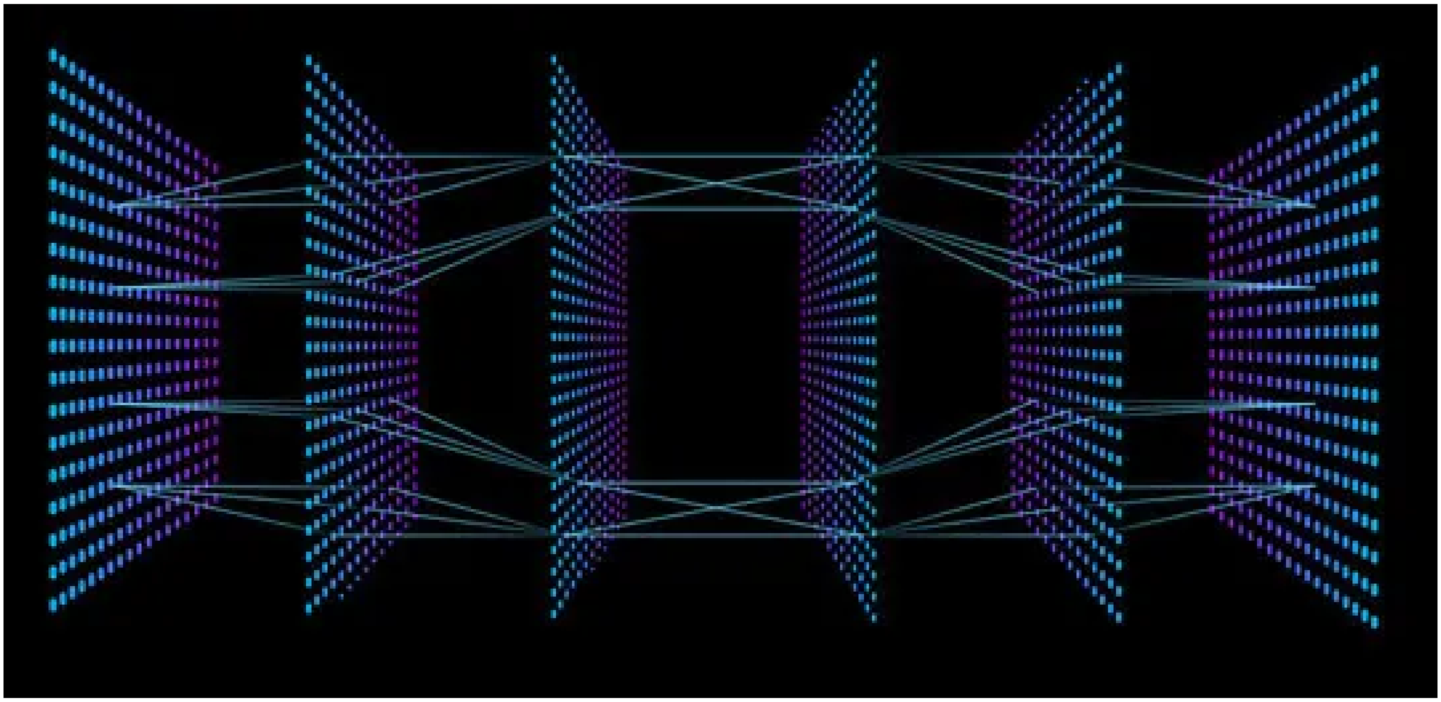
Quadruple Bam Corp.

Neural Networks



Presented by:

Lead: Scott

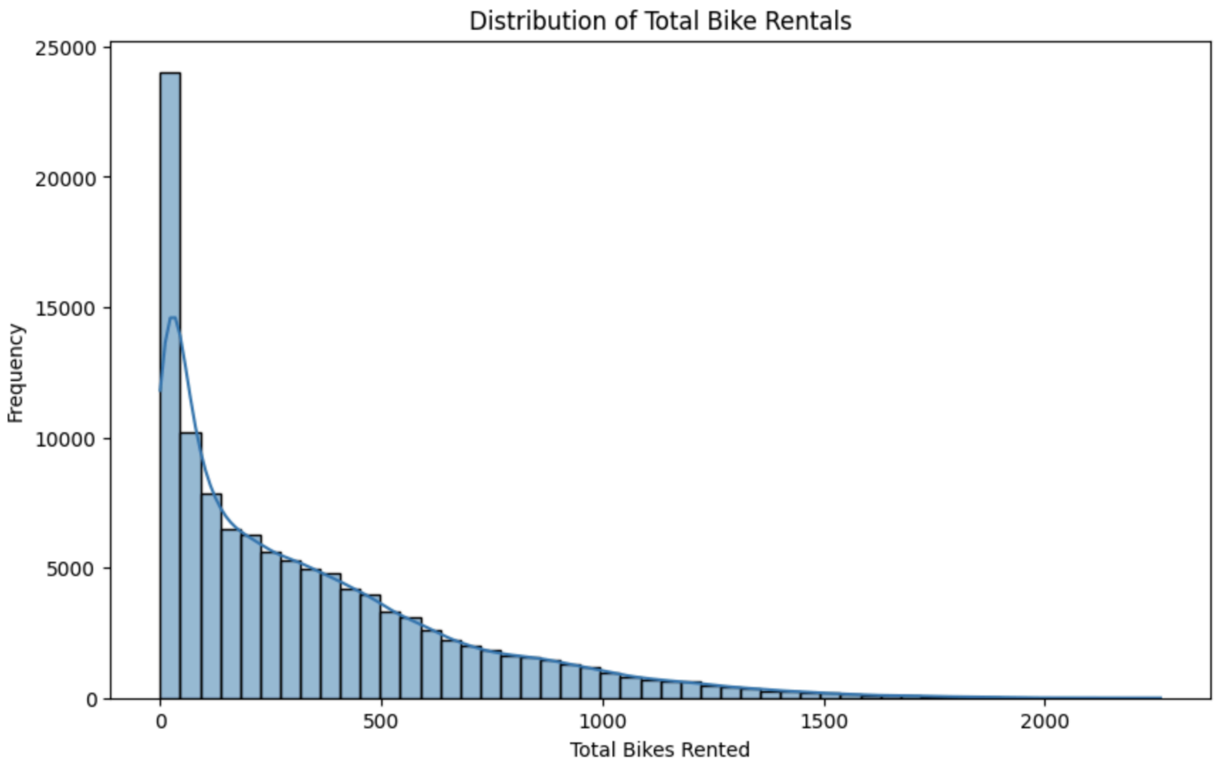
Model 1: Jarom

Model 2: Sam

Model 3: Luke

Introduction

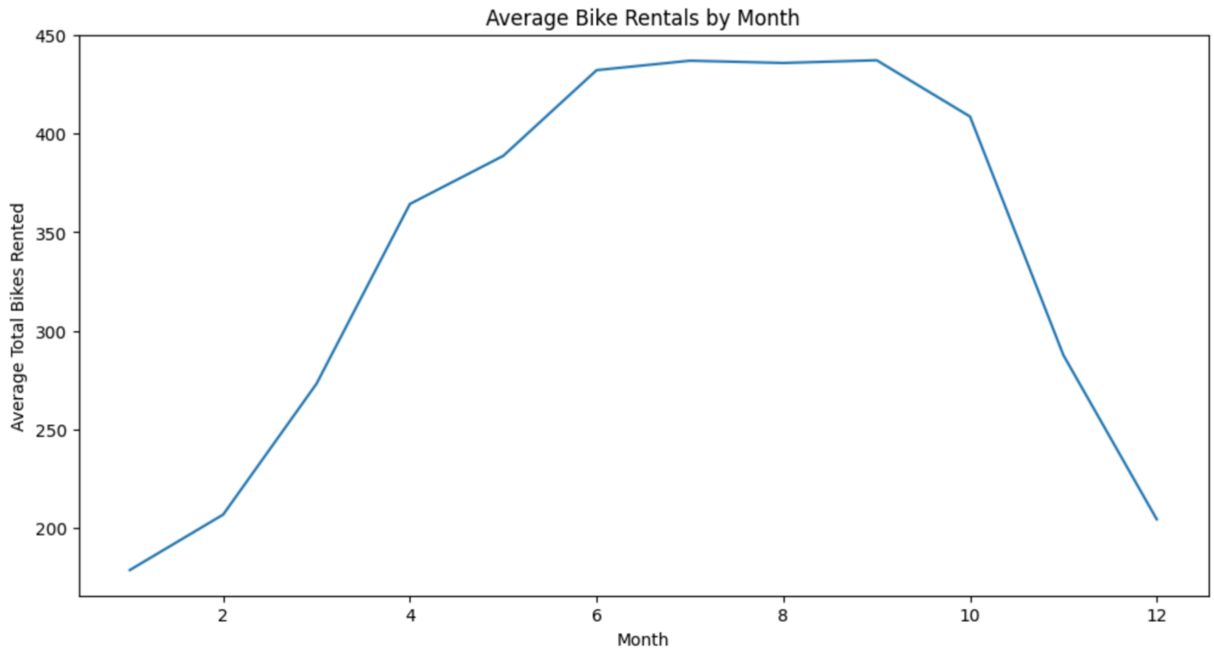
We were tasked with developing a machine learning model capable of predicting the number of rented bikes in use at any given time. This model serves as a critical tool for forecasting future demand, enabling our company to optimize bike availability, reduce inefficiencies, and ultimately maximize profits. After evaluating multiple machine learning approaches, we determined that a neural network would be the most effective model for this task, given its ability to recognize complex patterns in data.



Total bikes rented at different times

Methodology

Our first priority was data cleaning, as high-quality data is essential for building an accurate model. We began by reviewing the data dictionary and assessing the format in which the data was provided. Initially, we found that a significant portion of the dataset consisted of non-numerical values that were not immediately useful for machine learning. To address this, we systematically converted as much of the data as possible into numerical representations.

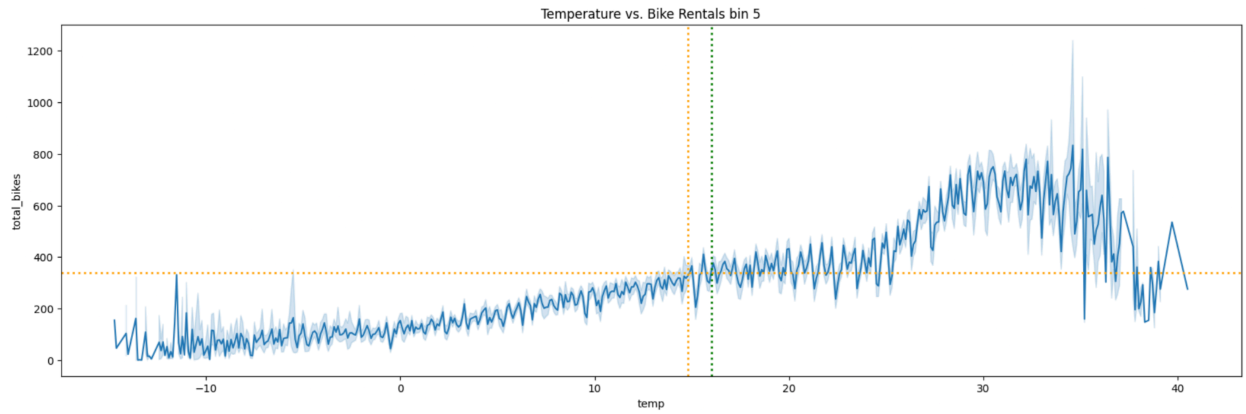


Bike rentals compared moth of the year

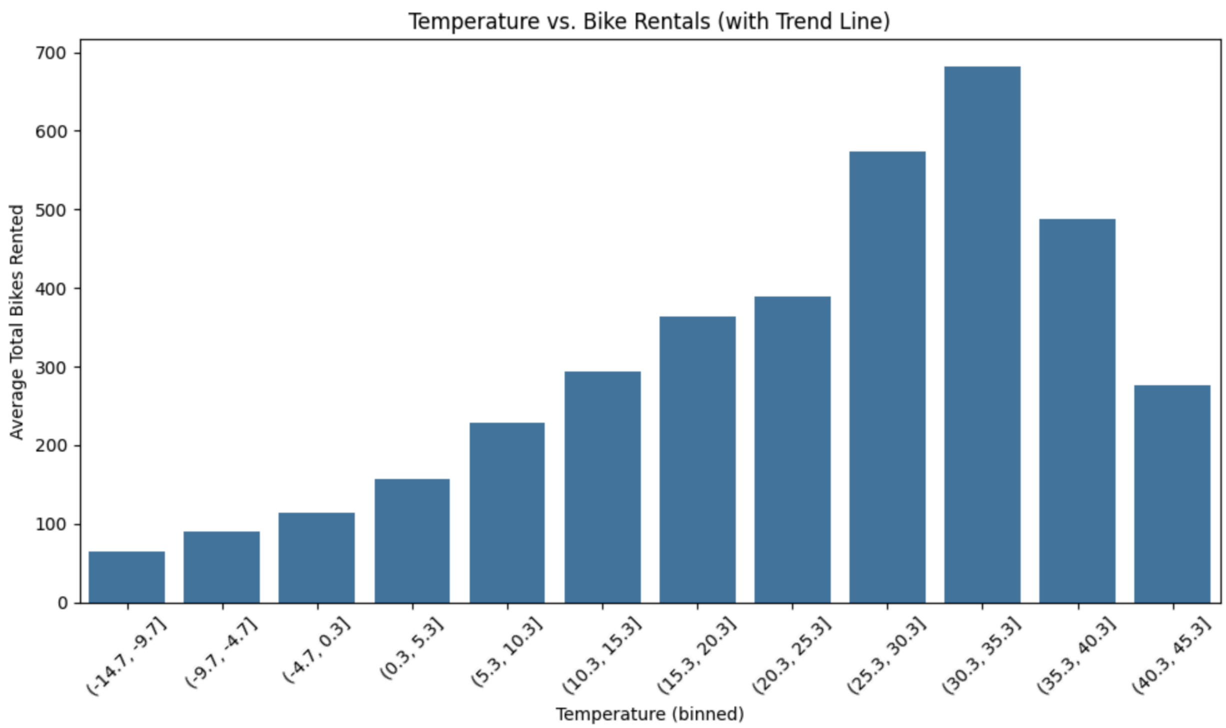
Beyond numerical conversion, we also refined the dataset by breaking broader data fields into more granular components. For example, instead of using raw date values such as "1/23/2013," we transformed the data to include separate columns for the month and day. Along with separating out many other data values in a similar way. This allowed the model to better capture seasonal and temporal trends, improving its ability to make accurate predictions.

Results

The impact of our data preprocessing was significant. Initially, an unrefined model trained on the raw data achieved an **R-squared value of approximately 0.60**, indicating a moderate correlation between predictions and actual results. However, after implementing our data cleaning techniques, the model's performance improved dramatically, reaching an **R-squared value of 0.89.**

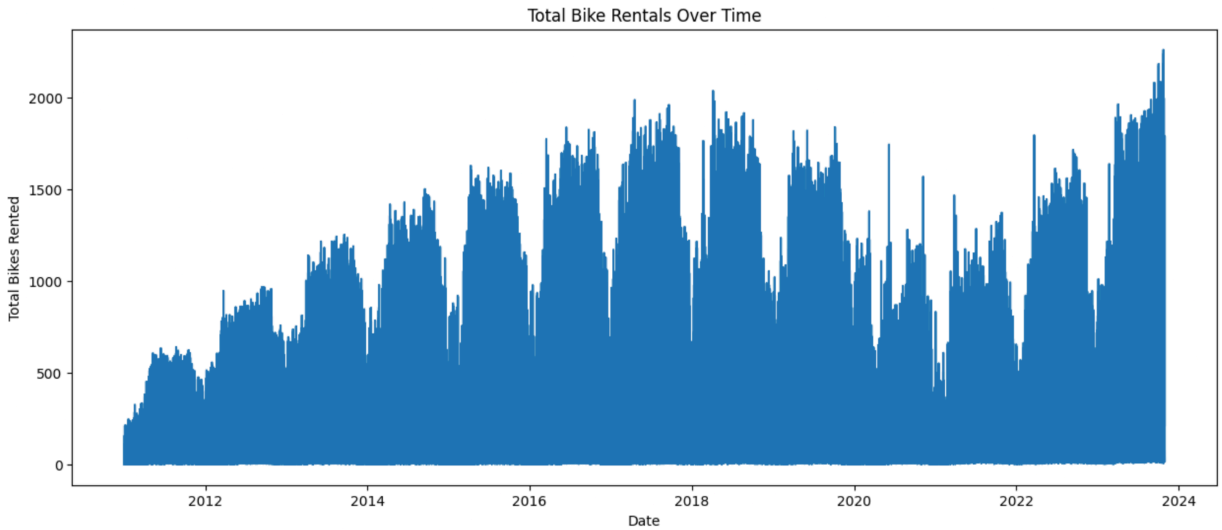


Bike rentals compared to temp



Bike rentals compared to temp

Recognizing the potential for further optimization, we fine-tuned the model through various adjustments, including modifying training parameters, optimizing the neural network structure, and setting epoch limits to prevent over-training. These refinements pushed our final model to achieve an **R-squared value of 0.955**, demonstrating a highly accurate ability to predict bike demand.



Bike rentals compared to year

Python Notebooks

https://colab.research.google.com/drive/1xNEGm5ok4Wfa0upW0U3sW4ZeKwZKAeIp?authuser=1

Conclusions

Through rigorous data preprocessing and strategic model optimization, we successfully developed a highly effective predictive model. The dramatic performance improvements resulting from data cleaning underscored the importance of high-quality input data in machine learning. Our findings suggest that with continued refinement—both in data processing and model tuning, we can further enhance predictive accuracy. This model provides a strong foundation for future demand forecasting, allowing us to efficiently allocate bike resources, reduce shortages, and improve overall operational efficiency.