Individual Capstone Project Assessment (DAB 322)

Project Title: Transit Windsor – Streamlining Operations and Improve Coordination

Student Name: Krupali Jayeshbhai Mewada (0792868)

Date: April 18th 2023

1) Executive Summary

A modern city's ability to provide safe, economical, and timely transportation for locals, visitors, and everyone else depends on its transit system. Given that Windsor is home to two significant educational institutions in southern Ontario, the need for the transit system to operate at capacity is increased by the thousands of international students who already reside there and use the transit system, as well as by a fresh batch at the beginning of each major intake when new students arrive in large numbers. Due to all of these issues, Transit Windsor is having trouble keeping up with the rising demand. The difficulties that are currently being faced are as follows:

In order to improve the transit system in Windsor, our team has achieved enormous milestones. Our machine learning model and Tableau dashboard gave us insights that allowed us to create plans and suggestions for tackling the problems we had found. The transit network should be expanded, service frequency and dependability should be increased, and new technology should be implemented to improve the passenger experience. However, there were several difficulties and obstacles we had to overcome along the way, including a lack of financing and resources as well as to work with existing players that are already in the industry. Despite these obstacles, we are still dedicated to achieving our objectives and providing the Windsor community with significant outcomes.

2) Progress Update

Our team want to be able to find an innovative and data-driven solution creating a lasting impact in the public transit domain. Firstly, we aim to enhance Transit Windsor's efficiency and reliability by tackling congestion and overcrowding. Talking in detail about approaching the proposed problem, our team went with the following steps:

Data Assessment:

Our team will be using descriptive analysis to gain a better understanding of the data and to identify patterns, trends, and relationships. This will help in optimization by looking at various data points in the dataset. Then, we started by performing an exploratory data analysis (EDA) on the transit data, which involved evaluating the data's quality, cleaning and converting it, and meaningfully visualising it.

Model Selection:

We chose the most suitable machine learning model to assist us in predicting the elements causing these difficulties. We started with implementing a classification problem which didn't workout as it was over-fitting the model. We then tried Linear Regression, decision tree, Random Forest Algorithms and settled on Random Forest as it gave us the best results.We created a prediction model that was able to pinpoint the issues' primary causes, such as busy periods and areas with high demand, and it assisted us in formulating suggestions for enhancing the transport service.

Dashboard:

After that, the insights gathered are communicated through a dashboard using various charts and other infographics that made it simple for stakeholders to see the outcomes of our prediction model and pinpoint the main problems and potential areas for development. The dashboard displayed the transport system's performance over time and gave information on the variables influencing passenger demand.

3) Key Progress

Right after the group discussion about the project, I conducted a business case analysis by taking into account various aspects as well as directions this project can take. I also conducted a risk assessment by considering different possibilities in terms of unanticipated technical / logistical problems, projects effectiveness and requirements related to the modification to the current transit system. I then constructed a project charter that had detailed explanation on the scope of the project, budget, requirements, communication, and weekly meetings.

Initially we started with doing EDAs with tableau but soon switched to python to support the final product implementation. We sourced the data available from Transit Windsor and cleaned it. We also added few rows to the dataset to allow for addition of calculated field from already existing fields. Categorizing the steps that were implemented and executed so far, following tasks were performed: Data gathering, cleaning, merging, exploratory data analysis and initial data visualization. Further, we established relationship between different variables and used different algorithms on the data including linear regression and random forest regression.

Additionally, building a Tableau dashboard to present the outcomes of our machine learning model in an clear and understandable manner. The dashboard gives insights into the current state of the transit system, as well as ideas for improvement based on our model's forecasts.

4) Data and Metrics

Using the data, we employed several techniques, such as Classification, Linear regression, Decision Tree and Random forest regression, to find relationships between various variables. With regards to classification, a random forest classifier was also tried but was chosen not to be gone forward with because of over fitting problem. The objective of the analysis is to find regularities and tendencies in the data that can be utilized to enhance the functioning of Transit Windsor.

Our accuracy in the preliminary stages of model development was approximately 0.6. Nevertheless, we were able to considerably raise the accuracy to above 0.8 after performing feature engineering on the dataset. In order to make sure that our model was correctly detecting patterns and forecasting rider counts, we also monitored precision, recall, and F1 score. We then developed a dashboard on Tableau that provides graphical representation of the key insights and important data measurements from the machine learning model. The result of this would enable stakeholders in getting patterns and insights in terms of how effective the machine learning model is in predicting rider counts.

5) Discuss any changes

In the upcoming term, our plan is to leverage the machine learning model in the development of user centric solutions such as a mobile application that provides real time data on routes and schedules.

6) Challenges

I think the hardest part in coming up with a solution for this problem was to be able to leverage the infrastructure that is already available with Transit Windsor. Working with the limited data that we have from Transit Windsor is also a major challenge owning to the fact that there is a significant difference in the way Transit Windsor used to operate and plan their routes and schedules pre vs post covid. This can affect the predictability of the resulting ML model.

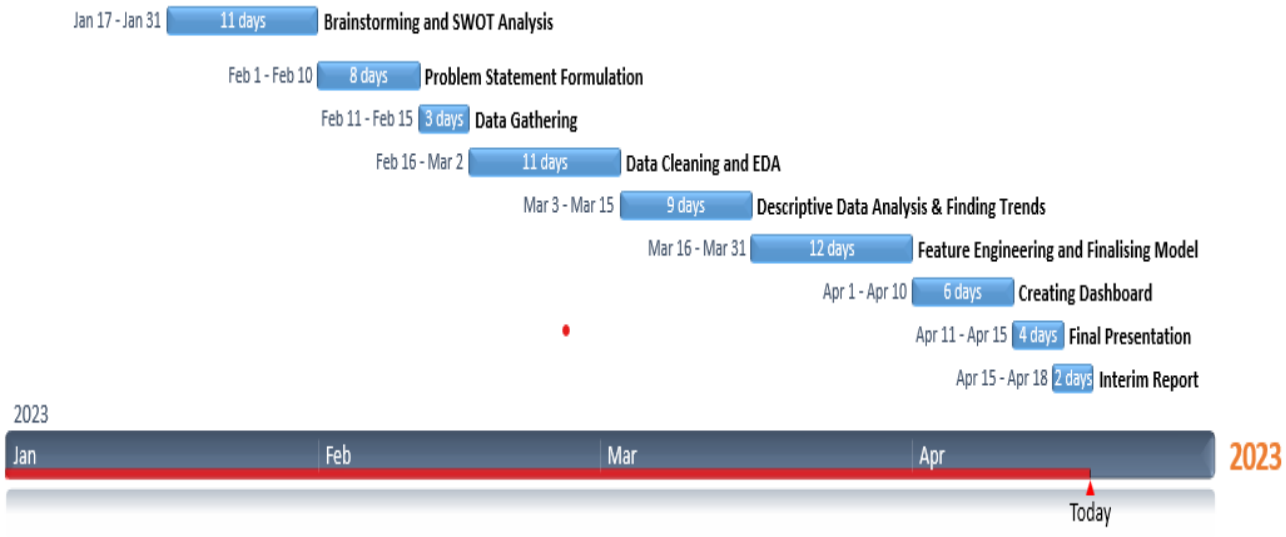
As an international student using Transit Windsor daily, we had firsthand experience when it comes to the challenges and problems that were being faced by the users of Transit Windsor. Despite being aware of the problem, the process of figuring about the approach and direction to take when it comes to the context of the project and also being able to reach out to appropriate individuals to get corresponding information and come up with a solution as a team was quite a major task that we have accomplished.

Initially, I struggled with technical aspects of the project that are required in proposing a solution. I was able to overcome this problem by revisiting the basics that I had learned throughout my bachelor's degree as well as reaching out to team mates in case I was stuck somewhere.

7) Results

Implementation of the dashboard as well as the machine learning model was successfully done. The accuracy of the machine learning model and how it was subsequently improved has already been discussed in the *Data & Metrics* section above. In addition, our initial goal was to concentrate only on the transit system's business-to-business component. Although we now understand the value of integrating the user's perspective as well, we still intend to include this element in our project plan in the future semester.

8) Timeline and Budget



This timeline has been maintained and updated right from the start. From data gathering to creating dashboard, we have always had it all planned. As soon as we got the data, we started working on what would the next checkpoint to achieve and when. As a team we successfully managed to complete all the tasks on time.

Budget

Resources	Cost
Mobile Application	70,000 CAD
Hardware (GPS & Others)	15,000 CAD
Database	20,000 CAD
Training	15,000 CAD
Technical Support	9,000 CAD
Testing	18,000 CAD
Total: 147,000 CAD	

- Mobile Application – Forms the base source which will be used by the consumers. Average yearly salary for a Jr. to Mid level developer is about 70,000.
- Hardware (GPS & Others) – Any trackable device to support the new infrastructure. Estimating that cost to be around 15,000.
- Database – Database will be used to store the data as well as meta-data. The cloud databases have different pricing models but using a SQL database like MySQL on a computing price pricing model is estimated to cost around 20,000.
- Training – Stakeholders will need training on working with and supporting the new implementation. That is estimated to cost around 15,000 CAD.
- Technical Support – Dedicated staff to support a client's infrastructure will be needed. This is estimated to cost around 9,000 CAD.
- Testing – Testing the new implementation along various stages will be needed. This will cost around 18,000 CAD.

9) Next Steps

Building better alliances with other transit agencies, companies, and organisations to develop a more connected transportation network and provide our riders with more comprehensive services might be one potential area of concentration for our project. The resulting implementation and system can also be expanded to other cities with similar transit systems who might be facing similar problems.

More availability of data could play a major role in widespread adoption of this solution. With more data the resulting ML model would have a better accuracy and would be able to provide better insights. In the future, there are various opportunities for further exploration and advancement, including the assessment of the effectiveness of the proposed changes, finding solutions to sustainability concerns, and exploring possibilities for integrating other transportation modes.

10) Conclusion

Conclusively speaking, our project to develop an optimized and efficient transit plan for Windsor has identified several key challenges facing the city's current transit system. These include overcrowding, longer waiting times, and accessibility issues, all of which are made worse by the growing population of international students and expanding urban areas. To address these challenges, we have conducted thorough research and analysis, taking into account the needs and preferences of different stakeholders, best practices in transit planning and management, and the existing infrastructure and resources available for the transit system.

Our proposed solutions involve optimizing bus routes, improving accessibility for all riders, and strengthening partnerships with other transit agencies, businesses, and organizations to create a more integrated transportation network. By implementing these solutions, we believe that we can create a more reliable, efficient, and accessible transit system for Windsor residents and visitors alike.

Acknowledgements

We would like to say a big thank you and acknowledge Transit Windsor and Professor John Ulakovich throughout the development life cycle of this project. Transit Windsor made it possible for us to proceed with this project by providing with the data. We would also like to say a big thank you to Professor Manoop Paliot who guided us and helped us throughout the project.