| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 10, Issue 8, August 2021 ||

| DOI:10.15680/IJIRSET.2021.1008188|

Social Network Analysis using BIGDATA

Dr. Sowmya Naik P T, Dr. Manjunath R, Dr. K.N. Narasimha Murthy[3], Ambika P.R[4]

Professor, Dept. of CSE., RR Institute f Technology, Visvesvaraya Technological University, Bangalore, India

Professor, Dept. of CSE., City Engineering College, Visveswaraya Technological University, Bangalore, India

Professor, Dept. of CSE., Christ University, Bangalore, India

Assistant Professor, Dept. of CSE., City Engineering College, Visvesvaraya Technological University, Bangalore, India

ABSTRACT: BigData refers to "A Growing Torrent"- the massive amounts and varieties of information, particularly in unstructured form being generated by websites, sensors, social media and others. Big Data recently has become a new ubiquitous term to describe large datasets. The main challenge in handling BigData lies in capturing data, providing a proper representation, managing and analysing data to extract meaningful information in the given domain in a fast and efficient manner. Effective management and analysis of Big Data would bring great benefits and unique opportunities to the users. Big data analytics is a technology-enabled strategy for gaining richer, deeper, and more accurate insights into customers, partners and the business and ultimately gaining competitive advantage. By processing a steady stream of real-time data, organizations can make time-sensitive decisions faster than ever before, monitor emerging trends, course-correct rapidly and jump on new business opportunities. In this paper, we present a brief history of BigData and a life cycle of BigData process model. In this paper , we have investigated many data analytics methods and finally we introduced the *Next Generation Analytics.*

KEYWORDS: Bigdata, Bigdata analytics, Social network analysis.

I. INTRODUCTION

"Big Data" is a term encompassing the use of techniques to capture, process, analyze and visualize potentially large datasets in a reasonable timeframe not accessible to standard IT technologies. By extension, the platform, tools and software used for this purpose are collectively called "Big Data technologies". Big data research is a vast field that connects with many enabling technologies. The new approach redefines the way data is managed and analyzed by leveraging the power of a distributed grid of computing resources. New technologies are emerging to make unstructured data analytics possible and cost-efficient.

1.1 CHARACTERISTICS OF BIGDATA

In this section, BigData characteristics are described using 3v's(Volume, Variety, Velocity) concept.
Volume: Big data implies enormous volumes of data. It used to be employees created data. Now that data is generated by machines, networks, sensors and human interaction on systems like social media the volume of data to be analyzed is massive.
Variety: Variety refers to many types of data both structured and unstructured. We used to store data from sources like spreadsheets and databases. Now data comes in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. This variety of unstructured data creates problems for storage, mining and analyzing data.
Velocity: Big Data Velocity deals with the pace at which data flows in from sources like business processes, machines, networks and human interaction with things like social media sites, mobile devices, etc. The flow of data is massive and continuous.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 10, Issue 8, August 2021 ||

| DOI:10.15680/IJIRSET.2021.1008188|

Social Network Analysis using BIGDATA

Dr. Sowmya Naik P T, Dr. Manjunath R, Dr. K.N. Narasimha Murthy[3], Ambika P.R[4]

Professor, Dept. of CSE., RR Institute f Technology, Visvesvaraya Technological University, Bangalore, India

Professor, Dept. of CSE., City Engineering College, Visveswaraya Technological University, Bangalore, India

Professor, Dept. of CSE., Christ University, Bangalore, India

Assistant Professor, Dept. of CSE., City Engineering College, Visvesvaraya Technological University, Bangalore, India

ABSTRACT: BigData refers to "A Growing Torrent"- the massive amounts and varieties of information, particularly in unstructured form being generated by websites, sensors, social media and others. Big Data recently has become a new ubiquitous term to describe large datasets. The main challenge in handling BigData lies in capturing data, providing a proper representation, managing and analysing data to extract meaningful information in the given domain in a fast and efficient manner. Effective management and analysis of Big Data would bring great benefits and unique opportunities to the users. Big data analytics is a technology-enabled strategy for gaining richer, deeper, and more accurate insights into customers, partners and the business and ultimately gaining competitive advantage. By processing a steady stream of real-time data, organizations can make time-sensitive decisions faster than ever before, monitor emerging trends, course-correct rapidly and jump on new business opportunities. In this paper, we present a brief history of BigData and a life cycle of BigData process model. In this paper , we have investigated many data analytics methods and finally we introduced the *Next Generation Analytics.*

KEYWORDS: Bigdata, Bigdata analytics, Social network analysis.

I. INTRODUCTION

"Big Data" is a term encompassing the use of techniques to capture, process, analyze and visualize potentially large datasets in a reasonable timeframe not accessible to standard IT technologies. By extension, the platform, tools and software used for this purpose are collectively called "Big Data technologies". Big data research is a vast field that connects with many enabling technologies. The new approach redefines the way data is managed and analyzed by leveraging the power of a distributed grid of computing resources. New technologies are emerging to make unstructured data analytics possible and cost-efficient.

1.1 CHARACTERISTICS OF BIGDATA

In this section, BigData characteristics are described using 3v's(Volume, Variety, Velocity) concept.
Volume: Big data implies enormous volumes of data. It used to be employees created data. Now that data is generated by machines, networks, sensors and human interaction on systems like social media the volume of data to be analyzed is massive.
Variety: Variety refers to many types of data both structured and unstructured. We used to store data from sources like spreadsheets and databases. Now data comes in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. This variety of unstructured data creates problems for storage, mining and analyzing data.
Velocity: Big Data Velocity deals with the pace at which data flows in from sources like business processes, machines, networks and human interaction with things like social media sites, mobile devices, etc. The flow of data is massive and continuous.

A Proposed Modern Manufacturing Technique by using Raspberry Pi based Microcontroller System

Volume-69 Issue-5

Year of Publication : 2021

Authors : Sumitha Manoj, Shalini Prasad

DOI :10.14445/22315381/IJETT-V69I5P233

(/assets/Volume-69/Issue-5/IJETT-V69I5P233.pdf)

How to Cite?

Sumitha Manoj, Shalini Prasad, "A Proposed Modern Manufacturing Technique by using Raspberry Pi based Microcontroller System," *International Journal of Engineering Trends and Technology* , vol. 69, no. 5, pp. 257-261, 2021. Crossref, **https://doi.org/10.14445/22315381/IJETT-V69I5P233 (https://doi.org/10.14445/22315381/IJETT-V69I5P233)**

Abstract

In the current global scenario, automation manufacturing systems are playing a vital role in global upward demand in a market where flexible production systems could be essential to meet the customer demand by the manufacturing industries. It improves the reliability, repeatability, and accuracy of the output at any given point in time; but is successful only if the whole manufacturing system is designed to take care of variations in the environment. Internet of Things (IoT) helps in bringing a closed-loop monitoring system that monitors the output even outside the production environment and applies corrective measures to bring control over the variations. The monitoring system also can alert the operator on the wrong output and guide him to correct the settings and/or discord the wrong output. The system explained in this refers to a complex component output from a CNC machine and how the closed-loop automation system can support an error-free output, as well as store the parameters of the output in a cloud for tracking. This concept can be applied to any machine/component/production process with a well-defined input-output matrix.

Keywords

Automation, IoT, Closed Loop System, Medium Scale Industries, Production Process

Reference

[1] Wahlster, W., From industry 1.0 to industry 4.0: towards the 4th industrial revolution, Forum Business Meets Research. 10(2) (2012) 38-45.

[2] Antonio Grieco et al., An industry 4.0 case study in fashion manufacturing, Procedia Manufacturing. 11 (2017) 871-877.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 11 | Nov 2021 www.irjet.net p-ISSN: 2395-0072

OPTIMUM SOLUTION FOR EFFECTIVE MEDICAL WASTE MANAGEMENT

Jyothi. P[1], Vatsala G.A[2], Radha Gupta[3]

[1]Professor & HOD, Dept. of Mathematics, City Engineering College, Bangalore, Karnataka, INDIA
[2]Associate Professor, Dept. of Mathematics, DSATM, Karnataka, Bangalore, INDIA
[3]Professor & HOD, Dept. of Mathematics, Dayananda Sagar college of Engineering, Bangalore, Karnataka, INDIA

---***---

Abstract- *Medical wastes are highly hazardous and it lead people under risk of incurable diseases. Medical waste management is very important as medical wastes are increasing day by day. During Health care activities like treating, diagnosing, immunizing human being and animal & while during research activities wastes are generated. Mainly hospitals, clinics, diagnosis centres, healthcare institutions, dental offices, medical laboratories generate medical wastes. This paper deals with application of GP technique which is one of the optimization technique to control the expenditure of medical waste management by considering various factors like segregation, storage, transportation, treatment, by-products after recycling. This paper also helps to the systematic disposal of medical waste and environmental protection.*

Keywords: By-products, Goal programming, Optimization, Overachievement, Underachievement

1. INTRODUCTION

Clinical squanders can be named general waste, pathological waste, sharps, Infectious squanders, compound squanders and Radio-active squanders, pharmaceutical squanders, pressurized compartments, Genotoxic squanders and so forth. Risk wastes are sharps, chemical, pathological, pressurized container, infectious, pharmaceutical, Genotoxic wastes etc. Medical Wastes are generated by Government hospital, private hospitals, nursing homes, physician's office, Dental office, Dispensaries, mortuaries, blood bank and collection Centre, Animal houses, Laboratories, Research organizations. Categories of humans exposed to chance of contamination are sanitation employees, scientific and paramedical workforce, sufferers and site visitors. Standard Operating procedures of the system are Generation, segregation, collection, storage, transportation, treatment of wastes. Non-hazardous and hazardous wastes are generated in the hospitals/public authority. Non-hazardous wastes are generated in the sit of office, kitchen, administration, hostels, stores, restrooms etc. Hazardous wastes are generated by wards, treatment rooms, ICU, labor room, dressing room, dialysis room, CT scan in hospital itself. WHO estimates 85% of hospital wastes are non-hazardous and 10% is infectious, 5% are non-infectious. In health care there are two type of wastes. First variety is Non risk wastes (75%-90%) & the other variety is Risk wastes (10%-25%). Developed countries produce 1-5 kg/bed/day/with variation among countries. In India 1-2kg/bed/day with variation among government and private establishments. Estimated 506.74tons/day wastes generated, out of which 57%wastes undergoes proper disposal. Medical waste handling involves some best practices. To avoid most medical waste problems Health care workers must adhere a few key best practices. Workers should mindful of the laws, they classify and separate all loss by type into the right shading coded squander holders. Isolated squanders should be marked relying upon its classification and the correct documentation ought to go with all holders during travel. Utilize the clinical garbage removal shading code. Recruit the correct garbage removal organization

2. Review of the literature

Details of hazardous management is explained in [1]. [2] Gives the explanation of medical waste management and control. Case study of medical waste management of Korea is explained detail in [3]. Detailed process and Device for the disposal of medical waste is given in [4]. Detailed study of GP model for Rubber Wood manufacturing factory in Tripura is mentioned in [5]. GP approach for food product distribution of small and medium enterprises is mentioned in [6]. [7] Gives the explanation of how to use analytic network process and goal programming for interdependent information system project selection. [8] Explains the process of development of model based on Linear Programming to solve resource allocation task with emphasis on financial aspects.

3. Objectives of the study

a) Distinguishing deviation in the objectives
b) Boosting the benefit of reused item
c) Environmental assurance by compelling the executives of clinical garbage removal

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903222 |

Designing a Scalable Cloud-based ETL Framework for Big Data Aggregation and Feature Engineering

Archana Bhat[1], Laxmi M C[2], Sriraksha S[3], Sreevidya G[4], Harish R[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The success of businesses largely hinges on their ability to effectively analyze data and derive actionable insights. The Extract-Transform-Load (ETL) process plays a crucial role in achieving these objectives, but it demands considerable effort, particularly when dealing with Big Data. Previous research has struggled to formalize, integrate, and assess the ETL process for Big Data challenges in a scalable and cost-efficient manner. In this paper, we introduce a cloud-based ETL framework designed for the fusion and aggregation of data from various sources. We then outline three scenarios for data aggregation during the ETL process: (i) ETL without aggregation; (ii) aggregation based on specified columns or time intervals; and (iii) aggregation within individual user sessions over varying time frames. The third scenario is particularly beneficial for feature engineering, enabling the definition of features such as "the time since the last occurrence of event X." We assessed the scalability of our framework using Amazon AWS Hadoop clusters.

KEYWORDS: ETL (Extract-Transform-Load), Big Data, Cloud Computing, Data Aggregation, Feature Engineering

I. INTRODUCTION

The prevalence of smart devices, sensors, and social media has resulted in massive amounts of data being generated. Concurrently, consumers have become accustomed to receiving personalized services instantly. Numerous companies, healthcare providers, and institutions have dedicated substantial resources to meet this demand. In the era defined by Big Data, organizations are increasingly under pressure to store and analyze all collected data to maintain a competitive edge in the data-driven market. The intrinsic characteristics of Big Data include volume, velocity, and variety. Recently, other aspects such as variability, veracity, visualization, and value have been recognized as equally important. Together, these elements represent the 7 Vs of Big Data, highlighting the significant complexity faced by those who seek to process, analyze, and leverage this information. Additionally, companies encounter challenges with data analysis outcomes such as joins, transformations, and aggregations—as well as the integration of data with other system components.

To present the data in a practical format, several essential steps must be undertaken [5, 6]: conducting analyses and modeling to uncover all relationships and the relevant business context; gathering data; and performing Extract-Transform-Load (ETL) processes, which typically involve lengthy development and execution phases. After the data is processed and integrated into a data warehouse, it should be accessible for reporting, visualization, analytics, and decision support purposes [1]. According to [7], data warehousing (DW) and business intelligence (BI) face numerous challenges regarding Big Data, including issues related to size, complexity, design and data modeling, computing methodologies, query languages, usability, end-user performance, data consistency and lineage, as well as adapting traditional tools for data exploration, visualization, and analytics and their integration into existing DW/BI solutions and platforms. As proposed in [7], future research directions concerning DW and OLAP in the context of Big Data should strive to provide se...

The usability of data throughout various processing stages is closely linked to its consistency, which presents substantial challenges within Big Data systems. Adopting strong consistency models can severely hinder the scalability and performance of these systems. Conversely, weak and eventual consistency models allow for greater availability and reduced latencies, but they may notably detract from the quality of the information obtained and diminish its usability. In the context of cloud computing, which has become a key paradigm offering a range of cost-effective hardware and software solutions well-suited for deploying Big Data systems, new challenges also arise. These include architectural

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903222 |

Designing a Scalable Cloud-based ETL Framework for Big Data Aggregation and Feature Engineering

Archana Bhat[1], Laxmi M C[2], Sriraksha S[3], Sreevidya G[4], Harish R[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The success of businesses largely hinges on their ability to effectively analyze data and derive actionable insights. The Extract-Transform-Load (ETL) process plays a crucial role in achieving these objectives, but it demands considerable effort, particularly when dealing with Big Data. Previous research has struggled to formalize, integrate, and assess the ETL process for Big Data challenges in a scalable and cost-efficient manner. In this paper, we introduce a cloud-based ETL framework designed for the fusion and aggregation of data from various sources. We then outline three scenarios for data aggregation during the ETL process: (i) ETL without aggregation; (ii) aggregation based on specified columns or time intervals; and (iii) aggregation within individual user sessions over varying time frames. The third scenario is particularly beneficial for feature engineering, enabling the definition of features such as "the time since the last occurrence of event X." We assessed the scalability of our framework using Amazon AWS Hadoop clusters.

KEYWORDS: ETL (Extract-Transform-Load), Big Data, Cloud Computing, Data Aggregation, Feature Engineering

I. INTRODUCTION

The prevalence of smart devices, sensors, and social media has resulted in massive amounts of data being generated. Concurrently, consumers have become accustomed to receiving personalized services instantly. Numerous companies, healthcare providers, and institutions have dedicated substantial resources to meet this demand. In the era defined by Big Data, organizations are increasingly under pressure to store and analyze all collected data to maintain a competitive edge in the data-driven market. The intrinsic characteristics of Big Data include volume, velocity, and variety. Recently, other aspects such as variability, veracity, visualization, and value have been recognized as equally important. Together, these elements represent the 7 Vs of Big Data, highlighting the significant complexity faced by those who seek to process, analyze, and leverage this information. Additionally, companies encounter challenges with data analysis outcomes such as joins, transformations, and aggregations—as well as the integration of data with other system components.

To present the data in a practical format, several essential steps must be undertaken [5, 6]: conducting analyses and modeling to uncover all relationships and the relevant business context; gathering data; and performing Extract-Transform-Load (ETL) processes, which typically involve lengthy development and execution phases. After the data is processed and integrated into a data warehouse, it should be accessible for reporting, visualization, analytics, and decision support purposes [1]. According to [7], data warehousing (DW) and business intelligence (BI) face numerous challenges regarding Big Data, including issues related to size, complexity, design and data modeling, computing methodologies, query languages, usability, end-user performance, data consistency and lineage, as well as adapting traditional tools for data exploration, visualization, and analytics and their integration into existing DW/BI solutions and platforms. As proposed in [7], future research directions concerning DW and OLAP in the context of Big Data should strive to provide se...

The usability of data throughout various processing stages is closely linked to its consistency, which presents substantial challenges within Big Data systems. Adopting strong consistency models can severely hinder the scalability and performance of these systems. Conversely, weak and eventual consistency models allow for greater availability and reduced latencies, but they may notably detract from the quality of the information obtained and diminish its usability. In the context of cloud computing, which has become a key paradigm offering a range of cost-effective hardware and software solutions well-suited for deploying Big Data systems, new challenges also arise. These include architectural

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 8, August 2021 |

| DOI:10.15680/IJMRSET.2021.0408027 |

Real - Time Emotion Recognition using Spiking Neural Networks on Wearable Edge Devices

Sowmya Naik P T[1], Deepika R[2], Channabasappa[3], Nanda Kumar A N[4], Sowmya L D[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,5]

Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[4]

ABSTRACT: Real-time emotion recognition through facial expressions offers significant potential for improving human-computer interactions and personalizing user experiences. This research introduces a novel approach using Spiking Neural Networks (SNN) on wearable edge devices to achieve this. The approach incorporates key technologies like Open Neural Network Exchange (ONNX), Message Queuing Telemetry Transport (MQTT), and Long Short-Term Memory (LSTM) networks to boost the efficiency and accuracy of emotion recognition systems in real-world scenarios. ONNX enables seamless model interchange and deployment across various hardware platforms, ensuring scalability and flexibility. The optimized model conversion for wearable edge devices enhances interoperability and efficiency in real-time emotion recognition. MQTT acts as a lightweight, reliable protocol for smooth data exchange between wearable devices and external systems, facilitating real-time transmission of facial expression data and inference results. This setup supports collaborative processing and decision-making across distributed networks, improving system responsiveness and scalability. Additionally, the use of LSTM networks helps capture temporal dependencies in facial expressions, enhancing the accuracy and robustness of emotion recognition systems by effectively modeling sequential data and long-term dependencies.

KEYWORDS: Spiking Neural Networks (SNNs), Open Neural Network Exchange (ONNX), Long Short-Term Memory (LSTM), and Message Queuing Telemetry Transport (MQTT)

I. INTRODUCTION

Emotion recognition from facial expressions is an intriguing area with applications ranging from human-computer interaction to mental health monitoring .With the advent of wearable technology and the increasing need for real-time processing, there is a growing interest in using advanced neural network architectures, such as SNNs, to perform emotion recognition tasks directly on wearable edge devices .This integration holds the potential to enhance both accuracy and efficiency, providing seamless and personalized user experiences across various domains. The use of Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) play a crucial role in enhancing the performance and efficiency of real-time emotion recognition systems based on SNN .These hardware platforms provide customizable and parallel processing capabilities, making them ideal for implementing complex neural network architectures optimized for low-latency inference on edge devices. By utilizing the inherent parallelism and reconfigure ability of FPGA and the specialized hardware design of ASIC, researchers and developers can create and deploy efficient SNN-based emotion recognition systems that meet the stringent requirements of real-time processing. Additionally, the adoption of standardized formats like ONNX enables seamless model deployment and interoperability across various hardware platforms and software frameworks. ONNX facilitates the conversion and exchange of trained neural network models between different deep learning frameworks, allowing developers to leverage pre-trained models and optimize them for deployment on FPGA and ASIC-based edge devices.

The protocols such as MQTT are essential for enabling seamless interaction between wearable edge devices and protocols such as MQTT are essential for facilitating seamless interaction between wearable edge devices and external systems. MQTT provides lightweight and reliable messaging communication suited for resource-constrained environments, enabling real-time data exchange between edge devices and cloud servers or other edge devices. By using MQTT for data transmission, SNN-based emotion recognition systems can integrate smoothly with existing infrastructure and enable collaborative processing and decision-making across distributed networks. Additionally, advanced neural network architectures such as LSTM and Temporal Convolutional Neural Networks (TCNN) offer complementary capabilities for capturing temporal dependencies and spatial features in facial expressions, improving the accuracy and robustness of emotion recognition systems. LSTM networks are adept at modeling sequential data and capturing long-term dependencies, making them suitable for analysing temporal patterns in facial expressions over

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 8, August 2021 |

| DOI:10.15680/IJMRSET.2021.0408027 |

Real - Time Emotion Recognition using Spiking Neural Networks on Wearable Edge Devices

Sowmya Naik P T[1], Deepika R[2], Channabasappa[3], Nanda Kumar A N[4], Sowmya L D[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,5]

Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[4]

ABSTRACT: Real-time emotion recognition through facial expressions offers significant potential for improving human-computer interactions and personalizing user experiences. This research introduces a novel approach using Spiking Neural Networks (SNN) on wearable edge devices to achieve this. The approach incorporates key technologies like Open Neural Network Exchange (ONNX), Message Queuing Telemetry Transport (MQTT), and Long Short-Term Memory (LSTM) networks to boost the efficiency and accuracy of emotion recognition systems in real-world scenarios. ONNX enables seamless model interchange and deployment across various hardware platforms, ensuring scalability and flexibility. The optimized model conversion for wearable edge devices enhances interoperability and efficiency in real-time emotion recognition. MQTT acts as a lightweight, reliable protocol for smooth data exchange between wearable devices and external systems, facilitating real-time transmission of facial expression data and inference results. This setup supports collaborative processing and decision-making across distributed networks, improving system responsiveness and scalability. Additionally, the use of LSTM networks helps capture temporal dependencies in facial expressions, enhancing the accuracy and robustness of emotion recognition systems by effectively modeling sequential data and long-term dependencies.

KEYWORDS: Spiking Neural Networks (SNNs), Open Neural Network Exchange (ONNX), Long Short-Term Memory (LSTM), and Message Queuing Telemetry Transport (MQTT)

I. INTRODUCTION

Emotion recognition from facial expressions is an intriguing area with applications ranging from human-computer interaction to mental health monitoring .With the advent of wearable technology and the increasing need for real-time processing, there is a growing interest in using advanced neural network architectures, such as SNNs, to perform emotion recognition tasks directly on wearable edge devices .This integration holds the potential to enhance both accuracy and efficiency, providing seamless and personalized user experiences across various domains. The use of Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) play a crucial role in enhancing the performance and efficiency of real-time emotion recognition systems based on SNN .These hardware platforms provide customizable and parallel processing capabilities, making them ideal for implementing complex neural network architectures optimized for low-latency inference on edge devices. By utilizing the inherent parallelism and reconfigure ability of FPGA and the specialized hardware design of ASIC, researchers and developers can create and deploy efficient SNN-based emotion recognition systems that meet the stringent requirements of real-time processing. Additionally, the adoption of standardized formats like ONNX enables seamless model deployment and interoperability across various hardware platforms and software frameworks. ONNX facilitates the conversion and exchange of trained neural network models between different deep learning frameworks, allowing developers to leverage pre-trained models and optimize them for deployment on FPGA and ASIC-based edge devices.

The protocols such as MQTT are essential for enabling seamless interaction between wearable edge devices and protocols such as MQTT are essential for facilitating seamless interaction between wearable edge devices and external systems. MQTT provides lightweight and reliable messaging communication suited for resource-constrained environments, enabling real-time data exchange between edge devices and cloud servers or other edge devices. By using MQTT for data transmission, SNN-based emotion recognition systems can integrate smoothly with existing infrastructure and enable collaborative processing and decision-making across distributed networks. Additionally, advanced neural network architectures such as LSTM and Temporal Convolutional Neural Networks (TCNN) offer complementary capabilities for capturing temporal dependencies and spatial features in facial expressions, improving the accuracy and robustness of emotion recognition systems. LSTM networks are adept at modeling sequential data and capturing long-term dependencies, making them suitable for analysing temporal patterns in facial expressions over

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 8, August 2021 |

| DOI:10.15680/IJMRSET.2021.0408027 |

Real - Time Emotion Recognition using Spiking Neural Networks on Wearable Edge Devices

Sowmya Naik P T[1], Deepika R[2], Channabasappa[3], Nanda Kumar A N[4], Sowmya L D[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,5]

Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[4]

ABSTRACT: Real-time emotion recognition through facial expressions offers significant potential for improving human-computer interactions and personalizing user experiences. This research introduces a novel approach using Spiking Neural Networks (SNN) on wearable edge devices to achieve this. The approach incorporates key technologies like Open Neural Network Exchange (ONNX), Message Queuing Telemetry Transport (MQTT), and Long Short-Term Memory (LSTM) networks to boost the efficiency and accuracy of emotion recognition systems in real-world scenarios. ONNX enables seamless model interchange and deployment across various hardware platforms, ensuring scalability and flexibility. The optimized model conversion for wearable edge devices enhances interoperability and efficiency in real-time emotion recognition. MQTT acts as a lightweight, reliable protocol for smooth data exchange between wearable devices and external systems, facilitating real-time transmission of facial expression data and inference results. This setup supports collaborative processing and decision-making across distributed networks, improving system responsiveness and scalability. Additionally, the use of LSTM networks helps capture temporal dependencies in facial expressions, enhancing the accuracy and robustness of emotion recognition systems by effectively modeling sequential data and long-term dependencies.

KEYWORDS: Spiking Neural Networks (SNNs), Open Neural Network Exchange (ONNX), Long Short-Term Memory (LSTM), and Message Queuing Telemetry Transport (MQTT)

I. INTRODUCTION

Emotion recognition from facial expressions is an intriguing area with applications ranging from human-computer interaction to mental health monitoring .With the advent of wearable technology and the increasing need for real-time processing, there is a growing interest in using advanced neural network architectures, such as SNNs, to perform emotion recognition tasks directly on wearable edge devices .This integration holds the potential to enhance both accuracy and efficiency, providing seamless and personalized user experiences across various domains. The use of Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) play a crucial role in enhancing the performance and efficiency of real-time emotion recognition systems based on SNN .These hardware platforms provide customizable and parallel processing capabilities, making them ideal for implementing complex neural network architectures optimized for low-latency inference on edge devices. By utilizing the inherent parallelism and reconfigure ability of FPGA and the specialized hardware design of ASIC, researchers and developers can create and deploy efficient SNN-based emotion recognition systems that meet the stringent requirements of real-time processing. Additionally, the adoption of standardized formats like ONNX enables seamless model deployment and interoperability across various hardware platforms and software frameworks. ONNX facilitates the conversion and exchange of trained neural network models between different deep learning frameworks, allowing developers to leverage pre-trained models and optimize them for deployment on FPGA and ASIC-based edge devices.

The protocols such as MQTT are essential for enabling seamless interaction between wearable edge devices and protocols such as MQTT are essential for facilitating seamless interaction between wearable edge devices and external systems. MQTT provides lightweight and reliable messaging communication suited for resource-constrained environments, enabling real-time data exchange between edge devices and cloud servers or other edge devices. By using MQTT for data transmission, SNN-based emotion recognition systems can integrate smoothly with existing infrastructure and enable collaborative processing and decision-making across distributed networks. Additionally, advanced neural network architectures such as LSTM and Temporal Convolutional Neural Networks (TCNN) offer complementary capabilities for capturing temporal dependencies and spatial features in facial expressions, improving the accuracy and robustness of emotion recognition systems. LSTM networks are adept at modeling sequential data and capturing long-term dependencies, making them suitable for analysing temporal patterns in facial expressions over

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 8, August 2021 |

| DOI:10.15680/IJMRSET.2021.0408027 |

Real - Time Emotion Recognition using Spiking Neural Networks on Wearable Edge Devices

Sowmya Naik P T[1], Deepika R[2], Channabasappa[3], Nanda Kumar A N[4], Sowmya L D[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,5]

Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[4]

ABSTRACT: Real-time emotion recognition through facial expressions offers significant potential for improving human-computer interactions and personalizing user experiences. This research introduces a novel approach using Spiking Neural Networks (SNN) on wearable edge devices to achieve this. The approach incorporates key technologies like Open Neural Network Exchange (ONNX), Message Queuing Telemetry Transport (MQTT), and Long Short-Term Memory (LSTM) networks to boost the efficiency and accuracy of emotion recognition systems in real-world scenarios. ONNX enables seamless model interchange and deployment across various hardware platforms, ensuring scalability and flexibility. The optimized model conversion for wearable edge devices enhances interoperability and efficiency in real-time emotion recognition. MQTT acts as a lightweight, reliable protocol for smooth data exchange between wearable devices and external systems, facilitating real-time transmission of facial expression data and inference results. This setup supports collaborative processing and decision-making across distributed networks, improving system responsiveness and scalability. Additionally, the use of LSTM networks helps capture temporal dependencies in facial expressions, enhancing the accuracy and robustness of emotion recognition systems by effectively modeling sequential data and long-term dependencies.

KEYWORDS: Spiking Neural Networks (SNNs), Open Neural Network Exchange (ONNX), Long Short-Term Memory (LSTM), and Message Queuing Telemetry Transport (MQTT)

I. INTRODUCTION

Emotion recognition from facial expressions is an intriguing area with applications ranging from human-computer interaction to mental health monitoring .With the advent of wearable technology and the increasing need for real-time processing, there is a growing interest in using advanced neural network architectures, such as SNNs, to perform emotion recognition tasks directly on wearable edge devices .This integration holds the potential to enhance both accuracy and efficiency, providing seamless and personalized user experiences across various domains. The use of Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) play a crucial role in enhancing the performance and efficiency of real-time emotion recognition systems based on SNN .These hardware platforms provide customizable and parallel processing capabilities, making them ideal for implementing complex neural network architectures optimized for low-latency inference on edge devices. By utilizing the inherent parallelism and reconfigure ability of FPGA and the specialized hardware design of ASIC, researchers and developers can create and deploy efficient SNN-based emotion recognition systems that meet the stringent requirements of real-time processing. Additionally, the adoption of standardized formats like ONNX enables seamless model deployment and interoperability across various hardware platforms and software frameworks. ONNX facilitates the conversion and exchange of trained neural network models between different deep learning frameworks, allowing developers to leverage pre-trained models and optimize them for deployment on FPGA and ASIC-based edge devices.

The protocols such as MQTT are essential for enabling seamless interaction between wearable edge devices and protocols such as MQTT are essential for facilitating seamless interaction between wearable edge devices and external systems. MQTT provides lightweight and reliable messaging communication suited for resource-constrained environments, enabling real-time data exchange between edge devices and cloud servers or other edge devices. By using MQTT for data transmission, SNN-based emotion recognition systems can integrate smoothly with existing infrastructure and enable collaborative processing and decision-making across distributed networks. Additionally, advanced neural network architectures such as LSTM and Temporal Convolutional Neural Networks (TCNN) offer complementary capabilities for capturing temporal dependencies and spatial features in facial expressions, improving the accuracy and robustness of emotion recognition systems. LSTM networks are adept at modeling sequential data and capturing long-term dependencies, making them suitable for analysing temporal patterns in facial expressions over

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 10, October 2021 ||

| DOI:10.15680/IJIRSET.2021.1110116 |

Integration of Data Mining Techniques with Geographic Information Systems (GIS) for Enhanced Spatial Analysis of Geographic Data

Savitri Kulkarni[1], Ramesh B[2], Nagashree R A[3], Gangappa Demannavar[4], Vinutha H M[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: This paper explores the integration of data mining techniques with Geographic Information Systems (GIS) to enhance spatial analysis of geographic data. The research begins by delving into the data mining functions specifically designed for geographic data and contrasts them with traditional data mining methods. Geographic data, characterized by spatial relationships and geographical coordinates, demands specialized analytical approaches that differ from conventional data mining techniques. The paper then reviews two predominant research approaches in this domain: the first approach focuses on learning within spatial databases, leveraging large-scale geographic datasets to uncover patterns and trends. The second approach is rooted in spatial statistics, emphasizing the application of statistical methods to spatial data for analytical insights. Through an in-depth comparison of these approaches, we highlight the unique benefits and limitations of each method, while also exploring the commonalities that unify them in their goal to extract meaningful insights from geographic data. These findings, emphasizing the key distinctions between the two approaches, such as the role of spatial relationships in determining patterns. It also identifies shared elements, such as the use of geographic data structures and algorithms, offering a comprehensive perspective on how data mining enhances GIS-based spatial analysis.

KEYWORDS: Spatial Data Mining, Spatial Databases, Rules Induction, Spatial Statistics, Spatial Neighborhood

I. INTRODUCTION

The rise in map production is generating vast amounts of data that are challenging to analyze. To address this, applying data mining techniques to spatial data is increasingly relevant. Unlike traditional data mining, spatial data mining must consider spatial relationships between objects. Spatial data mining is used in decision-making areas such as Geo marketing, environmental studies, and risk analysis. For example, Geo marketing can help retailers determine trade areas and analyze customer profiles based on geographic data. In our project, spatial data mining is applied to traffic risk analysis, using historical accident data combined with information on the road network, population, and buildings. The aim is to identify high-risk areas and understand these risks in their geographic context. Geographic data analysis traditionally relies on conventional statistics and multidimensional methods, which often overlook spatial aspects. Spatial statistics, by contrast, considers the interdependence of nearby observations, forming a distinct field of study. This includes techniques like Geo statistics, Exploratory Spatial Data Analysis (ESDA), and the Geographical Analysis Machine (GAM). Recent advancements have extended multidimensional methods to include spatial contiguity, integrating spatial statistics into data mining. Two notable research teams have made significant contributions: the DB Research Lab at Simon Fraser University with Geo Miner, and Munich University with various algorithms and clustering methods. Additionally, the University of Laval has developed data warehouses for spatial data. This paper will outline data mining methods for Geographic Information Systems, focusing on their application in spatial data analysis. It will cover statistical and database-based approaches, compare their similarities and differences, and discuss current research issues.

II. DEFINITION OF SPATIAL DATA MINING

Spatial Data Mining (SDM) involves extracting knowledge, identifying spatial relationships, and uncovering properties not explicitly stored in databases. It aims to discover implicit patterns and relationships between spatial and non-spatial data. The distinct feature of SDM is its focus on spatial interactions. A geographical database represents a spatio-temporal continuum where the attributes of a location are often interconnected with its surroundings. This emphasizes the importance of spatial relationships in analysis. While temporal aspects of spatial data are significant, they are often

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 10, October 2021 ||

| DOI:10.15680/IJIRSET.2021.1110116 |

Integration of Data Mining Techniques with Geographic Information Systems (GIS) for Enhanced Spatial Analysis of Geographic Data

Savitri Kulkarni[1], Ramesh B[2], Nagashree R A[3], Gangappa Demannavar[4], Vinutha H M[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: This paper explores the integration of data mining techniques with Geographic Information Systems (GIS) to enhance spatial analysis of geographic data. The research begins by delving into the data mining functions specifically designed for geographic data and contrasts them with traditional data mining methods. Geographic data, characterized by spatial relationships and geographical coordinates, demands specialized analytical approaches that differ from conventional data mining techniques. The paper then reviews two predominant research approaches in this domain: the first approach focuses on learning within spatial databases, leveraging large-scale geographic datasets to uncover patterns and trends. The second approach is rooted in spatial statistics, emphasizing the application of statistical methods to spatial data for analytical insights. Through an in-depth comparison of these approaches, we highlight the unique benefits and limitations of each method, while also exploring the commonalities that unify them in their goal to extract meaningful insights from geographic data. These findings, emphasizing the key distinctions between the two approaches, such as the role of spatial relationships in determining patterns. It also identifies shared elements, such as the use of geographic data structures and algorithms, offering a comprehensive perspective on how data mining enhances GIS-based spatial analysis.

KEYWORDS: Spatial Data Mining, Spatial Databases, Rules Induction, Spatial Statistics, Spatial Neighborhood

I. INTRODUCTION

The rise in map production is generating vast amounts of data that are challenging to analyze. To address this, applying data mining techniques to spatial data is increasingly relevant. Unlike traditional data mining, spatial data mining must consider spatial relationships between objects. Spatial data mining is used in decision-making areas such as Geo marketing, environmental studies, and risk analysis. For example, Geo marketing can help retailers determine trade areas and analyze customer profiles based on geographic data. In our project, spatial data mining is applied to traffic risk analysis, using historical accident data combined with information on the road network, population, and buildings. The aim is to identify high-risk areas and understand these risks in their geographic context. Geographic data analysis traditionally relies on conventional statistics and multidimensional methods, which often overlook spatial aspects. Spatial statistics, by contrast, considers the interdependence of nearby observations, forming a distinct field of study. This includes techniques like Geo statistics, Exploratory Spatial Data Analysis (ESDA), and the Geographical Analysis Machine (GAM). Recent advancements have extended multidimensional methods to include spatial contiguity, integrating spatial statistics into data mining. Two notable research teams have made significant contributions: the DB Research Lab at Simon Fraser University with Geo Miner, and Munich University with various algorithms and clustering methods. Additionally, the University of Laval has developed data warehouses for spatial data. This paper will outline data mining methods for Geographic Information Systems, focusing on their application in spatial data analysis. It will cover statistical and database-based approaches, compare their similarities and differences, and discuss current research issues.

II. DEFINITION OF SPATIAL DATA MINING

Spatial Data Mining (SDM) involves extracting knowledge, identifying spatial relationships, and uncovering properties not explicitly stored in databases. It aims to discover implicit patterns and relationships between spatial and non-spatial data. The distinct feature of SDM is its focus on spatial interactions. A geographical database represents a spatio-temporal continuum where the attributes of a location are often interconnected with its surroundings. This emphasizes the importance of spatial relationships in analysis. While temporal aspects of spatial data are significant, they are often

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 8, August 2021 |

| DOI:10.15680/IJMRSET.2021.0408027 |

Real - Time Emotion Recognition using Spiking Neural Networks on Wearable Edge Devices

Sowmya Naik P T[1], Deepika R[2], Channabasappa[3], Nanda Kumar A N[4], Sowmya L D[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,5]

Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[4]

ABSTRACT: Real-time emotion recognition through facial expressions offers significant potential for improving human-computer interactions and personalizing user experiences. This research introduces a novel approach using Spiking Neural Networks (SNN) on wearable edge devices to achieve this. The approach incorporates key technologies like Open Neural Network Exchange (ONNX), Message Queuing Telemetry Transport (MQTT), and Long Short-Term Memory (LSTM) networks to boost the efficiency and accuracy of emotion recognition systems in real-world scenarios. ONNX enables seamless model interchange and deployment across various hardware platforms, ensuring scalability and flexibility. The optimized model conversion for wearable edge devices enhances interoperability and efficiency in real-time emotion recognition. MQTT acts as a lightweight, reliable protocol for smooth data exchange between wearable devices and external systems, facilitating real-time transmission of facial expression data and inference results. This setup supports collaborative processing and decision-making across distributed networks, improving system responsiveness and scalability. Additionally, the use of LSTM networks helps capture temporal dependencies in facial expressions, enhancing the accuracy and robustness of emotion recognition systems by effectively modeling sequential data and long-term dependencies.

KEYWORDS: Spiking Neural Networks (SNNs), Open Neural Network Exchange (ONNX), Long Short-Term Memory (LSTM), and Message Queuing Telemetry Transport (MQTT)

I. INTRODUCTION

Emotion recognition from facial expressions is an intriguing area with applications ranging from human-computer interaction to mental health monitoring .With the advent of wearable technology and the increasing need for real-time processing, there is a growing interest in using advanced neural network architectures, such as SNNs, to perform emotion recognition tasks directly on wearable edge devices .This integration holds the potential to enhance both accuracy and efficiency, providing seamless and personalized user experiences across various domains. The use of Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) play a crucial role in enhancing the performance and efficiency of real-time emotion recognition systems based on SNN .These hardware platforms provide customizable and parallel processing capabilities, making them ideal for implementing complex neural network architectures optimized for low-latency inference on edge devices. By utilizing the inherent parallelism and reconfigure ability of FPGA and the specialized hardware design of ASIC, researchers and developers can create and deploy efficient SNN-based emotion recognition systems that meet the stringent requirements of real-time processing. Additionally, the adoption of standardized formats like ONNX enables seamless model deployment and interoperability across various hardware platforms and software frameworks. ONNX facilitates the conversion and exchange of trained neural network models between different deep learning frameworks, allowing developers to leverage pre-trained models and optimize them for deployment on FPGA and ASIC-based edge devices.

The protocols such as MQTT are essential for enabling seamless interaction between wearable edge devices and protocols such as MQTT are essential for facilitating seamless interaction between wearable edge devices and external systems. MQTT provides lightweight and reliable messaging communication suited for resource-constrained environments, enabling real-time data exchange between edge devices and cloud servers or other edge devices. By using MQTT for data transmission, SNN-based emotion recognition systems can integrate smoothly with existing infrastructure and enable collaborative processing and decision-making across distributed networks. Additionally, advanced neural network architectures such as LSTM and Temporal Convolutional Neural Networks (TCNN) offer complementary capabilities for capturing temporal dependencies and spatial features in facial expressions, improving the accuracy and robustness of emotion recognition systems. LSTM networks are adept at modeling sequential data and capturing long-term dependencies, making them suitable for analysing temporal patterns in facial expressions over

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903222 |

Designing a Scalable Cloud-based ETL Framework for Big Data Aggregation and Feature Engineering

Archana Bhat[1], Laxmi M C[2], Sriraksha S[3], Sreevidya G[4], Harish R[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The success of businesses largely hinges on their ability to effectively analyze data and derive actionable insights. The Extract-Transform-Load (ETL) process plays a crucial role in achieving these objectives, but it demands considerable effort, particularly when dealing with Big Data. Previous research has struggled to formalize, integrate, and assess the ETL process for Big Data challenges in a scalable and cost-efficient manner. In this paper, we introduce a cloud-based ETL framework designed for the fusion and aggregation of data from various sources. We then outline three scenarios for data aggregation during the ETL process: (i) ETL without aggregation; (ii) aggregation based on specified columns or time intervals; and (iii) aggregation within individual user sessions over varying time frames. The third scenario is particularly beneficial for feature engineering, enabling the definition of features such as "the time since the last occurrence of event X." We assessed the scalability of our framework using Amazon AWS Hadoop clusters.

KEYWORDS: ETL (Extract-Transform-Load), Big Data, Cloud Computing, Data Aggregation, Feature Engineering

I. INTRODUCTION

The prevalence of smart devices, sensors, and social media has resulted in massive amounts of data being generated. Concurrently, consumers have become accustomed to receiving personalized services instantly. Numerous companies, healthcare providers, and institutions have dedicated substantial resources to meet this demand. In the era defined by Big Data, organizations are increasingly under pressure to store and analyze all collected data to maintain a competitive edge in the data-driven market. The intrinsic characteristics of Big Data include volume, velocity, and variety. Recently, other aspects such as variability, veracity, visualization, and value have been recognized as equally important. Together, these elements represent the 7 Vs of Big Data, highlighting the significant complexity faced by those who seek to process, analyze, and leverage this information. Additionally, companies encounter challenges with data analysis outcomes such as joins, transformations, and aggregations—as well as the integration of data with other system components.

To present the data in a practical format, several essential steps must be undertaken [5, 6]: conducting analyses and modeling to uncover all relationships and the relevant business context; gathering data; and performing Extract-Transform-Load (ETL) processes, which typically involve lengthy development and execution phases. After the data is processed and integrated into a data warehouse, it should be accessible for reporting, visualization, analytics, and decision support purposes [1]. According to [7], data warehousing (DW) and business intelligence (BI) face numerous challenges regarding Big Data, including issues related to size, complexity, design and data modeling, computing methodologies, query languages, usability, end-user performance, data consistency and lineage, as well as adapting traditional tools for data exploration, visualization, and analytics and their integration into existing DW/BI solutions and platforms. As proposed in [7], future research directions concerning DW and OLAP in the context of Big Data should strive to provide se...

The usability of data throughout various processing stages is closely linked to its consistency, which presents substantial challenges within Big Data systems. Adopting strong consistency models can severely hinder the scalability and performance of these systems. Conversely, weak and eventual consistency models allow for greater availability and reduced latencies, but they may notably detract from the quality of the information obtained and diminish its usability. In the context of cloud computing, which has become a key paradigm offering a range of cost-effective hardware and software solutions well-suited for deploying Big Data systems, new challenges also arise. These include architectural

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 10, October 2021 ||

| DOI:10.15680/IJIRSET.2021.1110116 |

Integration of Data Mining Techniques with Geographic Information Systems (GIS) for Enhanced Spatial Analysis of Geographic Data

Savitri Kulkarni[1], Ramesh B[2], Nagashree R A[3], Gangappa Demannavar[4], Vinutha H M[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: This paper explores the integration of data mining techniques with Geographic Information Systems (GIS) to enhance spatial analysis of geographic data. The research begins by delving into the data mining functions specifically designed for geographic data and contrasts them with traditional data mining methods. Geographic data, characterized by spatial relationships and geographical coordinates, demands specialized analytical approaches that differ from conventional data mining techniques. The paper then reviews two predominant research approaches in this domain: the first approach focuses on learning within spatial databases, leveraging large-scale geographic datasets to uncover patterns and trends. The second approach is rooted in spatial statistics, emphasizing the application of statistical methods to spatial data for analytical insights. Through an in-depth comparison of these approaches, we highlight the unique benefits and limitations of each method, while also exploring the commonalities that unify them in their goal to extract meaningful insights from geographic data. These findings, emphasizing the key distinctions between the two approaches, such as the role of spatial relationships in determining patterns. It also identifies shared elements, such as the use of geographic data structures and algorithms, offering a comprehensive perspective on how data mining enhances GIS-based spatial analysis.

KEYWORDS: Spatial Data Mining, Spatial Databases, Rules Induction, Spatial Statistics, Spatial Neighborhood

I. INTRODUCTION

The rise in map production is generating vast amounts of data that are challenging to analyze. To address this, applying data mining techniques to spatial data is increasingly relevant. Unlike traditional data mining, spatial data mining must consider spatial relationships between objects. Spatial data mining is used in decision-making areas such as Geo marketing, environmental studies, and risk analysis. For example, Geo marketing can help retailers determine trade areas and analyze customer profiles based on geographic data. In our project, spatial data mining is applied to traffic risk analysis, using historical accident data combined with information on the road network, population, and buildings. The aim is to identify high-risk areas and understand these risks in their geographic context. Geographic data analysis traditionally relies on conventional statistics and multidimensional methods, which often overlook spatial aspects. Spatial statistics, by contrast, considers the interdependence of nearby observations, forming a distinct field of study. This includes techniques like Geo statistics, Exploratory Spatial Data Analysis (ESDA), and the Geographical Analysis Machine (GAM). Recent advancements have extended multidimensional methods to include spatial contiguity, integrating spatial statistics into data mining. Two notable research teams have made significant contributions: the DB Research Lab at Simon Fraser University with Geo Miner, and Munich University with various algorithms and clustering methods. Additionally, the University of Laval has developed data warehouses for spatial data. This paper will outline data mining methods for Geographic Information Systems, focusing on their application in spatial data analysis. It will cover statistical and database-based approaches, compare their similarities and differences, and discuss current research issues.

II. DEFINITION OF SPATIAL DATA MINING

Spatial Data Mining (SDM) involves extracting knowledge, identifying spatial relationships, and uncovering properties not explicitly stored in databases. It aims to discover implicit patterns and relationships between spatial and non-spatial data. The distinct feature of SDM is its focus on spatial interactions. A geographical database represents a spatio-temporal continuum where the attributes of a location are often interconnected with its surroundings. This emphasizes the importance of spatial relationships in analysis. While temporal aspects of spatial data are significant, they are often

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 10, October 2021 ||

| DOI:10.15680/IJIRSET.2021.1110116 |

Integration of Data Mining Techniques with Geographic Information Systems (GIS) for Enhanced Spatial Analysis of Geographic Data

Savitri Kulkarni[1], Ramesh B[2], Nagashree R A[3], Gangappa Demannavar[4], Vinutha H M[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: This paper explores the integration of data mining techniques with Geographic Information Systems (GIS) to enhance spatial analysis of geographic data. The research begins by delving into the data mining functions specifically designed for geographic data and contrasts them with traditional data mining methods. Geographic data, characterized by spatial relationships and geographical coordinates, demands specialized analytical approaches that differ from conventional data mining techniques. The paper then reviews two predominant research approaches in this domain: the first approach focuses on learning within spatial databases, leveraging large-scale geographic datasets to uncover patterns and trends. The second approach is rooted in spatial statistics, emphasizing the application of statistical methods to spatial data for analytical insights. Through an in-depth comparison of these approaches, we highlight the unique benefits and limitations of each method, while also exploring the commonalities that unify them in their goal to extract meaningful insights from geographic data. These findings, emphasizing the key distinctions between the two approaches, such as the role of spatial relationships in determining patterns. It also identifies shared elements, such as the use of geographic data structures and algorithms, offering a comprehensive perspective on how data mining enhances GIS-based spatial analysis.

KEYWORDS: Spatial Data Mining, Spatial Databases, Rules Induction, Spatial Statistics, Spatial Neighborhood

I. INTRODUCTION

The rise in map production is generating vast amounts of data that are challenging to analyze. To address this, applying data mining techniques to spatial data is increasingly relevant. Unlike traditional data mining, spatial data mining must consider spatial relationships between objects. Spatial data mining is used in decision-making areas such as Geo marketing, environmental studies, and risk analysis. For example, Geo marketing can help retailers determine trade areas and analyze customer profiles based on geographic data. In our project, spatial data mining is applied to traffic risk analysis, using historical accident data combined with information on the road network, population, and buildings. The aim is to identify high-risk areas and understand these risks in their geographic context. Geographic data analysis traditionally relies on conventional statistics and multidimensional methods, which often overlook spatial aspects. Spatial statistics, by contrast, considers the interdependence of nearby observations, forming a distinct field of study. This includes techniques like Geo statistics, Exploratory Spatial Data Analysis (ESDA), and the Geographical Analysis Machine (GAM). Recent advancements have extended multidimensional methods to include spatial contiguity, integrating spatial statistics into data mining. Two notable research teams have made significant contributions: the DB Research Lab at Simon Fraser University with Geo Miner, and Munich University with various algorithms and clustering methods. Additionally, the University of Laval has developed data warehouses for spatial data. This paper will outline data mining methods for Geographic Information Systems, focusing on their application in spatial data analysis. It will cover statistical and database-based approaches, compare their similarities and differences, and discuss current research issues.

II. DEFINITION OF SPATIAL DATA MINING

Spatial Data Mining (SDM) involves extracting knowledge, identifying spatial relationships, and uncovering properties not explicitly stored in databases. It aims to discover implicit patterns and relationships between spatial and non-spatial data. The distinct feature of SDM is its focus on spatial interactions. A geographical database represents a spatio-temporal continuum where the attributes of a location are often interconnected with its surroundings. This emphasizes the importance of spatial relationships in analysis. While temporal aspects of spatial data are significant, they are often

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903222 |

Designing a Scalable Cloud-based ETL Framework for Big Data Aggregation and Feature Engineering

Archana Bhat[1], Laxmi M C[2], Sriraksha S[3], Sreevidya G[4], Harish R[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The success of businesses largely hinges on their ability to effectively analyze data and derive actionable insights. The Extract-Transform-Load (ETL) process plays a crucial role in achieving these objectives, but it demands considerable effort, particularly when dealing with Big Data. Previous research has struggled to formalize, integrate, and assess the ETL process for Big Data challenges in a scalable and cost-efficient manner. In this paper, we introduce a cloud-based ETL framework designed for the fusion and aggregation of data from various sources. We then outline three scenarios for data aggregation during the ETL process: (i) ETL without aggregation; (ii) aggregation based on specified columns or time intervals; and (iii) aggregation within individual user sessions over varying time frames. The third scenario is particularly beneficial for feature engineering, enabling the definition of features such as "the time since the last occurrence of event X." We assessed the scalability of our framework using Amazon AWS Hadoop clusters.

KEYWORDS: ETL (Extract-Transform-Load), Big Data, Cloud Computing, Data Aggregation, Feature Engineering

I. INTRODUCTION

The prevalence of smart devices, sensors, and social media has resulted in massive amounts of data being generated. Concurrently, consumers have become accustomed to receiving personalized services instantly. Numerous companies, healthcare providers, and institutions have dedicated substantial resources to meet this demand. In the era defined by Big Data, organizations are increasingly under pressure to store and analyze all collected data to maintain a competitive edge in the data-driven market. The intrinsic characteristics of Big Data include volume, velocity, and variety. Recently, other aspects such as variability, veracity, visualization, and value have been recognized as equally important. Together, these elements represent the 7 Vs of Big Data, highlighting the significant complexity faced by those who seek to process, analyze, and leverage this information. Additionally, companies encounter challenges with data analysis outcomes such as joins, transformations, and aggregations—as well as the integration of data with other system components.

To present the data in a practical format, several essential steps must be undertaken [5, 6]: conducting analyses and modeling to uncover all relationships and the relevant business context; gathering data; and performing Extract-Transform-Load (ETL) processes, which typically involve lengthy development and execution phases. After the data is processed and integrated into a data warehouse, it should be accessible for reporting, visualization, analytics, and decision support purposes [1]. According to [7], data warehousing (DW) and business intelligence (BI) face numerous challenges regarding Big Data, including issues related to size, complexity, design and data modeling, computing methodologies, query languages, usability, end-user performance, data consistency and lineage, as well as adapting traditional tools for data exploration, visualization, and analytics and their integration into existing DW/BI solutions and platforms. As proposed in [7], future research directions concerning DW and OLAP in the context of Big Data should strive to provide se...

The usability of data throughout various processing stages is closely linked to its consistency, which presents substantial challenges within Big Data systems. Adopting strong consistency models can severely hinder the scalability and performance of these systems. Conversely, weak and eventual consistency models allow for greater availability and reduced latencies, but they may notably detract from the quality of the information obtained and diminish its usability. In the context of cloud computing, which has become a key paradigm offering a range of cost-effective hardware and software solutions well-suited for deploying Big Data systems, new challenges also arise. These include architectural

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 8, August 2021 |

| DOI:10.15680/IJMRSET.2021.0408027 |

Real - Time Emotion Recognition using Spiking Neural Networks on Wearable Edge Devices

Sowmya Naik P T[1], Deepika R[2], Channabasappa[3], Nanda Kumar A N[4], Sowmya L D[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,5]

Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[4]

ABSTRACT: Real-time emotion recognition through facial expressions offers significant potential for improving human-computer interactions and personalizing user experiences. This research introduces a novel approach using Spiking Neural Networks (SNN) on wearable edge devices to achieve this. The approach incorporates key technologies like Open Neural Network Exchange (ONNX), Message Queuing Telemetry Transport (MQTT), and Long Short-Term Memory (LSTM) networks to boost the efficiency and accuracy of emotion recognition systems in real-world scenarios. ONNX enables seamless model interchange and deployment across various hardware platforms, ensuring scalability and flexibility. The optimized model conversion for wearable edge devices enhances interoperability and efficiency in real-time emotion recognition. MQTT acts as a lightweight, reliable protocol for smooth data exchange between wearable devices and external systems, facilitating real-time transmission of facial expression data and inference results. This setup supports collaborative processing and decision-making across distributed networks, improving system responsiveness and scalability. Additionally, the use of LSTM networks helps capture temporal dependencies in facial expressions, enhancing the accuracy and robustness of emotion recognition systems by effectively modeling sequential data and long-term dependencies.

KEYWORDS: Spiking Neural Networks (SNNs), Open Neural Network Exchange (ONNX), Long Short-Term Memory (LSTM), and Message Queuing Telemetry Transport (MQTT)

I. INTRODUCTION

Emotion recognition from facial expressions is an intriguing area with applications ranging from human-computer interaction to mental health monitoring .With the advent of wearable technology and the increasing need for real-time processing, there is a growing interest in using advanced neural network architectures, such as SNNs, to perform emotion recognition tasks directly on wearable edge devices .This integration holds the potential to enhance both accuracy and efficiency, providing seamless and personalized user experiences across various domains. The use of Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC) play a crucial role in enhancing the performance and efficiency of real-time emotion recognition systems based on SNN .These hardware platforms provide customizable and parallel processing capabilities, making them ideal for implementing complex neural network architectures optimized for low-latency inference on edge devices. By utilizing the inherent parallelism and reconfigure ability of FPGA and the specialized hardware design of ASIC, researchers and developers can create and deploy efficient SNN-based emotion recognition systems that meet the stringent requirements of real-time processing. Additionally, the adoption of standardized formats like ONNX enables seamless model deployment and interoperability across various hardware platforms and software frameworks. ONNX facilitates the conversion and exchange of trained neural network models between different deep learning frameworks, allowing developers to leverage pre-trained models and optimize them for deployment on FPGA and ASIC-based edge devices.

The protocols such as MQTT are essential for enabling seamless interaction between wearable edge devices and protocols such as MQTT are essential for facilitating seamless interaction between wearable edge devices and external systems. MQTT provides lightweight and reliable messaging communication suited for resource-constrained environments, enabling real-time data exchange between edge devices and cloud servers or other edge devices. By using MQTT for data transmission, SNN-based emotion recognition systems can integrate smoothly with existing infrastructure and enable collaborative processing and decision-making across distributed networks. Additionally, advanced neural network architectures such as LSTM and Temporal Convolutional Neural Networks (TCNN) offer complementary capabilities for capturing temporal dependencies and spatial features in facial expressions, improving the accuracy and robustness of emotion recognition systems. LSTM networks are adept at modeling sequential data and capturing long-term dependencies, making them suitable for analysing temporal patterns in facial expressions over

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 10, October 2021 ||

| DOI:10.15680/IJIRSET.2021.1110116 |

Integration of Data Mining Techniques with Geographic Information Systems (GIS) for Enhanced Spatial Analysis of Geographic Data

Savitri Kulkarni[1], Ramesh B[2], Nagashree R A[3], Gangappa Demannavar[4], Vinutha H M[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: This paper explores the integration of data mining techniques with Geographic Information Systems (GIS) to enhance spatial analysis of geographic data. The research begins by delving into the data mining functions specifically designed for geographic data and contrasts them with traditional data mining methods. Geographic data, characterized by spatial relationships and geographical coordinates, demands specialized analytical approaches that differ from conventional data mining techniques. The paper then reviews two predominant research approaches in this domain: the first approach focuses on learning within spatial databases, leveraging large-scale geographic datasets to uncover patterns and trends. The second approach is rooted in spatial statistics, emphasizing the application of statistical methods to spatial data for analytical insights. Through an in-depth comparison of these approaches, we highlight the unique benefits and limitations of each method, while also exploring the commonalities that unify them in their goal to extract meaningful insights from geographic data. These findings, emphasizing the key distinctions between the two approaches, such as the role of spatial relationships in determining patterns. It also identifies shared elements, such as the use of geographic data structures and algorithms, offering a comprehensive perspective on how data mining enhances GIS-based spatial analysis.

KEYWORDS: Spatial Data Mining, Spatial Databases, Rules Induction, Spatial Statistics, Spatial Neighborhood

I. INTRODUCTION

The rise in map production is generating vast amounts of data that are challenging to analyze. To address this, applying data mining techniques to spatial data is increasingly relevant. Unlike traditional data mining, spatial data mining must consider spatial relationships between objects. Spatial data mining is used in decision-making areas such as Geo marketing, environmental studies, and risk analysis. For example, Geo marketing can help retailers determine trade areas and analyze customer profiles based on geographic data. In our project, spatial data mining is applied to traffic risk analysis, using historical accident data combined with information on the road network, population, and buildings. The aim is to identify high-risk areas and understand these risks in their geographic context. Geographic data analysis traditionally relies on conventional statistics and multidimensional methods, which often overlook spatial aspects. Spatial statistics, by contrast, considers the interdependence of nearby observations, forming a distinct field of study. This includes techniques like Geo statistics, Exploratory Spatial Data Analysis (ESDA), and the Geographical Analysis Machine (GAM). Recent advancements have extended multidimensional methods to include spatial contiguity, integrating spatial statistics into data mining. Two notable research teams have made significant contributions: the DB Research Lab at Simon Fraser University with Geo Miner, and Munich University with various algorithms and clustering methods. Additionally, the University of Laval has developed data warehouses for spatial data. This paper will outline data mining methods for Geographic Information Systems, focusing on their application in spatial data analysis. It will cover statistical and database-based approaches, compare their similarities and differences, and discuss current research issues.

II. DEFINITION OF SPATIAL DATA MINING

Spatial Data Mining (SDM) involves extracting knowledge, identifying spatial relationships, and uncovering properties not explicitly stored in databases. It aims to discover implicit patterns and relationships between spatial and non-spatial data. The distinct feature of SDM is its focus on spatial interactions. A geographical database represents a spatio-temporal continuum where the attributes of a location are often interconnected with its surroundings. This emphasizes the importance of spatial relationships in analysis. While temporal aspects of spatial data are significant, they are often

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903222 |

Designing a Scalable Cloud-based ETL Framework for Big Data Aggregation and Feature Engineering

Archana Bhat[1], Laxmi M C[2], Sriraksha S[3], Sreevidya G[4], Harish R[5]

Assistant Professors, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The success of businesses largely hinges on their ability to effectively analyze data and derive actionable insights. The Extract-Transform-Load (ETL) process plays a crucial role in achieving these objectives, but it demands considerable effort, particularly when dealing with Big Data. Previous research has struggled to formalize, integrate, and assess the ETL process for Big Data challenges in a scalable and cost-efficient manner. In this paper, we introduce a cloud-based ETL framework designed for the fusion and aggregation of data from various sources. We then outline three scenarios for data aggregation during the ETL process: (i) ETL without aggregation; (ii) aggregation based on specified columns or time intervals; and (iii) aggregation within individual user sessions over varying time frames. The third scenario is particularly beneficial for feature engineering, enabling the definition of features such as "the time since the last occurrence of event X." We assessed the scalability of our framework using Amazon AWS Hadoop clusters.

KEYWORDS: ETL (Extract-Transform-Load), Big Data, Cloud Computing, Data Aggregation, Feature Engineering

I. INTRODUCTION

The prevalence of smart devices, sensors, and social media has resulted in massive amounts of data being generated. Concurrently, consumers have become accustomed to receiving personalized services instantly. Numerous companies, healthcare providers, and institutions have dedicated substantial resources to meet this demand. In the era defined by Big Data, organizations are increasingly under pressure to store and analyze all collected data to maintain a competitive edge in the data-driven market. The intrinsic characteristics of Big Data include volume, velocity, and variety. Recently, other aspects such as variability, veracity, visualization, and value have been recognized as equally important. Together, these elements represent the 7 Vs of Big Data, highlighting the significant complexity faced by those who seek to process, analyze, and leverage this information. Additionally, companies encounter challenges with data analysis outcomes such as joins, transformations, and aggregations—as well as the integration of data with other system components.

To present the data in a practical format, several essential steps must be undertaken [5, 6]: conducting analyses and modeling to uncover all relationships and the relevant business context; gathering data; and performing Extract-Transform-Load (ETL) processes, which typically involve lengthy development and execution phases. After the data is processed and integrated into a data warehouse, it should be accessible for reporting, visualization, analytics, and decision support purposes [1]. According to [7], data warehousing (DW) and business intelligence (BI) face numerous challenges regarding Big Data, including issues related to size, complexity, design and data modeling, computing methodologies, query languages, usability, end-user performance, data consistency and lineage, as well as adapting traditional tools for data exploration, visualization, and analytics and their integration into existing DW/BI solutions and platforms. As proposed in [7], future research directions concerning DW and OLAP in the context of Big Data should strive to provide se...

The usability of data throughout various processing stages is closely linked to its consistency, which presents substantial challenges within Big Data systems. Adopting strong consistency models can severely hinder the scalability and performance of these systems. Conversely, weak and eventual consistency models allow for greater availability and reduced latencies, but they may notably detract from the quality of the information obtained and diminish its usability. In the context of cloud computing, which has become a key paradigm offering a range of cost-effective hardware and software solutions well-suited for deploying Big Data systems, new challenges also arise. These include architectural

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901045 |

Enhancing Credit Card Fraud Detection using Machine Learning and Blockchain: A Novel Approach with Anomaly Detection and Deep Learning Techniques

Madhavi J Kulkarni[1], Katuri Sravanthi Ravi[2], Geethanjali V[3], S K L Narayana[4], Krishna K S[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Credit card fraud represents a major obstacle for financial institutions and consumers globally. Traditional methods of fraud detection frequently struggle to keep pace with the increasingly sophisticated nature of fraudulent activities. This research introduces a novel strategy that leverages cutting-edge machine learning (ML) techniques and blockchain technology to improve the capabilities of fraud detection. The study analyzes a comprehensive dataset that includes a wide range of transactional features, such as transaction amount, location, and time. Several ML models are utilized, including anomaly detection, supervised learning approaches (Random Forest and Gradient Boosting with ensemble methods), and deep learning techniques (custom Recurrent Neural Networks in conjunction with xgboost). Initial experiments yield encouraging accuracy results, with anomaly detection achieving nearly 99.9% accuracy, 99.8% recall, 99.9% sensitivity, and an F1 score of 99.9% in fraud identification.

KEYWORDS: Transactions, unified, unchangeable, precision, distributed

I. INTRODUCTION

The rise of digital commerce has significantly changed the landscape of financial transactions, providing unparalleled convenience. However, this swift move toward digitization has also created an environment conducive to fraudulent activities, with credit card fraud representing a major risk for both consumers and financial institutions. To address this growing issue, advanced fraud detection systems are essential.

In traditional credit card transactions, a complex relationship exists among cardholders, merchants, and financial institutions. Each transaction creates a set of data, which includes transaction amount, location, time, and cardholder information. By scrutinizing these data points, institutions can uncover patterns that suggest fraudulent behavior. Nonetheless, the continuously evolving tactics of fraud require the adoption of more sophisticated strategies. Blockchain technology, known for its characteristics of immutability, transparency, and decentralization, presents a promising option for improving fraud detection mechanisms. By fusing the inventive potential of block chain technology with conventional machine learning approaches, this research seeks to create a reliable credit card fraud detection system. We aim to develop a model that can reliably detect fraudulent transactions, minimize financial losses, and preserve consumer confidence by utilizing the advantages of both approaches. A crucial field of study has been credit card fraud detection (CCFD) because of the growth in fraudulent activity and the number of financial transactions. Technological developments such as blockchain, machine learning, and federated learning have created new opportunities to enhance the efficacy and precision of fraud detection systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score.

II. LITERATURE REVIEW

Recent developments in machine learning, blockchain, and federated learning have created new opportunities to boost the efficacy and accuracy of credit card fraud detection (CCFD) systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score. This paper, written by Pushpita Chatterjee, Debashis Das, and Danda Rawat, examines how federated learning and blockchain technology might be combined to improve the security and precision of fraud detection systems using a private credit card transaction dataset. These two technologies are combined by the approach to produce The key findings of this robust fraud detection framework reveal that the integrated approach markedly enhances detection accuracy while safeguarding user privacy. The reported performance metrics are an accuracy of 95.3% and an F1 score of 94.8%. In the research conducted by Baabdullah, Tahani;

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901045 |

Enhancing Credit Card Fraud Detection using Machine Learning and Blockchain: A Novel Approach with Anomaly Detection and Deep Learning Techniques

Madhavi J Kulkarni[1], Katuri Sravanthi Ravi[2], Geethanjali V[3], S K L Narayana[4], Krishna K S[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Credit card fraud represents a major obstacle for financial institutions and consumers globally. Traditional methods of fraud detection frequently struggle to keep pace with the increasingly sophisticated nature of fraudulent activities. This research introduces a novel strategy that leverages cutting-edge machine learning (ML) techniques and blockchain technology to improve the capabilities of fraud detection. The study analyzes a comprehensive dataset that includes a wide range of transactional features, such as transaction amount, location, and time. Several ML models are utilized, including anomaly detection, supervised learning approaches (Random Forest and Gradient Boosting with ensemble methods), and deep learning techniques (custom Recurrent Neural Networks in conjunction with xgboost). Initial experiments yield encouraging accuracy results, with anomaly detection achieving nearly 99.9% accuracy, 99.8% recall, 99.9% sensitivity, and an F1 score of 99.9% in fraud identification.

KEYWORDS: Transactions, unified, unchangeable, precision, distributed

I. INTRODUCTION

The rise of digital commerce has significantly changed the landscape of financial transactions, providing unparalleled convenience. However, this swift move toward digitization has also created an environment conducive to fraudulent activities, with credit card fraud representing a major risk for both consumers and financial institutions. To address this growing issue, advanced fraud detection systems are essential.

In traditional credit card transactions, a complex relationship exists among cardholders, merchants, and financial institutions. Each transaction creates a set of data, which includes transaction amount, location, time, and cardholder information. By scrutinizing these data points, institutions can uncover patterns that suggest fraudulent behavior. Nonetheless, the continuously evolving tactics of fraud require the adoption of more sophisticated strategies. Blockchain technology, known for its characteristics of immutability, transparency, and decentralization, presents a promising option for improving fraud detection mechanisms. By fusing the inventive potential of block chain technology with conventional machine learning approaches, this research seeks to create a reliable credit card fraud detection system. We aim to develop a model that can reliably detect fraudulent transactions, minimize financial losses, and preserve consumer confidence by utilizing the advantages of both approaches. A crucial field of study has been credit card fraud detection (CCFD) because of the growth in fraudulent activity and the number of financial transactions. Technological developments such as blockchain, machine learning, and federated learning have created new opportunities to enhance the efficacy and precision of fraud detection systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score.

II. LITERATURE REVIEW

Recent developments in machine learning, blockchain, and federated learning have created new opportunities to boost the efficacy and accuracy of credit card fraud detection (CCFD) systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score. This paper, written by Pushpita Chatterjee, Debashis Das, and Danda Rawat, examines how federated learning and blockchain technology might be combined to improve the security and precision of fraud detection systems using a private credit card transaction dataset. These two technologies are combined by the approach to produce The key findings of this robust fraud detection framework reveal that the integrated approach markedly enhances detection accuracy while safeguarding user privacy. The reported performance metrics are an accuracy of 95.3% and an F1 score of 94.8%. In the research conducted by Baabdullah, Tahani;

ATTESTED COPY

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901045 |

Enhancing Credit Card Fraud Detection using Machine Learning and Blockchain: A Novel Approach with Anomaly Detection and Deep Learning Techniques

Madhavi J Kulkarni[1], Katuri Sravanthi Ravi[2], Geethanjali V[3], S K L Narayana[4], Krishna K S[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Credit card fraud represents a major obstacle for financial institutions and consumers globally. Traditional methods of fraud detection frequently struggle to keep pace with the increasingly sophisticated nature of fraudulent activities. This research introduces a novel strategy that leverages cutting-edge machine learning (ML) techniques and blockchain technology to improve the capabilities of fraud detection. The study analyzes a comprehensive dataset that includes a wide range of transactional features, such as transaction amount, location, and time. Several ML models are utilized, including anomaly detection, supervised learning approaches (Random Forest and Gradient Boosting with ensemble methods), and deep learning techniques (custom Recurrent Neural Networks in conjunction with xgboost). Initial experiments yield encouraging accuracy results, with anomaly detection achieving nearly 99.9% accuracy, 99.8% recall, 99.9% sensitivity, and an F1 score of 99.9% in fraud identification.

KEYWORDS: Transactions, unified, unchangeable, precision, distributed

I. INTRODUCTION

The rise of digital commerce has significantly changed the landscape of financial transactions, providing unparalleled convenience. However, this swift move toward digitization has also created an environment conducive to fraudulent activities, with credit card fraud representing a major risk for both consumers and financial institutions. To address this growing issue, advanced fraud detection systems are essential.

In traditional credit card transactions, a complex relationship exists among cardholders, merchants, and financial institutions. Each transaction creates a set of data, which includes transaction amount, location, time, and cardholder information. By scrutinizing these data points, institutions can uncover patterns that suggest fraudulent behavior. Nonetheless, the continuously evolving tactics of fraud require the adoption of more sophisticated strategies. Blockchain technology, known for its characteristics of immutability, transparency, and decentralization, presents a promising option for improving fraud detection mechanisms. By fusing the inventive potential of block chain technology with conventional machine learning approaches, this research seeks to create a reliable credit card fraud detection system. We aim to develop a model that can reliably detect fraudulent transactions, minimize financial losses, and preserve consumer confidence by utilizing the advantages of both approaches. A crucial field of study has been credit card fraud detection (CCFD) because of the growth in fraudulent activity and the number of financial transactions. Technological developments such as blockchain, machine learning, and federated learning have created new opportunities to enhance the efficacy and precision of fraud detection systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score.

II. LITERATURE REVIEW

Recent developments in machine learning, blockchain, and federated learning have created new opportunities to boost the efficacy and accuracy of credit card fraud detection (CCFD) systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score. This paper, written by Pushpita Chatterjee, Debashis Das, and Danda Rawat, examines how federated learning and blockchain technology might be combined to improve the security and precision of fraud detection systems using a private credit card transaction dataset. These two technologies are combined by the approach to produce The key findings of this robust fraud detection framework reveal that the integrated approach markedly enhances detection accuracy while safeguarding user privacy. The reported performance metrics are an accuracy of 95.3% and an F1 score of 94.8%. In the research conducted by Baabdullah, Tahani;

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901045 |

Enhancing Credit Card Fraud Detection using Machine Learning and Blockchain: A Novel Approach with Anomaly Detection and Deep Learning Techniques

Madhavi J Kulkarni[1], Katuri Sravanthi Ravi[2], Geethanjali V[3], S K L Narayana[4], Krishna K S[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Credit card fraud represents a major obstacle for financial institutions and consumers globally. Traditional methods of fraud detection frequently struggle to keep pace with the increasingly sophisticated nature of fraudulent activities. This research introduces a novel strategy that leverages cutting-edge machine learning (ML) techniques and blockchain technology to improve the capabilities of fraud detection. The study analyzes a comprehensive dataset that includes a wide range of transactional features, such as transaction amount, location, and time. Several ML models are utilized, including anomaly detection, supervised learning approaches (Random Forest and Gradient Boosting with ensemble methods), and deep learning techniques (custom Recurrent Neural Networks in conjunction with xgboost). Initial experiments yield encouraging accuracy results, with anomaly detection achieving nearly 99.9% accuracy, 99.8% recall, 99.9% sensitivity, and an F1 score of 99.9% in fraud identification.

KEYWORDS: Transactions, unified, unchangeable, precision, distributed

I. INTRODUCTION

The rise of digital commerce has significantly changed the landscape of financial transactions, providing unparalleled convenience. However, this swift move toward digitization has also created an environment conducive to fraudulent activities, with credit card fraud representing a major risk for both consumers and financial institutions. To address this growing issue, advanced fraud detection systems are essential.

In traditional credit card transactions, a complex relationship exists among cardholders, merchants, and financial institutions. Each transaction creates a set of data, which includes transaction amount, location, time, and cardholder information. By scrutinizing these data points, institutions can uncover patterns that suggest fraudulent behavior. Nonetheless, the continuously evolving tactics of fraud require the adoption of more sophisticated strategies. Blockchain technology, known for its characteristics of immutability, transparency, and decentralization, presents a promising option for improving fraud detection mechanisms. By fusing the inventive potential of block chain technology with conventional machine learning approaches, this research seeks to create a reliable credit card fraud detection system. We aim to develop a model that can reliably detect fraudulent transactions, minimize financial losses, and preserve consumer confidence by utilizing the advantages of both approaches. A crucial field of study has been credit card fraud detection (CCFD) because of the growth in fraudulent activity and the number of financial transactions. Technological developments such as blockchain, machine learning, and federated learning have created new opportunities to enhance the efficacy and precision of fraud detection systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score.

II. LITERATURE REVIEW

Recent developments in machine learning, blockchain, and federated learning have created new opportunities to boost the efficacy and accuracy of credit card fraud detection (CCFD) systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score. This paper, written by Pushpita Chatterjee, Debashis Das, and Danda Rawat, examines how federated learning and blockchain technology might be combined to improve the security and precision of fraud detection systems using a private credit card transaction dataset. These two technologies are combined by the approach to produce The key findings of this robust fraud detection framework reveal that the integrated approach markedly enhances detection accuracy while safeguarding user privacy. The reported performance metrics are an accuracy of 95.3% and an F1 score of 94.8%. In the research conducted by Baabdullah, Tahani;

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 2, February 2021 |

| DOI:10.15680/IJMRSET.2021.0402006 |

Autonomous Machine-to-Machine (M2m) Collaboration for Indoor Search and Rescue Missions: A Multimodal Approach using Ground Vehicles and Drones

S Harsha Vardhan U[1], Vijay Kumar[2], Sampath H P[3], Rakesh Y D[4], Anil Kumar R[5]

Assistant Professors, Department of Mechanical Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Machine-to-Machine (M2M) collaboration has unlocked new possibilities for systems to autonomously and efficiently solve complex technical challenges. In recent years, M2M applications have expanded across various industries, particularly in areas where autonomous decision-making and real-time collaboration are crucial. This research project presents the development of an M2M system that enables autonomous collaboration between two ground vehicles and a drone, managed by a central base station. The system is designed specifically for indoor search and rescue operations, where traditional human intervention may be challenging or hazardous. The primary function of the drone in this system is to perform an autonomous aerial search, while the ground vehicles handle obstacle navigation and ground-level data collection. The base station serves as the control hub, managing both the drone's flight path and the vehicles' navigation, ensuring real-time coordination. A machine learning model is employed to optimize the drone's flight path, achieving near-perfect accuracy in navigating the indoor environment. This accuracy improves significantly with an increase in training data, highlighting the importance of robust training in machine learning-based navigation systems. For ground vehicles, machine learning models are also implemented to navigate complex environments with obstacles. During field testing, the integrated system demonstrated high levels of accuracy and efficiency in coordinating tasks between the drone and ground vehicles. This successful implementation of M2M collaboration showcases its potential in critical applications like search and rescue, offering a glimpse into future advancements in autonomous systems for emergency response operations. The project's findings underscore the growing importance of M2M technology in enhancing the effectiveness of autonomous systems across various domains.

KEYWORDS: Search and Rescue, Image Processing, Navigation Systems, Autonomous Systems, and Object Detection

I. INTRODUCTION

Machine-to-machine (M2M) collaboration became possible with the introduction of Cyber-Physical Systems (CPS) and the Internet of Things (IoT), together with significant advancements in networking, cloud computing, and machine learning. Road transportation, system automation, surveillance, search and rescue operations, and many other situations make use of the idea of M2M collaboration. The Internet of Things (IoT) offers a network of items with integrated sensors, software, and other technology designed to link and exchange data with other systems and devices. Machine learning is one of the key components of M2M collaboration. Researchers have used machine learning for a variety of purposes in a number of efforts. In one project, Chirra and his colleagues used face state observation to apply a behavioral technique to identify driver fatigue. In order to ascertain whether or not the driver is sleepy, it used the Viola-Jones identification algorithm to identify the face and extract the eye region from the facial image. To determine whether the driver is asleep or not, a Convolutional Neural Network (CNN) classifier employs a SoftMax layer. Better accuracy was reported by this work as compared to conventional CNN. Munawar and associates devised a categorization scheme for flooding management to organize the several technologies under examination [5]. They discovered that there were not many hybrid models—which fuse machine learning and image processing—for flood control. Furthermore, it was discovered that there was little use of machine learning-based techniques in the aftermath of a disaster. Therefore, to enable efficient and comprehensive catastrophe management throughout all phases, future efforts must concentrate on integrating disaster management expertise, image processing techniques, and machine learning tools.

Semeraro and colleagues provided a survey of the literature on machine-learning approaches used in human-robot collaboration. In order to do work grouping based on the kinds of collaborative activities, evaluation measures, and

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901045 |

Enhancing Credit Card Fraud Detection using Machine Learning and Blockchain: A Novel Approach with Anomaly Detection and Deep Learning Techniques

Madhavi J Kulkarni[1], Katuri Sravanthi Ravi[2], Geethanjali V[3], S K L Narayana[4], Krishna K S[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Credit card fraud represents a major obstacle for financial institutions and consumers globally. Traditional methods of fraud detection frequently struggle to keep pace with the increasingly sophisticated nature of fraudulent activities. This research introduces a novel strategy that leverages cutting-edge machine learning (ML) techniques and blockchain technology to improve the capabilities of fraud detection. The study analyzes a comprehensive dataset that includes a wide range of transactional features, such as transaction amount, location, and time. Several ML models are utilized, including anomaly detection, supervised learning approaches (Random Forest and Gradient Boosting with ensemble methods), and deep learning techniques (custom Recurrent Neural Networks in conjunction with xgboost). Initial experiments yield encouraging accuracy results, with anomaly detection achieving nearly 99.9% accuracy, 99.8% recall, 99.9% sensitivity, and an F1 score of 99.9% in fraud identification.

KEYWORDS: Transactions, unified, unchangeable, precision, distributed

I. INTRODUCTION

The rise of digital commerce has significantly changed the landscape of financial transactions, providing unparalleled convenience. However, this swift move toward digitization has also created an environment conducive to fraudulent activities, with credit card fraud representing a major risk for both consumers and financial institutions. To address this growing issue, advanced fraud detection systems are essential.

In traditional credit card transactions, a complex relationship exists among cardholders, merchants, and financial institutions. Each transaction creates a set of data, which includes transaction amount, location, time, and cardholder information. By scrutinizing these data points, institutions can uncover patterns that suggest fraudulent behavior. Nonetheless, the continuously evolving tactics of fraud require the adoption of more sophisticated strategies. Blockchain technology, known for its characteristics of immutability, transparency, and decentralization, presents a promising option for improving fraud detection mechanisms. By fusing the inventive potential of block chain technology with conventional machine learning approaches, this research seeks to create a reliable credit card fraud detection system. We aim to develop a model that can reliably detect fraudulent transactions, minimize financial losses, and preserve consumer confidence by utilizing the advantages of both approaches. A crucial field of study has been credit card fraud detection (CCFD) because of the growth in fraudulent activity and the number of financial transactions. Technological developments such as blockchain, machine learning, and federated learning have created new opportunities to enhance the efficacy and precision of fraud detection systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score.

II. LITERATURE REVIEW

Recent developments in machine learning, blockchain, and federated learning have created new opportunities to boost the efficacy and accuracy of credit card fraud detection (CCFD) systems. The important research contributions in this field are examined in this overview of the literature, with an emphasis on the datasets utilized, the techniques, the main conclusions, and performance metrics like accuracy and F1 score. This paper, written by Pushpita Chatterjee, Debashis Das, and Danda Rawat, examines how federated learning and blockchain technology might be combined to improve the security and precision of fraud detection systems using a private credit card transaction dataset. These two technologies are combined by the approach to produce The key findings of this robust fraud detection framework reveal that the integrated approach markedly enhances detection accuracy while safeguarding user privacy. The reported performance metrics are an accuracy of 95.3% and an F1 score of 94.8%. In the research conducted by Baabdullah, Tahani;

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 2, February 2021 |

| DOI:10.15680/IJMRSET.2021.0402006 |

Autonomous Machine-to-Machine (M2m) Collaboration for Indoor Search and Rescue Missions: A Multimodal Approach using Ground Vehicles and Drones

S Harsha Vardhan U[1], Vijay Kumar[2], Sampath H P[3], Rakesh Y D[4], Anil Kumar R[5]

Assistant Professors, Department of Mechanical Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Machine-to-Machine (M2M) collaboration has unlocked new possibilities for systems to autonomously and efficiently solve complex technical challenges. In recent years, M2M applications have expanded across various industries, particularly in areas where autonomous decision-making and real-time collaboration are crucial. This research project presents the development of an M2M system that enables autonomous collaboration between two ground vehicles and a drone, managed by a central base station. The system is designed specifically for indoor search and rescue operations, where traditional human intervention may be challenging or hazardous. The primary function of the drone in this system is to perform an autonomous aerial search, while the ground vehicles handle obstacle navigation and ground-level data collection. The base station serves as the control hub, managing both the drone's flight path and the vehicles' navigation, ensuring real-time coordination. A machine learning model is employed to optimize the drone's flight path, achieving near-perfect accuracy in navigating the indoor environment. This accuracy improves significantly with an increase in training data, highlighting the importance of robust training in machine learning-based navigation systems. For ground vehicles, machine learning models are also implemented to navigate complex environments with obstacles. During field testing, the integrated system demonstrated high levels of accuracy and efficiency in coordinating tasks between the drone and ground vehicles. This successful implementation of M2M collaboration showcases its potential in critical applications like search and rescue, offering a glimpse into future advancements in autonomous systems for emergency response operations. The project's findings underscore the growing importance of M2M technology in enhancing the effectiveness of autonomous systems across various domains.

KEYWORDS: Search and Rescue, Image Processing, Navigation Systems, Autonomous Systems, and Object Detection

I. INTRODUCTION

Machine-to-machine (M2M) collaboration became possible with the introduction of Cyber-Physical Systems (CPS) and the Internet of Things (IoT), together with significant advancements in networking, cloud computing, and machine learning. Road transportation, system automation, surveillance, search and rescue operations, and many other situations make use of the idea of M2M collaboration. The Internet of Things (IoT) offers a network of items with integrated sensors, software, and other technology designed to link and exchange data with other systems and devices. Machine learning is one of the key components of M2M collaboration. Researchers have used machine learning for a variety of purposes in a number of efforts. In one project, Chirra and his colleagues used face state observation to apply a behavioral technique to identify driver fatigue. In order to ascertain whether or not the driver is sleepy, it used the Viola-Jones identification algorithm to identify the face and extract the eye region from the facial image. To determine whether the driver is asleep or not, a Convolutional Neural Network (CNN) classifier employs a SoftMax layer. Better accuracy was reported by this work as compared to conventional CNN. Munawar and associates devised a categorization scheme for flooding management to organize the several technologies under examination [5]. They discovered that there were not many hybrid models—which fuse machine learning and image processing—for flood control. Furthermore, it was discovered that there was little use of machine learning-based techniques in the aftermath of a disaster. Therefore, to enable efficient and comprehensive catastrophe management throughout all phases, future efforts must concentrate on integrating disaster management expertise, image processing techniques, and machine learning tools.

Semeraro and colleagues provided a survey of the literature on machine-learning approaches used in human-robot collaboration. In order to do work grouping based on the kinds of collaborative activities, evaluation measures, and

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 12, December 2021 ||

| DOI:10.15680/IJIRSET.2021.1012188 |

Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance: A Comprehensive Analysis of Online Portfolio Strategies

Gayatri Annasagaram[1], Vanitha G R[2], Nagendra Naik K[3], Rekha R[4], Jyothi[5]

Assistant Professors, Department of Mathematics, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The pattern matching method has long been recognized as a traditional and effective classification approach within the domain of online portfolio selection strategies. This article delves into the core elements of this method, focusing on two key components: the assessment of similarity and the selection of similarity sets. To address the growing complexity of financial markets, this study introduces an enhanced approach known as the Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance (PMDI). The aim is to optimize portfolio selection by integrating a more comprehensive analysis of historical stock price patterns. PMDI leverages a dual-information framework, combining both direction and distance metrics to create a more accurate and refined pattern-matching process. The method incorporates various well-known techniques, such as the Euclidean distance, Chebyshev distance, and the correlation coefficient, to capture the essential information embedded in historical stock price movements. By analyzing both the directional trends (the way prices move over time) and the relative distances between stock prices, PMDI identifies a similarity set that closely mirrors the desired investment strategy. This similarity set forms the foundation of the portfolio selection algorithm. The novelty of PMDI lies in its ability to not only assess patterns based on price direction but also to quantify how far stock prices deviate from one another. This dual approach enhances the system's capability to filter out noise and irrelevant data, thereby selecting only the most relevant patterns for further processing. As a result, the system can make more informed decisions when constructing an optimal investment portfolio, balancing between maximizing returns and minimizing risks. To validate the effectiveness of PMDI, extensive experiments were carried out using two real-world stock market datasets. The results of these tests demonstrated that PMDI outperformed other traditional algorithms, achieving a superior balance between returns and risk management. In comparison to single-metric models, which may focus only on direction or distance, PMDI's dual-information framework proves to be more robust and adaptable to the ever-changing dynamics of financial markets. In the increasingly volatile world of finance, having a strategy that can dynamically respond to market shifts is crucial. The PMDI method offers a significant improvement over existing portfolio selection techniques by utilizing a richer set of data inputs. Its ability to handle both directional trends and price deviations makes it particularly well-suited for real-world financial environments, where adaptability and precision are key to achieving sustainable long-term returns. As financial markets evolve, methods like PMDI will likely become essential tools for investors seeking to maximize performance while effectively managing risk.

KEYWORDS: Portfolio Selection, Pattern Matching, Dual Information, Financial Markets, Risk Management

I. INTRODUCTION

The allocation of limited capital in global financial markets has consistently garnered significant interest from both investors and researchers, particularly in the quest to optimize returns. The intricate nature and high interconnectivity of these markets complicate the ability to forecast market movements. Additionally, elements such as financial innovation, diverse trading strategies, and regulatory frameworks contribute to the prevailing uncertainty within the market. Consequently, devising effective capital allocation strategies within this multifaceted environment has emerged as a pressing global issue. Portfolio Selection (PS) stands as a fundamental endeavour within the financial sector, focusing on the distribution of capital across a defined array of assets to meet specific investment objectives. The roots of this challenge can be traced back to the mean-variance theory introduced by Markowitz in 1952, which employs mean-variance analysis to identify the optimal investment portfolio by balancing anticipated returns against associated risks. Nevertheless, as market dynamics evolve and investor preferences diversify, conventional portfolio selection techniques have increasingly revealed their shortcomings.

In recent years, the swift advancement of artificial intelligence and machine learning has introduced innovative approaches to address portfolio selection challenges. These technologies possess the capability to scrutinize historical data, uncovering patterns and trends that facilitate more accurate predictions of future market behaviour. By processing

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 2, February 2021 |

| DOI:10.15680/IJMRSET.2021.0402006 |

Autonomous Machine-to-Machine (M2m) Collaboration for Indoor Search and Rescue Missions: A Multimodal Approach using Ground Vehicles and Drones

S Harsha Vardhan U[1], Vijay Kumar[2], Sampath H P[3], Rakesh Y D[4], Anil Kumar R[5]

Assistant Professors, Department of Mechanical Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Machine-to-Machine (M2M) collaboration has unlocked new possibilities for systems to autonomously and efficiently solve complex technical challenges. In recent years, M2M applications have expanded across various industries, particularly in areas where autonomous decision-making and real-time collaboration are crucial. This research project presents the development of an M2M system that enables autonomous collaboration between two ground vehicles and a drone, managed by a central base station. The system is designed specifically for indoor search and rescue operations, where traditional human intervention may be challenging or hazardous. The primary function of the drone in this system is to perform an autonomous aerial search, while the ground vehicles handle obstacle navigation and ground-level data collection. The base station serves as the control hub, managing both the drone's flight path and the vehicles' navigation, ensuring real-time coordination. A machine learning model is employed to optimize the drone's flight path, achieving near-perfect accuracy in navigating the indoor environment. This accuracy improves significantly with an increase in training data, highlighting the importance of robust training in machine learning-based navigation systems. For ground vehicles, machine learning models are also implemented to navigate complex environments with obstacles. During field testing, the integrated system demonstrated high levels of accuracy and efficiency in coordinating tasks between the drone and ground vehicles. This successful implementation of M2M collaboration showcases its potential in critical applications like search and rescue, offering a glimpse into future advancements in autonomous systems for emergency response operations. The project's findings underscore the growing importance of M2M technology in enhancing the effectiveness of autonomous systems across various domains.

KEYWORDS: Search and Rescue, Image Processing, Navigation Systems, Autonomous Systems, and Object Detection

I. INTRODUCTION

Machine-to-machine (M2M) collaboration became possible with the introduction of Cyber-Physical Systems (CPS) and the Internet of Things (IoT), together with significant advancements in networking, cloud computing, and machine learning. Road transportation, system automation, surveillance, search and rescue operations, and many other situations make use of the idea of M2M collaboration. The Internet of Things (IoT) offers a network of items with integrated sensors, software, and other technology designed to link and exchange data with other systems and devices. Machine learning is one of the key components of M2M collaboration. Researchers have used machine learning for a variety of purposes in a number of efforts. In one project, Chirra and his colleagues used face state observation to apply a behavioral technique to identify driver fatigue. In order to ascertain whether or not the driver is sleepy, it used the Viola-Jones identification algorithm to identify the face and extract the eye region from the facial image. To determine whether the driver is asleep or not, a Convolutional Neural Network (CNN) classifier employs a SoftMax layer. Better accuracy was reported by this work as compared to conventional CNN. Munawar and associates devised a categorization scheme for flooding management to organize the several technologies under examination [5]. They discovered that there were not many hybrid models—which fuse machine learning and image processing—for flood control. Furthermore, it was discovered that there was little use of machine learning-based techniques in the aftermath of a disaster. Therefore, to enable efficient and comprehensive catastrophe management throughout all phases, future efforts must concentrate on integrating disaster management expertise, image processing techniques, and machine learning tools.

Semeraro and colleagues provided a survey of the literature on machine-learning approaches used in human-robot collaboration. In order to do work grouping based on the kinds of collaborative activities, evaluation measures, and

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 12, December 2021 ||

| DOI:10.15680/IJIRSET.2021.1012188 |

Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance: A Comprehensive Analysis of Online Portfolio Strategies

Gayatri Annasagaram[1], Vanitha G R[2], Nagendra Naik K[3], Rekha R[4], Jyothi[5]

Assistant Professors, Department of Mathematics, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The pattern matching method has long been recognized as a traditional and effective classification approach within the domain of online portfolio selection strategies. This article delves into the core elements of this method, focusing on two key components: the assessment of similarity and the selection of similarity sets. To address the growing complexity of financial markets, this study introduces an enhanced approach known as the Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance (PMDI). The aim is to optimize portfolio selection by integrating a more comprehensive analysis of historical stock price patterns. PMDI leverages a dual-information framework, combining both direction and distance metrics to create a more accurate and refined pattern-matching process. The method incorporates various well-known techniques, such as the Euclidean distance, Chebyshev distance, and the correlation coefficient, to capture the essential information embedded in historical stock price movements. By analyzing both the directional trends (the way prices move over time) and the relative distances between stock prices, PMDI identifies a similarity set that closely mirrors the desired investment strategy. This similarity set forms the foundation of the portfolio selection algorithm. The novelty of PMDI lies in its ability to not only assess patterns based on price direction but also to quantify how far stock prices deviate from one another. This dual approach enhances the system's capability to filter out noise and irrelevant data, thereby selecting only the most relevant patterns for further processing. As a result, the system can make more informed decisions when constructing an optimal investment portfolio, balancing between maximizing returns and minimizing risks. To validate the effectiveness of PMDI, extensive experiments were carried out using two real-world stock market datasets. The results of these tests demonstrated that PMDI outperformed other traditional algorithms, achieving a superior balance between returns and risk management. In comparison to single-metric models, which may focus only on direction or distance, PMDI's dual-information framework proves to be more robust and adaptable to the ever-changing dynamics of financial markets. In the increasingly volatile world of finance, having a strategy that can dynamically respond to market shifts is crucial. The PMDI method offers a significant improvement over existing portfolio selection techniques by utilizing a richer set of data inputs. Its ability to handle both directional trends and price deviations makes it particularly well-suited for real-world financial environments, where adaptability and precision are key to achieving sustainable long-term returns. As financial markets evolve, methods like PMDI will likely become essential tools for investors seeking to maximize performance while effectively managing risk.

KEYWORDS: Portfolio Selection, Pattern Matching, Dual Information, Financial Markets, Risk Management

I. INTRODUCTION

The allocation of limited capital in global financial markets has consistently garnered significant interest from both investors and researchers, particularly in the quest to optimize returns. The intricate nature and high interconnectivity of these markets complicate the ability to forecast market movements. Additionally, elements such as financial innovation, diverse trading strategies, and regulatory frameworks contribute to the prevailing uncertainty within the market. Consequently, devising effective capital allocation strategies within this multifaceted environment has emerged as a pressing global issue. Portfolio Selection (PS) stands as a fundamental endeavour within the financial sector, focusing on the distribution of capital across a defined array of assets to meet specific investment objectives. The roots of this challenge can be traced back to the mean-variance theory introduced by Markowitz in 1952, which employs mean-variance analysis to identify the optimal investment portfolio by balancing anticipated returns against associated risks. Nevertheless, as market dynamics evolve and investor preferences diversify, conventional portfolio selection techniques have increasingly revealed their shortcomings.

In recent years, the swift advancement of artificial intelligence and machine learning has introduced innovative approaches to address portfolio selection challenges. These technologies possess the capability to scrutinize historical data, uncovering patterns and trends that facilitate more accurate predictions of future market behaviour. By processing

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 7.122|| A Monthly Peer Reviewed & Referred Journal |

||Volume 10, Issue 1, January 2021||

DOI:10.15662/IJAREEIE.2021.1001028

Discovering Spatial Correlations through Spatial Data Mining: A Framework for Geospatial Data Analytics

P Rajshekar[1], Sujatha[2], Nagashree G[3], Ashwini Hindiholi[4], Sunitha N[5]

Professor, Department of Chemistry, Engineering College, Bengaluru, Karnataka, India[1]

Professor, Department of Physics, City Engineering College, Bengaluru, Karnataka, India[2]

Assistant Professors, Department of Basic Sciences, City Engineering College, Bengaluru, Karnataka, India[3,4,5]

ABSTRACT: Spatial data holds immense potential for uncovering geographical correlations, which can provide valuable insights into patterns across various fields such as urban planning, environmental monitoring, and social media analysis. Spatial Data Mining (SDM) offers powerful methods for analyzing these correlations, and this paper introduces a novel approach to enhance this analytical process. Specifically, we propose a framework for geospatial data analytics that integrates the G statistic and ZG score computations for spatial correlation discovery. The proposed algorithm, Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD), is designed to perform comprehensive spatial correlation analysis while also incorporating Principal Component Analysis (PCA) to uncover trends and reduce dimensionality. By applying SDM-SCD to Twitter data, we demonstrate how spatial correlations can be effectively analyzed based on the origin location of tweets and specific keywords. The analysis process begins by collecting geolocated Twitter data based on a set of predefined keywords. The algorithm then computes the G statistic and ZG score to detect significant spatial correlations between different geographic regions. These correlations are mapped and analyzed to identify hotspots, clusters, or patterns of interest. By applying PCA, the algorithm reduces the complexity of the data while retaining the most relevant components, enabling more efficient trend identification and visualization. Our experimental results demonstrate that the SDM-SCD framework effectively identifies spatial correlations in Twitter data, providing meaningful insights into how topics or trends evolve across geographical regions. For instance, the algorithm can reveal the spatial spread of discussions related to public health, natural disasters, or political events, offering valuable real-time insights for decision-makers. The SDM-SCD algorithm and framework provide a robust approach for geospatial data analysis, with the ability to uncover spatial patterns and trends from large-scale social media data. By leveraging the power of spatial data mining and advanced statistical techniques, this framework opens new avenues for exploring spatial correlations in various fields, with potential applications in social media analytics, public health, and beyond. The experimental validation highlights its efficacy in providing actionable insights from geospatial data.\

KEYWORDS: Spatial Data Mining, Geospatial Data Analysis, G Statistic, ZG Score Computations, Spatial Correlation Analysis

I. INTRODUCTION

Spatial data analysis has become a crucial research area in the modern era. It has diverse applications such as traffic forecasting, weather updates, and more. Spatial data often includes non-spatial observations that play a significant role in knowledge discovery. Various techniques are discussed. Qinjun et al. proposed a text mining approach coupled with spatial data processing for generating spatial analysis results in geoscience reports. Senzhang et al. utilized deep learning to discover spatial features from datasets. Maria et al. applied various techniques to big spatial data for emergency management in business systems. Wesley analyzed climate data and the associated challenges. Fernandez et al. investigated SDM for situational analysis in maritime contexts. The literature reveals many techniques for spatial data analysis, considering the temporal domain as well. This paper sheds light on spatial correlation discovery using geographical datasets and specific keywords. Our contributions are as follows:

- We proposed a framework focused on discovering spatial correlations based on G statistic and ZG score computations.
- We developed an algorithm known as Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD).

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 2, February 2021 |

| DOI:10.15680/IJMRSET.2021.0402006 |

Autonomous Machine-to-Machine (M2m) Collaboration for Indoor Search and Rescue Missions: A Multimodal Approach using Ground Vehicles and Drones

S Harsha Vardhan U[1], Vijay Kumar[2], Sampath H P[3], Rakesh Y D[4], Anil Kumar R[5]

Assistant Professors, Department of Mechanical Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Machine-to-Machine (M2M) collaboration has unlocked new possibilities for systems to autonomously and efficiently solve complex technical challenges. In recent years, M2M applications have expanded across various industries, particularly in areas where autonomous decision-making and real-time collaboration are crucial. This research project presents the development of an M2M system that enables autonomous collaboration between two ground vehicles and a drone, managed by a central base station. The system is designed specifically for indoor search and rescue operations, where traditional human intervention may be challenging or hazardous. The primary function of the drone in this system is to perform an autonomous aerial search, while the ground vehicles handle obstacle navigation and ground-level data collection. The base station serves as the control hub, managing both the drone's flight path and the vehicles' navigation, ensuring real-time coordination. A machine learning model is employed to optimize the drone's flight path, achieving near-perfect accuracy in navigating the indoor environment. This accuracy improves significantly with an increase in training data, highlighting the importance of robust training in machine learning-based navigation systems. For ground vehicles, machine learning models are also implemented to navigate complex environments with obstacles. During field testing, the integrated system demonstrated high levels of accuracy and efficiency in coordinating tasks between the drone and ground vehicles. This successful implementation of M2M collaboration showcases its potential in critical applications like search and rescue, offering a glimpse into future advancements in autonomous systems for emergency response operations. The project's findings underscore the growing importance of M2M technology in enhancing the effectiveness of autonomous systems across various domains.

KEYWORDS: Search and Rescue, Image Processing, Navigation Systems, Autonomous Systems, and Object Detection

I. INTRODUCTION

Machine-to-machine (M2M) collaboration became possible with the introduction of Cyber-Physical Systems (CPS) and the Internet of Things (IoT), together with significant advancements in networking, cloud computing, and machine learning. Road transportation, system automation, surveillance, search and rescue operations, and many other situations make use of the idea of M2M collaboration. The Internet of Things (IoT) offers a network of items with integrated sensors, software, and other technology designed to link and exchange data with other systems and devices. Machine learning is one of the key components of M2M collaboration. Researchers have used machine learning for a variety of purposes in a number of efforts. In one project, Chirra and his colleagues used face state observation to apply a behavioral technique to identify driver fatigue. In order to ascertain whether or not the driver is sleepy, it used the Viola-Jones identification algorithm to identify the face and extract the eye region from the facial image. To determine whether the driver is asleep or not, a Convolutional Neural Network (CNN) classifier employs a SoftMax layer. Better accuracy was reported by this work as compared to conventional CNN. Munawar and associates devised a categorization scheme for flooding management to organize the several technologies under examination [5]. They discovered that there were not many hybrid models—which fuse machine learning and image processing—for flood control. Furthermore, it was discovered that there was little use of machine learning-based techniques in the aftermath of a disaster. Therefore, to enable efficient and comprehensive catastrophe management throughout all phases, future efforts must concentrate on integrating disaster management expertise, image processing techniques, and machine learning tools.

Semeraro and colleagues provided a survey of the literature on machine-learning approaches used in human-robot collaboration. In order to do work grouping based on the kinds of collaborative activities, evaluation measures, and

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 12, December 2021 ||

| DOI:10.15680/IJIRSET.2021.1012188 |

Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance: A Comprehensive Analysis of Online Portfolio Strategies

Gayatri Annasagaram[1], Vanitha G R[2], Nagendra Naik K[3], Rekha R[4], Jyothi[5]

Assistant Professors, Department of Mathematics, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The pattern matching method has long been recognized as a traditional and effective classification approach within the domain of online portfolio selection strategies. This article delves into the core elements of this method, focusing on two key components: the assessment of similarity and the selection of similarity sets. To address the growing complexity of financial markets, this study introduces an enhanced approach known as the Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance (PMDI). The aim is to optimize portfolio selection by integrating a more comprehensive analysis of historical stock price patterns. PMDI leverages a dual-information framework, combining both direction and distance metrics to create a more accurate and refined pattern-matching process. The method incorporates various well-known techniques, such as the Euclidean distance, Chebyshev distance, and the correlation coefficient, to capture the essential information embedded in historical stock price movements. By analyzing both the directional trends (the way prices move over time) and the relative distances between stock prices, PMDI identifies a similarity set that closely mirrors the desired investment strategy. This similarity set forms the foundation of the portfolio selection algorithm. The novelty of PMDI lies in its ability to not only assess patterns based on price direction but also to quantify how far stock prices deviate from one another. This dual approach enhances the system's capability to filter out noise and irrelevant data, thereby selecting only the most relevant patterns for further processing. As a result, the system can make more informed decisions when constructing an optimal investment portfolio, balancing between maximizing returns and minimizing risks. To validate the effectiveness of PMDI, extensive experiments were carried out using two real-world stock market datasets. The results of these tests demonstrated that PMDI outperformed other traditional algorithms, achieving a superior balance between returns and risk management. In comparison to single-metric models, which may focus only on direction or distance, PMDI's dual-information framework proves to be more robust and adaptable to the ever-changing dynamics of financial markets. In the increasingly volatile world of finance, having a strategy that can dynamically respond to market shifts is crucial. The PMDI method offers a significant improvement over existing portfolio selection techniques by utilizing a richer set of data inputs. Its ability to handle both directional trends and price deviations makes it particularly well-suited for real-world financial environments, where adaptability and precision are key to achieving sustainable long-term returns. As financial markets evolve, methods like PMDI will likely become essential tools for investors seeking to maximize performance while effectively managing risk.

KEYWORDS: Portfolio Selection, Pattern Matching, Dual Information, Financial Markets, Risk Management

I. INTRODUCTION

The allocation of limited capital in global financial markets has consistently garnered significant interest from both investors and researchers, particularly in the quest to optimize returns. The intricate nature and high interconnectivity of these markets complicate the ability to forecast market movements. Additionally, elements such as financial innovation, diverse trading strategies, and regulatory frameworks contribute to the prevailing uncertainty within the market. Consequently, devising effective capital allocation strategies within this multifaceted environment has emerged as a pressing global issue. Portfolio Selection (PS) stands as a fundamental endeavour within the financial sector, focusing on the distribution of capital across a defined array of assets to meet specific investment objectives. The roots of this challenge can be traced back to the mean-variance theory introduced by Markowitz in 1952, which employs mean-variance analysis to identify the optimal investment portfolio by balancing anticipated returns against associated risks. Nevertheless, as market dynamics evolve and investor preferences diversify, conventional portfolio selection techniques have increasingly revealed their shortcomings.

In recent years, the swift advancement of artificial intelligence and machine learning has introduced innovative approaches to address portfolio selection challenges. These technologies possess the capability to scrutinize historical data, uncovering patterns and trends that facilitate more accurate predictions of future market behaviour. By processing

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 7.122|| A Monthly Peer Reviewed & Referred Journal |

||Volume 10, Issue 1, January 2021||

DOI:10.15662/IJAREEIE.2021.1001028

Discovering Spatial Correlations through Spatial Data Mining: A Framework for Geospatial Data Analytics

P Rajshekar[1], Sujatha[2], Nagashree G[3], Ashwini Hindiholi[4], Sunitha N[5]

Professor, Department of Chemistry, Engineering College, Bengaluru, Karnataka, India[1]

Professor, Department of Physics, City Engineering College, Bengaluru, Karnataka, India[2]

Assistant Professors, Department of Basic Sciences, City Engineering College, Bengaluru, Karnataka, India[3,4,5]

ABSTRACT: Spatial data holds immense potential for uncovering geographical correlations, which can provide valuable insights into patterns across various fields such as urban planning, environmental monitoring, and social media analysis. Spatial Data Mining (SDM) offers powerful methods for analyzing these correlations, and this paper introduces a novel approach to enhance this analytical process. Specifically, we propose a framework for geospatial data analytics that integrates the G statistic and ZG score computations for spatial correlation discovery. The proposed algorithm, Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD), is designed to perform comprehensive spatial correlation analysis while also incorporating Principal Component Analysis (PCA) to uncover trends and reduce dimensionality. By applying SDM-SCD to Twitter data, we demonstrate how spatial correlations can be effectively analyzed based on the origin location of tweets and specific keywords. The analysis process begins by collecting geolocated Twitter data based on a set of predefined keywords. The algorithm then computes the G statistic and ZG score to detect significant spatial correlations between different geographic regions. These correlations are mapped and analyzed to identify hotspots, clusters, or patterns of interest. By applying PCA, the algorithm reduces the complexity of the data while retaining the most relevant components, enabling more efficient trend identification and visualization. Our experimental results demonstrate that the SDM-SCD framework effectively identifies spatial correlations in Twitter data, providing meaningful insights into how topics or trends evolve across geographical regions. For instance, the algorithm can reveal the spatial spread of discussions related to public health, natural disasters, or political events, offering valuable real-time insights for decision-makers. The SDM-SCD algorithm and framework provide a robust approach for geospatial data analysis, with the ability to uncover spatial patterns and trends from large-scale social media data. By leveraging the power of spatial data mining and advanced statistical techniques, this framework opens new avenues for exploring spatial correlations in various fields, with potential applications in social media analytics, public health, and beyond. The experimental validation highlights its efficacy in providing actionable insights from geospatial data.\

KEYWORDS: Spatial Data Mining, Geospatial Data Analysis, G Statistic, ZG Score Computations, Spatial Correlation Analysis

I. INTRODUCTION

Spatial data analysis has become a crucial research area in the modern era. It has diverse applications such as traffic forecasting, weather updates, and more. Spatial data often includes non-spatial observations that play a significant role in knowledge discovery. Various techniques are discussed. Qinjun et al. proposed a text mining approach coupled with spatial data processing for generating spatial analysis results in geoscience reports. Senzhang et al. utilized deep learning to discover spatial features from datasets. Maria et al. applied various techniques to big spatial data for emergency management in business systems. Wesley analyzed climate data and the associated challenges. Fernandez et al. investigated SDM for situational analysis in maritime contexts. The literature reveals many techniques for spatial data analysis, considering the temporal domain as well. This paper sheds light on spatial correlation discovery using geographical datasets and specific keywords. Our contributions are as follows:

- We proposed a framework focused on discovering spatial correlations based on G statistic and ZG score computations.
- We developed an algorithm known as Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD).

ATTESTED COPY
PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 12, December 2021 ||

| DOI:10.15680/IJIRSET.2021.1012188 |

Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance: A Comprehensive Analysis of Online Portfolio Strategies

Gayatri Annasagaram[1], Vanitha G R[2], Nagendra Naik K[3], Rekha R[4], Jyothi[5]

Assistant Professors, Department of Mathematics, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The pattern matching method has long been recognized as a traditional and effective classification approach within the domain of online portfolio selection strategies. This article delves into the core elements of this method, focusing on two key components: the assessment of similarity and the selection of similarity sets. To address the growing complexity of financial markets, this study introduces an enhanced approach known as the Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance (PMDI). The aim is to optimize portfolio selection by integrating a more comprehensive analysis of historical stock price patterns. PMDI leverages a dual-information framework, combining both direction and distance metrics to create a more accurate and refined pattern-matching process. The method incorporates various well-known techniques, such as the Euclidean distance, Chebyshev distance, and the correlation coefficient, to capture the essential information embedded in historical stock price movements. By analyzing both the directional trends (the way prices move over time) and the relative distances between stock prices, PMDI identifies a similarity set that closely mirrors the desired investment strategy. This similarity set forms the foundation of the portfolio selection algorithm. The novelty of PMDI lies in its ability to not only assess patterns based on price direction but also to quantify how far stock prices deviate from one another. This dual approach enhances the system's capability to filter out noise and irrelevant data, thereby selecting only the most relevant patterns for further processing. As a result, the system can make more informed decisions when constructing an optimal investment portfolio, balancing between maximizing returns and minimizing risks. To validate the effectiveness of PMDI, extensive experiments were carried out using two real-world stock market datasets. The results of these tests demonstrated that PMDI outperformed other traditional algorithms, achieving a superior balance between returns and risk management. In comparison to single-metric models, which may focus only on direction or distance, PMDI's dual-information framework proves to be more robust and adaptable to the ever-changing dynamics of financial markets. In the increasingly volatile world of finance, having a strategy that can dynamically respond to market shifts is crucial. The PMDI method offers a significant improvement over existing portfolio selection techniques by utilizing a richer set of data inputs. Its ability to handle both directional trends and price deviations makes it particularly well-suited for real-world financial environments, where adaptability and precision are key to achieving sustainable long-term returns. As financial markets evolve, methods like PMDI will likely become essential tools for investors seeking to maximize performance while effectively managing risk.

KEYWORDS: Portfolio Selection, Pattern Matching, Dual Information, Financial Markets, Risk Management

I. INTRODUCTION

The allocation of limited capital in global financial markets has consistently garnered significant interest from both investors and researchers, particularly in the quest to optimize returns. The intricate nature and high interconnectivity of these markets complicate the ability to forecast market movements. Additionally, elements such as financial innovation, diverse trading strategies, and regulatory frameworks contribute to the prevailing uncertainty within the market. Consequently, devising effective capital allocation strategies within this multifaceted environment has emerged as a pressing global issue. Portfolio Selection (PS) stands as a fundamental endeavour within the financial sector, focusing on the distribution of capital across a defined array of assets to meet specific investment objectives. The roots of this challenge can be traced back to the mean-variance theory introduced by Markowitz in 1952, which employs mean-variance analysis to identify the optimal investment portfolio by balancing anticipated returns against associated risks. Nevertheless, as market dynamics evolve and investor preferences diversify, conventional portfolio selection techniques have increasingly revealed their shortcomings.

In recent years, the swift advancement of artificial intelligence and machine learning has introduced innovative approaches to address portfolio selection challenges. These technologies possess the capability to scrutinize historical data, uncovering patterns and trends that facilitate more accurate predictions of future market behaviour. By processing

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 2, February 2021 |

| DOI:10.15680/IJMRSET.2021.0402006 |

Autonomous Machine-to-Machine (M2m) Collaboration for Indoor Search and Rescue Missions: A Multimodal Approach using Ground Vehicles and Drones

S Harsha Vardhan U[1], Vijay Kumar[2], Sampath H P[3], Rakesh Y D[4], Anil Kumar R[5]

Assistant Professors, Department of Mechanical Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: Machine-to-Machine (M2M) collaboration has unlocked new possibilities for systems to autonomously and efficiently solve complex technical challenges. In recent years, M2M applications have expanded across various industries, particularly in areas where autonomous decision-making and real-time collaboration are crucial. This research project presents the development of an M2M system that enables autonomous collaboration between two ground vehicles and a drone, managed by a central base station. The system is designed specifically for indoor search and rescue operations, where traditional human intervention may be challenging or hazardous. The primary function of the drone in this system is to perform an autonomous aerial search, while the ground vehicles handle obstacle navigation and ground-level data collection. The base station serves as the control hub, managing both the drone's flight path and the vehicles' navigation, ensuring real-time coordination. A machine learning model is employed to optimize the drone's flight path, achieving near-perfect accuracy in navigating the indoor environment. This accuracy improves significantly with an increase in training data, highlighting the importance of robust training in machine learning-based navigation systems. For ground vehicles, machine learning models are also implemented to navigate complex environments with obstacles. During field testing, the integrated system demonstrated high levels of accuracy and efficiency in coordinating tasks between the drone and ground vehicles. This successful implementation of M2M collaboration showcases its potential in critical applications like search and rescue, offering a glimpse into future advancements in autonomous systems for emergency response operations. The project's findings underscore the growing importance of M2M technology in enhancing the effectiveness of autonomous systems across various domains.

KEYWORDS: Search and Rescue, Image Processing, Navigation Systems, Autonomous Systems, and Object Detection

I. INTRODUCTION

Machine-to-machine (M2M) collaboration became possible with the introduction of Cyber-Physical Systems (CPS) and the Internet of Things (IoT), together with significant advancements in networking, cloud computing, and machine learning. Road transportation, system automation, surveillance, search and rescue operations, and many other situations make use of the idea of M2M collaboration. The Internet of Things (IoT) offers a network of items with integrated sensors, software, and other technology designed to link and exchange data with other systems and devices. Machine learning is one of the key components of M2M collaboration. Researchers have used machine learning for a variety of purposes in a number of efforts. In one project, Chirra and his colleagues used face state observation to apply a behavioral technique to identify driver fatigue. In order to ascertain whether or not the driver is sleepy, it used the Viola-Jones identification algorithm to identify the face and extract the eye region from the facial image. To determine whether the driver is asleep or not, a Convolutional Neural Network (CNN) classifier employs a SoftMax layer. Better accuracy was reported by this work as compared to conventional CNN. Munawar and associates devised a categorization scheme for flooding management to organize the several technologies under examination [5]. They discovered that there were not many hybrid models—which fuse machine learning and image processing—for flood control. Furthermore, it was discovered that there was little use of machine learning-based techniques in the aftermath of a disaster. Therefore, to enable efficient and comprehensive catastrophe management throughout all phases, future efforts must concentrate on integrating disaster management expertise, image processing techniques, and machine learning tools.

Semeraro and colleagues provided a survey of the literature on machine-learning approaches used in human-robot collaboration. In order to do work grouping based on the kinds of collaborative activities, evaluation measures, and

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 7.122|| A Monthly Peer Reviewed & Referred Journal |

||Volume 10, Issue 1, January 2021||

DOI:10.15662/IJAREEIE.2021.1001028

Discovering Spatial Correlations through Spatial Data Mining: A Framework for Geospatial Data Analytics

P Rajshekar[1], Sujatha[2], Nagashree G[3], Ashwini Hindiholi[4], Sunitha N[5]

Professor, Department of Chemistry, Engineering College, Bengaluru, Karnataka, India[1]

Professor, Department of Physics, City Engineering College, Bengaluru, Karnataka, India[2]

Assistant Professors, Department of Basic Sciences, City Engineering College, Bengaluru, Karnataka, India[3,4,5]

ABSTRACT: Spatial data holds immense potential for uncovering geographical correlations, which can provide valuable insights into patterns across various fields such as urban planning, environmental monitoring, and social media analysis. Spatial Data Mining (SDM) offers powerful methods for analyzing these correlations, and this paper introduces a novel approach to enhance this analytical process. Specifically, we propose a framework for geospatial data analytics that integrates the G statistic and ZG score computations for spatial correlation discovery. The proposed algorithm, Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD), is designed to perform comprehensive spatial correlation analysis while also incorporating Principal Component Analysis (PCA) to uncover trends and reduce dimensionality. By applying SDM-SCD to Twitter data, we demonstrate how spatial correlations can be effectively analyzed based on the origin location of tweets and specific keywords. The analysis process begins by collecting geolocated Twitter data based on a set of predefined keywords. The algorithm then computes the G statistic and ZG score to detect significant spatial correlations between different geographic regions. These correlations are mapped and analyzed to identify hotspots, clusters, or patterns of interest. By applying PCA, the algorithm reduces the complexity of the data while retaining the most relevant components, enabling more efficient trend identification and visualization. Our experimental results demonstrate that the SDM-SCD framework effectively identifies spatial correlations in Twitter data, providing meaningful insights into how topics or trends evolve across geographical regions. For instance, the algorithm can reveal the spatial spread of discussions related to public health, natural disasters, or political events, offering valuable real-time insights for decision-makers. The SDM-SCD algorithm and framework provide a robust approach for geospatial data analysis, with the ability to uncover spatial patterns and trends from large-scale social media data. By leveraging the power of spatial data mining and advanced statistical techniques, this framework opens new avenues for exploring spatial correlations in various fields, with potential applications in social media analytics, public health, and beyond. The experimental validation highlights its efficacy in providing actionable insights from geospatial data.\

KEYWORDS: Spatial Data Mining, Geospatial Data Analysis, G Statistic, ZG Score Computations, Spatial Correlation Analysis

I. INTRODUCTION

Spatial data analysis has become a crucial research area in the modern era. It has diverse applications such as traffic forecasting, weather updates, and more. Spatial data often includes non-spatial observations that play a significant role in knowledge discovery. Various techniques are discussed. Qinjun et al. proposed a text mining approach coupled with spatial data processing for generating spatial analysis results in geoscience reports. Senzhang et al. utilized deep learning to discover spatial features from datasets. Maria et al. applied various techniques to big spatial data for emergency management in business systems. Wesley analyzed climate data and the associated challenges. Fernandez et al. investigated SDM for situational analysis in maritime contexts. The literature reveals many techniques for spatial data analysis, considering the temporal domain as well. This paper sheds light on spatial correlation discovery using geographical datasets and specific keywords. Our contributions are as follows:

- We proposed a framework focused on discovering spatial correlations based on G statistic and ZG score computations.
- We developed an algorithm known as Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD).

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 7.122|| A Monthly Peer Reviewed & Referred Journal |

||Volume 10, Issue 1, January 2021||

DOI:10.15662/IJAREEIE.2021.1001028

Discovering Spatial Correlations through Spatial Data Mining: A Framework for Geospatial Data Analytics

P Rajshekar[1], Sujatha[2], Nagashree G[3], Ashwini Hindiholi[4], Sunitha N[5]

Professor, Department of Chemistry, Engineering College, Bengaluru, Karnataka, India[1]

Professor, Department of Physics, City Engineering College, Bengaluru, Karnataka, India[2]

Assistant Professors, Department of Basic Sciences, City Engineering College, Bengaluru, Karnataka, India[3,4,5]

ABSTRACT: Spatial data holds immense potential for uncovering geographical correlations, which can provide valuable insights into patterns across various fields such as urban planning, environmental monitoring, and social media analysis. Spatial Data Mining (SDM) offers powerful methods for analyzing these correlations, and this paper introduces a novel approach to enhance this analytical process. Specifically, we propose a framework for geospatial data analytics that integrates the G statistic and ZG score computations for spatial correlation discovery. The proposed algorithm, Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD), is designed to perform comprehensive spatial correlation analysis while also incorporating Principal Component Analysis (PCA) to uncover trends and reduce dimensionality. By applying SDM-SCD to Twitter data, we demonstrate how spatial correlations can be effectively analyzed based on the origin location of tweets and specific keywords. The analysis process begins by collecting geolocated Twitter data based on a set of predefined keywords. The algorithm then computes the G statistic and ZG score to detect significant spatial correlations between different geographic regions. These correlations are mapped and analyzed to identify hotspots, clusters, or patterns of interest. By applying PCA, the algorithm reduces the complexity of the data while retaining the most relevant components, enabling more efficient trend identification and visualization. Our experimental results demonstrate that the SDM-SCD framework effectively identifies spatial correlations in Twitter data, providing meaningful insights into how topics or trends evolve across geographical regions. For instance, the algorithm can reveal the spatial spread of discussions related to public health, natural disasters, or political events, offering valuable real-time insights for decision-makers. The SDM-SCD algorithm and framework provide a robust approach for geospatial data analysis, with the ability to uncover spatial patterns and trends from large-scale social media data. By leveraging the power of spatial data mining and advanced statistical techniques, this framework opens new avenues for exploring spatial correlations in various fields, with potential applications in social media analytics, public health, and beyond. The experimental validation highlights its efficacy in providing actionable insights from geospatial data.\

KEYWORDS: Spatial Data Mining, Geospatial Data Analysis, G Statistic, ZG Score Computations, Spatial Correlation Analysis

I. INTRODUCTION

Spatial data analysis has become a crucial research area in the modern era. It has diverse applications such as traffic forecasting, weather updates, and more. Spatial data often includes non-spatial observations that play a significant role in knowledge discovery. Various techniques are discussed. Qinjun et al. proposed a text mining approach coupled with spatial data processing for generating spatial analysis results in geoscience reports. Senzhang et al. utilized deep learning to discover spatial features from datasets. Maria et al. applied various techniques to big spatial data for emergency management in business systems. Wesley analyzed climate data and the associated challenges. Fernandez et al. investigated SDM for situational analysis in maritime contexts. The literature reveals many techniques for spatial data analysis, considering the temporal domain as well. This paper sheds light on spatial correlation discovery using geographical datasets and specific keywords. Our contributions are as follows:

- We proposed a framework focused on discovering spatial correlations based on G statistic and ZG score computations.
- We developed an algorithm known as Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD).

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 7.569| A Monthly Peer Reviewed & Referred Journal |

|| Volume 10, Issue 12, December 2021 ||

| DOI:10.15680/IJIRSET.2021.1012188 |

Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance: A Comprehensive Analysis of Online Portfolio Strategies

Gayatri Annasagaram[1], Vanitha G R[2], Nagendra Naik K[3], Rekha R[4], Jyothi[5]

Assistant Professors, Department of Mathematics, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: The pattern matching method has long been recognized as a traditional and effective classification approach within the domain of online portfolio selection strategies. This article delves into the core elements of this method, focusing on two key components: the assessment of similarity and the selection of similarity sets. To address the growing complexity of financial markets, this study introduces an enhanced approach known as the Portfolio Selection Method based on Pattern Matching with Dual Information of Direction and Distance (PMDI). The aim is to optimize portfolio selection by integrating a more comprehensive analysis of historical stock price patterns. PMDI leverages a dual-information framework, combining both direction and distance metrics to create a more accurate and refined pattern-matching process. The method incorporates various well-known techniques, such as the Euclidean distance, Chebyshev distance, and the correlation coefficient, to capture the essential information embedded in historical stock price movements. By analyzing both the directional trends (the way prices move over time) and the relative distances between stock prices, PMDI identifies a similarity set that closely mirrors the desired investment strategy. This similarity set forms the foundation of the portfolio selection algorithm. The novelty of PMDI lies in its ability to not only assess patterns based on price direction but also to quantify how far stock prices deviate from one another. This dual approach enhances the system's capability to filter out noise and irrelevant data, thereby selecting only the most relevant patterns for further processing. As a result, the system can make more informed decisions when constructing an optimal investment portfolio, balancing between maximizing returns and minimizing risks. To validate the effectiveness of PMDI, extensive experiments were carried out using two real-world stock market datasets. The results of these tests demonstrated that PMDI outperformed other traditional algorithms, achieving a superior balance between returns and risk management. In comparison to single-metric models, which may focus only on direction or distance, PMDI's dual-information framework proves to be more robust and adaptable to the ever-changing dynamics of financial markets. In the increasingly volatile world of finance, having a strategy that can dynamically respond to market shifts is crucial. The PMDI method offers a significant improvement over existing portfolio selection techniques by utilizing a richer set of data inputs. Its ability to handle both directional trends and price deviations makes it particularly well-suited for real-world financial environments, where adaptability and precision are key to achieving sustainable long-term returns. As financial markets evolve, methods like PMDI will likely become essential tools for investors seeking to maximize performance while effectively managing risk.

KEYWORDS: Portfolio Selection, Pattern Matching, Dual Information, Financial Markets, Risk Management

I. INTRODUCTION

The allocation of limited capital in global financial markets has consistently garnered significant interest from both investors and researchers, particularly in the quest to optimize returns. The intricate nature and high interconnectivity of these markets complicate the ability to forecast market movements. Additionally, elements such as financial innovation, diverse trading strategies, and regulatory frameworks contribute to the prevailing uncertainty within the market. Consequently, devising effective capital allocation strategies within this multifaceted environment has emerged as a pressing global issue. Portfolio Selection (PS) stands as a fundamental endeavour within the financial sector, focusing on the distribution of capital across a defined array of assets to meet specific investment objectives. The roots of this challenge can be traced back to the mean-variance theory introduced by Markowitz in 1952, which employs mean-variance analysis to identify the optimal investment portfolio by balancing anticipated returns against associated risks. Nevertheless, as market dynamics evolve and investor preferences diversify, conventional portfolio selection techniques have increasingly revealed their shortcomings.

In recent years, the swift advancement of artificial intelligence and machine learning has introduced innovative approaches to address portfolio selection challenges. These technologies possess the capability to scrutinize historical data, uncovering patterns and trends that facilitate more accurate predictions of future market behaviour. By processing

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 7.122|| A Monthly Peer Reviewed & Referred Journal |

||Volume 10, Issue 1, January 2021||

DOI:10.15662/IJAREEIE.2021.1001028

Discovering Spatial Correlations through Spatial Data Mining: A Framework for Geospatial Data Analytics

P Rajshekar[1], Sujatha[2], Nagashree G[3], Ashwini Hindiholi[4], Sunitha N[5]

Professor, Department of Chemistry, Engineering College, Bengaluru, Karnataka, India[1]

Professor, Department of Physics, City Engineering College, Bengaluru, Karnataka, India[2]

Assistant Professors, Department of Basic Sciences, City Engineering College, Bengaluru, Karnataka, India[3,4,5]

ABSTRACT: Spatial data holds immense potential for uncovering geographical correlations, which can provide valuable insights into patterns across various fields such as urban planning, environmental monitoring, and social media analysis. Spatial Data Mining (SDM) offers powerful methods for analyzing these correlations, and this paper introduces a novel approach to enhance this analytical process. Specifically, we propose a framework for geospatial data analytics that integrates the G statistic and ZG score computations for spatial correlation discovery. The proposed algorithm, Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD), is designed to perform comprehensive spatial correlation analysis while also incorporating Principal Component Analysis (PCA) to uncover trends and reduce dimensionality. By applying SDM-SCD to Twitter data, we demonstrate how spatial correlations can be effectively analyzed based on the origin location of tweets and specific keywords. The analysis process begins by collecting geolocated Twitter data based on a set of predefined keywords. The algorithm then computes the G statistic and ZG score to detect significant spatial correlations between different geographic regions. These correlations are mapped and analyzed to identify hotspots, clusters, or patterns of interest. By applying PCA, the algorithm reduces the complexity of the data while retaining the most relevant components, enabling more efficient trend identification and visualization. Our experimental results demonstrate that the SDM-SCD framework effectively identifies spatial correlations in Twitter data, providing meaningful insights into how topics or trends evolve across geographical regions. For instance, the algorithm can reveal the spatial spread of discussions related to public health, natural disasters, or political events, offering valuable real-time insights for decision-makers. The SDM-SCD algorithm and framework provide a robust approach for geospatial data analysis, with the ability to uncover spatial patterns and trends from large-scale social media data. By leveraging the power of spatial data mining and advanced statistical techniques, this framework opens new avenues for exploring spatial correlations in various fields, with potential applications in social media analytics, public health, and beyond. The experimental validation highlights its efficacy in providing actionable insights from geospatial data.\

KEYWORDS: Spatial Data Mining, Geospatial Data Analysis, G Statistic, ZG Score Computations, Spatial Correlation Analysis

I. INTRODUCTION

Spatial data analysis has become a crucial research area in the modern era. It has diverse applications such as traffic forecasting, weather updates, and more. Spatial data often includes non-spatial observations that play a significant role in knowledge discovery. Various techniques are discussed. Qinjun et al. proposed a text mining approach coupled with spatial data processing for generating spatial analysis results in geoscience reports. Senzhang et al. utilized deep learning to discover spatial features from datasets. Maria et al. applied various techniques to big spatial data for emergency management in business systems. Wesley analyzed climate data and the associated challenges. Fernandez et al. investigated SDM for situational analysis in maritime contexts. The literature reveals many techniques for spatial data analysis, considering the temporal domain as well. This paper sheds light on spatial correlation discovery using geographical datasets and specific keywords. Our contributions are as follows:

- We proposed a framework focused on discovering spatial correlations based on G statistic and ZG score computations.
- We developed an algorithm known as Spatial Data Mining for Spatial Correlations Discovery (SDM-SCD).

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 4, April 2021 |

| DOI:10.15680/IJMRSET.2021.0404025 |

Predicting Bitcoin Prices using Machine Learning: A Comparative Analysis of Algorithms and Feature Engineering Techniques

Chandranaik G[1], Nanditha H G[2], Aurbindo Koti[3], Shalini Prasad[4], Sheetal Patted[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: A lot of attention has been paid to cryptocurrency price prediction as digital assets become more and more important in the financial system. This paper provides a thorough analysis of the use of machine learning algorithms in bit coin price prediction. Historical open-source data from multiple bit coin exchanges is used. To deal with missing data, interpolation techniques are used, guaranteeing the dataset's trustworthiness and completeness. Four technical indicators have been chosen as prediction features. The study looks at how five machine learning algorithms might be used to identify intricate patterns in the wildly unstable bit coin market. The results show the benefits and drawbacks of the various strategies, emphasizing the role that feature engineering and algorithm selection play in producing reliable predictions of bit coin prices. As digital assets like cryptocurrencies become increasingly integrated into the financial system, predicting their prices has garnered significant attention from traders, investors, and researchers. This paper focuses on the application of machine learning algorithms to predict Bitcoin prices, a highly volatile and unpredictable market. By leveraging historical data from various Bitcoin exchanges, the study ensures data integrity by employing interpolation techniques to handle missing values, resulting in a trustworthy and complete dataset for analysis. The research explores the use of four key technical indicators as predictive features, carefully selected based on their relevance to cryptocurrency markets. These features serve as inputs to five different machine learning algorithms, which are employed to uncover complex patterns and relationships within the Bitcoin market. The paper offers a comparative analysis of these algorithms, examining their strengths and weaknesses in predicting Bitcoin prices under volatile market conditions. A key focus of the study is on the importance of feature engineering and algorithm selection in achieving accurate and reliable predictions. The results demonstrate that certain algorithms outperform others in terms of prediction accuracy, while highlighting the critical role that feature engineering plays in enhancing model performance. The insights from this study provide valuable guidance for traders and investors navigating the rapidly evolving world of cryptocurrency markets, offering a practical framework for using machine learning to improve prediction accuracy in financial decision-making.

KEYWORDS: Cryptocurrency Price Prediction, Machine Learning Algorithms, Feature Engineering, Performance Metrics globally

I. INTRODUCTION

The financial system is seeing a shift due to cryptocurrencies, which are bringing decentralized digital assets based on block chain technology. The original cryptocurrency, Bitcoin, is driving a global wave of digital currencies that are giving rise to a plethora of substitute cryptocurrencies, or altcoins. The increasing prevalence of cryptocurrencies is garnering substantial interest from traders, investors, and financial establishments. In the meantime, cryptocurrencies are becoming a more attractive and fascinating asset class due to its decentralized structure, possibility for large profits, and distinct market dynamics. The bit coin sector is becoming more and more legitimate as a result of mainstream organizations and businesses adopting cryptocurrencies. Furthermore, as governments are looking at digital alternatives to their fiat currencies, the development of digital currencies issued by central banks is a big step forward. Sustainability issues, interoperability across many block chain networks, and industry's progress is also being shaped by regulatory frameworks, but investors looking to take advantage of market opportunities face significant obstacles due to the tremendous volatility and unpredictability of bit coin values. Therefore, there is a strong need for reliable and accurate predictive models to help investors make wise decisions in this quickly changing financial environment.

The investigation of novel techniques to predict future price movements is driven by the intrinsic complexity and volatility of cryptocurrency markets. The distinctive qualities of cryptocurrencies are frequently difficult for traditional financial models to represent, which leads researchers to consider machine learning algorithms as a possible remedy.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 4, April 2021 |

| DOI:10.15680/IJMRSET.2021.0404025 |

Predicting Bitcoin Prices using Machine Learning: A Comparative Analysis of Algorithms and Feature Engineering Techniques

Chandranaik G[1], Nanditha H G[2], Aurbindo Koti[3], Shalini Prasad[4], Sheetal Patted[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: A lot of attention has been paid to cryptocurrency price prediction as digital assets become more and more important in the financial system. This paper provides a thorough analysis of the use of machine learning algorithms in bit coin price prediction. Historical open-source data from multiple bit coin exchanges is used. To deal with missing data, interpolation techniques are used, guaranteeing the dataset's trustworthiness and completeness. Four technical indicators have been chosen as prediction features. The study looks at how five machine learning algorithms might be used to identify intricate patterns in the wildly unstable bit coin market. The results show the benefits and drawbacks of the various strategies, emphasizing the role that feature engineering and algorithm selection play in producing reliable predictions of bit coin prices. As digital assets like cryptocurrencies become increasingly integrated into the financial system, predicting their prices has garnered significant attention from traders, investors, and researchers. This paper focuses on the application of machine learning algorithms to predict Bitcoin prices, a highly volatile and unpredictable market. By leveraging historical data from various Bitcoin exchanges, the study ensures data integrity by employing interpolation techniques to handle missing values, resulting in a trustworthy and complete dataset for analysis. The research explores the use of four key technical indicators as predictive features, carefully selected based on their relevance to cryptocurrency markets. These features serve as inputs to five different machine learning algorithms, which are employed to uncover complex patterns and relationships within the Bitcoin market. The paper offers a comparative analysis of these algorithms, examining their strengths and weaknesses in predicting Bitcoin prices under volatile market conditions. A key focus of the study is on the importance of feature engineering and algorithm selection in achieving accurate and reliable predictions. The results demonstrate that certain algorithms outperform others in terms of prediction accuracy, while highlighting the critical role that feature engineering plays in enhancing model performance. The insights from this study provide valuable guidance for traders and investors navigating the rapidly evolving world of cryptocurrency markets, offering a practical framework for using machine learning to improve prediction accuracy in financial decision-making.

KEYWORDS: Cryptocurrency Price Prediction, Machine Learning Algorithms, Feature Engineering, Performance Metrics globally

I. INTRODUCTION

The financial system is seeing a shift due to cryptocurrencies, which are bringing decentralized digital assets based on block chain technology. The original cryptocurrency, Bitcoin, is driving a global wave of digital currencies that are giving rise to a plethora of substitute cryptocurrencies, or altcoins. The increasing prevalence of cryptocurrencies is garnering substantial interest from traders, investors, and financial establishments. In the meantime, cryptocurrencies are becoming a more attractive and fascinating asset class due to its decentralized structure, possibility for large profits, and distinct market dynamics. The bit coin sector is becoming more and more legitimate as a result of mainstream organizations and businesses adopting cryptocurrencies. Furthermore, as governments are looking at digital alternatives to their fiat currencies, the development of digital currencies issued by central banks is a big step forward. Sustainability issues, interoperability across many block chain networks, and industry's progress is also being shaped by regulatory frameworks, but investors looking to take advantage of market opportunities face significant obstacles due to the tremendous volatility and unpredictability of bit coin values. Therefore, there is a strong need for reliable and accurate predictive models to help investors make wise decisions in this quickly changing financial environment.

The investigation of novel techniques to predict future price movements is driven by the intrinsic complexity and volatility of cryptocurrency markets. The distinctive qualities of cryptocurrencies are frequently difficult for traditional financial models to represent, which leads researchers to consider machine learning algorithms as a possible remedy.

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 4, April 2021 |

| DOI:10.15680/IJMRSET.2021.0404025 |

Predicting Bitcoin Prices using Machine Learning: A Comparative Analysis of Algorithms and Feature Engineering Techniques

Chandranaik G[1], Nanditha H G[2], Aurbindo Koti[3], Shalini Prasad[4], Sheetal Patted[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: A lot of attention has been paid to cryptocurrency price prediction as digital assets become more and more important in the financial system. This paper provides a thorough analysis of the use of machine learning algorithms in bit coin price prediction. Historical open-source data from multiple bit coin exchanges is used. To deal with missing data, interpolation techniques are used, guaranteeing the dataset's trustworthiness and completeness. Four technical indicators have been chosen as prediction features. The study looks at how five machine learning algorithms might be used to identify intricate patterns in the wildly unstable bit coin market. The results show the benefits and drawbacks of the various strategies, emphasizing the role that feature engineering and algorithm selection play in producing reliable predictions of bit coin prices. As digital assets like cryptocurrencies become increasingly integrated into the financial system, predicting their prices has garnered significant attention from traders, investors, and researchers. This paper focuses on the application of machine learning algorithms to predict Bitcoin prices, a highly volatile and unpredictable market. By leveraging historical data from various Bitcoin exchanges, the study ensures data integrity by employing interpolation techniques to handle missing values, resulting in a trustworthy and complete dataset for analysis. The research explores the use of four key technical indicators as predictive features, carefully selected based on their relevance to cryptocurrency markets. These features serve as inputs to five different machine learning algorithms, which are employed to uncover complex patterns and relationships within the Bitcoin market. The paper offers a comparative analysis of these algorithms, examining their strengths and weaknesses in predicting Bitcoin prices under volatile market conditions. A key focus of the study is on the importance of feature engineering and algorithm selection in achieving accurate and reliable predictions. The results demonstrate that certain algorithms outperform others in terms of prediction accuracy, while highlighting the critical role that feature engineering plays in enhancing model performance. The insights from this study provide valuable guidance for traders and investors navigating the rapidly evolving world of cryptocurrency markets, offering a practical framework for using machine learning to improve prediction accuracy in financial decision-making.

KEYWORDS: Cryptocurrency Price Prediction, Machine Learning Algorithms, Feature Engineering, Performance Metrics globally

I. INTRODUCTION

The financial system is seeing a shift due to cryptocurrencies, which are bringing decentralized digital assets based on block chain technology. The original cryptocurrency, Bitcoin, is driving a global wave of digital currencies that are giving rise to a plethora of substitute cryptocurrencies, or altcoins. The increasing prevalence of cryptocurrencies is garnering substantial interest from traders, investors, and financial establishments. In the meantime, cryptocurrencies are becoming a more attractive and fascinating asset class due to its decentralized structure, possibility for large profits, and distinct market dynamics. The bit coin sector is becoming more and more legitimate as a result of mainstream organizations and businesses adopting cryptocurrencies. Furthermore, as governments are looking at digital alternatives to their fiat currencies, the development of digital currencies issued by central banks is a big step forward. Sustainability issues, interoperability across many block chain networks, and industry's progress is also being shaped by regulatory frameworks, but investors looking to take advantage of market opportunities face significant obstacles due to the tremendous volatility and unpredictability of bit coin values. Therefore, there is a strong need for reliable and accurate predictive models to help investors make wise decisions in this quickly changing financial environment.

The investigation of novel techniques to predict future price movements is driven by the intrinsic complexity and volatility of cryptocurrency markets. The distinctive qualities of cryptocurrencies are frequently difficult for traditional financial models to represent, which leads researchers to consider machine learning algorithms as a possible remedy.

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 4, April 2021 |

| DOI:10.15680/IJMRSET.2021.0404025 |

Predicting Bitcoin Prices using Machine Learning: A Comparative Analysis of Algorithms and Feature Engineering Techniques

Chandranaik G[1],Ianditha H G[2], Aurbindo Koti[3], Shalini Prasad[4], Sheetal Patted[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: A lot of attention has been paid to cryptocurrency price prediction as digital assets become more and more important in the financial system. This paper provides a thorough analysis of the use of machine learning algorithms in bit coin price prediction. Historical open-source data from multiple bit coin exchanges is used. To deal with missing data, interpolation techniques are used, guaranteeing the dataset's trustworthiness and completeness. Four technical indicators have been chosen as prediction features. The study looks at how five machine learning algorithms might be used to identify intricate patterns in the wildly unstable bit coin market. The results show the benefits and drawbacks of the various strategies, emphasizing the role that feature engineering and algorithm selection play in producing reliable predictions of bit coin prices. As digital assets like cryptocurrencies become increasingly integrated into the financial system, predicting their prices has garnered significant attention from traders, investors, and researchers. This paper focuses on the application of machine learning algorithms to predict Bitcoin prices, a highly volatile and unpredictable market. By leveraging historical data from various Bitcoin exchanges, the study ensures data integrity by employing interpolation techniques to handle missing values, resulting in a trustworthy and complete dataset for analysis. The research explores the use of four key technical indicators as predictive features, carefully selected based on their relevance to cryptocurrency markets. These features serve as inputs to five different machine learning algorithms, which are employed to uncover complex patterns and relationships within the Bitcoin market. The paper offers a comparative analysis of these algorithms, examining their strengths and weaknesses in predicting Bitcoin prices under volatile market conditions. A key focus of the study is on the importance of feature engineering and algorithm selection in achieving accurate and reliable predictions. The results demonstrate that certain algorithms outperform others in terms of prediction accuracy, while highlighting the critical role that feature engineering plays in enhancing model performance. The insights from this study provide valuable guidance for traders and investors navigating the rapidly evolving world of cryptocurrency markets, offering a practical framework for using machine learning to improve prediction accuracy in financial decision-making.

KEYWORDS: Cryptocurrency Price Prediction, Machine Learning Algorithms, Feature Engineering, Performance Metrics globally

I. INTRODUCTION

The financial system is seeing a shift due to cryptocurrencies, which are bringing decentralized digital assets based on block chain technology. The original cryptocurrency, Bitcoin, is driving a global wave of digital currencies that are giving rise to a plethora of substitute cryptocurrencies, or altcoins. The increasing prevalence of cryptocurrencies is garnering substantial interest from traders, investors, and financial establishments. In the meantime, cryptocurrencies are becoming a more attractive and fascinating asset class due to its decentralized structure, possibility for large profits, and distinct market dynamics. The bit coin sector is becoming more and more legitimate as a result of mainstream organizations and businesses adopting cryptocurrencies. Furthermore, as governments are looking at digital alternatives to their fiat currencies, the development of digital currencies issued by central banks is a big step forward. Sustainability issues, interoperability across many block chain networks, and industry's progress is also being shaped by regulatory frameworks, but investors looking to take advantage of market opportunities face significant obstacles due to the tremendous volatility and unpredictability of bit coin values. Therefore, there is a strong need for reliable and accurate predictive models to help investors make wise decisions in this quickly changing financial environment.

The investigation of novel techniques to predict future price movements is driven by the intrinsic complexity and volatility of cryptocurrency markets. The distinctive qualities of cryptocurrencies are frequently difficult for traditional financial models to represent, which leads researchers to consider machine learning algorithms as a possible remedy.

ATTESTED COPY

PRINCIPAL
CITY ENGINEERING COLLEGE
Kanakapura Main Road, Bangalore-560061

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988| Monthly, Peer Reviewed & Referred Journal|

| Volume 4, Issue 4, April 2021 |

| DOI:10.15680/IJMRSET.2021.0404025 |

Predicting Bitcoin Prices using Machine Learning: A Comparative Analysis of Algorithms and Feature Engineering Techniques

Chandranaik G[1], Nanditha H G[2], Aurbindo Koti[3], Shalini Prasad[4], Sheetal Patted[5]

Assistant Professors, Department of Electronics and Communication Engineering, City Engineering College, Bengaluru, Karnataka, India[1,2,3,4,5]

ABSTRACT: A lot of attention has been paid to cryptocurrency price prediction as digital assets become more and more important in the financial system. This paper provides a thorough analysis of the use of machine learning algorithms in bit coin price prediction. Historical open-source data from multiple bit coin exchanges is used. To deal with missing data, interpolation techniques are used, guaranteeing the dataset's trustworthiness and completeness. Four technical indicators have been chosen as prediction features. The study looks at how five machine learning algorithms might be used to identify intricate patterns in the wildly unstable bit coin market. The results show the benefits and drawbacks of the various strategies, emphasizing the role that feature engineering and algorithm selection play in producing reliable predictions of bit coin prices. As digital assets like cryptocurrencies become increasingly integrated into the financial system, predicting their prices has garnered significant attention from traders, investors, and researchers. This paper focuses on the application of machine learning algorithms to predict Bitcoin prices, a highly volatile and unpredictable market. By leveraging historical data from various Bitcoin exchanges, the study ensures data integrity by employing interpolation techniques to handle missing values, resulting in a trustworthy and complete dataset for analysis. The research explores the use of four key technical indicators as predictive features, carefully selected based on their relevance to cryptocurrency markets. These features serve as inputs to five different machine learning algorithms, which are employed to uncover complex patterns and relationships within the Bitcoin market. The paper offers a comparative analysis of these algorithms, examining their strengths and weaknesses in predicting Bitcoin prices under volatile market conditions. A key focus of the study is on the importance of feature engineering and algorithm selection in achieving accurate and reliable predictions. The results demonstrate that certain algorithms outperform others in terms of prediction accuracy, while highlighting the critical role that feature engineering plays in enhancing model performance. The insights from this study provide valuable guidance for traders and investors navigating the rapidly evolving world of cryptocurrency markets, offering a practical framework for using machine learning to improve prediction accuracy in financial decision-making.

KEYWORDS: Cryptocurrency Price Prediction, Machine Learning Algorithms, Feature Engineering, Performance Metrics globally

I. INTRODUCTION

The financial system is seeing a shift due to cryptocurrencies, which are bringing decentralized digital assets based on block chain technology. The original cryptocurrency, Bitcoin, is driving a global wave of digital currencies that are giving rise to a plethora of substitute cryptocurrencies, or altcoins. The increasing prevalence of cryptocurrencies is garnering substantial interest from traders, investors, and financial establishments. In the meantime, cryptocurrencies are becoming a more attractive and fascinating asset class due to its decentralized structure, possibility for large profits, and distinct market dynamics. The bit coin sector is becoming more and more legitimate as a result of mainstream organizations and businesses adopting cryptocurrencies. Furthermore, as governments are looking at digital alternatives to their fiat currencies, the development of digital currencies issued by central banks is a big step forward. Sustainability issues, interoperability across many block chain networks, and industry's progress is also being shaped by regulatory frameworks, but investors looking to take advantage of market opportunities face significant obstacles due to the tremendous volatility and unpredictability of bit coin values. Therefore, there is a strong need for reliable and accurate predictive models to help investors make wise decisions in this quickly changing financial environment.

The investigation of novel techniques to predict future price movements is driven by the intrinsic complexity and volatility of cryptocurrency markets. The distinctive qualities of cryptocurrencies are frequently difficult for traditional financial models to represent, which leads researchers to consider machine learning algorithms as a possible remedy.