CMPT 3830: Project Proposal Template

1. Project Title:

PriceVision: Vehicle Pricing Optimization Tool Using Regression Based on Key Vehicle Characteristics

2. Project Overview:

- Objective:

To develop a predictive pricing model using machine learning that estimates the optimal price for vehicles based on characteristics like year, make, model, and mileage, aimed at optimizing sales strategies for Go Auto.

- Background:

This project is part of a broader effort by Go Auto, one of the largest car dealerships in Canada, to leverage machine learning to improve its business operations. The dataset includes information about vehicles (both sold and listed) across multiple dealerships, utilizing APIs from the Canadian Black Book (CBB). Using this data, we aim to create a pricing model that helps Go Auto to maximize sales while maintaining competitiveness.

- Scope:

This section outlines the essential tasks that must be accomplished to achieve the primary objectives of the project, while staying aligned with the business goals that initially drove the need for its development. The scope statements define the boundaries of these objectives and business drivers, providing clear criteria for determining when the project can be considered complete.

In Scope:

- Develop a regression-based machine learning model to predict vehicle prices, using features such as year, make, model, and mileage.
- Collect, clean, and preprocess vehicle sales data, ensuring that key features are optimized for model training.
- Handle outliers, validate data accuracy, and ensure reliable predicted values by identifying and addressing anomalies.
- Create interactive visualizations and dashboards using tools like PowerBI or Google Looker Studio to showcase key insights, including predicted vs. actual prices.
- Deliver actionable insights aligned with Go Auto's business goals to optimize pricing strategies and improve sales performance.

Out-of-Scope:

- Predict vehicle sales volume or other unrelated vehicle attributes, such as transmission type.
- Integrate the developed pricing model with Go Auto's internal software or sales platforms.
- Test or deploy the model in a real-time production environment; this will not be part of the project.
- Provide ongoing support, maintenance, or updates to the model after the project's completion.
- Create other predictive models, such as those for sales volume or customer behavior, which is beyond the scope of this project.

3. Project Deliverables:

List the specific outputs or deliverables expected from the project.

1. **Deliverable 1: Team Formation and Initial Client Meeting:** Formation of the team, and an initial meeting with Go Auto to align project goals, expectations, and scope.
2. **Deliverable 2: Dataset Exploration and Preprocessing:** Thorough exploration of the dataset, cleaning, and preparation for model training, ensuring data quality and readiness.
3. **Deliverable 3: Team Charter Submission:** Submission of the Team Charter, outlining team roles, responsibilities, communication protocols, and project goals.
4. **Deliverable 4: Project Proposal Submission:** Submission of the Project Charter, detailing key deliverables, project scope (including in-scope and out-of-scope items), risks, timeline, etc.
5. **Deliverable 5: Machine Learning Model Development:** Creation of a regression model for predicting vehicle pricing using key features such as year, make, model, and mileage.
6. **Deliverable 6: Model Evaluation:** Evaluation of the machine learning model using metrics like Mean Squared Error to assess its accuracy and performance.
7. **Deliverable 7: Project Demo 1:** Presentation of dataset exploration, preprocessing, and initial model development, with client feedback for further refinement.
8. **Deliverable 8: Phase 1 Project Report:** A detailed report on dataset exploration, model development, and initial insights, including visualizations and recommendations.

9. **Deliverable 9: Insights and Recommendations:** Key insights derived from the model and recommendations on how Go Auto can apply them to improve pricing strategies.
10. **Deliverable 10: Project Demo 2:** Final model demonstration, showcasing improvements, performance evaluation, and visualizations, with client feedback.
11. **Deliverable 11: Phase 2 Project Report:** Comprehensive report covering the entire modeling process, challenges, solutions, insights, and recommendations.
12. **Deliverable 12: Data Visualizations:** Creation of interactive dashboards and visualizations of pricing predictions, key metrics, and insights using PowerBI or Looker Studio.
13. **Deliverable 13: Final Report Submission:** A complete project report, summarizing findings, lessons learned, and final recommendations for Go Auto.

4. Project Timeline:

Milestone	Completion Date
Milestone 1: Finish cleaning, processing, and encoding data.	Oct 06]
Milestone 2: Have a basic ML model for the data deployed.	[Nov 05]
Milestone 3: Finish adjusting model and deliver findings.	[Dec 10]

6. Project Plan:

- Tasks and Activities:

Provide a breakdown of the tasks needed to complete the project.

Task Category	Task	Assigned To	Progress	Start Date	End Date
Initiation and Planning	Set up team	Everyone	100%	9/10/2024	9/13/2024
	Select problem statement	Everyone	100%	9/10/2024	9/13/2024
	Dataset exploration	Everyone	70%	9/10/2024	9/25/2024
	Team Charter template	Everyone	100%	9/10/2024	9/13/2024
	Project proposal	Everyone	60%	9/13/2024	9/24/2024
	Create schedule	Everyone	60%	9/14/2024	9/23/2024
Execution and Evaluation	Model evaluation	Everyone	0%	9/22/2024	10/8/2024
	Insights and Recommendations	Everyone	0%	9/24/2024	10/8/2024
	Evaluate progress	Everyone	20%	9/10/2024	12/10/2024
	Testing and validation	Everyone	0%	10/8/2024	10/20/2024
	Demo 1	Everyone	0%	10/8/2024	10/8/2024
	Phase 1 report	Everyone	0%	10/8/2024	10/15/2024
	Demo 2	Everyone	0%	11/5/2024	11/5/2024
	Phase 2 report	Everyone	0%	11/5/2024	11/19/2024
	Final presentation	Everyone	0%	11/19/2024	12/10/2024

7. Resources Required:

List the resources necessary for completing the project.

Resource	Description	Estimated Cost

1. | Resource 1: Google Collab | Allows users to write and execute Python code through their browser and work with others on the same file. This includes importing and pip installing other python modules, allowing for the use of data editing with modules such as pandas. $0 |

2. | Resource 2: Pandas, Scikit-learn, etc. | Various python modules that are vital to processing the data and deploying a ML model using said data. $0 |

3. | Resource 3: Various coding documentations | Each module used has its own documentation that can be used to gain insight into different methods or meanings into the python module it relates to. $0 |

4. | Resource 4: PowerBI and Google Looker | Necessary software used for data visualizations. $0

5. | Resource 5: Laptops | Personal computers that the project will be completed on. | $0 |

8. Risk Management Plan:

Identify potential risks and strategies to mitigate them. See the table below for reference. Generate a similar table with the header mentioned below and complete Risk Management Plan.

Risk	Likelihood	Impact	Mitigation Strategy
Technical delays	Medium	High	Build in extra time for each task and keep a detailed schedule to track progress.
Risk of encountering "unknowns" that may arise during its execution	Medium	Moderate	Do a deep dive into the data to clean it up and fix any errors before building the model.
Model Performance Not Good Enough	Medium	High	Keep improving the model, try different techniques, and use cross-validation to check accuracy.
Lack of Expertise for Analysis	High	High	Do additional research, ask for help from the instructor or peers,

			and look up online tutorials.
Not Enough Data for Good Results	Low	High	Explain the limitations during the Demos.
Coding Does not Work	Medium	High	Test code in small chunks, fix issues as you go, and regularly save working versions of the code.
Dashboard Is Poorly Designed	Medium	Moderate	Get feedback early on the dashboard and make changes so it is easy to use and understand.
Not Meeting Client Expectations	Medium	High	Have regular meetings with the client to get feedback and make sure you're on the right track.
Unequal Collaboration in the Group	High	High	Clearly assign tasks to each team member, keep communication open, and do peer evaluations to ensure everyone is contributing.

9. Budget:

Provide an overview of the total budget for the project. If your project does not have any cost, then write “No external cost”.

This project does not have a formal budget as it is part of the Machine Learning Analysts program at NorQuest College. However, the time and effort contributed by both students and professors will be compensated by the valuable knowledge and practical experience gained through the completion of this project.