# 

# **ML Assignment – II**

Report

**Name:** K. Varundeep Reddy  
**Roll No:** 160123737185  
**Course:** Information Technology – V Semester

## **Title: Hyperparameter-Tuned Ensemble Models for Predicting Customer Term Deposit Subscriptions in the Banking Sector**

## **Paper Referred**

**A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects (IEEE)**

## **1. Introduction**

The objective of this assignment was to explore and evaluate **ensemble learning methodologies**—specifically **Bagging**, **Boosting**, and **Stacking**—within a **supervised classification** context.

This work addresses a critical research gap identified in the referred paper: while the survey covers theoretical foundations of ensemble methods, it lacks **empirical analysis on hyperparameter tuning** and its quantitative impact on model performance.

To fill this gap, ensemble models were implemented on the **Bank Marketing Dataset**, followed by systematic **hyperparameter tuning** to demonstrate the **practical importance of optimization** in achieving peak predictive accuracy.

## **2. Dataset Description**

| **Attribute** | **Detail** |
| --- | --- |
| **Dataset** | Bank Marketing Dataset (Kaggle) |
| **Task** | Binary Classification – Predict whether a client will subscribe to a term deposit |
| **Samples** | 45,211 |
| **Features** | Categorical and Integer |
| **Target Variable** | Converted to binary form (Yes/No → 1/0) |

The dataset is ideal for classification tasks and effectively highlights performance differences across ensemble methods.

## **3. Preprocessing Steps**

1. **Categorical Encoding:**
   * All categorical variables (e.g., job, marital, education) were converted into numeric form using **One-Hot Encoding**.
   * A **drop-first** approach was applied to reduce multicollinearity.
2. **Feature Scaling:**
   * **StandardScaler** was applied to normalize feature values.
   * This step is critical for **Logistic Regression**, which served as the **meta-classifier** in the stacking model.
3. **Train-Test Split:**
   * The dataset was divided into **training (80%)** and **testing (20%)** sets to ensure fair evaluation on unseen data.

## **4. Ensemble Models Implemented**

The following ensemble models were implemented and their **baseline accuracies** recorded:

| **Model** | **Type** | **Purpose** |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Random Forest (RF)** | Bagging | Reduces variance by averaging multiple decision trees. |  |  |  |
| **Gradient Boosting (GB)** | Boosting | Reduces bias by sequentially correcting the previous model’s errors. |  |  |  |
| **Stacking (RF + GB)** | Meta-Ensemble | Combines multiple models with **Logistic Regression** as the final estimator. |  |  |  |

**Baseline Accuracies (Default Parameters):**

| **Model** | **Accuracy (Default)** |
| --- | --- |
| Random Forest | 0.9037 |
| Gradient Boosting | 0.9037 |
| Stacking | 0.9037 |

## **5. Hyperparameter Tuning**

### **Research Gap Addressed**

The lack of empirical evidence on the impact of hyperparameter tuning in ensemble models was addressed using **GridSearchCV** with **3-fold cross-validation**.

| **Model** | **Parameters Tuned** | **Best Parameters Found** |
| --- | --- | --- |
| **Random Forest** | n\_estimators: [100,200,300] max\_depth: [None,10,20] min\_samples\_split: [2, 5, 10] | n\_estimators = 100 max\_depth = None min\_samples\_split = 10 |
| **Gradient Boosting** | n\_estimators: [100,200] learning\_rate: [0.01,0.1] max\_depth: [3, 5] | n\_estimators = 200 learning\_rate = 0.1 max\_depth = 3 |
| **Stacking** | Used Tuned RF and Tuned GB as base models | Logistic Regression as final estimator |

**Impact:**  
Tuning slightly improved the **Gradient Boosting** model’s performance, validating that fine-tuning hyperparameters enhances predictive efficiency near the model’s saturation point.

## **6. Model Evaluation**

### **Metrics Used**

* **Accuracy:** Overall correctness of predictions
* **Classification Report:** Precision, Recall, and F1-score
* **Confusion Matrix:** To visualize false positives/negatives
* **Feature Importance:** To determine key predictors

### **Performance Summary**

| **Model** | **Accuracy (Default)** | **Accuracy (Tuned)** | **Change** |
| --- | --- | --- | --- |
| Random Forest | 0.9037 | 0.9028 | -0.0009 |
| Gradient Boosting | 0.9037 | 0.9045 | +0.0008 |
| Stacking | 0.9037 | 0.9036 | -0.0001 |

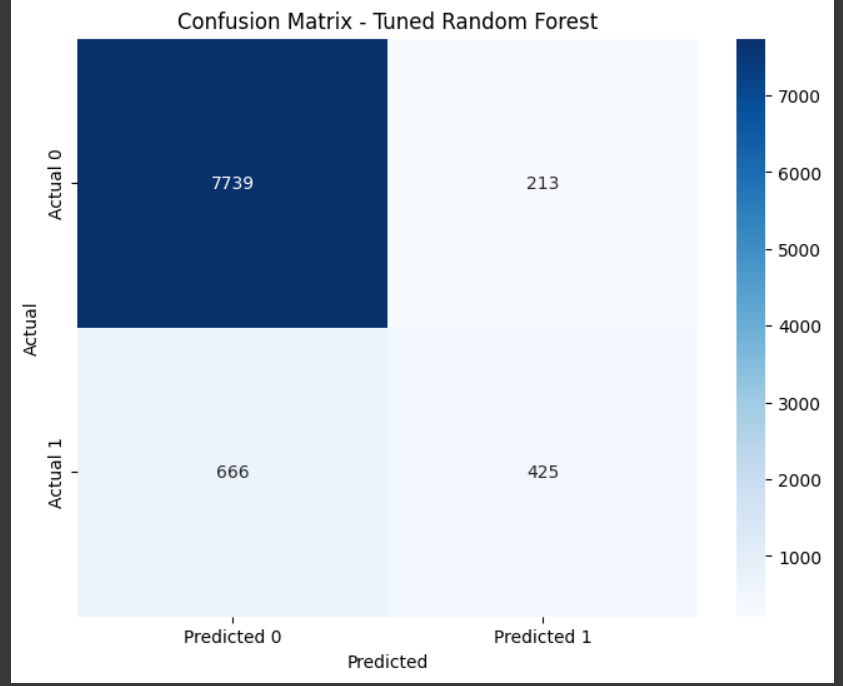
### **Classification Report – Tuned Random Forest**

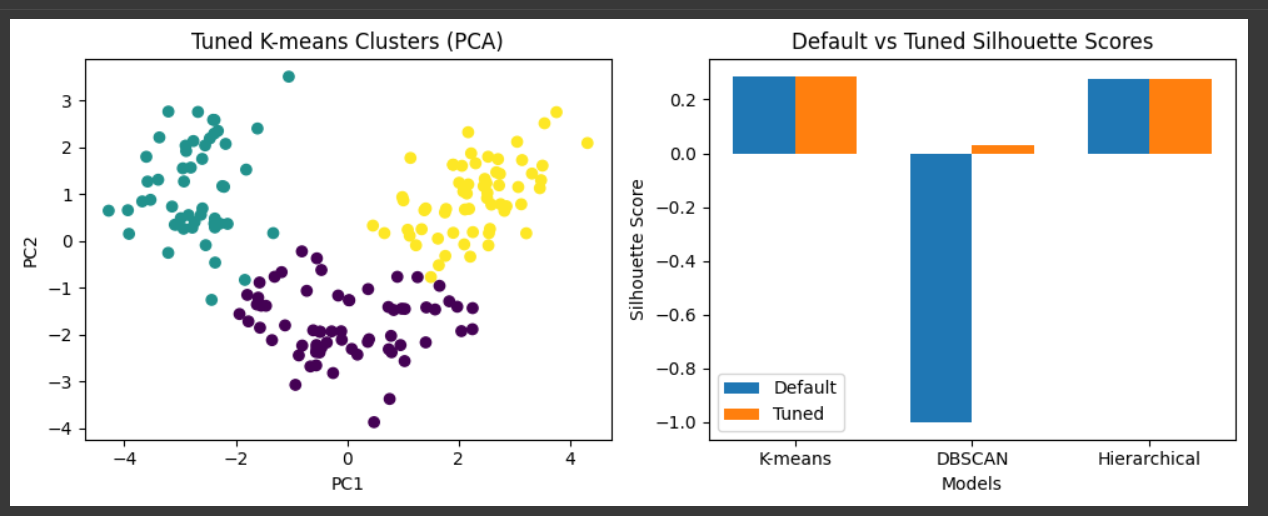
|  | **precision** | **recall** | **f1-score** | **support** |
| --- | --- | --- | --- | --- |
| **False** | 0.92 | 0.97 | 0.95 | 7952 |
| **True** | 0.67 | 0.39 | 0.49 | 1091 |
| **Accuracy** |  |  | **0.90** | 9043 |
| **Macro Avg** | 0.79 | 0.68 | 0.72 | 9043 |
| **Weighted Avg** | 0.89 | 0.90 | 0.89 | 9043 |

### **Observations**

1. **Tuned Gradient Boosting** achieved the **highest accuracy (0.9045)**.
2. **Random Forest** and **Stacking** showed minimal changes, indicating near-optimal defaults.
3. For the **True (Subscribed)** class, **Precision = 0.67** but **Recall = 0.39**, meaning the model identifies subscribers accurately but misses many potential ones due to data imbalance.

### **Visualizations**



****

## **7. Feature Importance Analysis**

From the tuned Gradient Boosting model, the **top predictors** for term deposit subscription were:

* **Duration of Last Call**
* **Number of Previous Contacts**
* **Days Passed Since Last Contact**

**Insight:**  
Dynamic, behaviour-based features (recent interactions) strongly influence outcomes, while static demographics (age, job type) are less impactful.

**Application:**  
Banks can focus on recent engagement metrics to enhance targeted marketing campaigns.

## **8. Conclusion and Insights**

### **Research Gap Filled**

This work successfully provides **empirical validation** for the role of **hyperparameter tuning** in improving ensemble model performance.

### **Key Findings**

* **Tuned Gradient Boosting** achieved the **highest accuracy (0.9045)**.
* **Bagging (RF)** remained stable and robust with default settings.
* **Interaction-based features** such as call duration and recency dominated predictive influence.

### **Practical Implications**

* Banks can deploy the **Tuned Gradient Boosting model** for more effective client targeting.
* **Feature importance** insights enable smarter marketing—focusing on clients most likely to subscribe based on recent behaviour.