**Report file HW2**

**(Xingyu Huang, Chunlei Wang, Wenbo Zhao, and Soyeon Park)**

**# The multinomial Naive Bayes algorithm**

* Data1

Accuracy on test set) 94.77%

* Data2

Accuracy on test set: 94.74%

* Data3

Accuracy on test set: 85.08%

→ We coded the algorithm based on the multinomial Naïve Bayes and Laplace smoothing. Additionally, we utilized stop-words for better accuracy on the test sets, leading the high accuracy above 85% in all data sets.

**# The MCAP Logistic Regression algorithm with L2 regularization**

|  |  |  |  |
| --- | --- | --- | --- |
| Lambda | Accuracy % | | |
| Dataset 1 | Dataset 2 | Dataset 3 |
| 0.001 | 92.16 | 91.47 | 91.25 |
| 0.01 | 93.52 | 92.96 | 93.54 |
| 0.1 | 94.24 | 93.74 | 94.69 |
| 1 | 94.97 | 93.29 | 94.69 |

* Data1

Lambda = 1

Eta = 0.0001

Iterations = 40

Accuracy is 94.97%

* Data2

Lambda = 1

Eta = 0.0001

Iterations = 40

Accuracy is 93.29%

* Data3
* Lambda = 1
* Eta = 0.0001
* Iterations = 40
* Accuracy is 94.69%

**# The perceptron algorithm**

|  |  |  |  |
| --- | --- | --- | --- |
| Iterations | Accuracy % | | |
| Dataset 1 | Dataset 2 | Dataset 3 |
| 10 | 73.51 | 70.25 | 65.34 |
| 20 | 84.58 | 87.68 | 78.12 |
| 30 | 89.24 | 90.12 | 87.35 |
| 40 | 92.45 | 91.25 | 90.38 |

* Data1
* Eta = 0.0001

Iterations = 40

Accuracy is 92.45%

* Data2:
* Eta = 0.0001

Iterations = 40

Accuracy is 91.25%

* Data3
* Eta = 0.0001
* Iterations = 40
* Accuracy is 90.38%