**Presentation Script – Accident Severity Prediction Analysis**

**1. Introduction**

*Good afternoon, everyone! I’ll walk you through our findings on predicting accident severity using machine learning.*

*As my colleagues said, we analyzed road accident data from France and applied different machine-learning models to classify accident severity. Our goal was to understand what factors contribute to severe injuries and how we can improve road safety through data-driven insights.*

*So, Let’s dive into the models, their performance, and what we learned.*

**2. Model Performance Overview**

*We tested multiple models, but I’ll focus on the three that performed best:* ***Random Forest, XGBoost, and LightGBM****.*

**Random Forest:**

*This model performed well, especially in distinguishing non-fatal cases. It achieved an* ***F1-score of 0.97 for non-fatal injuries****, showing a good balance of precision and recall. However, it struggled slightly with more severe cases. As we can see in the confusion matrix.*

**XGBoost:**

*XGBoost was our strongest model for non-fatal cases, achieving an F1-score of* ***0.98****. It also handled severe injuries better than Random Forest, with an F1-score of* ***0.59****. Its strength lies in how it processes imbalanced data, something crucial for accident prediction. This feature importance plot highlights that location and safety equipment are the most influential factors in predicting accident severity.*

**LightGBM:**

*LightGBM delivered competitive performance across all categories. It was* ***fast*** *and* ***efficient****, making it an attractive option for real-time applications. However, its recall for severe injuries was slightly lower compared to XGBoost. The ROC curve shows that the LightGBM model has a high AUC score, indicating strong performance in distinguishing between classes.*

*Now, predicting accident severity isn’t just about accuracy—it’s about understanding* ***why*** *a model makes certain predictions. That’s where interpretability comes in.*

**3. Feature Importance and Interpretability**

*We used* ***SHAP*** *and* ***LIME*** *to interpret our models, and the results were insightful!*

**XGBoost Insights:**

*XGBoost showed that the* ***most important factors*** *in accident severity are* ***safety equipment, vehicle type, location, and speed****. Interestingly,* ***higher speeds strongly correlate with more severe injuries****.*

**LightGBM Insights:**

*LightGBM highlighted* ***road conditions, lighting, and collision type*** *as key predictors. It also revealed strong interactions between factors—for example,* ***poor lighting combined with high speed increased severity significantly****.*

**Random Forest Insights:**

*Random Forest confirmed similar trends but placed* ***more emphasis on user category and vehicle type****—meaning that the role of the person in the accident (driver, passenger, pedestrian) plays a significant role in injury severity.*

*So, what does all this mean in real life?*

**4. Real-Life Recommendations**

*Based on our analysis, we identified four key areas where changes can reduce accident severity:*

**Speed Regulations:** Enforcing speed limits in high-risk areas can significantly reduce the number of severe injuries.

**Safety Equipment Awareness:** Promoting the use of **seat belts and airbags**—especially for passengers—could lower injury severity.

**Improving Road Infrastructure:** Better **lighting, clearer road signs, and safer intersections** in accident-prone areas can have a big impact.

**Driver Training Programs:** Refresher courses for **older drivers** and safety programs for **young drivers** could help reduce risk.

*These insights can help policymakers and city planners make data-driven decisions to improve road safety.*

**5. Future Enhancements & Conclusion**

*Our models performed well, but there’s always room for improvement! Here are some areas we’d like to explore next:*

**Adding Real-Time Data:** Incorporating **weather conditions, traffic congestion, and vehicle safety ratings** can improve predictions.

**Driver Behavior Analysis:** If we can track **speeding history and phone usage**, we may get even better risk assessments.

**Fine-Tuning Model Performance:** By balancing the dataset further and experimenting with ensemble methods, we can enhance predictions for severe injuries.

*In conclusion, machine learning has the potential to transform road safety. By analyzing accident data, we can identify risk factors, inform policies, and ultimately* ***save lives****.*

To finalize our presentation, we propose an Accident Severity Prediction Tool based on XGBoost to help assess accident severity based on various conditions.

On the left, we separate the accident details into three parts:

TIME AND LOCATION

VEHICLE AND DRIVER

ENVIRONMENTAL CONDITIONS

So, after selecting those details, we can just press the button Predict Severity and see the result.

*Thank you for your time, and we’d be happy to answer any questions you may have!*