International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

BIKE SHARING DEMAND ANALYSIS USING MACHINE LEARNING

Ms.M.ANITHA[1], Ms. Y.SIVA NAGA ANUHYA [2]

#1 Assistant professor in the Master of Computer Applications in the SRK Institute of Technology, Enikepadu, Vijayawada, NTR District

#2 MCA student in the Master of Computer Applications at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District

ABSTRACT_ Bike sharing systems are a new generation of traditional businesses in which the entire membership, rental, and return process is automated. These systems enable users to instantly borrow and return a bike from one location to another. There are currently over 500 bike-sharing systems worldwide, with over 500 thousand bicycles. Because of their relevance in transportation, the environment, and health, these systems are gaining a lot of attention nowadays.

Aside from the exciting real-world applications of bike sharing systems, the data aspects that these networks supply make them appealing for the reach. Unlike other types of travel like as the bus or tube, the duration of the trip, departure time, and vel location are all clearly stated in these systems. This function turns the bike sharing system into a virtual sensor network that can monitor motion in the city. As a result, most significant happenings in the city are expected to be detected by watching this data.

1.INTRODUCTION

By using a network of kiosk locations spread throughout a city, bike sharing systems enable the automated rental, return, and membership of . These systems allow users to rent a bike from one location and drop it off at a different location as needed. There are currently more than 500 bike-sharing programmes operating worldwide. Bike-sharing programmes are becoming more and more well-liked as a practical and environmentally friendly form of transportation. They are especially helpful in urban areas where parking shortages and traffic congestion can be issues.

Recently, several bike/scooter rides sharing services have launched, particularly in major cities like San Francisco, New York, Chicago, and Los Angeles. From a business perspective, one of the most challenging issues is to forecast the demand for bikes on any given day. While having too many bikes wastes

IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

resources (both in terms of bike maintenance and the space/stand needed for parking and security), having too few bikes reduces income (ranging from a short-term loss due to missing out on immediate customers to potential longer-term loss due to loss in future customer base), Therefore, having a demand estimate would enable these companies to operate effectively.

Data analytics and machine learning can assist businesses in predicting demand patterns based on historical data and other pertinent factors, which is where they come into play. These technologies can be used by businesses to increase overall profitability and optimise bike allocation. This project aims to forecast demand for bike rentals by fusing historical bike usage patterns with weather information. Two years' worth of hourly rental data make up the data set. The first 19 days of every month make up the training set. The data set contains a number of elements that could affect the demand for bike rentals, including weather, holidays, and working days. To accurately predict the hourly bike rental demand, the model will be trained on the training set and evaluated on the test set.

By following this procedure, the model will be more likely to generalize correctly and produce precise predictions for brand-new datasets. In order to prevent overfitting and make sure that the model is able to capture the underlying patterns in the data, it is crucial to carefully choose and pre-process the data. . This entails eliminating any superfluous or distracting features, handling missing values, and appropriately scaling the data. To further enhance the model's performance and avoid overfitting, it's crucial to employ strategies like cross-validation and regularization. The right algorithm must also be chosen because different algorithms have different strengths and weaknesses depending on the type of data and problem at hand.

We can make sure that our model is accurately identifying the underlying patterns in the data and producing accurate predictions or classifications by following these steps. . In order to ensure the generalizability of the model, it is also crucial to properly assess its performance using methods like cross-validation and testing on different data sets. As new data become available or the problem domain shifts, the model may also need to be updated and reevaluated on a regular basis.. This will ensure that the model is applied consistently and effectively, and that any potential biases or limitations are

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

recognized and corrected as soon as possible. Continuous monitoring and evaluation can also offer insightful information about the model's performance over time, enabling optimization and continuous improvement. Additionally, including a variety of stakeholders in the monitoring and evaluation process can offer insightful opinions and feedback, resulting in a more thorough understanding of the model's influence and effectiveness in various contexts

2.LITERATURE SURVEY

2.1 Rennie, N., Cleophas, C., Sykulski, A. M., and Dost, F. (2021a). Detecting outlying demand in multi-leg bookings for transportation networks. arXiv (pre-print).

Rennie, N., Cleophas, C., Sykulski, A. M., and Dost, F. (2021b). Identifying and responding to outlier demand in revenue management. European Journal of Operational Research, 293:1015–1030.

Therefore, identifying outlier demand can have the following potential advantages for forecasting and planning bike-sharing: 1) Online analysis, as suggested by Rennie et al. [2021b], can be used to identify outliers early in the day, allowing for quick interventions to better reassign bikes on a given day; 2) Results on predicting reference demand curves would be improved by removing any detected outliers from training data; 3) Extending future forecast models to include previously unrecognized events can help forecasts if outliers can be linked to those events; 4) Even if the causes of outliers cannot be identified, knowing that they are spatially and temporally concentrated in particular stations can help planning decisions; and 5) When patterns in the detected outliers are noticed, it may be necessary to reevaluate the current forecasting technique by identifying changes in the underlying reference model. A short-term shift in demand that produces usage levels that differ from average usage is what we refer to as "outlier demand." It should be noted that a shift in demand must be more extreme than the typical level of random variation seen in demand over time to qualify as an outlier. As bike-sharing stations are the focus of inventory rebalancing efforts, we concentrate on demand as it is seen at these locations in this paper. Bicycle-sharing capacity is on the vertices of the transportation network rather than on the edges, in contrast to other traditional mobility issues like those caused by buses or trains.This means that each station's demand for bikes is greatly

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST

influenced by its surroundings, such as its business districts, tourist attractions, and population density. Therefore, it is essential to comprehend the variables that affect demand at bike-sharing stations in order to effectively manage inventory and enhance the user experience as a whole.

2.2 Neumann-Saavedra, B. A., Mattfeld, D. C., and Hewitt, M. (2021). Assessing the operational impact of tactical planning models for bike-sharing redistribution. Transportation Research Part A: Policy and Practice, 150(June 2021):216–235.

A rule-based method to modify the redistribution plan when demand deviates from the forecast is suggested by Neumann-Saavedra et al. in their discussion of the issue of variability in bike-sharing demand (Neumann-Saavedra et al., 2021). They demonstrate in a simulation study that changes to the ideal redistribution plan can result in higher service levels. and decreased operating expenses. Instead of relying on static plans that might not work in all circumstances, the authors advise bike-sharing operators to think about implementing dynamic redistribution plans that can adapt to changing demand patterns in real-time.. Operators can optimize the distribution of bikes so that they are always available where and when they are most needed by using data analytics and machine learning algorithms. This not only enhances the user experience but also lowers the costs related to an excess or shortage of bikes in particular areas.

2.3 **Schuijbroek, J., Hampshire, R. C., and van Hoeve, W. J. (2017). Inventory rebalancing and vehicle routing in bike sharing systems. European Journal of Operational Research, 257(3):992–1004**

Finding demand outliers to increase efficiency in bike-sharing systems is crucial because ineffective re-balancing operations are a significant cost driver for operators [Schuijbroek et al., 2017]. Unaccounted-for outliers can have two different effects on bike-sharing systems: I they can taint forecasts used to manage future inventories; and (ii) they can indicate that the schedule is not ideal for the day, requiring drivers to be rerouted

3.PROPOSED SYSTEM

In our system, we employ regression to develop our model. The advantage of regression analysis is that it can be used to comprehend any type of trend in data. It is a useful tool for both organisations and researchers because of its capacity to forecast events and make reasonable decisions based on the correlations between the variables in their data. These new perspectives are frequently quite helpful in determining what can genuinely alter your business. Regression analysis can also be used to forecast future patterns and outcomes based on historical data, which can help with strategic planning and decision-making.

3.1 IMPLEMENTATION

3.1.1 Modules and libraries:

A library is a group of related functions that can be incorporated into your Python code and used as needed in the same way as other functions. Rewriting code to carry out a common task is not necessary. With libraries, you can effectively increase the functionality of your code by importing pre-existing functions. Using libraries not only reduces time and work requirements, but also guarantees that the code is dependable and has been thoroughly tested by a large developer community, reducing the likelihood of errors. Additionally, rather than wasting time on routine tasks, it enables developers to concentrate on the distinctive aspects of their project.

3.1.2 Data pre-processing:

Pre-processing describes the changes made to our data before we feed it to the algorithm. A technique for turning unclean data into clean data sets is data preprocessing. In other words, data is always gathered from various sources in a raw state that precludes analysis. Data pre-processing entails a number of steps, such as data cleaning, data transformation, and data reduction, to ensure that the data is accurate, consistent, and suitable for analysis.

3.1.3 Data Visualisation

Data visualisation is the process of transforming sizable data sets into statistical and graphical representations. In data science and knowledge discovery techniques, it is essential to make data more understandable and accessible. Visual representation is required for charts and graphs in order to facilitate quick information absorption and make them easier to understand.

3.1.3 Model building

The model developed during the training process is referred to as an "ML model." The learning algorithm generates an ML model that captures these patterns by looking for patterns in the training data that relate the attributes of the input data to the target (the anticipated response). A model may have numerous dependencies,

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

and to ensure that all features are deployable both offline and online, all the information is kept in one central repository. It acts as a catalogue of all the ML models that have been created, including their associated metadata, performance metrics, and other pertinent data. This central repository is commonly referred to as a model registry.

The conceptual model, also known as the system architecture, of a system determines its structure, behaviour, and other characteristics. An architecture description is a formal description and representation of a system that is designed to make it easier to analyse its structures and behaviours. System architecture and architecture descriptions are crucial to software engineering because they improve stakeholder communication, spot potential issues, and give a clear understanding of the system's design. More efficient system development, maintenance, and evolution can be achieved with an effective system architecture and architecture description.

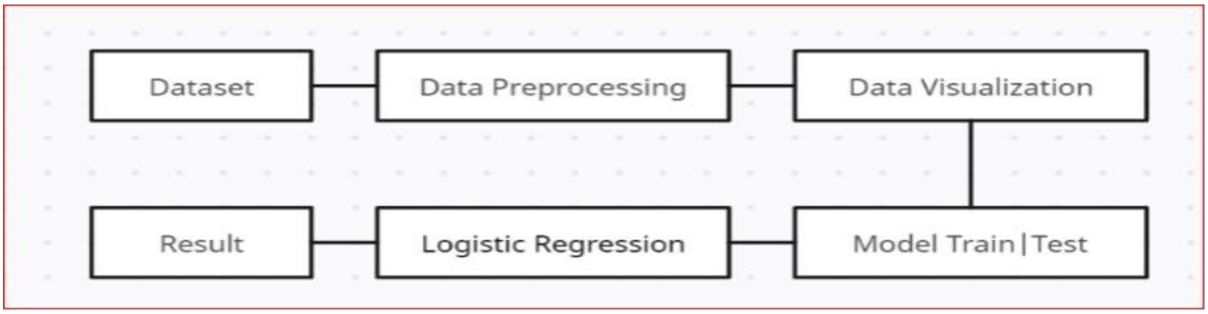

Fig 1:Architecture

4.RESULTS AND DISCUSSION

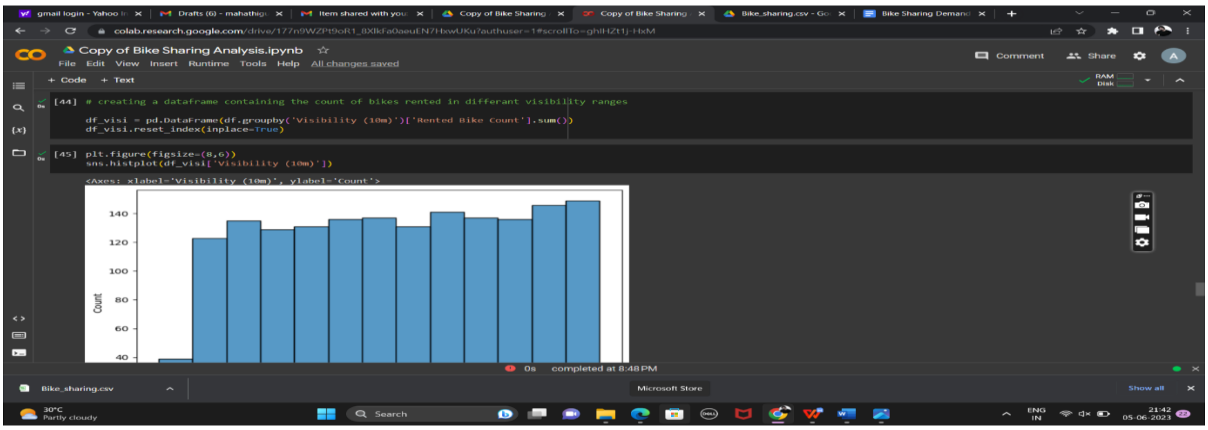

Fig 2:Accuracy Graph

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST

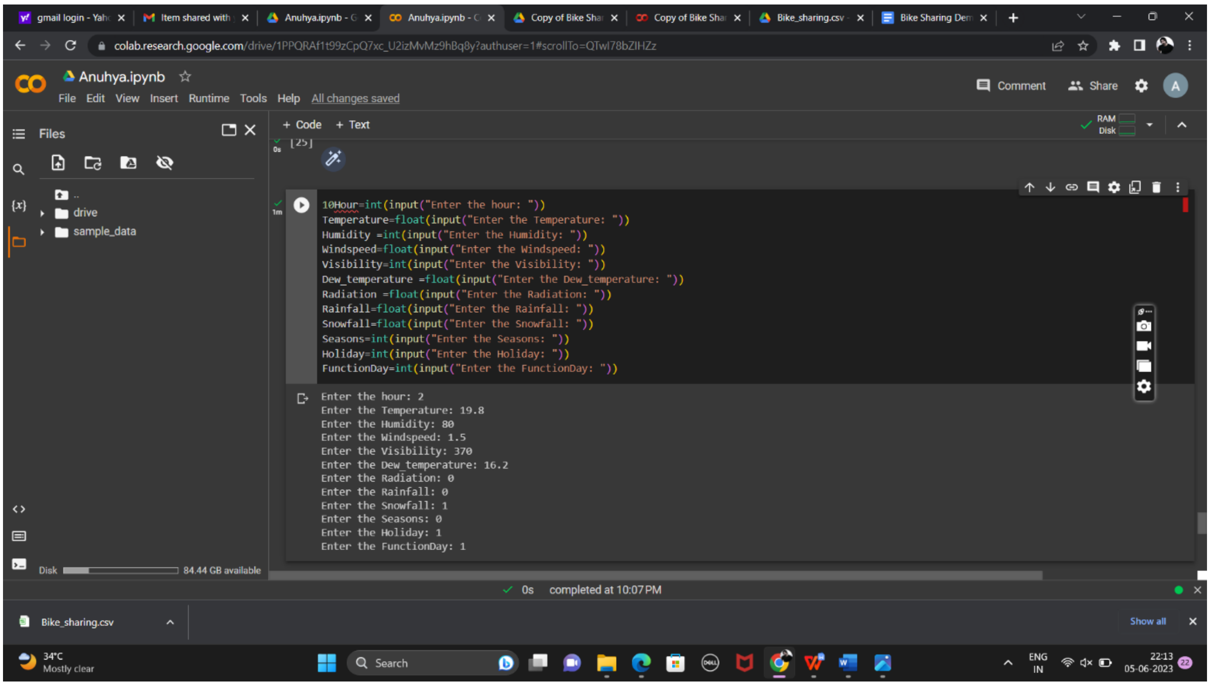

Fig 3:Prediction

5.CONCLUSION

Using the dataset's attributes, we created a model that predicted how many bikes will be shared in a day. A separate set of data was utilised to assess the accuracy of the model, and the findings were promising with a limited margin of error. It is critical to keep in mind that outside elements such as weather and events might have an impact on the actual number of bikes shared each day and are not taken into account in the dataset.As a result, additional study and changes may be required to improve the model's capacity to forecast the future and account for these external elements. The model's accuracy could also be improved by include new data sources, such as traffic patterns and public transportation timetables.

REFERENCES

1. Rennie, N., Cleophas, C., Sykulski, A. M., and Dost, F. (2021a). Detecting outlying demand in multi-leg bookings for transportation networks. arXiv (pre-print).
2. Rennie, N., Cleophas, C., Sykulski, A. M., and Dost, F. (2021b). Identifying and responding to outlier demand in revenue management. European Journal of Operational Research, 293:1015–1030.
3. Neumann-Saavedra, B. A., Mattfeld, D. C., and Hewitt, M. (2021). Assessing the operational impact of

tactical planning models for bike-sharing redistribution.

4. Transportation Research Part A: Policy and Practice, 150(June 2021):216–235. . Transportation research part C: emerging technologies, 135:103515.

5. Schuijbroek, J., Hampshire, R. C., and van Hoeve, W. J. (2017). Inventory rebalancing and vehicle routing in bike sharing systems. European Journal of Operational Research, 257(3):992–1004

6. Luo, H., Kou, Z., Zhao, F., and Cai, H. (2019). Comparative life cycle assessment of station-based and dock-less bike sharing systems. Resources, Conservation and Recycling, 146:180–189.

7. Teixeira, J. F., Silva, C., and Moura e Sá, F. (2021). Empirical evidence on the impacts of bikesharing: a literature review. Transport reviews, 41(3):329–351

8. Blazanin, G., Mondal, A., Asmussen, K. E., and Bhat, C. R. (2022). E-scooter sharing and bikesharing systems: An individual-level analysis of factors affecting first-use and suse frequency

9. Basole, R., Bendoly, E., Chandrasekaran, A., and Linderman, K. (2021). Visualization in operations management research. INFORMS Journal on Data Science.

AUTHOR PROFILES

Ms. M. ANITHA completed her Master of Computer Applications and Masters of Technology. Currently working as an Assistant professor in the Department of Masters of Computer Applications in the SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. Her area of interest includes Machine Learning with Python and DBMS.

Ms. Y.SIVA NAGA ANUHYA is an MCA student in the Department of Master Of Computer Applications at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. She has a Completed Degree in B.Sc.(computers) from S.V.L kranthi Degree college Avanigadda. Her areas of interest are DBMS, Java Script, and Machine Learning with Python.

International Journal For Advanced Research
In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362