Q. Can each model provide explainable insights to support its predictions?

A. When evaluating the explainability of different machine learning models used for predicting loan approvals, it’s important to consider how each model can provide insights into the decision-making process.

**Logistic Regression**

**Explainability: High**

**Feature Coefficients:** Logistic regression models are inherently interpretable because they provide coefficients for each feature. These coefficients indicate the strength and direction of the relationship between each feature and the likelihood of loan approval. E.g; a positive coefficient for a feature like income suggests that higher income increases the probability of loan approval.

**Odds Ratios:** The coefficients can be exponentiated to obtain odds ratios, which can be more intuitive. An odds ratio greater that 1 indicates that the feature increases the odds of loan approval, while an odds ratio less than 1 indicates a decrease.

**Gradient Boosting**

**Explainability: Moderate to High (with additional techniques)**

**Feature Importance:** Gradient boosting models, while more complex, can still provide useful insights through feature importance scores. These scores indicate how much each feature contributes to the model’s predictions. The model can rank features based on their importance, helping to identify the most influential factors in loan approval decisions.

**Partial Dependence Plots (PDP’s):** PDPs show the marginal effect of a feature on the predicted outcome, holding other features constant. This helps in understanding how changes in a particular feature impact the probability of loan approvaal.

**SHAP Values:** SHAP (Shapley Additive exPlanations) values provide a way to explain individual predictions. SHAP values quantify the contribution of each feature to a specific prediction, making it easier to understand why the model made a particular decision.

**Decision Tree Classifier**

**Explainability: High**

**Tree Structure:** Decision trees are highly interpretable because the decision -making process is explicitly represented in the tree structure. Each node in the tree represents a decision based ona feature vale, leading to a clear, step-by-step explanation of how a prediction is made.

**Rules Extraction:** The tree can be converted into a set of if-then rules, which are easy to understand. E.g: a rule might state “if income > $50,000 and credit score > 700, then approve the loan”

**Feature Importance:** Similar to gradient boosting, decision trees can also feature importance scores, indicating which features are most influential in the decision making process.

**Discuss using Our Gradient Boosting Optimised Model**

**Explainability of the Gradient Boosting Model**

Given the high performance of the gradient boosting model, it’s crucial to ensure that the model’s predictions can be explainable in a way that stakeholders can understand and trust.

The gradient boosting model, with its high performance as shown by the optimised results, can effectively predict loan approvals. By leveraging explainability techniques such as feature importance, PDPs, SHAP values, and LIME, transparent, understandable, and actionable insights can be provided for the model’s decision-making process. This ensures that stakeholders can trust the model’s predictions and use them to make informed decisions in the loan approval process.