**Case Study for Data Science Lead Candidate**

**Title:** *Predicting Fast-Track Eligibility for Motor Insurance Claims*

**Business Context –** An Insurance company handles thousands of motor insurance claims each month. Some claims are simple, low-risk, and suitable for **fast-track processing** — meaning they can be approved with minimal manual intervention. This reduces operational cost and improves customer satisfaction.

**Goal -** As the Data Science Lead, you are tasked with building a classification model to predict the **is\_fast\_tracked** label (1 = eligible, 0 = not eligible), determining whether a claim qualifies for fast-track processing. The model should leverage both structured metadata (e.g., vehicle information, claim history) and unstructured data (e.g., free-text damage descriptions).

Additionally, we ask that you present the model’s findings, insights, and recommendations to the leadership team in the form of a clear and concise presentation.

**Dataset Overview -** motor\_claims\_fasttrack.csv

You are provided with a sample dataset (motor\_claims\_fasttrack.csv) containing 10,000 rows of historical claims data. Each row represents one claim.

**Structured Columns**

|  |  |  |
| --- | --- | --- |
| Feature | Type | Description |
| claim\_id | String | Unique claim ID |
| vehicle\_make | categorical | Make of the vehicle |
| vehicle\_model | categorical | Model of the vehicle |
| vehicle\_year | integer | Year of manufacture |
| vehicle\_mileage | integer | Mileage at time of claim |
| accident\_location\_type | categorical | {Urban, Rural, Highway} |
| damage\_level\_reported | categorical | {Minor, Moderate, Severe} |
| customer\_tenure | float | Customer's tenure in years |
| historical\_claims\_count | integer | Previous claims by customer |
| garage\_estimate\_provided | binary | Was a garage repair estimate submitted? |
| days\_between\_accident\_and\_claim | integer | Time taken to file claim |
| is\_fast\_tracked | binary | Target label (1 = eligible, 0 = not eligible) |

**Unstructured Column**

|  |  |  |
| --- | --- | --- |
| Feature | Type | Description |
| damage\_description | text | Free-text description of the damage written by customer or claims adjuster |

**Your Tasks**

**1. Exploratory Data Analysis (EDA) & Insights**

**2. Feature Engineering (**what other interesting features can be created**)**

**3. Model Development**

**4. Model Evaluation**

**5. Model Explainability**

**Deliverables from Candidate**

Please submit:

* A **Jupyter Notebook** with clean, well-documented code
* A short **executive summary** (PDF or slides) explaining:
  + Problem statement
  + Data and challenges
  + Model results
  + Model interpretation and how can they use it
  + Next implementation steps / improvements