AI Time Slot and Route Prediction for Buses using Machine Learning

Anshu Mishra

Submitted on: 21.06.2023

Abstract:

Transportation is one of the most important industries for any economy. A vast population daily use these services to move from one place to another. One of the major modes of transportation is roadways and buses are a major means of transportation by roads. As it occupies such a large market, modern technologies and applications can be a great asset for this industry. One such technology is Artificial Intelligence. AI is nowadays a part of almost every sector and industry. In transportation industry, AI tools like chatbots and virtual assistants are already prevalent. The product discussed below can also act as an important AI tool which can greatly help the roadways mainly buses. The product will use AI and Machine Learning models to predict the time slots and routes which are most preferred by the commuters of buses. This product aims towards improving business efficiency and make it convenient for the commuters to have buses readily available on their preferred routes and time slots

Problem Statement:

Roadways are one of the most important and convenient modes of transportation for people and products alike. People prefer roadways due to its flexible schedule and high availability. Buses are on of the most common medium of travelling by roads. Due to its high demand many private bus services are emerging in the market along with the government buses. This has provided the travelers with an array of choices to choose from when they decide to travel via buses. This has led to a high competition among the service providers and they aim towards capturing as much market as possible. Hence, they always work towards catering customers' needs in order to make more profits. One way of capturing a greater market share is to run more of their buses on customer preferred time slots and routes. Machine Learning can be a great asset for performing such type of predictive analysis. Manually determining the appropriate time slots and routes can be a very tedious and burdensome task, while with the help of appropriate machine learning model, this task can be performed very easily in a short amount of time.

Market/Customer/Business Need Assessment

On an average almost 18% of India's population uses public transportation for commuting. Buses being a major mean of public transportation also have a fairly large share of population among this 18%. Customers have specific needs and preferences while they use buses. They might prefer a specific time slot in which they travel more often like the office timings for daily commuting or a particular time of the day they prefer while travelling from one city to another. They might also prefer a specific route for either daily commuting or outstation travelling. For example, a customer might prefer that a bus route consist of major and important cities or parts of a city.

According to such customer preferences, bus service providers either private or public can adjust their bus schedules and routes to attract more travelers and earn more profits and revenue. Availability of more buses during preferred time slots and on preferred routes will provide a more ease of travelling for the customers and in return be beneficial for service providers.

Using the proposed model, the bus service providers can make accurate predictions about the more popular time slots and routes and can alter their strategies accordingly. The model will use the data from customers like when they are using the bus service (time of day, number of days in a week, etc.) and from where they board the bus also other demographic data can be collected to get a clearer picture of their preferences.

Target Specifications/Characterizations

The following can be the targets of a company for using this product:

1. Cater to customers and increase their customer base.
2. Experience an increase in revenue and profits.
3. Capture a larger market share.

Primary data from customers will be the input data for this model.

The target customers will be the frequent users of the bus services who commute through buses on a frequent basis.

External Search

1. https://towardsdatascience.com/how-machine-learning-and-ai-can-improve-travel-services-3fc8a88664c4
2. https://www.mdpi.com/2071-1050/13/1/225
3. https://link.springer.com/article/10.1007/s12469-022-00319-y
4. https://pdfs.semanticscholar.org/947f/8fca2e8f3e2a6570cf0982626d109b94ec03.pdf
5. https://mitsloan.mit.edu/ideas-made-to-matter/creating-better-bus-routes-algorithms

Applicable Constraints

1. Availability of Data: Reliable and accurate data will be difficult to get as all the customers might not be willing to provide such data. Also, the data provided might not be accurate.

2. Traffic Issues: Running of many such large vehicles at the same time might lead to traffic jams on the road causing wastage of time and resources.

3. Lack of Developers: The service providers might not always have the proper availability of developers to maintain and run the product.

4. Lack of proper means to Collect Data: Companies might not necessarily have a proper and fast way to collect data.

5. Continuous Collection of Data: It will be a huge cost to the companies to deploy the means to continuously collect data for the product.

Applicable Regulations

1. Government Regulations: With the increase in the number of buses, the company has to obtain permit for each vehicle which can be tedious and there are many other regulations while operating public transports.

2. Environmental Regulations: Regulations to control pollution and traffic conditions.

3. Data Privacy: Privacy of users' data is also stressed upon by the government.

Business Model
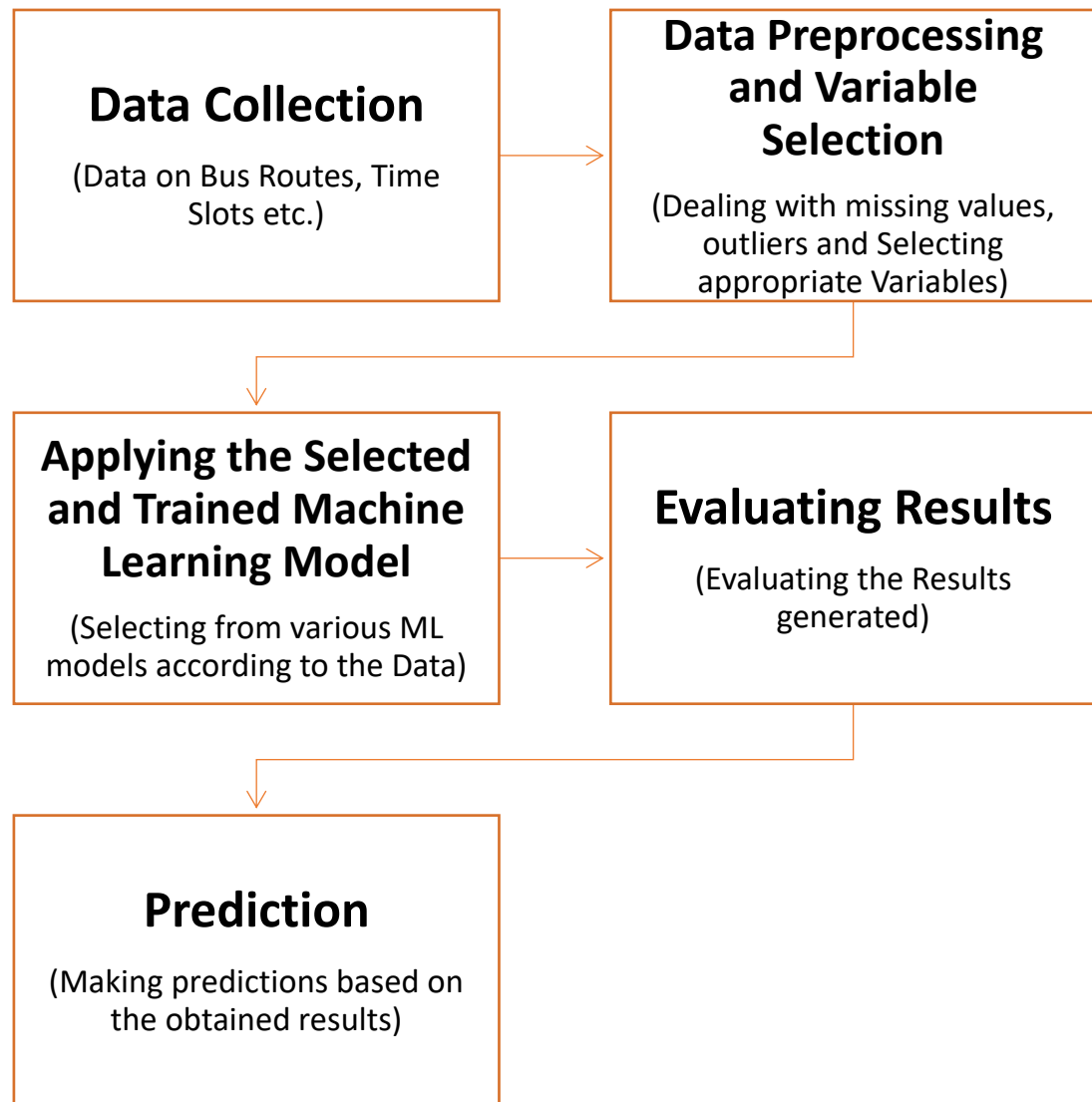
The product can use freemium method to get monetized:

1. Free Version: The free version will provide businesses with some basic features like the preferences of limited number of customers based on a limited number of categories. But the free version will provide enough information to attract the businesses to use the premium versions.

2. Premium Version: The premium version can be tiered that is it will have various plans and will give businesses access to the features according to the price category of the premium version. With premium version, the businesses will have full access to the product where they can get complete information about customer preferences about time slots and routes.

3. Free Trial to Premium Features: In order to promote premium version and encourage more businesses to subscribe the premium versions a limited day free trial of premium features can be provided applicable once per user like for the duration of 7 or 10 days.

Concept Development

AI is one of the most up and coming technology which is attracting businesses and general users alike. Most businesses nowadays are adopting AI services in one way or another to enhance their business. Most customers are also attracted to AI generated services of companies. This AI product will use machine learning algorithms, programming languages (e.g., Python) and database management software (e.g., SQL) to derive meaningful insights from the given data. The product will give the bus service providers the most customer preferred time slots and routes based on the data provided by the customers and from their travelling history so that they can use this to increase their business efficiency. The product will also give the businesses real time updates about any changes in the previous information so that they can keep up with the varying customer demands. Using this product, the service providers can increase their customer base and increase their revenue. They can also further increase their users by partnering with online ticket booking services and show the availability of their buses for the popular time slots and routes to a wider population. The product will be available for both free of cost and subscription fees. The free version will have limited features while full access can be obtained by paying the subscription fees.

Product Prototype (with Schematic Diagram)

This AI product aims towards eliminating the difficult and tedious task of evaluating the customers' choices and preferences for businesses and provide them with definitive results which would help them to cater better to their customers and in return earn more revenue and profits. This product will also help in expanding businesses and introduce the service providers to new technologies and the world of AI. The product will derive results based on customer provided data and historical data from partner businesses. The product will evaluate the results derived from various machine learning algorithms and then provide the most appropriate results.

Data Collection

(Data on Bus Routes, Time Slots etc.)

→

Data Preprocessing and Variable Selection

(Dealing with missing values, outliers and Selecting appropriate Variables)

Applying the Selected and Trained Machine Learning Model

(Selecting from various ML models according to the Data)

→

Evaluating Results

(Evaluating the Results generated)

Prediction

(Making predictions based on the obtained results)

How the Product Works

1. **Data Collection**: Gathering historical data on bus routes, time slots, and customer preferences. This could include data on when buses are most crowded, when customers

typically use the bus, customer feedback on preferred times and routes, and so on. Survey forms filled by the customers can also act as a data source.

2. **Data Preprocessing**: Cleaning and preprocessing the data to make it suitable for machine learning. This could involve handling missing values, dealing with outliers, and converting categorical variables into a format that a machine learning algorithm can understand.

3. **Feature Engineering**: Identifying the features (variables) that are likely to be predictive of the ideal time slots and routes. These could include time of day, day of the week, location, weather conditions, and so on.

4. **Model Selection**: Choosing a suitable machine learning model. Given that we are trying to predict both time slots (a continuous variable) and routes (a categorical variable), we might need to use a combination of regression and classification models. Algorithms such as Linear Regression, Decision Trees etc. can be used for the model. Reinforcement Learning could also be a good approach, as it can handle complex problems and optimize decisions based on rewards (in this case, customer satisfaction or usage).

5. **Model Training**: Training the model on the preprocessed data. This involves feeding the data into the model so that it can learn the relationships between the features and the target variables (ideal time slots and routes).

6. **Model Evaluation**: Evaluating the performance of the model using appropriate metrics. For regression tasks, this could be mean absolute error or root mean squared error. For classification tasks, this could be accuracy, precision, recall, or F1 score.

7. **Model Optimization**: Tuning the model's parameters to improve its performance. This could involve techniques like grid search or random search.

8. **Prediction**: Using the trained model to predict the most ideal time slots and routes based on customer preferences.

9. **Deployment and Monitoring**: Deploying the model in a real-world setting, monitoring its performance, and making adjustments as necessary.

10. **Making the Final Product available in the Market**: After the product is fully tested, it will be monetized and made available in the market for the use of relevant businesses. The interested parties can then subscribe accordingly and take advantage of the of the product's services.

Algorithms, Frameworks and Software Needed

1. **Algorithms:** The model will use algorithms like Linear Regression, Decision Trees etc. depending on specific use cases and the learning method can be Reinforcement Learning as it will provide better results.
2. **Frameworks:** Various Python libraries like TensorFlow, scikit-learn and PyTorch can be used for the machine learning model and algorithms.

3. **Software:** Software like SQL which is a database management software can be used for storing the vast amount of data and to make it easily compatible with the model.

Team Required to Develop the Product

The following team will be required to develop and maintain the product

1. Data Analysts and Scientists
2. Software Developers
3. AI Developers
4. Domain Experts

Conclusion

This AI product will help the businesses involved in providing transportation services through buses to get a better insight into the preferences of their customers regarding the routes they prefer to travel on and the time slots they generally commute in. This will enable the businesses to make their strategies according to such results and get better outcomes. Such strategies will also make travelling convenient for the general public hence increasing the customer base of these service providers. It will also convenience to the businesses as they don't require such tedious and constant data collection for analysis on their personal scale and can thus avoid a huge cost which could be used for other ventures.