

Apply all three techniques of feature selection in dataset(<https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset>) and also apply one machine learning model and get above 90% accuracy

1. Add Task Description:

Objective:

The goal is to apply three feature selection techniques on the breast cancer dataset and build a machine learning model with above 90% accuracy. Feature selection improves model performance by selecting the most relevant features, reducing dimensionality, and mitigating overfitting.

2. Attach Screenshot Of Output.:

Accuracy using Filter (SelectKBest): 85.96%
Accuracy using Wrapper (RFE): 95.61%
Accuracy using Embedded (Random Forest): 95.61%

3. Describe Widget/Algorithm Used In Task:

1. **Filter Method:**

- **Algorithm:** SelectKBest (using chi-square test).
- **Description:** Selects the top K features with the highest correlation to the target variable. It is computationally efficient and independent of the ML model.

2. **Wrapper Method:**

- **Algorithm:** Recursive Feature Elimination (RFE) using Logistic Regression.
- **Description:** Iteratively fits a model and removes the least significant features based on their coefficients or importance scores. RFE is computationally expensive but effective for feature ranking.

3. **Embedded Method:**

- **Algorithm:** Random Forest feature importance.
- **Description:** Random Forest assigns importance scores to features during training, helping to identify the most influential features.

4. **Machine Learning Model:**

- **Algorithm:** Random Forest Classifier.
- **Description:** An ensemble learning method that combines multiple decision trees to enhance prediction accuracy and robustness. It is robust to overfitting and works well with selected features.