

Project Proposal for the New York City TLC Taxi Fare Estimation Model

1. Project Overview

Project Name: New York City TLC Taxi Fare Estimation Model

Client: New York City Taxi and Limousine Commission (TLC)

Consulting Firm: Automatidata

Project Goal: Develop a regression model to estimate taxi fares before the ride using TLC data.

2. Project Milestones and Tasks

Milestone 1: Project Planning and Stakeholder Engagement

Tasks:

- **T1.1:** Define project scope and objectives (PACE: Plan)
- **T1.2:** Identify key stakeholders and establish communication channels (PACE: Plan)
- **T1.3:** Develop a project timeline with key deliverables (PACE: Plan)
- **T1.4:** Conduct a kickoff meeting with TLC and internal teams (PACE: Plan)

Deliverables:

- Project scope document
 - Stakeholder communication plan
 - Detailed project timeline
-

Milestone 2: Data Collection and Initial Analysis

Tasks:

- **T2.1:** Gather and consolidate data from TLC's taxi and limousine records (PACE: Analyze)
- **T2.2:** Assess data quality, identify gaps, and handle missing data (PACE: Analyze)
- **T2.3:** Conduct exploratory data analysis (EDA) to understand data patterns and trends (PACE: Analyze)
- **T2.4:** Provide initial insights and potential challenges to stakeholders (PACE: Analyze)

Deliverables:

- Data collection report
 - Data quality assessment report
 - EDA summary with initial insights
-

Milestone 3: Model Development and Testing

Tasks:

- **T3.1:** Define model requirements and select appropriate regression techniques (PACE: Construct)
- **T3.2:** Preprocess data for model training, including feature engineering (PACE: Construct)
- **T3.3:** Develop the regression model(s) and fine-tune parameters (PACE: Construct)
- **T3.4:** Evaluate model performance using relevant metrics (PACE: Construct)

Deliverables:

- Model design document
- Preprocessed dataset
- Initial regression model with performance metrics

Milestone 4: Model Deployment and Execution

Tasks:

- **T4.1:** Finalize the regression model and prepare it for deployment (PACE: Execute)
- **T4.2:** Integrate the model into TLC's existing systems for fare estimation (PACE: Execute)
- **T4.3:** Conduct user testing and gather feedback from stakeholders (PACE: Execute)
- **T4.4:** Implement improvements based on feedback and finalize deployment (PACE: Execute)

Deliverables:

- Deployed regression model
 - User feedback report
 - Finalized model with post-deployment improvements
-

3. Stakeholder Identification

Primary Stakeholders:

- **New York City TLC:** Primary client, responsible for providing data and using the fare estimation model.
- **Automatidata Data Analysis Team:** Responsible for model development, testing, and deployment.

Secondary Stakeholders:

- **Taxi and Limousine Drivers:** End-users of the fare estimation model.
- **New York City Government:** Interested in ensuring the accuracy and fairness of taxi fare estimations.

- **TLC's IT Department:** Collaborates on integrating the model into existing systems.
-

4. PACE Workflow Classification

- **Plan:** Focused on setting project goals, identifying stakeholders, and defining deliverables.
 - **Analyze:** Involves data collection, quality assessment, and exploratory analysis to lay the foundation for model development.
 - **Construct:** Centered on model design, data preprocessing, and initial testing of the regression model.
 - **Execute:** Involves finalizing, deploying, and integrating the model into TLC's systems, followed by testing and iteration based on user feedback.
-

5. Conclusion

This project proposal outlines the key milestones, tasks, stakeholders, and the PACE workflow necessary to successfully develop and deploy the taxi fare estimation model for the New York City TLC. By following this structured approach, Automatidata aims to deliver a robust and accurate solution that meets the needs of the TLC and its stakeholders.

Prepared by: Syed Ali Raza Naqvi

Date: 27, September, 2024

For: Automatidata - New York City TLC Project