**Before hyperparameter turning**

**1. Logistic Regression**

* **Train Performance**: Moderate accuracy and balanced recall, precision, and F1 score.
* **Test Performance**: Similar performance on the test set.
* **Conclusion**: Consistent performance without overfitting, but accuracy and F1 score are lower compared to other models.

**2. Support Vector Machine (SVM)**

* **Train Performance**: Good accuracy and F1 score.
* **Test Performance**: Best performance among all models in terms of accuracy (70.94%), precision (89.06%), and F1 score (77.03%).
* **Conclusion**: SVM is the most reliable model with high generalization ability, maintaining good test performance without overfitting.

**3. K-Nearest Neighbor (KNN)**

* **Train Performance**: High accuracy and F1 score.
* **Test Performance**: Lower accuracy (64.96%) and F1 score (72.48%) compared to training, indicating slight overfitting.
* **Conclusion**: Not the best choice due to overfitting and lower test accuracy.

**4. Decision Tree**

* **Train Performance**: Perfect accuracy and F1 score (100%).
* **Test Performance**: Significant drop in accuracy (65.81%) and F1 score (76.74%).
* **Conclusion**: Highly overfitted. Poor generalization on test data.

**5. Bagging**

* **Train Performance**: High accuracy and F1 score.
* **Test Performance**: Relatively lower accuracy (60.68%) and F1 score (70.89%).
* **Conclusion**: Underfitting, as the test performance is lower than expected.

**6. Random Forest**

* **Train Performance**: Perfect accuracy and F1 score (100%).
* **Test Performance**: Moderate accuracy (64.10%) and F1 score (74.70%).
* **Conclusion**: Overfitting is evident due to the drop in test performance compared to training.

**7. Gradient Boosting**

* **Train Performance**: High accuracy and balanced metrics.
* **Test Performance**: Moderate accuracy (64.10%) and F1 score (75.00%).
* **Conclusion**: Slight overfitting, but better generalization compared to Decision Tree and Random Forest.

**8. Extreme Gradient Boosting (XGBoost)**

* **Train Performance**: Perfect accuracy and F1 score (100%).
* **Test Performance**: Slightly lower accuracy (63.25%) and F1 score (74.56%).
* **Conclusion**: Overfitting with a notable drop in test performance.

**After Hyperparameter Tuning:**

1. **Gradient Boosting** achieved the highest F1 score on the test set (76.74%) with improved precision and recall, indicating a better balance after tuning.
2. **Extreme Gradient Boosting** also showed good performance with an F1 score of 75.90%, closely following Gradient Boosting.
3. **Support Vector Machine (SVM)** and **K-Nearest Neighbor (KNN)** maintained relatively high accuracy but did not outperform Gradient Boosting.
4. **Ensemble Methods** still showed high training accuracy but improved generalization on the test set.
5. **Decision Tree** and **Bagging** showed moderate performance, but they were not the best in terms of test accuracy or F1 score.

**Conclusion:**

After hyperparameter tuning, **Gradient Boosting** is the best model due to its highest F1 score on the test set, indicating a good balance between precision and recall. **Extreme Gradient Boosting** is also a strong contender with similar performance.