RA2311047010016

S.No.	Date	Title	Signature
1	07.08.2025	Exploring the Deep Learning platforms	
2.	07.08.2025	Implement a classifier using open source dataset	
3.	14.08.2025	Study of classifiers with respect to Statistical parameters	
4.	22.08.2025	To build and train a simple feed forward Network (~~FFNN~~) on MNIST dataset.	
5.	22.8.25	Study different activation function used in NN	
6.	9.9.25	Implement Gradient Descent & Backpropagation in ~~NN~~	
7.	16.09.2025	Build a CNN model to classify Cat and dog image	

19.08.2025

3. Study of classifiers with respect to statistical parameters.

Aim: To evaluate and compare performance of classifier respect to statistical parameters.

Objectives:

→ Load and preprosses Iris dataset

→ Apply KNN, Decision tree and SVM classifiers

→ Evaluate each model using accuracy, precision and f1score.

→ Compare the performance of the classifiers using tabular format

→ To draw observations and Calculate which classifier performs best on dataset.

TP : True positive.

TN : True Negatives

FP : false positive

FN : False negative

Comparision table.

~~Classifier~~

~~KNN~~

	KNN (K=5)	Decision Tree	SVM.
Accuracy	89%.	92%.	94%.
Precision	86%.	90%.	93%.
Recall	84%.	88%.	91%.
F1 Score	85%.	89%.	92%.
Training Time.	Low	Medium	High.

Observations :

→ KNN is simple and effective for small datasets, but slows down with large datasets.

→ Decision Trees are interpretable and fast but prone to overfitting.

→ SVM gives high accuracy especially in high-dimensional spaces but is computationally expensive.

→ Precision is important in applications like spam detection or medical diagnosis where false positive are costly.

→ Accuracy alone is not reliable when classes are imbalanced - F1-Score or precision / Recall give better insights.

14/8/25

Result:

Successfully completed the study of classifiers using statistical parameters.

2. Else:

a. For each feature, calculate information gain and Gini index.

b.) Select the feature with the highest gain

c.) Split the dataset based on this feature.

d.) Recursively build subtrees for each split.

3.) Return the root node of the tree.

3. Support Vector Machine (SVM).

Input: Training dataset.

Output: Optimal hyperplane and classifier mode.

1. Map data to high-dimensional Space (if needed)

2. Identify the hyperplane that maximizes margin between classes.

3. Use optimization (e.g., quadratic programming) to find support vectors.

4. For new input:

a. Calculate on which side of the hyperplane the input lies

b) classify based on the side.

Statistical parameters.

Metric	Formula	Explanation.
Accuracy	(TP + TN) / (TP + TN + FP + FN)	proportion of correct predictions
precision	TP/(TP + FP)	How many predicted positives were true positives
Recall	TP / (TP + FN)	How many actual positives were correctly predicted
f1-score	2 * (precision * Recall) / (precision + Recall)	Harmonic mean of precision recall.

pseudocode - KNN .

Input : Training data, value of k, test instance

output : Predicted class .

1. For each instance in the training data :
   a) Calculate the distance to the test instance (e.g., Euclidean distance)
2. Sort the distance in ascending order
3. Select the top k nearest neighbors.
4. Count the class labels among the k neighbors .
5. Return the most frequent class label as prediction

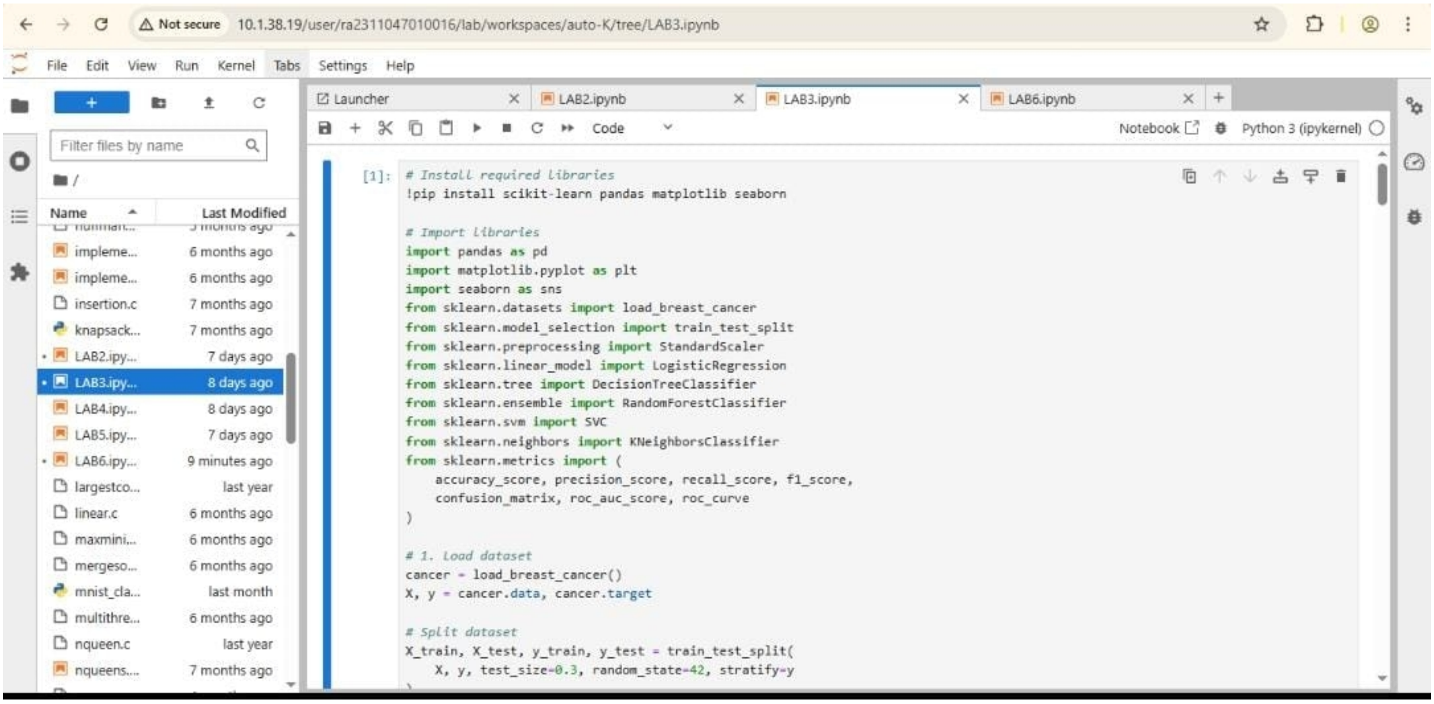
Decision Tree :

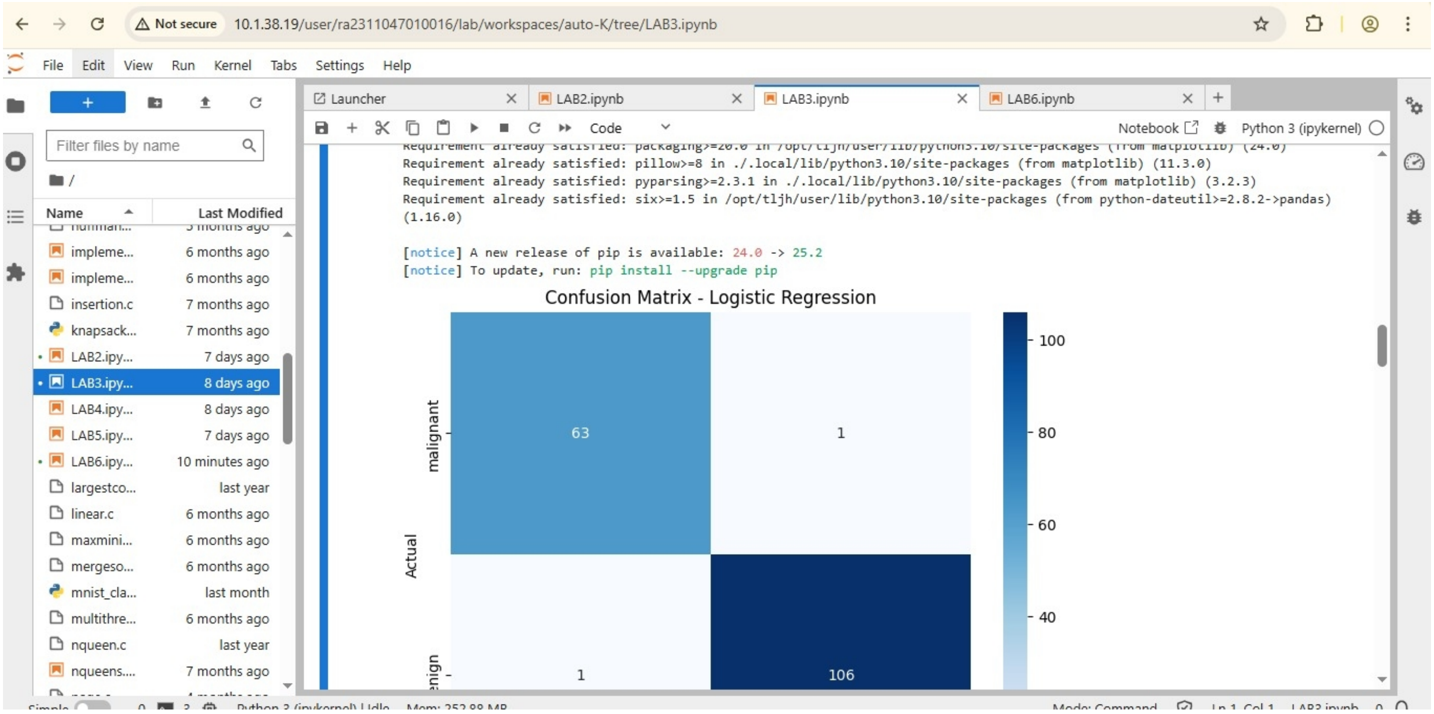
Pseudocode:

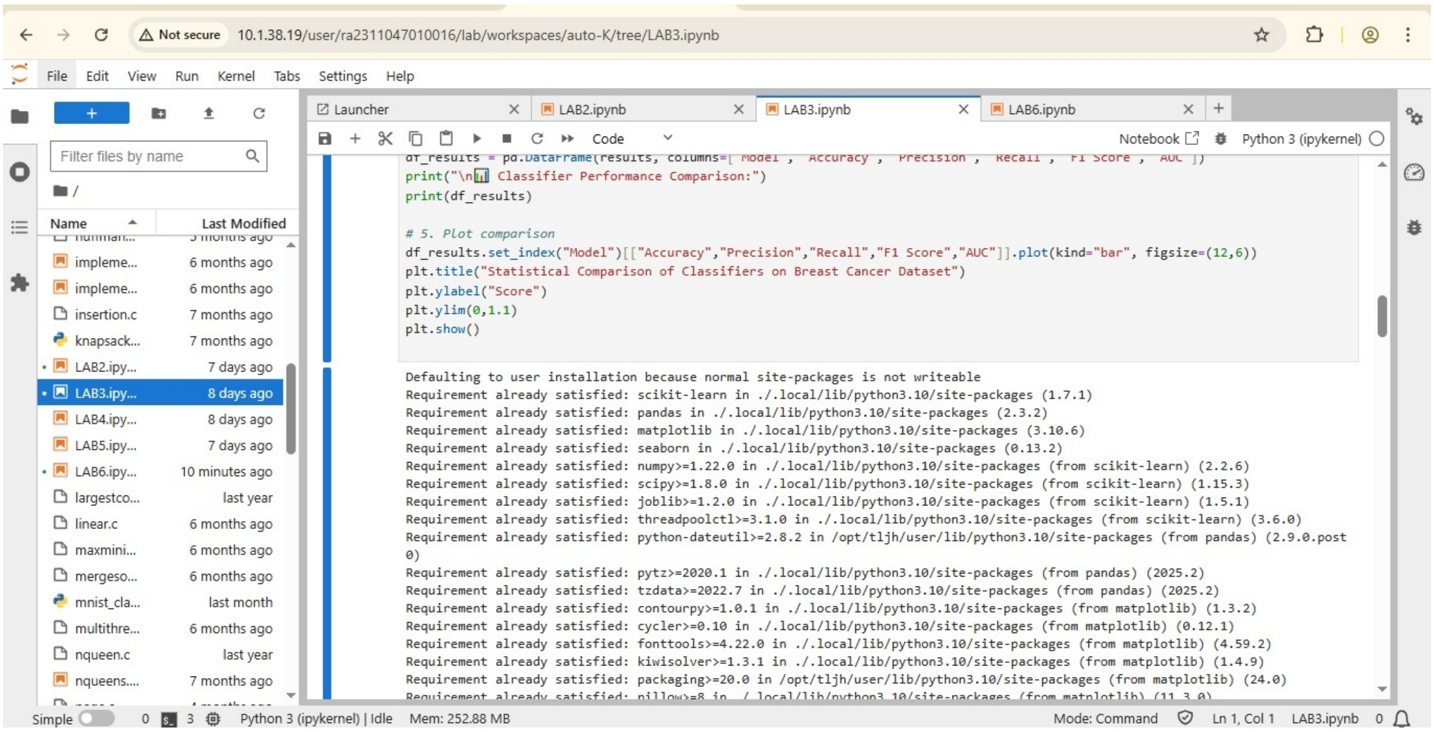
Input : Training dataset .

Output : Decision Tree ,

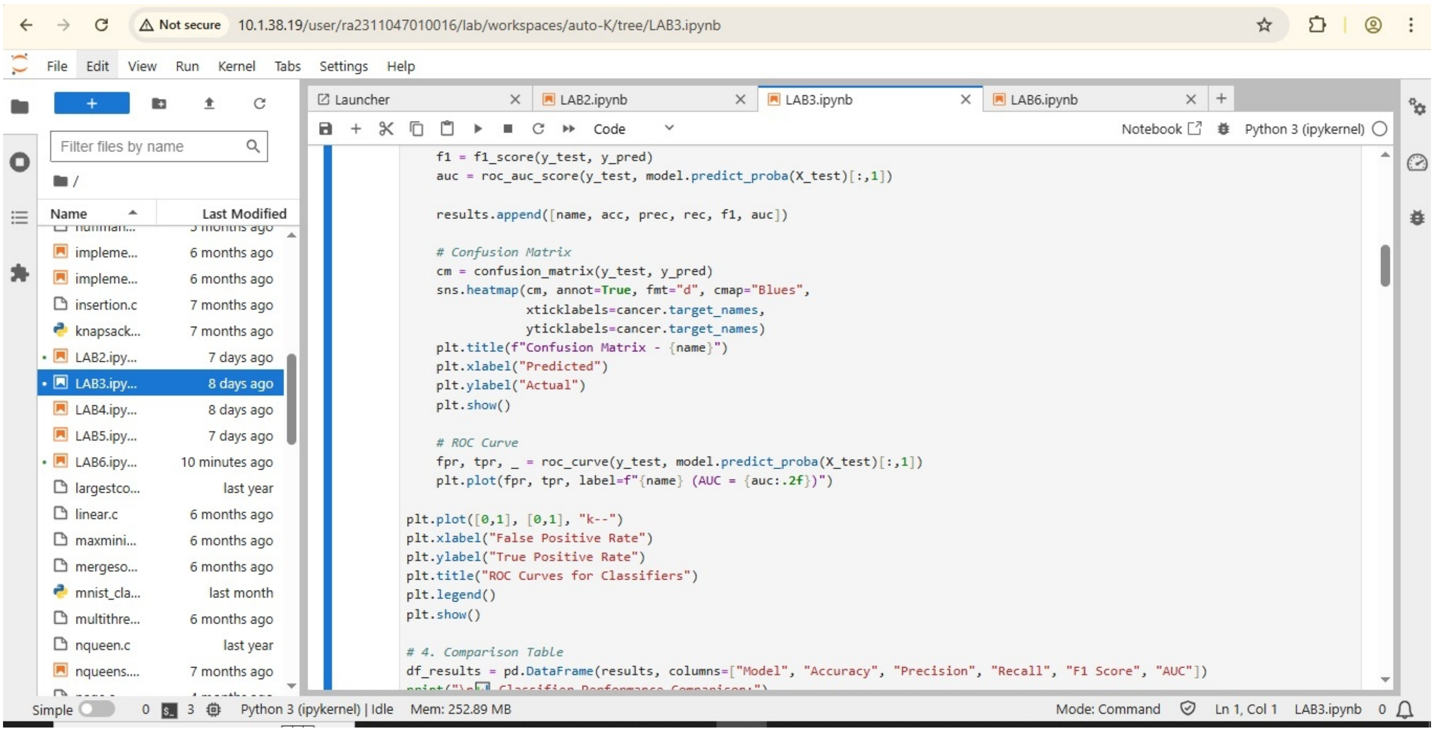
1. If all instance belong to one class ;
   Return a leaf node with that class

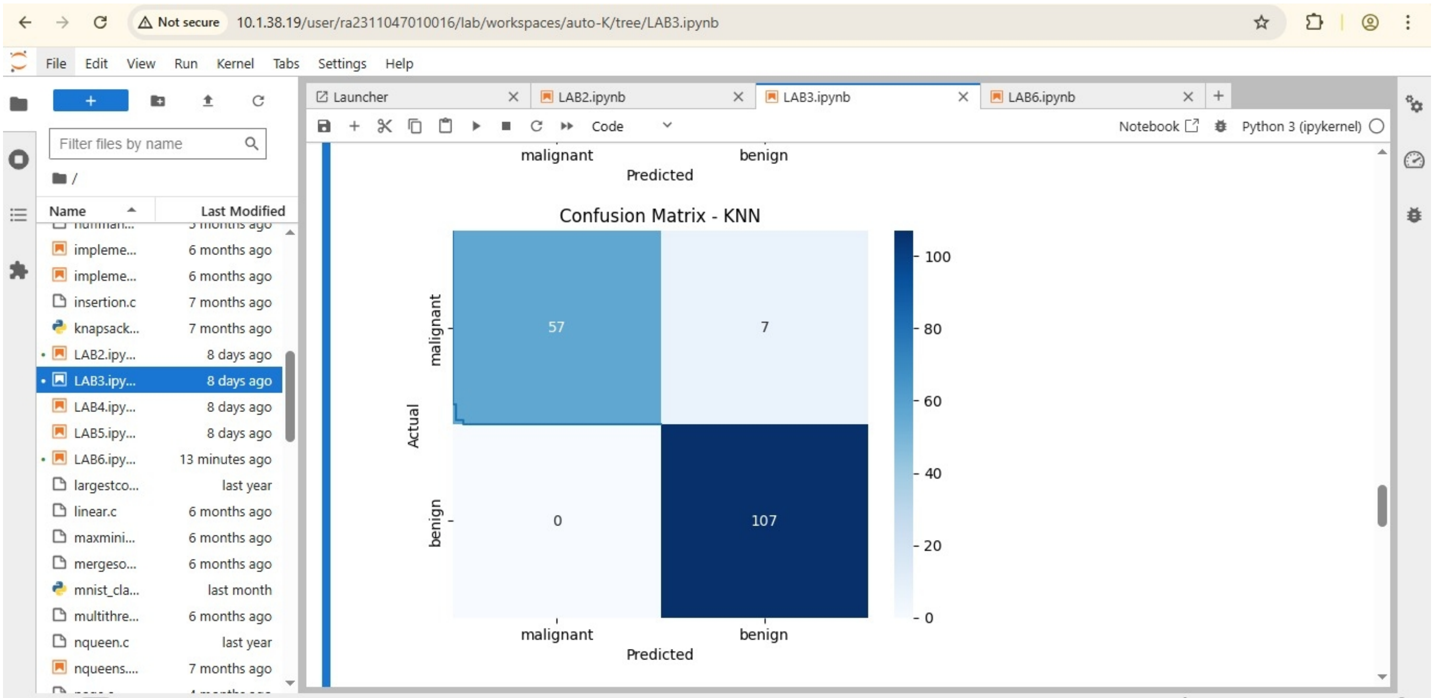

Not secure 10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File Edit View Run Kernel Tabs Settings Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 7 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 7 days ago
LAB6.ipy... 9 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
[1]:
Install required libraries
!pip install scikit-learn pandas matplotlib seaborn
# Import libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import (
accuracy_score, precision_score, recall_score, f1_score,
confusion_matrix, roc_auc_score, roc_curve
)
# 1. Load dataset
cancer = load_breast_cancer()
X, y = cancer.data, cancer.target
Split dataset
X_train, X_test, y_train, y_test = train_test_split(
X, y, test_size=0.3, random_state=42, stratify=y

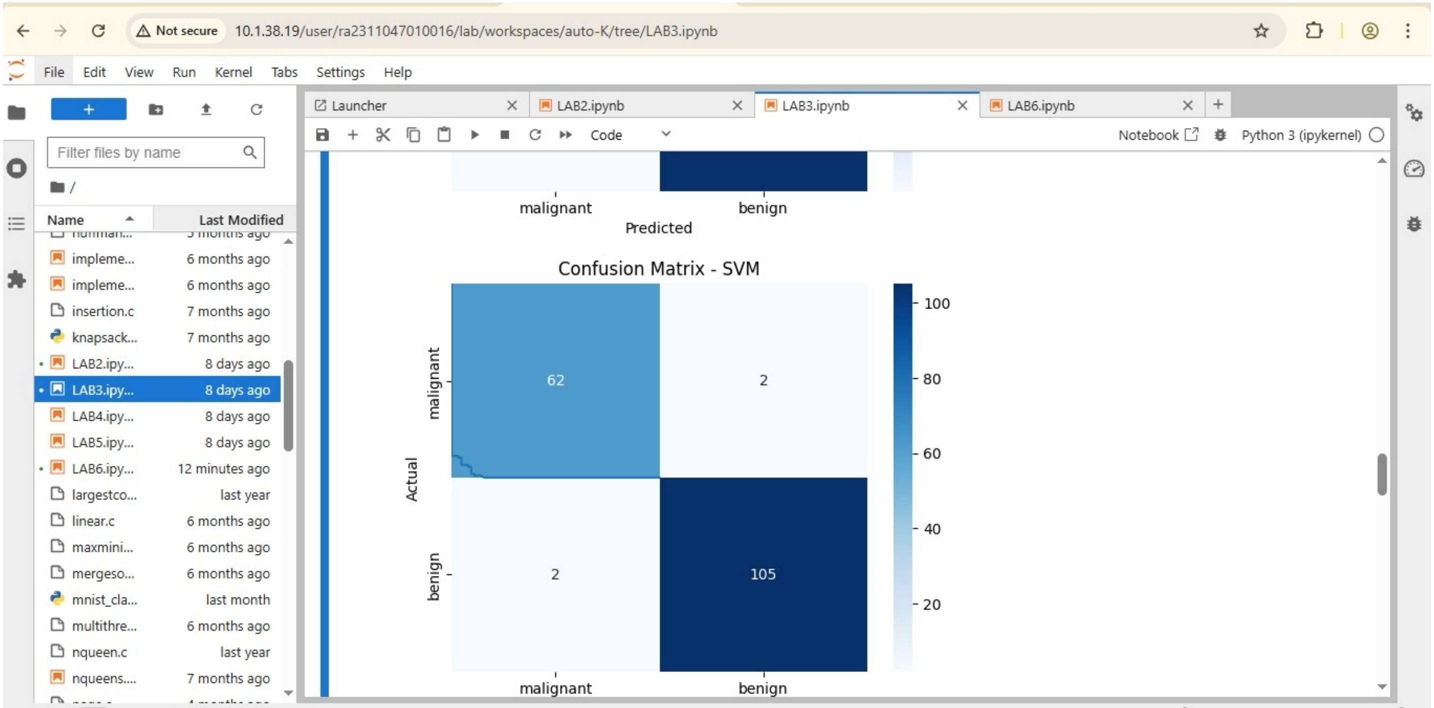

10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File Edit View Run Kernel Tabs Settings Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 7 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 7 days ago
LAB6.ipy... 10 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
Code
Notebook
Python 3 (ipykernel)
Requirement already satisfied: pillow>=8 in ./.local/lib/python3.10/site-packages (from matplotlib) (11.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in ./.local/lib/python3.10/site-packages (from matplotlib) (3.2.3)
Requirement already satisfied: six>=1.5 in /opt/tljh/user/lib/python3.10/site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
[notice] A new release of pip is available: 24.0 -> 25.2
[notice] To update, run: pip install --upgrade pip
Confusion Matrix - Logistic Regression
malignant
Actual
63
1
1
106
100
80
60
40



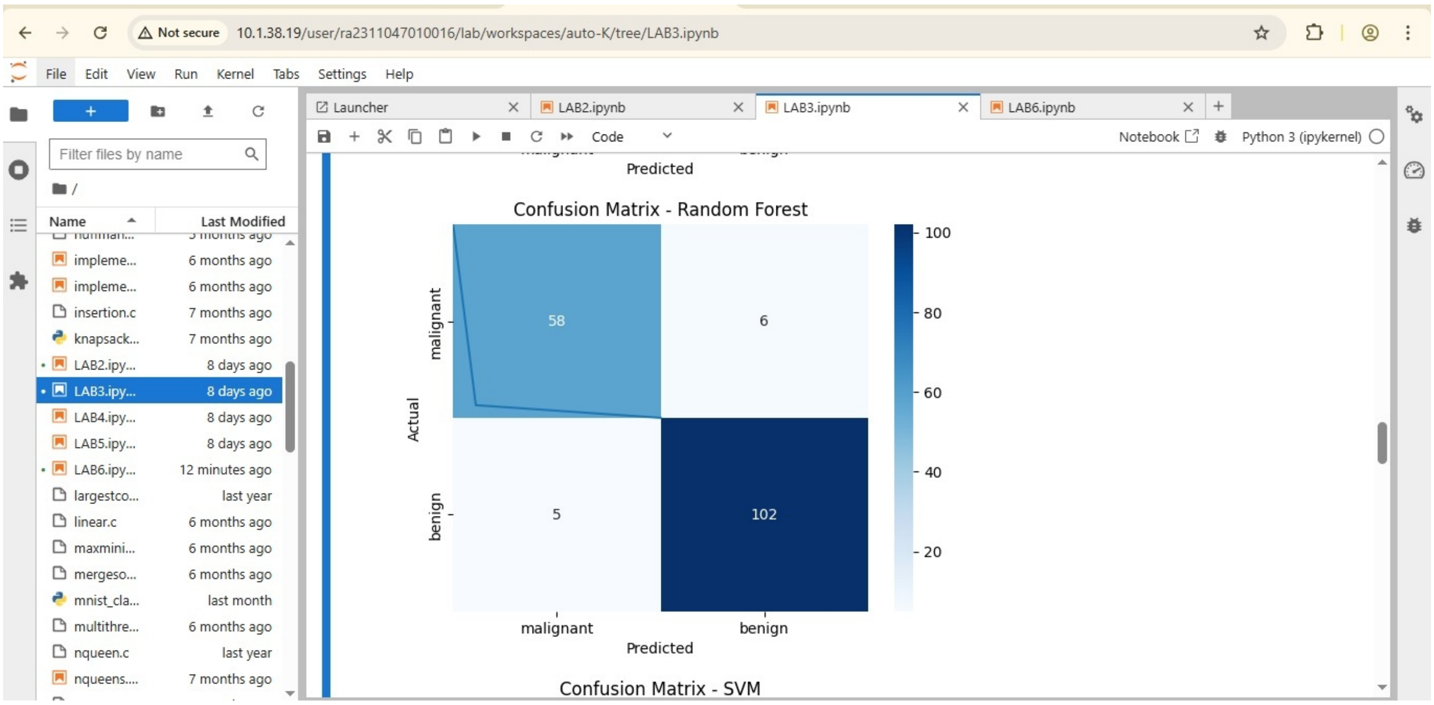
df_results = pd.DataFrame(results, columns=["Model","Accuracy","Precision","Recall","F1 Score","AUC"])
print("\n📊 Classifier Performance Comparison:")
print(df_results)
# 5. Plot comparison
df_results.set_index("Model")[["Accuracy","Precision","Recall","F1 Score","AUC"]].plot(kind="bar", figsize=(12,6))
plt.title("Statistical Comparison of Classifiers on Breast Cancer Dataset")
plt.ylabel("Score")
plt.ylim(0,1.1)
plt.show()
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: scikit-learn in ./.local/lib/python3.10/site-packages (1.7.1)
Requirement already satisfied: pandas in ./.local/lib/python3.10/site-packages (2.3.2)
Requirement already satisfied: matplotlib in ./.local/lib/python3.10/site-packages (3.10.6)
Requirement already satisfied: seaborn in ./.local/lib/python3.10/site-packages (0.13.2)
Requirement already satisfied: numpy>=1.22.0 in ./.local/lib/python3.10/site-packages (from scikit-learn) (2.2.6)
Requirement already satisfied: scipy>=1.8.0 in ./.local/lib/python3.10/site-packages (from scikit-learn) (1.15.3)
Requirement already satisfied: joblib>=1.2.0 in ./.local/lib/python3.10/site-packages (from scikit-learn) (1.5.1)
Requirement already satisfied: threadpoolctl>=3.1.0 in ./.local/lib/python3.10/site-packages (from scikit-learn) (3.6.0)
Requirement already satisfied: python-dateutil>=2.8.2 in /opt/tljh/user/lib/python3.10/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in ./.local/lib/python3.10/site-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in ./.local/lib/python3.10/site-packages (from pandas) (2025.2)
Requirement already satisfied: contourpy>=1.0.1 in ./.local/lib/python3.10/site-packages (from matplotlib) (1.3.2)
Requirement already satisfied: cycler>=0.10 in ./.local/lib/python3.10/site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in ./.local/lib/python3.10/site-packages (from matplotlib) (4.59.2)
Requirement already satisfied: kiwisolver>=1.3.1 in ./.local/lib/python3.10/site-packages (from matplotlib) (1.4.9)
Requirement already satisfied: packaging>=20.0 in /opt/tljh/user/lib/python3.10/site-packages (from matplotlib) (24.0)


Not secure
10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File Edit View Run Kernel Tabs Settings Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 7 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 7 days ago
LAB6.ipy... 10 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
f1 = f1_score(y_test, y_pred)
auc = roc_auc_score(y_test, model.predict_proba(X_test)[:,1])
results.append([name, acc, prec, rec, f1, auc])
# Confusion Matrix
cm = confusion_matrix(y_test, y_pred)
sns.heatmap(cm, annot=True, fmt="d", cmap="Blues",
xticklabels=cancer.target_names,
yticklabels=cancer.target_names)
plt.title(f"Confusion Matrix - {name}")
plt.xlabel("Predicted")
plt.ylabel("Actual")
plt.show()
# ROC Curve
fpr, tpr, _ = roc_curve(y_test, model.predict_proba(X_test)[:,1])
plt.plot(fpr, tpr, label=f"{name} (AUC = {auc:.2f})")
plt.plot([0,1], [0,1], "k--")
plt.xlabel("False Positive Rate")
plt.ylabel("True Positive Rate")
plt.title("ROC Curves for Classifiers")
plt.legend()
plt.show()
# 4. Comparison Table
df_results = pd.DataFrame(results, columns=["Model", "Accuracy", "Precision", "Recall", "F1 Score", "AUC"])
Simple
Python 3 (ipykernel) | Idle
Mem: 252.89 MB
Mode: Command
Ln 1, Col 1
LAB3.ipynb

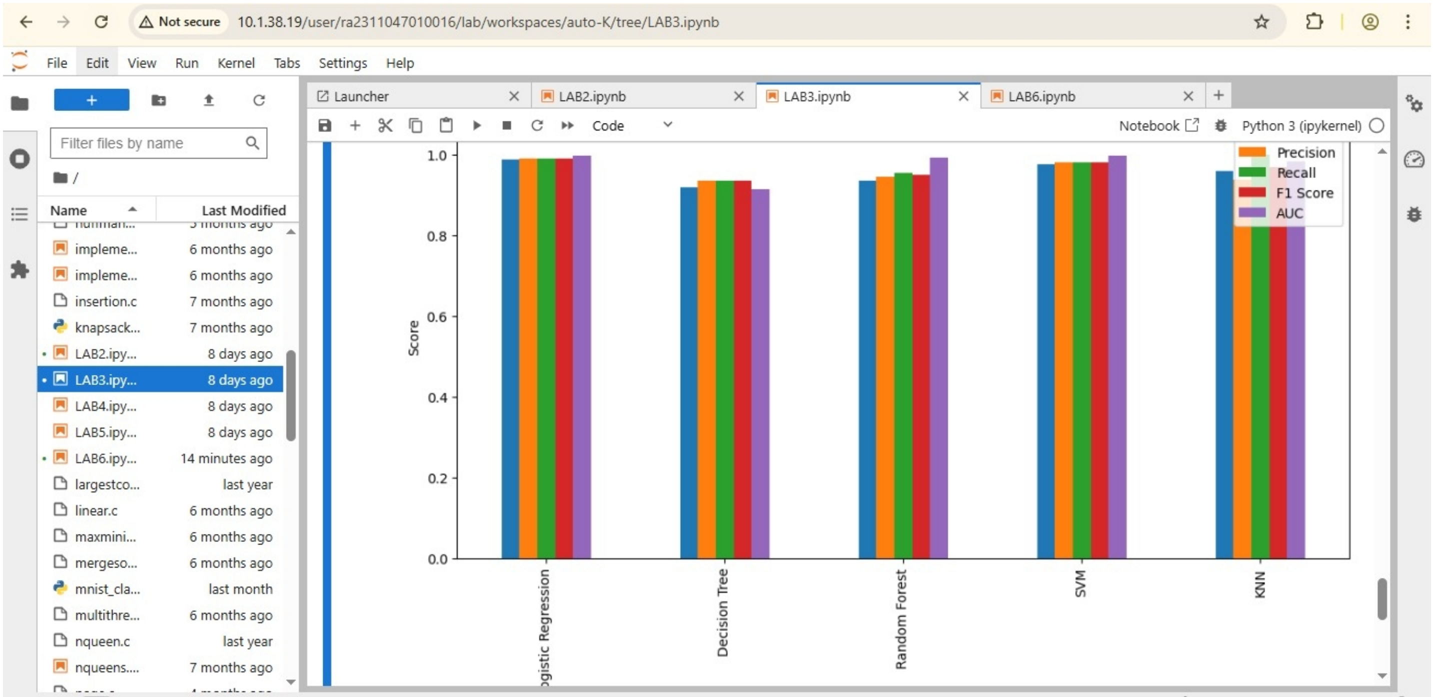

Not secure
10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File
Edit
View
Run
Kernel
Tabs
Settings
Help
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 8 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 8 days ago
LAB6.ipy... 13 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
malignant
benign
Predicted
Confusion Matrix - KNN
Actual
malignant
benign
57
7
0
107
malignant
benign
Predicted
100
80
60
40
20
0



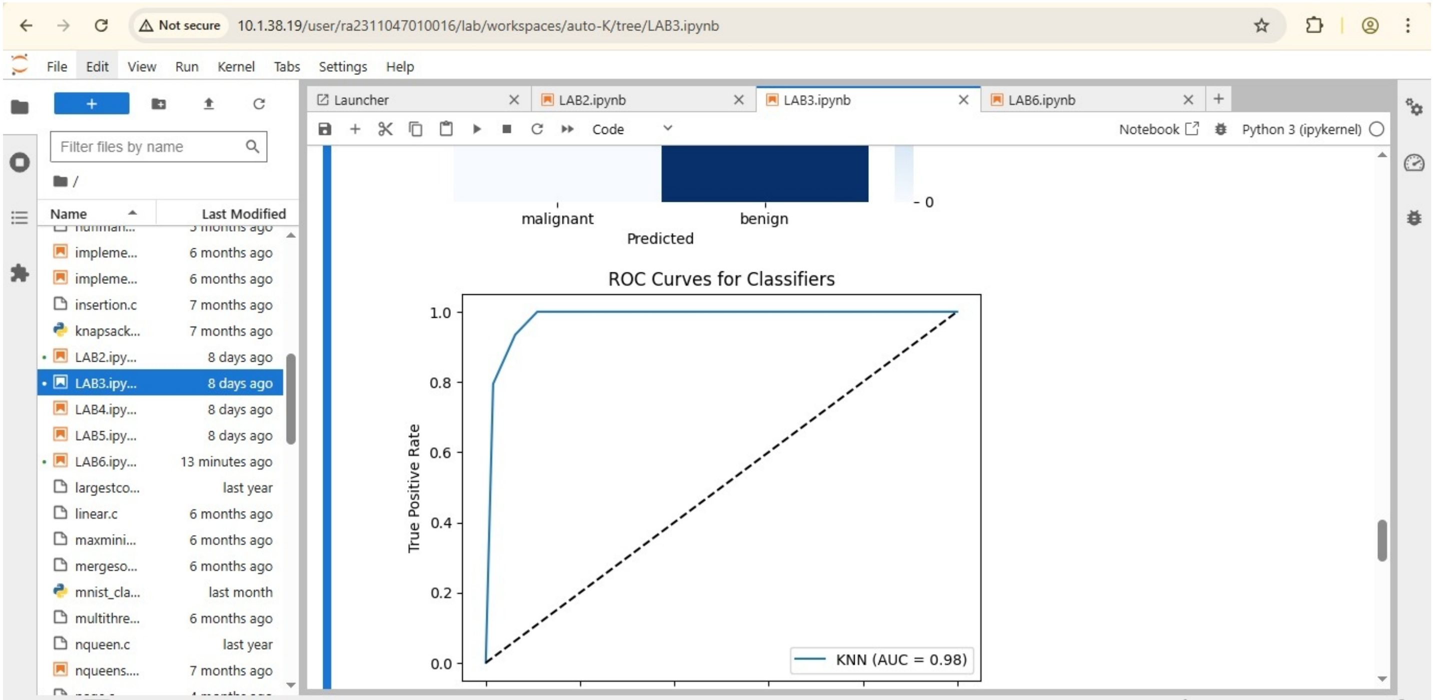
Not secure 10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File Edit View Run Kernel Tabs Settings Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 8 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 8 days ago
LAB6.ipy... 12 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
malignant
benign
Predicted
Confusion Matrix - SVM
Actual
malignant
benign
62
2
2
105
100
80
60
40
20
malignant
benign



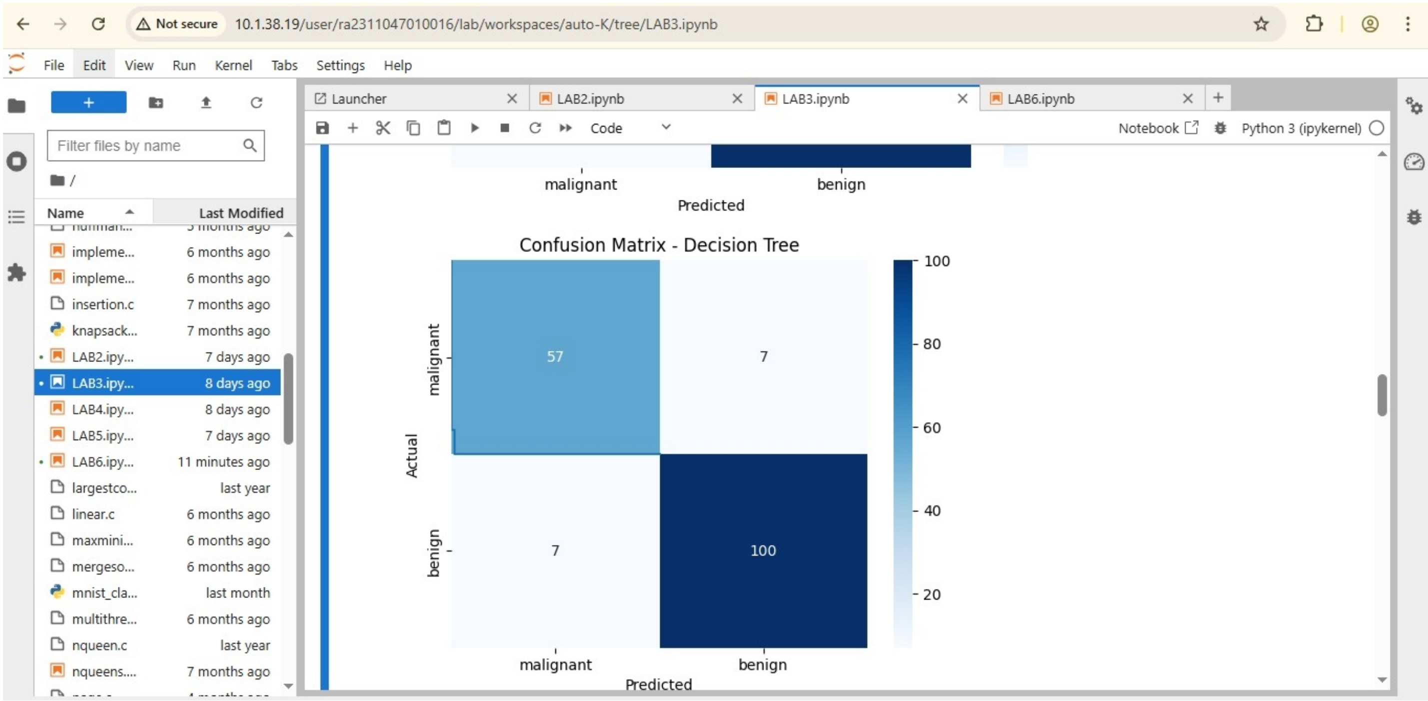
10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File Edit View Run Kernel Tabs Settings Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 8 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 8 days ago
LAB6.ipy... 12 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
Code
Notebook
Python 3 (ipykernel)
Predicted
Confusion Matrix - Random Forest
Actual
malignant
benign
58
6
5
102
Predicted
Confusion Matrix - SVM

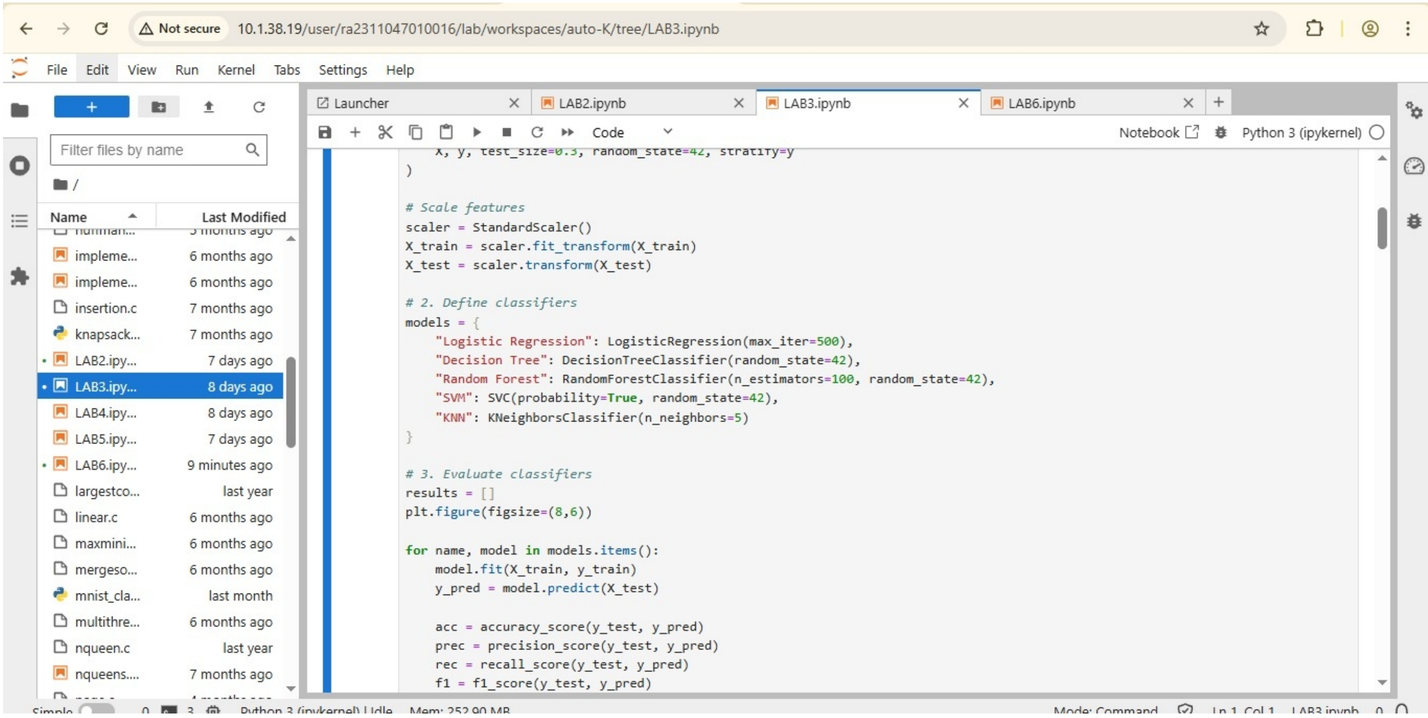


Not secure
10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File
Edit
View
Run
Kernel
Tabs
Settings
Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 8 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 8 days ago
LAB6.ipy... 14 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
Precision
Recall
F1 Score
AUC
Score
1.0
0.8
0.6
0.4
0.2
0.0
ogistic Regression
Decision Tree
Random Forest
SVM
KNN

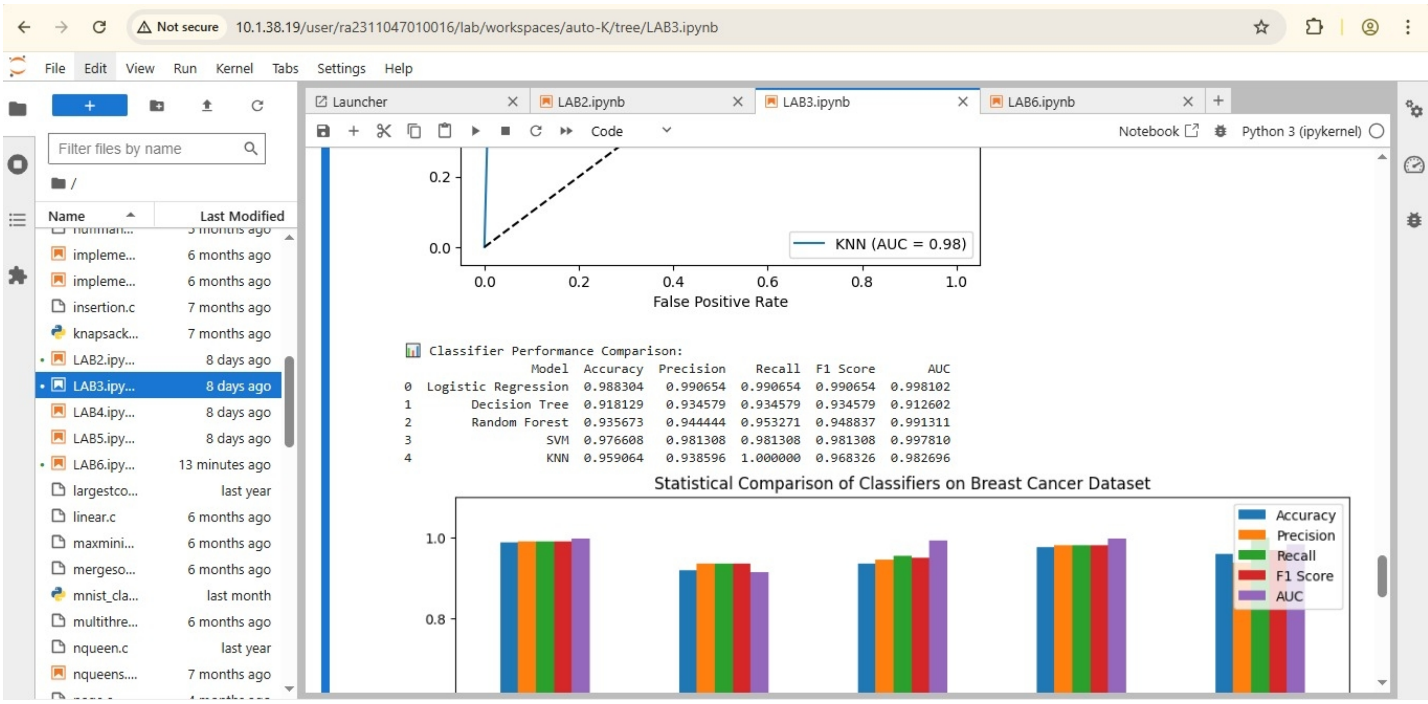


Not secure
10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File Edit View Run Kernel Tabs Settings Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
Filter files by name
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 8 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 8 days ago
LAB6.ipy... 13 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
malignant
benign
Predicted
- 0
ROC Curves for Classifiers
True Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
KNN (AUC = 0.98)


Not secure
10.1.38.19/user/ra2311047010016/lab/workspaces/auto-K/tree/LAB3.ipynb
File
Edit
View
Run
Kernel
Tabs
Settings
Help
Launcher
LAB2.ipynb
LAB3.ipynb
LAB6.ipynb
Code
Notebook
Python 3 (ipykernel)
Filter files by name
/
Name
Last Modified
impleme... 6 months ago
impleme... 6 months ago
insertion.c 7 months ago
knapsack... 7 months ago
LAB2.ipy... 7 days ago
LAB3.ipy... 8 days ago
LAB4.ipy... 8 days ago
LAB5.ipy... 7 days ago
LAB6.ipy... 11 minutes ago
largestco... last year
linear.c 6 months ago
maxmini... 6 months ago
mergeso... 6 months ago
mnist_cla... last month
multithre... 6 months ago
nqueen.c last year
nqueens.... 7 months ago
malignant
benign
Predicted
Confusion Matrix - Decision Tree
Actual
malignant
benign
57
7
7
100
malignant
benign
Predicted
100
80
60
40
20



```python
    X, y, test_size=0.3, random_state=42, stratify=y
)

# Scale features
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

# 2. Define classifiers
models = {
    "Logistic Regression": LogisticRegression(max_iter=500),
    "Decision Tree": DecisionTreeClassifier(random_state=42),
    "Random Forest": RandomForestClassifier(n_estimators=100, random_state=42),
    "SVM": SVC(probability=True, random_state=42),
    "KNN": KNeighborsClassifier(n_neighbors=5)
}

# 3. Evaluate classifiers
results = []
plt.figure(figsize=(8,6))

for name, model in models.items():
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    acc = accuracy_score(y_test, y_pred)
    prec = precision_score(y_test, y_pred)
    rec = recall_score(y_test, y_pred)
    f1 = f1_score(y_test, y_pred)
```



	Model	Accuracy	Precision	Recall	F1 Score	AUC
0	Logistic Regression	0.988304	0.990654	0.990654	0.990654	0.998102
1	Decision Tree	0.918129	0.934579	0.934579	0.934579	0.912602
2	Random Forest	0.935673	0.944444	0.953271	0.948837	0.991311
3	SVM	0.976608	0.981308	0.981308	0.981308	0.997810
4	KNN	0.959064	0.938596	1.000000	0.968326	0.982696