Acknowledgment

We, as a team, wholeheartedly express our gratitude to the individuals who have been pivotal in our journey to develop a GIS based Power Distribution System using Machine Learning. Our collective motivation is that the existing system suffers from manual surveys that are time-consuming and prone to errors, rampant electricity theft, and inefficiencies in determining optimal paths for consumers from the electricity poles.To overcome these challenges by harnessing the power of modern technology.Primary objectives are to eliminate the need for manual surveys through automation, implement accurate load forecasting through machine learning, and establish optimal consumer-to-pole pathways.Our foremost acknowledgment goes to our project supervisor Dr. D. R. Patil for their unwavering guidance, which has been instrumental in shaping our project. The project's feasibility and scope have the potential to change the existing Power Distribution System. In conclusion, we extend our heartfelt appreciation to all who have contributed to this project.

Abstract

The conventional power distribution network faces challenges such as resource inefficiency, suboptimal allocation, and increased energy losses, leading to longer outage restoration times and higher operational costs. In India, the power industry is expanding to meet rising demand due to population growth. The integration of GIS technology is significantly benefiting power system planning and management. This project advocates a conventional yet innovative approach, combining GIS data with Machine Learning (ML) to optimize power distribution systems. The GIS-based system, leveraging ML algorithms, predicts demand, identifies faults, and improves energy efficiency. By analyzing historical data, the model optimizes maintenance schedules and forecasts equipment failures, enhancing overall network reliability. The objective is to create a robust framework using GIS spatial data to optimize power distribution network operations and maintenance. The system aims to improve grid reliability, optimize load balancing, and minimize energy losses. Advanced machine learning algorithms enable the model to learn from historical and real-time data, predicting electricity demand patterns and potential fault locations. This initiative contributes to building a smarter, more sustainable power infrastructure by harnessing GIS and ML capabilities in energy management.

Synopsis

- **Title: GIS based Power Distribution System using ML**
- **Domain and Sub-domain: Machine Learning and Power Distribution**
- **Objectives:**
  1. To manage power distribution of the electrical network remotely.
  2. To develop Machine Learning models to accurately predict consumer power demand patterns, aiding in optimal load management and resource planning.
  3. To identify inefficiencies and potential cost-saving opportunities in the power distribution system.
  4. Optimize power distribution systems using Geographic Information System (GIS) data combined with Machine Learning.
- **Abstract:**

The traditional power distribution network often struggles with resource inefficiency and suboptimal allocation. This results in higher energy losses, longer restoration times during outages, and increased operational costs. Aging equipment and limited visibility into the network exacerbate these issues, leading to decreased reliability. In India, the electric power industries have been developing power transmission system to follow up with a rapid growth of the power demand due to increasing population pressure on this sector. With the advancement of GIS technology, electricity asset system in India is going to be much more benefited in power system planning and management. The integration of GIS with electric utilities is tremendously improving the planning and operation of the power system. The project proposes a traditional approach to optimize power distribution systems using Geographic Information System (GIS) data combined with Machine Learning. A GIS-based power distribution system using machine learning (ML) involves integrating geographical information system (GIS) data with ML algorithms to optimize power distribution.

ML can help predict demand, identify potential faults, and enhance energy efficiency. It can analyze historical data to optimize maintenance schedules and predict equipment failures, improving the overall reliability of the power distribution network. The objective is to develop a robust framework that leverages spatial data from GIS platforms to optimize the operation and maintenance of power distribution networks. The proposed system aims to enhance grid reliability, optimize load balancing, and minimize energy losses. Through the application of advanced neural network architectures, the model learns from historical data and real-time inputs to predict electricity demand patterns and potential fault locations. It contributes to the ongoing efforts towards building a smarter and more sustainable power infrastructure by harnessing the capabilities of GIS and Machine Learning in the domain of energy management.

- **Keywords:**

  1. Power Distribution
  2. ML:- Machine Learning
  3. GIS :- Geographic Information System
  4. MBC :- Metering, Billing and Collections
  5. QGIS :- Quantum Geographic Information System

- **Problem Definition:**

  The adoption of GIS in power distribution aims to enhance overall system efficiency, reliability, and sustainability while providing utilities with data-driven insights for better decision-making and improved customer service. GIS-based power distribution system that integrates Machine Learning Libraries as the underlying technology to optimize and improve the efficiency of power distribution networks. Several usages of GIS in power utilities such as automated route selection for the construction of new power lines which uses a dynamic programming model for route optimization, load forecasting and optimizing planning of substation's location and capacity with comprehensive algorithm which involves an accurate small-area electric load forecasting procedure and simulates the different cost functions of substations.

- **List of Modules:**
  1. Data Collection and Preprocessing
  2. GIS Integration Module
  3. Load Forecasting Module
  4. Fault Detection and Localization Module
- **Current Market Survey:**
  In India, the electric power industries have been developing power transmission system to follow up with a rapid growth of the power demand due to increasing population pressure on this sector. The electric distribution system is dedicated to delivering electrical energy to end users. Many companies and electrical departments in India are using GIS techniques for asset mapping and consumer indexing. Use of GIS will simplify easily updatable and manageable database to cable to the needs of monitoring and sustain reliable quality power supply, effectual MBC (Metering, billing and collections), comprehensive energy audit, theft recognition and a decrease of transmission and distribution losses.All these measures will ultimately advance the overall internal effectual and help accelerate attaining commercial feasibility. A new period of higher implication has arrived for the global positioning system (GPS) and GIS functions at electric utilities mapping. Improved hardware, software, and networking technology have made prospects for the utility industry to form and benefit from more comprehensive and sophisticated GIS. GIS applications have changed from their foundation in map production to advanced analysis tools for planning and operations. To a degree never equaled before, utility managers are looking to their GIS programs, occupied with progressively accurate data collected by global positioning system (GPS) technology, before creating any kind of decisions in the urban and rural areas. Power Distribution Company ensures availability of electricity to end consumers.
  1. It deals with HT and LT electrical networks and related consumers spread over wide geographical area.

2. Its business operation cycle: Power purchase – supply – metering - billing - revenue collection.
3. In past, Discoms have been using conventional Physical Area Maps for proper O and M and management of the system.

- **Scope of the Project:**

1. Predicting power demand and load forecasting in specific geographic regions using historical data.
2. Identifying optimal locations for new power infrastructure, such as substations or transformers, using GIS and machine learning techniques.
3. Building a system to optimize power distribution routes for efficient energy transfer.
4. Developing a geospatial visualization platform to monitor and manage power distribution assets.

- **Literature Survey:**

1. Emanuele Fabbiani "Machine learning approaches for energy distribution and planning" PhD Thesis –cycle XXXIII November 2020
2. Ning Zhao , Fengqi You , "New York State's 100electricity transition planning under uncertainty using a data-driven multistage adaptive robust optimization approach with machine learning" Available on 17 March 2021
3. Alessandro Bosisio, Matteo Moncecchi, Andrea Morotti, Marco Merlo "Machine Learning and GIS Approach for Electrical Load Assessment to Increase Distribution Networks Resilience" article Published on 8 July 2021.
4. Aanand Kumbhar, Pravin G. Dhawale , Shobha Kumbhar, Uday Patil, Pravin Magdum "A comprehensive review: Machine learning and its application in integrated power system " Review article 7 September 2021

5. WenBin Wang, Jie Wei , Yu Zhu , ShiYang Zhou " Power distribution equipment and defect identification technology based on deep learning" . Journaal of Physics: Conference Series, Volume 2030, 2021 International Conference on Electrical Engineering and Computer Technology (ICEECT 2021) 20-22 August 2021, Qingdao, China.

- **Software and Hardware Requirement of the Project:**

  Software:

  1. QGIS
  2. Python

  Hardware:

  1. CPU / GPU
  2. RAM: 8 GB
  3. 1GB hard drive Space

- **Contribution to Society:**

  1. Reliable Power Supply: By analyzing data from GIS and using machine learning techniques, the project can identify potential faults, predict power outages, and help in quicker restoration, ensuring a more reliable power supply to consumers.
  2. Cost Savings: Improved efficiency and reduced downtime can result in cost savings for both energy providers and consumers, potentially leading to more affordable electricity rates.
  3. Grid Resilience: The project can contribute to making the power grid more resilient to natural disasters and other emergencies, ensuring a stable power supply during critical situations.

- **Probable Date of Project Completion:** April 2024

- **Outcome of the Project:**

  1. By optimizing power distribution based on consumer demand patterns, consumers experience more reliable electricity supply, leading to increased satisfaction.
  2. GIS analytics can segment consumers based on geographical location and usage patterns, allowing power companies to design tailored services and demand response program.
  3. By implementing cost-effective strategies, such as load balancing and demand response, utilities can reduce operational expenses and pass on the benefits to consumers. To manage and map the location of millions of connections over transformers remotely to reduce the cost.

Abbreviation

ML	Machine Learning
GIS	Geographic Information System
MBC	Metering, Billing and Collections
QGIS	Quantum Geographic Information System

List of Figures

List of Tables

INDEX

CHAPTER 1

INTRODUCTION

1.1 PROJECT IDEA

The adoption of GIS in power distribution aims to enhance overall system efficiency, reliability, and sustainability while providing utilities with data-driven insights for better decision-making and improved customer service. GIS-based power distribution system that integrates Machine Learning Libraries as the underlying technology to optimize and improve the efficiency of power distribution networks.

Several usages of GIS in power utilities such as automated route selection for the construction of new power lines which uses a dynamic programming model for route optimization, load forecasting and optimizing planning of substation's location and capacity with comprehensive algorithm which involves an accurate small-area electric load forecasting procedure and simulates the different cost functions of substations.

Nowadays, for developing large scale electrical network geodatabases, robust methodologies that integrate traditional data conversion techniques with sophisticated project administration mechanism to enable efficient workflow management and quality control is required. Technological advancement is taking place much faster in generation and transmission sectors. With the radical changes that the electric utility industry is facing, customer choice has become the buzzword for the entire country. Nearly every stage is implementing limited choice programs, choice pilots or at least debating choice. Nowadays, in many developing countries several power utilities are continuing with old conventional manual systems in the distribution sector. The network asset maps in many cities are not properly updated, the customer bills and related data is inaccurate and the details of network assets and facilities are unavailable and not in an appropriate format so that it's proper management are very difficult to the electrical department. [1]

Many companies and electrical departments in India are using GIS techniques for asset mapping and consumer indexing, which explains the concept, the systems' requirements and varied application of this tool in asset management, load predicting, delivery system planning and customer care etc. Use of GIS will simplify easily updatable and manageable database to cable to the needs of monitoring and sustain reliable quality power supply, effectual MBC (Metering, billing and collections), comprehensive energy audit, theft recognition and a decrease of transmission and distribution losses. All these measures will ultimately advance the overall internal effectual and help accelerate attaining commercial feasibility. [1]

A new period of higher implication has arrived for the global positioning system (GPS) and GIS functions at electric utilities mapping. Improved hardware, software, and networking technology have made prospects for the utility industry to form and benefit from more comprehensive and sophisticated GIS. GIS applications have changed from their foundation in map production to advanced analysis tools for planning and operations. To a degree never equaled before, utility managers are looking to their GIS programs, occupied with progressively accurate data collected by global positioning system (GPS) technology, before creating any kind of decisions in the urban and rural areas. [1]

In an era characterized by increasing energy demands and the imperative for sustainable development, the power distribution sector faces unprecedented challenges. Traditional power distribution systems often grapple with inefficiencies, outages, and the need for real-time adaptability to dynamic demand patterns. To address these issues, this report introduces a groundbreaking approach that leverages Geographic Information Systems (GIS) and machine learning techniques to revolutionize power distribution systems.

GIS technology enables the spatial analysis and visualization of complex geographical data, providing a comprehensive understanding of the distribution network. By integrating machine learning algorithms into this GIS framework, we aim to develop an intelligent decision support system capable of optimizing power distribution operations. This amalgamation of GIS and machine learning holds the potential to enhance fault detection, predictive maintenance, load forecasting, and overall system resilience.

This report delves into the key components of our GIS-based power distribution system, shedding light on the integration of machine learning algorithms for predictive analytics and automated decision-making. Through an in-depth exploration of case studies and practical implementations, we aim to showcase the transformative impact of this approach on the efficiency, reliability, and sustainability of power distribution networks.

By harnessing the power of spatial data analytics and machine learning, our project strives to usher in a new era of intelligent power distribution systems that can adapt to evolving demands, reduce downtime, and contribute to a more sustainable and resilient energy infrastructure. The subsequent sections of this report will delve into the methodology, results, and implications of our GIS-based power distribution system, illustrating the potential for innovation and positive change in the realm of energy distribution.

1.2 MOTIVATION OF THE PROJECT

1. The existing system suffers from manual surveys that are time-consuming and prone to errors, rampant electricity theft, and inefficiencies in determining optimal paths for consumers from the electricity poles.

2. To overcome these challenges by harnessing the power of modern technology.

3. Primary objectives are to eliminate the need for manual surveys through automation, implement accurate load forecasting through machine learning, and establish optimal consumer-to-pole pathways.

4. The motivation behind the development of a GIS-based power distribution system using machine learning stems from the pressing need to address challenges inherent in conventional power distribution networks. Traditional systems often struggle with issues such as inefficient asset management, prolonged downtimes, and an inability to adapt swiftly to dynamic changes in demand and environmental conditions.

5. GIS technology provides a spatial framework that allows for a holistic understanding of the distribution network, incorporating geographical data to enhance decision-making. Integrating machine learning into this GIS infrastructure introduces a predictive and adaptive dimension, enabling the power distribution system to learn from historical data, anticipate future demands, and optimize operations in real-time.

6. The motivation for this project lies in the potential to revolutionize the power distribution landscape, offering solutions that go beyond reactive measures to proactively manage and enhance the reliability, efficiency, and resilience of the electrical grid. By leveraging GIS and machine learning, we aim to empower utilities and operators to make data-driven decisions, reduce downtime, improve maintenance strategies, and ultimately contribute to the development of a smarter, more sustainable energy infrastructure.

7. This fusion of spatial intelligence and machine learning capabilities holds the promise of transforming the traditional power distribution paradigm into a dynamic and adaptive system capable of meeting the evolving needs of our energy landscape.

8. Conventional power distribution systems face inefficiencies and challenges in adapting to dynamic demands. The integration of GIS technology and machine learning aims to revolutionize this landscape. GIS provides a spatial understanding of the network, while machine learning offers predictive capabilities.

9. The motivation lies in enhancing the reliability, efficiency, and resilience of power distribution by enabling data-driven decision-making, reducing downtime, and contributing to a smarter, more sustainable energy infrastructure. This fusion of spatial intelligence and machine learning has the potential to transform traditional systems into adaptive, proactive networks aligned with the evolving needs of the energy sector.

CHAPTER 2

LITERATURE SURVEY

2.1 LITERATURE SURVEY

Paper Name: GIS in electrical asset mapping: The case of bhadohi, India.

Author Name: Praveen Kumar Rai, Chandan SINGH

Year: 2016

Technology: GIS, GPS

Main Finding/Summary: Finding shortest optimum path for power distribution line.It is also very helpful to improve the customer complaints, long-standing faults brought to a minimum, to stop power theft (in local language it is called Katia practice), and to provide better facilities to the each consumer[1]

Paper Name: Power distribution equipment and defect identification technology based on deep learning

Author Name: WenBin Wang

Year: 2021

Technology: M210 UAV airborne visible light camera

Main Finding/Summary: In this paper, a method of distribution network equipment and fault identification based on edge calculation is proposed. The experimental results show that the accuracy of the proposed algorithm reaches 87.9 percentage and the recall rate reaches 83.3 percentage.[2]

Paper Name: A comprehensive review: Machine learning and its application in integrated power system

Author Name: Aanand Kumbhar, Pravin Dhavle, Uday Patil

Year: 2021

Technology: Types and Applications of ML

Main Finding/Summary: This paper given review of application of ML in integrated power system.This paper also gives detailed information about the types of machine learning techniques. It summarize the uses of ML to blame finding in the force framework, which is the essential advance of the force framework security and control.[3]

Paper Name: Machine Learning and GIS Approach for Electrical Load Assessment to Increase Distribution Networks Resilience

Author Name: Alessandro Bosisio, Matteo Moncecchi, Andrea Morotti, Marco Merlo

Year: 2021

Technology: ML Techniques, GIS

Main Finding/Summary: The study explores the possibility of predicting the power profile of MV/LV substations using machine learning approaches and GIS information. [4]

Paper Name: Platform-Independent Web Application for Short-Term Electric Power Load Forecasting on 33/11 kV Substation Using Regression Tree

Author Name: Venkataramana Veeramsetty , Sai Pavan Kumar and Surender Reddy Salkuti

Year: 2022

Technology: Regression Algorithm

Main Finding/Summary: Developed to forecast the load with good accuracy irrespective of the weather conditions and types of the day by incorporating features such as season, temperature, humidity, and day-status.[5]

Sr. No.	Paper Title	Author	Year of Publication	Technology
1	GIS in electrical asset mapping: The case of bhadohi, India.	Praveen Kumar Rai	2016	GIS and GPS
2	Power distribution equipment and defect identification technology based on deep learning	WenBin Wang	2021	M210 UAV airborne visible light camera
3	A comprehensive review: Machine learning and its application in integrated power system	Aanand Kumbhar, Pravin Dhavle, Uday Patil	2021	Types and Applications of ML
4	Machine Learning and GIS Approach for Electrical Load Assessment to Increase Distribution Networks Resilience	Alessandro Bosisio, Matteo Moncecchi, Andrea Morotti, Marco Merlo	2021	ML Techniques, GIS
5	Platform-Independent Web Application for Short-Term Electric Power Load Forecasting on 33/11 kV Substation Using Regression Tree	Venkataramana Veeramsetty, Sai Pavan Kumar and Surender Reddy Salkuti	2022	Regression Algorithm

Table 2.1: Comparative Analysis

CHAPTER 3

PROBLEM DEFINITION AND SCOPE

3.1 PROBLEM STATEMENT

The adoption of GIS in power distribution aims to enhance overall system efficiency, reliability, and sustainability while providing utilities with data-driven insights for better decision-making and improved customer service. GIS-based power distribution system that integrates Machine Learning Libraries as the underlying technology to optimize and improve the efficiency of power distribution networks. Several usages of GIS in power utilities such as automated route selection for the construction of new power lines which uses a dynamic programming model for route optimization, load forecasting and optimizing planning of substation's location and capacity with comprehensive algorithm which involves an accurate small-area electric load forecasting procedure and simulates the different cost functions of substations.

3.1.1 Goals and Objectives

1. Goal

   - Remote Power Distribution Management:
     Goal: Implement a remote power distribution management system that allows real-time monitoring and control of the electrical network, ensuring efficient and reliable power supply.
   - Machine Learning Models for Power Demand Prediction:
     Goal: Develop and deploy accurate Machine Learning models that predict consumer power demand patterns with a high degree of precision. These models should assist in optimizing load management and resource planning, leading to cost savings and better resource allocation.
   - Inefficiency Identification and Cost-saving Opportunities:
     Goal: Identify and address inefficiencies in the power distribution system through data analysis and process optimization. Identify cost-saving opportunities by minimizing losses, improving maintenance, and enhancing operational efficiency.
   - GIS and Machine Learning Integration for Optimization:
     Goal: Integrate Geographic Information System (GIS) data with Ma-

chine Learning algorithms to optimize power distribution systems. This integration should help improve network reliability, reduce downtime, and enhance overall system performance.

2. Objectives

   - To manage power distribution of the electrical network remotely.
   - To develop Machine Learning models to accurately predict consumer power demand patterns, aiding in optimal load management and resource planning.
   - To identify inefficiencies and potential cost-saving opportunities in the power distribution system.
   - Optimize power distribution systems using Geographic Information System (GIS) data combined with Machine Learning.

3.1.2 Statement of scope

1. Power Demand Prediction and Load Forecasting: Collect historical power consumption data for specific geographic regions. Develop Accurate Machine Learning models that analyze this date to predict future power demand and load patterns. Utilize these predictions for efficient load management and resource planning to ensure a reliable power supply.

2. Optimal Infrastructure Location Identification: Combine Geographic Information System (GIS) data with machine learning techniques to identify the best locations for new power infrastructure, such as substations and transformers. Consider factors like current grid capacity, future demand growth, and environmental impact in the decision-making process.

3. Power Distribution Route Optimization: Develop an intelligent system that optimizes the routing of power distribution to maximize energy transfer efficiency. Take into account factors such as load balancing, minimizing losses, and considering the condition of existing infrastructure.Implement real-time adjustments to adapt to changing conditions.

4. Geospatial Visualization Platform: Create a user-friendly geospatial visualization platform for monitoring and managing power distribution assets. Integrate GIS data, real-time sensor data, and data from other relevant sources to provide a comprehensive view of the power distribution network.Include features for asset tracking, fault detection, and maintenance scheduling.

5. Data Integration and Management: Implement a robust data management system that efficiently collects, stores, and updates data from various sources, including historical records, GIS databases, and real-time sensors. Ensure data quality, consistency, and security to support accurate decision-making and system operation.

3.2 SOFTWARE CONTEXT

1. Data Analytics and Machine Learning: Python with libraries such as Pandas, NumPy, and Scikit-Learn for data analysis and machine learning model development. Jupyter Notebooks for data exploration and model development. TensorFlow or PyTorch for deep learning applications in load forecasting.

2. Geospatial Analysis and GIS: Geographic Information System (GIS) software like ArcGIS, QGIS, or open-source GIS libraries for spatial data analysis. GIS databases to store and manage geospatial data. Python libraries like GeoPandas and Fiona for geospatial data manipulation.

3. Real-Time Data Processing: Apache Kafka or Apache Pulsar for real-time data streaming and processing. Apache Spark for large-scale data processing.

4. Optimization and Routing: Optimization libraries like Google OR-Tools or Gurobi for solving routing and load balancing problems. Graph databases like Neo4j for representing and analyzing distribution networks.

5. Database Management: Relational databases (e.g., PostgreSQL, MySQL) for structured data storage. Time-series databases (e.g., InfluxDB) for handling time-series data from sensors. NoSQL databases (e.g., MongoDB) for flexibility in data storage.

6. Visualization and Monitoring: Web-based visualization frameworks like D3.js, Leaflet, or Mapbox for geospatial data visualization. Custom dashboards developed with libraries like Flask or Django. Real-time monitoring and alerting using tools like Grafana or Kibana.

7. Security and Access Control: Implement robust security protocols to safeguard data and systems. Role-based access control (RBAC) to manage user access to sensitive information.

8. Collaboration and Communication: Collaboration tools (e.g., Slack, Microsoft Teams) for team communication and project management. Version control systems like Git (GitHub, GitLab) for code management and collaboration.

9. Documentation and Reporting: Documentation tools (e.g., Confluence, Jira) for keeping track of project progress and requirements. Reporting and data visualization tools (e.g., Tableau, Power BI) for generating reports and insights.

3.3 MAJOR CONSTRAINTS

Implementing a project of the scale and complexity described in the context of power distribution optimization and management can encounter several major constraints and challenges. Some of the key constraints include:

1. Regulatory and Compliance Constraints: Adherence to local, regional, and national regulations and standards related to power distribution, environmental impact, and safety can be a significant constraint. Non-compliance can lead to legal and financial repercussions.

2. Data Quality and Availability: Data quality is crucial for accurate predictions and decision-making. Incomplete, inaccurate, or outdated data can hinder project success. Additionally, data privacy and security regulations may limit access to certain datasets.

3. Integration Challenges: Integrating various software systems, data sources, and legacy infrastructure can be technically challenging and time-consuming.

Incompatibility between existing systems and new technologies is a common constraint.

4. Scalability: Ensuring that the system can scale to handle increased power demand and infrastructure expansion over time is a significant challenge. The project must be designed to accommodate future growth.

5. Budget Constraints: Budget limitations can constrain the project in terms of technology choices, resource allocation, and the pace of implementation. Balancing cost-effectiveness with project goals is essential.

6. Expertise and Talent: A shortage of skilled professionals with expertise in geospatial analysis, machine learning, and power distribution systems can be a constraint. Recruiting and retaining top talent can be challenging.

7. Stakeholder Alignment: Ensuring alignment among all stakeholders, including utility companies, regulatory bodies, and local communities, can be difficult. Differing interests and priorities may pose challenges in project approval and implementation.

8. Environmental and Social Impact: Projects that involve the placement of new power infrastructure can face environmental and social constraints. Concerns about the ecological impact, land use, and community resistance can slow down or block project implementation.

9. Security and Privacy: Protecting sensitive data and ensuring the security of the power distribution system against cyber threats are crucial. Security constraints can add complexity to the project and increase costs.

10. Data Transmission and Latency: Real-time data transmission and low-latency requirements for power distribution systems can be challenging, particularly in remote or less-connected areas.

3.4 METHODOLOGIES OF PROBLEM SOLVING AND EFFICIENCY ISSUES

Certainly, let's delve into methodologies of problem-solving and efficiency issues for the project described: Methodologies of Problem Solving:

1. Data-Driven Approach:
    - Data Collection: Collect and curate relevant historical and real-time data, including power consumption, infrastructure information, and geospatial data.
    - Data Analysis: Employ data analysis techniques to identify trends, patterns, and anomalies in the data, helping to make informed decisions. Predictive Modeling: Develop machine learning models for load forecasting, asset optimization, and route planning based on the data analysis.
2. Cross-Disciplinary Collaboration:
    - Interdisciplinary Teams: Form cross-functional teams with experts in power engineering, data science, geospatial analysis, and software development to bring diverse perspectives to problem-solving.
    - Stakeholder Engagement: Collaborate with stakeholders, including utility companies, regulatory bodies, and local communities, to understand their concerns and incorporate their feedback.
3. Risk Management:
    - Risk Assessment: Identify potential risks and constraints and assess their impact on the project's objectives and timelines.
    - Risk Mitigation: Develop strategies to mitigate and manage risks, such as redundancy planning for critical infrastructure components.
4. Agile Project Management:

- Iterative Development: Adopt agile methodologies for project management to facilitate continuous improvement and adaptability in response to changing requirements.
- Scrum or Kanban: Consider using Scrum or Kanban frameworks to manage tasks and streamline project workflows.

5. Efficiency Issues and Solutions:
   - Data Efficiency: Data Compression: Implement data compression techniques to reduce storage and transmission requirements, especially for large geospatial and time-series data.
   - Data Cleaning and Preprocessing: Automate data cleaning and preprocessing to enhance data quality and reduce manual effort.
6. Scalability and Performance:
   - Cloud Resources: Utilize cloud computing platforms to dynamically allocate resources as needed for scalability and improved system performance.
   - Load Balancing: Implement load balancing techniques to distribute computing workloads efficiently across servers.
7. Energy Efficiency:
   - Green Infrastructure: Consider the use of energy-efficient and environmentally friendly infrastructure components, such as smart transformers and substations.
   - Energy Optimization Algorithms: Develop algorithms for optimizing power distribution routes and load balancing to minimize energy losses.
8. Security and Privacy:
   - Cybersecurity Protocols: Implement robust cybersecurity protocols to protect the power distribution system and sensitive data from cyber threats.

- Privacy Measures: Employ data anonymization and encryption techniques to protect user privacy while utilizing data for analysis.

9. Time Efficiency:

   - Real-time Data Processing: Utilize stream processing technologies to handle real-time data efficiently, ensuring low-latency responses for monitoring and control.
   - Automated Decision Support: Develop decision support systems that automate routine decision-making processes, saving time for human operators.

10. Resource Optimization:

    - Resource Allocation Algorithms: Develop optimization algorithms for resource allocation, such as scheduling maintenance tasks to minimize downtime.
    - Supply Chain Efficiency: Streamline the supply chain for infrastructure components to reduce lead times and procurement costs.

11. Community Engagement:

    - Communication Strategies: Develop communication and engagement strategies to address community concerns and gain public support for the project.
    - Transparency: Maintain transparency in project activities and decisions to build trust within the community.
    - Efficiently solving the complex problems in power distribution optimization requires a multi-faceted approach, including data-driven decision-making, cross-disciplinary collaboration, risk management, and a focus on addressing the identified efficiency issues. Flexibility and adaptability are essential to navigate unexpected challenges and evolving project requirements.

3.5 SCENARIO IN WHICH MULTI-CORE, EMBEDDED AND DISTRIBUTED COMPUTING USED

In a power distribution optimization multi-core, embedded, and distributed computing can be used in various scenarios to improve efficiency, reliability, and performance. Here's a scenario that combines these computing paradigms:

3.5.1 Scenario 1 : Real-Time Load Balancing and Fault Handling

In this scenario, the project aims to enhance the real-time load balancing and fault handling capabilities of the power distribution system. Multi-core, embedded, and distributed computing technologies are utilized to achieve this goal:

1. Multi-Core Computing:
   - Load Balancing Algorithms: Multi-core processors are employed to run load balancing algorithms, which continuously analyze the power demand across the distribution network. Multiple threads or cores can perform load analysis concurrently, allowing for quicker response to load changes.
   - Predictive Load Forecasting: Multi-core servers are used to execute advanced machine learning models for load forecasting. These models take historical data and real-time information to predict future power demand accurately. The parallel processing power of multi-core systems speeds up the modeling process.
2. Embedded Computing:
   - Distributed Intelligent Controllers: Embedded systems with real-time capabilities are placed within substations and transformers. These controllers are equipped with embedded processors and sensors that monitor local load conditions, voltage levels, and equipment health.
   - Local Fault Detection: Embedded systems detect local faults or anomalies within their respective substations or equipment. When a fault is

detected, the embedded controller initiates automatic isolation or corrective actions to minimize the impact on the wider grid.

- Communication Protocols: These embedded systems use communication protocols to exchange information with the central control system. They can provide real-time updates on local conditions and operational status.

3. Distributed Computing:

- Grid-Wide Coordination: Distributed computing clusters at the central control center process data from embedded controllers across the distribution network. They continuously evaluate the status of different substations and transformers to make real-time load balancing decisions.
- Fault Resolution: Distributed computing systems play a key role in coordinating fault resolution efforts. When a fault is detected, the central system communicates with relevant embedded controllers to isolate the affected area and reroute power to minimize disruptions.
- Redundancy and Failover: Distributed computing systems can incorporate redundancy and failover mechanisms to ensure the continuous operation of the control center. This helps maintain control even in the event of hardware or network failures.
- Historical Data Analysis: Distributed systems can store and analyze historical fault data to identify patterns and proactively address recurring issues in the distribution network.

3.5.2 Scenario 2: Grid Resilience and Demand Response Management

In this scenario, the project aims to enhance the resilience of the power distribution grid and improve demand response capabilities. This involves efficiently managing energy demand during peak periods, reducing energy losses, and ensuring uninterrupted power supply, especially during extreme weather events or equipment failures.

1. Multi-Core Computing:

    - Real-Time Load Balancing: Multi-core servers are utilized in the data center to process real-time data from various substations and endpoints. Load balancing algorithms run on these servers, optimizing energy distribution across the grid by considering factors like current demand, available capacity, and equipment health.
    - Predictive Maintenance: Machine learning models for predictive maintenance run on multi-core processors to continuously analyze sensor data from embedded devices within transformers and substations. These models predict equipment failures and trigger maintenance actions before critical issues arise.

2. Embedded Computing:

    - Local Demand Response: Embedded control systems installed in smart meters and IoT devices at consumer premises play a crucial role in managing energy demand. During peak periods or grid instability, these embedded systems communicate with central servers to adjust the power consumption of specific devices or appliances, supporting demand response initiatives.
    - Resilient Substations: Substations are equipped with embedded control systems that can automatically isolate faulty components, reroute power, and engage backup power sources to maintain power supply in the event of equipment failures or natural disasters.

3. Distributed Computing:

    - Distributed Energy Resources (DERs) Integration: Distributed computing clusters aggregate data from various DERs, such as solar panels and wind turbines, to optimize their contribution to the grid. This involves balancing energy generation, storage, and consumption across different locations to enhance grid stability.

- Fault Detection and Recovery: A distributed computing infrastructure connects substations and central control systems. It enables fault detection and coordination in real time. When a fault is detected, the distributed system collaborates to isolate the affected area and reroute power to minimize disruptions.

4. Geospatial Data Analysis:

   - Optimal Infrastructure Placement: Geospatial data analysis using multi-core servers assists in identifying optimal locations for new infrastructure components, such as substations and transformers. These locations are strategically chosen to improve the grid's ability to withstand adverse weather events.
   - Geospatial Visualization: A distributed geospatial visualization platform helps operators monitor the distribution assets and their geospatial context in real time. It provides a comprehensive view of the grid's status, aiding in situational awareness during emergencies.

This scenario highlights how a combination of multi-core, embedded, and distributed computing technologies is integral to ensuring grid resilience and efficient demand response management in the power distribution optimization project. These computing paradigms work together to enable real-time decision-making, equipment health monitoring, and data-driven grid control, ultimately enhancing the reliability and adaptability of the power distribution system.

3.6 OUTCOME

1. Enhanced Reliability and Consumer Satisfaction: Reduced Outages:

   - By optimizing power distribution based on consumer demand patterns, the system becomes more resilient to fluctuations and faults, reducing the frequency and duration of power outages.

- Improved Voltage Control: With better load management, voltage levels are more stable, leading to fewer voltage fluctuations that can harm sensitive electronic devices.
- Increased Satisfaction: Consumers experience fewer disruptions, improved power quality, and a more reliable electricity supply. This, in turn, results in higher consumer satisfaction and trust in the utility company.

2. Tailored Services and Demand Response:

   - Geographical Segmentation: GIS analytics segment consumers based on their geographical location and usage patterns. This enables power companies to offer tailored services and demand response programs to different consumer groups.
   - Energy Efficiency: Consumers benefit from personalized energy-saving recommendations and incentives, resulting in lower energy bills and reduced environmental impact.
   - Peak Load Management: Utilities can implement demand response initiatives to reduce peak load periods, resulting in lower costs and potentially passing on the savings to consumers.

3. Cost Reduction and Remote Management:

   - Load Balancing: Implementing cost-effective strategies like load balancing and demand response helps utilities reduce operational expenses. Load balancing optimizes the use of existing infrastructure, delaying the need for costly upgrades.
   - Operational Efficiency: The ability to manage and map the location of millions of connections and transformers remotely enhances operational efficiency. This reduces the need for physical site visits, resulting in significant cost savings.
   - Economic Benefits: Cost savings translate into lower operational costs for the utility, which can be passed on to consumers through reduced electricity rates or more modest rate increases.

4. Sustainability and Environmental Benefits:

   - Renewable Integration: The project may facilitate the integration of renewable energy sources into the grid more efficiently, reducing reliance on fossil fuels and decreasing greenhouse gas emissions.
   - Energy Conservation: Demand response programs and energy-saving initiatives encourage consumers to use electricity more efficiently, contributing to energy conservation and reduced environmental impact.

5. Grid Resilience and Disaster Recovery:

   - Improved Resilience: Optimal infrastructure placement and fault detection capabilities enhance the grid's resilience against natural disasters and equipment failures.
   - Faster Recovery: In the event of outages, the system can quickly identify and isolate faults, allowing for faster restoration of power to affected areas, benefiting consumers in times of need.

6. Data-Driven Decision-Making:

   - Efficient Planning: Data analytics and geospatial visualization tools support efficient planning for maintenance and upgrades, reducing downtime and minimizing disruptions to consumers.
   - Proactive Maintenance: Predictive maintenance models enable utilities to address potential equipment failures before they occur, improving reliability and reducing service interruptions.

3.7 APPLICATIONS

1. Tata Power
2. Adani Power
3. Reliance Infrastructure
4. NTPC (National Thermal Power Corporation)

5. BSES Rajdhani Power Limited (BRPL)
6. BSES Yamuna Power Limited (BYPL)
7. Delhi Transco Limited (DTL)
8. Uttar Pradesh Power Corporation Limited (UPPCL)
9. Maharashtra State Electricity Distribution Company Limited (MSEDCL)
10. Gujarat Urja Vikas Nigam Limited (GUVNL)
11. West Bengal State Electricity Distribution Company Limited (WBSEDCL)
12. Kerala State Electricity Board (KSEB)
13. Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO)
14. Karnataka Power Transmission Corporation Limited (KPTCL)
15. Andhra Pradesh Eastern Power Distribution Company Limited (APEPDCL)
16. Punjab State Power Corporation Limited (PSPCL)
17. Haryana Vidyut Prasaran Nigam Limited (HVPNL)
18. Rajasthan Rajya Vidyut Utpadan Nigam Limited (RVUNL)

3.8 SOFTWARE RESOURCES REQUIRED

Platform :

1. Operating System
2. IDE
3. Programming Language

3.9 HARDWARE RESOURCES REQUIRED

Sr. No.	Parameter	Minimum Requirement	Justification
1	CPU / GPU	2.5 GHz	The CPU is responsible for initial data preprocessing tasks, including data cleaning, transformation, and feature engineering. It prepares the data for input into machine learning algorithms.
2	RAM	8 GB	GIS data, especially when dealing with spatial information, can be large and complex. Sufficient RAM is essential for efficiently loading, storing, and managing this data. Insufficient RAM may lead to slow data access and processing, impacting system performance.
3	Hard Drive Space	1 GB	GIS data, including geographical information and attributes, can be extensive. Large datasets for spatial analysis, power distribution network data, and related information would quickly consume more than 1 GB of storage space.

Table 3.1: Hardware Requirements

CHAPTER 4

SOFTWARE REQUIREMENT SPECIFICATION

4.1 INTRODUCTION

4.1.1 Purpose and Scope of Document

4.1.1.1 Purpose

A Geographic Information System (GIS) based power distribution system combined with machine learning can serve several purposes in the field of energy and power management. Here are some of the key purposes and benefits of using machine learning in a GIS-based power distribution system: [6]

1. Predictive Maintenance: Machine learning algorithms can analyze historical data and sensor information from the power distribution network to predict when equipment, such as transformers or circuit breakers, is likely to fail. This helps utilities schedule maintenance activities more efficiently and reduce downtime.

2. Fault Detection and Localization: Machine learning models can detect and locate faults in the power distribution system by analyzing data from sensors and smart meters. This information can be displayed on GIS maps, allowing for quicker response and restoration of service.

3. Load Forecasting: Machine learning can be used to forecast energy demand and load patterns in different areas of the distribution network. This helps utilities optimize resource allocation, plan for peak demand, and reduce energy wastage.[4]

4. Energy Theft Detection: Machine learning can identify irregular consumption patterns that may be indicative of energy theft or tampering with meters. This helps utilities reduce revenue losses and maintain the integrity of their distribution system.

5. Grid Optimization: Machine learning can optimize the distribution grid by reconfiguring it in real-time to minimize energy losses, improve power quality, and enhance the reliability of the system.

6. Asset Management: Machine learning can assist in asset management by prioritizing the replacement or upgrade of aging infrastructure components based on data-driven insights and cost-effectiveness.
7. Outage Management: GIS-based maps, combined with machine learning, can help utilities track and manage outages more effectively. Machine learning can predict the impact of outages and suggest optimal routes for repair crews.
8. Renewable Energy Integration: Machine learning can help utilities integrate renewable energy sources, such as solar and wind, into the distribution system by forecasting generation patterns and optimizing grid operations accordingly.
9. Customer Engagement: By analyzing customer data, machine learning can help utilities offer personalized services, such as energy-saving tips or demand response programs, to improve customer satisfaction and energy efficiency.
10. Regulatory Compliance: Machine learning can assist utilities in complying with regulatory requirements, such as reporting on environmental impacts and monitoring grid reliability.

4.1.1.2 Scope

1. The scope of GIS-based power distribution systems combined with machine learning is broad and continues to expand as technology and data capabilities advance. This integration has the potential to transform the power distribution industry by improving efficiency, reliability, and sustainability. Here are some aspects of the scope for GIS-based power distribution systems with machine learning:
2. Predictive Maintenance: Machine learning can predict equipment failures, reducing downtime and maintenance costs.
3. Grid Optimization: Real-time optimization of the distribution grid can lead to reduced energy losses and improved power quality.
4. Load Forecasting: Accurate load forecasting helps utilities plan for peak demand and optimize resource allocation.

5. Outage Management: Improved outage detection and management can reduce the duration of disruptions.
6. Asset Management: Machine learning can assist in prioritizing asset replacement and upgrade projects.
7. Energy Theft Detection: Identifying irregular consumption patterns helps prevent revenue losses.
8. Renewable Energy Integration: Machine learning can optimize the integration of renewable energy sources into the grid.
9. Regulatory Compliance: Compliance with regulatory requirements can be enhanced through data analysis and reporting.
10. Customer Engagement: Machine learning can provide personalized services to customers, enhancing satisfaction and energy efficiency.
11. Data Analysis and Visualization: GIS-based systems allow for the visualization of data, which aids in decision-making and system monitoring.
12. Smart Grid Development: Machine learning can support the development of smart grids by enabling real-time data analysis and control.
13. Environmental Impact Assessment: Utilities can assess and minimize the environmental impact of their operations.
14. Cybersecurity: Machine learning can be used to enhance the cybersecurity of power distribution systems, protecting against cyber threats.
15. Microgrid Management: Machine learning can optimize the operation of microgrids, which are becoming more common for local energy generation and distribution.
16. Integration with IoT Devices: GIS and machine learning can work with IoT devices and sensors to gather real-time data and improve system management.
17. Workforce Optimization: Machine learning can optimize the allocation of maintenance crews and resources.

18. Research and Development: Ongoing research in machine learning techniques and data analysis methods can lead to continuous improvement in power distribution systems.

19. The scope of GIS-based power distribution system machine learning is dynamic and evolving. As technology advances, the integration of GIS and machine learning will likely lead to even more sophisticated applications and capabilities. Utilities and organizations in the energy sector are increasingly recognizing the value of these technologies to improve operations and provide better services to consumers.

4.1.2 Overview of responsibilities of Developer

1. Algorithm Selection: Choosing appropriate Machine Learning algorithms for specific tasks, such as regression for demand forecasting or anomaly detection for fault prediction.

2. Model Development: Designing and implementing machine learning models, including deep learning models for tasks like image recognition or time-series forecasting.Fine-tuning model hyperparameters and architecture for optimal performance.

3. Training and Evaluation: Splitting data into training, validation, and test sets for model training and evaluation.Measuring model performance using metrics relevant to the specific task, such as Mean Absolute Error (MAE) or F1-score.

4. GIS Integration: Integrating GIS mapping tools and features into the system to provide spatial context for the data and results.

5. Testing and Quality Assurance: Conducting thorough testing, including unit testing, integration testing, and system testing to ensure the system functions correctly.Implementing quality assurance processes to catch and resolve bugs and issues.

6. Maintenance and Support: Providing ongoing support and maintenance for the system, addressing issues, and making updates as necessary.

4.2 FUNCTIONAL REQUIREMENTS

- Visualize and Analyze the Power Distribution Network:
  1. GIS Mapping: Create a GIS-based map that displays the entire power distribution network, including substations, transformers, power lines, and customer connections.
  2. Data Integration: Integrate geographical data, network topology, and real-time sensor data to provide a comprehensive view of the network.
  3. Geographic Information System (GIS) software: Utilize GIS software to visualize, query, and analyze the network, allowing operators to identify key network components and their status.
- Predict Power Demand:
  1. Machine Learning Models: Develop ML models to predict power demand based on historical data, weather conditions, and other relevant factors.
  2. Data Sources: Collect historical load data, weather data, and customer information to train and update the prediction models.
  3. Real-time Monitoring: Implement real-time monitoring to adjust predictions as conditions change.
- Detect and Identify Faults:
  1. Fault Detection Algorithms: Implement ML algorithms to detect faults in the power distribution system, such as line failures, transformer issues, or other equipment malfunctions.
  2. Sensor Data: Use data from sensors and remote monitoring systems to identify anomalies or deviations from normal operation.
  3. Fault Localization: Develop algorithms that can pinpoint the exact location of the fault for faster response and maintenance.

- Optimize Power Distribution:

    1. Network Optimization Models: Create ML models that optimize the distribution of power based on various factors, including load, reliability, and cost.
    2. Predictive Maintenance: Use predictive maintenance models to optimize the maintenance schedule, reducing downtime and improving overall system efficiency.
    3. Load Balancing: Ensure that power distribution is balanced to prevent overloads and minimize energy losses.

- Support Decision-Making:

    1. Decision Support Systems (DSS): Implement DSS tools that provide recommendations and insights to operators and decision-makers.
    2. Scenario Analysis: Allow users to simulate different scenarios, such as adding new substations, adjusting load profiles, or introducing renewable energy sources, to assess their impact on the network.

4.3 EXTERNAL INTERFACE REQUIREMENTS (IF ANY)

4.3.1 User Interfaces

1. Dashboard: Create a user-friendly dashboard that provides an overview of the power distribution system. This should include key information such as system status, alerts, and performance metrics.

2. Map View: Integrate a map interface that displays the power grid, including substations, power lines, and other infrastructure. Users should be able to interact with the map to view specific locations and assets.

3. Machine Learning Controls: Incorporate controls for configuring and managing machine learning models. Users may need to select algorithms, set parameters, and monitor model performance.

4. Data Visualization: Use charts and graphs to visualize power consumption trends, grid efficiency, and other relevant data. These visualizations can help users make informed decisions.

5. User Management: Include user authentication and access control features. Administrators should be able to manage user roles and permissions.

4.3.2 Hardware Interfaces

1. GPS and Location Services: If mobile devices or vehicles are part of the system, integrate with GPS and location services for asset tracking and route optimization.

2. Communication Protocols: Support various communication protocols used for device-to-device and device-to-server communication, such as Modbus, MQTT, or RESTful APIs.

3. Data Storage and Servers: Interface with the hardware components responsible for data storage, backup, and the server infrastructure hosting the software components.

4. Control Systems: If applicable, interface with control systems that manage power distribution, including switches, transformers, and other equipment.

4.3.3 Software Interfaces

1. GIS Integration: Interface with Geographic Information Systems (GIS) software to retrieve and display spatial data. This integration should support popular GIS formats and protocols.

2. Machine Learning Framework: Connect with the machine learning framework used for predictive maintenance, load forecasting, or fault detection. This interface allows for model training and execution.

3. Database Integration: Interface with databases that store power grid data, historical records, and machine learning model parameters. This includes SQL or NoSQL databases, as well as cloud-based data storage solutions.

4. APIs: Provide APIs (Application Programming Interfaces) to allow integration with third-party software and hardware, such as weather data services, IoT sensors, or maintenance management systems.

4.3.4 Communication Interfaces

1. GIS Data Interface:
   - GIS APIs: Utilize GIS APIs (e.g., ArcGIS API for Python, OpenLayers) to interact with the GIS database and access spatial data for mapping and analysis.
   - Standard GIS Formats: Support common GIS file formats (e.g., shapefiles, GeoJSON) to import and export geospatial data.
2. Data Ingestion and Integration:
   - Data Ingestion APIs: Develop APIs for ingesting data from various sources, such as sensors, SCADA systems, weather feeds, and customer information systems.
   - ETL (Extract, Transform, Load) Interfaces: Implement ETL processes for cleansing, transforming, and loading data into the system's databases.
3. Machine Learning Model Deployment:
   - Model Serving APIs: Create APIs or endpoints for deploying machine learning models, enabling real-time predictions and analysis.
   - Model Version Control: Establish a system for managing and updating machine learning model versions to ensure accuracy and consistency.
4. Performance Optimization:
   - Optimize communication protocols and data exchange processes to minimize latency and maximize data throughput.
   - Scalability: Ensure that the communication infrastructure can scale to accommodate increased data volume and user load as the system grows.

4.4 NONFUNCTIONAL REQUIREMENTS

4.4.1 Performance Requirements

- Performance: Response Time: The system should provide real-time or near-real-time responses to user queries and data updates.
- Scalability: The system must handle an increasing amount of data and users without significant performance degradation.
- Throughput: It should support a high volume of data processing and analysis, especially during peak usage periods.

4.4.2 Safety Requirements

- Data Security: Protect sensitive customer information and operational data from unauthorized access or breaches.
- Authentication and Authorization: Implement robust user authentication and authorization mechanisms to control access to system features and data.
- Compliance: Adhere to industry and regulatory standards for data security and privacy.

4.4.3 Security Requirements

- Data Integrity: Implement mechanisms to verify data integrity to detect and prevent unauthorized alterations to data during storage or transmission.Use hash functions or digital signatures to verify data authenticity.

4.4.4 Software Quality Attributes

- Reliability: The system should be dependable and available 24/7, minimizing downtime and service interruptions.It should be resilient to faults, hardware failures, and network disruptions.

- Scalability: The system should scale both horizontally and vertically to accommodate increasing data loads and user demands without sacrificing performance.

- Maintainability: The system should be easy to maintain, update, and extend with new features, data sources, and machine learning models.Code should be well-documented and adhere to best practices in software development.

4.5 SYSTEM REQUIREMENTS

4.5.1 Database Requirements

1. Software Requirements(Platform Choice)

    - Operating System: A server operating system, such as Linux (e.g., Ubuntu, CentOS, Red Hat) or Windows Server, is commonly used to host the system.
    - Database Management System (DBMS): A relational or NoSQL DBMS to store and manage GIS data, operational data, and machine learning models. Options include PostgreSQL with PostGIS, MySQL, or MongoDB.
    - GIS Software:GIS software for spatial data management and analysis, like ArcGIS, QGIS, or open-source alternatives such as GeoServer and MapServer.
    - Machine Learning Frameworks: Frameworks for developing and deploying machine learning models, such as TensorFlow, PyTorch, scikit-learn, or specialized libraries for time series forecasting.
    - Security Tools: Security tools and software for data encryption, access control, and protection against cyber threats, including firewalls and intrusion detection systems.

2. Hardware Requirements

- Server Hardware: High-performance servers with multicore processors and ample RAM to handle data processing, machine learning computations, and web server operations.
- Storage: Large storage capacity, including fast SSDs for storing GIS data, operational data, and machine learning models. The exact storage requirements will depend on the volume of data.
- GPU (Graphics Processing Unit): If deep learning models are used, consider GPUs for accelerating machine learning computations. GPUs are essential for training large neural networks efficiently.
- Networking: High-speed network infrastructure to support data transfer, especially when dealing with real-time data streams and remote devices.
- Redundancy and Failover: Redundant hardware components and failover mechanisms to ensure system availability and reliability.

4.6 ANALYSIS MODELS: SDLC MODEL TO BE APPLIED

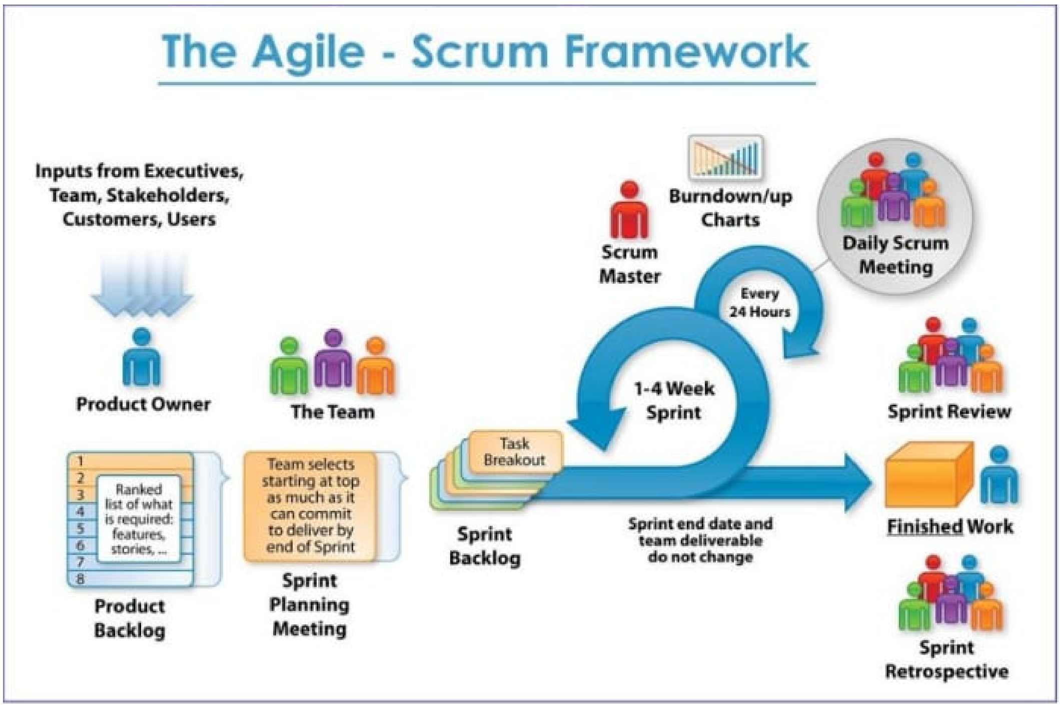

Figure 4.1: Agile Software Development Model

In our project we are going to use Agile software model because

- In a project like the development of a blockchain-based pharmaceutical supply chain track-and-trace system, where regulations, industry standards, and customer needs may evolve, Agile's adaptability is highly advantageous. You can accommodate changes in requirements and pivot quickly to address emerging challenges.

- Agile promotes continuous engagement with stakeholders, which is crucial in the healthcare supply chain. Regular feedback from manufacturers, distributors, healthcare providers, and regulatory bodies ensures that the system aligns with their evolving needs.

- Agile's incremental approach allows you to deliver valuable features and functionality sooner. This can be particularly valuable when addressing urgent supply chain challenges or regulatory changes.

- By conducting frequent testing and integration, Agile helps to identify and address issues early in the development process, reducing the risk of significant problems emerging later.

- Agile provides transparency into the development process through techniques like daily stand-up meetings and visible project backlogs. This transparency can instill confidence in stakeholders and help with project tracking.

- Agile enables you to focus on developing and delivering the most critical features first. This ensures that the core functionality of supply chain track-and trace system is available as soon as possible.

- Agile encourages a culture of continuous improvement. Regular retrospectives allow your team to assess what's working well and what could be enhanced, promoting ongoing refinement of the project.

4.7 SYSTEM IMPLEMENTATION PLAN:

Software implementation plan is a detailed roadmap to plan, manage, and control a software release through various stages. Software implementation project plan aims to improve the quality, speed, and efficiency of software delivery. This is an essential issue allowing your team to get the right information at the right time and thus promote a successful development. A well-structured system implementation plan is essential for ensuring a successful transition from development to production, minimizing disruptions, and maintaining the integrity and reliability of the software system. It serves as a roadmap in the software development process.

Following Table 4.1 outlines the System Implementation Plan according to Project Calendar:

1. Project Title Finalization: This phase involves finalizing the project's scope and defining its overall purpose. It ensures that the team is aligned on the project's direction and objectives.

2. Literature Survey: This phase focuses on conducting a thorough literature review to gather relevant knowledge and insights from previous research. This information will be crucial for designing and developing the system.

3. Requirement Analysis: This phase involves identifying and documenting the specific requirements for the project. This includes understanding the needs of stakeholders, defining functional and non-functional requirements, and prioritizing them.

4. System Architecture and UML Diagrams: This phase focuses on designing the system architecture and creating UML diagrams (Unified Modeling Language) to represent the system's components, their relationships, and their interactions. This will provide a clear blueprint for the system's implementation.

5. Project Stage-1 Report: This phase involves preparing the first project report, summarizing the work completed in the first four phases. It serves as an update for stakeholders and helps maintain project visibility.

6. Designing Smart Contracts: This phase involves designing and developing smart contracts for the project. Smart contracts are self-executing contracts that enable secure and transparent transactions on a blockchain. They form the core of many decentralized applications.

7. Implementation: This phase focuses on implementing the project's design and coding the system. It involves translating the system architecture and UML diagrams into actual software code.

8. Validation and Testing: This phase involves conducting rigorous testing to ensure the system's functionality, performance, and adherence to requirements. It includes unit testing, integration testing, and system testing.

9. Project Stage-2 Report: This phase involves preparing the final project report, summarizing the work completed in all phases. It provides a comprehensive overview of the project's development and outcomes.

Sr. No.	Activity	Work to be Accomplished	Duration
1	Project title Finalization	Deciding the project's Domain, Subdomain and Title	3 Weeks
2	Literature survey	To study research papers relevant to project title	3 Weeks
3	Requirement Analysis	Gathering and analyzing all the necessary requirements	4 Weeks
4	System Architecture and UML Diagrams	All UML diagrams to be developed, along with system architecture	5 Weeks
5	Project Stage-1 Report	Project Stage-1 report completion and submission	4 Weeks
6	Designing	Developing smart contracts for each stakeholder	5 Weeks
7	Implementation Web Application	Implementation and coding of the project	6 Weeks
8	Validation and Testing	To perform various tests on the system	4 Weeks
9	Project Stage-2 Report	Project Stage-2 report completion and submission	4 Weeks

Table 4.1: Comparative Analysis

CHAPTER 5

SYSTEM DESIGN

5.1 SYSTEM ARCHITECTURE

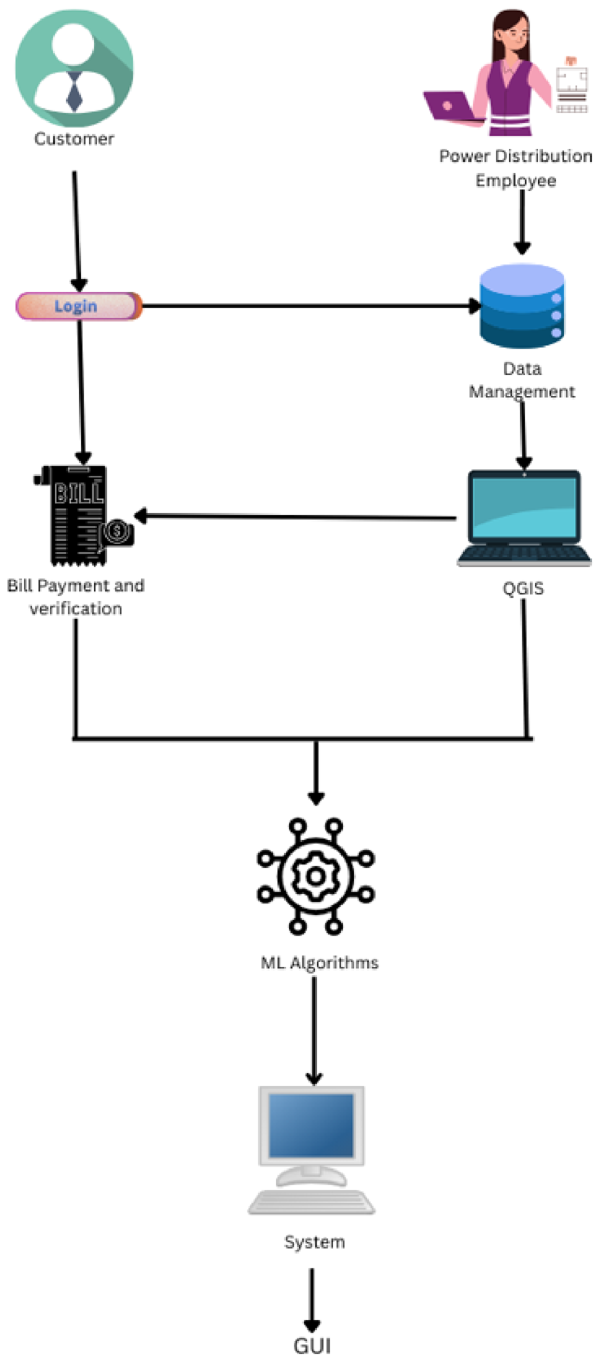

Figure 5.1: System Architecture

A System Architecture Diagram provides a high-level overview of the structure and organization of a system. It illustrates the key components, their relationships, and how they interact to achieve the overall functionality of the system. This type of diagram is a fundamental tool in software engineering and systems design, aiding in the communication of the system's architecture among stakeholders.

The system architecture for the power distribution involves a structured arrangement of components to enable efficient data flow and decision-making. At its core, the architecture comprises processes for QGIS mapping, route optimization, load forecasting, and a graphical user interface (GUI) for user interaction. The QGIS operation process takes GIS spatial data and system configuration data as inputs, generating QGIS mapping data as output. This mapping data serves as a crucial input for both route optimization and load forecasting processes.

The route optimization process utilizes dynamic programming models, incorporating QGIS mapping data and information on new power line construction to produce optimized route data. Simultaneously, the load forecasting process employs machine learning algorithms, leveraging QGIS mapping data and historical usage data to generate load forecast data. These outputs from route optimization and load forecasting are then visualized on the GUI, which integrates data such as load forecast information, optimized route details, GIS mapping data, and system configuration data.

The GUI acts as a user-friendly interface, allowing stakeholders to interact with the system, make informed decisions, and gain insights into the optimized power distribution network. The data store, housing GIS spatial data and system configuration data, serves as a centralized repository accessed by the QGIS operation, route optimization, and load forecasting processes. This system architecture ensures a cohesive and collaborative workflow, harnessing GIS technology, machine learning, and user interface design to enhance the efficiency, reliability, and sustainability of the power distribution network.

5.2 DATA FLOW DIAGRAMS

5.2.1 Data Flow Diagram Level 0

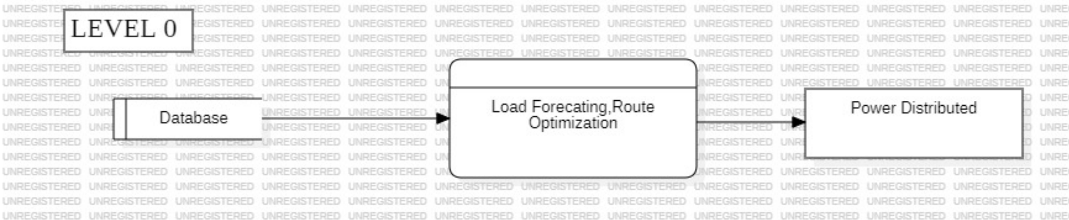

Figure 5.2: DFD Level 0

- Data Store: Data: Historical usage data, GIS spatial data, system configuration data Accessed by: Load Forecasting, Route Optimization, and Power Distribution processes.
- Load Forecasting Process: Input Flow: Historical usage data is fed into the Load Forecasting process. GIS spatial data may be used to enhance the accuracy of the load forecasting model. Output Flow: Load forecast data is generated and passed to the Power Distribution process.
- Route Optimization Process: Input Flow: GIS spatial data is used for route optimization. Information on new power line construction is considered. Output Flow: Optimized route data is produced and sent to the Power Distribution process.
- Power Distribution Process: Input Flow: Load forecast data, optimized route data, GIS spatial data, and current system status are inputs. The system configuration data is accessed from the data store. Output Flow: Efficiently distributed power is sent to customers. Maintenance schedule updates are generated and stored in the data store.

5.2.2 Data Flow Diagram Level 1

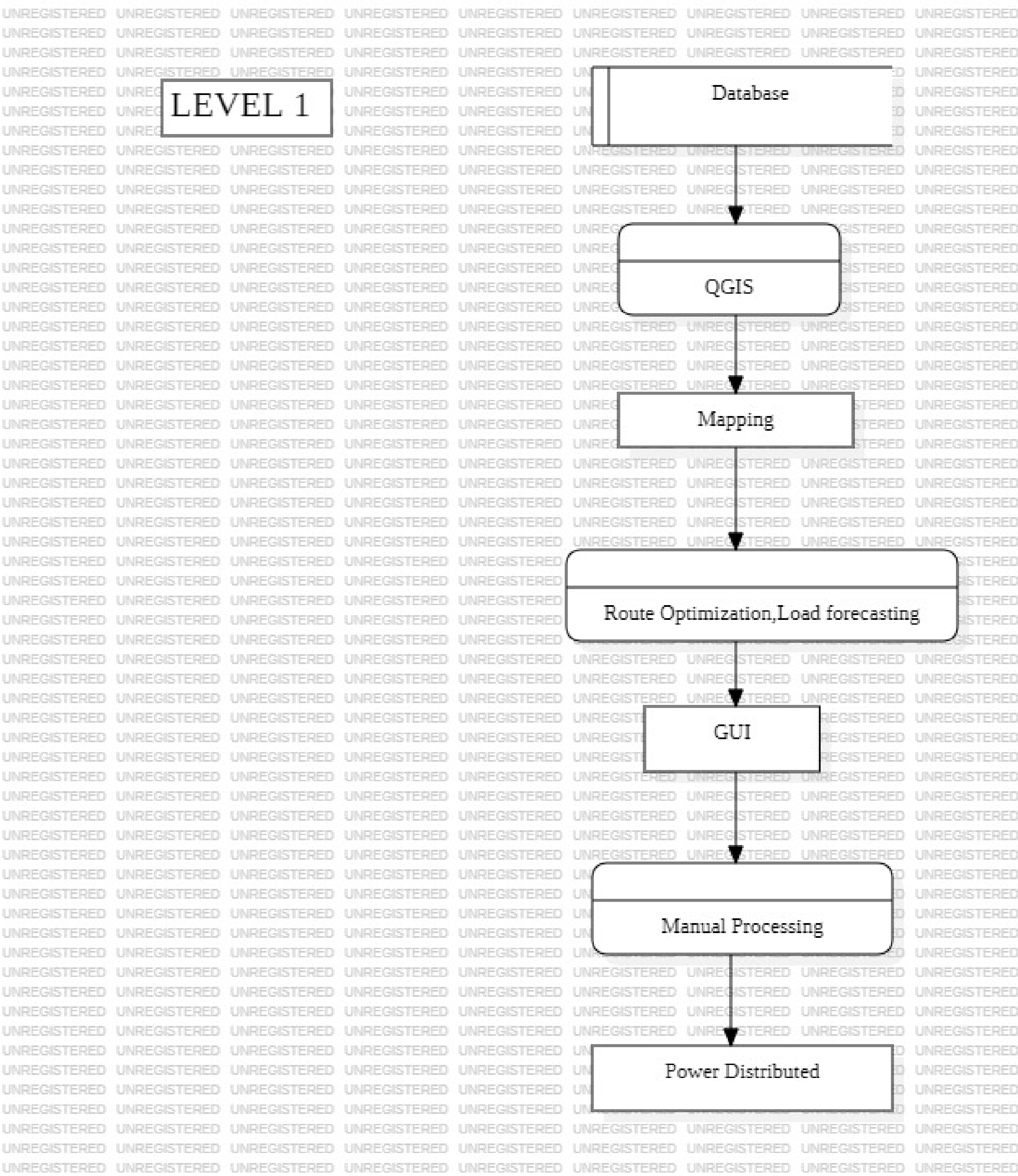

Figure 5.3: DFD Level 1

1. Data Store:
   - Data: GIS spatial data, system configuration data
   - Accessed by: QGIS Operation, Route Optimization, and Load Forecasting processes.
2. QGIS Operation Process:
   - Input: GIS spatial data, System configuration data
   - Process: QGIS processes spatial data and maps it, generating a mapped output.
   - Output: QGIS Mapping Data
3. Route Optimization :
   - Input: QGIS Mapping Data, Information on new power line construction
   - Process: Utilizes dynamic programming models to optimize the route for new power lines.
   - Output: Optimized Route Data
4. Load Forecasting Process:
   - Input: QGIS Mapping Data, Historical usage data
   - Process: Applies Machine Learning algorithms to forecast electricity demand based on historical and spatial data.
   - Output: Load Forecast Data
5. GUI (Graphical User Interface):
   - Input: Load Forecast Data, Optimized Route Data, GIS Mapping Data, System Configuration Data
   - Process: Displays the output to the user, allowing interaction and decision-making.
   - Output: Visualization on the GUI.

5.3 ENTITY RELATIONSHIP DIAGRAM

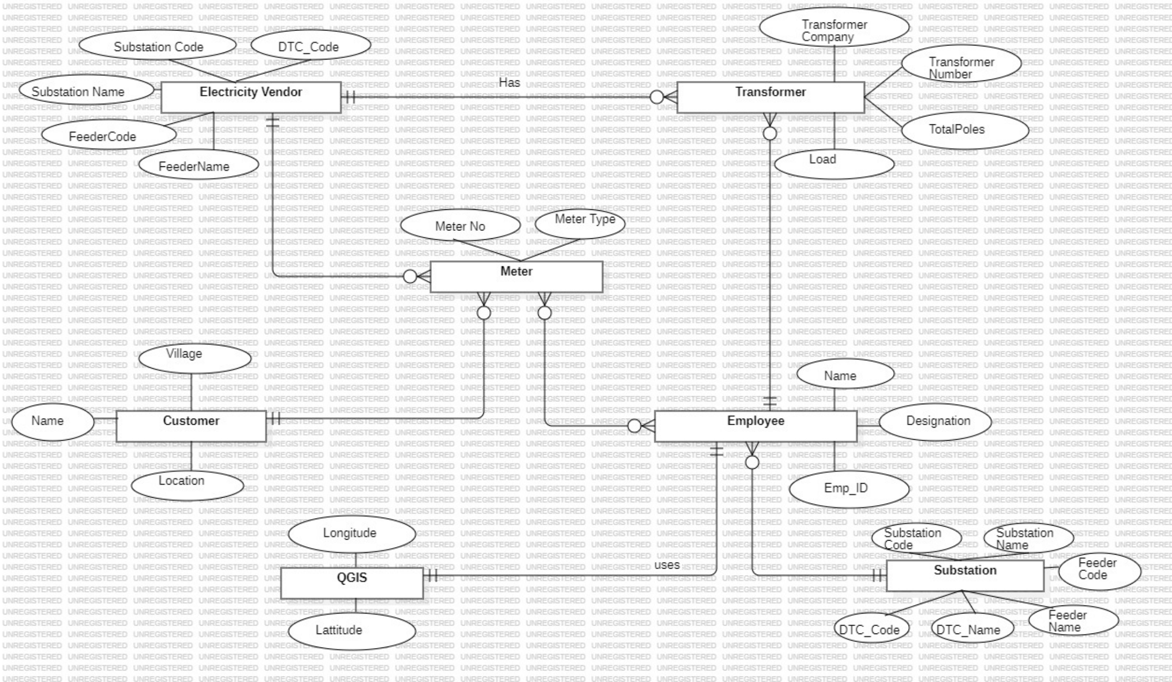

Figure 5.4: Entity Relationship Diagram

An Entity-Relationship (ER) diagram is a type of diagram in Unified Modeling Language (UML) used to visually represent the structure of a database in terms of entities and their relationships. The ER diagram is primarily used in database design and is a high-level conceptual data model.

The ER diagram you sent shows the following entities:

- Transformer: A transformer is a device that converts electrical energy from one voltage to another.
- Substation: A substation is a facility that contains electrical equipment to step up or step down voltage, switch circuits, and protect equipment.
- Feeder: A feeder is a cable or overhead line that distributes electricity from a substation to customers.
- Customer: A customer is a consumer of electricity.

The relationships between the entities are as follows:

- A substation has one or more transformers.
- A transformer is connected to one substation.
- A feeder is connected to one substation and one or more transformers.
- A transformer is connected to one feeder and one or more customers.
- A customer is connected to one transformer.

5.4 UML DIAGRAMS

5.4.1 Use Case Diagram

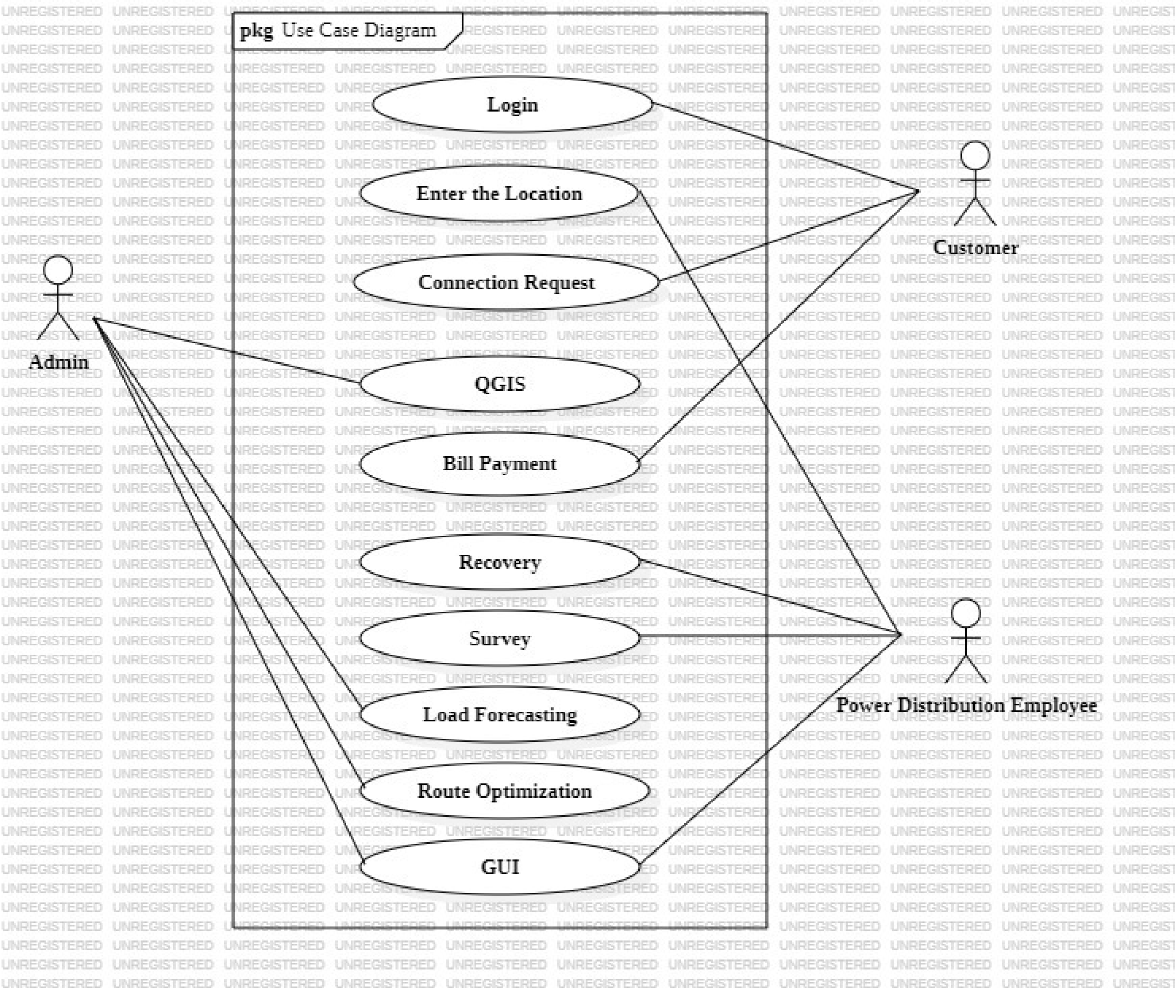

Figure 5.5: Use Case Diagram

A use case diagram is a type of behavioral diagram in the Unified Modeling Language (UML) that is used to capture the functional requirements of a system. It shows the interactions between different users or roles, known as actors, and the system's functionalities, represented as use cases. Use case diagrams provide a high-level overview of the system's behavior and can be used to identify and communicate the system's requirements.

The use case diagram shows the following use cases for a Power management system:

- Register: This use case allows users to register for an account with the system.
- Login: This use case allows users to log in to their accounts. Enter the Location: This use case allows users to enter their current location into the system.
- Exit the Location: This use case allows users to exit their current location from the system.
- Connection Request: This use case allows users to request a connection to another user in the system.
- Bill Payment: This use case allows users to pay their bills through the system.
- Recovery: This use case allows users to recover their accounts in case they forget their passwords.
- Survey: This use case allows users to participate in surveys.
- Load Forecasting: This use case allows users to forecast the load on the system.
- Route Optimization: This use case allows users to optimize their routes.

The use case diagram also shows the following actors:

- Customer: A customer is a user of the system who wants to track their location and use the system's other features.
- Admin: An admin is a user of the system who is responsible for configuring and managing the system.

The use case diagram shows the following relationships between the actors and use cases:

- Customer: A customer can register for an account, log in to their account, enter their current location, exit their current location, request a connection to another user, pay their bills, recover their account, participate in surveys, and forecast the load on the system.
- Admin: An admin can perform all of the same use cases as a customer, as well as optimize routes and manage the system.

5.4.2 Activity Diagram

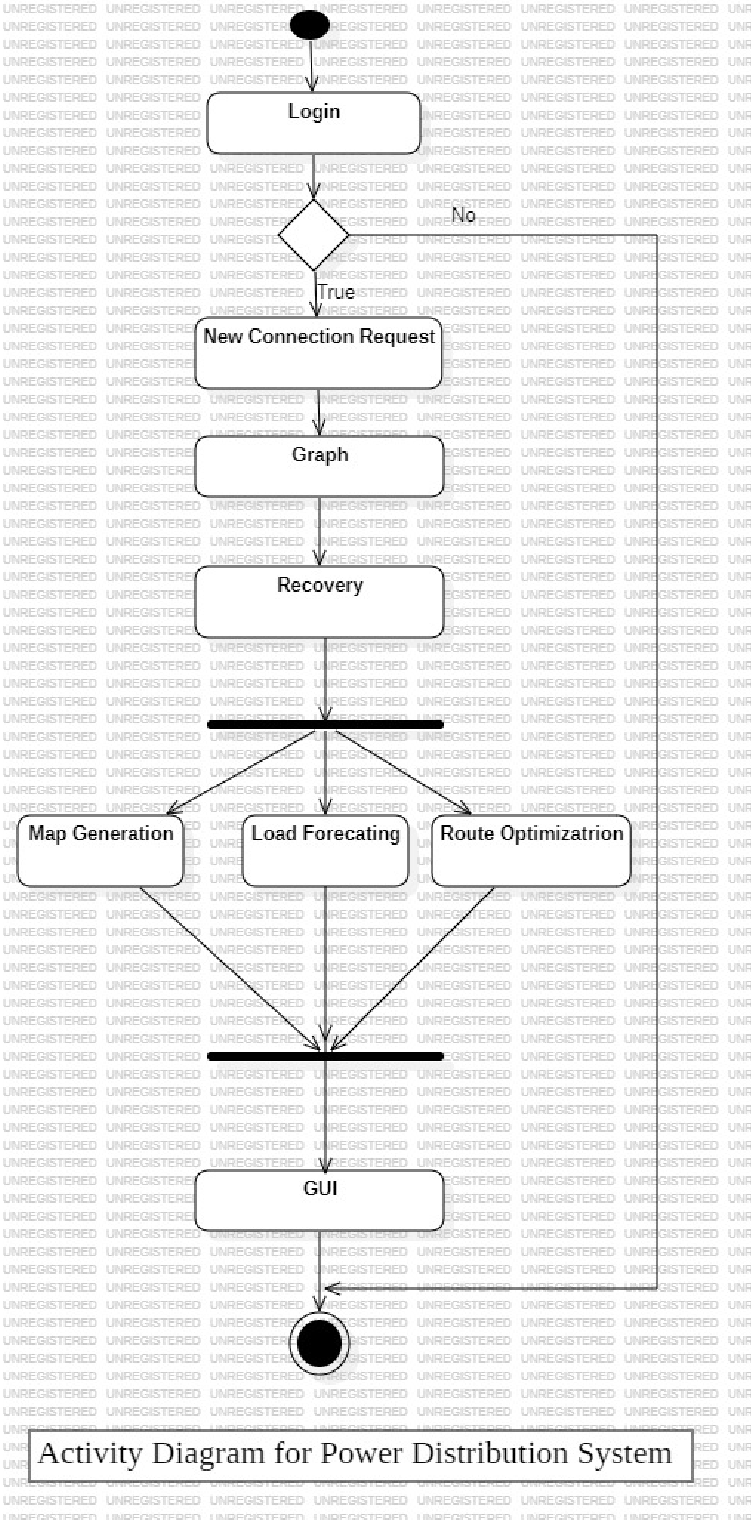

Figure 5.6: Activity Diagram

In Unified Modeling Language (UML), an activity diagram is a graphical representation of the dynamic aspects of a system, focusing on the flow of control and the sequence of activities. It is a behavioral diagram that illustrates the workflow of a system or a specific process within a system. Activity diagrams are particularly useful for modeling business processes, system workflows, and the interaction between different components or actors in a system.

1. The customer logs in to the power distribution system.
2. The customer enters their customer details.
3. The Customer's location graph is generated using QGIS.
4. Existing bill Payment recovery is done by Power Distribution Employee.
5. Map Generation, Load Forecasting, Route Optimization is done .
6. The GUI is shown to the Power Distribution employee.
7. The system installs the new connection at the customer's location.
8. The system completes the connection and sends a confirmation to the customer.

5.4.3 Class Diagram

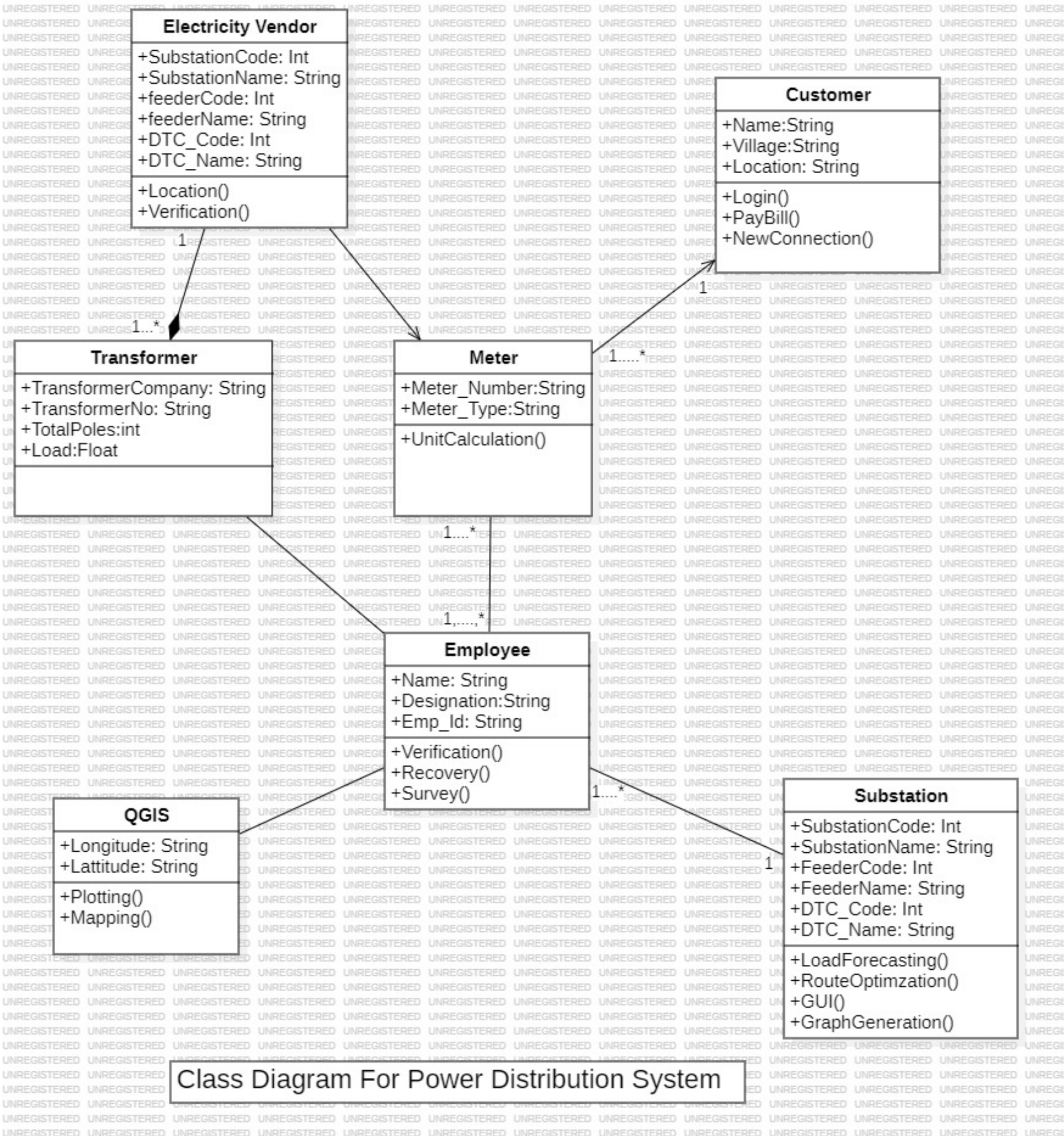

Figure 5.7: Class Diagram

A class diagram in Unified Modeling Language (UML) is a type of static structure diagram that provides a visual representation of the classes and their relationships within a system. It is a fundamental tool for modeling the structure of object-oriented systems and is widely used in software engineering for visualizing and designing complex systems.

1. Electricity Vendor Class:

   - Represents the entity responsible for supplying electricity.

- Attributes may include vendor details, contract information, and service-related data.
- Methods could involve managing electricity supply contracts, handling billing, and providing data to optimize power distribution.

2. Customer Class:
   - Represents the end-users or consumers of electricity.
   - Attributes may include customer details, usage history, and billing information.
   - Methods could involve managing customer accounts, processing payments, and accessing usage information.
3. Meter Class:
   - Represents the electricity meter used to measure consumption at a customer's location.
   - Attributes may include meter ID, location, and usage data.
   - Methods could involve recording consumption, providing usage data to the system, and handling meter-related maintenance.
4. Employee Class:
   - Represents individuals involved in the management and maintenance of the power distribution system.
   - Attributes including employee details, roles, and responsibilities.
   - Methods involve tasks related to system maintenance, fault resolution, and data management.
5. Transformer Class:
   - Represents the transformers used in the power distribution network.
   - Attributes may include transformer ID, location, and capacity.

- Methods could involve monitoring transformer status, handling maintenance, and providing data for load forecasting.

6. QGIS Class:

   - Represents the integration of Geographic Information System (QGIS) technology.
   - Attributes may include GIS-related configurations, data sources, and system parameters.
   - Methods could involve handling GIS data, providing spatial information for route optimization, and supporting decision-making.

7. Substation Class:

   - Represents the substations in the power distribution system.
   - Attributes include substation ID, location, and capacity.
   - Methods could involve managing substation operations, handling load balancing, and interacting with the GIS for planning and optimization.

5.4.4 Object Diagram

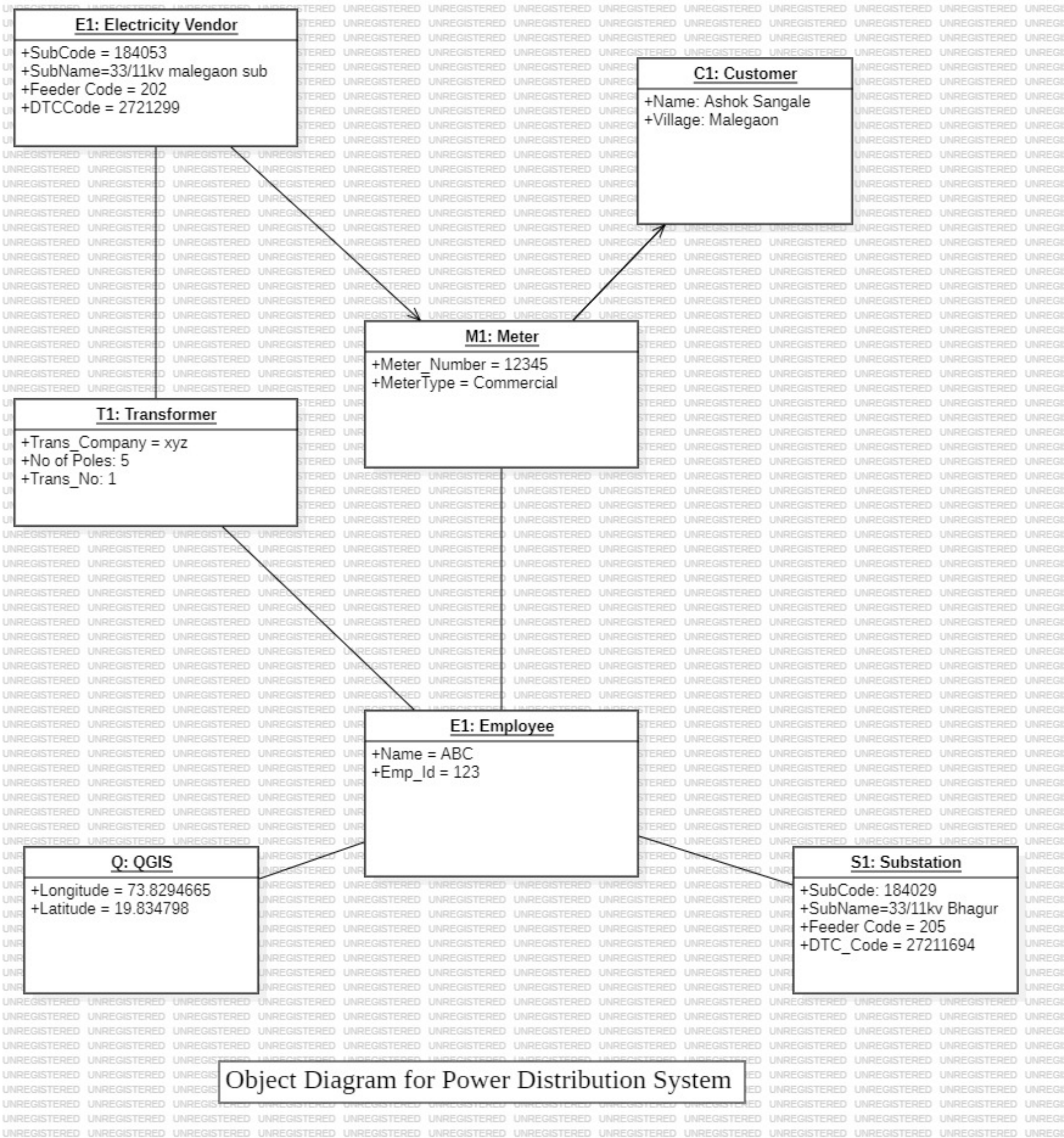

Figure 5.8: Object Diagram

In Unified Modeling Language (UML), an object diagram is a structural diagram that represents a set of objects and their relationships at a specific point in time. It provides a snapshot of the instances of classes in a system and the relationships between these instances. The object diagram shows a network of objects connected to each other. The objects represent different components of a power distribution system, such as customers, substations, feeders, and distribution transformers. The relationships between the objects represent the physical and logical connections between the different components of the system. For example, the object diagram shows that a customer is connected to a substation, which is connected to a feeder, which is connected to a distribution transformer. The object diagram also shows that a substation can be connected to multiple feeders, and a feeder can be connected to multiple distribution transformers.

The object diagram also shows the following information about each object:

1. Substation: The substation object has a substation code and name. It is also connected to one or more feeders.

2. Feeder: The feeder object has a feeder code and name. It is connected to a substation and one or more distribution transformers.

3. Distribution Transformer: The distribution transformer object has a distribution transformer code and name. It is connected to a feeder and one or more customers.

4. Customer: The customer object has a customer code and name. It is connected to a distribution transformer. The object diagram provides a high-level overview of the topology of the power distribution system. It shows the different components of the system and how they are connected to each other. The object diagram does not show the detailed logic of how the system works, but it does provide a good starting point for understanding the system's topology.

A power distribution system is a network of electrical components that delivers electricity from power plants to consumers. The object diagram shows a simplified view of a power distribution system, with the following components:

1. Substation: A substation is a facility that steps down the voltage of electricity from high voltage transmission lines to medium voltage distribution lines. Substations also contain other equipment, such as transformers, circuit breakers, and switches, to protect and control the flow of electricity.

2. Feeder: A feeder is a cable or overhead line that distributes electricity from a substation to customers.

3. Distribution Transformer: A distribution transformer steps down the voltage of electricity from medium voltage distribution lines to low voltage service lines. Service lines deliver electricity to homes and businesses.

4. Customer: A customer is a consumer of electricity. The object diagram shows that a customer is connected to a distribution transformer, which is connected to a feeder, which is connected to a substation. Substations can be connected to multiple feeders, and feeders can be connected to multiple distribution transformers.

5.4.5 Sequence Diagram

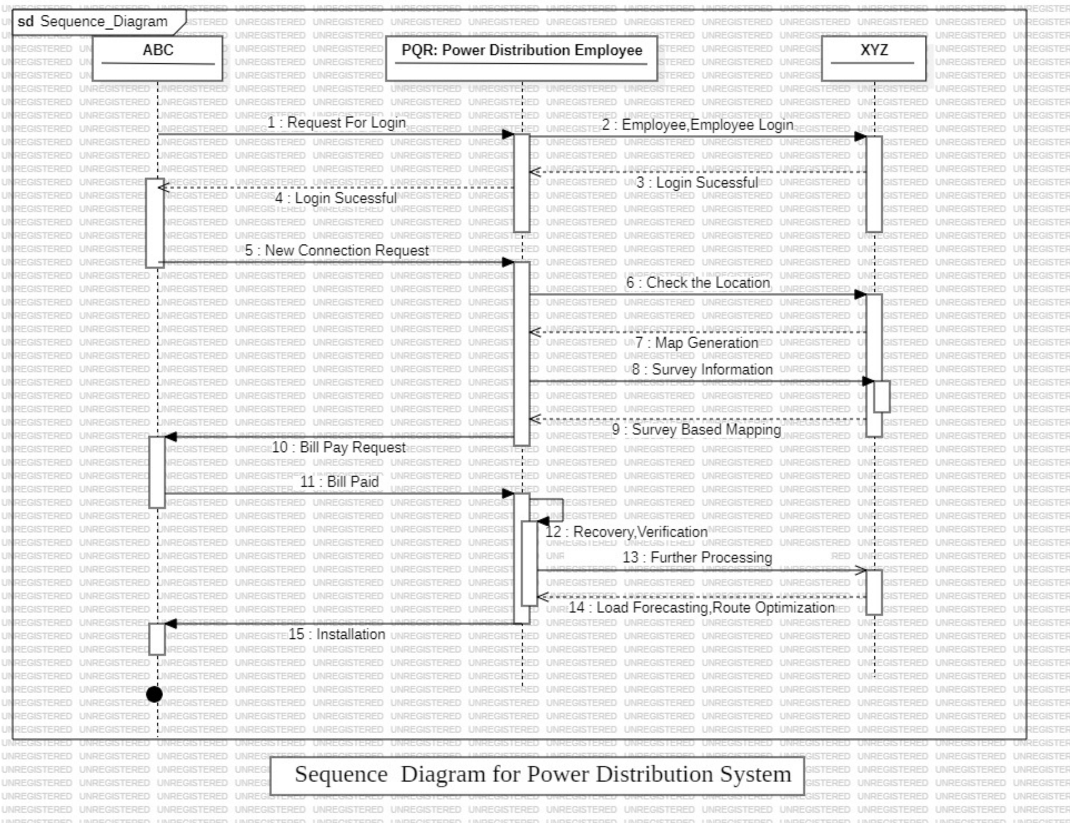

Figure 5.9: Sequence Diagram

A sequence diagram in Unified Modeling Language (UML) is a type of interaction diagram that visualizes the interactions and messages exchanged between different objects or components within a system over a specific period of time. It represents the dynamic aspects of a system, emphasizing the flow of control and the sequence of events that occur during the execution of a particular scenario or use case.

The sequence diagram shows the interaction between a customer and a power distribution system to.

1. The customer sends a new connection request to the power distribution system.
2. The power distribution system receives the new connection request and verifies that the customer has provided all of the required information.
3. The power distribution system checks to see if there is a connection available at the customer's location.
4. If there is a connection available, the power distribution system reserves the connection for the customer.
5. The power distribution system sends a confirmation to the customer that the new connection request has been received and processed.
6. The power distribution system does the QGIS Mapping and the ML algorithms are applied for the New connection.
7. The power distribution system schedules a time to install the new connection.
8. The power distribution system installs the new connection at the customer's location.
9. The customer begins using the new connection.

CHAPTER 6

OTHER SPECIFICATION

6.1 ADVANTAGES

1. Enhanced Data Visualization and Analysis: GIS allows for the spatial representation of power distribution infrastructure, making it easier to visualize and analyze the network's components and performance.Machine learning can process vast amounts of data to identify patterns, anomalies, and trends that might not be apparent through traditional methods, enabling better decision-making.[1]

2. Improved Asset Management: GIS helps in tracking and managing assets (transformers, substations, lines, etc.) more efficiently. It allows for better asset condition monitoring, which can lead to reduced maintenance costs and improved reliability.Machine learning can predict equipment failures, helping in proactive maintenance and replacement planning, thereby reducing downtime.

3. Optimized Network Operation: Machine learning algorithms can optimize power distribution by intelligently routing electricity through various lines and transformers to minimize losses and improve energy efficiency.It supports load forecasting, which aids in efficient load shedding during peak demand periods.[7]

4. Real-time Monitoring and Control: GIS, coupled with machine learning, provides real-time monitoring and control of the power distribution system, allowing rapid response to outages, faults, and other operational issues.Automated alerts and predictive maintenance improve system reliability and uptime.

5. Accurate Predictive Analytics: Machine learning models can predict power demand with high accuracy, considering various factors such as weather conditions, historical data, and customer behavior.These predictions help in efficient resource allocation and load management.[1]

6. Fault Detection and Identification: Machine learning algorithms can detect and identify faults in the power distribution system, including line faults and

equipment failures.Rapid fault detection and isolation lead to quicker restoration of service and reduced outage times.[8]

6.2 LIMITATIONS

1. Dependency on Data Sources: The system's effectiveness heavily relies on the quality and availability of data. Data source failures or inaccuracies can impact system performance.[1]

2. Human Expertise: The system requires personnel with expertise in machine learning, GIS, data management, and power distribution, making it necessary to attract and retain skilled professionals.[6]

3. Initial Data Collection and Labeling: Gathering historical data and labeling it for training machine learning models can be labor-intensive and time-consuming, especially for identifying fault conditions.[1]

4. Complexity of Machine Learning Models: Developing, training, and maintaining machine learning models can be complex and resource-intensive. Model selection and tuning are critical, and errors in model development can lead to inaccurate predictions.[4]

5. Data Quality and Availability: Data quality is crucial for machine learning models to perform accurately.

6.3 APPLICATIONS

1. A GIS-based power distribution system integrated with machine learning holds immense potential for a wide range of applications, transforming the way power utilities manage and deliver electricity. One key application is real-time monitoring and analysis of the power distribution network. By harnessing spatial data and machine learning algorithms, this system can provide utilities with a comprehensive view of the network's status, pinpointing areas of congestion, outages, or potential issues. This real-time monitoring enhances situ-

ational awareness, enabling swift responses to faults and optimizing resource allocation for more reliable power distribution.

2. Predictive analytics is another powerful application. Machine learning models can analyze historical data, weather patterns, and consumer behaviors to forecast power demand accurately. Utilities can use these predictions to proactively optimize the distribution of power, minimize losses, and prevent overloads during peak usage, ensuring a smoother and more efficient energy supply.

3. Fault detection and identification represent another critical application. Machine learning algorithms can automatically detect and locate faults in the power distribution network. This capability reduces downtime and enables quicker responses from maintenance teams, leading to a more resilient and reliable network.

4. Optimizing power distribution is a fundamental application as well. Machine learning models can dynamically route power through the network to ensure efficient and balanced distribution. During peak demand, the system can automatically shed loads or balance power to maintain stability, ultimately improving the overall performance of the network.

CHAPTER 7

SUMMARY AND CONCLUSION

The project titled ”GIS based Power Distribution System using ML” is aimed at revolutionizing the power distribution sector by integrating Geographic Information System (GIS) technology with Machine Learning (ML) to optimize and improve the efficiency and reliability of power distribution networks. The project's objectives encompass managing power distribution remotely, developing ML models for accurate power demand prediction, identifying inefficiencies and cost-saving opportunities, and optimizing power distribution systems.The project ”GIS based Power Distribution System using ML” offers a promising solution to the challenges faced by traditional power distribution networks. By leveraging GIS and ML technologies, it aims to optimize power distribution and enhance the reliability and efficiency of electrical grids. The integration of GIS data with machine learning models enables accurate power demand predictions, identification of inefficiencies, and effective load management.As power distribution is a critical component of modern society, the success of this project can have a substantial impact. A more reliable and efficient power distribution network can lead to reduced energy losses, cost savings, and enhanced grid resilience during emergencies. By tailoring services and optimizing infrastructure, the project contributes to the advancement of the power distribution sector, ultimately benefiting both utilities and consumers. In conclusion, the project represents a significant step toward building a more sustainable and intelligent power infrastructure, making power distribution more efficient and responsive to the needs of a growing population and evolving energy demands.

CHAPTER 8

REFERENCES

[1] Chandan SINGH Praveen Kumar Rai. *"GIS in electrical asset mapping: The case of bhadohi, India.", uropean Journal of Geography Volume 7, Number 4:19 - 33, December 2016 ©Association of European Geographers European Journal of Geography-ISSN 1792-1341*. 2016.

[2] WenBin Wang. "power distribution equipment and defect identification technology based on deep learning", iceect. 2021.

[3] Uday Patil Aanand Kumbhar, Pravin Dhavle. "a comprehensive review: Machine learning and its application in integrated power system", school of electrical engineering, sanjay ghodawat university, kolhapur, india. 2021.

[4] Andrea Morotti Marco Merlo Alessandro Bosisio, Matteo Moncecchi. "machine learning and gis approach for electrical load assessment to increase distribution networks resilience", energies 2021, 14, 4133. 2021.

[5] Sai Pavan Kumar Venkataramana Veeramsetty and Surender Reddy Salkuti. " platform-independent web application for short-term electric power load forecasting on 33/11 kv substation using regression tree",. 2022.

[6] A. Roosta T. Niknam N. Rezaee, M Nayeripour. "role of gis in distribution power systems", world academy of science, engineering and technology international journal of electrical and computer engineering vol:3. no: 12. 2009.

[7] Atsushi Shiota Yaser Qudaih Yasunori Mitani Yusuke Kakumoto, Yuki Koyamatsu. "application geographic information sysytem to power distribution system analysis", peer-review under responsibility of the organizing committee of cpese 2016. 2016.

[8] V Phani Kuma P Sree Gayathri. "gis the future of utility management", october 17-18, 3rd international conference on emerging technologies in computer science engineering (icetcse 2016) v. r. siddhartha engineering college, vijayawada, india, october 17-18. 2016.

ANNEXURE A

PROBLEM STATEMENT FEASIBILITY

1. Technical Feasibility: The integration of QGIS and machine learning tools is technically feasible, given the availability of open-source geospatial software and the capability of machine learning algorithms to optimize power distribution.

2. Operational Feasibility: The project is operationally feasible as it streamlines power distribution processes, reducing manual efforts, and enhancing system efficiency.

3. Economic Feasibility: The cost savings from reduced downtime and improved resource allocation make this project economically viable for power distribution companies.

4. Legal Feasibility: Adhering to data privacy and regulatory compliance in the utilization of customer data is essential, but existing legal frameworks support the responsible use of such data.

5. Schedule Feasibility: The project can be completed within a reasonable timeframe, provided that resources and expertise are readily available, making it schedule-feasible for timely implementation.

ANNEXURE B

DETAILS OF THE PAPERS REFERRED

1. N. Rezaee, M Nayeripour, A. Roosta, T. Niknam, ”Role of GIS in Distribution Power Systems”, World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering Vol:3. No: 12, 2009.

   Summary: A new methodology of automated optimal routing for new power lines by dynamic programming (DP) and economic corridors which include suitable economic routing alternatives for new power lines and optimizing substation’s location and capacity.

2. Praveen Kumar RAI, Chandan SINGH, ”GIS IN ELECTRICAL ASSET MAPPING: THE CASE OF BHADOHI, INDIA.”, European Journal of Geography Volume 7, Number 4:19 - 33, December 2016 ©Association of European Geographers European Journal of Geography-ISSN 1792-1341, December 2016

   Summary: Finding shortest optimum path for power distribution line.It is also very helpful to improve the customer complaints, long-standing faults brought to a minimum, to stop power theft (in local language it is called Katia practice), and to provide better facilities to the each consumer

3. Aanand Kumbhar, Pravin G. Dhawale, Shobha Kumbhar, Uday Patil, Pravin Magdum ”A comprehensive review: Machine learning and its application in integrated power system ” School of Electrical Engineering, Sanjay Ghodawat University, Kolhapur, India , 31 may 2021.

   Summary: This paper given review of application of ML in integrated power system.This paper also gives detailed information about the types of machine learning techniques. It summarize the uses of ML to blame finding in the force framework, which is the essential advance of the force framework security and control.

4. Venkataramana Veeramsetty , Sai Pavan Kumar and Surender Reddy Salkuti, ”Platform-Independent Web Application for Short-Term Electric Power Load Forecasting on 33/11 kV Substation Using Regression Tree ”,Computers 2022.

   Summary: Developed to forecast the load with good accuracy irrespective of the weather conditions and types of the day by incorporating features such as season, temperature, humidity, and day-status.

5. P Sree Gayathri, V Phani Kumar ”GIS the future of Utility Management”, October 17-18, 3rd International Conference on Emerging Technologies in Computer Science Engineering (ICETCSE 2016) V. R. Siddhartha Engineering College, Vijayawada, India, October 17-18, 2016.
   Summary: The GIS applications extend over various assorted fields, but it is computer science which holds the key to understand and make advancements of the underlying spatial database and programming for custom applications.The applications finally developed will facilitate online query with geographical display, by showing particular assets and their attributes related to utilities.

6. Alessandro Bosisio, Matteo Moncecchi, Andrea Morotti and Marco Merlo, ”Machine Learning and GIS Approach for Electrical Load Assessment to Increase Distribution Networks Resilience”, Energies 2021, 14, 4133.
   Summary: The study explores the possibility of predicting the power profile of MV/LV substations using machine learning approaches and GIS information.

7. WenBin Wang, Jie Wei, Yu Zhu, ShiYang Zhou, ”Power distribution equipment and defect identification technology based on deep learning”, ICEECT 2021.
   Summary: In this paper, a method of distribution network equipment and fault identification based on edge calculation is proposed. The experimental results show that the accuracy of the proposed algorithm reaches 87.9 percentage and the recall rate reaches 83.3 percentage.

ANNEXURE C

PLAGIARISM REPORT FOR THIS REPORT

Plagiarism checking is the process of detecting whether written work has been copied from another source without proper attribution. It is an important tool for ensuring academic integrity and preventing copyright infringement. Plagiarism checking is the process of examining a piece of writing or content to determine if it contains elements that have been copied or closely imitated from other sources without proper attribution. The goal of plagiarism checking is to maintain academic and professional integrity by ensuring that individuals present original work and give credit to the original creators of ideas, words, or concepts.The average Plagiarism percentage of our report is 7.4

C.1 CHAPTER 1

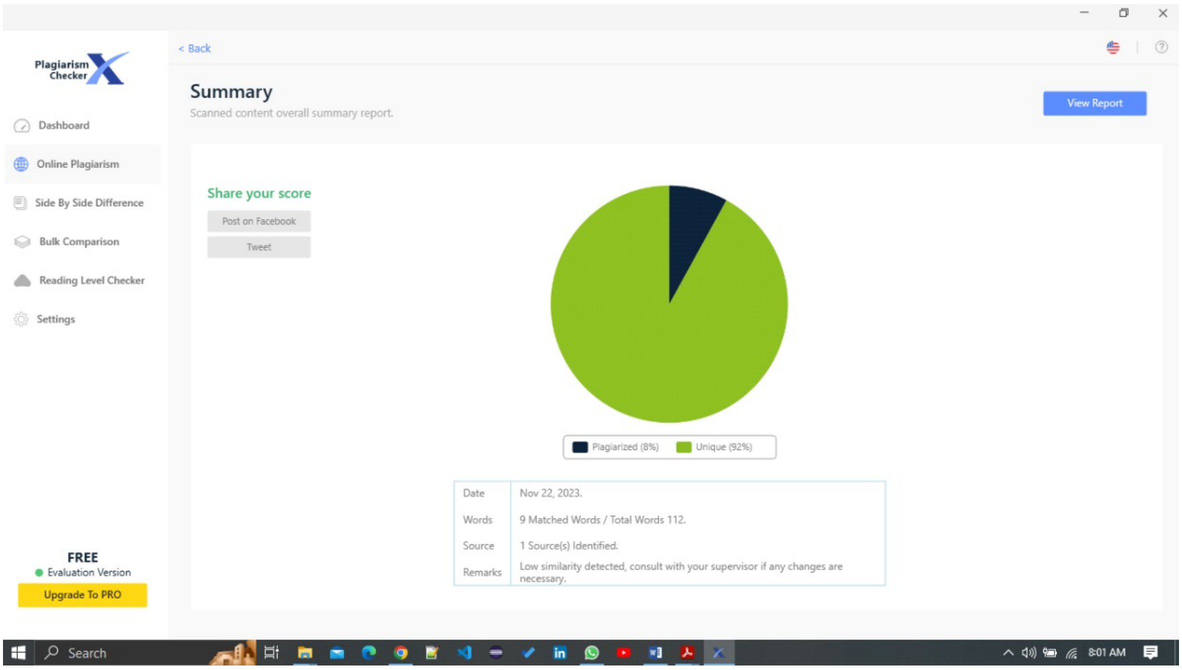

Figure C.1: Chapter 1 Plagiarism

C.2 CHAPTER 3

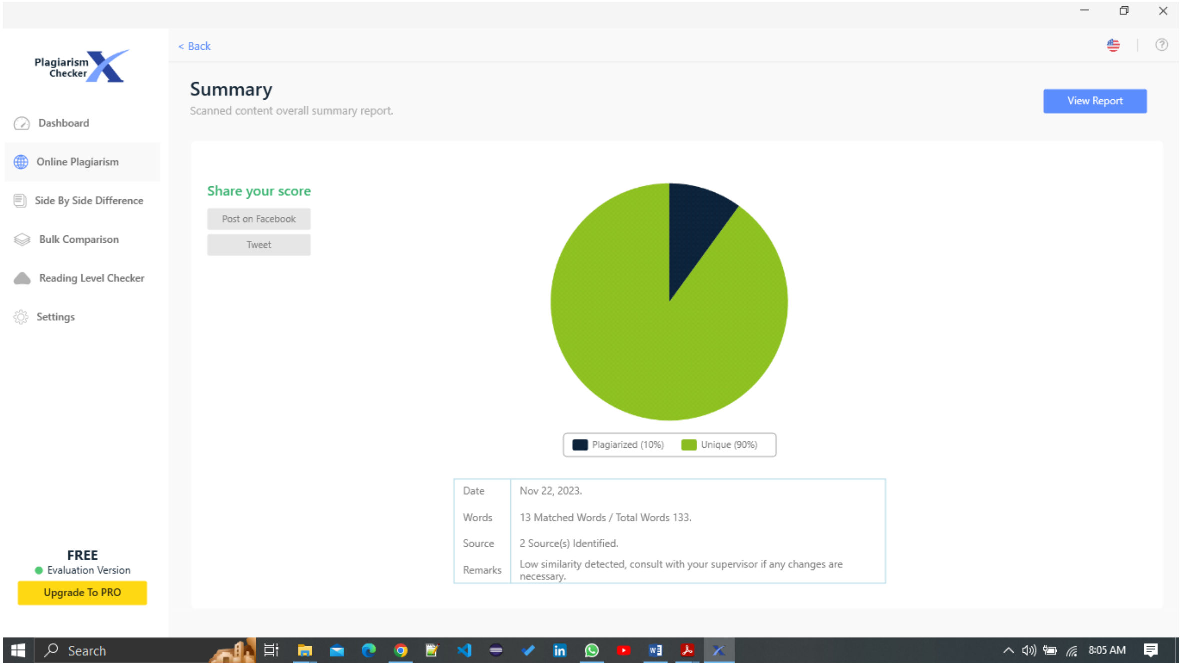

Figure C.2: Chapter 3 Plagiarism

C.3 CHAPTER 4

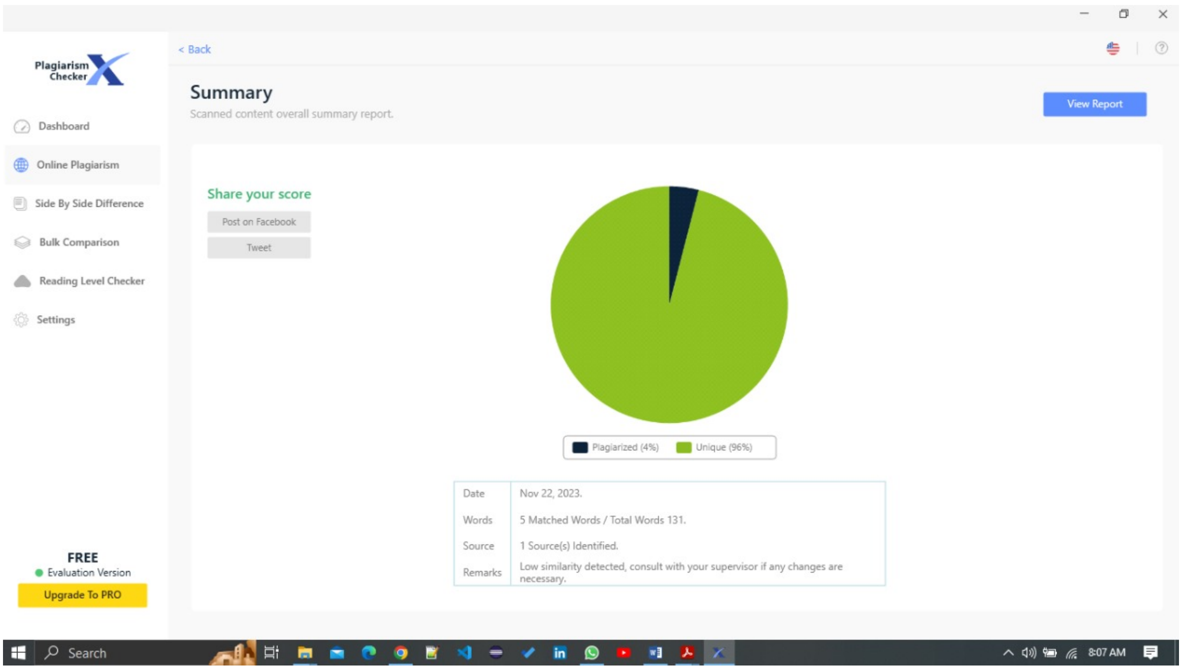

Figure C.3: Chapter 4 Plagiarism