**Predicting Customer Churn with Stacked Ensemble and SHAP Interpretation**

**Objective:**

The aim of this project is to predict which customers are likely to leave a telecom provider using a stacked ensemble of classifiers. Additionally, we utilize SHAP (SHapley Additive exPlanations) to understand the key factors influencing each customer's churn probability. The insights derived from SHAP are used to design actionable retention strategies.

**Methodology:**

**1. Data Preparation:**

* **Data Cleaning:** Addressed missing values and outliers to maintain data integrity.
* **Feature Encoding:** Applied OneHotEncoding for categorical variables.
* **Scaling:** Standardized numerical features for consistency.
* **Train-Test Split:** Divided the dataset into training and testing sets for model evaluation.

**2. Model Development:**

* **Base Models:** Trained Logistic Regression, Random Forest, and XGBoost classifiers.
* **Stacked Ensemble:** Combined base models using a StackingClassifier, with Logistic Regression as the meta-learner.
* **Evaluation Metrics:** Used Accuracy, Precision, Recall, F1-score, and ROC-AUC for evaluation.
* **Cross-validation:** Ensured model robustness through 5-fold cross-validation.

**3. SHAP Analysis:**

* **SHAP Explainer:** Kernel Explainer was used due to the ensemble model structure.
* **Summary Plot:** Identified the most impactful features influencing customer churn.
* **Force Plot:** Visualized individual customer predictions to understand personalized risk factors.

**Key Insights:**

1. **Critical Features Influencing Churn:**
   * **Contract Type:** Customers with month-to-month contracts showed higher churn rates.
   * **Tenure:** Short-tenured customers exhibited higher churn probabilities.
   * **Monthly Charges:** Higher monthly charges correlated with increased churn risk.
   * **Technical Support:** Lack of technical support service increased the likelihood of churn.
2. **Retention Strategies:**
   * **Targeted Offers:** Provide incentives for customers with short tenure and high monthly charges.
   * **Contract Improvements:** Promote long-term contracts to reduce monthly billing volatility.
   * **Enhanced Support:** Strengthen technical support services for better customer satisfaction.

**Conclusion:**

The stacked ensemble model demonstrated strong predictive performance, effectively identifying customers at high risk of leaving. SHAP values provided valuable transparency into the model's decision-making process, enabling tailored retention strategies.