

For Dataset 1:

1. A brief description of the dataset (what is the task, what are the features and the target)

Dataset Name: Bank Marketing dataset, ID 1461, from OpenML.

Task: The task is to predict whether a client will subscribe to a term deposit.

Features: age, job, marital, education, default, balance, housing, loan, contact, day, month, duration, campaign, pdays, previous, poutcome.

Target: A binary classification target variable indicating whether a client subscribed to a term deposit (yes or no).

2. Describe the parameters you tuned and with what values

Decision Tree Classifier (DecisionTreeClassifier):

Parameter tuned: min_samples_leaf

Values: [5, 20, 50, 200, 500]

K-Nearest Neighbors Classifier (KNeighborsClassifier):

Parameter tuned: n_neighbors

Values: [5, 20, 50, 80, 100]

Multinomial Naive Bayes Classifier (MultinomialNB):

Used default parameters without tuning.

Logistic Regression Classifier (LogisticRegression):

Used default parameters without tuning.

Dummy Classifier (DummyClassifier):

Strategy: most_frequent, which always predicts the most common class.

3.A table showing means and standard deviations of the cross-validation results of all the methods

Model	Best Mean AUC	AUC Standard Deviation
Decision Tree Classifier (DTC)	0.7944957481049605	0.06109118201619647
K-Nearest Neighbors (KNN)	0.8281142072757728	0.08180008950592577
Multinomial Naive Bayes (MNB)	0.5732458309960312	0.2073225198723102
Logistic Regression (LR)	0.7240584992037207	0.18299335485913484
Dummy Classifier (DC)	0.5	0.0

4. Discussion of the results

K-Nearest Neighbors (KNN) achieved the highest performance, with a **Best Mean AUC of 0.828** and a **Standard Deviation of 0.0818**. This result indicates that KNN has strong discriminatory ability on this dataset, but the higher standard deviation suggests that its performance may vary significantly across different data splits. This variability may be attributed to KNN's sensitivity to feature scaling and neighborhood composition, which affects its stability.

Decision Tree Classifier (DTC) showed competitive performance, with a **Best Mean AUC of 0.794** and a **Standard Deviation of 0.0611**. This relatively low standard deviation suggests that the model is more stable than KNN, though slightly less effective in discriminatory power. Decision Trees generally handle mixed data types well and capture non-linear relationships, making them suitable for this dataset.

Logistic Regression (LR) performed reasonably well, with a **Best Mean AUC of 0.724** and a **Standard Deviation of 0.1830**. While the AUC is relatively high, the large standard deviation indicates a degree of instability in its performance, possibly due to variability in categorical feature encoding or sensitivity to outliers in numerical features.

Multinomial Naive Bayes (MNB) had a **Best Mean AUC of 0.573** and a **Standard Deviation of 0.2073**. This model performed significantly lower than the other models. Naive Bayes relies on the assumption of feature independence, which may not hold in this dataset, resulting in reduced performance and higher variability.

Dummy Classifier (DC) served as a baseline model, with a **Best Mean AUC of 0.5** and **zero standard deviation**, as it always predicts the most frequent class. This

outcome reflects no true predictive power and confirms its role as a baseline for comparison.

For Dataset 2:

1. A brief description of the dataset (what is the task, what are the features and the target)

Dataset Name: Abalone (ID 720) from OpenML

Task: Classification task to predict if the target variable (ring count) exceeds a certain age threshold.

Features: Contains 8 features: Sex (categorical): M, F, I. Length, Diameter, Height, Whole weight, Shucked weight, Viscera weight, Shell weight: numerical variables.

Target: Ring count, typically representing age, is converted to a categorical target for classification purposes.

2. Describe the parameters you tuned and with what values

Decision Tree Classifier (DecisionTreeClassifier):

Parameter tuned: min_samples_leaf

Values: [5, 20, 50, 200, 500]

K-Nearest Neighbors Classifier (KNeighborsClassifier):

Parameter tuned: n_neighbors

Values: [5, 20, 50, 80, 100]

Multinomial Naive Bayes Classifier (MultinomialNB):

Used default parameters without tuning.

Logistic Regression Classifier (LogisticRegression):

Used default parameters without tuning.

Dummy Classifier (DummyClassifier):

Strategy: most_frequent, which always predicts the most common class.

3.A table showing means and standard deviations of the cross-validation results of all the methods

Model	Best Mean AUC	AUC Standard Deviation
Decision Tree Classifier (DTC)	0.8530307854237595	0.02330320935685634
K-Nearest Neighbors (KNN)	0.8622121894620951	0.016511183189846387
Multinomial Naive Bayes (MNB)	0.8056960823320052	0.03941028325850498
Logistic Regression (LR)	0.8640101086586544	0.029553370090355083
Dummy Classifier (DC)	0.5	0.0

4. Discussion of the results

Logistic Regression (LR) demonstrated the highest performance with a **Best Mean AUC of 0.864** and a **Standard Deviation of 0.0296**. This result indicates that Logistic Regression provides strong discriminatory ability with relatively stable performance. The model's ability to handle mixed feature types (numerical and one-hot encoded categorical) contributed to its success on this dataset.

K-Nearest Neighbors (KNN) performed very well, with a **Best Mean AUC of 0.862** and a **Standard Deviation of 0.0165**. This model also achieved a high AUC score, with a lower standard deviation than previously observed, suggesting that KNN has benefited from optimal parameter tuning. Its performance, however, may still be sensitive to neighborhood composition and feature scaling.

Decision Tree Classifier (DTC) achieved a **Best Mean AUC of 0.853** and a **Standard Deviation of 0.0233**. The relatively low standard deviation shows that the Decision Tree model provides stable performance with high discriminatory power, making it a reliable choice for this dataset. Decision Trees are known for handling non-linear relationships well, which likely contributed to its success here.

Multinomial Naive Bayes (MNB) had a **Best Mean AUC of 0.806** and a **Standard Deviation of 0.0394**. This model performed lower than the others, which may be due to the Naive Bayes assumption of feature independence, which doesn't hold well in this dataset, leading to reduced predictive power and increased variability.

Dummy Classifier (DC), serving as a baseline, had a **Best Mean AUC of 0.5** with **zero standard deviation**, indicating no predictive power and providing a reference for comparison with other models.

