# Instructions for ACL 2023 Proceedings

Anonymous ACL submission

## Abstract

This is abstract

## 1 Introduction

## 2 Previous studies

## 3 Methodology

## 4 Evaluating Sentence Embeddings with Gaussian Mixture

This section presents an unsupervised approach for evaluating the separability of sentence embeddings. We measure label separability pairwisely using a Gaussian Mixture model and calculate the F-score of the unsupervised clustered labels with the true labels.

The Gaussian Mixture model is a probabilistic model that assumes that each cluster follows a Gaussian (or normal) distribution and estimates the weight of the density function for each cluster (Reynolds et al., 2009; Singh et al., 2010). We assume that the sentence vectors of two distinct classes should achieve high accuracy with Gaussian Mixture if they are displayed in a Gaussian distribution in space and are separable from each other.

However, there is a potential limitation to using this method. The separability and accuracy score may be underestimated if two clusters are not normally distributed, as illustrated in Figure 1.

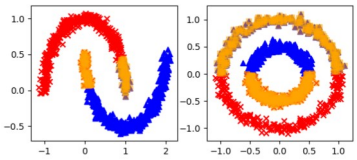

Figure 1: Cases that Gaussian Mixture Underestimate the Separability of Two Clusters

Our evaluation results show that out of 78 pairs of derivation classes, 7 pairs achieved an accuracy score above 90%, 24 pairs above 80%, and 47 pairs above 60%, as depicted in Figure 2.

In particular, the class 'ban' demonstrates good separability with many other classes, achieving an accuracy score of 0.982 with 'past', 0.980 with 'formal sentence', 0.942 with 'minimal change', and 0.937 with 'future', among other pairs.

Additionally, our evaluation results reveal that tenses are generally well-separated, with an accuracy score of 0.90 for 'past' and future' classes, and an accuracy score of 0.924 for 'simple sentence' and 'future' classes.

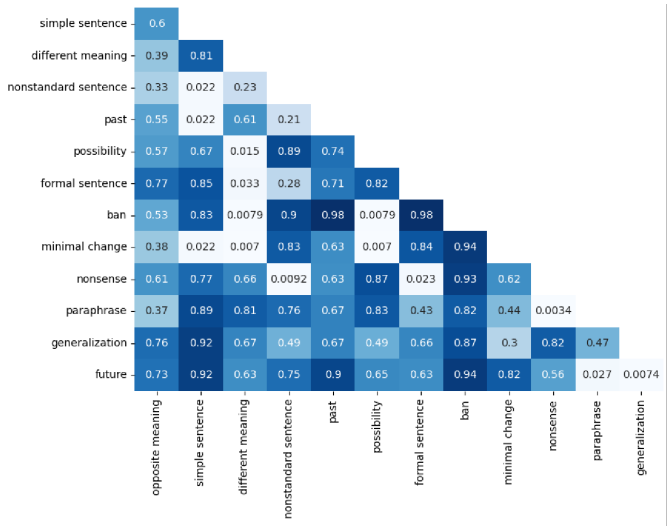


Figure 2: Accuracy of Measured Separability with Gaussian Mixture

## References

Douglas A Reynolds et al. 2009. Gaussian mixture models. *Encyclopedia of biometrics*, 741(659-663).

Ravindra Singh, Bikash C. Pal, and Rabih A. Jabr. 2010. Statistical representation of distribution system loads using gaussian mixture model. *IEEE Transactions on Power Systems*, 25(1):29–37.