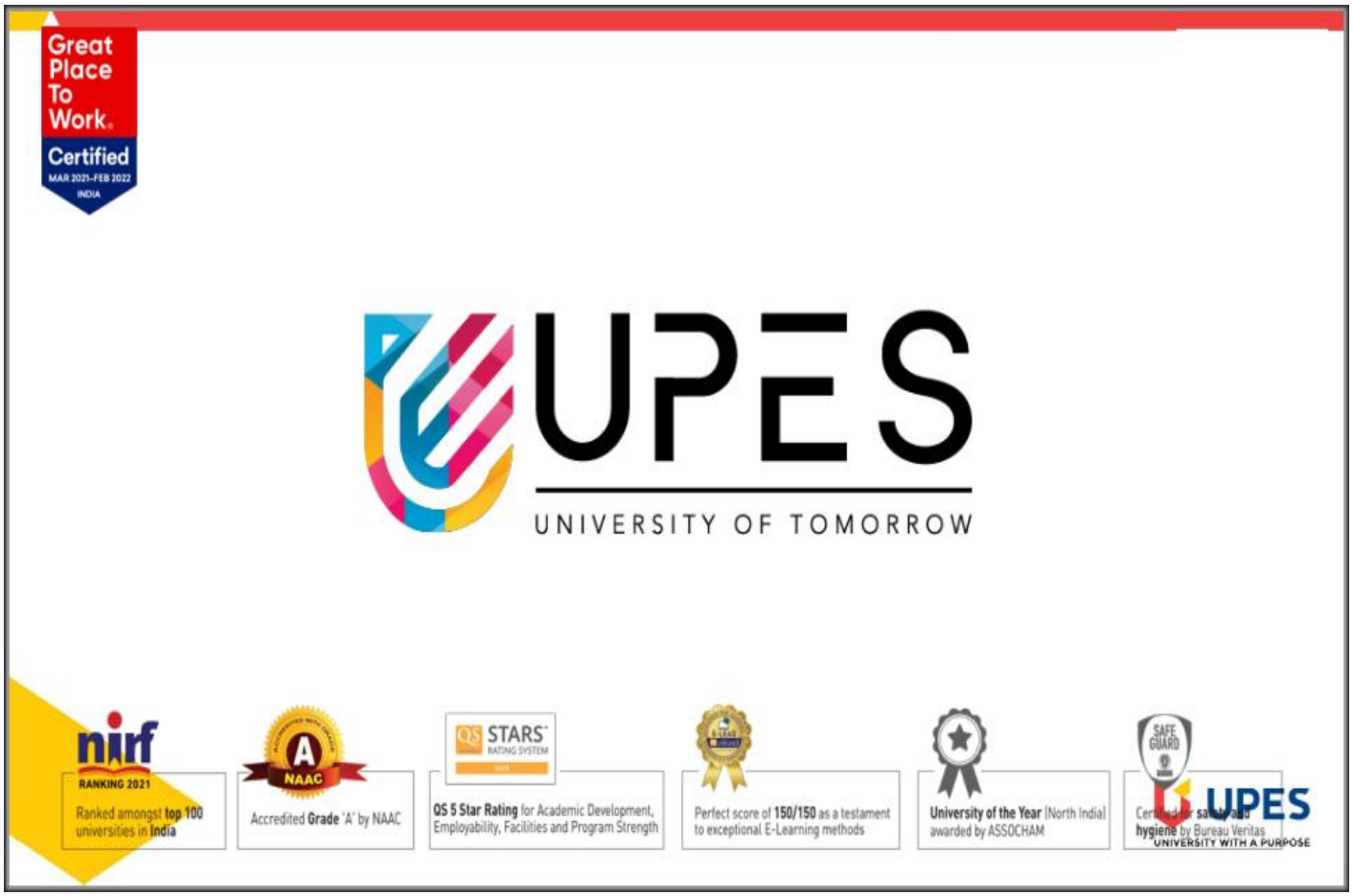
OVERVIEW OF DIFFERENT VALIDATION METHODS FOR DIFFERENT CLASSIFICATION MODELS



Name:
Kapil Singh Negi

Course / Department:
B.Tech CSE – School of Computer Science

Institution / Organization:
UPES, Dehradun, Uttarakhand, Bharat

Date:
01 December 2025

Overview

The `code.py` file contains a Python script that implements a command-line interface (CLI) for evaluating and comparing various machine learning classification models. The script uses the Iris dataset, which it loads from `archive/Iris.csv`.

The main functionalities of the script are:

- Allows the user to select a cross-validation method (Holdout, LOOCV, Stratified K-Fold, K-Fold, Repeated K-Fold).
- Provides an option to shuffle the dataset.
- Runs evaluations of three different models: K-Nearest Neighbours, Logistic Regression, and a Decision Tree.
- Displays the accuracy of each model.
- Visualizes the confusion matrices for each model using `matplotlib`.

Dependencies

The script relies on the following external libraries:

- `pandas`: For data manipulation and reading the CSV file.
- `numpy`: For numerical operations.
- `scikit-learn`: For machine learning models, metrics, and cross-validation strategies.
- `matplotlib`: For plotting the confusion matrices.

Execution Analysis

The script is intended to be run as a standalone Python script from the command line.

```
#!bash
python code.py
```

When executed, it will start an interactive menu in the console. The user can then select different options to configure the evaluation and run the models. The script depends on the `archive/Iris.csv` file being present in the correct location relative to the script.

GitHub link

https://github.com/Exist-Er/VALIDATION-METHODS.git