Work Integrated Learning Programmes Division
M.Tech (AIML/ DSE)
Machine Learning

Assignment - 2

Marks: 15   Submission Deadline: 15-Feb-2026

1. Overview

In this assignment, you are required to: - Implement multiple classification models - Build an interactive Streamlit web application to demonstrate your models - Deploy the app on Streamlit Community Cloud (FREE) - Share clickable links for evaluation

You will learn real-world end-to-end ML deployment workflow: modeling, evaluation, UI design, and deployment.

The assignment has to be performed on BITS Virtual Lab. If you are still facing any issue with this, please send an email to neha.vinayak@pilani.bits-pilani.ac.in with subject as "ML Assignment 2: BITS Lab issue" and get it resolved at the earliest.

2. Mandatory Submission Links

Each submission must be a single PDF file with the following (maintain the order):

1. GitHub Repository Link containing
   o Complete source code
   o requirements.txt
   o A clear README.md
2. Live Streamlit App Link
   o Deployed using Streamlit Community Cloud
   o Must open an interactive frontend when clicked
3. Screenshot
   o Upload screenshot of assignment execution on BITS Virtual Lab

4. The Github README content (details mentioned in Section 3 - Step 5) should also be part of the submitted PDF file.

As you are comfortable with the BITS Virtual Lab and Taxila Assignment submission process now, <u>only ONE submission will be accepted in Assignment 2</u>.
No Resubmission requests will be accepted.

3. Assignment Details

Step 1: Dataset choice

Choose <u>ONE</u> classification dataset of your choice from any public repository - Kaggle or UCI. It may be a binary classification problem or a multi-class classification problem.

Minimum Feature Size: 12
Minimum Instance Size: 500

Step 2: Machine Learning Classification models and Evaluation metrics

Implement the following classification models using the dataset chosen above. All the 6 ML models have to be implemented on the same dataset.

1. Logistic Regression
2. Decision Tree Classifier
3. K-Nearest Neighbor Classifier
4. Naive Bayes Classifier - Gaussian or Multinomial
5. Ensemble Model - Random Forest
6. Ensemble Model - XGBoost

For each of the models above, calculate the following evaluation metrics:

1. Accuracy
2. AUC Score
3. Precision
4. Recall
5. F1 Score

6. Matthews Correlation Coefficient (MCC Score)

The assignment has to be performed on BITS Virtual Lab and a (ONE) screenshot has to be uploaded as a proof of that. [1 mark]

Step 3: Prepare Your GitHub Repository

Your repository must contain:

```
project-folder/
|-- app.py   (or streamlit_app.py)
|-- requirements.txt
|-- README.md
|-- model/ (saved model files for all implemented models - *.py or *.ipynb)
```

Step 4: Create requirements.txt

Example:

```
streamlit
scikit-learn
numpy
pandas
matplotlib
seaborn
```

Missing dependencies are the #1 cause of deployment failure.

Step 5: README.md with the following structure. This README content should also be part of the submitted PDF file. Follow the required structure carefully.

a. Problem statement
b. Dataset description [1 mark]
c. Models used: [6 marks - 1 marks for all the metrics for each model]
   Make a Comparison Table with the evaluation metrics calculated for all the 6 models as below:

ML Model Name	Accuracy	AUC	Precision	Recall	F1	MCC
Logistic Regression						

Decision Tree						
kNN						
Naive Bayes						
Random Forest (Ensemble)						
XGBoost (Ensemble)						

- Add your observations on the performance of each model on the chosen dataset. [3 marks]

ML Model Name	Observation about model performance
Logistic Regression	
Decision Tree	
kNN	
Naive Bayes	
Random Forest (Ensemble)	
XGBoost (Ensemble)	

Step 6: Deploy on Streamlit Community Cloud
1. Go to https://streamlit.io/cloud
2. Sign in using GitHub account
3. Click "New App"
4. Select your repository
5. Choose branch (usually main)
6. Select app.py
7. Click Deploy

Within a few minutes, your app will be live.

Your Streamlit app must include at least the following features : -

    a. Dataset upload option (CSV) [As streamlit free tier has limited capacity, upload only test data] [1 mark]
    b. Model selection dropdown (if multiple models) [1 mark]
    c. Display of evaluation metrics [1 mark]
    d. Confusion matrix or classification report [1 mark]

5. Anti-Plagiarism & Academic Integrity Guidelines

To ensure originality we will be performing the following checks. **Any plagiarism found will result in ZERO (0) marks.**

Code-Level Checks
- GitHub commit history will be reviewed
- Identical repo structure & variable names may be flagged

UI-Level Checks
- Copy-paste Streamlit templates without customization may be penalized

Model-Level Checks
- Same dataset + same model + same outputs across students will be investigated

Using AI tools is allowed only for learning support, not for direct copy-paste submissions.

8. Final Submission Checklist (Before You Submit)

☐ GitHub repo link works

☐ Streamlit app link opens correctly

☐ App loads without errors

☐ All required features implemented

☐ README.md updated and added in the submitted PDF.

The assignment is for **15 Marks** - Model implementation and uploading on Github (10 marks) ; Streamlit App Development (4 marks). Additional 1 mark is for performing the assignment on BITS Lab and uploading a screenshot of the same.

No extension of deadlines will be provided. Please submit within the deadline - 15 Feb 23:59 PM.

No DRAFT submissions will be accepted. Please remember to SUBMIT your assignment.

Note: There is no leaderboard for this assignment and there will be no comparison of model performance across students.