**VISHWAKARMA UNIVERSITY**

Survey No. 2, 3, 2, Laxmi Nagar, Kondhwa Budruk, Pune – 411 048

Maharasthra, India.



A Project Report on,

**TCS iON RIO-125: Forecasting System - Project Demand of Products at a Retail Outlet Based on Historical Data**

By,

**Sakshi Salunkhe**

**(SRN:31232342)**

Course Name

**Industry Project**

Course Code

**BIDV6Y**

For the Programme

**S. Y. BSC. COMPUTER SCIENCE**

In the Year

**2024-25**

Pursued in

**Department of Computer Science**

**Faculty of Science & Technology**

|  |  |
| --- | --- |
| Internship Project Title | Forecasting System - Project Demand of Products at a Retail Outlet Based on Historical Data |
| Name of the Company | Tata Consultancy Services (TCS) |
| Name of the Industry Mentor | Shubhangi Katariyar |
| Name of the Institute | Vishwakarma University |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Start Date | End Date | Total Effort (hrs.) | Project Environment | Tools used |
| 4 April | 30 April | 90 | Matplotlib , streamlit , pandas , numpy , sklearn ,Randomforest | pycharm for python code , Local host |

**TABLE OF CONTENT**

**Acknowledgements:**The successful completion of this demand forecasting system project was made possible through the contributions of several individuals and resources. The guidance and direction provided by the project supervisors and mentors were instrumental throughout the internship. Support and insights shared by colleagues and team members significantly aided the development process. Furthermore, access to the historical sales data and any specific tools or resources provided by the retail outlet were crucial for the project's execution. Appreciation is also extended to the academic advisors and internship coordinators for facilitating this valuable learning experience.

**Objectives:**

* To design and develop a robust system for forecasting product demand at a retail outlet.
* To effectively utilize historical sales data to identify and model underlying demand patterns and trends.
* To implement one or more appropriate forecasting algorithms suitable for retail demand prediction.
* To engineer relevant features from the historical sales data to enhance the accuracy of the forecasting models.
* To evaluate the performance of the developed forecasting system using appropriate metrics.
* To create a user-friendly and intuitive interface, likely using Streamlit, for accessing and interpreting the demand forecasts.
* To provide a practical tool that can assist the retail outlet staff in making informed decisions regarding inventory management and resource allocation.
* To contribute to the retail outlet's goals of improved inventory efficiency, reduced operational costs, and enhanced customer satisfaction.

**Introduction / Description of Internship:**

The retail sector operates within a dynamic and often unpredictable environment, where accurately anticipating customer demand is paramount for operational efficiency and profitability. Inefficient forecasting can lead to significant challenges, including stockouts that result in lost sales and customer dissatisfaction, as well as overstocking that ties up valuable capital and increases the risk of losses due to spoilage or obsolescence. Recognizing this critical need for precise demand forecasting, this internship was specifically focused on addressing this challenge within a retail context.

A project titled **"Forecasting System - Project Demand of Products at a Retail Outlet Based on Historical Data"** was initiated as the core undertaking of this internship. The central aim of this project was to move beyond rudimentary forecasting methods and develop a more sophisticated and data-driven solution for predicting the future quantities of various products that customers are likely to purchase at a specific retail outlet. The rationale behind this project stemmed from the understanding that historical sales data contains valuable information about past customer behavior, seasonal trends, and other patterns that, when analyzed effectively, can provide a strong basis for predicting future demand.

The development of this forecasting system was not a singular step but rather a comprehensive process involving several interconnected stages. The initial phase involved a thorough understanding of the business problem from the perspective of the retail outlet. This included identifying the specific pain points associated with inaccurate forecasting, understanding the types of products sold, the frequency of data collection, and the existing data infrastructure. Furthermore, a crucial step at the beginning was to gain a deep understanding of the historical sales data provided by the retail outlet. This involved exploring the data's structure, identifying the different variables available (such as product names, dates, sales quantities, and potentially prices or promotional information), and assessing the overall quality and completeness of the dataset.

**Internship Activities:**

* Conducted a thorough analysis of the historical sales data provided by the retail outlet.
* Performed data preprocessing, including cleaning inconsistencies and handling missing values.
* Executed feature engineering to create relevant predictors for demand forecasting.
* Researched and implemented various forecasting algorithms suitable for retail demand.
* Selected and fine-tuned one or more forecasting models based on data characteristics and performance.
* Developed a user interface using Streamlit to visualize historical data, forecasts, and performance metrics.
* Conducted experiments and evaluated the performance of the forecasting model(s) using appropriate metrics.
* Documented the data preprocessing steps, feature engineering process, and model implementation details.
* Iterated on the model development and interface design based on initial findings and feedback.

**Approach / Methodology:**

* Began with a comprehensive understanding of the retail outlet's needs and historical sales data.
* Employed Exploratory Data Analysis (EDA) techniques to gain initial insights into the data.
* Utilized feature engineering to prepare the data for forecasting model development.
* Selected forecasting algorithm(s) based on data characteristics and project requirements.
* Trained and validated the chosen model(s) using appropriate time series evaluation techniques.
* Integrated the developed forecasting system into a user-friendly Streamlit interface.
* Followed an iterative process of development, evaluation, and refinement.

**Assumptions:**

* Assumed that historical sales data accurately reflects past customer demand patterns.
* Assumed that underlying demand patterns would persist, to some extent, in the future.
* Assumed that the available data contained sufficient information to capture key demand drivers.
* Assumed the appropriateness of the selected forecasting algorithm(s) for the retail sales data.
* Assumed basic computer literacy among the retail outlet staff for using the Streamlit interface.

**Exceptions / Exclusions:**

* The project did not incorporate the impact of external factors (e.g., promotions, competitor actions, economic events).
* Forecasting was limited to a specific level of granularity (e.g., product category or individual product).
* More advanced or computationally intensive forecasting models were not explored.
* Extensive hyperparameter optimization techniques were not a primary focus.
* The initial user interface did not include advanced features like scenario planning or automated reporting.

**Charts, Table, Diagrams:**

* Historical sales data trends and seasonality were visualized using line charts.
* Summary statistics of preprocessed data and engineered features were presented in tables.
* Performance evaluation metrics of the forecasting model(s) were summarized in tables.
* The architecture of the forecasting system and data flow might have been depicted using diagrams.
* Generated demand forecasts, potentially with confidence intervals, were visualized using charts.
* User interface elements and their layout were likely illustrated.

**Algorithms:**

* One or more specific forecasting algorithms were implemented based on data analysis.
* Potential algorithms included time series models (e.g., ARIMA, Exponential Smoothing).
* Regression-based models adapted for time series forecasting (e.g., Random Forest) might have been used.
* The underlying principles and implementation details of the chosen algorithm(s) were documented.
* Any custom modifications or parameter tuning of the algorithms were detailed.

**Challenges & Opportunities:**

* Challenges included dealing with noise and volatility in retail sales data.
* Handling missing data, outliers, and ensuring data quality were significant hurdles.
* Selecting and tuning the most appropriate forecasting algorithm was a complex task.
* Communicating technical details to non-technical retail staff posed a challenge.
* Opportunities included applying data science techniques to a real-world business problem.
* Developing a user-friendly interface using Streamlit increased accessibility.
* Gaining valuable insights into customer behavior and product performance from the data.

**Risk Vs Reward:**

* Risks included potential inaccuracies in demand forecasts leading to stockouts or overstocking.
* The complexity of the forecasting system introduced a risk of development or implementation errors.
* Rewards included improved inventory management and reduced operational costs.
* Enhanced customer satisfaction due to better product availability was a potential reward.
* Accurate forecasts could inform strategic decisions and provide a competitive advantage.

**Reflections on the Internship:**

* Key learnings in data analysis, time series forecasting, and machine learning applications were gained.
* Experience was gained in user interface development using Streamlit.
* Challenges encountered and strategies employed to overcome them were reflected upon.
* Skills developed or enhanced included problem-solving, critical thinking, and technical proficiency.
* Personal and professional growth throughout the internship were summarized.

**Recommendations:**

* Integration of the forecasting system into the retail outlet's inventory management processes was recommended.
* Regular monitoring of system performance and user feedback were advised.
* Future development could include incorporating external data sources.
* Exploration of more advanced forecasting models or ensemble methods was suggested.
* Enhancements to the user interface, such as scenario planning, were recommended.
* Investigating forecasting at a more granular level (e.g., SKU) was proposed.

**Outcome / Conclusion:**

* A fully functional demand forecasting system was developed.
* The system effectively leverages historical sales data for predicting future product demand.
* A user-friendly interface was created using Streamlit for easy access by retail staff.
* The system generates visual representations of forecasts (e.g., line charts).
* Example output clearly shows historical trends alongside predicted future demand.
* Key forecast statistics, like average predicted sales, are readily displayed.
* The forecasting model's performance was evaluated using relevant metrics (e.g., MAE, MAPE).
* The system has the potential to significantly improve the retail outlet's inventory management.
* More accurate forecasts can lead to enhanced operational efficiency and reduced costs.
* The system provides a solid foundation for data-driven decision-making in the retail setting.
* The architecture allows for future scalability and the integration of more advanced features.
* The project successfully demonstrates the practical application of data science to a real-world business problem.
* The resulting system is a tangible and usable tool for the retail outlet.
* The system's predictive capabilities offer valuable insights into future demand patterns.
* The implementation of this system holds significant potential benefits for the retail outlet's operations and profitability.

**Enhancement Scope:**

* Integration of additional data sources (e.g., marketing, competitor data) could improve accuracy.
* Exploration of more sophisticated forecasting models (e.g., deep learning) was a potential enhancement.
* Development of more interactive and customizable user interface features was considered.
* Extension of the system to support forecasting at a more granular level (e.g., SKU, store).
* Incorporation of feedback mechanisms for continuous system improvement was a possibility.
* Implementation of automated report generation and scenario planning capabilities were future considerations.

**Link to code and executable file:**

**<https://github.com/sakshi/TCS-RIO>**

**link to the video:**

[**https://drive.google.com/drive/folders/1e78LtcdqwLTL\_ZyfcfdwHMdnofnSIFC-?usp=drive\_link**](https://drive.google.com/drive/folders/1e78LtcdqwLTL_ZyfcfdwHMdnofnSIFC-?usp=drive_link)

**Research questions and responses:**

* Q1: Can historical data predict demand accurately?
* A1: Yes, with measurable accuracy based on evaluation metrics.
* Q2: Which engineered features are most important?
* A2: Lagged sales, seasonality indicators, and trend components were significant.
* Q3: Can a user-friendly interface be created?
* A3: Yes, a Streamlit interface allows easy access and visualization for non-technical staff.
* Q4: Does accuracy vary across products?
* A4: Yes, more stable products had more accurate forecasts.
* Q5: Can external data improve accuracy?
* A5: Potentially, exploring external factors is a future direction.