

ML ISA WEEKLY ASSIGNMENT #6

1. Question: An insurance company wants to predict the likelihood of claims based on customer demographics and policy details. Implement a Random Forest model and explain how bagging helps improve the model's robustness and accuracy. Compare the Random Forest model's performance with a single decision tree.

Objective: Apply bagging through Random Forest to enhance model performance and stability.

2. A retail company wants to forecast sales for the next quarter using historical sales data and external factors (e.g., economic indicators, seasonal effects). Develop an ensemble model combining bagging and boosting techniques. Explain your approach, including data preprocessing, model training, and performance evaluation using appropriate metrics.

Objective: Apply both bagging and boosting techniques to a real-world forecasting problem and evaluate their effectiveness.

3. Imagine you are tasked with developing a fraud detection system for an online payment platform. Compare the effectiveness of ensemble methods such as Random Forest, XGBoost, and a custom-stacking model that combines several classifiers. Discuss the impact of K-fold cross-validation on model selection and the importance of performance metrics like precision, recall, and ROC-AUC in this context.

Objective: Compare different ensemble methods for a critical application and evaluate their performance using comprehensive metrics.