***Title***:- Testing Various Machine Learning Models and comparing them with the Neural Tangent Kernel.

***Description:-*** The code trains various models on 105 datasets and outputs their accuracies to a csv file called output\_trial.csv.

***Dependencies:-***

* os :- Provides a way to interact with the os and access files.
* Arff :- Helps with handling the .arff format files.
* Pandas:- Library for data manipulation and analysis.
* Numpy:- Library for numerical calculations.
* scikit-learn:- Library that provides various classifiers and evaluation metrics.
* Tensorflow:- Library that helps in building Neural Nets.

***Functions:-***

1. DTC(X\_train, y\_train, X\_test, y\_test) :

Trains a Decision Tree Classifier and returns the accuracy of the model on the test set.

Parameters:-

* 1. X\_train :- Training data features.
  2. Y\_train :- Training data labels.
  3. X\_test:- Test data features.
  4. Y\_test:- Test data labels.

1. RFC(X\_train, y\_train, X\_test, y\_test) :

Trains a Random Forest with 500 trees and returns the accuracy of the model on the test set.

Parameters:-

* 1. X\_train :- Training data features.
  2. Y\_train :- Training data labels.
  3. X\_test:- Test data features.
  4. Y\_test:- Test data labels.

1. svmlin(X\_train, y\_train, X\_test, y\_test) :

Trains a Support Vector Machine with a linear kernel and returns the accuracy of the model on the test set.

Parameters:-

* 1. X\_train :- Training data features.
  2. Y\_train :- Training data labels.
  3. X\_test:- Test data features.
  4. Y\_test:- Test data labels.

1. Svmpolydeg2(X\_train, y\_train, X\_test, y\_test) :

Trains a Support Vector Machine with polynomial features of degree 2 and returns the accuracy of the model on the test set.

Parameters:-

* 1. X\_train :- Training data features.
  2. Y\_train :- Training data labels.
  3. X\_test:- Test data features.
  4. Y\_test:- Test data labels.

1. Svmpolydeg3(X\_train, y\_train, X\_test, y\_test) :

Trains a Support Vector Machine with polynomial features of degree 3 and returns the accuracy of the model on the test set.

Parameters:-

* 1. X\_train :- Training data features.
  2. Y\_train :- Training data labels.
  3. X\_test:- Test data features.
  4. Y\_test:- Test data labels.

1. svmgaussian(X\_train, y\_train, X\_test, y\_test) :

Trains a Support Vector Machine with a Gaussian RBF kernel and returns the accuracy of the model on the test set.

Parameters:-

* 1. X\_train :- Training data features.
  2. Y\_train :- Training data labels.
  3. X\_test:- Test data features.
  4. Y\_test:- Test data labels.

The datasets have been chosen as they have a wide variety of datasets with varying sizes. The datasets are also pre-scaled meaning that no sort of further data manipulation and feature engineering was required and the raw data itself could be used. Some datasets were of different structure and had to be trained separately.

The datasets were of .arff format and had to be opened using the arff library.

The datasets were taken from :