IS 584 Course Term Project Final Report

Özkan Tuğberk Kartal
Electrical-Electronics Engineering
METU
Ankara, TURKEY
tugberk.kartal@metu.edu.tr

***Abstract*—This document deals with the problem of predicting the acceptance/rejection decision of conference articles by involving capsule layer architecture. The proposed model is compared with Global Vectors for Word Representation (GloVe) pre-trained word vectors. Accuracy, F1-Score, Area Under the Receiver Operating Characteristic (ROC) Curve and Confusion Matrix are calculated and stored to evaluate the performance of both of GloVe embeddings and Capsule Layer Architecture. Results are saved into GitHub and WANDB (Weights and Biases) experiment page. Comparing the results of GloVe and Capsule Neural Network, similar performances are reached.**

***Index Terms*—Deep Learning, Neural Network, Capsule Layer, Capsule Neural Network, MLP, GloVe, ICLR, NIPS, NLTK, EDA, WANDB, GitHub**

I. INTRODUCTION

The purpose of this document is to utilize a capsule layer architecture on a dataset about conference articles in order to predict the acceptance/rejection decision of these conference articles.

By using Capsule Neural Network, hierarchical and semantic relationships in the dataset can be captured in a much advanced manner and due to the dynamic routing mechanism of Capsule Neural Network, unseen review patterns can be handled much better [1], [2]. Capsule Neural Networks can provide better robustness to the variations in the dataset, can guarantee less information loss and requires fewer training samples [1], [2].

As a baseline comparison model, GloVe can be used [3]. In order to use GloVe pre-trained word vectors, they should be downloaded from the link [4] and unzipped.

The outline of this document is as follows: In Section II, dataset is explained with EDA (Exploratory Data Analysis) and quality checks. Section III presents the used GloVe architecture and capsule layer architecture separately. In Section IV, evaluation methods and performance metrics are mentioned. Section V gives the results of GloVe architecture and capsule layer architecture. Finally, the paper is concluded by Section VI.

II. DATASET

The notebook file named ”EDA_and_Quality_Checks.ipynb” was written in order to make EDA and Quality Checks and saved under the folder named ”Notebooks” in [5]. Moreover, by using this notebook file various histograms were plotted and saved in the formats ”eps”, ”jpg” and ”svg” under the directory named ”figures/histograms” in [5]. In this document, we can mention some of these analyses.

The dataset is about ICLR (International Conference on Learning Representations) papers from 2017-2020 and NIPS (Neural Information Processing Systems) papers from 2016-2019 and information related to these papers are located in their corresponding files in ”json” format [6]. All of these files in ”json” format were loaded in order to form pandas data frame variables ”ICLR_data” (only ICLR data) having 5194 rows, ”NIPS_data” (only NIPS data) having 3934 rows and ”data” (ICLR and NIPS data together) having 9128 rows. The column names of the variable ”data”, the format of each column, unique values related to each column can be summarized in Table I and II. It must be emphasized that there exists NaN and None values used for missing data row-column values. Regarding decision column, the histogram can be seen in Figures 1.

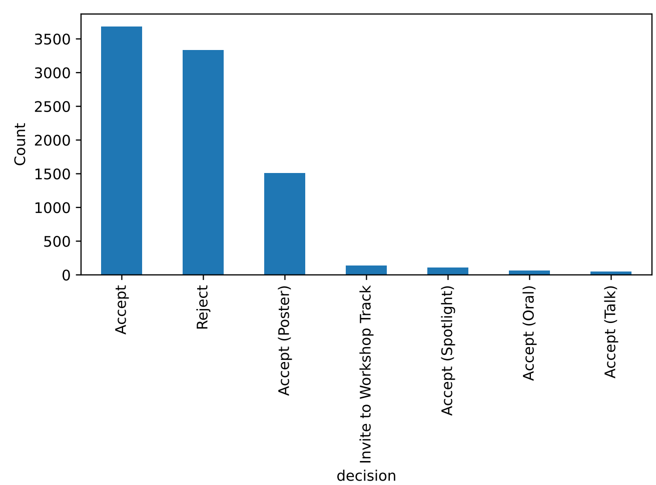

Fig. 1. Decision Histogram

The texts to be used as inputs in this dataset are titles combined with meta reviews. By using Natural Language Toolkit (NLTK), these input texts are tokenized and English stop words, punctuation marks, None, nan, NA and N\A values are removed from these tokens. The texts to be used as outputs in this dataset are decisions of seven different types which are Accept, Accept (Oral), Accept (Poster), Accept (Spotlight), Accept (Talk), Invite to Workshop Track, Reject.

TABLE I
DATA SUMMARY

Column Name	Format	Unique Value
name	string Some Examples ICLR_2017_1.pdf, NIPS_2016_1.pdf	8856 unique values
id	string Some Examples ICLR_2017_1, NIPS_2016_1	each of them
metadata.source	string	CRF, META, NaN
metadata.title	string	5459 unique values (None and NaN for missing 3652 time occurrences, RECURRENT NEURAL NETWORKS 5 time occurrences, DEEP REINFORCEMENT LEARNING 3 times occurrences, some other titles occur more than once, most of the titles occur only once)
metadata.authors	list An Example (Jonathon Cai, Richard Shin, Dawn Song)	not applicable
metadata.emails	list An Example (jonathon@ cs.berkeley.edu, ricshin@ cs.berkeley.edu, dawnsong@ cs.berkeley.edu)	not applicable
metadata.sections	list of dictionaries	not applicable
metadata.references	list of dictionaries	not applicable
metadata. referenceMentions	list of dictionaries	not applicable
metadata.year	int	0, 1969, 2016, 2017, 2018, 2019, 2020, NaN
metadata.abstractText	string	8844 unique values (None and NaN for missing 281 time occurrences, some abstracts occur twice due to either uploading to the same conference twice or to the two distinct conferences, most of the abstracts occur only once)
metadata.creator	string, An Example Microsoft® Word 2016	59 unique values
conference	string	ICLR, NIPS, NaN
decision	string	Accept Accept (Oral) Accept (Poster) Accept (Spotlight) Accept (Talk) Invite to Workshop Track Reject NaN

TABLE II
DATA SUMMARY (CONTINUATION)

Column Name	Format	Unique Value
url	string	8878 unique values
hasContent	boolean	true, false, NaN
hasReview	boolean	true, false, NaN
title	string	8826 unique values (None, NaN, NA and N\A for missing 279 time occurrences, some titles occur twice due to either uploading to the same conference twice or to the two distinct conferences, most of the titles occur only once)
authors	list An Example (Jonathon Cai, Richard Shin, Dawn Song)	not applicable
reviews	list of dictionaries	not applicable
metaReview	string	5876 unique values

III. MODELING

The classifier trained with GloVe embeddings is a Multi-layer Perceptron (MLP) classifier. MLP classifier has a hidden layer of size 128.

Capsule Layer Architecture has the following parameters

- num capsules: Number of output capsules in the layer = 2
- dim capsule: Number of dimensions in each output capsule's vector = 16
- input dim: Number of dimensions in each input capsule's vector = 3
- input len: Number of input capsules = 4
- routing iters: Number of routing iterations = 3

Capsule Layer is preceded with an embedding layer and two convolutional layers and is followed by a linear output layer. The embedding layer is of type vocabulary size x embedding dimension (20002 x 32), the first convolutional layer is of type embedding dimension x 64 (32 x 64), the second convolutional layer serving as primary capsules is of type 64 x (input dim x input len) (64 x 12), the capsule layer serving as dense capsules is of type 1 x 24 x 2 x 16 x 3 and the linear output layer is of type 512 x output length (512 x 7).

Both of GloVe embeddings and Capsule Layer Architecture are implemented in Python.

IV. EVALUATION

For GloVe embeddings, training and testing ratios are 0.8 and 0.2, respectively. Training is done for maximum 50 epochs with early stopping patience being equal to 3. In every epoch, Classifier Checkpoints, Accuracy, F1-Score, Area Under the ROC Curve and Confusion Matrix are calculated and stored to evaluate the performance of the MLP classifier.

For the capsule neural network architecture, training and testing ratios are 0.8 and 0.2, respectively. Training is done for 50 epochs with learning rate = 1e-3 and batch size = 128. In every epoch, Model Checkpoints, Cross Entropy Loss, Accuracy, F1-Score, Area Under the ROC Curve and Confusion Matrix are calculated and stored to evaluate the performance of the capsule neural network architecture.

V. Results

The Github link of the repository is given by [5]. The link of the WANDB experiment page is given by [7].

For GloVe embeddings, after the training is done; the final values for Accuracy, F1-Score, and Area Under the ROC Curve are equal to 0.62218, 0.58854, and 0.73483, respectively, and the Confusion Matrix is given by Figure 2. Much detailed results can be found in [5] and [7]. Importantly, the related .ipynb file can be seen in [8].

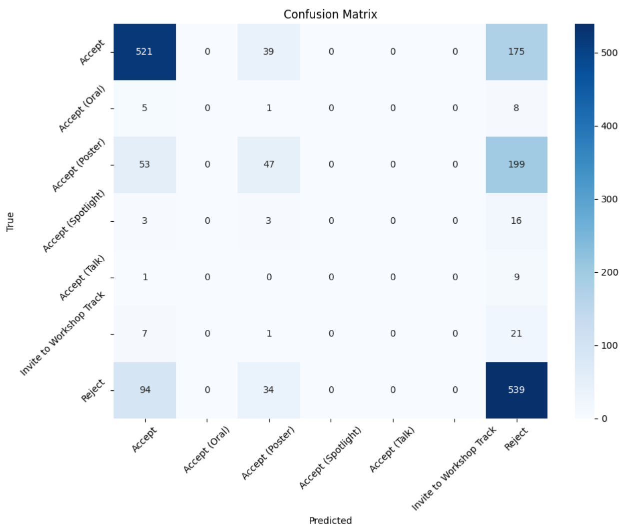

Fig. 2. GloVe Confusion Matrix

For the capsule neural network architecture, after the training is done; the final values for Cross Entropy Loss, Accuracy, F1-Score, and Area Under the ROC Curve are equal to 0.37507, 0.54561, 0.53338, and 0.71804, respectively, and the Confusion Matrix is given by Figure 3. Much detailed results can be found in [5] and [7]. Importantly, the related .ipynb file can be seen in [9].

VI. Discussion & Conclusion

Comparing the results of GloVe and Capsule Neural Network in terms of Accuracy, F1-Score, Area Under the ROC Curve and Confusion Matrix, it could be said that similar performances are reached.

The purpose of this document was to utilize a capsule layer architecture on a dataset about conference articles in order to predict the acceptance/rejection decision of these conference articles and dataset, modeling, evaluation and results were mentioned.

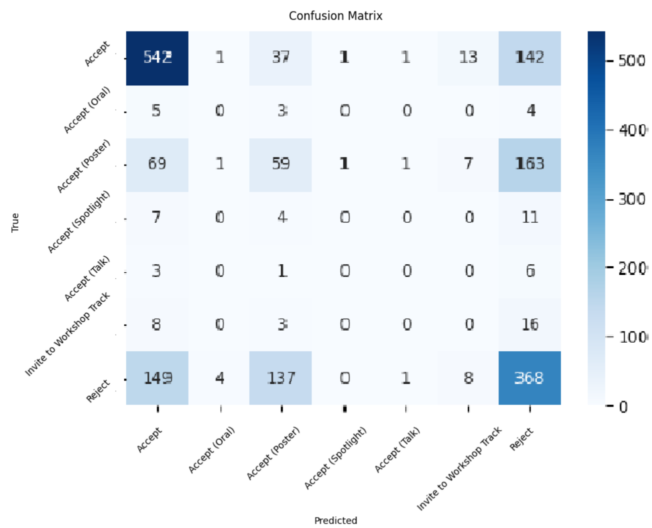

Fig. 3. Capsule Neural Network Confusion Matrix

References

[1] Wikipedia contributors, Capsule neural network — Wikipedia, The Free Encyclopedia. 2024. [Online]. Available: https://en.wikipedia.org/wiki/Capsule_neural_network

[2] S. Sabour, N. Frosst, and G. E. Hinton, Dynamic Routing Between Capsules. 2017. [Online]. Available: https://arxiv.org/abs/1710.09829

[3] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," in Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532–1543. [Online]. Available: http://www.aclweb.org/anthology/D14-1162

[4] J. Pennington, R. Socher, and C. D. Manning, "GloVe pretrained word vectors," 2014. [Online]. Available: http://nlp.stanford.edu/data/glove.6B.zip

[5] Ö. T. Kartal, IS584CourseTermProject. Online. Available: https://github.com/oztuka/IS584CourseTermProject/

[6] W. Yuan, P. Liu and G. Neubig, "ASAP-Review Dataset," 2023. Available: https://drive.google.com/file/d/1nJdljy468roUcKLbVwWUhMs7teirah75/view (Original: https://www.kaggle.com/datasets/jonauskis/asap-review)

[7] Ö. T. Kartal, IS 584 Course Term Project. Online. Available: https://wandb.ai/eabu2ss-metu-middle-east-technical-university/IS%20584%20Course%20Term%20Project

[8] Ö. T. Kartal, IS584CourseTermProject. Online. Available: https://github.com/oztuka/IS584CourseTermProject/blob/main/notebooks/Baselines.ipynb

[9] Ö. T. Kartal, IS584CourseTermProject. Online. Available: https://github.com/oztuka/IS584CourseTermProject/blob/main/notebooks/Preliminary_Results.ipynb