



MAJOR REVIEW

Capsule layer and attention layer augmentation for review categorization

**SUBMITTED BY:- SHARAT SACHIN (03111502716)
ABHA TRIPATHI (41211502716)
NAVYA MAHAJAN (35411502716)
SHIVANI AGGARWAL (35611502716)**



CONTENTS

- ❑ Introduction
- ❑ Objectives
- ❑ Definitions
 - ❑ Long Short Term Memory
 - ❑ Gated Recurrent Unit
 - ❑ Convolutional Neural Network
 - ❑ Attention Mechanism
 - ❑ Capsule Networks
- ❑ Literature Review
- ❑ Implementation & Results
- ❑ Future Scope/Conclusion
- ❑ References



INTRODUCTION

Sentiment analysis is an important branch of NLP.

It is a machine learning problem made demanding due to the varying nature of sentences, different lengths of the paragraphs of text, contextual understanding, sentiment ambiguity and the use of sarcasm and comparatives.

In this paper, the main focus is on comparing the results of using baseline unidirectional and bidirectional LSTM and GRU to ones augmented using a convolutional layer and self-attention layer.

This can also influence decisions made concerning customers. It helps companies to analyse and improve their brand image.



OBJECTIVES

- This paper is a continuation of the previous paper we performed in which we study simple RNN unidirectional and bidirectional models such as lstm, gru, bigru and bilstm.
- In the recent paper, we added layers to the previous used models such as attention layer, cnn layer, CNN + Attention layer, in order to know if there is a significant change in the results.
- Also in this paper, we performed 2-class categorization as well as we added Capsule layer, CNN layer, CNN+Attention layer and Attention layer.



CONVOLUTIONAL NEURAL NETWORK

- Deep, feed-forward artificial network, or ANN which is inculcated majorly to refine the processing to recognise image to voice recognition [10], or in classification, segmentation.
- It is a type of ANN that usually comprises a set of layers namely, convolutional, fully connected, and pooling layers.



LONG TERM SHORT MEMORY

- developed by Hochreiter and Schmidhuber
- consists of a memory cell
- computes the input, output and forget gate
- help in propagating an important feature that came early in the input
- **Bidirectional LSTM** : recognize long-term dependencies and contextual features from previous and future states
- maintain and use details from both backwards in time and forwards in time.



GATED RECURRENT UNIT

- gating units control movement of data within the unit
 - without making use of any separate memory cells
 - calculates two important gates called update and reset gates
-
- **Bidirectional GRU** : allows for the use of details from both past and future to make predictions about the current state
 - combine hidden state and cell state into one, resulting in faster training



ATTENTION MECHANISM

Self-attention is the process in which we apply attention mechanisms in each concerned point of the main input sequence. In this type of attention, for each sequence position, we create 3 vectors namely, query, key, value .



CAPSULE NETWORK

- developed by Geoffrey Hinton
- It is a type of network that contains numerous capsules which contain groups of neurons which are used to learn or detect an object.
- Also, capsNets have a unique feature in which lower-level capsules interact with higher-level capsules by sending them an output.
- Capsules are very promising in applications such as emotion detection, hand-written and text recognition, MT and many more.



LITERATURE REVIEW

Attention Mechanism Based:-

Research	Model Used	Dataset	Result
1. Feature Fusion Text Classification Model Combining CNN and BiGRU with Multi-Attention Mechanism	MATT(CNN) along with the combination of Bi-directional GRU (Bi-GRU)	SemEval2016	79.22%
2. Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. Neurocomputing	They used MF-BiLSTM and SA-BiLSTM	MR, SST-5, SST-2 datasets	89%



LITERATURE REVIEW

Capsule Based:-

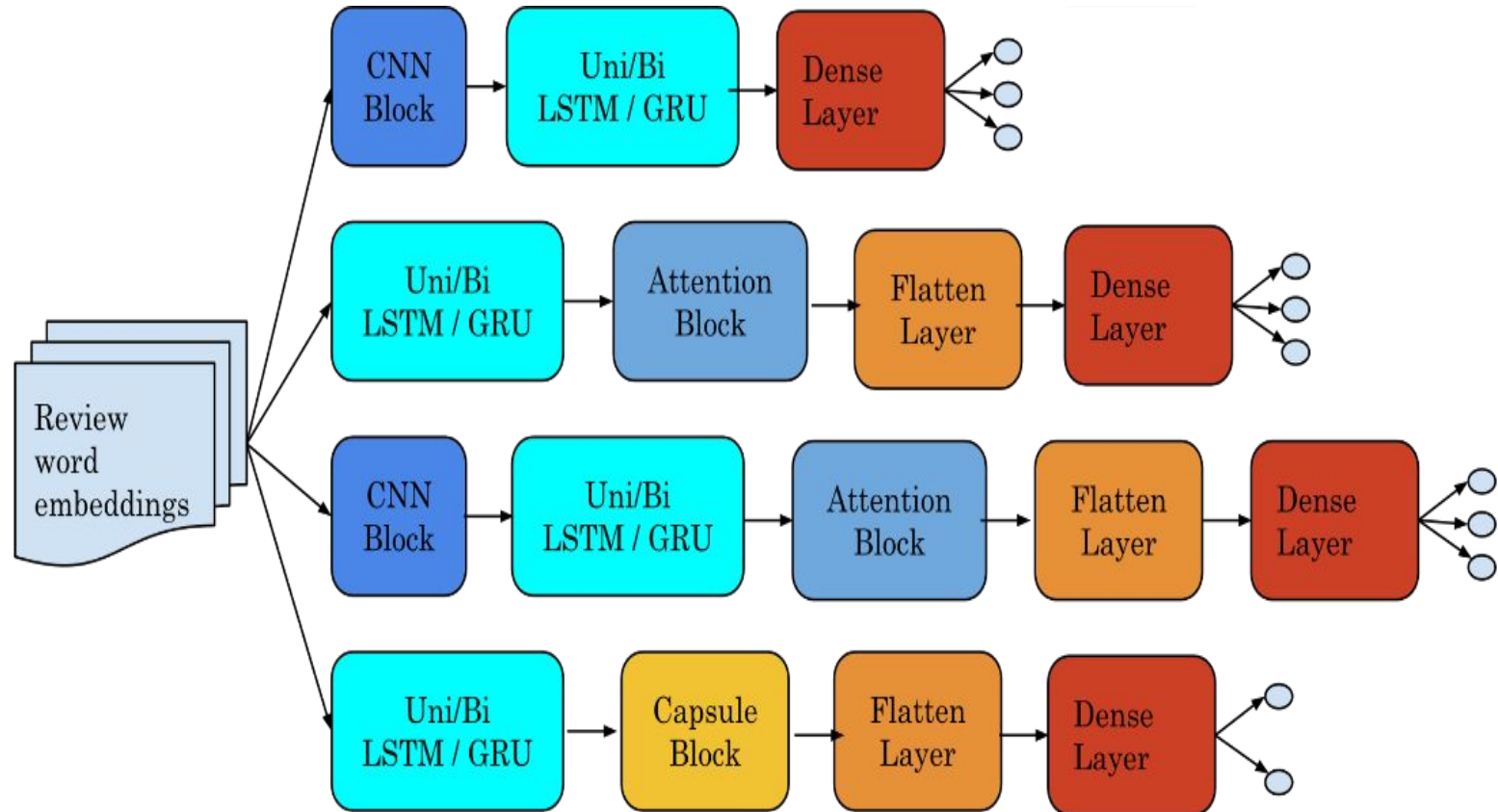
Research	Model Used	Dataset	Result
1. A Sentiment Classification Model Based on Bi-Channel Capsule Network	3 parts of modules, G-word vector and Bi-channel capsule representation	2 datasets of IMDB and NLPCC	92.38%
2. A novel capsule-based hybrid neural network for sentiment classification	CapsNet with addition to BiGRU	two text datasets	82.5%



DATASET USED

- Implemented the baseline models for LSTM, GRU and Bi-LSTM and Bi-GRU along with attention, CNN and capsule layer on an Amazon review dataset [5].
- For the 3 class variant->
 - ❑ classified reviews based on the Electronics section into 3 classes: positive, negative and neutral
 - ❑ used a balanced dataset with 40,000 reviews in each class, that is, with 120,000 reviews total
 - ❑ 5-core reviews only used - i.e. reviews written by users with at least 5 reviews
- For the 2-class variant,
 - ❑ we have only two classes- positive and negative.
 - ❑ We choose 100 thousand reviews from each class and split this into 2 parts - 150 thousand for training and 50 thousand for testing.

IMPLEMENTATION





IMPLEMENTATION

on input reviews,

- we perform tokenization and padding, followed by word embeddings on the data column
- one-hot encoding on the target variables.

description of layers :

Capsule block - number of capsules - 10, dimension of capsules - 10, routings - 5

CNN layer - filters - 128, kernel_size = 5,

Attention layer - sigmoid activation

[Link to code](#)



RESULTS

Accuracies for different deep learning baseline methods on the Amazon dataset for sentiment based classification

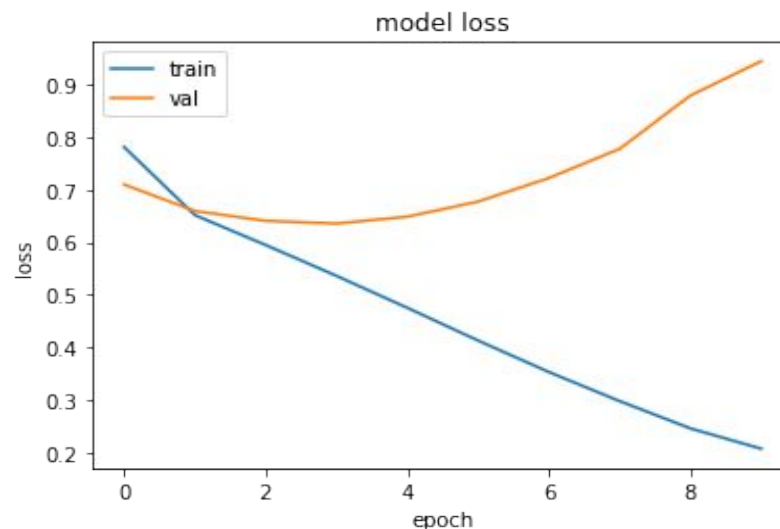
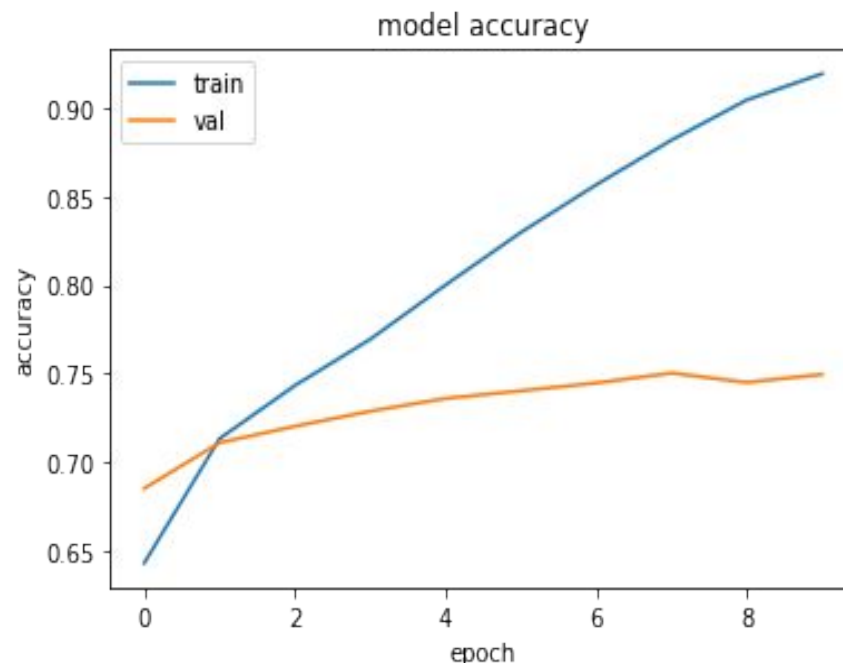
3 Class			2 Class	
		Acc.		Acc.
CNN	+LSTM	73.58	+LSTM	93.47
Attention	+BiLSTM	74.96	+BiLSTM	93.56
CNN+Attention	+LSTM	72.91	+LSTM	92.78
Capsule	-	-	+BiLSTM	94.87

Department of Computer Science, BVCOE, New Delhi
Subject: MAJOR REVIEW ,Mentor:: Ms. Preeti Nagrath

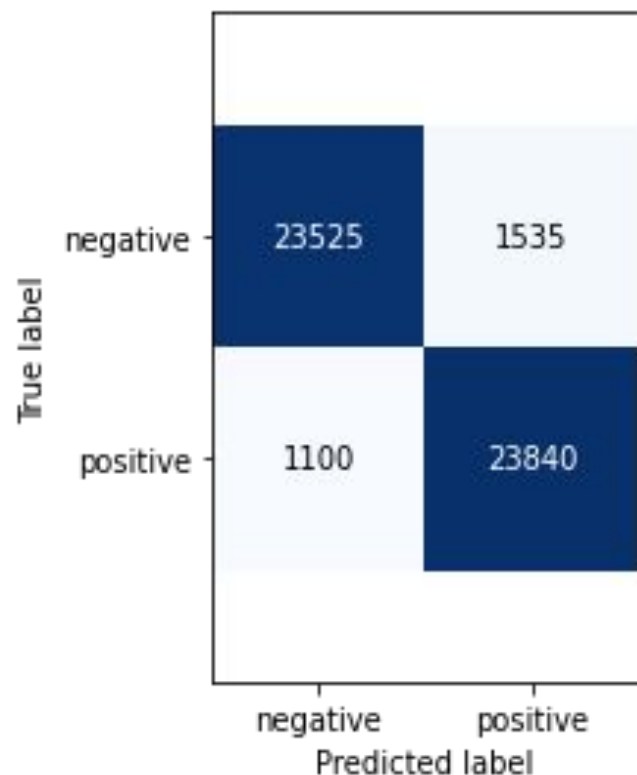


True label	Predicted label		
	negative	neutral	positive
	negative	7653	1863
neutral	2200	6313	1522
positive	428	1048	8522

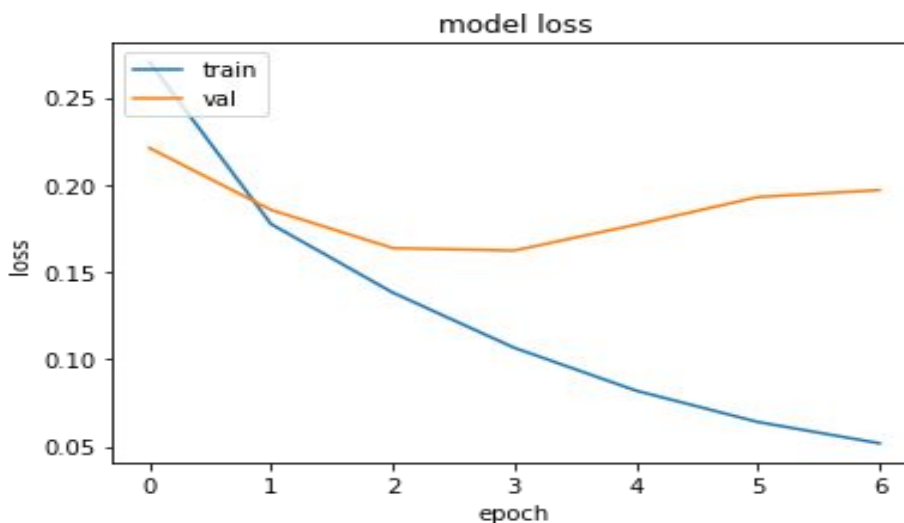
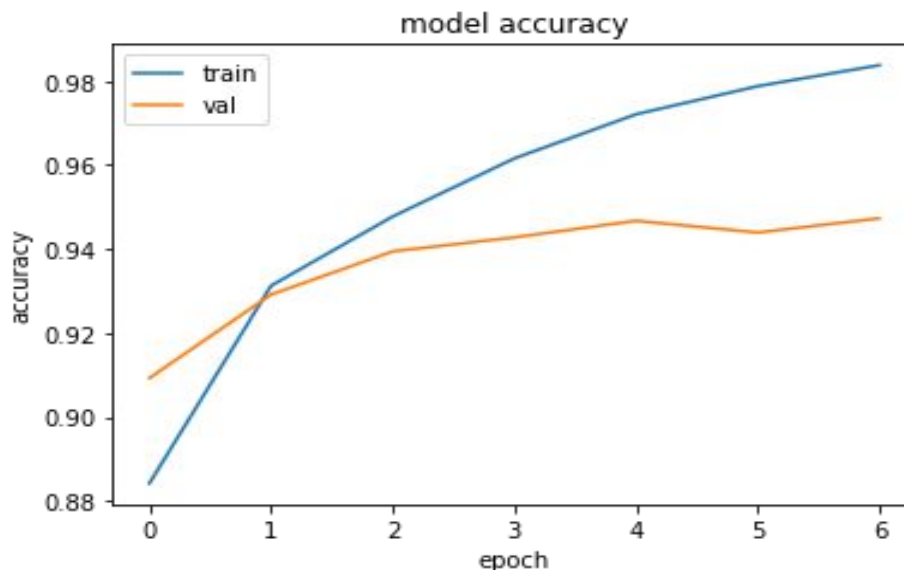
Accuracy and loss graphs,
confusion matrix of Attention +
BiLSTM(3-class)



Department of Computer Science, BVCOE, New Delhi
Subject: MAJOR REVIEW ,Mentor:: Ms. Preeti Nagrath



Accuracy and loss graphs,
confusion matrix of
Capsule + BiLSTM(2 class)





FUTURE SCOPE & CONCLUSION

- create document representation over various complex discourse tree structures
- concentrate on the perception of the special contexts in reviews
- put forth tree-based methods to model the intrinsic logic of sentiment contained in these reviews
- see if such models can obtain tricky implicit knowledge of human interaction such as humor or sarcasm
- assess the performance of these models on data from other domains
- it is clear that capsule network and convolutional and attention layers have their inherent advantages when used in comparison with base models of unidirectional and bidirectional LSTM and GRU



REFERENCES

1. Zhang, J., Liu, F. A., Xu, W., & Yu, H. (2019). Feature Fusion Text Classification Model Combining CNN and BiGRU with Multi-Attention Mechanism. *Future Internet*, 11(11), 237.
2. Li, W., Qi, F., Tang, M., & Yu, Z. (2020). Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*.
3. Zhang, K., Jiao, M., Chen, X., Wang, Z., Liu, B., & Liu, L. (2019). SC-BiCapsNet: A Sentiment Classification Model Based on Bi-Channel Capsule Network. *IEEE Access*, 7, 171801-171813.
4. Du, Y., Zhao, X., He, M., & Guo, W. (2019). A novel capsule-based hybrid neural network for sentiment classification. *IEEE Access*, 7, 39321-39328.
5. Ni, J., Li, J., & McAuley, J. (2019, November). Justifying Recommendations using Distantly-Labeled Reviews and Fine-Grained Aspects. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 188-197).