HiPool, BERT - 20 Epochs

\*\*\* avg\_loss : 0.67, time : ~14.0 min (873.03 sec) \*\*\*

==> evaluation : avg\_loss = 1.23, time : 58.87 sec

=====> {'accuracy': 0.536, 'nb exemple': 1000, 'true\_prediction': 536, 'false\_prediction': 464}

###############################################################################

GCN, BERT - 20 Epochs

\*\*\* avg\_loss : 0.20, time : ~14.0 min (865.24 sec) \*\*\*

==> evaluation : avg\_loss = 2.18, time : 60.19 sec

=====> {'accuracy': 0.57, 'nb exemple': 1000, 'true\_prediction': 570, 'false\_prediction': 430}

§§ model has been saved §§

###############################################################################

Graphsage, BERT - 20 Epochs

\*\*\* avg\_loss : 0.13, time : ~14.0 min (859.54 sec) \*\*\*

==> evaluation : avg\_loss = 2.17, time : 58.48 sec

=====> {'accuracy': 0.563, 'nb exemple': 1000, 'true\_prediction': 563, 'false\_prediction': 437}

###############################################################################

GAT, BERT - 20 Epochs

\*\*\* avg\_loss : 0.05, time : ~14.0 min (867.72 sec) \*\*\*

==> evaluation : avg\_loss = 2.56, time : 58.80 sec

=====> {'accuracy': 0.582, 'nb exemple': 1000, 'true\_prediction': 582, 'false\_prediction': 418}

§§ model has been saved §§

Robustness Results - If the misclassified examples have very small differences in probability between neighboring classes, it indicates that the model's decision boundaries for those examples are near the decision boundary between those classes. This scenario can make the classification task challenging, and small changes in the input or model parameters might lead to different predictions.

1. Model : GCN Variant on Amazon 2048

Accuracy - {'accuracy': 0.585, 'nb exemple': 1000, 'true\_prediction': 585, 'false\_prediction': 415}

Jumbling - 0.8

Dropping Words - 1.0

Misspelled Words - 0.8

Verdict : The randomly selected sample of 10 correctly classified examples demonstrates an accuracy ranging between 70-80%. Based on this sample, we observe that the sample mean accuracy is indicative of the population mean accuracy for the classification task. This inference is completely following the constraints, the random selection was unbiased and representative of the overall population.

1. Model : Graph Sage Variant on Amazon 2048

Accuracy - {'accuracy': 0.534, 'nb exemple': 1000, 'true\_prediction': 534, 'false\_prediction': 466}

Jumbling - 0.6

Dropping Words - 0.8

Misspelled Words - 0.8

Verdict : Graph Sage was vulnerable when the words were jumbled. Despite the vulnerability to jumbled words, the model achieved an accuracy level of around 80% in dropping the words or misspelling words. Jumbling words disrupts the natural order and structure of sentences, challenging the model's ability to understand and classify the content correctly. In the context of long document sentence classification, it implies that the model is structured to handle sentences within the context of a longer document, possibly by considering relationships or dependencies between sentences.

1. Model : GAT Variant on Amazon 2048

Accuracy - {'accuracy': 0.579, 'nb exemple': 1000, 'true\_prediction': 579,

'false\_prediction': 421}

Jumbling - 0.6

Dropping Words - 0.4

Misspelled Words - 0.7

Verdict : The lower accuracy (0.6) might be due to the fact that your GAT variant is sensitive to the order of words in long sequences. GATs, like other graph-based models, often rely on the sequential information of input data. Jumbling disrupts this sequential order, leading to a decrease in accuracy. Graph-based models, including GAT, rely on relationships between nodes (words in this context). Dropping words could break important connections, leading to a loss of crucial context for classification. The lower accuracy (0.4) might be a result of this loss.The GAT variant might still perform reasonably well (0.7 accuracy) because it can capture semantic relationships between words even with some noise. The model might have learned to generalize and recognize patterns despite the presence of misspelled words.

1. Model : Hi Pool Variant

Accuracy - {'accuracy': 0.595, 'nb exemple': 1000, 'true\_prediction': 585, 'false\_prediction': 415}

Jumbling - 0.8

Dropping Words - 0.8

Misspelled Words - 0.7

Verdict :

Hierarchical Pooling: The use of hierarchical pooling allows the model to aggregate information at different levels of granularity, potentially capturing important patterns in long sequences.

Cross-Layer Attention: The introduction of cross-layer attention between different layers helps the model attend to relevant information across layers. This attention mechanism may contribute to the model's ability to handle variations introduced by jumbling, dropping words, and misspelled words.

The hierarchical pooling mechanism may help maintain informative features even with reduced input size, leading to higher accuracy

The HiPool variant might be more robust to changes in word order introduced by jumbling. The model's ability to capture important information across different word positions could contribute to higher accuracy

The hierarchical pooling and attention mechanisms might allow the model to focus on relevant information and mitigate the impact of misspellings to some extent.

For 2048 Tokens, both GCN and HiPool have performed robustly well.

Robustness - A512

1. Model - GCN Variant:

{'accuracy': 0.6305058336183635,

'nb exemple': 52626,

'true\_prediction': 33181,

'false\_prediction': 19445}

Jumbling - 0.7

Dropping Words - 0.7

Misspelled Words - 0.8

Verdict : The results suggest that the model performs relatively better with misspelled words as compared to jumbling or dropping words. Noisy or mislabeled data can negatively impact model performance due to the fact that, For graph-based models, the construction of the graph (relations between nodes) is crucial and need to ensure that the graph structure captures relevant relationships in the data. A more complex model may capture intricate patterns in the data but could also be prone to overfitting or be sensitive to specific types of variations.

1. Model - GAT Variant:

{'accuracy': 0.6329950974803329,

'nb exemple': 52626,

'true\_prediction': 33312,

'false\_prediction': 19314}

Jumbling - 0.6

Dropping Words - 0.7

Misspelling - 0.8

Verdict : GAT is effective in learning expressive node representations. The attention mechanism helps nodes to selectively aggregate information from their neighbors, leading to more informative embeddings. It can adaptively assign attention weights to different edges, allowing the model to capture complex relationships in the data. This might be the reason why GAT exhibits better robustness to irregularities or noise in the graph structure, contributing to improved generalization on the test set.

1. Model - Graph Sage Variant

{'accuracy': 0.6308858739026337,

'nb exemple': 52626,

'true\_prediction': 33201, 'false\_prediction': 19425}

In [12]:

Jumbling - 0.6

Dropping Words - 0.7

Misspelled Words - 0.7

Verdict : GraphSAGE is known for its ability to generate node representations by sampling and aggregating information from local neighborhoods. The effectiveness of this strategy depends on how well it captures the global characteristics of the graph. If the graph has long-range dependencies or complex relationships that extend beyond the sampled neighborhoods, it might affect the model's ability to make accurate predictions. If the perturbations are too severe, they might hinder the model's ability to generalize.

1. Model - HiPool Variant

{'accuracy': 0.6454224147759663,

'nb exemple': 52626,

'true\_prediction': 33966,

'false\_prediction': 18660}

Jumbling - 0.8

Dropping Words - 0.8

Misspelled Words - 0.8

Verdict: The use of hierarchical pooling allows the model to aggregate information at different levels of granularity. This hierarchical representation might help the model capture important patterns in long sequences. The pooling mechanism could contribute to maintaining informative features even when the input size is reduced, leading to higher accuracy. The hierarchical pooling mechanism, combined with cross-layer attention, could make the HiPool variant more robust to changes in word order introduced by jumbling. The model's ability to capture important information across different word positions may contribute to its higher accuracy in scenarios where the word order is altered. The combination of hierarchical pooling and cross-layer attention creates a robust architecture that excels in capturing both local and global dependencies in long sequences. This robustness contributes to the model's effectiveness across various perturbations

Multilinguality

A2048

GAT:

4th epoch

\*\*\*\*\*\*\*\* Running time this step.. 700.5385463237762

\*\*\* avg\_loss : 1.01, time : ~11.0 min (700.54 sec) \*\*\*

==> evaluation : avg\_loss = 1.13, time : 47.85 sec

=====> {'accuracy': 0.538, 'nb exemple': 500, 'true\_prediction': 269, 'false\_prediction': 231}

accuracy is 0.538

§§ Best model saved at epoch 4 §§

10th epoch:

\*\*\*\*\*\*\*\* Running time this step.. 700.0875301361084

\*\*\* avg\_loss : 0.33, time : ~11.0 min (700.09 sec) \*\*\*

==> evaluation : avg\_loss = 2.06, time : 47.75 sec

=====> {'accuracy': 0.454, 'nb exemple': 500, 'true\_prediction': 227, 'false\_prediction': 273}

**Hipool with Xlmroberta**

9th epoch

\*\*\*\*\*\*\*\* Running time this step.. 892.7600862979889

\*\*\* avg\_loss : 1.02, time : ~14.0 min (892.76 sec) \*\*\*

==> evaluation : avg\_loss = 1.23, time : 47.30 sec

=====> {'accuracy': 0.53, 'nb exemple': 500, 'true\_prediction': 265, 'false\_prediction': 235}

Best loss epoch: 10:

\*\*\*\*\*\*\*\* Running time this step.. 890.4723942279816

\*\*\* avg\_loss : 0.96, time : ~14.0 min (890.47 sec) \*\*\*

==> evaluation : avg\_loss = 1.35, time : 47.19 sec

=====> {'accuracy': 0.51, 'nb exemple': 500, 'true\_prediction': 255, 'false\_prediction': 245}

Best accuracy epoch:

\*\*\*\*\*\*\*\* Running time this step.. 893.6397342681885

\*\*\* avg\_loss : 1.05, time : ~14.0 min (893.64 sec) \*\*\*

==> evaluation : avg\_loss = 1.25, time : 47.20 sec

=====> {'accuracy': 0.54, 'nb exemple': 500, 'true\_prediction': 270, 'false\_prediction': 230}

**BERT: HIPool:**

\*\*\*\*\*\*\*\* Running time this step.. 700.099689245224

\*\*\* avg\_loss : 0.24, time : ~11.0 min (700.10 sec) \*\*\*

==> evaluation : avg\_loss = 2.17, time : 47.58 sec

=====> {'accuracy': 0.456, 'nb exemple': 500, 'true\_prediction': 228, 'false\_prediction': 272}

accuracy is 0.456

§§ Best model saved at epoch 6 §§

future తరాలు shakes head పాపం ఇది కథ isn t గురించి అమ్మ లేదా i ఉంటుంది కలిగి ఉంటాయి stopped reading అది after a కొన్ని chapters i కలిగి ఉంటాయి ఒక పాతది స్నేహితుడు whose ఫోన్ calls i dread because ఆమె కలిగి ఉంది ఉంది పై a మార్గం కు నాశనం ఆమె life కోసం a while మరియు every సమయం