**Observations Criteria:**

Note:

1. **preprocessed\_data** was used for training data processing. The training data/tweets are converted into tokens of words after punctuation removal and lowercase conversion.

* **MLP model accuracy with pre-trained embeddings from the second task:**
* **MLPClassifier – For better understanding, didn’t change the default features and used the default solver(adam) and activation function (ReLu)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Vector size** | **Number of neurons** | **Accuracy** | **Prediction** |
| **Layer 1** | 500 | 50 | 70.3% | Off and Non-offensive |
| **Layer 2** | 500 | 30,50 | **71.74%** | Off and Non-offensive |
| **Layer 3** | 500 | 30,50,50 | **71.74%** | Off and Non-offensive |

* **MLP model accuracy with fine-trained embeddings:**
* **MLPClassifier – For better understanding, didn’t change the default features and used the default solver(adam) and activation function (ReLu)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Vector size** | **Number of neurons** | **Accuracy** | **Prediction** |
| **Layer 1** | 500 | 50 | 67.09% | Off and Non-offensive |
| **Layer 2** | 500 | 30,50 | 60% | Off and Non-offensive |
| **Layer 3** | 500 | 30,50,50 | 57.09% | Off and Non-offensive |

* **With default layer settings, vector size: 500:**

Accuracy: 60.23%

* **Default layer settings, vector size: 300:**

Accuracy: 72.21

**SUMMARY OF ANALYSIS OF RESULTS:**

1. The MLP model with pre-trained embeddings has 71.74% of accuracy when the model is trained on two and three layers.
2. Surprisingly, when fine-tuned embeddings are used with three layers of the same configuration as the pre-trained embedding MLP model, the accuracy dropped to 57.09%.
3. But with default layers, the accuracy is around 72.2% better than the MLP model trained on pre-trained embeddings, when the dimensionality of the word vector is 300.
4. Fine-tuned embeddings are a transfer learning technique because the model already learned so much about the dataset and can relate better as compared to the MLP model trained on the pre-trained embeddings.
5. When the word vector size is increased to 500 with default layer settings the accuracy was 60.23%, which could be obvious because the model already learned most of the information from the dataset, and advancing the word vector doesn’t make sense.
6. Overall, it seems fine-tuned model performed well as compared to the pre-trained model without any further additional layers and settings for the fine-tuned model. As it is already learned so much about the dataset through fine-tuning of pre-trained embeddings.