**Subject:** Detailed Attribute Conditions for verdict Prediction Models

As we are moving towards the model development phase, I wanted to share my understanding and request further inputs to ensure clarity:

1. **Scope of Models**
   * Our work will focus on two predictive models:
     + **Model 1 (Classification):** Predict whether the court decision will favor the claimant or the defendant.[ claims that go to court, the **final verdict** will favor the **Claimant (Plaintiff)** or the **Defendant (Insurance Company) depending on case context).]**
     + **Model 2 (Regression):** Predict the verdict amount (in case the matter proceeds to litigation).
   * Before model building, I need to apply a filter to select only those claims that are eligible for litigation. For this, I need clarity on the business conditions that define when a claim goes to court. (Example: claim type, severity, dispute flag, amount threshold, etc.). and that attribute should be there in final layer list .
2. **Key Data Requirements [Explained below]**
   * Attributes/fields that indicate whether a claim proceeded to court.
   * Attributes that capture **final verdict outcome** (claimant vs defendant).
   * Attributes related to **verdict amount**.
   * Supporting attributes such as: claim type, policy type, loss details, injury severity, attorney involvement, jurisdiction, timelines, and financial reserves.
   * Any **flags or labels used internally** to classify claim disputes or litigation status.
3. **Additional Clarifications Needed**
   * What is the **time horizon** we should consider (e.g., last 3 years of claims)?
4. **Model Development Environment**
   * Where will the coding environment be set up? (Local Jupyter notebook)
   * Will I have the ability to install Python packages directly, or do I need prior approvals?
5. **Evaluation Metrics**
   * For the classification model: What matrices we are considering accuracy/precision/recall.
   * For the amount prediction R2, RMSE.

**Suggested Next Steps**

* Business team to provide the **litigation condition**.
* Data team to confirm availability of the required attributes (X,Y)and labels.
* Tech team to confirm environment setup, access, and package installation process.

**Key Data Requirements for Model Development**

To build the models, I will need some specific data attributes

**1. Condition for Filtering Claims Going to Court**

* Before we build any model, we first need to identify **which claims proceed to court/legal proceedings**.
* For this, I will need a **specific attribute/field** (such as a flag, status) that clearly indicates whether the claim is proceeding to court.
* This attribute will act as the **filter** for all downstream modeling.

**2. Model 1 – Classification Model (Final Verdict: Claimant vs Defendant)**

* **Target attribute:** An attribute that captures the **final verdict outcome**, i.e. whether the decision was in favor of the claimant or in favor of the defendant.

**3. Model 2 – Regression Model (Prediction of Verdict/Settlement Amount)**

* Target attribute: An attribute that records the **final verdict amount** (the monetary outcome of the court case).
* Supporting attributes: Features that influence the amount (claim characteristics, claimant demographics, policy details, jurisdiction, litigation duration etc.).

**4. General Information Needed for Model Development**

* Clarification on where the Python Jupyter environment will be (platform) and whether I will have direct access to install necessary packages or will need approvals.