

DATA MINING

Final Project Report

Professor: Yasser Abduallah

Department of Computer Science

Teja Kulkarni

UCID – tk55

Gitlink - [https://github.com/TejaKulkarni/
FinalDataMining](https://github.com/TejaKulkarni/FinalDataMining)

1. Which Option Selected ?

—> I have selected the 1st option. The 4 different classification algorithms which I would like to implement are :

- Naïve Bayes
- Random Forest
- LSTM

2. What software you are planning to use? —> The software I am planning to use are :

- Python
- NumPy
- Pandas
- Jupyter Notebook

3. Hardware configuration.

—> The Hardware configuration are :

- Hard Disk – 1TB
- RAM – 8 GB
- OS – MAC OS
- Processor – Dual-Core Intel Core i5

4. Framework, if applied.

—> The following is the framework I applied,

- Import required packages and libraries
- Loading datasets Cleaning and editing data

- Features Engineering
- Data visualization and analysis
- Testing and training of data
- Modelling 3 algorithms
- Comparing the outputs and analysis of results of all algorithms

DATASET USED

train.csv

JUPYTER NOTEBOOK

Implementing existing Classifier libraries to the above dataset.

SCREENSHOTS IS ATTACHED IN THE PDF COPY.

