Proc. IEEE Int. Conf. Technol. Homeland Secur. (HST), Nov. 2013, pp. 504-510.
[13] "Leveraging Data Science to Combat COVID-19: A Comprehensive Review," in W. Choi, H. W. Hamacher, and S. Tufekci, "Eur. J. Operar. Res., vol. 35, no. 1, pp. 98-110, 1988.
[14] Parallel or Intersecting Lines? by D. Contreras, T. Blaschke, S. Kienberger, and P. Zeil. Int. J. Disaster Risk Reduction, vol. 8, pp. 125–142, June 2014. "Intelligent Bibliometrics for Investigating the Involvement of Data Science in Policy Analysis."
[15] "Passing the Data Baton: A Retrospective Analysis on Data Science Work and Workers," by M. Dolce and A. Goretti. 2015, August, Bull. Earthq. Eng., vol. 13, no. 8, pp. 2241-2264.
[16] "Proactive Scheduling and Resource Management for Connected Autonomous Vehicles: A Data Science Perspective," Transp. Res. C, Emerg. Technol., vol. 37, Dec. 2013, pp. 193-209. D. C. Duives, W. Daamen, and S. P. Hoogendoorn.
[17] Processes Meet Big Data: Connecting Data Science with Process Science, by M. S. Eid and I. H. El-Adaway, 17 J. Infrastructure Systems, September 2018, Vol. 24, No. 3, Art. No. 04018009.
[18] "SNP Data Science for Classification of Bipolar Disorder I and Bipolar Disorder," in Proc. 23rd iSTEAMS Conf., 2020, pp. 117–186. B. Eze and O. Olaiya.
[19] M. M. Fischer, "Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data," in Spatial Analysis and GeoComputation: Selected Essays, 2006, pages 43–60.
[20] "Toward Effective Response to Natural Disasters: A Data Science Approach," by P. Ghannad, Y.-C. Lee, C. J. Friedland, J. O. Choi, and E. Yang. 2020, July, J. Manage. Eng., vol. 36, no. 4, art.no. 04020038.
[21] ThirumoorthyPalanisamy, D. Alghazzawi, S. Bhatia, A. A. Malibari, P. Dadheeche and colleagues, "Improved energy-based multi-sensor object recognition in wireless sensor networks," Intelligent Automation & Soft Computing, vol. 33, no.1, pp. 227–244, 2022.
[22] P. "Improved key agreement based kerberos protocol for m-health security," Thirumoorthy, K. S. Bhuvaneshwari, C. Kamalanathan, P. Sunita, E. Prabhu et al., Computer Systems Science and Engineering, vol. 42, no.2, pp. 577–587, 2022.
[23] M.Deepika, S.Prabhu, M.Parvathi, and S.Hemalatha, "Cloud Task Scheduling Using Dynamic Algorithm", Gradiva Review Journal, Vol.8, No. 11, pages 53-60, 2022.
[24] S. "A research on routing strategy for in-network aggregation in wireless sensor networks," Sudha, B. Manimegalai, and P. Thirumoorthy, in Proc, IEEE ICCCI, Coimbatore, India, pp.1-4, 2014.
[25] P. N. Thirumoorthy and Thirumoorthy "Bayes node energy polynomial distribution to optimise routing in wireless sensor networks," K. Karthikeyan, PLoS ONE, vol. 10, no. 10: e0138932, pp.1-15, 2015.