

Lab#2 AIDI-2004: AI in Enterprise System  
100943365 – Akshay Bharadva

1. Initialized the repository and added the remote location.

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2
$ git init
Initialized empty Git repository in C:/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2/.git/
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git remote add origin https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git remote -v
origin https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git (fetch)
origin https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git (push)
```

2. Added the data given for the Lab#2 and then committed it.

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git status
On branch master

No commits yet

Untracked files:
  (use "git add <file>..." to include in what will be committed)
    data/

nothing added to commit but untracked files present (use "git add" to track)
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git add .
warning: in the working copy of 'data/wdbc.data', LF will be replaced by CRLF the next time Git touches it
warning: in the working copy of 'data/wdbc.names', LF will be replaced by CRLF the next time Git touches it
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git commit -m 'initialized the repo with the wdbc data'
[master (root-commit) 1e9361a] initialized the repo with the wdbc data
 2 files changed, 709 insertions(+)
 create mode 100644 data/wdbc.data
 create mode 100644 data/wdbc.names
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git log
commit 1e9361a8c5597d002113e199c9bc85a16682635d (HEAD -> master)
Author: Akshay Bharadva <akshaybharadva19@gmail.com>
Date: Tue Feb 6 10:38:48 2024 -0500

    initialized the repo with the wdbc data
```

### 3. Now, I'm creating the v1 model, steps included are

- Installing libraries
- Importing packages
- Importing data into pandas
- Splitting the data
- Selecting the model
- Training the model
- Evaluating it and checking the accuracy

#### Installing libraries

```
pip install pandas scikit-learn
```

✓ 0.8s

Requirement already satisfied: pandas in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: scikit-learn in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: numpy<2, ≥1.23.2 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: python-dateutil ≥2.8.2 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: pytz ≥2020.1 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: tzdata ≥2022.7 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: scipy ≥1.6.0 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: joblib ≥1.2.0 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: threadpoolctl ≥2.0.0 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

Requirement already satisfied: six ≥1.5 in c:\durham college\aidi\term 2\202441.11562-aidi-2004-01 - ai in

#### Importing the necessary libraries

```
> import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, classification_report
```

✓ 0.6s

#### Load the dataset into a pandas DataFrame

```
url = 'https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/wdbc.data'
column_names = ['id', 'diagnosis', 'radius_mean', 'texture_mean', 'perimeter_mean', 'area_mean', 'smoothness_mean', 'compactness_mean']
data = pd.read_csv(url, header=None, names=column_names)
```

✓ 1.0s

```
data.head()
```

✓ 0.0s

	id	diagnosis	radius mean	texture mean	perimeter mean	area mean	smoothness mean	compactness mean
0	842302	M	17.99	10.38	122.80	1001.0	0.11840	0.27760
1	842517	M	20.57	17.77	132.90	1326.0	0.08474	0.07864
2	84300903	M	19.69	21.25	130.00	1203.0	0.10960	0.15990
3	84348301	M	11.42	20.38	77.58	386.1	0.14250	0.28390
4	84358402	M	20.29	14.34	135.10	1297.0	0.10030	0.13280

5 rows × 9 columns

#### Preprocessing the data

```
X = data.drop(['id', 'diagnosis'], axis=1)
y = data['diagnosis']
y = y.map({'M': 1, 'B': 0})
```

✓ 0.0s

## Split the data into training and testing sets

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

✓ 0.0s

## Feature scaling

```
scaler = StandardScaler()  
X_train = scaler.fit_transform(X_train)  
X_test = scaler.transform(X_test)
```

✓ 0.0s

## Model building

```
model = LogisticRegression()  
model.fit(X_train, y_train)
```

✓ 0.0s

▼ LogisticRegression ⓘ ?  
LogisticRegression()

## Model evaluation

```
y_pred = model.predict(X_test)  
print("Accuracy:", accuracy_score(y_test, y_pred))  
print(classification_report(y_test, y_pred))
```

✓ 0.0s

Accuracy: 0.9736842105263158

	precision	recall	f1-score	support
0	0.97	0.99	0.98	71
1	0.98	0.95	0.96	43
accuracy			0.97	114
macro avg	0.97	0.97	0.97	114
weighted avg	0.97	0.97	0.97	114

#### 4. Committing the akshay\_bharadva\_model\_v1.ipynb to the master branch and pushing to the remote location.

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git push
fatal: The current branch master has no upstream branch.
To push the current branch and set the remote as upstream, use

    git push --set-upstream origin master

To have this happen automatically for branches without a tracking
upstream, see 'push.autoSetupRemote' in 'git help config'.

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git push --set-upstream origin master
Enumerating objects: 8, done.
Counting objects: 100% (8/8), done.
Delta compression using up to 12 threads
Compressing objects: 100% (7/7), done.
Writing objects: 100% (8/8), 56.45 KiB | 6.27 MiB/s, done.
Total 8 (delta 0), reused 0 (delta 0), pack-reused 0
To https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git
 * [new branch]      master -> master
branch 'master' set up to track 'origin/master'.
```

#### Checking to the remote.

The screenshot shows the GitHub interface for a repository named 'aidi-2004-lab-2-wdbc' by user 'akshay-bharadva'. The repository is public and has 1 branch (master) and 0 tags. The commit history shows two commits: 'initializing the repo with the wdbc data' (24 minutes ago) and 'adding trained model v1 ipynb' (3 minutes ago). The file list shows 'data' and 'akshay\_bharadva\_model\_v1.ipynb'. The README section is empty, with a prompt to 'Add a README'.

aidi-2004-lab-2-wdbc Public

master 1 Branch 0 Tags

Go to file Add file Code

akshay-bharadva adding trained model v1 ipynb 23f9ee0 · 3 minutes ago 2 Commits

File	Commit Message	Time
data	initialized the repo with the wdbc data	24 minutes ago
akshay_bharadva_model_v1.ipynb	adding trained model v1 ipynb	3 minutes ago

README

Add a README

Help people interested in this repository understand your project by adding a README.

Add a README

5. Created random-forest branch from the master branch by using the command,\
  - a. Git checkout -b random-forest

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git checkout -b random-forest
Switched to a new branch 'random-forest'

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git branch
  master
* random-forest
```

6. Adding the v2 model ipynb with the Random Forest Classifier Algorithm.

akshay\_bharadva\_model\_v2.ipynb U X

akshay\_bharadva\_model\_v2.ipynb > Importing the necessary libraries > import pandas as pd

+ Code + Markdown | ▶ Run All ↺ Restart ≡ Clear All Outputs | [V] Variables ≡ Outline ...

## Model building with Random Forest

```
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
```

[7] ✓ 0.2s

... **RandomForestClassifier** ⓘ ?

```
RandomForestClassifier(random_state=42)
```

## Model evaluation

```
y_pred = model.predict(X_test)
print("Random Forest Accuracy:", accuracy_score(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

[8] ✓ 0.0s

... Random Forest Accuracy: 0.9649122807017544

	precision	recall	f1-score	support
0	0.96	0.99	0.97	71
1	0.98	0.93	0.95	43
accuracy			0.96	114
macro avg	0.97	0.96	0.96	114
weighted avg	0.97	0.96	0.96	114

## Lab#2 AIDI-2004: AI in Enterprise System

100943365 – Akshay Bharadva

### 7. Committing akshay\_bharadva\_model\_v2.ipynb and pushing it to the remote.

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git status
On branch random-forest
Untracked files:
  (use "git add <file>..." to include in what will be committed)
    akshay_bharadva_model_v2.ipynb

nothing added to commit but untracked files present (use "git add" to track)

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git add .
warning: in the working copy of 'akshay_bharadva_model_v2.ipynb', LF will be replaced by CRLF the next time Git touches it

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git status
On branch random-forest
Changes to be committed:
  (use "git restore --staged <file>..." to unstage)
    new file:   akshay_bharadva_model_v2.ipynb

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git commit -m 'adding random-forest-classification model in v2 ipynb'
[random-forest 2c8cca1] adding random-forest-classification model in v2 ipynb
1 file changed, 823 insertions(+)
create mode 100644 akshay_bharadva_model_v2.ipynb
```

### Checking git log,

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git log
commit 2c8cca14a0e76b47a3b92c52cdc69b1e4a50409b (HEAD -> random-forest)
Author: Akshay Bharadva <akshaybharadva19@gmail.com>
Date:   Tue Feb 6 11:18:32 2024 -0500

    adding random-forest-classification model in v2 ipynb

commit 23f9ee03b1f4067a61e4fbc5a3a7bf48587f1d04 (origin/master, master)
Author: Akshay Bharadva <akshaybharadva19@gmail.com>
Date:   Tue Feb 6 11:00:32 2024 -0500

    adding trained model v1 ipynb

commit 1e9361a8c5597d002113e199c9bc85a16682635d
Author: Akshay Bharadva <akshaybharadva19@gmail.com>
Date:   Tue Feb 6 10:38:48 2024 -0500

    initialized the repo with the wdbc data
```

### Pushing to the remote,

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git push
fatal: The current branch random-forest has no upstream branch.
To push the current branch and set the remote as upstream, use

    git push --set-upstream origin random-forest

To have this happen automatically for branches without a tracking
upstream, see 'push.autoSetUpRemote' in 'git help config'.

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git push --set-upstream origin random-forest
Enumerating objects: 4, done.
Counting objects: 100% (4/4), done.
Delta compression using up to 12 threads
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 6.18 KiB | 6.18 MiB/s, done.
Total 3 (delta 0), reused 0 (delta 0), pack-reused 0
remote:
remote: Create a pull request for 'random-forest' on GitHub by visiting:
remote:   https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc/pull/new/random-forest
remote:
To https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git
 * [new branch]      random-forest -> random-forest
branch 'random-forest' set up to track 'origin/random-forest'.
```

## Lab#2 AIDI-2004: AI in Enterprise System

100943365 – Akshay Bharadva

### 8. Checking the remote location

The screenshot shows the GitHub interface for the repository 'aidi-2004-lab-2-wdbc'. At the top, there's a notification that 'random-forest' had recent pushes 39 seconds ago, with a 'Compare & pull request' button. Below this, the repository name 'random-forest' is shown with '2 Branches' and '0 Tags'. A search bar 'Go to file' and buttons 'Add file' and '<> Code' are visible. A status bar indicates 'This branch is 1 commit ahead of master'. The commit history shows three commits by 'akshay-bharadva':

- Commit 2c8cca1 (1 minute ago): 'adding random-forest-classification model in v2 ipynb' (3 Commits)
- Commit 41 minutes ago: 'initialized the repo with the wdbc data' (Folder: data)
- Commit 19 minutes ago: 'adding trained model v1 ipynb' (File: akshay\_bharadva\_model\_v1.ipynb)
- Commit 1 minute ago: 'adding random-forest-classification model in v2 ipynb' (File: akshay\_bharadva\_model\_v2.ipynb)

### 9. Merging the random-forest branch into master branch so that master have all the code from random-forest branch.

```
aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (random-forest)
$ git checkout master
Switched to branch 'master'
Your branch is up to date with 'origin/master'.

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git merge random-forest
Updating 23f9ee0..2c8cca1
Fast-forward
 akshay_bharadva_model_v2.ipynb | 823 +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
 1 file changed, 823 insertions(+)
 create mode 100644 akshay_bharadva_model_v2.ipynb

aksha@YOGA MINGW64 /c/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git push
Total 0 (delta 0), reused 0 (delta 0), pack-reused 0
To https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git
 23f9ee0..2c8cca1 master -> master
```

Checking the remote repository after merging and pushing,

The screenshot shows the GitHub repository page after the merge and push. The 'random-forest' branch is no longer visible, and the 'master' branch is selected. The commit history now shows three commits by 'akshay-bharadva':

- Commit 2c8cca1 (3 minutes ago): 'adding random-forest-classification model in v2 ipynb' (3 Commits)
- Commit 42 minutes ago: 'initialized the repo with the wdbc data' (Folder: data)
- Commit 21 minutes ago: 'adding trained model v1 ipynb' (File: akshay\_bharadva\_model\_v1.ipynb)
- Commit 3 minutes ago: 'adding random-forest-classification model in v2 ipynb' (File: akshay\_bharadva\_model\_v2.ipynb)

A 'README' file is also visible at the bottom of the commit list.

## Lab#2 AIDI-2004: AI in Enterprise System

100943365 – Akshay Bharadva

### 10. Creating the README.md file for the description of the Lab#2 and tasks performed.

```
aksha@YOGA MINGW64 /c:/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git status
On branch master
Your branch is up to date with 'origin/master'.

Untracked files:
  (use "git add <file>..." to include in what will be committed)
        README.md

nothing added to commit but untracked files present (use "git add" to track)

aksha@YOGA MINGW64 /c:/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git add .

aksha@YOGA MINGW64 /c:/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git status
On branch master
Your branch is up to date with 'origin/master'.

Changes to be committed:
  (use "git restore --staged <file>..." to unstage)
        new file:   README.md

aksha@YOGA MINGW64 /c:/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git commit -m 'added README.md file'
[master 8a742c4] added README.md file
1 file changed, 12 insertions(+)
create mode 100644 README.md

aksha@YOGA MINGW64 /c:/Durham College/AIDI/Term 2/202441.11562-AIDI-2004-01 - AI IN ENTERPRISE SYSTEMS/Lab#2 (master)
$ git push
Enumerating objects: 4, done.
Counting objects: 100% (4/4), done.
Delta compression using up to 12 threads
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 555 bytes | 555.00 KiB/s, done.
Total 3 (delta 1), reused 0 (delta 0), pack-reused 0
remote: Resolving deltas: 100% (1/1), completed with 1 local object.
To https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc.git
2c8ccal..8a742c4 master -> master
```

The screenshot shows the GitHub repository page for 'aidi-2004-lab-2-wdbc' by user 'akshay-bharadva'. The repository is public and has 2 branches and 0 tags. The file list shows: 'data' (initialized the repo with the wdbc data, 49 minutes ago), 'README.md' (added README.md file, now), 'akshay\_bharadva\_model\_v1.ipynb' (adding trained model v1 ipynb, 27 minutes ago), and 'akshay\_bharadva\_model\_v2.ipynb' (adding random-forest-classification model in v2 ipynb, 9 minutes ago). The README file is selected, showing the title 'Lab#2 - Create two branches with different ML model from given data and merge'. The content includes a bulleted list of steps: 1. master branch (contains all the files), 2. random-forest branch (created this branch from master, created v2 ML model ipynb with RandomForestClassifier), 3. afterwards, merged random-forest branch into master, 4. created README.md file with proper description. A 'Reference' section at the bottom lists the data source: <https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic>.

GitHub URL: <https://github.com/akshay-bharadva/aidi-2004-lab-2-wdbc>