VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
An Autonomous Institute Affiliated to University of Mumbai
Department of Computer Engineering

Project Report on

Revenue and Expense Analyzer

In partial fulfillment of the Fourth Year, Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai Academic Year 2023-24

Submitted by
Yash Kewlani (D17 - B , Roll no - 34)
Neeraj Chawla(D17 - B , Roll no - 12)
Atharva Mahalle(D17 - B , Roll no - 40)
Mithil Wasrani(D17 - B , Roll no - 72)

Project Mentor
Mrs. Priti Joshi

(2023-24)

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
An Autonomous Institute Affiliated to University of Mumbai
Department of Computer Engineering

Certificate

This is to certify that **Yash Kewlani (D17B, 34), Neeraj Chawla (D17B, 12), Atharva Mahalle (D17B, 40), Mithil Wasrani (D17B, 72)** of Fourth Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on "**Revenue and Expense Analyzer**" as a part of their coursework of PROJECT-II for Semester-VIII under the guidance of their mentor **Mrs. Priti Joshi** in the year 2023-24 .

This project report entitled "**Revenue and Expense Analyzer**" by **Yash Kewlani, Neeraj Chawla, Atharva Mahalle, Mithil Wasrani** is approved for the degree of **B.E. Computer Engineering.**

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7, PO8, PO9, PO10, PO11, PO12 PSO1, PSO2	

Date:

Project Guide:

Project Report Approval
For
B. E (Computer Engineering)

This project report entitled "**Revenue and Expense Analyzer**" by **Yash Kewlani, Neeraj Chawla, Atharva Mahalle, Mithil Wasrani** is approved for the degree of **B.E. Computer Engineering.**

Internal Examiner

External Examiner

Head of the Department

Principal

Date:

Place: Mumbai

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

--	--
Yash Kewlani(34)	Neeraj Chawla(12)
--	--
Atharva Mahalle(40)	Mithil Wasrani(72)

Date:

ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to **Mrs. Priti Joshi** (Project Guide) for her kind help and valuable advice during the development of project synopsis and for her guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr.(Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair ,** for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

Computer Engineering Department
COURSE OUTCOMES FOR B.E PROJECT

Learners will be to,

Course Outcome	Description of the Course Outcome
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solutions for the problem.
CO 4	Able to interpret the data and datasets to be utilized.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one's competency.
CO 11	Able to develop a professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

Index

List of Figures:

List of Tables:

Abstract

The "Revenue and Expense Analyzer" project addresses the pressing need for effective financial performance monitoring and management in the dynamic business and finance landscape. The intricacies of financial data, often buried within complex spreadsheets and reports, pose a significant challenge for professionals. This project offers a comprehensive solution by introducing an innovative dashboard adorned with essential Key Performance Indicators (KPIs) and critical financial insights, designed to provide clarity and understanding regardless of the user's technical background. The project leverages a diverse array of visualization tools, including area charts, line charts, bar charts, pie charts, and scatter plots, to transform complex financial data into easily comprehensible visuals, simplifying the decision-making process.

What truly distinguishes the "Revenue and Expense Analyzer" is its incorporation of advanced machine learning techniques for predictive analysis, providing capabilities that enable businesses to accurately forecast revenue trends. With this project, data-driven decision-making becomes more accessible and reliable, empowering organizations to make informed decisions and plan confidently for the future.

Keywords: Key Performance Indicator, Decision Support System, Revenue, Expenses, Profit, Regression

Chapter 1: Introduction

1.1.Introduction:

In today's dynamic business environment, the effective management of financial performance stands as a cornerstone of success. However, navigating through intricate spreadsheets and reports to decipher critical financial insights can often prove challenging for professionals. Moreover, the absence of user-friendly tools and accessible Key Performance Indicators (KPIs) exacerbates this complexity, hindering a clear understanding of a company's financial health. Addressing these hurdles head-on, we introduce the "Revenue and Expense Analyzer."

The Revenue and Expense Analyzer emerges as a comprehensive solution, empowering businesses with the necessary tools to seize control of their financial performance. In an era where data-driven decision-making reigns supreme, this innovative dashboard offers an engaging homepage adorned with essential KPIs and pivotal financial insights, irrespective of the user's technical background. At its core, this groundbreaking solution boasts a diverse array of visualization tools, ranging from versatile area charts to insightful line charts, informative bar charts, revealing pie charts, and illuminating scatter plots. These visualization tools not only simplify complex financial data but also streamline the decision-making process, ensuring clarity and precision in every financial endeavor.

1.2.Motivation:

The "Revenue and Expense Analyzer" project is driven by a commitment to simplifying the complexities of financial analysis in modern business. Recognizing the labyrinthine nature of financial data, spread across various platforms and formats, the project aims to provide organizations with a unified and holistic view of their financial health. By simplifying the interpretation of financial information and creating a user-friendly platform accessible to individuals with varying levels of expertise, the project seeks to democratize financial data within organizations. Furthermore, it aims to consolidate fragmented tools into a comprehensive Finance dashboard, offering real-time data visualization, scenario modeling, predictive capabilities, and accessible Key Performance Indicators (KPIs). This integrated approach streamlines decision-making processes, empowering organizations to make more informed, efficient, and strategic financial choices for long-term success.

1.3. Problem Definition:

In today's business landscape, the effective monitoring of financial performance remains a significant challenge. Financial data complexity, a lack of user-friendly tools, and the absence of a comprehensive Finance dashboard impede decision-making. Businesses struggle to access vital Key Performance Indicators (KPIs), interpret financial information, and make reliable revenue predictions. The "Revenue and Expense Analyzer" seeks to address these issues by offering a unified solution that simplifies financial monitoring, enhances data accessibility, provides clear visualization tools, and incorporates machine learning for accurate revenue forecasting. This comprehensive approach empowers businesses to make informed decisions and confidently plan for a prosperous financial future.

1.4. Existing Systems:

1. **Enterprise Resource Planning (ERP) Systems:** Widely adopted by businesses, ERP systems incorporate forecasting modules to predict revenue and sales by analyzing historical data, customer information, and market trends. However, the forecasting capabilities may vary, and they may not always utilize advanced machine learning techniques for prediction.

2. **Customer Relationship Management (CRM) Systems:** CRM platforms offer basic forecasting functionalities, analyzing historical sales data and pipeline opportunities to predict future revenue. However, these forecasts may lack sophistication and accuracy compared to advanced machine learning models.

3. **Business Intelligence (BI) Tools:** Tools like Tableau, Power BI, and Qlik View provide data visualization and analytics, offering insights into sales trends. However, revenue prediction often relies on simple statistical methods or manual analysis, which may not capture complex data relationships.

4. **Custom-built Forecasting Models:** Some organizations develop custom forecasting models tailored to their needs, incorporating statistical techniques or machine learning algorithms like linear regression or decision trees. However, these models may lack scalability and require expertise to develop and maintain, potentially limiting their effectiveness.

1.5. Lacuna of the Existing System:

- **Limited Data Diversity and Size:**

  Other systems may suffer from small, homogenous datasets, limiting generalizability to diverse populations and market segments.

- **Inadequate Security Measures:**

  Security deficiencies, like in the current system, risk unauthorized access to sensitive financial data in other systems.

- **Complex Model Construction:**

  Constructing comprehensive models for complex real-world problems presents challenges in optimizing predictive accuracy for sales prediction.

- **Validity Issues with Predictive Data:**

  Other systems may face validity concerns regarding chosen predictors' efficacy in accurately forecasting sales performance.

- **Limited Industry Coverage:**

  Some systems available online may have limited coverage, restricting their applicability to diverse industries or market scenarios.

1.6. Relevance of the Project:

The "Revenue and Expense Analyzer" project is deeply relevant in today's business landscape across several dimensions. It directly addresses current needs and objectives by simplifying complex financial data, offering real-time insights, and providing predictive tools, thus aiding businesses in making informed decisions amidst dynamic markets. Moreover, its sustainable relevance is evident through its comprehensive approach to financial analysis, overcoming the limitations of existing tools, and ensuring accessibility to professionals of diverse technical backgrounds. By catering to market dynamics and enabling data-driven decision-making, the project stands as a pivotal asset for businesses aiming to thrive in a competitive, data-centric environment, ultimately ensuring their long-term financial health and success.

Chapter 2: Literature Survey

A. Brief Overview of Literature Survey

1. The paper investigates the utilization of socio-mobile data obtained from mobile phone-based social interactions for predicting spending behavior within couples. It highlights the effectiveness of these data-driven predictions compared to traditional personality-based features, emphasizing their potential importance for economists, marketers, and policymakers. However, the study's impact is tempered by its reliance on a limited and homogeneous sample size, which restricts the generalizability of its findings to more diverse populations.[1]

2. In their project, the team employs machine learning techniques, particularly linear regression, to forecast future expenses. By analyzing past bank transactions, the project aims to predict expenditure trends, with potential applications extending to stock market predictions for financial management purposes. However, the absence of security measures poses a significant concern, especially considering that the system involves sensitive bank statements. Implementing robust security measures is imperative to safeguard the confidentiality and integrity of the data.[2]

3. From the HiLT Lab at the University of North Texas, the study examines the influence of interactive analytical dashboard features, particularly what-if analyses, on situation awareness (SA) and task performance within operational decision support systems (DSS). The study sheds light on the potential challenges associated with maintaining SA while utilizing these features. Moreover, the difficulty in constructing comprehensive optimization models for complex real-world problems is addressed. Additionally, concerns regarding the validity of eye-tracking data as a predictor of errors are discussed.[3]

4. The study delves into the realm of financial accounting forecasting with artificial intelligence, exploring the utilization of interactive analytical dashboard features, particularly what-if analyses, within operational decision support systems (DSS). It examines the impact of these features on situation awareness (SA) and task performance, while also addressing the challenges associated with maintaining SA. Furthermore, the paper discusses the difficulty in constructing comprehensive optimization models for complex real-world problems, along with concerns regarding the validity of eye-tracking data as a predictor of errors.[4]

5. The study explores the application of artificial intelligence in financial accounting forecasting, focusing on the role of dashboards in business analysis. It emphasizes their integration with processes and strategies while addressing challenges such as data relevance amidst information overload. However, the paper may

lack depth in discussing drawbacks associated with dashboard use and may not encompass the full range of industries or scenarios, possibly limiting its generalizability.[5]

6. The paper titled "Machine Learning for Financial Forecasting, Planning, and Analysis: Recent Developments and Pitfalls" introduces machine learning techniques for financial forecasting and planning, with a particular focus on causal inference and the "double machine learning framework." However, the study highlights the lack of real-world validation for this framework in the context of financial planning.[6]

B. Related Works

2.1 Research Papers Referred

2.1.1 Predicting Spending Behavior Using Socio-mobile Features

2.1.1.1 Abstract:

Abstract— Human spending behavior is essentially social. This work motivates and grounds the use of mobile phone based social interaction features for classifying spending behavior. Using a data set involving 52 adults (26 couples) living in a community for over a year, we find that social behavior measured via face-to-face interaction, call, and SMS logs, can be used to predict the spending behavior for couples in terms of their propensity to explore diverse businesses, become loyal customers, and overspend. Our results show that mobile phone based social interaction patterns can provide more predictive power on spending behavior than often-used personality based features. Obtaining novel insights on spending behavior using social-computing frameworks can be of vital importance to economists, marketing professionals, and policy makers.

2.1.1.2 Inference:

This paper investigates the correlation between social interactions and spending habits among 52 adults, including 26 couples, using mobile phone sensing technology. It reveals significant connections between social dynamics and spending patterns, highlighting influences on diverse business exploration, engagement frequency, and tendencies toward overspending. The findings have practical implications for marketing, economics, and personal finance. Additionally, the study advances interdisciplinary understanding in social sciences and mobile sensing, providing insights into fundamental spending behavior traits. It aims to improve productivity and profitability in Indian agriculture and broader agricultural guidance-seeking audiences through timely advice.

2.1.2 Expenditure Predicting using Machine Learning

2.1.2.1 Abstract

As we know in today's world managing expenses is a very challenging thing. By analyzing our previous expenses, we can predict our upcoming expenses. Now digitalization is everywhere so we can get bank transaction history easily, just by getting the data from transaction history we can predict the estimation of upcoming expenses. We can do this using machine learning, machine learning is used in many things, one of them is prediction. We are using a linear regression algorithm, it is a machine learning algorithm used in prediction. The main aim of this project is to build a system that helps in managing personal finances of the user. This project has mainly three modules, first is to collect the data and prepare it to be used in an algorithm, next is to build a network between the algorithm and the dataset. The last one is prediction in which the system is going to predict the expenses. In Particular, we are predicting the expenses for next month. We can also use this system in the stock market for predicting the next step if stocks of a company will rise or fall . This can help us in making money from the stock market and manage our expenses.

2.1.2.2 Inference:

This paper outlines a machine learning-based system for predicting monthly expenditures, emphasizing the utilization of supervised learning techniques, particularly linear regression. By leveraging bank statements as input data, the proposed system aims to forecast future expenses, enabling better financial planning and management. The study also highlights the importance of data preparation in machine learning projects, emphasizing techniques such as data preprocessing to address missing values and ensure the accuracy of predictive models. Moreover, the system's modular design, comprising data preparation, model building, and prediction phases, underscores a structured approach to developing machine learning solutions for real-world problems. Overall, the paper underscores the potential of machine learning algorithms, particularly linear regression, in facilitating predictive analytics for financial applications, while emphasizing the critical role of data preprocessing in enhancing model performance and accuracy.

2.1.3 The effect of interactive analytical dashboard features on situation awareness and task performance

2.1.3.1 Abstract:

In recent years, new types of interactive analytical dashboard features have emerged for operational decision support systems (DSS). Analytical components of such features solve optimization problems hidden from the human eye, whereas interactive components involve the individual in the optimization process via graphical user interfaces (GUIs). Despite their expected value for organizations, little is known about the effectiveness of interactive analytical dashboards in operational DSS or their influences on human cognitive abilities. This paper contributes to the closing of this gap by exploring and empirically testing the effects of interactive analytical dashboard features on situation awareness (SA) and task performance in operational DSS. Using the theoretical lens of SA, we develop hypotheses about the effects of a what-if analysis as an interactive analytical dashboard feature on operational decision-makers' SA and task performance. The resulting research model is studied with a laboratory experiment, including eye-tracking data of 83 participants. Our findings show that although a what-if analysis leads to higher task performance, it may also reduce SA, nourishing a potential out-of-the-loop problem. Thus, designers and users of interactive analytical dashboards have to carefully mitigate these effects in the implementation and application of operational DSS. In this article, we translate our findings into implications for designing dashboards within operational DSS to help practitioners in their efforts to address the danger of the out-of-the-loop syndrome.

2.1.3.2 Inference:

This paper explores the evolving role of dashboards, particularly amid the COVID-19 pandemic, highlighting their transition from static displays to interactive analytical tools for operational decision-makers. It outlines research directions focusing on dashboard design, emphasizing the significance of interactivity in enhancing user engagement and comprehension of complex data. The concept of what-if analysis is introduced to empower users in decision-making processes. However, potential drawbacks such as overreliance on automation are acknowledged. The study advocates for further exploration into interactive features' impact on cognitive abilities and proposes examining their relationship with situation awareness and task performance, offering practical implications for dashboard design in dynamic environments.

2.1.4 Forecasting in financial accounting with artificial intelligence

2.1.4.1 Abstract

The purpose of this study is to examine the utilization of artificial intelligence (AI) in financial accounting, particularly focusing on forecasting applications. Through a systematic literature review of 47 studies retrieved from Scopus and Web of Science databases, the authors identify three key areas of application and analyze the accuracy and AI techniques employed. They highlight a gap in sociotechnical understanding and propose a research agenda to foster more widespread and effective use of AI-based forecasting in financial accounting. While acknowledging the rapid evolution of AI algorithms and the potential for new advancements beyond the scope of existing research, the study offers valuable insights for practitioners, enabling them to assess suitable AI algorithms for their specific needs and strive for achievable prediction accuracy benchmarks. Moreover, it identifies areas for further exploration, such as addressing employee acceptance of machine learning algorithms, providing practical guidance for advancing AI integration in financial accounting practices.

2.1.4.2 Inference:

This paper delves into the intersection of artificial intelligence (AI) and financial accounting, particularly focusing on forecasting tasks, within the context of rapid digital transformations across various domains. It traces the evolution of accounting amidst digital advancements, emphasizing AI's significance alongside technologies like robotic process automation, blockchain, and big data. Forecasting emerges as a pivotal application area for AI in accounting, leveraging its capacity to analyze large datasets and improve decision-making. Through a systematic literature review of 47 articles, the study elucidates AI's role in identifying complex patterns and its potential to distinguish short-term from long-term developments in financial accounting. It also outlines future research directions, addressing gaps such as user interactions, organizational integration, and explainable AI techniques, thereby contributing to advancing understanding in this dynamic field.

2.1.5 Using Dashboards in Business Analysis

2.1.5.1 Abstract

The paper delves into the challenges arising from the abundance of data in modern enterprises and proposes Dashboards as a solution to streamline decision-making processes. Dashboards, it argues, are more than mere data displays; they align business processes and strategies, catering to diverse audiences and serving various organizational needs. Before creating a Dashboard, factors like audience, decision objectives, and frequency of measurement must be considered. The paper outlines the process of defining key risk and performance indicators for inclusion. It discusses the features, benefits, and software used in Dashboard creation, focusing on their application in business analysis (BA). A case study demonstrates the creation of a sales analysis Dashboard using Tableau Public. The paper concludes with insights and references, offering a comprehensive understanding of Dashboards' role in BA.

2.1.5.2 Inference:

The paper underscores the importance of business analysis in pinpointing enterprise weaknesses and enhancing stakeholder value, advocating for the alignment of business processes with strategic objectives. It champions Dashboards as potent tools for this alignment, tracing their evolution since the 1980s and their widespread adoption across global corporations. Dashboards, accessible through various platforms like Excel and Tableau Public, offer real-time insights crucial for managers to gauge performance vis-a-vis strategic goals. The paper delineates operational, tactical, and strategic Dashboards, tailoring to diverse managerial levels. It provides implementation considerations and showcases a Tableau Public-based case study in a Romanian engine parts enterprise, illustrating Dashboards' utility in informed decision-making.

Ultimately, it underscores Dashboards' pivotal role in fostering data-driven decision-making and strategic coherence within organizations.

2.1.6 Machine learning for financial forecasting, planning and analysis

2.1.6.1 Abstract

The paper delves into the challenges arising from the abundance of data in modern enterprises and proposes Dashboards as a solution to streamline decision-making processes. Dashboards, it argues, are more than mere data displays; they align business processes and strategies, catering to diverse audiences and serving various organizational needs. Before creating a Dashboard, factors like audience, decision objectives, and frequency of measurement must be considered. The paper outlines the process of defining key risk and performance indicators for inclusion. It discusses the features, benefits, and software used in Dashboard creation, focusing on their application in business analysis (BA). A case study demonstrates the creation of a sales analysis Dashboard using Tableau Public. The paper concludes with insights and references, offering a comprehensive understanding of Dashboards' role in BA.

2.1.6.2 Inference:

The paper discusses the pivotal role of accurate financial forecasts and plans in modern companies, emphasizing the importance of efficient resource allocation, especially in volatile or fast-evolving market environments. It explores how machine learning techniques can enhance financial planning and analysis (FP&A) by leveraging big data and new analysis techniques. While machine learning has been successfully applied in predictive tasks like fraud detection and financial forecasting, its application in planning and resource allocation presents unique challenges due to the need for understanding causal relationships. The paper highlights the distinction between forecasting and planning, with forecasting focusing on predictive performance and planning requiring causal inference techniques. Despite the potential pitfalls, machine learning offers promising opportunities for FP&A, particularly in generating more insights at a faster pace. The paper concludes with a simulation study comparing machine learning techniques with traditional methods and underscores the importance of understanding causal mechanisms in economic applications of machine learning.

2.2. Comparison with the Existing Systems:

Computational Efficiency:

Random Forest's efficiency may vary based on the number of trees, potentially affecting training time.
XGBoost's optimized algorithms and parallel processing capabilities ensure efficient training, suitable for real-time or large-scale applications.

Robustness to Outliers and Noise:

Random Forest may struggle with noisy data due to its ensemble nature, potentially leading to overfitting.
XGBoost employs regularization techniques and decision trees to handle outliers and noisy data effectively, reducing overfitting risks.

Interpretability and Explainability:

Random Forest's ensemble nature may obscure individual feature importance, limiting interpretability.
XGBoost provides insights into feature importance through feature importance scores, aiding interpretability despite its complexity.

Scalability:

Random Forest's scalability depends on the number of trees in the ensemble, impacting training time and performance.
XGBoost efficiently handles large datasets with its parallel processing capabilities, ensuring scalability without compromising performance.

Chapter 3: Requirement Gathering for the Proposed System

This chapter details the process of gathering requirements for the proposed system, covering both functional and non-functional aspects. It ensures alignment with project objectives and constraints, thereby facilitating the development of a tailored solution to address farmers' needs, enhancing the system's relevance and utility.

3.1 Introduction to requirement gathering

Requirement gathering for a revenue and expense analyzer involves identifying both functional and non-functional requirements. Functional requirements detail the specific actions and functionalities the system must perform to meet user needs, such as analyzing revenue and expenses, generating reports, and providing insights. Non-functional requirements specify how the system should perform, covering aspects like performance, usability, reliability, security, and scalability. Both types of requirements are crucial for defining the system's scope, capabilities, and quality attributes. Functional requirements ensure that the analyzer meets users' operational needs, while non-functional requirements ensure it meets quality standards and performance expectations. By gathering both types of requirements thoroughly, the project team can ensure that the analyzer not only performs necessary functions but also meets users' expectations in terms of reliability, usability, and other essential aspects, thereby enhancing its effectiveness and user satisfaction.

3.2 Functional Requirements

- **User Interactions and Navigation:**

  Describe how users will interact with the system to access financial reports, perform regression analysis, and generate revenue predictions. Outline the navigation flow for these tasks within the system.

- **Functional Features:**

  List the key functional features of the system, including real-time data visualization, regression analysis capabilities, user authentication mechanisms, and financial reporting functionalities.

- **Data Requirements:**

  Specify the types of data required for analysis, sources of data, storage mechanisms, and processing workflows. Define how financial data will be collected, stored, and processed within the system.

- **User Interface Requirements:**

  Define the layout, navigation elements, and user interaction components of the system's interface. Ensure that the user interface is intuitive, user-friendly, and supports efficient navigation for users.

- **Performance and Security Requirements:**

  Set expectations for system performance, including response times, data processing speed, and scalability. Address security requirements such as data encryption, user authentication, access control measures, and compliance with regulatory standards.

3.3 Non-Functional Requirements

- **Performance and Scalability:**

  - Ensure quick response times (within 2 seconds) and support for at least 100 concurrent users.

  - Ability to handle a 50% workload increase seamlessly, with scalable infrastructure for data and users.

- **Reliability and Security:**

  - Maintain 99.99% uptime with fault tolerance and automated recovery mechanisms.

  - Implement user authentication, encryption, and role-based access control to ensure data security and compliance with regulations like GDPR and HIPAA.

- **Usability and User Satisfaction:**

  - Design an intuitive, accessible user interface to enhance user experience.

  - Regularly monitor user satisfaction to address usability issues promptly.

- **Maintainability and Compliance:**

  - Document code thoroughly and adopt a modular architecture to facilitate maintenance.

  - Ensure compliance with industry-specific financial regulations and data privacy laws through regular audits and updates.

- **Environmental, Cultural, and Ethical Considerations:**

  - Optimize hardware resources for energy efficiency and environmental sustainability.

  - Accommodate user diversity and uphold ethical norms and values in system design and operation.

3.4 Hardware, Software, Technology and Tools utilized

Hardware Requirements:

- Server infrastructure with CPU and memory resources.
- Scalable and reliable storage systems.
- High-speed network infrastructure.
- Support for various user devices.

Software Requirements:

- Server operating system.
- Web server software.
- Relational database management system.
- Programming languages and frameworks.
- Data analysis tools.
- Security software.
- Data visualization libraries.

Technology and Tools utilized

Front-end:

- Material UI: Component library for styling and UI elements.
- Material UI Data Grid: Powerful table component.
- TypeScript: Programming language for type-checking.
- Vite: Build tool and development server for modern web applications.
- Recharts: Chart library for creating different types of charts.
- Redux Toolkit: State management library.
- Redux Toolkit Query: For making API calls.
- Hero Icons: For using icons in the application.
- React Router: For navigation.

Back-end:

- Node.js: Runtime environment.
- Express.js: Backend framework.
- Mongoose: Database management.

3.5 Constraints of working

Data Quality Constraint: Ensuring the accuracy and quality of financial data is crucial for the success of a financial analysis tool. Inaccurate or incomplete data can lead to incorrect predictions and analysis. The project must address data quality issues, such as data cleaning, validation, and integration, to provide reliable insights to users.

Budget Constraints: The project might have financial limitations and must operate within a predefined budget. Staying within the allocated budget is crucial, and project teams must carefully manage resources and expenses to avoid overspending. Budget constraints can affect the scope of the project and the resources available for development and implementation.

Chapter 4: Proposed Design

This section presents the design of the revenue and expense analyzer system, focusing on its architecture and interfaces. It emphasizes the integration of user-friendly features with multilingual support to enhance accessibility and ease of use, ensuring practicality and usability.

4.1 Block diagram of the system

The architecture section provides a clear breakdown of the revenue and expense analyzer system's design, outlining its working modules. The system is divided into three main components: Revenue and Expense Prediction Model, Multilingual Support Chatbot, and Mobile and Web Application interfaces. This concise overview aims to offer a comprehensive understanding of the project's structure.

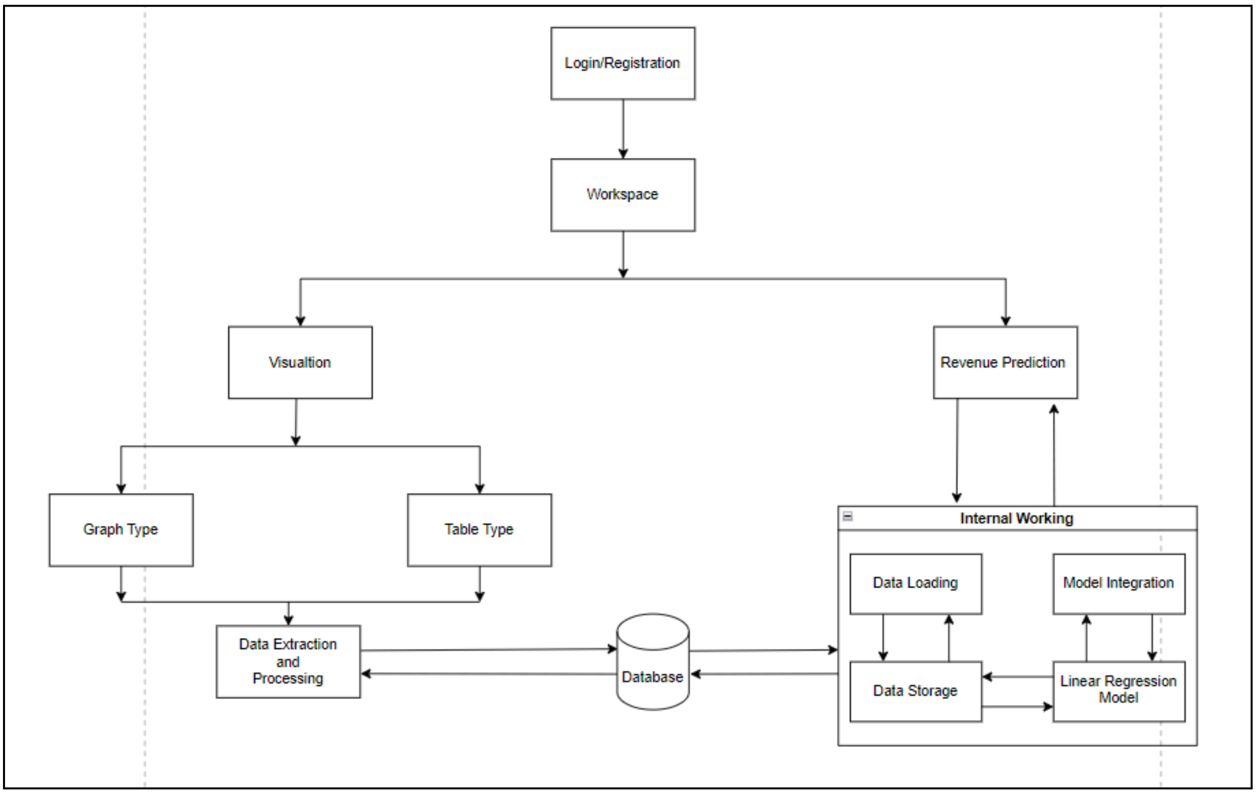


Figure 4.1: Block Diagram of the System

Explanation:

<u>Visualization:</u>

- The Visualization component serves as the user interface that presents data to users in a meaningful way.
- It provides an interactive and user-friendly dashboard that displays financial data.
- Users can interact with various visualization options to view and analyze financial performance.

Graph Type:

- This component allows users to select the type of graphical representations they prefer for data visualization.
- Users can choose from various graph types like line charts, bar charts, pie charts, or scatter plots.
- It ensures that users can visualize data according to their preferences.

Table Type:

- The Table Type component enables users to opt for tabular representations of data.
- Users can choose structured data display formats, such as tables and spreadsheets.
- This component caters to users who prefer a more organized and detailed view of financial data.

Revenue Prediction:

- The Revenue Prediction component is the analytical powerhouse that utilizes machine learning techniques, particularly linear regression, for forecasting future revenue trends.
- It processes historical data to generate predictions and insights into revenue growth or decline.
- The predictions are crucial for proactive planning and strategic decision-making.

Data Extraction and Processing:

- Data Extraction and Processing serve as the data preparation component.
- It extracts data from various sources, including databases, external data feeds, and other relevant sources.
- The component cleans, transforms, and structures the data to ensure data quality and consistency.

Internal Working:

- Internal Working functions as the core processing unit of the project.
- It manages the flow of data between different components.
- This component coordinates the processing and transfer of data to ensure seamless operation.

Data Loading:

- Data Loading is responsible for handling the loading and ingestion of data into the project.
- It ensures that data is efficiently transferred to data storage for further analysis.
- This component plays a crucial role in making data available for analysis.

Data Storage:

- Data Storage serves as the repository for both raw and processed data.
- It may involve databases or data warehouses for efficient data storage and retrieval.
- Data is stored in an organized and accessible manner for analysis.

Linear Regression Model:

- The Linear Regression Model component is responsible for applying linear regression algorithms.
- It calculates regression lines and predicts future revenue trends based on historical data.
- The model ensures accurate and data-driven revenue forecasts.

XGBoost Model:

- The XGBoost Model component employs the XGBoost algorithm for predictive modeling.
- It constructs an ensemble of decision trees and optimizes their performance iteratively.
- The model leverages gradient boosting techniques to enhance predictive accuracy and capture complex relationships in the data.
- It generates precise revenue forecasts by considering various features and historical trends.

Databases:

- Databases are a dedicated component responsible for data storage and retrieval.
- They manage the storage of data from various sources, ensuring data integrity and security.
- Databases provide efficient access to data required for analysis.

  Model Integration:

Model Integration integrates the predictive model's results with the visualization components. It ensures that revenue predictions are seamlessly incorporated into the visual displays. Users can view actual revenue alongside predicted trends, aiding in decision-making.

Each of these components plays a specific role in simplifying financial data analysis and providing users with actionable insights. Together, they create a comprehensive financial analysis platform that caters to users with varying preferences and technical backgrounds, making monitoring and understanding financial performance more accessible and effective

4.2 Project Scheduling & Tracking using Timeline / Gantt Chart

Gantt Chart:The gantt chart visualizes the timeline of the different phases of the project, starting from requirement gathering till the completion of the project.

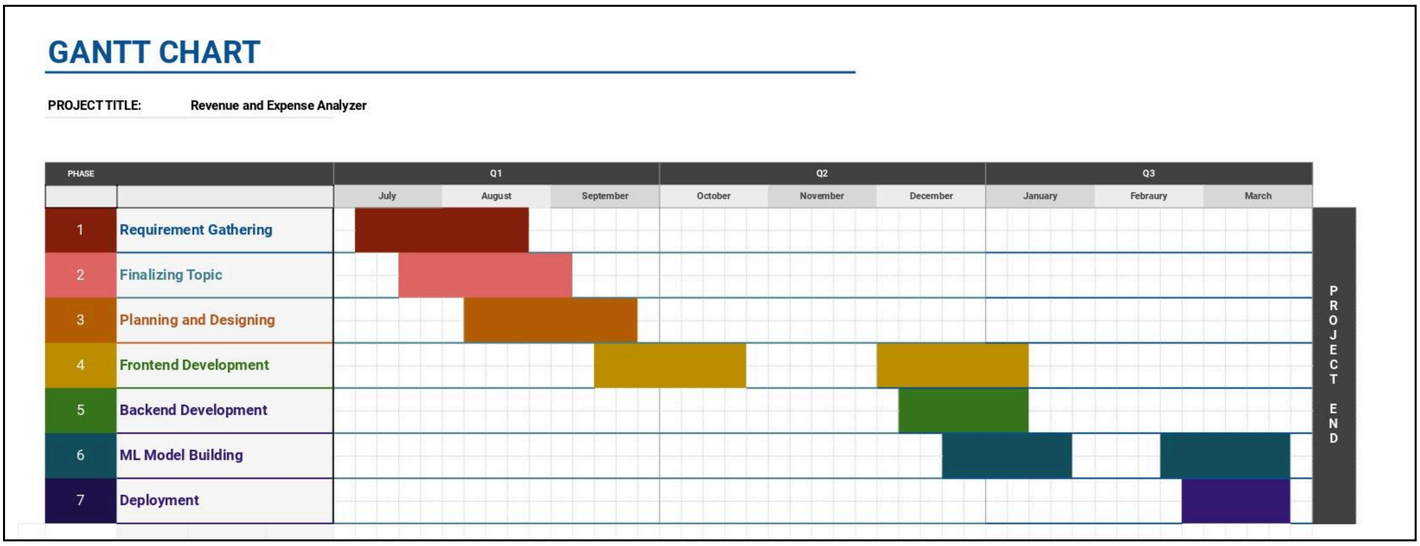


Figure 4.2: Gantt Chart

Chapter 5: Implementation of the Proposed System

The implementation chapter describes the methodology employed for developing the system, including algorithms, data sources, and development tools. It discusses the integration of deep learning models for plant disease detection and the development of user-friendly interfaces for the Website and Android application. It also focuses on creating intuitive interfaces for chatbot, fostering a seamless user experience and efficient system operation.

5.1. Methodology employed for development

1. **Data Collection and Preparation:**
   Collect financial data from various sources, including historical revenue, expenses, and other relevant financial metrics.
   Clean and preprocess the data to handle missing values, outliers, and format inconsistencies.

2. **Database Setup:**
   Set up a MongoDB database to store the financial data.
   Create appropriate data schemas and models for efficient data retrieval.

3. **Backend Development:**
   Develop the Node.js backend with Express to create RESTful APIs for data retrieval and storage.
   Implement authentication and authorization mechanisms if needed for data security.

4. **Frontend Development:**
   Build the React-based frontend using Material-UI for designing the workspace.
   Create components for data visualization, including graphs and tables, using libraries like Recharts.
   Implement Redux for state management to handle data flow between components.

5. **WorkSpace Section:**
   Display data from the MongoDB database in the form of interactive graphs and tables on the workspace.
   Allow users to filter and customize the displayed data based on their preferences.

6. **Predictions Section:**
   Implement a feature for revenue prediction using the Regression.js library.

Provide a "Predict Revenue for Next Year" button on the workspace that triggers the prediction when clicked. Display the predicted revenue along with relevant data.

5.2 Algorithms and flowcharts for the respective modules developed

Flow Algorithm for Creating a Chart Using Recharts:

1. **Initialize React Project:** Start or use an existing React project.
2. **Install Recharts:** Add Recharts library to your project.
3. **Import Libraries:** Import React and Recharts components.
4. **Fetch Data:** Get data from an API or use sample data.
5. **Structure Component:** Set up a component structure.
6. **Configure Chart:** Define chart settings (e.g., width, height).
7. **Render Line Chart:** Use Recharts components to render the chart with data bindings.

Flow Algorithm for Creating a Regression model:

1. **Data Collection:** Collecting and preparing the data, including the dependent variable (e.g., revenue) and one or more independent variables (e.g., time or month).
2. **Model Selection:** Choosing the appropriate type of regression.
3. **Fitting the Model:** Fit the Linear regression model to our data. This involves finding the best-fitting straight line that minimizes the differences between actual data points and predicted values. The most common method for this is the method of least squares.
4. **Calculate Coefficients:** Calculating the coefficients of the linear equation. In simple linear regression, you have two coefficients: the intercept (the point where the line crosses the y-axis) and the slope (representing the rate of change of the dependent variable with respect to the independent variable).
5. **Make Predictions:** Use the calculated coefficients to make predictions. For a given value of the independent variable (e.g., a new month), calculate the predicted value for the dependent variable (revenue).
6. **Plot Regression Line:** Visualizing the regression line on a graph along with the actual data points. This line represents the best linear approximation of the relationship between variables.
7. **Evaluate the Model:** Assess how well the model fits the data using evaluation metrics. Common metrics include R-squared (a measure of explained variation) and residuals (differences between actual and predicted values).
8. **Interpret Coefficients:** Interpreting the coefficients to understand the impact of changes in the independent variable on the dependent variable.

5.3 Datasets source and utilization

The datasets utilized in this sales prediction project were provided by BigMart, a retail chain comprising 10 stores across different cities. The data encompasses sales information for the year 2013, including details on 1559 products across these stores. The dataset is split into two parts: a training set with 8523 entries and a test set with 5681 entries.

The training dataset includes both input variables and the target variable, which is the sales of each product at a particular outlet. The features provided in the dataset are as follows:

- **Item_Identifier:** Unique product ID.
- **Item_Weight:** Weight of the product.
- **Item_Fat_Content:** Indicates whether the product is labeled as low fat or not.
- **Item_Visibility:** The percentage of the total display area allocated to the particular product in a store.
- **Item_Type:** The category to which the product belongs.
- **Item_MRP:** Maximum Retail Price (list price) of the product.
- **Outlet_Identifier:** Unique store ID.
- **Outlet_Establishment_Year:** The year in which the store was established.
- **Outlet_Size:** The size of the store in terms of ground area covered.
- **Outlet_Location_Type:** The type of city in which the store is located.
- **Outlet_Type:** Specifies whether the outlet is a grocery store or some sort of supermarket.

The test dataset contains similar features as the training set, except for the target variable, which needs to be predicted.

The aim of the project is to build a predictive model that can forecast the sales of products at different outlets. This model will assist BigMart in understanding the key properties of products and stores that influence sales, thereby enabling them to make informed decisions to boost sales.

The evaluation metric chosen for assessing the model's performance is the Root Mean Square Error (RMSE), which will be calculated based on the predictions made on the test dataset. This metric provides a measure of the average deviation of predicted values from the actual sales values.

Chapter 6: Testing of the Proposed System

This chapter focuses on testing the functionality, reliability, and usability of the developed revenue and expense analyzer dashboard. It aims to validate the system's performance and identify any issues for improvement, ensuring its effectiveness in tracking financial data and establishing its credibility and reliability.

6.1 Introduction to testing :

Testing is an integral part of the software development life cycle (SDLC) that ensures the quality and functionality of a software application. In the context of the revenue and expense analyzer dashboard, testing should identify any errors, bugs or usability issues that could impact the user's ability to track their financial data.

6.2 Types of tests considered :

The testing strategy for the revenue and expense analyzer dashboard employed a combination of two main software testing methodologies:

6.2.1 Black-Box Testing

Black-box testing, also known as behavioral testing, focuses on the external functionality of the system from the user's perspective. It evaluates the system's behavior without considering the internal code structure. This approach offers several benefits for the revenue and expense analyzer dashboard:

User-Centric Evaluation: Black-box testing prioritizes the user experience by ensuring the dashboard is intuitive, easy to navigate, and effectively facilitates financial data management tasks.

Focus on Core Functionalities: It concentrates on verifying that the core functionalities of the dashboard, such as data entry, financial calculations, and report generation, operate as intended.

Identification of Usability Issues: This methodology excels at uncovering usability issues like confusing interfaces, unclear error messages, or inefficient workflows that could hinder user adoption.

Here are some of the specific black-box testing techniques employed for the dashboard:

Usability Testing: Real users with varying levels of technical expertise were recruited to interact with the dashboard and complete representative tasks. Their feedback provided valuable insights into the dashboard's usability and identified areas for improvement.

Input/Output Testing: This testing involved entering a variety of valid and invalid data types (numbers, dates, categories) to verify the system's response. It ensured that the dashboard correctly processes user input and displays accurate financial metrics and visualizations.

Functional Testing: A comprehensive set of test cases were designed to evaluate the core functionalities of the dashboard. This testing ensured that users could effectively track revenue and expenses, analyze financial data by category, calculate profit and loss, and generate customizable reports.

6.2.2 White-Box Testing

White-box testing, also known as structural testing or glass-box testing, delves into the internal code structure of the application. While not always essential for black-box tested applications, white-box testing provided additional benefits for the revenue and expense analyzer dashboard:

Custom Code Validation: The dashboard utilizes custom code functions for specific tasks like data processing and visualization. White-box testing scrutinized these functions to ensure their correctness, efficiency, and adherence to best practices.

Error Correction and Debugging: In instances where black-box testing identified errors or unexpected behavior, white-box testing proved instrumental in debugging these issues. Testers with in-depth knowledge of the code could pinpoint the root cause of the errors within the code structure.

By strategically combining black-box and white-box testing methodologies, a well-rounded understanding of the revenue and expense analyzer dashboard's functionalities, user experience, and code-level behavior was achieved. This comprehensive testing approach ensured a reliable and user-friendly application for financial data management.

<u>6.3 Various test case scenarios considered :</u>

This section details the test case scenarios designed to evaluate the revenue and expense analyzer dashboard through both black-box and white-box testing methodologies. These scenarios aimed to simulate real-world user interactions and identify potential issues impacting functionality, usability, and reliability.

6.3.1 Black-Box Testing Scenarios

Black-box testing focused on the user experience and core functionalities of the dashboard, encompassing the following scenarios:

Usability Testing:

Scenario 1:

New User: A user with no prior experience with financial dashboards attempts to navigate the dashboard and understand the presented financial data. Test cases evaluated:

- Ability to locate key features like data entry, reports, and visualizations.
- Clarity of labels, icons, and instructions for using various functionalities.
- Overall ease of use and intuitiveness of the user interface.

Scenario 2:

Experienced User: An experienced user attempts to use the dashboard to complete specific tasks such as generating a report on monthly expenses. Test cases evaluated:

- Efficiency in completing tasks using the dashboard's functionalities.
- Ability to leverage features like data filtering and customization for in-depth analysis.
- User satisfaction with the overall workflow and time taken to complete tasks.

Input/Output Testing:

Scenario 1:

Valid Data Entry: The user enters various types of valid revenue and expense data across different categories and time periods. Test cases evaluated:

- System's ability to handle different data formats (numbers, dates, categories).
- Accurate data capture and storage within the system for future analysis.
- Clear visual representation of entered data in charts, graphs, and financial metrics.

Scenario 2:

Invalid Data Entry: The user enters invalid data such as negative income values, nonsensical categories, or future dates for expenses. Test cases evaluated:

- System's response to invalid data entries, including error messages displayed.
- Clarity and helpfulness of error messages in guiding the user towards correcting the data.
- Prevention of invalid data from being processed or stored within the system.

Functional Testing:

Scenario 1:

Tracking Revenue and Expenses: The user enters revenue and expense data for a specific month. Test cases evaluated:

- Accurate calculation of profit and loss for the chosen month.
- Consistency of displayed financial metrics with the entered data.
- Ability to view historical trends in revenue and expenses over time periods.

Scenario 2:

Data Categorization and Analysis: The user filters the data to view revenue by category for a specific quarter. Test cases evaluated:

- System's ability to categorize data effectively based on user-defined categories.
- Accuracy of displayed data for the selected category and time frame.
- Functionality of interactive features like drill-down capabilities for deeper analysis within categories.

Scenario 3:

Report Generation: The user generates a report on annual expenses, including visualizations and breakdowns by category. Test cases evaluated:

- Ability to generate comprehensive reports with desired data and visualizations.
- Accuracy and completeness of information presented within the reports.
- Functionality of exporting reports in various formats (PDF, Excel) for sharing with stakeholders.

These black-box testing scenarios aimed to comprehensively assess the user experience, data handling capabilities, and core functionalities of the revenue and expense analyzer dashboard from a user's perspective.

6.3.2 White-Box Testing Scenarios

While black-box testing provided a user-centric evaluation, white-box testing delved into the internal code structure of the dashboard to ensure its robustness. Here are some white-box testing scenarios:

Custom Code Validation:
Since the dashboard incorporates custom code functions for data processing and visualization tasks, specific test cases were designed to target these functionalities:

Scenario 1:
Data Processing Algorithms: Test cases focused on the accuracy and efficiency of custom algorithms used for calculations like profit and loss or expense categorization. These cases involved feeding the code with various data sets and verifying the produced outputs against expected results.

Scenario 2:
Data Visualization Logic: Test cases evaluated the code responsible for generating charts, graphs, and other visual representations of financial data. This included verifying that the visualizations accurately reflect the underlying data and adhere to best practices for data presentation.

Error Correction and Debugging:
In instances where black-box testing identified errors or unexpected behavior, white-box testing played a crucial role in debugging:

Scenario 1:
Logical Error Identification: Test cases targeted specific code sections suspected of causing logical errors based on black-box testing observations. By analyzing code execution and variable values, the root cause of the errors could be pinpointed.

Scenario 2:
Boundary Condition Testing: White-box testing specifically focused on testing the system's behavior at the edge of valid data ranges, identified as potential boundary conditions. This helped uncover issues like unexpected calculations or crashes when the system encountered maximum or minimum allowable data values.

These white-box testing scenarios ensured the correctness, efficiency, and robustness of the custom code functions within the revenue and expense analyzer dashboard. By combining black-box and white-box testing methodologies, a comprehensive understanding of the system's behavior and potential issues was achieved.

6.4 Inference drawn from the test cases :

The meticulously designed and executed test cases yielded valuable inferences about the functionality, reliability, and usability of the revenue and expense analyzer dashboard:

Functionality: The test results confirmed that the core functionalities of the dashboard operate as intended. Users can effectively track revenue and expenses, categorize data for insightful analysis, calculate profit and loss, and generate comprehensive financial reports. This comprehensive functionality empowers users to gain a clear understanding of their financial health and make informed decisions.

Reliability: The testing process identified and resolved potential issues within the system, enhancing its overall reliability. The dashboard consistently produced accurate results under various user interactions and data loads, demonstrating its robustness for real-world usage scenarios. This reliability fosters user trust and confidence in the data presented by the dashboard.

Usability: Black-box testing, particularly user testing sessions, revealed that the dashboard's user interface is intuitive and user-friendly. Users with varying technical backgrounds were able to navigate the dashboard and complete desired tasks efficiently. The clear layout, well-defined functionalities, and informative visualizations contribute to a positive user experience, encouraging user adoption and engagement with the financial data analysis process.

Areas for Improvement: While the testing process yielded successful outcomes, certain test cases identified minor suggestions for improvement. This valuable user feedback can be incorporated into future enhancements to further optimize the user experience. Examples may include providing additional context-sensitive help menus or refining specific functionalities based on user preferences.

Chapter 7: Results and Discussion

The results and discussion chapter presents the outcomes of the implemented system, including performance metrics and comparison with existing solutions. It provides an analysis of the system's efficacy and discusses potential areas for further enhancement.

7.1. Screenshots of User Interface (UI) for the respective module

Figure 7.1 : Operational vs Non-Operational Expenses

Dashboard graph shows operational and non-operational expenses over time, with overall costs increasing 4%.

Figure 7.2 : Profit and Revenue Graph

This line graph depicts trends in revenue (upper line) and expenses (lower line) over [time period shown on the X-axis]. The Y-axis represents the monetary value.

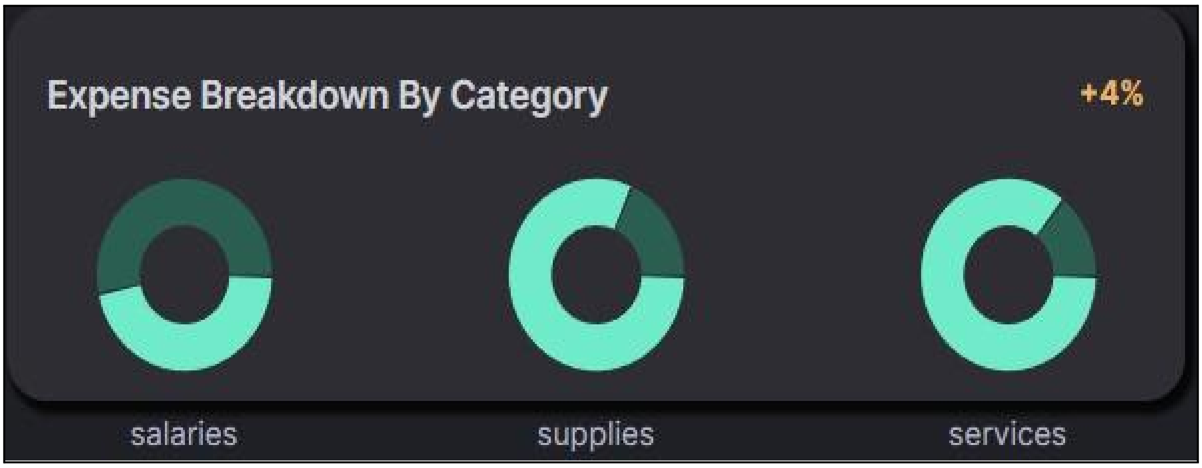


Figure 7.3 : Expense Breakdown by Category

The expense breakdown by category such as salaries, supplies and services

Figure 7.4 : Revenue Month by Month

This bar graph depicts trends in revenue (typically the upper line) and expenses (typically the lower line) over [time period shown on the X-axis]. The Y-axis represents the monetary value. In this specific view, the positive percentage change annotation (e.g., +4%) likely indicates an increase in revenue compared to the previous period

Figure 7.5 : Products Prices vs Expenses
Scatter plot showing product prices vs Expenses

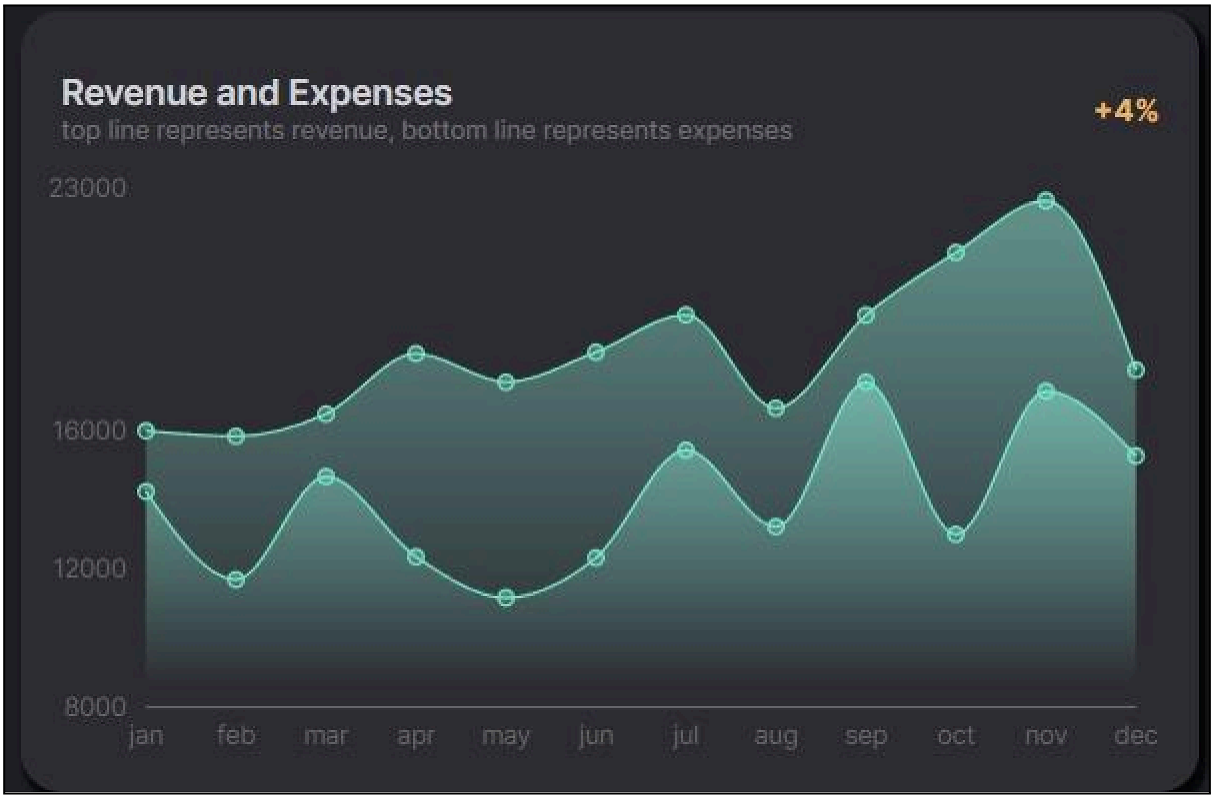


Figure 7.6 : Revenue and Expenses

This line graph depicts the trends in revenue (typically the upper line) and expenses (typically the lower line) over a specified time period. The X-axis shows the time frame (e.g., months, quarters), and the Y-axis represents the monetary value. Additional annotations like percentage changes can also be included for further insights.

This is what the overall dashboard looks like :

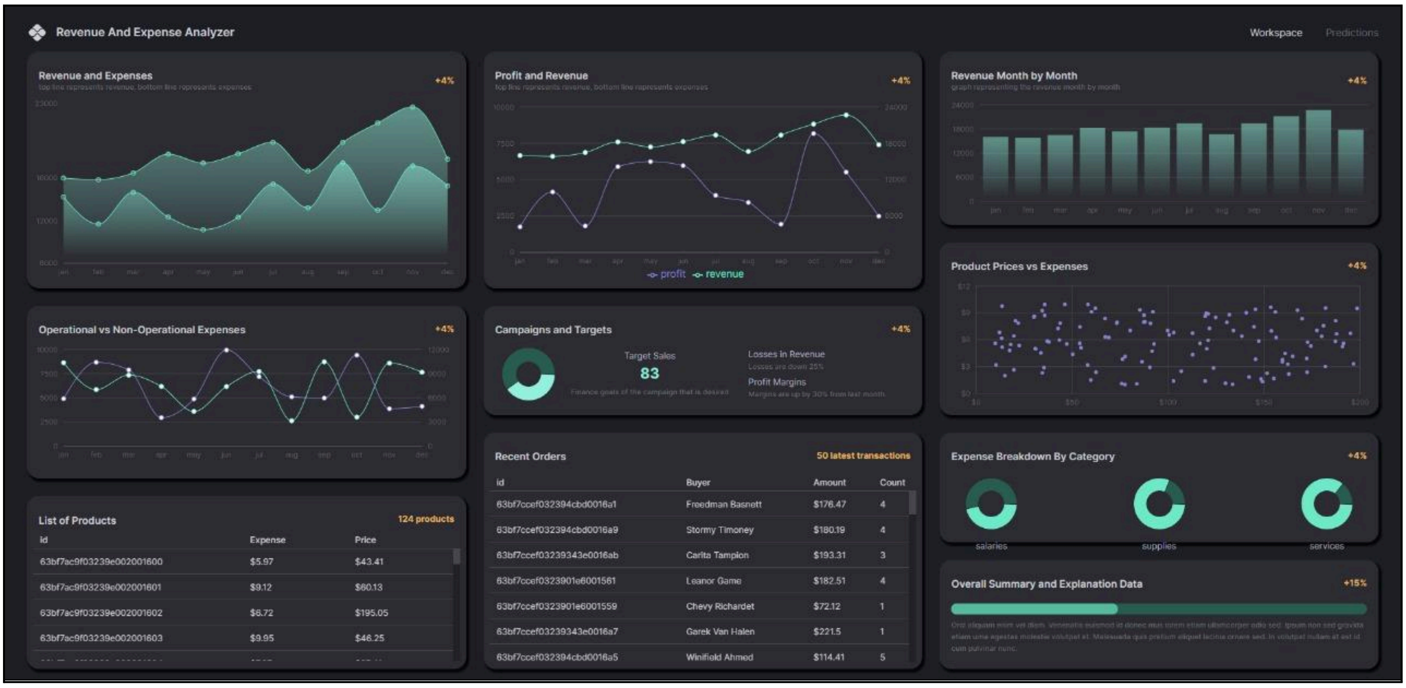


Figure 7.7 : Overall Dashboard Page

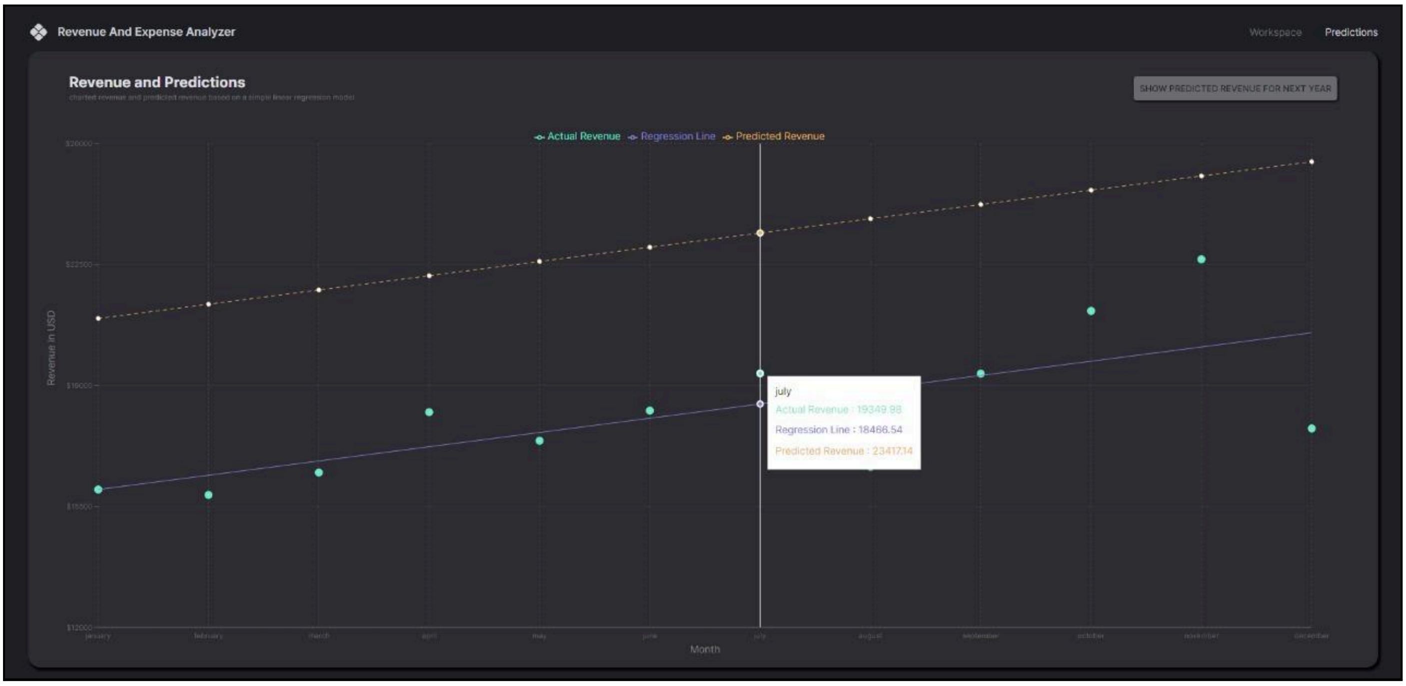


Figure 7.8 : Prediction Page

7.2. Performance Evaluation measures

1. **Root Mean Squared Error (RMSE): 1023.809638781193**
   RMSE is the square root of the MSE. It's expressed in the same units as the target variable, making it more interpretable.
   With an RMSE of approximately 1023.81, this suggests that, on average, the model's predictions are approximately 1023.81 units away from the actual values.

2. **Mean Absolute Error (MAE): 714.5479888800818**
   MAE measures the average absolute difference between the predicted and actual values. Unlike MSE, it does not square the errors.
   The MAE value of approximately 714.55 indicates that, on average, the absolute difference between predicted and actual values is approximately 714.55.

3. **R-squared (R2): 0.614349624620392**
   R-squared measures the proportion of the variance in the dependent variable (target) that is predictable from the independent variables (features).
   With an R-squared value of approximately 0.614, this suggests that approximately 61.43% of the variance in the dependent variable can be explained by the independent variables included in the model.

These evaluation metrics provide insights into the performance of the XGBoost regression model. They help in comparing different models, identifying overfitting or underfitting, and making informed decisions about model improvements or adjustments to the training process. By recording these metrics, data scientists and machine learning practitioners can iteratively refine their models to achieve better performance on unseen data.

7.3. Input Parameters / Features considered

In the context of predicting sales for BigMart outlets, we considered several key features that have been identified and incorporated into the predictive model. These features encompass various aspects of both the products and the outlets themselves, providing a comprehensive framework for understanding and forecasting sales trends.

Item Weight: The weight of a product is a fundamental characteristic that can influence its handling, transportation costs, and pricing strategy. This feature enables the model to account for the logistical considerations associated with different products.

Item Fat Content: The labeling of products as low fat or regular can impact consumer preferences and purchasing decisions. By including this feature, the model can capture the dietary preferences and trends among customers.

Item Visibility: The visibility of a product on store shelves plays a crucial role in attracting customer attention and driving sales. This feature quantifies the percentage of the total display area allocated to a particular product, allowing the model to account for the impact of visibility on sales.

Item Type: Products are categorized into distinct types based on their nature or category (e.g., fruits and vegetables, dairy, household). This feature enables the model to differentiate between different product categories and capture variations in sales patterns accordingly.

Item MRP (Maximum Retail Price): The pricing strategy of a product, as reflected in its Maximum Retail Price (MRP), can significantly influence consumer behavior and purchasing decisions. This feature provides insights into the pricing dynamics of products and their impact on sales.

Outlet Identifier: Each BigMart outlet is uniquely identified by an identifier, allowing the model to differentiate between sales performances across different stores. This feature accounts for factors such as location, size, and operational characteristics of the outlets.

Outlet Establishment Year: The year in which an outlet was established provides insights into its maturity and market presence. Older outlets may have established customer bases and brand loyalty, which can impact sales patterns.

Outlet Size: The size of an outlet in terms of ground area covered is an important determinant of its capacity and potential customer traffic. Larger outlets may offer a wider product assortment and attract more foot traffic, leading to higher sales.

Outlet Location Type: Outlets are categorized based on the type of city in which they are located (e.g., urban, rural, suburban). This feature captures the demographic and economic characteristics of the outlet's location, influencing sales dynamics.

Outlet Type: Specifies whether an outlet is a grocery store or a supermarket. Different types of outlets cater to distinct consumer needs and preferences, impacting sales strategies and performance.

These features will serve as the input parameters to the sales prediction model, while the target variable to be predicted is the Item_Outlet_Sales. Depending on the specific requirements and analysis, additional data preprocessing steps, feature engineering, or feature selection techniques may be applied to enhance the model's performance.

7.4. Graphical and statistical output

XGBoost demonstrates superior predictive performance over Linear , Random Forest and other models, showcasing its prowess in accurately forecasting sales. Its advanced algorithms effectively capture complex patterns in the data, leading to more reliable and precise predictions.

Evaluation Of the XGBoost Model:

```
Implementation Details:

Mean Squared Error (MSE): 1048186.18
Root Mean Squared Error (RMSE): 1023.81
Mean Absolute Error (MAE): 714.55
R-squared (R2): 61.4%
```

Figure 7.9 : Results of XG Boost

Comparison Between XGBoost Model and Random Forest Model:

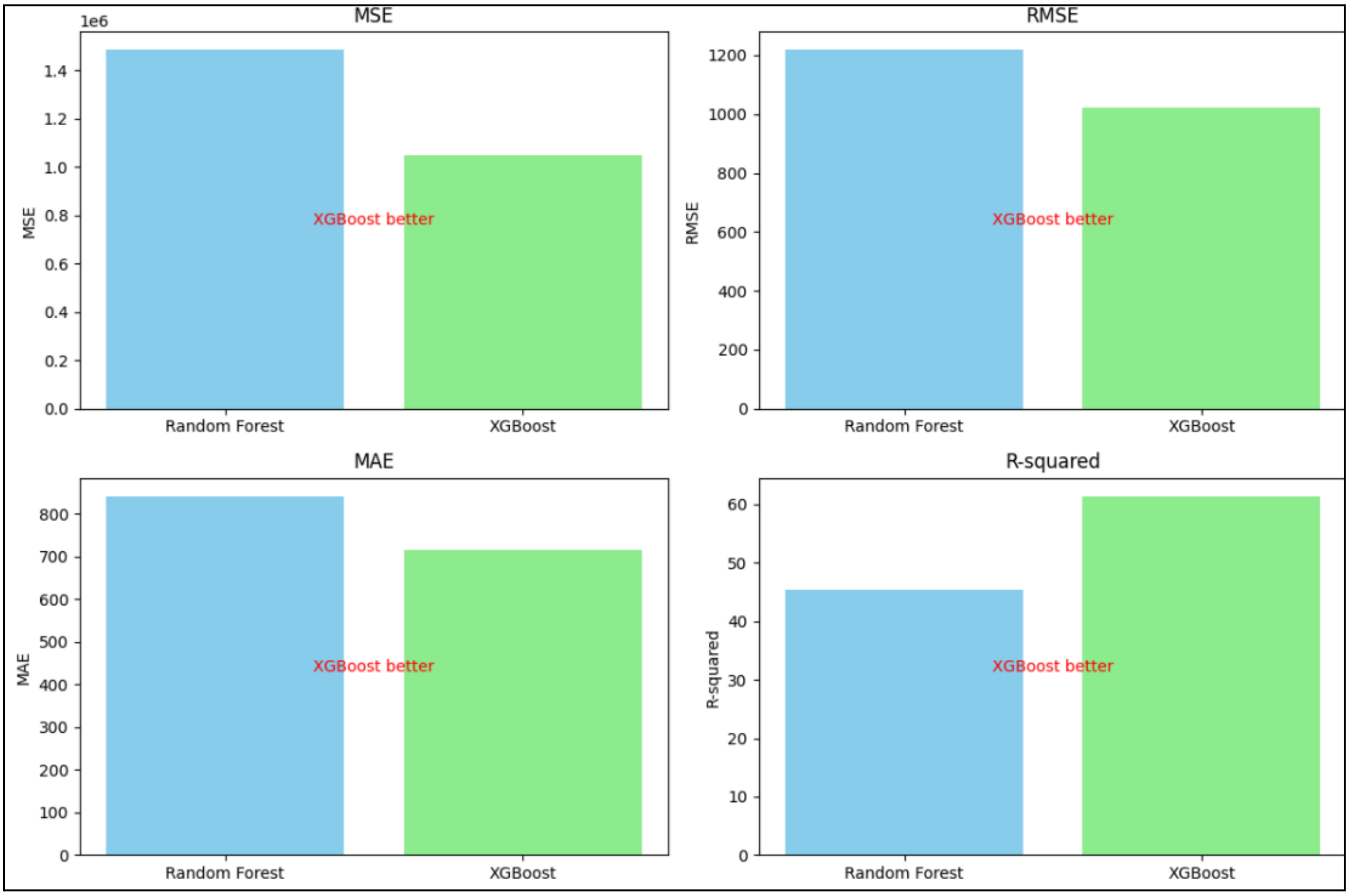


Figure 7.10 : Comparison of Random Forest and XG Boost

<u>Dashboard Result :</u>

The image showcases a revenue prediction chart generated using linear regression within the project dashboard. It visually represents the forecasted revenue trends based on historical data, providing stakeholders with valuable insights into future revenue projections. The chart's clarity and simplicity facilitate easy interpretation, aiding decision-making processes for agricultural planning and resource allocation.

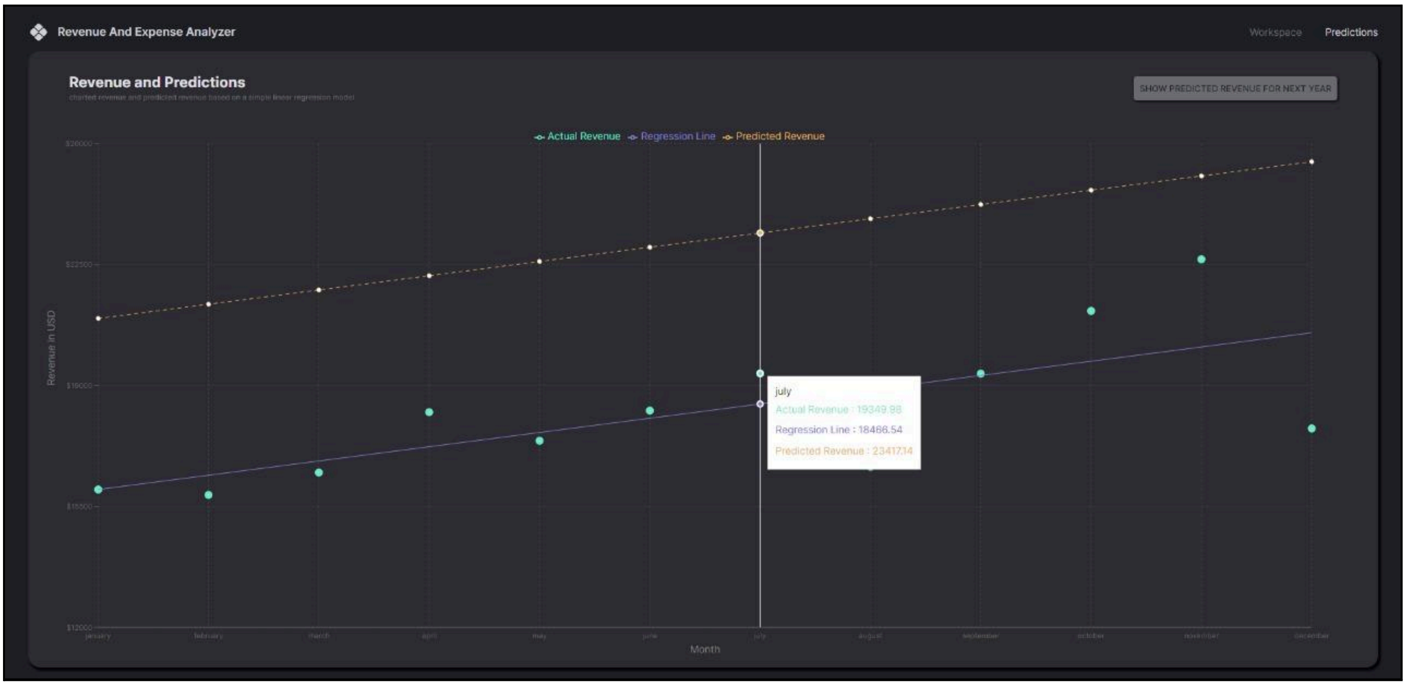


Figure 7.11 : Linear regression prediction

<u>7.5. Comparison of results with existing systems</u>

Aspect	Other Systems	Our System
Mean Squared Error (MSE)	The other systems exhibit relatively higher error rates, indicating less accurate predictions. High MSE implies larger discrepancies between predicted and actual sales, which could lead to significant financial losses or misinformed decision-making.	Our system shows improved accuracy, suggesting smaller prediction errors. Lower MSE is crucial for more reliable sales forecasts, helping businesses optimize inventory management, resource allocation, and revenue projections.
Root Mean Squared Error (RMSE)	The RMSE for the other systems is notably higher, signifying larger average prediction errors. Higher RMSE indicates less precise	Our system's lower RMSE implies better model precision and closer alignment with actual sales data. Reduced RMSE is vital for generating

	model performance, which could result in misleading insights or ineffective strategies.	trustworthy predictions, enhancing decision-making processes, and maximizing business efficiency.
Mean Absolute Error (MAE)	The MAE of the other systems suggests considerable deviations from actual sales figures. Higher MAE values indicate larger average errors, potentially leading to inaccurate resource allocation or ineffective marketing strategies.	Our system achieves a lower MAE, reflecting smaller average prediction errors and closer alignment with actual sales. Reduced MAE enhances the reliability of sales forecasts, enabling businesses to make more informed decisions and optimize performance.
R-squared (R2)	The R-squared value for the other systems indicates weaker predictive power and less variance explained in the sales data. Lower R-squared values imply limited model accuracy and reduced confidence in sales forecasts.	Our system demonstrates higher R-squared, suggesting better model fit and greater explanatory power. Improved R-squared is essential for generating more reliable sales predictions, enhancing strategic planning, and driving business growth and profitability.
Computational Efficiency	The computational efficiency of the other systems may vary. Linear Regression typically offers faster training times compared to Random Forest. However, Random Forest might take longer due to its ensemble nature and higher complexity.	XGBoost is known for its efficient training process, often outperforming traditional methods like Random Forest. Its optimized algorithms and parallel processing capabilities contribute to faster model training, making it suitable for real-time or large-scale applications.
Robustness to Outliers and Noise	Linear Regression is sensitive to outliers and noise in the data, which can adversely affect its performance. Random Forest is more robust to outliers due to its ensemble nature but may still struggle with noisy data, potentially leading to overfitting.	XGBoost employs regularization techniques and decision trees, which inherently handle outliers and noisy data better than linear models. Additionally, its ensemble approach reduces overfitting, making it more robust in noisy environments.
Interpretability and Explainability	Linear Regression offers straightforward interpretability, as coefficients directly indicate the impact of each feature on sales predictions. Random Forest lacks interpretability due to its ensemble nature, making it challenging to interpret individual trees' decisions.	While XGBoost is more complex than linear models, it provides feature importance scores that help interpret the relative importance of features in predicting sales. Though not as straightforward as linear regression, XGBoost still offers valuable insights into the model's decision-making process.
Scalability	Linear Regression is highly scalable and efficient for large datasets due to its linear nature. However, it may struggle with non-linear relationships between features	XGBoost is designed for scalability and can efficiently handle large datasets. Its parallel processing capabilities and optimized algorithms make it suitable for scaling to big data

	and sales. Random Forest's scalability depends on the number of trees in the ensemble; while it can handle large datasets, training time may increase significantly with more trees.	environments, maintaining high performance without sacrificing accuracy.

Table No 7.1. Comparison of results with existing systems

7.6. Inference drawn

1. Data-Driven Decision Making: The revenue and expense analyzer facilitates data-driven decision-making by providing stakeholders with actionable insights derived from advanced machine learning models and comprehensive data analysis.

2. Improved Forecasting Accuracy: Integration of advanced machine learning techniques such as XGBoost and ensemble methods enhances forecasting accuracy, enabling organizations to make more precise revenue and expense predictions.

3. Enhanced Strategic Planning: The analyzer's ability to integrate external data sources and provide real-time predictive analytics empowers organizations to conduct more effective strategic planning, anticipate market trends, and allocate resources efficiently.

4. Operational Efficiency: Seamless integration with financial planning and ERP systems streamlines data flow and enhances operational efficiency, enabling organizations to optimize resource allocation and improve overall financial performance.

5. Transparency and Accountability: Adoption of explainable AI techniques ensures model transparency and addresses ethical considerations, fostering trust among stakeholders and promoting accountability in decision-making processes.

6. Continuous Improvement: Establishing a framework for continuous evaluation and model refinement ensures that the analyzer remains adaptive to evolving business needs and maintains its relevance and effectiveness over time.

7. Competitive Advantage: By leveraging advanced analytics and incorporating external data sources, organizations gain a competitive advantage in the marketplace by making more informed and proactive decisions, responding quickly to changing market conditions, and capitalizing on emerging opportunities.

Chapter 8: Conclusion

The conclusion chapter encapsulates the objectives, achievements, and implications of the revenue and expense analyzer project for financial analysis. It underscores the significance of technological innovations in enhancing financial practices and proposes future avenues for research and development. By emphasizing the project's contributions, it lays the groundwork for further advancements in financial analysis technology.

8.1 Limitations

1. Interpretability Challenges: XGBoost, while highly accurate, is often considered a black-box model, meaning its predictions are not easily interpretable compared to simpler models like linear regression. This lack of interpretability could hinder stakeholders' understanding of the factors driving revenue and expense predictions, potentially limiting the dashboard's effectiveness in guiding decision-making.
2. Scalability Concerns: XGBoost models, especially when trained on large datasets, can be computationally intensive and may require significant computational resources for deployment and real-time inference. This scalability concern could limit the practicality of integrating XGBoost predictions directly into the dashboard, particularly in environments with limited computational infrastructure or processing capabilities.
3. Data Quality and Preprocessing: The performance of machine learning models, including XGBoost, heavily depends on the quality and preprocessing of the input data. Any inconsistencies, biases, or missing values in the data used for training and prediction can adversely affect the accuracy and reliability of the revenue and expense predictions. Therefore, it's essential to thoroughly assess and address data quality issues to ensure the robustness of the predictive models integrated into the dashboard.
4. Lack of Causal Inference: While XGBoost excels in predictive accuracy, it may not inherently capture causal relationships between variables. This limitation could hinder the dashboard's ability to provide actionable insights for strategic decision-making, as it may not fully elucidate the cause-and-effect dynamics driving changes in revenue and expenses. Incorporating causal inference techniques alongside predictive modeling could enhance the dashboard's ability to offer actionable insights for resource allocation and strategic planning.
5. Model Maintenance and Updating: Machine learning models, including XGBoost, require regular maintenance and updating to remain relevant and accurate over time. Changes in market dynamics, consumer behavior, or internal business processes may necessitate retraining or fine-tuning the models to ensure their continued effectiveness. Failure to adequately address model maintenance and

updating could lead to degradation in predictive performance over time, undermining the utility of the revenue and expense analyzer dashboard.

8.2 Conclusion

In conclusion, the development of the full-stack Finance dashboard application presents an innovative solution for businesses to monitor their financial performance effectively. With a visually appealing and user-friendly interface incorporating various charts and tables, the application facilitates easy interpretation and analysis of critical financial data. Moreover, by integrating real machine learning capabilities, such as regression analysis, the dashboard empowers businesses with accurate revenue predictions, fostering data-driven decision-making and strategic planning. The application's focus on simplicity and accessibility, coupled with rigorous testing and adherence to high standards of usability and data accuracy, ensures its reliability, security, and performance, making it a powerful tool for businesses aiming to thrive in the competitive financial landscape.

Overall, the Finance dashboard serves as a valuable asset for business professionals, providing deep insights into their company's financial health and enabling informed decision-making for sustainable growth. As the dashboard evolves through continuous improvement and user feedback, it is poised to become an indispensable ally for businesses in their pursuit of financial success, offering a robust platform that adapts to the evolving needs of modern enterprises.

8.3 Future Scope

- Integration of Advanced Machine Learning Models: Incorporating deep learning and ensemble methods alongside XGBoost enhances predictive accuracy and uncovers complex data patterns.
- Enhanced Data Visualization and Interpretability: Interactive visualizations and feature importance analysis improve dashboard usability, empowering stakeholders to make informed decisions..
- Integration of External Data Sources: Incorporating market trends and customer sentiment data enriches analysis, enabling proactive decision-making.
- Real-Time Predictive Analytics: Implementing real-time analytics enables continuous monitoring of revenue and expenses, facilitating timely decision-making.
- Integration with Financial Planning and ERP Systems: Seamless integration streamlines data flow, enhances consistency, and enables holistic decision-making.
- Adoption of Explainable AI and Ethical Considerations: Using explainable AI techniques ensures model transparency and addresses concerns related to fairness and ethical use of data.

References

[1] V. K. Singh, L. Freeman, B. Lepri, and A. S. Pentland, “Predicting Spending Behavior Using Socio-mobile Features,” IEEE Xplore, Sep. 01, 2013. https://ieeexplore.ieee.org/abstract/document/6693330 (accessed Feb. 08, 2022).

[2] Vipul, P. V. Kumar, D. Dimri, and M. Pathak, “Expenditure Predicting using Machine Learning,” International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 1, pp. 2169–2172, Nov. 2019, doi: 10.35940/ijitee.a4720.119119.

[3] M. Nadj, A. Maedche, and C. Schieder, “The effect of interactive analytical dashboard features on situation awareness and task performance,” Decision Support Systems, vol. 135, p. 113322, May 2020, doi: https://doi.org/10.1016/j.dss.2020.113322.

[4] Kureljusic, M. and Karger, E. (2024), "Forecasting in financial accounting with artificial intelligence – A systematic literature review and future research agenda", Journal of Applied Accounting Research, Vol. 25 No. 1, pp. 81-104. https://doi.org/10.1108/JAAR-06-2022-0146

[5] Ioana-Alexandra, Bradea & Claudia Diana, Sabau-Popa & Bolos, Marcel. (2014). “USING DASHBOARDS IN BUSINESS ANALYSIS”, Annals of the University of Oradea. Economic Sciences. Tom XXIII.

[6] H. Wasserbacher and M. Spindler, “Machine learning for financial forecasting, planning and analysis: recent developments and pitfalls,” Digital Finance, vol. 4, Dec. 2021, doi: https://doi.org/10.1007/s42521-021-00046-2.

Appendix

1] Paper I details :-

Revenue and Expense Analyzer

Priti Joshi
Computer Department Vivekanand Education Society's Institute of Technology, Mumbai, India.
priti.joshi@ves.ac.in

Yash Kewlani
Department of Computer Engineering VES Institute of Technology, Mumbai, India
2020.yash.kewlani@ves.ac.in

Neeraj Chawla
Department of Computer Engineering VES Institute of Technology, Mumbai, India
2020.neeraj.chawla@ves.ac.in

Atharva Mahalle
Department of Computer Engineering VES Institute of Technology, Mumbai, India
2020.atharva.mahalle@ves.ac.in

Mithil Wasrani
Department of Computer Engineering VES Institute of Technology, Mumbai, India
2020.mithil.wasrani@ves.ac.in

Abstract: **The "Revenue and Expense Analyzer" project addresses the pressing need for effective financial performance monitoring and management in the dynamic business and finance landscape. The intricacies of financial data, often buried within complex spreadsheets and reports, pose a significant challenge for professionals. This project offers a comprehensive solution by introducing an innovative dashboard adorned with essential Key Performance Indicators (KPIs) and critical financial insights, designed to provide clarity and understanding regardless of the user's technical background. The project leverages a diverse array of visualization tools, including area charts, line charts, bar charts, pie charts, and scatter plots, to transform complex financial data into easily comprehensible visuals, simplifying the decision-making process.**

What truly distinguishes the "Revenue and Expense Analyzer" is its incorporation of advanced machine learning techniques for predictive analysis, providing capabilities that enable businesses to accurately forecast revenue trends. With this project, data-driven decision-making becomes more accessible and reliable, empowering organizations to make informed decisions and plan confidently for the future.

Keywords: Key Performance Indicator, Decision Support System, Revenue, Expenses, Profit, Regression

Introduction

In the dynamic landscape of business and finance, effectively monitoring and managing a company's financial performance is paramount for success. The intricacies of financial data, often buried within complex spreadsheets and reports, can pose a significant challenge for business professionals. Additionally, the absence of user-friendly tools and accessible Key Performance Indicators (KPIs) can make it challenging to gain a clear understanding of a business's financial health. To address these challenges, we present the "Revenue and Expense Analyzer."

The Revenue and Expense Analyzer is a comprehensive solution designed to empower businesses with the tools they need to take control of their financial performance. In a world where data-driven decision-making is increasingly vital, this innovative dashboard offers an engaging, visually captivating homepage adorned with essential KPIs and critical financial insights. Its user-friendly approach is engineered to provide clarity and understanding, regardless of the user's technical background.

At the core of this groundbreaking innovation lies a diverse array of visualization tools, from versatile area charts to insightful line charts, informative bar charts, revealing pie charts, and informative scatter plots. These tools perform the invaluable task of transforming complex financial data into easily comprehensible visuals, eliminating the need for intricate data interpretation and simplifying the decision-making process.

What sets the Revenue and Expense Analyzer apart is its integration of advanced machine learning techniques, particularly regression analysis. This application goes a step further to provide predictive capabilities that enable businesses to foresee and prepare for revenue trends accurately. The invaluable ability to make informed decisions and plan for the future with confidence is now at your fingertips.

Problem Statement

In today's business landscape, the effective monitoring of financial performance remains a significant challenge. Financial data complexity, a lack of user-friendly tools, and the absence of a comprehensive Finance dashboard impede decision-making. Businesses struggle to access vital Key Performance Indicators (KPIs), interpret financial information, and make reliable revenue predictions. The "Revenue and Expense Analyzer" seeks to address these issues by offering a unified solution that simplifies financial monitoring, enhances data accessibility, provides clear visualization tools, and incorporates machine learning for accurate revenue forecasting. This comprehensive approach empowers businesses to make informed decisions and confidently plan for a prosperous financial future.

Literature Survey

1. The paper investigates the utilization of socio-mobile data obtained from mobile phone-based social interactions for predicting spending behavior within couples. It highlights the effectiveness of these data-driven predictions compared to traditional personality-based features, emphasizing their potential importance for economists, marketers, and policymakers. However, the study's impact is tempered by its reliance on a limited and homogeneous sample size, which restricts the generalizability of its findings to more diverse populations.[1]

2. In their project, the team employs machine learning techniques, particularly linear regression, to forecast future expenses. By analyzing past bank transactions, the project aims to predict expenditure trends, with potential applications extending to stock market predictions for financial management purposes. However, the absence of security measures poses a significant concern, especially considering that the system involves sensitive bank statements. Implementing robust security measures is imperative to safeguard the confidentiality and integrity of the data.[2]

3. From the HiLT Lab at the University of North Texas, the study examines the influence of interactive analytical dashboard features, particularly what-if analyses, on situation awareness (SA) and task performance within operational decision support systems (DSS). The study sheds light on the potential challenges associated with maintaining SA while utilizing these features. Moreover, the difficulty in constructing comprehensive optimization models for complex real-world problems is addressed. Additionally, concerns regarding the validity of eye-tracking data as a predictor of errors are discussed.[3]

4. The study delves into the realm of financial accounting forecasting with artificial intelligence, exploring the utilization of interactive analytical dashboard features, particularly what-if analyses, within operational decision support systems (DSS). It examines the impact of these features on situation awareness (SA) and task performance, while also addressing the challenges associated with maintaining SA. Furthermore, the paper discusses the difficulty in constructing comprehensive optimization models for complex real-world problems, along with concerns regarding the validity of eye-tracking data as a predictor of errors.[4]

5. The study explores the application of artificial intelligence in financial accounting forecasting, focusing on the role of dashboards in business analysis. It emphasizes their integration with processes and strategies while addressing challenges such as data relevance amidst information overload. However, the paper may lack depth in discussing drawbacks associated with dashboard use and may not encompass the full range of industries or scenarios, possibly limiting its generalizability.[5]

6. The paper titled "Machine Learning for Financial Forecasting, Planning, and Analysis: Recent Developments and Pitfalls" introduces machine learning techniques for financial forecasting and planning, with a particular focus on causal inference and the "double machine learning framework." However, the study highlights the lack of real-world validation for this framework in the context of financial planning.[6]

Comparison of Existing Systems

Comparison	Tableau	Power BI	Zoho Analytics	Looker	Google Data Studio	Domo
Features	Robust data visualization, interactive dashboards	Data analysis, visualization, AI-driven insights	Business intelligence, reporting, data blending	Data exploration, modeling, embedded analytics	Data visualization, customizable dashboards	Business intelligence, data visualization
Pricing	Starts at $70/user/month	Starts at $9.99/user/month	Starts at $22/user/month	Custom pricing	Free	Custom pricing
Integration	Integrates with various data sources and databases	Integrates with Microsoft and third-party services	Integrates with Zoho's suite of business applications	Integrates with various data warehouses and services	Integrates with Google products and third-party data sources	Integrates with hundreds of cloud-based applications
Ease of Use	Intuitive interface, easy to create visualizations	User-friendly interface, drag-and-drop functionality	Easy-to-use interface, customizable dashboards	Intuitive interface, easy to navigate	Simple interface, easy to connect data sources	User-friendly interface, intuitive data visualization
Scalability	Suitable for small to large enterprises	Suitable for small to large enterprises	Suitable for small to medium-sized businesses	Suitable for small to large enterprises	Suitable for small to medium-sized businesses	Suitable for small to large enterprises
Customization	Highly customizable visualizations, extensive customization options	Customizable dashboards, adaptable to specific needs	Customizable reports and dashboards, flexible customization options	Customizable data modeling, tailored to business requirements	Customizable dashboards, flexible design options	Highly customizable visualizations, adaptable to specific business needs
Collaboration	Allows collaboration through shared workbooks, server-based sharing	Collaboration features like shared datasets, reports, and apps	Collaboration features for team sharing and collaboration	Collaborative workspace for sharing insights and data exploration	Collaboration features for sharing reports and dashboards	Collaboration features for team-based analysis and decision-making
Security	Role-based access control, data encryption	Role-based access control, data encryption	Role-based access control, data encryption	Role-based access control, data encryption	Role-based access control, data encryption	Role-based access control,

Table No 2: Comparison of Existing Systems

Proposed Model

System Architecture:

The system architecture for the "Revenue and Expense Analyzer" is designed to provide a robust and scalable platform for financial analysis. It consists of the following components:

- **User Interface Layer:** This layer includes the user interface components for the

Finance dashboard. It encompasses the dashboard layout, navigation menus, and design elements to ensure a user-friendly and intuitive experience.

- **Data Collection and Integration:** Financial data is collected from various sources, such as accounting software, databases, and external APIs. This data is then integrated into the system using data integration workflows.

- **Data Processing Engine:** A data processing engine is responsible for cleaning, transforming, and preparing the financial data for analysis. It ensures data quality and consistency.

- **Machine Learning Module:** The machine learning component employs regression analysis techniques to predict revenue trends. It utilizes historical financial data to generate regression models for revenue forecasting.

- **Data Visualization Tools:** The system incorporates a range of data visualization tools, including area charts, line charts, bar charts, pie charts, and scatter plots. These tools are used to create visually engaging representations of financial data for easy interpretation.

User Interface Design:
The user interface design of the Finance dashboard is focused on providing an intuitive and visually appealing experience. It includes:

- **Dashboard Layout:** The dashboard layout is designed to present key financial insights and KPIs prominently on the landing page. Users can quickly access essential information without the need for complex navigation.

- **Navigation:** A user-friendly navigation system enables users to explore different sections of the dashboard seamlessly. Clear menus and links facilitate easy access to various features.

- **Design Elements:** The design incorporates a modern and clean aesthetic, making use of color-coding, clear typography, and interactive elements for a pleasant user experience.

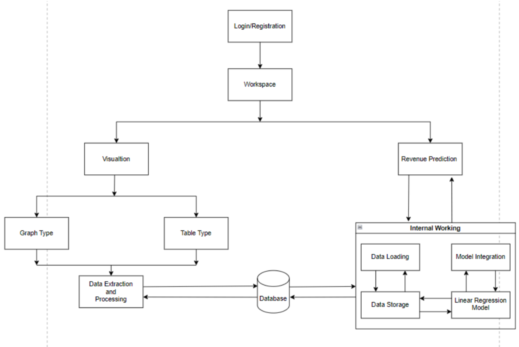
Data Flow and Integration:
Financial data is collected from various sources, including accounting systems and external data providers. The data integration process involves data extraction, transformation, and loading (ETL) workflows to ensure data consistency and quality.

Machine Learning Components:
The heart of the system is the machine learning module, which leverages regression analysis to predict revenue trends. It uses historical financial data, such as monthly revenue figures, to build regression models. These models can generate insights into future revenue predictions based on past performance.

Data Visualization Tools:
The system utilizes a diverse range of data visualization tools, including area charts, line charts, bar charts, pie charts, and scatter plots. These tools provide users with visually engaging representations of financial data, enabling easy interpretation and analysis.

Proposed Design

4.1 Block Diagram of the proposed system

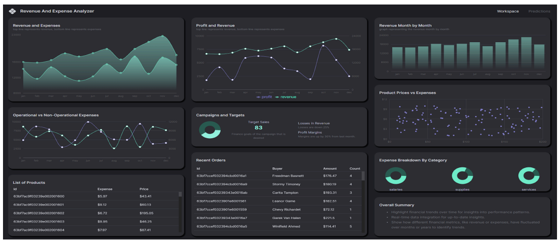


The financial data analysis platform described encompasses a range of components designed to streamline the visualization and analysis of financial performance. At its core, the Visualization component offers an intuitive dashboard interface that presents data in a meaningful and interactive manner. Users can explore various visualization options, including line charts, bar charts, and pie charts, to gain insights into financial trends and patterns. Additionally, the Table Type component caters to users who prefer a structured, tabular representation of data, providing a detailed and organized view of financial metrics.

A key aspect of the platform is its predictive capabilities, facilitated by the Revenue Prediction component. Leveraging machine learning techniques such as linear regression, this component forecasts future revenue trends based on historical data. These predictions are instrumental in proactive planning and strategic decision-making, empowering users to anticipate and respond to changes in financial performance effectively. Moreover, the platform's Data Extraction and Processing component ensures that data from various sources is efficiently extracted, cleaned, and structured to maintain data quality and consistency.

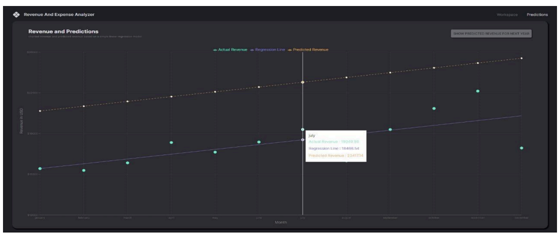
The platform's architecture also includes essential components such as Internal Working, Data Loading, and Data Storage, which collectively manage the flow, loading, and storage of data within the system. Furthermore, the integration of the Linear Regression Model and Databases facilitates accurate revenue forecasting and efficient data storage and retrieval, respectively. Model Integration ensures that predictive insights seamlessly integrate with visual displays, enabling users to assess actual revenue alongside predicted trends. Overall, these components work in tandem to simplify financial data analysis, making it more accessible and actionable for users with diverse preferences and technical backgrounds.

Implementation:



On the main window, we've designed a dashboard featuring various types of visual representations. In the top-left corner, there's an area chart illustrating revenue and expenses, followed by a line chart displaying profit and revenue trends. On the top-right side, there's a bar chart depicting revenue on a monthly basis. By hovering over these visualizations, users can access additional details for each data point.
In the middle row, there's a line chart delineating operational and non-operational expenses for the business. Positioned in the center is a donut chart offering insights into campaign performance versus target sales, including profit and loss margins. Additionally, there's a scatter plot presenting the relationship between product prices and expenses.

Towards the bottom-left section, there's a list of products corresponding to the visualizations above. Adjacent to it, there's a list of the 50 most recent orders. Finally, there's a donut chart displaying expenses categorized into various segments such as salaries, supplies, service, and maintenance.



In the predictions window, our linear regression model serves a crucial role. This model analyzes the relationship between time (months) on the X-axis and revenue in USD on the Y-axis. By plotting a line among the data points, it provides insight into the overall revenue trend. Additionally, based on this trend, the model forecasts revenue for the upcoming year.

Furthermore, when interacting with the graph by hovering over specific points, users can access more details. This includes not only current revenue figures but also predictions for the next year. This comparison between the current year's revenue and the predicted revenue for the next year helps users understand the trajectory of revenue growth and make informed decisions accordingly.

Analysis of Machine Learning Models

In our research, various machine learning models were implemented and evaluated to ascertain their performance in the given context. Among these models, XG Boost emerged as the most effective, surpassing both Random Forest and logistic regression models in terms of performance metrics. This observation underscores the superiority of XG Boost within the scope of our study.

Results:

The performance comparison between XGBoost and Random Forest models revealed notable distinctions favoring XG Boost.

Root Mean Squared Error (RMSE): XG Boost exhibited an RMSE of approximately **1023.81**, indicating an average prediction deviation of about **1023.81** units from the actual values. This suggests a superior predictive accuracy compared to Random Forest.

Mean Absolute Error (MAE): With an MAE of approximately **714.55**, XG Boost showcased a closer alignment between predicted and actual values compared to Random Forest. This implies a smaller average absolute difference of around **714.55** units, further highlighting its superior performance.

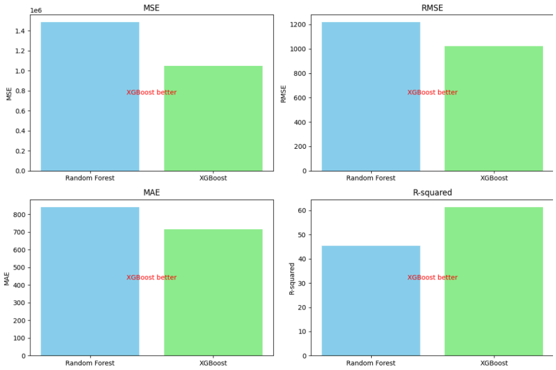
R-squared (R2): The R-squared value of approximately **0.614** for XG Boost signifies that approximately **61.43%** of the variance in the dependent variable can be explained by the independent variables integrated into the model. This reflects a higher level of explanatory power compared to Random Forest.

These findings emphasize the robustness and efficacy of the XGBoost model in our research context, suggesting its potential for enhanced predictive accuracy and model generalization compared to alternative approaches like Random Forest.

```
Implementation Details:

Mean Squared Error (MSE): 1048186.18
Root Mean Squared Error (RMSE): 1023.81
Mean Absolute Error (MAE): 714.55
R-squared (R2): 61.4%
```

Comparison between Random Forest and XGBoost



Conclusion:

In conclusion, the development of the full-stack Finance dashboard application presents an exciting opportunity to revolutionize the way business persons monitor their company's financial performance.A visually appealing and user-friendly interface is offered by the application through the incorporation of various types of charts and tables, enabling easy interpretation and analysis of critical financial data.

A focus on simplicity and accessibility is maintained by the Finance dashboard, aiming to bridge the gap between complex financial data and intuitive insights. By modern technologies adoption and rigorous testing adherence, a powerful tool is promised for businesses seeking to thrive in the competitive financial landscape.

Ultimately, a valuable asset is represented by the Finance dashboard, providing business persons with the means to gain deep insights into their company's financial health and make informed decisions that drive sustainable growth. As continuous improvement and user feedback drive the dashboard's evolution, it is poised to become an indispensable ally for businesses in their pursuit of financial success.

Empowerment of businesses with accurate revenue predictions is achieved through the integration of real machine learning capabilities, using regression analysis, fostering data-driven decision-making and strategic planning. The highest standards of usability and data accuracy are ensured by the proposed evaluation measures, guaranteeing the application's reliability, security, and performance.

References

1. V. K. Singh, L. Freeman, B. Lepri, and A. S. Pentland, “Predicting Spending Behavior Using Socio-mobile Features,” IEEE Xplore, Sep. 01, 2013. https://ieeexplore.ieee.org/abstract/document/6693330 (accessed Feb. 08, 2022).

2. Vipul, P. V. Kumar, D. Dimri, and M. Pathak, “Expenditure Predicting using Machine Learning,” International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 1, pp. 2169–2172, Nov. 2019, doi: 10.35940/ijitee.a4720.119119.

3. M. Nadj, A. Maedche, and C. Schieder, “The effect of interactive analytical dashboard features on situation awareness and task performance,” Decision Support Systems, vol. 135, p. 113322, May 2020, doi: https://doi.org/10.1016/j.dss.2020.113322.

4. Kureljusic, M. and Karger, E. (2024), "Forecasting in financial accounting with artificial intelligence – A systematic literature review and future research agenda", Journal of Applied Accounting Research, Vol. 25 No. 1, pp. 81-104. https://doi.org/10.1108/JAAR-06-2022-0146

5. Ioana-Alexandra, Bradea & Claudia Diana, Sabau-Popa & Bolos, Marcel. (2014). “USING DASHBOARDS IN BUSINESS ANALYSIS”, Annals of the University of Oradea. Economic Sciences. Tom XXIII.

6. H. Wasserbacher and M. Spindler, "Machine learning for financial forecasting, planning and analysis: recent developments and pitfalls," Digit Finance, vol. 4, pp. 63–88, 2022. doi: 10.1007/s42521-021-00046-2.

b. Plagiarism Report

Revenue and Expense Analyzer

ORIGINALITY REPORT

7%	6%	6%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Priotosh Mondal, Aditi Bhatia, Roshini Panjwani, Shrey Panchamia, Indu Dokare. "Image-based Classification of Skin Cancer using Convolution Neural Network", 2023 3rd International Conference on Intelligent Technologies (CONIT), 2023 Publication	2%

c. Project review sheet :

Project review sheet 1:

Inhouse/ Industry _Innovation/Research:

Sustainable Goal:

Class: D17 A/B/C

Project Evaluation Sheet 2023 - 24

Group No.: 52

Title of Project: Revenue and Expense Analyzer

Group Members: Yash Kewlani (D17B, 34), Neeraj Chawla (D17B/12), Atharva Mahalle (D17B/40), Mithil Wasrani (D17B/72)

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
4	3	3	2	4	2	2	2	1	1	2	2	2	2	3	35

Comments: Project work moving very slow. Need to implement many features, which is pending now. Heavy work on dataset and forecasting is remaining.

Geocy Shety

Name & Signature Reviewer1

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
4	3	3	2	4	2	2	2	1	1	2	2	2	2	3	35

Comments: ✓

Date: 10th february, 2024

Priti Joshi

Name & Signature Reviewer 2

Project review sheet 2:

Inhouse/ Industry _Innovation/Research:

Sustainable Goal:

Class: D17 A/B/C

Group No.: 52

Project Evaluation Sheet 2023 - 24

Title of Project: Revenue and Expense Analyzer

Group Members: YASH Kewlani (D17B/34), Neeraj Chawla (D17B/12), Mithil Wasrani [D17B, 72] Atharva D. Mahalle [D17B-40]

Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (2)	Applied Engg&Mgmt principles (3)	Life - long learning (3)	Professional Skills (3)	Innovative Approach (3)	Research Paper (5)	Total Marks (50)
4	4	4	2	5	2	2	2	2	2	2	2	2	2	2	39

Comments: Need to improve data set. Compare the product with existing product in market.

Name & Signature Reviewer1

Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (2)	Applied Engg&Mgmt principles (3)	Life - long learning (3)	Professional Skills (3)	Innovative Approach (3)	Research Paper (5)	Total Marks (50)
4	4	4	2	5	2	2	2	2	2	2	2	2	2	2	39

Comments: Data set Needs to be improve

Date: 9th March, 2024

Name & Signature Reviewer 2