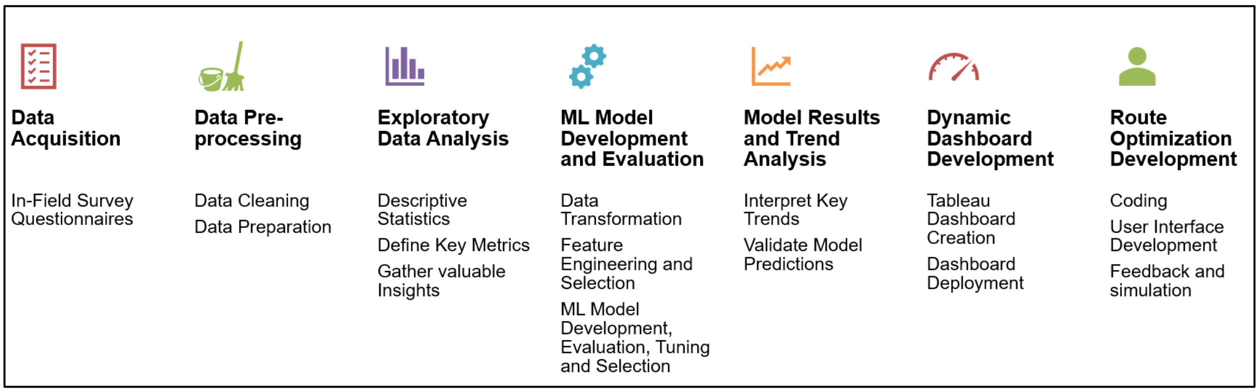
Project Progress Report

Project Title: Edmonton Food Drive 2024

Project Phase:

Figure 1: Project Methodology Process



Team Members' Name with specific roles:

- **Kendrick Moreno:** Project Manager, Software Developer, Data Scientist
- **Catrina Llamas:** Project Manager, Data Analyst, Machine Learning Analyst
- **Roe Alincastre:** Project Manager, Data Analyst, Machine Learning Analyst

Reporting Period:

- Sep 17th: Project Kick-off Meeting
- Oct 8th: Exploratory Data Analysis (EDA) Demo
- Nov 12th: Model Iteration Demo
- Dec 11th: Final Demo and Deployment
- Dec 18th: Documentation & Portfolio Updates

Project Overview: The goal of this project is to develop a machine learning solution to optimize the management of food donation activities in Edmonton, AB. The project aims to enhance the efficiency and effectiveness of drop-off and pick-up processes, streamline route planning, and improve resource allocation.

Problem Statement: The current food donation management system in Edmonton faces challenges in coordinating drop-off locations, pick-up processes, and route planning. There is a need to automate and optimize these processes to ensure timely collection of donations and minimize logistical complexities.

Objectives:

- Enhance Data Collection and Analysis Processes.
- Provide Comparative Insights to support Data-Driven Decision-Making.

- Predict Future Donation Trends and identify which wards or stakes will be most effective over the next 5 years.
- Streamline and Optimized manual process of map generation.

Proposed Solutions:

- *Enhance Data-Collection Forms for Faster Response Times*: Refine the existing data-acquisition forms to simplify the process, enabling volunteers to answer questions as quickly and efficiently as possible.
- *Conduct Year-over-Year Analysis to Identify Key Trends and Performance Metrics*: Use data visualization and statistical techniques to perform a year-over-year analysis, revealing critical trends and performance indicators.
- *Design Interactive Dashboards for Data Exploration and Comparison*: Create user-friendly, interactive dashboards that allow stakeholders to easily explore and compare data, facilitating more informed decision-making.
- *Develop a Predictive Model to Estimate Donation Patterns and Trends*: Implement machine learning techniques to develop a predictive model that forecasts donation patterns and identifies emerging trends.
- *Forecast the Efficiency of Wards or Stakes Using Predictive Modeling*: Build a predictive model to estimate which Wards or Stakes will have the greatest impact in terms of efficiency over the coming years.
- *Develop or Integrate a Digital Route Mapping Application for Volunteers*: Create a route digitization application that automatically generates optimized maps for volunteers, improving operational efficiency.
- *Create a Long-Term Route Optimization Solution Based on Hot-Zone Addresses*: Develop a route mapping application that generates interactive maps for volunteers, focusing on high-demand or hot-zone addresses for long-term operational efficiency.

Accomplishments:

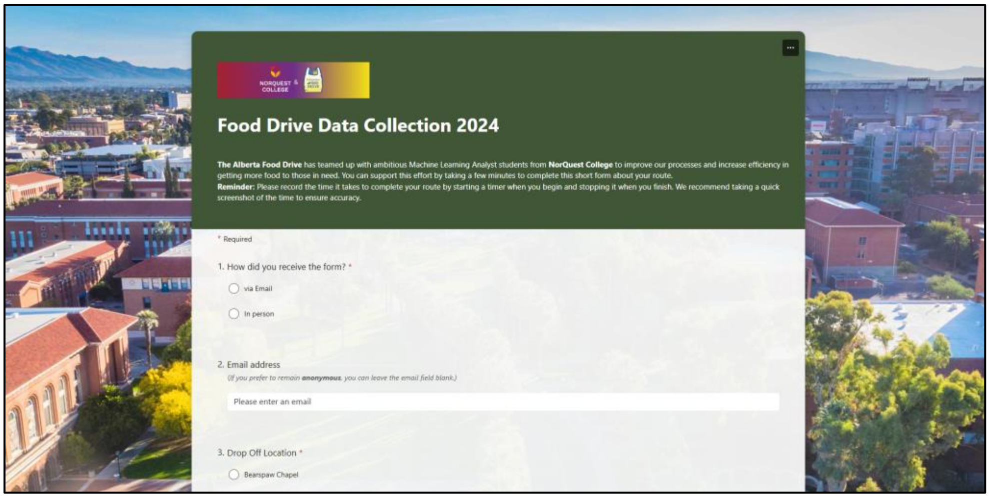
- **Data collection form creation: Date:** August 16 – 17, 2024
- **Data Collection:** August 28, 2024
- **Exploratory Data Analysis (EDA):** August 29, 2024 – October 4, 2024
- **EDA reporting to stakeholders:** October 8, 2024
- **Data Preprocessing:** October 7 – 14, 2024
- **Data Modelling and Evaluation:** October 14 – 21, 2024
- **Model Tuning and Selection:** October 21 – 28, 2024
- **Model Iteration Demo:** November 11, 2024
- **Model Deployment:** November 18 – 30, 2024
- **Final Demo:** December 11, 2024

Progress timeline Link: [LINK]

Data Acquisition Form:

We enhanced the existing Google Form by adding restrictions on the data input, while ensuring that the time required for volunteers to complete the form remains as short as before.

Figure 2: Data Acquisition Form for EFD 2024



Attach Link to Data Acquisition Form: [LINK]

Datasets:

Food Drive Data Collection 2023: [LINK]

Description: This dataset comprises data collected via a Google Form during the Edmonton Food Drive 2023.

Number of Columns:

Number of Rows: 454

Table 1: Feature Information of EFD 2023 dataset

Column Name	Description
Date	The date the food drive activity took place.
Location	The specific area or neighborhood where the food drive was conducted.
Stake	The organization or group responsible for managing the volunteers in the area.
# of Adult Volunteers	The number of adult volunteers who participated in the activity.
# of Youth Volunteers	The number of youth volunteers who participated in the activity.
Donation Bags Collected	The total number of donation bags collected during the activity.
Time to Complete (min)	The total time (in minutes) taken to complete the assigned route(s).
Completed More Than One Route	Indicates whether more than one route was completed (e.g., Yes/No).
Ward	The municipal ward where the food drive activity occurred.
Routes Completed	The total number of routes completed by the volunteers.
Doors in Route	The total number of doors covered within the assigned route.
Route Number/Name.1	The identifier or name of the route assigned to the volunteers.
Time Spent	The total duration volunteers spent during the food drive activity.

Food Drive Data Collection 2024: [LINK]

Description: This dataset contains information gathered through a Microsoft Form during this year's Edmonton Food Drive.

Number of Columns: 31

Number of Rows: 653

Table 2: Feature Information of EFD 2024 dataset

Column Name	Description
ID	A unique identifier assigned to each form submission.
Start time	The time the volunteer began filling out the form.
Completion time	The time the volunteer completed the form.
Email	The email address provided by the volunteer.
Name	The name of the volunteer.
How did you receive the form?	The method through which the volunteer received the form (e.g., email, link).
Email address	The contact email address for further communication.
Drop Off Location	The primary location where donations were dropped off.
Other Drop-off Locations	Additional locations where donations were dropped off.
Stake	The specific stake responsible for organizing the volunteer's participation.
Bonnie Doon Stake	Indicates involvement with the Bonnie Doon Stake.
Edmonton North Stake	Indicates involvement with the Edmonton North Stake.
Gateway Stake	Indicates involvement with the Gateway Stake.
Riverbend Stake	Indicates involvement with the Riverbend Stake.
Sherwood Park Stake	Indicates involvement with the Sherwood Park Stake.
YSA Stake	Indicates involvement with the Young Single Adults (YSA) Stake.
Route Number/Name	The identifier or name of the donation collection route.
Time Spent Collecting Donations	The total time spent collecting donations for the route.
# of Adult Volunteers who participated in this route	The number of adult volunteers involved in this specific route.
# of Youth Volunteers who participated in this route	The number of youth volunteers involved in this specific route.
# of Doors in Route	The total number of doors covered within the route.
# of Donation Bags Collected	The total number of donation bags collected from the route.
Did you complete more than 1 route?	Indicates whether the volunteer completed more than one route (e.g., Yes/No).
How many routes did you complete?	The total number of routes completed by the volunteer.
Additional Routes completed (2 routes)	Details about a second additional route completed, if applicable.
Additional routes completed (3 routes)	Details about a third additional route completed, if applicable.
Additional routes completed (3 routes)2	Details about another third route completed, if applicable.
Additional routes completed (More than 3 Routes)	Details about additional routes completed beyond three, if applicable.
Additional routes completed (More than 3 Routes)2	Further details about routes completed beyond three, if applicable.
Additional routes completed (More than 3 Routes)3	Further details about routes completed beyond three, if applicable.
Comments or Feedback	Any additional comments, suggestions, or feedback provided by the volunteer.

City of Edmonton Neighborhood Dataset: [LINK]

Description: The City of Edmonton Neighborhood Dataset provides comprehensive information about neighborhood boundaries, demographics, land use, and other characteristics for urban planning and analysis.

Number of Columns: 3

Number of Rows: 427

Table 3: Feature Information of City of Edmonton Neighborhood dataset

Column Name	Description
Neighbourhood Name	The official name of the neighborhood in the City of Edmonton.
Latitude	The geographic coordinate specifying the north-south position of the neighborhood.
Longitude	The geographic coordinate specifying the east-west position of the neighborhood.

Challenges Encountered:

- *Data Collection Challenges*: Acquiring complete data during the food drive was difficult due to the need for more personnel to be assigned to each ward, which slowed down the process of gathering all required information.
- *Dataset Structure Differences*: The structure of the datasets for 2023 and 2024 are notably different, requiring additional efforts to match the data before analysis.
- *Data Preparation Effort*: The dataset requires substantial cleaning and preprocessing to address inconsistencies and ensure it is in a format suitable for machine learning analysis.

Data Refinement:

For the EFD 2024 dataset, we identified the following issues and applied the respective methods to address them.

Table 4: Identified Issues in the EFD 2024 Dataset and Their Respective Solutions

Issues Detected	Refining Method
Too long column names	Rename column names for clarity
Inconsistent string formats	Removed leading and trailing spaces Converted to title format Removed unnecessary characters
Incorrect and inconsistent data types	Converted variables to the correct data types
Detected null values	Performed mean imputation to replace null values
Detected empty values	Tagged empty categorical fields with placeholders (e.g., "Unknown Routes")
Duplicated values	Dropped duplicated values and columns
Too many irrelevant data	Dropped irrelevant columns
Identified outliers	Detected using IQR method and imputed using mean

After performing data refining on the EFD 2024 dataset, we merged it with the EFD 2023 dataset and the City of Edmonton Neighborhood dataset. This allowed us to determine the geolocation of the wards.

Link of the Cleaned Dataset: [Link]

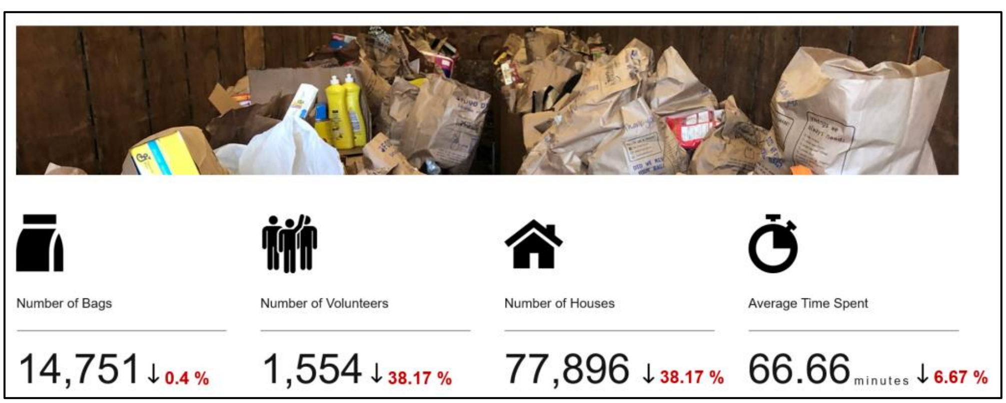
Exploratory Data Analysis (EDA) Highlights:

Our EDA strategy involves examining each feature individually and performing detailed analyses for each.

- *Drop-off Locations*: Analyze the frequency and distribution of different drop-off locations. Explore the relationship between drop-off locations and variables like the number of donation bags and volunteers.
- *Stake*: Assess the frequency of different "Stake" values. Examine the impact of "Stake" on numerical features such as the number of doors and donation bags.
- *Time Spent*: Explore the frequency of each time category. Analyze the distribution of time spent for each "Stake" and "Ward."
- *Ward*: Analyze the distribution of data across different wards. Examine how "Ward" influences other variables, such as the number of donation bags and routes.
- *Total Volunteers*: Check the distribution of volunteers. Analyze the correlation between adult volunteers and other numerical features.
- *Number of Doors*: Assess the distribution of the number of doors. Examine the relationship between the number of doors and other categorical variables. Calculate the average number of doors by "Stake."
- *Total Number of Donation Bags*: Explore relationships with other numerical variables. Compare the variation in donation bags across different locations and wards.
- *Number of Routes*: Examine the distribution of the number of routes. Compare the number of donation bags to the number of routes.
- *Years*: Analyze trends over the years for donation bags and total volunteers. Explore how the numerical features have changed over time.

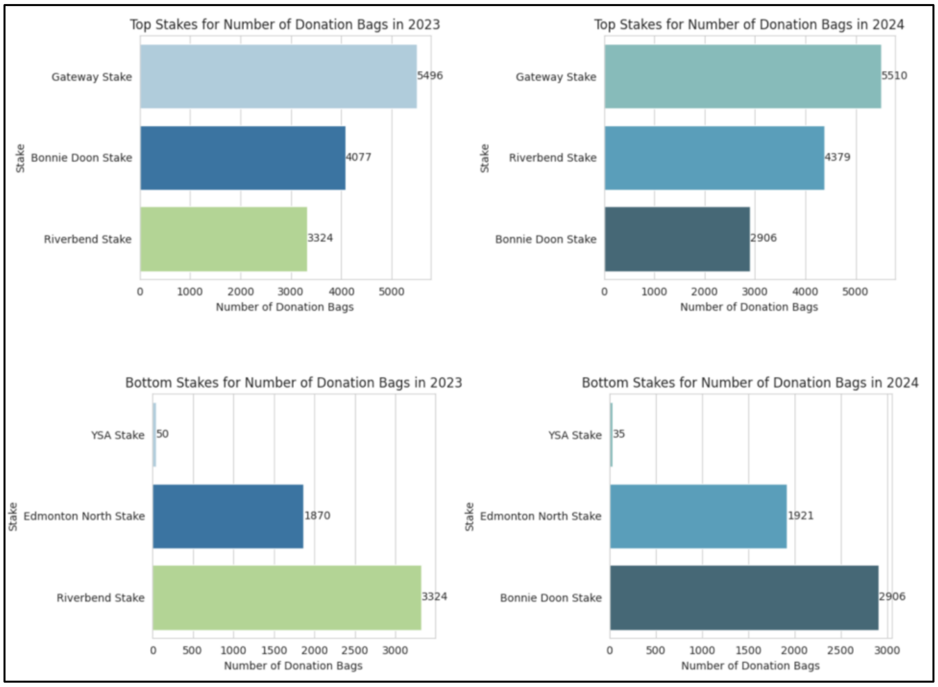
Below are the key insights derived from our exploratory data analysis (EDA).

Figure 3: 2024 vs 2023 EFD Highlights



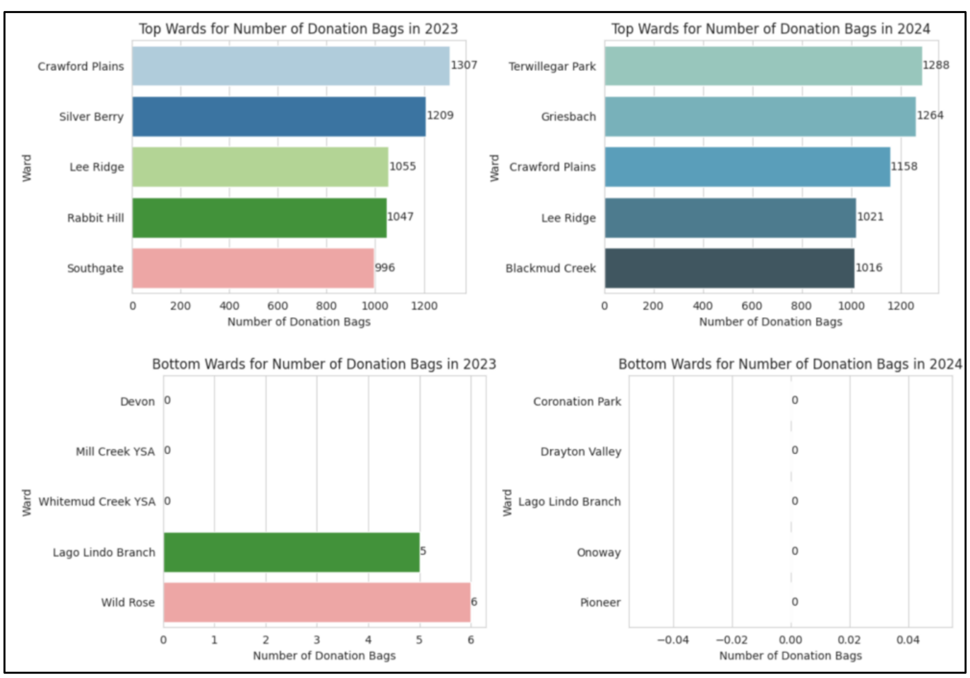
Compared to the 2023 food drive, the 2024 results showed a decrease in several key metrics: the number of donation bags, number of volunteers, number of houses, and average time spent per route decreased by 0.4%, 38.17%, 38.17%, and 6.67%, respectively.

Figure 4: Top and Bottom Three Stakes of 2023 and 2024



The top stakes in 2024 remained largely consistent with 2023, with Gateway, Bonnie Doon, Riverbend, Edmonton North, and YSA leading the rankings. However, Riverbend and Bonnie Doon swapped positions, indicating a slight shift in their relative performance between the two years

Figure 5: Top and Bottom Five Wards of 2023 and 2024

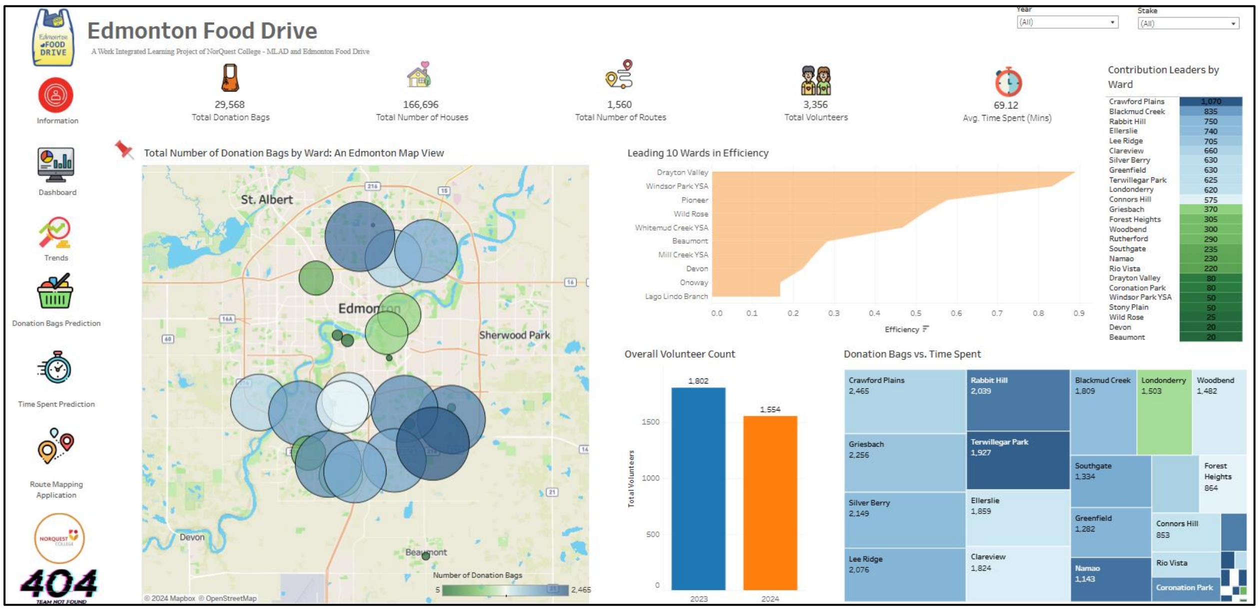


In 2024, Crawford Plains stayed in the top 5, just like in 2023. Some new wards, like Terwillegar Park and Griesbach, joined the top ranks. On the other hand, wards like Coronation Park, Drayton Valley, and Pioneer moved into the bottom 5 in 2024, replacing last year's bottom wards like Devon and Mill Creek YSA.

Data Visualization:

We created interactive visualizations using Tableau to make our EDA findings easy to understand. These visualizations allow users to explore the data and gain insights through dynamic charts and maps. The dashboard includes various charts and maps that present the key aspects of our analysis in a simple and clear way. Below, we list the visualizations included in the dashboard that help support our overall analysis.

Figure 6: Edmonton Food Drive 2024 Dashboard using Tableau



Here are the features of the dashboard.

- *KPI Card for Key Features*: Displays the total number of donation bags, houses, routes, volunteers, and average time spent, based on the selected criteria.
- *Total Number of Donation Bags by Ward*: This map of Edmonton shows the distribution of donation bags across different wards, providing a clear comparison of how they are spread throughout the city.
- *Leading 10 Wards in Efficiency*: Highlights the top 10 wards with the highest efficiency, showcasing their performance across key metrics.
- *Overall Volunteer Count*: A bar chart comparing volunteer counts over different years, offering insights into trends and changes over time.
- *Contribution Leaders by Ward*: A heatmap showing the contributions from each ward, using color gradients to highlight the areas with the highest and lowest contributions.
- *Donation Bags vs. Time Spent Chart*: A visualization comparing the number of donation bags to the time spent, providing insights into the efficiency of the donation process.

Dashboard Link: [Link]

Machine Learning Model:

Our team was given the task of predicting the total number of donation bags and the time spent for each ward over the next five years. To do this, we started by improving the data with feature engineering, then split it into training and testing sets. We prepared the data further by encoding categorical variables and normalizing numerical features to make everything consistent. Next, we built and tested six different machine learning models for each prediction to find the most accurate ones. The results of these models are shared below, showing how well they performed.

Table 5: Performance Metrics for Models Predicting Total Donation Bags

Model	MSE	RMSE	MAE	R^2	Adjusted R^2
Linear Regression	3393.986256	58.257929	26.828851	-0.100185	-0.168338
Polynomial Regression	49.838645	7.059649	2.388835	0.983844	1.146869
Decision Tree Regression	2356.665557	48.545500	8.232945	0.236070	0.188747
Random Forest Regression	1990.524740	44.615297	8.457754	0.354757	0.314786
Gradient Boosting Regression	2144.987415	46.314009	8.164502	0.304687	0.261615
K-Nearest Neighbors Regression	3092.228686	55.607811	17.474875	-0.002368	-0.064461

Based on the results, the best model for predicting total donation bags is *Polynomial Regression*, as it delivers the lowest RMSE (7.059649) and MAE (2.388835) while achieving the highest R^2 score (0.983844), indicating excellent accuracy and consistency.

Table 6: Performance Metrics for Models Predicting Time Spent

Model	MSE	RMSE	MAE	R^2	Adjusted R^2
Linear Regression	1.583989	1.258566	0.917151	0.075887	2.771216
Polynomial Regression	0.708581	0.841772	0.634814	0.586608	1.014787
Decision Tree Regression	0.192435	0.438674	0.356527	0.887732	1.215181
Random Forest Regression	0.216073	0.464836	0.377927	0.873941	1.241613
Gradient Boosting Regression	0.256885	0.506838	0.391840	0.850131	1.287249
K-Nearest Neighbors Regression	0.278344	0.527583	0.394887	0.837612	1.311244

For predicting time spent, the Decision Tree Regression model stands out as the best among the listed options. It achieves the lowest RMSE (0.438674) and MAE (0.356527), coupled with a high positive R^2 (0.887732) and Adjusted R^2 (1.215181), indicating superior accuracy and a strong fit to the data compared to the other models.

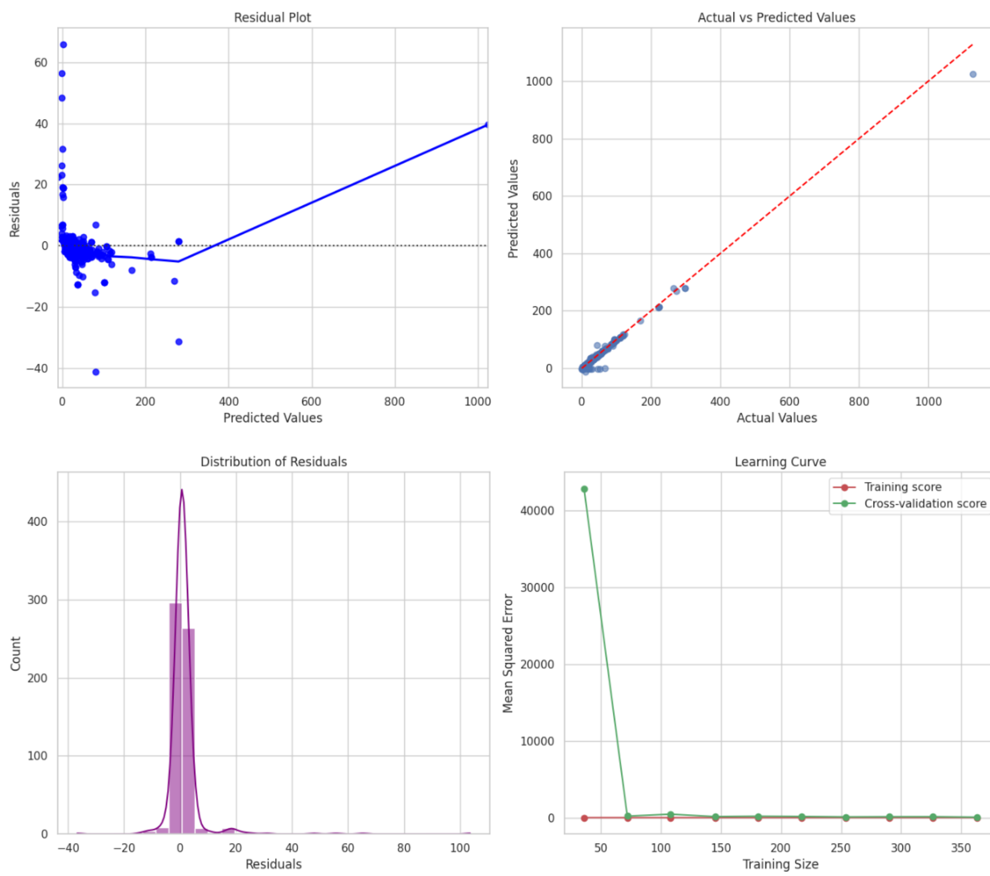
Link to Code: [Link]

Model Optimization:

For the Polynomial Regression model used to predict total donation bags, we opted not to perform additional tuning to avoid the risk of overfitting. Since the metrics are already acceptable, with an R^2 score of 0.98, further increasing model complexity could lead to diminished generalization and overfitting the training data.

Here are some useful graphs that can help visualize the behavior of the model:

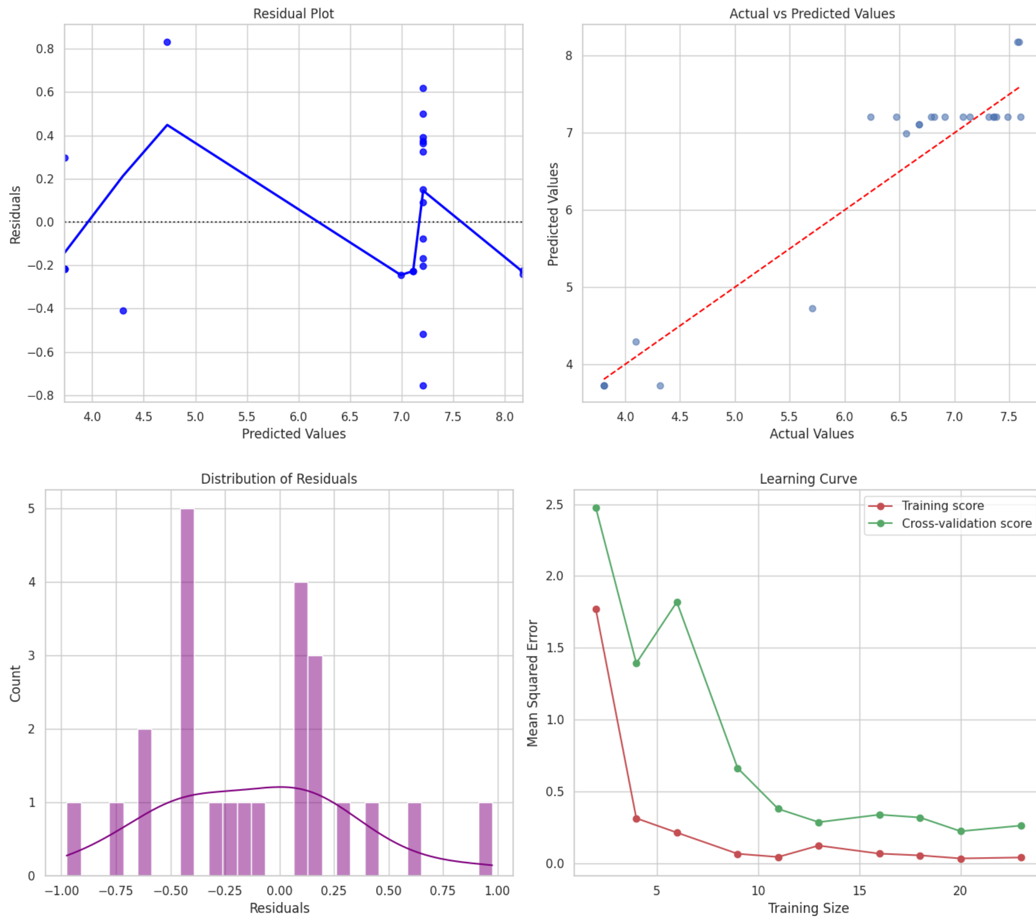
Figure 7: Visualizing the Behavior of Polynomial Regression: Key Plots



We performed hyperparameter tuning on the Decision Tree Regression model for predicting time spent, but it did not result in significant improvements. The tuned model achieved a Mean Squared Error (MSE) of 0.2041, Root Mean Squared Error (RMSE) of 0.4517, Mean Absolute Error (MAE) of 0.3652, R-squared (R^2) of 0.8810, and Adjusted R-squared of 1.2282.

Here are some useful graphs that can help visualize the behavior of the model:

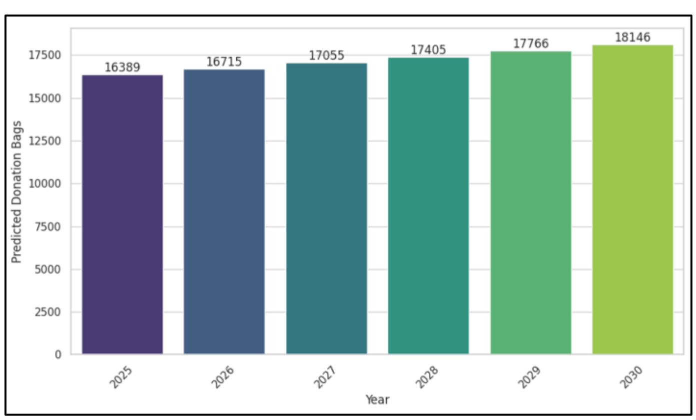
Figure 8: Visualizing the Behavior of Decision Tree Regressor: Key Plots



Advanced Analysis:

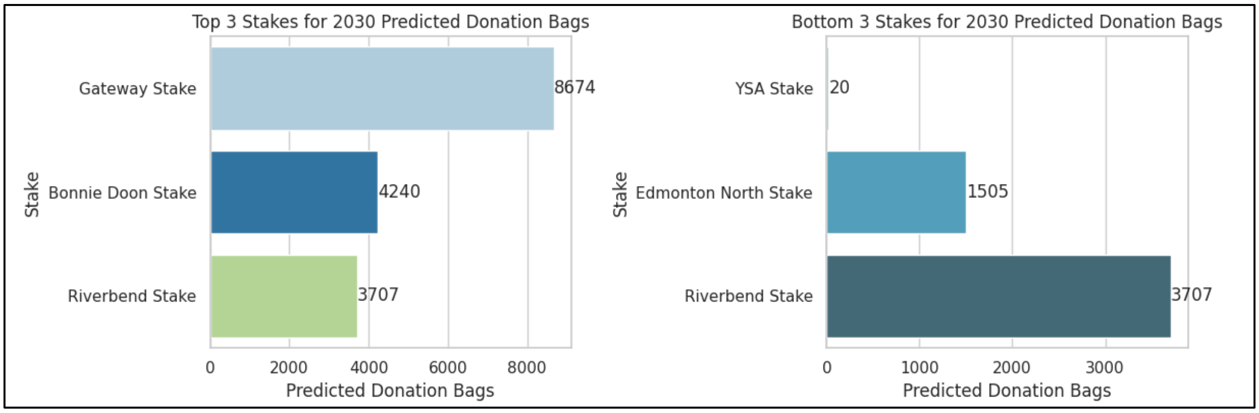
We used the Polynomial Regression and Decision Tree models to predict the number of donation bags and time spent per ward for each year until 2030. Below are some key insights based on the predicted values.

Figure 10: Projected Total Number of Predicted Donation Bags for the next 6 years



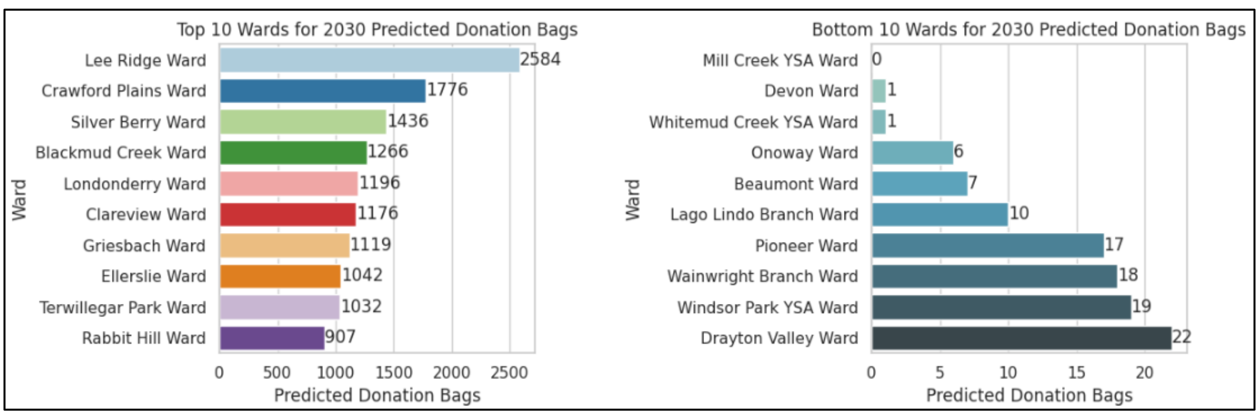
The predicted number of donation bags for the next six years shows a steady increase. From 2025 to 2030, the total number of donation bags is expected to grow, starting at 16,389 in 2025 and reaching 18,146 by 2030, reflecting an average annual increase of around 2%.

Figure 11: Six-Year Outlook of Donation Bags: Top and Bottom 3 Stakes



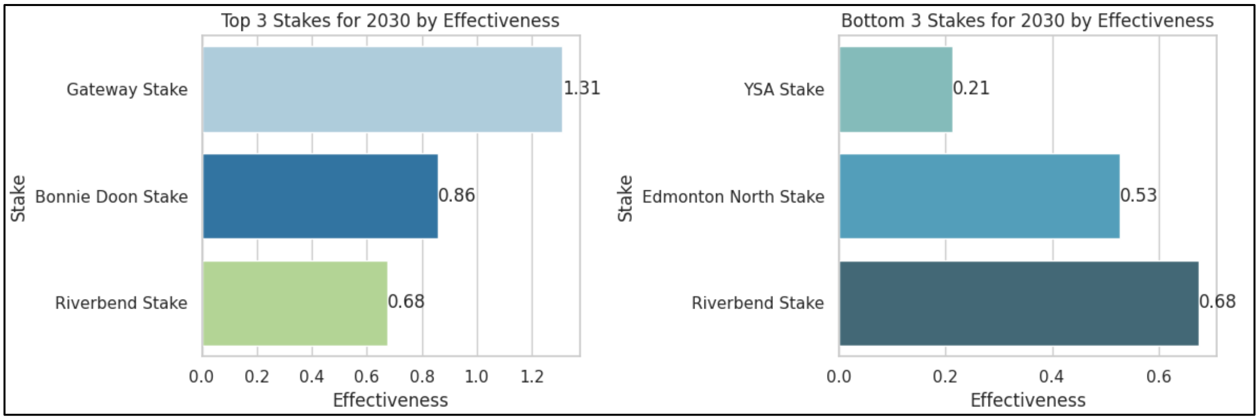
The six-year outlook for donation bags reveals the top and bottom-performing stakes. The top three stakes, which are expected to contribute the most to donation bags, are Gateway, Bonnie Doon, and Riverbend. On the other hand, the bottom three stakes, contributing fewer donation bags, are YSA, Edmonton North, and Riverbend.

Figure 12: Six-Year Outlook of Donation Bags: Top and Bottom 10 Wards



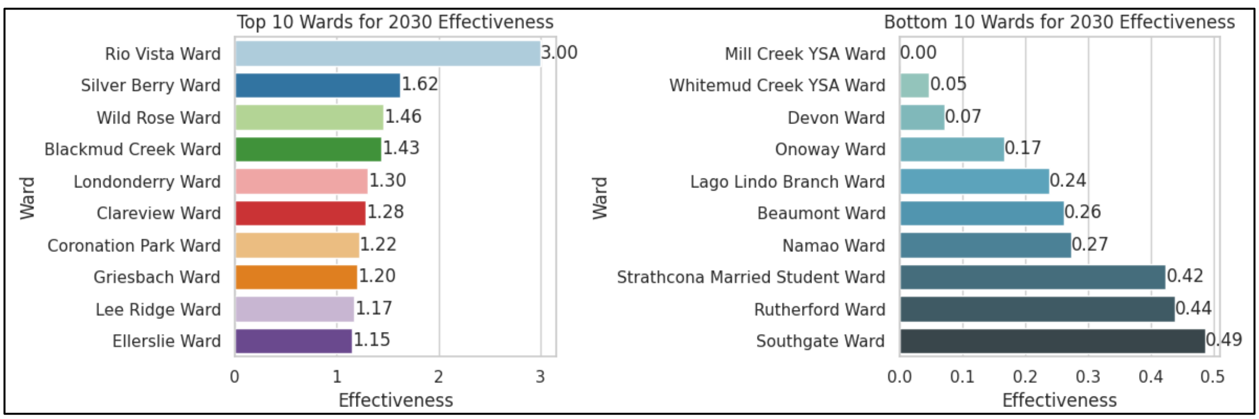
The six-year outlook for donation bags reveals the top and bottom-performing wards. The top 10 wards expected to contribute the most donation bags are Lee Ridge, Crawford Plains, Silver Berry, Blackmud Creek, Londonderry, Clareview, Griesbach, Ellerslie, Terwillegar, and Rabbit Hill. On the other hand, the bottom 10 wards, which are projected to contribute fewer donation bags, include Mill Creek YSA, Devon, Whitemud Creek YSA, Onoway, Beaumont, Lago Lindo, Pioneer, Wainwright, Windsor Park, and Drayton Valley. These insights show a notable variation in donation contributions across different wards.

Figure 13: Six-Year Outlook of Effectiveness: Top and Bottom 3 Stakes



The top 3 stakes with the highest effectiveness (i.e., they are expected to generate the most donation bags per unit of time spent) are Gateway, Bonnie Doon, and Riverbend. On the other hand, the bottom 3 stakes with the lowest effectiveness, meaning they are expected to have the least donation bags per unit of time spent, are YSA, Edmonton North, and Riverbend.

Figure 14: Six-Year Outlook of Effectiveness: Top and Bottom 10 Wards



The six-year outlook for effectiveness, defined as the predicted donation bags divided by the predicted time spent, highlights the top and bottom-performing wards. The top 10 wards with the highest effectiveness, meaning they are expected to generate the most donation bags per unit of time spent, are Rio Vista, Silver Berry, Wild Rose, Blackmud Creek, Londonderry, Clareview, Coronation Park, Griesbach, Lee Ridge, and Ellerslie. These wards are predicted to be more efficient in converting time spent into donation bags.

In contrast, the bottom 10 wards with the lowest effectiveness, meaning they are expected to have the least donation bags per unit of time spent, include Mill Creek YSA, Whitemud Creek, Devon, Onoway, Lago Lindo Branch, Beaumont, Namao, Strathcona Married Student, Rutherford, and Southgate. These wards are projected to require more time to achieve similar numbers of donation bags, reflecting a lower efficiency in their donation efforts.

Based on the insights gathered from the EDA and advanced analytics, here are our recommendations:

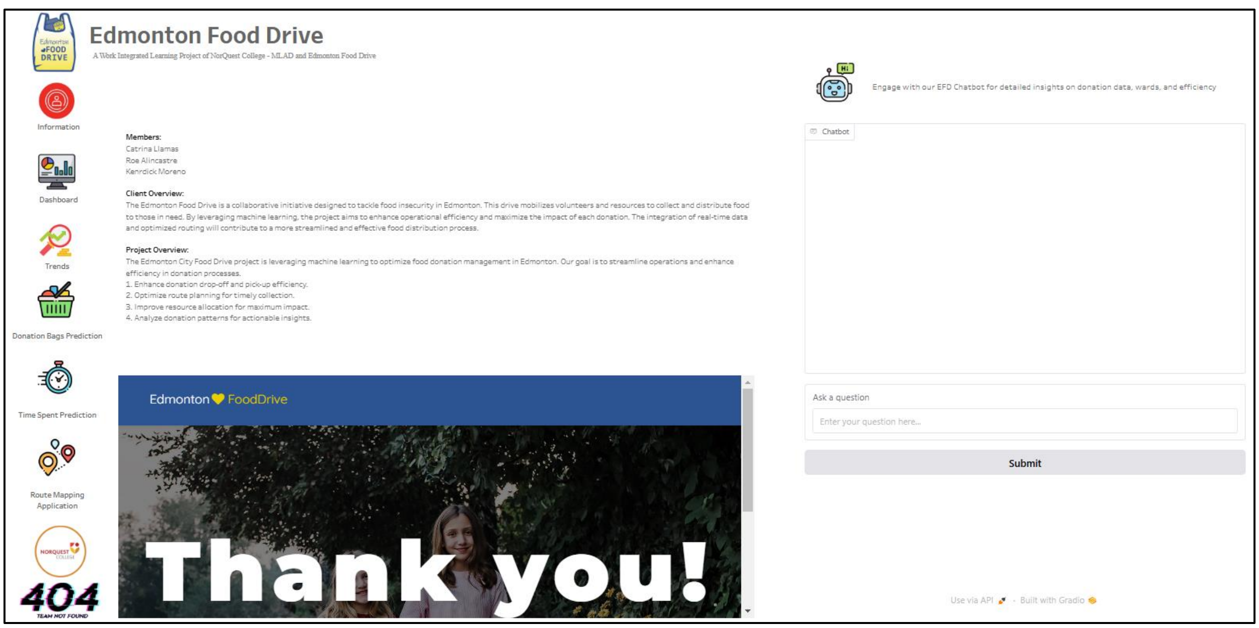
- Volunteers should be allocated more resources and time to the top-performing wards, such as Lee Ridge, Crawford Plains, and Silver Berry, as they are expected to generate the most donation bags.

- Focus on strengthening the donation efforts in the bottom 10 wards, including Mill Creek YSA, Devon, and Whitemud Creek, which are projected to contribute fewer donation bags, indicating areas needing improvement.
- Stakeholders in YSA, Edmonton North, and Riverbend should focus on improving the efficiency of their donation drives, as they have lower effectiveness in generating donation bags per unit of time spent.
- Greater attention should be given to the Gateway, Bonnie Doon, and Riverbend stakes, which are expected to be the most effective in converting time into donation bags, but still need further support to maintain their success.
- To enhance overall effectiveness, a more balanced allocation of volunteers should be considered, with a focus on both improving the performance of lower-performing areas and maintaining the momentum in top-performing wards and stakes.

Model Deployment:

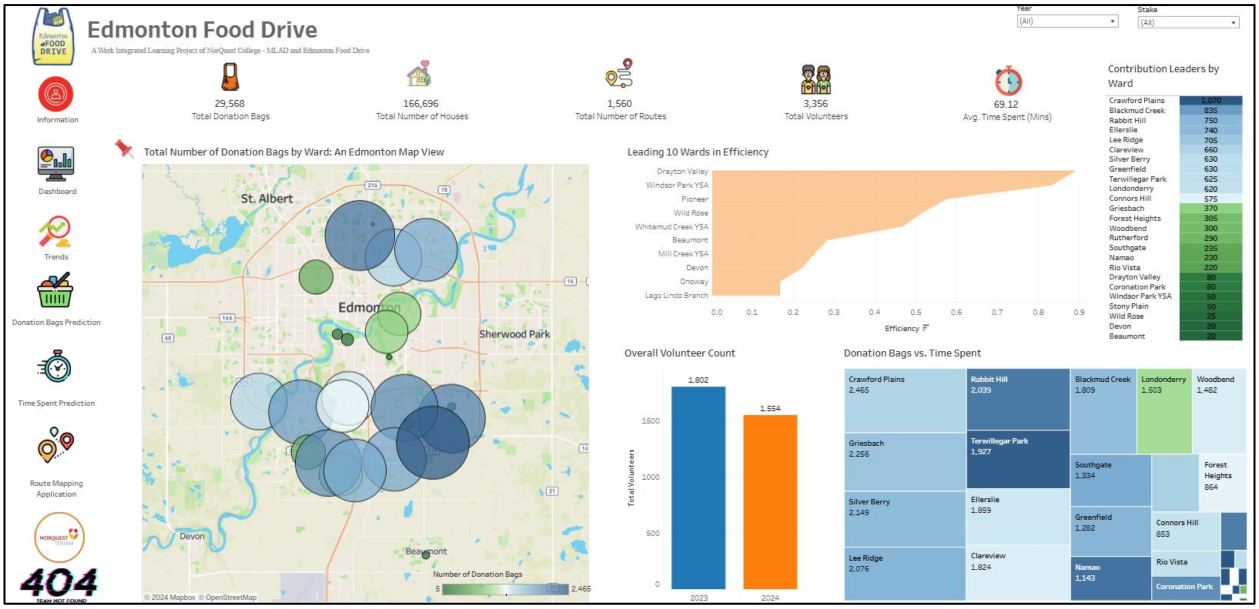
Our application is divided into six sections: the Information Page, Dashboard Page, Trends Page, Donation Bags Prediction Page, Time Spent Prediction Page, and Route Mapping Application Page. Each page has a distinct feature designed to deliver specific insights and valuable information to its users, ensuring a comprehensive experience. Together, these sections allow users to easily navigate through different functionalities, making data-driven decisions more accessible and efficient.

Figure 15: EFD Application: Information Page Overview



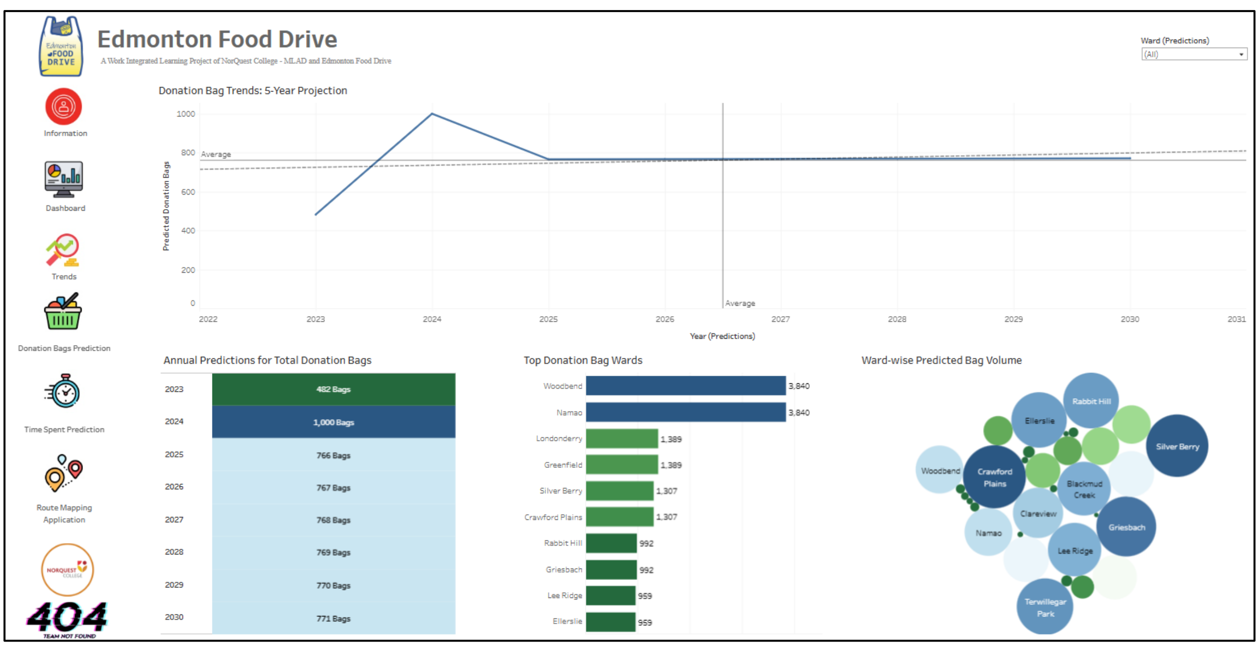
The Information Page acts as an introductory guide, providing details about the project team, client overview, and project scope. It also features an embedded link to the Edmonton Food Drive website. Additionally, a chatbot using DialogGPT from Microsoft is available to provide quick answers and insights, enhancing user interaction.

Figure 16: EFD Application: Dashboard Page Overview



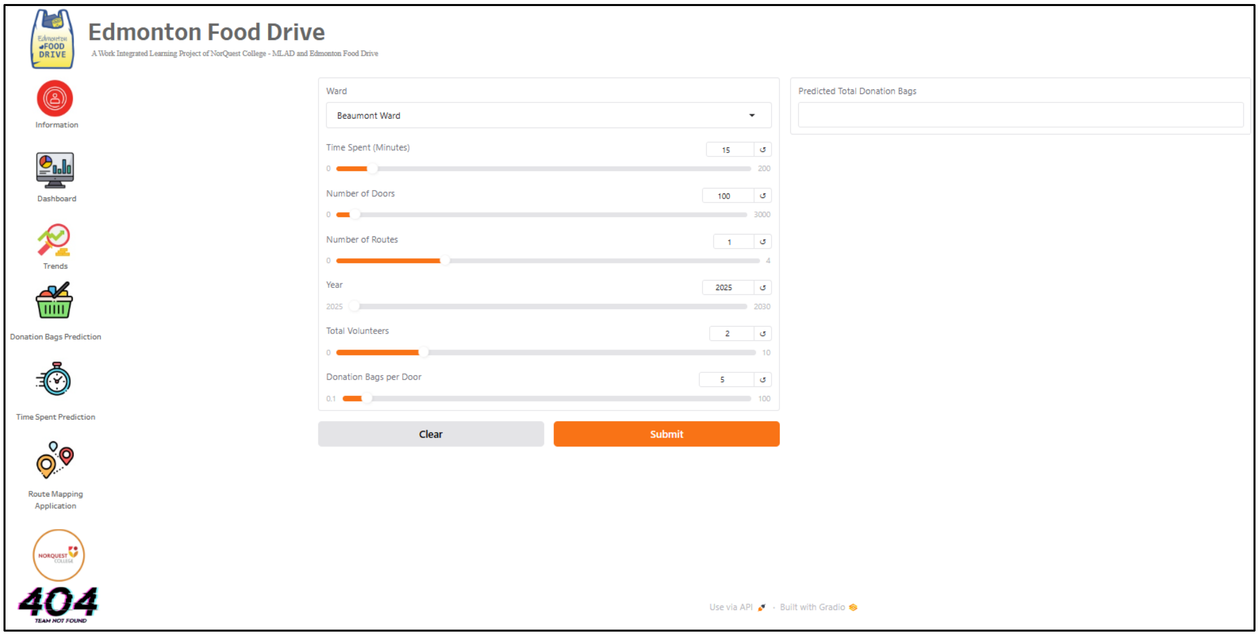
The Dashboard Page presents key metrics in real-time, featuring informative charts and visualizations that support our exploratory data analysis (EDA). It offers valuable insights to help users track and interpret important trends and performance indicators.

Figure 17: EFD Application: Trends Page Overview



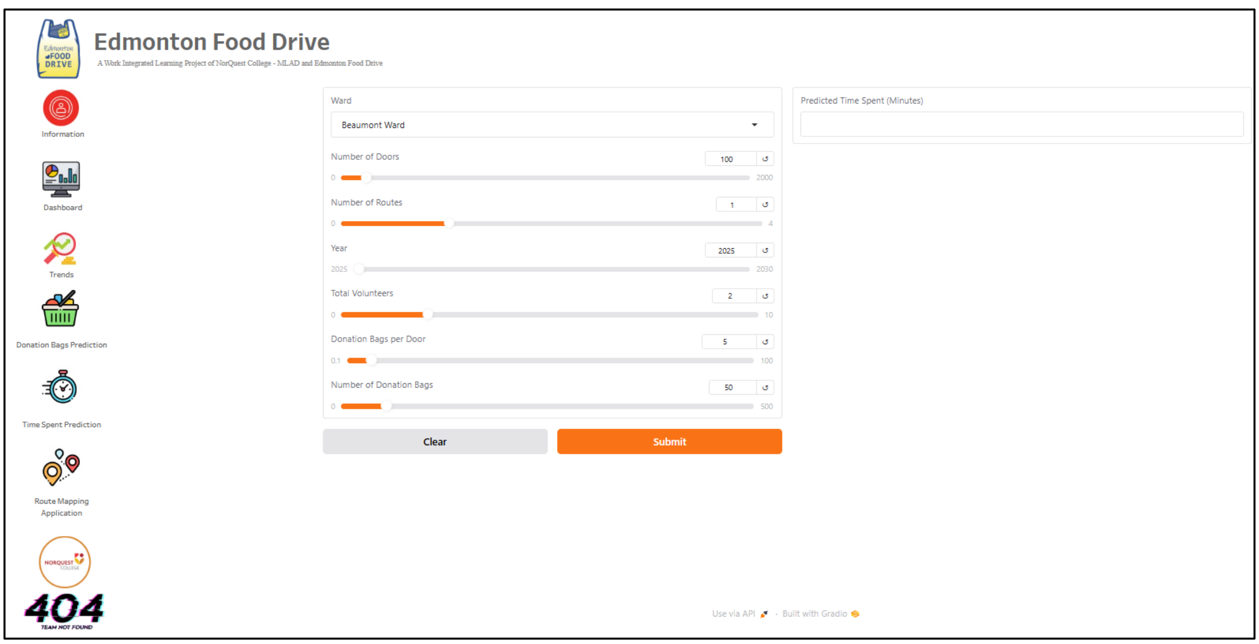
The Trends Page displays data on predicted donation bags and includes various charts to facilitate easy interpretation of the trends and patterns in the data. These visualizations help users quickly understand how donation bag numbers are expected to change over time.

Figure 18: EFD Application: Donation Bags Prediction Page Overview



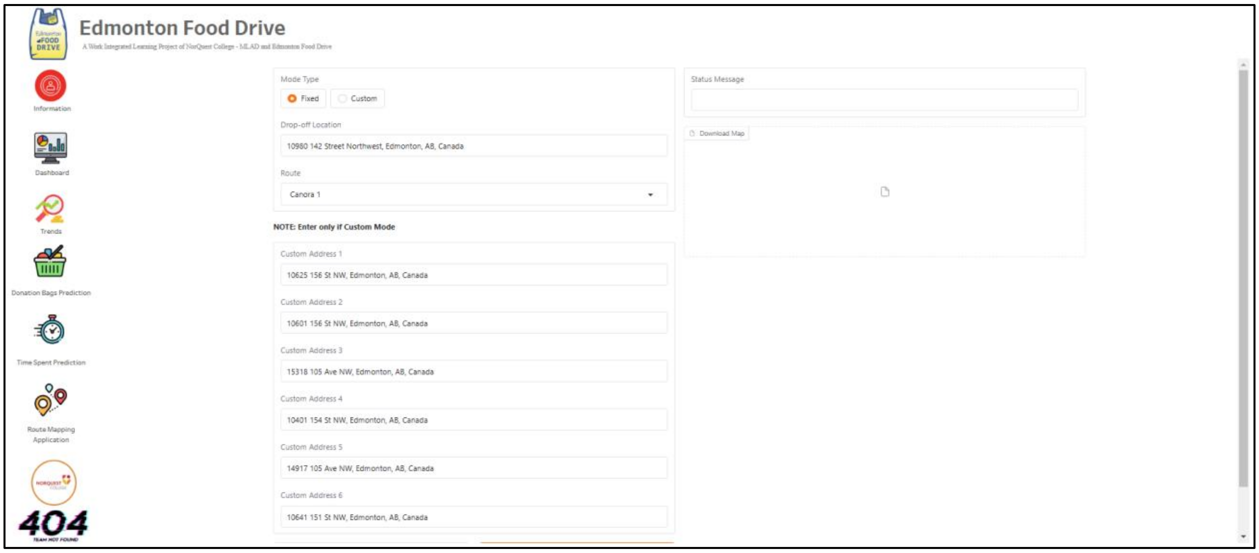
The Donation Bags Prediction Page allows users to predict the total number of donation bags based on parameters such as ward, time spent, number of doors, and other relevant factors. This feature provides users with a dynamic tool to estimate donation bag numbers under various scenarios.

Figure 19: EFD Application: Time Spent Prediction Page Overview



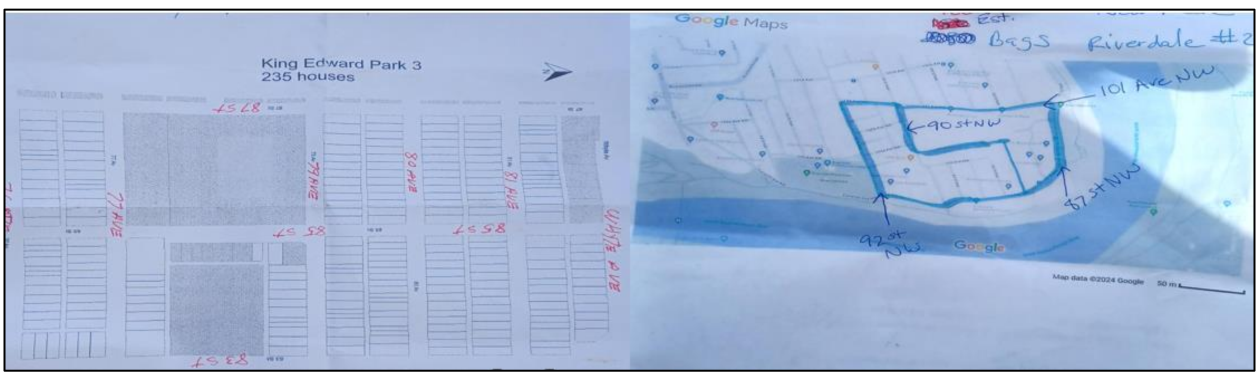
The Time Spent Prediction Page enables users to estimate the total time spent based on parameters such as ward, total donation bags, number of doors, and other relevant factors. This feature offers a dynamic tool to forecast time requirements under various conditions, helping users plan and allocate resources effectively.

Figure 20: EFD Application: Route Mapping Application Page Overview



The Route Mapping Application is an initiative developed by our team in response to the client's repeated challenges with generating maps. The existing process involves printing a selected portion of the Edmonton map, manually highlighting the route, and then distributing the map to volunteers. This manual process is time-consuming and inefficient. To address this, our application streamlines route generation and visualization, improving efficiency and reducing errors. Below are some images showing the manually printed maps previously used by the client.

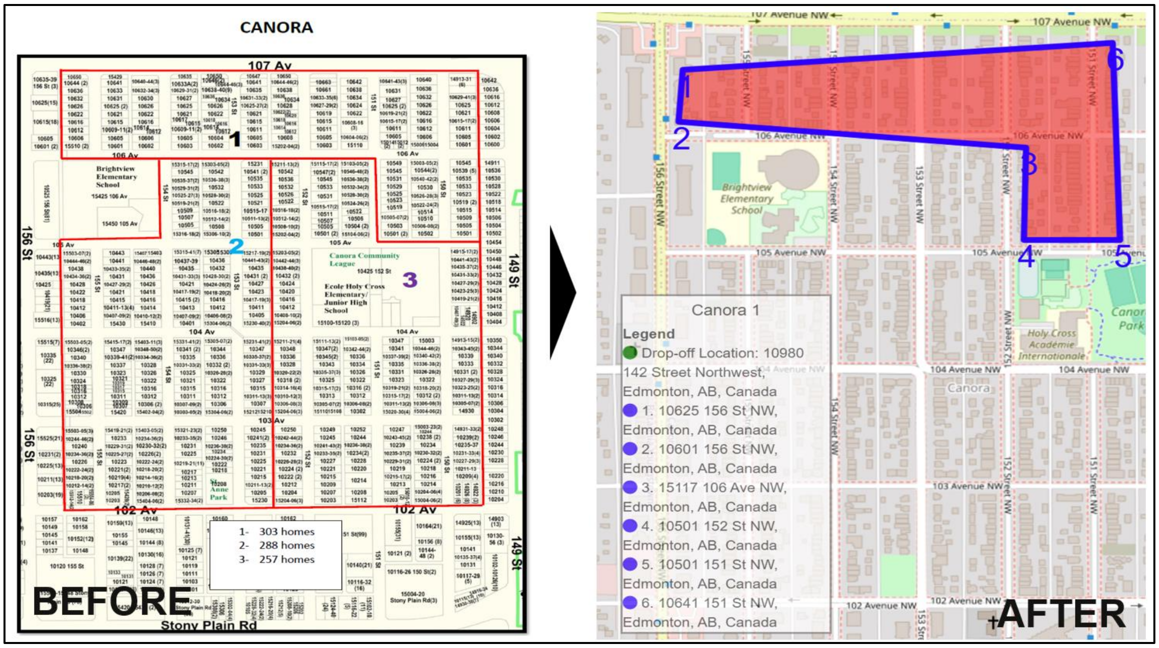
Figure 21: Manually Printed Maps for Route Planning



The application offers two modes: Fixed Mode and Custom Mode. The Fixed Mode aims to digitize the map generation process for our client, streamlining their workflow. The Custom Mode, on the other hand, is designed for long-term planning, generating maps based on identified hot zones to optimize route efficiency.

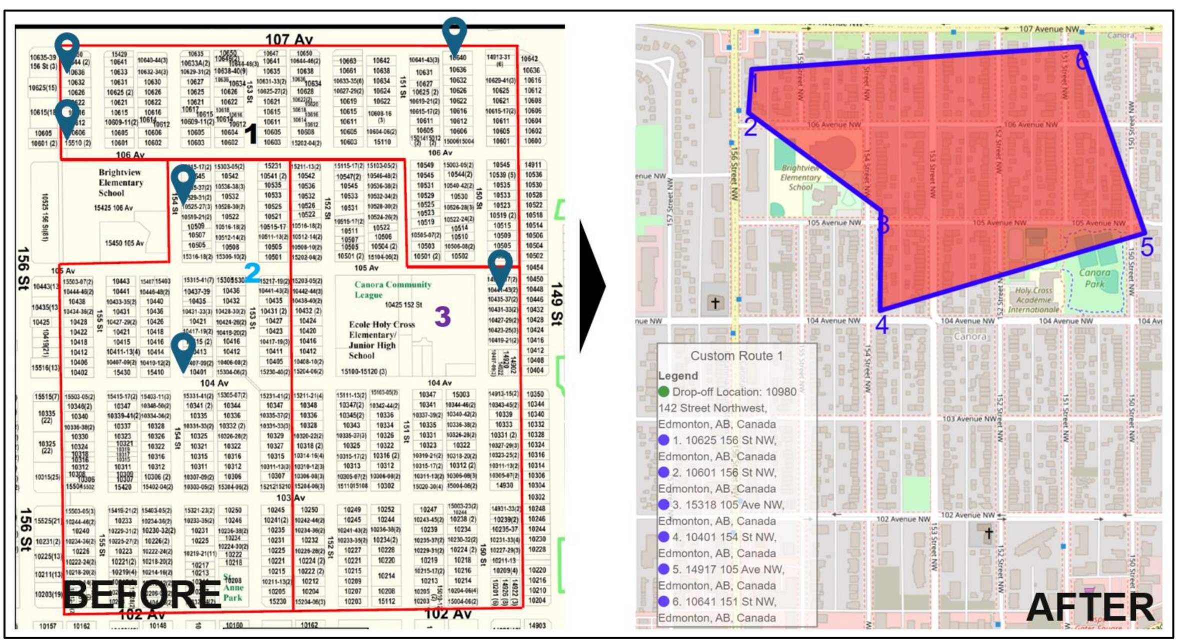
To generate maps in Fixed Mode, the client only needs to select the desired ward and route, click "Submit," download the generated map, and then easily email it to the volunteers. This streamlined process eliminates the need for manual map creation, saving time and effort. The provided image shows the before and after results of generating maps using Fixed Mode, highlighting the efficiency and ease of the new approach.

Figure 22: Before and After: Map Generation Comparison Using Fixed Mode



The provided image demonstrates how the map can now be digitally generated with accurate information, providing volunteers with clear details. On the other hand, the Custom Mode is designed for route mapping enhancement. The process involves identifying hot-zone addresses, inputting the required parameters into the application, generating the map, downloading it, and then distributing it to the volunteers. Hot-zone addresses refer to houses that consistently donate bags to our client, helping to streamline the donation collection process.

Figure 23: Before and After: Map Generation Comparison Using Custom Mode



In the provided image, the pins represent the hot zone addresses from Routes 1, 2, and 3. The client needs to input these six addresses, and the application will calculate the optimal route order based on the distance between them. This ensures that volunteers follow the most efficient path, saving time. Now, volunteers no longer need to cover all three routes; instead, they can focus on specific portions of each route, improving efficiency and streamlining the donation collection process.

Stakeholder Engagement:

Stakeholders were informed about key information such as number of donation bags, number of volunteers, number of houses, and average time spent for both 2023 and 2024 Edmonton Food Drive events. The information shows the percentage change from 2023 to 2024. While the figures suggest a decrease for all key metrics, the data these insights were based on are not indicative of the overall food drive event. Only select parts of Edmonton were covered by this year's data collection and must therefore be considered. The client stakeholder further verified this by mentioning that based from the Edmonton Food Bank and other internal sources, their figures were an increase from last year's records.

Additionally, stakeholders were shown which wards and stakes garnered the most donation bags in 2023 and 2024. The clients expressed appreciation for this information as it could help them in assigning volunteers to specific areas with records of high donation volume. The premise of a route mapping tool was also appreciated since this will be a significant upgrade from their traditional workflow of using paper maps. An interactive dashboard was also presented to stakeholders wherein they could further inspect the data, view 5-year projections, and in the final dashboard iteration; stakeholders will be able to interact with other features such as a chatbot, an app that predicts donation amount, and a route mapping tool.

Challenges Faced:

The primary challenge in this project comes from the data collection process; developing and deploying a machine learning model heavily relies on data which makes this phase significant. With the manpower available, not all participating areas of Edmonton were covered in the data collection. The partial coverage consists of the drop-off locations in Bearspaw, Londonderry, Riverbend, Gateway, and Bonnie Doon.

Due to limited data available, the predicted growth on donations resulted in figures that were not to the clients' expectations. According to the client's internal sources, their data indicated a general increase in 2024 in comparison to 2023. Data refilling was performed to make the data higher for 2024 and address the inconsistency between predicted and actual data.

Lessons Learned:

In this project, the group received first-hand experience of working with real-life data and client stakeholder group and learned ways to communicate the process and findings of this project. High-level communication is needed to convey the significance and value of this work to the stakeholders. This entails a balance between technical discussions on data preparation and model development and business storytelling that must be achieved. There is also the aspect of understanding the clients' needs and incorporating that to the machine learning problem.

By working on this project, the group improved their skills across the phases of data acquisition, data pre-processing, exploratory data analysis, machine learning model development and evaluation, model results and trend analysis, interactive dashboard development, and route mapping app development. Along with this, the group learned to anchor their decision-making on domain knowledge and client feedback.

Future Recommendations:

To enhance model performance, it is recommended that this project is continued for the following years to come. This is because current data volume of two years is not enough. Furthermore, there is partial data collection coverage due to the manpower available. Either a rework of the collection process or increased data access to other locations is highly suggested.

The group also recommends further development to go into the route mapping tool to transform it into a route optimization tool. In its current state, the route mapping tool has two modes: a fixed and a custom mode; both of which rely on fixed map information and user-inputted location data, respectively. The current mapping application digitalizes the system of generating maps but is not yet capable of generating maps on its own. In the future, a route optimization application would be able to predict hot spots in a given area and generate an efficient route. This optimization tool would alleviate event organizers the tedious process of collecting data on high-yielding locations and entering those into the mapping tool.

Another recommendation is further development for the chat bot application. The chat bot in current development can respond to simple inquiries on topics such as total donations for 2023 to 2024 and top and bottom stakes and wards in terms of efficiency and most amount of donations. The chat bot has the potential to hold more complex conversations with users in a way that users can converse with the bot as their assistant for event planning and organization.

Impact on the Community:

The Edmonton Food Drive 2024 project makes a positive difference in helping the community. By being involved in data collection and analysis, the project can help improve the way food donations are collected and managed. By using technology like predictive models and interactive tools, the project can save time and resources, e.g. volunteers are made to focus on the most important areas, so that more food is collected. The project also helps organizers make better decisions using clear data.

Project Conclusion:

The project successfully achieved its goals of recommending improvements in the food donation process in the Edmonton Food Drive. Tools to predict donation trends and time requirements were introduced, helping volunteers and organizers plan better. The route mapping application simplifies volunteer coordination and saves significant effort compared to the traditional manual processes. Additionally, interactive dashboards make it easier for stakeholders to understand and analyze the data, leading to better decision-making. Overall, the project streamlines operations, and contributes to a more effective and efficient food donation drive.

Acknowledgments:

We would like to express our deepest gratitude to Professor Uchenna for her invaluable guidance and support throughout this project. Her insights and expertise were instrumental in shaping the direction of our work and ensuring its success.

A special thanks goes to our classmates, whose efforts in collecting data and providing their input were essential to the success of this project.

We also extend our heartfelt thanks to the stakeholders of the Edmonton Food Drive for their collaboration and trust. Their feedback and participation provided us with the necessary context and

motivation to deliver a solution that meets the community's needs. Additionally, we are grateful to all the volunteers and sponsors who contributed their time and resources to make this project possible. Your efforts have been crucial in driving positive change within the community.

Appendices:

Data Collection Form

Project Dashboard

Project Notebook

Project Presentation

References

City of Edmonton. (n.d.). Neighbourhood dataset.

Edmonton Food Drive. (2023). Donation data for 2023.

Edmonton Food Drive. (2024). Donation data for 2024.

Tableau. (2024). Interactive data visualization software.