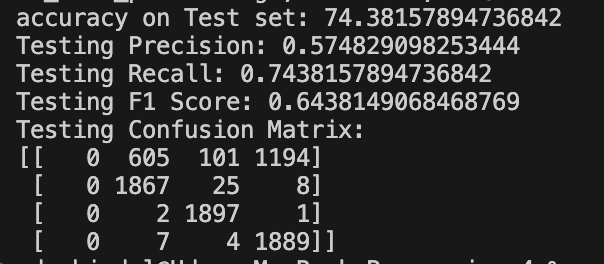
Assignment – 4

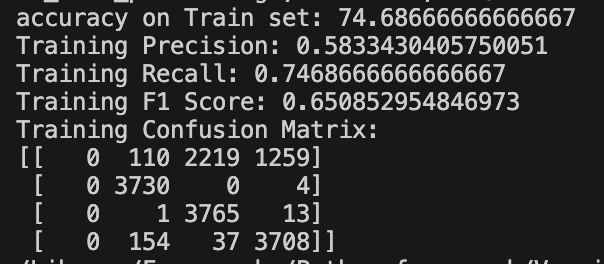
Train Accuracy: 74.6

Test Accuracy: 74.3

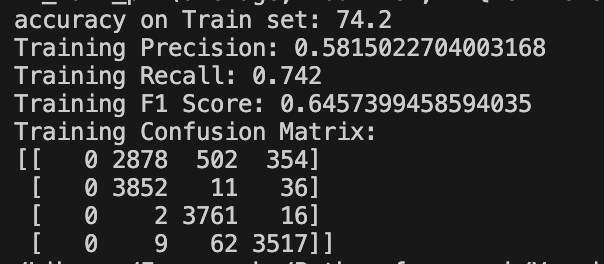
**Trainable λs:**

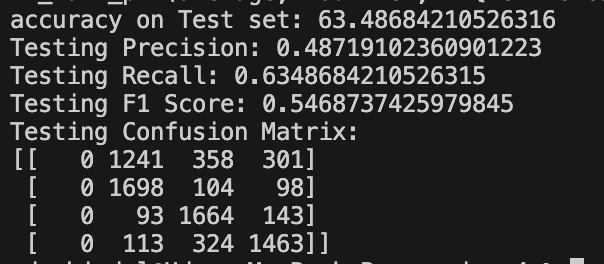
Best λs: [ 0.03010594 -0.01414574 -0.00537632]





**Frozen λs:**





**Analysis:**

**ELMo (Embeddings from Language Models):** ELMo embeddings are contextual word representations generated from pretraining on large text corpora using deep bidirectional LSTMs. They capture the contextual meaning of words based on their surrounding context.

**SVD (Singular Value Decomposition):** SVD is a dimensionality reduction technique used to generate word embeddings from a co-occurrence matrix of words in a corpus. It captures global statistical patterns in the data.

**Word2Vec:** Word2Vec is a shallow neural network-based model that learns word embeddings by predicting context words given a target word or vice versa. It captures semantic relationships between words.

**Performance Metrics:**

Test Set:

| **Model** | **Accuracy** | **F1 Score** | **Precision** | **Recall** |
| --- | --- | --- | --- | --- |
| Word2Vec | 50.2 | 0.48 | 0.77 | 0.50 |
| SVD | 69.92 | 0.69 | 0.72 | 0.69 |
| ELMo | 74.38 | 0.65 | 0.57 | 0.74 |

Train Set:

| **Model** | **Accuracy** | **F1 Score** | **Precision** | **Recall** |
| --- | --- | --- | --- | --- |
| Word2Vec | 54.3 | 0.53 | 0.58 | 0.54 |
| SVD | 69.92 | 0.75 | 0.77 | 0.75 |
| ELMo | 74.6 | 0.65 | 0.58 | 0.74 |

**ELMo vs. Word2Vec vs. SVD:**

* ELMo outperforms both Word2Vec and SVD across all evaluation metrics.
* ELMo captures more contextual information due to its deep bidirectional architecture, leading to better representations for sentiment analysis.
* Word2Vec performs better than SVD, possibly because it captures semantic relationships between words through distributed representations.
* ELMo's contextual embeddings capture nuances in language and context-specific meanings, making it more suitable for downstream tasks like sentiment analysis.
* The superior performance of ELMo can be attributed to its ability to capture syntactic and semantic information from large-scale corpora during pretraining.