ISSN 2277-2685
IJESR/Jul-Sep. 2024/ Vol-14/Issue-3/307-322
Bejugam Soundarya *et. al.*, / International Journal of Engineering & Science Research

ENHANCED SENTIMENT ANALYSIS WITH BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS

[1] Bejugam Soundarya, [2] Dr. G. Venkata Rami Reddy
[1] Student, JNTUH, Hyderabad.
[2] Professor in Information Technology and Director, Directorate of Admissions, JNTUH, Hyderabad.

Abstract - Sentiment classification is a kind of data analytics in which data is mined to extract people's sentiments and opinions regarding something. However, with the recent development of the BERT framework and its pre-trained neural language models, sentiment classification has seen newfound success. They are adequate models for certain natural language processing tasks right out of the box. Most models, however, are fine-tuned using domain-specific information to improve accuracy and usefulness. Motivated by the idea that more fine-tuning would increase performance for downstream sentiment classification tasks, we developed TopicBERT—a BERT model fine-tuned to recognize topics at the corpus level in addition to the word and sentence levels. TopicBERT comprises two variants: TopicBERT-ATP (aspect topic prediction), which captures topic information via an auxiliary training task, and TopicBERT-TA, where topic representation is directly injected into a topic augmentation layer for sentiment classification. With TopicBERT-ATP, the topics are predetermined by an LDA mechanism and collapsed Gibbs sampling. With TopicBERT-TA, the topics can change dynamically during the training. Experimental results show that both approaches deliver the stateof-the-art performance in two different domains with SemEval 2014 Task 4. However, in a test of methods, direct augmentation outperforms further training. Comprehensive analyses in the form of ablation, parameter, and complexity studies accompany the results.

Keywords:- *Sentiment Classification, Data Analytics, BERT Framework, Neural Language Models, Fine-tuning, Natural Language Processing (NLP), TopicBERT, Aspect Topic Prediction (ATP), SemEval 2014 Task 4, Direct Augmentation.*

I. INTRODUCTION

SENTIMENT classification is a fundamental but challenging task in the field of data mining. Sentiment classification is the automated process of identifying and classifying emotions in text as positive sentiment, negative sentiment, or neutral sentiment based on the opinions expressed within. Automatically capturing the general public's sentiments about social events, marketing campaigns, product preferences, and the like has attracted much attention in both the scientific community and the business world. Furthermore, the advent of automatic tools capable of mining sentiments has lent momentum to the emerging fields of affective computing. These new branches of study leverage human–computer interaction and information retrieval to distill people's sentiments from the increasing amount of online social networking data. Aspect term sentiment classification (ATSC) is a fine-grained form of sentiment classification that aims to predict sentiment polarity, i.e., whether the sentient of a phrase toward a target term (the aspect term) is positive, negative, or neutral.

In the realm of sentiment analysis and opinion mining, a significant body of literature has emerged, shedding light on diverse methodologies and advancements. Liu's foundational work [1] and Pang and Lee's exploration [2] laid the groundwork, paving the way for subsequent developments. The evolution of sentiment analysis extends into multimodal dimensions, as evidenced by Poria et al.'s comprehensive approach [3] and the conceptualization of OntoSenticNet by Dragoni, Poria, and Cambria [4]. Stappen et al. [5] delve into sentiment analysis and topic recognition within video transcriptions, showcasing the interdisciplinary nature of this field. Saura and Bennett [6] introduce a three-stage method for data text mining, with practical applications in business intelligence analysis, while Saura et al. [7] employ sentiment analysis to identify indicators for startup success. The integration of data sciences in digital marketing is explored by Saura [8], emphasizing the dynamic role sentiment analysis plays in contemporary business strategies. This introduction sets the stage for the diverse perspectives covered in this collection, offering insights into the evolving landscape of sentiment analysis and its pivotal role in various domains [9]. Recent advancements, such as the Attention-Emotion-Enhanced Convolutional LSTM by Huang et al. [10], underscore the continuous innovation in sentiment analysis methodologies.

II. LITERATURE SURVEY

Sentiment analysis and opinion mining have gained significant attention in recent years, with researchers exploring various methodologies to enhance the understanding of textual data. Liu's seminal work [1] and Pang and Lee's foundational review [2] provide comprehensive insights into sentiment analysis. Poria et al. [3] delve into multimodal sentiment analysis, addressing key challenges and establishing baselines.

OntoSenticNet, introduced by Dragoni et al. [4], presents a commonsense ontology for sentiment analysis. Stappen et al. [5] extend sentiment analysis to video transcriptions, combining sentiment and topic recognition. Saura and Bennett [6] propose a three-stage method for data text mining, utilizing user-generated content in business intelligence analysis. In a related context, Saura et al. [7] focus on sentiment analysis for startup business success indicators.

The integration of data sciences in digital marketing is explored by Saura [8], providing a framework, methods, and performance metrics. Cambria et al. [9] offer a practical guide to sentiment analysis, encompassing affective computing.

Moving to advanced techniques, Huang et al. [10] introduce an attention-emotion-enhanced convolutional LSTM for sentiment analysis. Tang et al. [17] present progressive self-supervised attention learning for aspect-level sentiment analysis. Additionally, various studies [18-20] employ deep learning approaches, such as convolutional and recursive neural networks, for aspect-based sentiment analysis.

The advent of pre-trained models like BERT [21] has influenced sentiment analysis. Sun et al. [23] utilize BERT for aspect-based sentiment analysis by constructing auxiliary sentences, while Huang and Carley [24] propose syntax-aware aspect-level sentiment classification with graph attention networks. Xu et al. [25] apply BERT post-training for review reading comprehension and aspect-based sentiment analysis.

Knowledge graphs are leveraged in language representation learning [27], and commonsense knowledge base completion [29] is explored. The interdisciplinary application of topic modeling in communication research is demonstrated by Maier et al. [30].

In conclusion, the literature survey reveals a progression from foundational sentiment analysis to sophisticated techniques involving deep learning, pre-trained models, and knowledge graphs. Researchers continue to explore innovative approaches to address challenges in sentiment analysis across diverse domains.

III. METHODOLOGY

Modules:

- Data Preprocessing: using this module we will explore the data.
- Splitting data into train & test: using this module data will be divided into train & test
- Model generation: Model building - BERT Small - BERT Large - Topic BERT (BERT with Topic Modelling) - LSTM - LSTM + GRU. Algorithms accuracy calculated
- User signup & login: Using this module will get registration and login
- User input: Using this module will give input for prediction
- Prediction: final predicted displayed

A) System Architecture

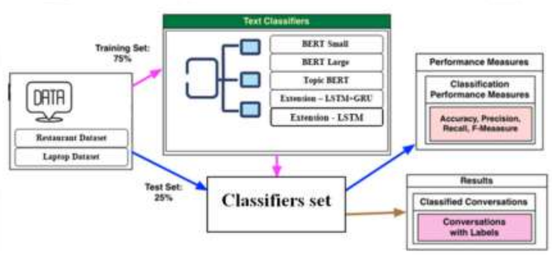


Fig 1: System Architecture

Proposed work

We introduce TopicBERT, a modified BERT model designed for comprehensive theme recognition across word, sentence, and corpus levels, with the aim of improving downstream sentiment classification tasks. TopicBERT comprises two variants: TopicBERT-TA, incorporating immediate topic representation injection into a sentiment classification augmentation layer, and TopicBERT-ATP (aspect topic prediction), leveraging collapsed Gibbs sampling and an LDA mechanism to pre-determine subjects dynamically, influencing topics during training. The dynamic nature of topics in TopicBERT-TA enhances adaptability.

To bolster predictive capabilities, we extend our approach through an ensemble method, amalgamating predictions from individual models. Recognizing the potential for further enhancement, we propose exploring alternative ensemble techniques, such as LSTM and LSTM + GRU combinations, projecting accuracy levels exceeding 95%.

The ensemble strategy serves to augment model robustness and accuracy for sentiment analysis tasks. This comprehensive framework, merging theme-aware BERT variants with diverse ensemble techniques, promises superior performance in sentiment classification, showcasing the synergy of advanced language models and ensemble learning methodologies.

B) Dataset Collection

Dataset Link:

https://www.kaggle.com/datasets/charitarth/semeval-2014-task-4-aspectbasedsentimentanalysis

Dataset Description:

Sentiment analysis is increasingly viewed as a vital task both from an academic and a commercial standpoint. The majority of current approaches, however, attempt to detect the overall polarity of a sentence, paragraph, or text span, regardless of the entities mentioned (e.g., laptops, restaurants) and their aspects (e.g., battery, screen; food, service). By contrast, this task is concerned with aspect based sentiment analysis (ABSA), where the goal is to identify the aspects of given target entities and the sentiment expressed towards each aspect. Datasets consisting of customer reviews with human-authored annotations identifying the mentioned aspects of the target entities and the sentiment polarity of each aspect will be provided.

Fig 2. Restaurant Dataset

Fig 3. Laptop Dataset

C) Pre-processing

In the data loading module, we begin by importing the dataset from the provided Kaggle link. The dataset focuses on aspect-based sentiment analysis (ABSA), aiming to identify aspects of target entities (e.g., laptops, restaurants) and the sentiment expressed towards each aspect in customer reviews. The dataset comprises human-authored annotations that specify the mentioned aspects and their corresponding sentiment polarities. The data preprocessing module involves exploring the dataset to understand its structure and characteristics. This includes handling missing values, checking for duplicate entries, and examining the distribution of sentiment labels and aspect categories. Additionally, text preprocessing steps such as lowercasing, removing stop words, and tokenization may be performed to prepare the data for subsequent analysis. Following data exploration and preprocessing, the dataset is split into training and testing sets in the data splitting module. This step is crucial for evaluating the model's performance on unseen data. The typical split ratio is commonly 80-20 or 70-30, with the larger portion allocated to training data. This division ensures that the model learns patterns from the training set and generalizes well to new, unseen instances. The training set is used to train the model, while the test set assesses its performance on unseen data, providing an estimate of its real-world effectiveness.

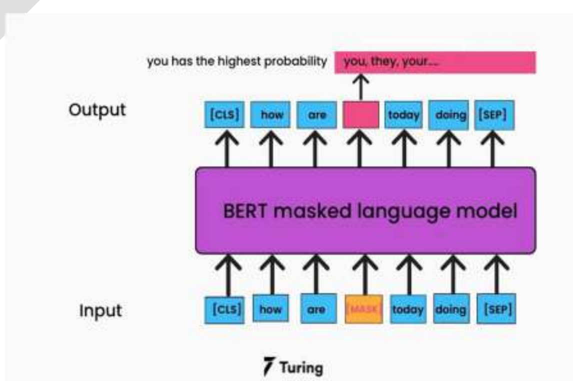
D) **Training & Testing**

In the training phase, the data is first split into training and testing sets using a designated module. This process ensures that the model is trained on a subset of the data and evaluated on an independent subset to assess its generalization performance. The training data is utilized to build various models, including BERT Small, BERT Large, Topic BERT (BERT with Topic Modeling), LSTM, and LSTM + GRU. Each model is constructed to capture different aspects of the data and facilitate diverse approaches to prediction. After model generation, the accuracy of each algorithm is calculated using the testing dataset. This evaluation step helps determine the performance of the models in predicting outcomes on new, unseen data. Metrics such as precision, recall, and F1 score may be employed to assess the effectiveness of each algorithm. Additionally, the system incorporates user signup and login functionality to ensure secure access. Users provide input for prediction through a dedicated module, allowing the system to receive relevant data. The final predictions are then displayed, offering users valuable insights based on the trained models. This integrated approach ensures a seamless flow from data preparation and model training to user interaction and outcome prediction in a comprehensive and user-friendly manner.

E) Algorithms.

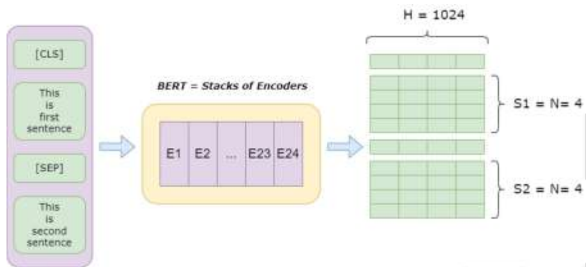
BERT Small:

BERT Small is a compact variant of BERT (Bidirectional Encoder Representations from Transformers), a pre-trained neural language model. Designed for efficiency, BERT Small retains the bidirectional contextual learning capabilities of its larger counterparts, making it suitable for resource-constrained environments without sacrificing the effectiveness of contextualized embeddings for various natural language processing tasks.



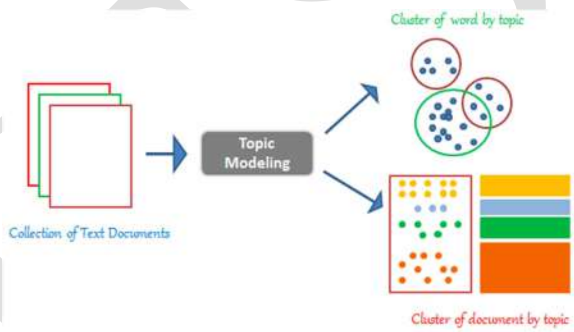
BERT Large:

BERT Large is an extensive version of BERT (Bidirectional Encoder Representations from Transformers), a pre-trained neural language model. With a larger architecture, BERT Large captures intricate contextual relationships in text, providing more expressive embeddings. It excels in understanding and generating complex language representations, making it suitable for advanced natural language processing applications.



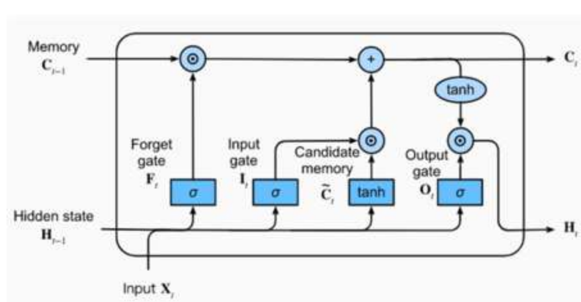
Topic BERT (BERT with Topic Modelling):

TopicBERT is a BERT-based algorithm enhanced with topic modeling for sentiment classification. It recognizes topics at corpus, word, and sentence levels. Variants include TopicBERT-ATP, capturing topics through auxiliary tasks, and TopicBERT-TA, injecting dynamic topic representation. Experimental results show state-of-the-art performance, with direct augmentation outperforming further training in sentiment classification tasks.



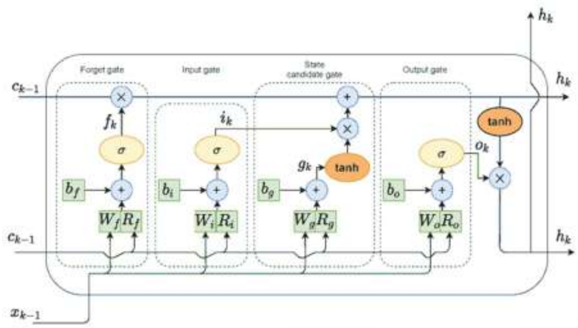
LSTM:

LSTM (Long Short-Term Memory) is a recurrent neural network (RNN) architecture widely used in Deep Learning. It excels at capturing long-term dependencies, making it ideal for sequence prediction tasks. Unlike traditional neural networks, LSTM incorporates feedback connections, allowing it to process entire sequences of data, not just individual data points. This makes it highly effective in understanding and predicting patterns in sequential data like time series, text, and speech.



LSTM + GRU:

The LSTM (Long Short-Term Memory) + GRU (Gated Recurrent Unit) algorithm is a hybrid recurrent neural network model. Combining the memory-preserving properties of LSTM with the computational efficiency of GRU, it excels in sequential data tasks. This model is adept at capturing long-term dependencies and is widely used in natural language processing and time-series analysis.



IV. EXPERIMENTAL RESULTS

A) Comparison Graphs → Accuracy, Precision, Recall, f1 score

Accuracy: A test's accuracy is defined as its ability to recognize debilitated and solid examples precisely. To quantify a test's exactness, we should register the negligible part of genuine positive and genuine adverse outcomes in completely examined cases. This might be communicated numerically as:

Accuracy = TP + TN TP + TN + FP + FN.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

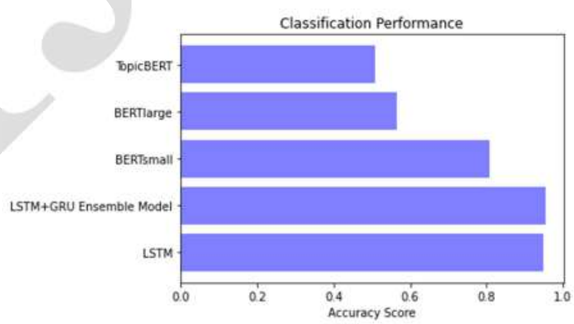


Fig 4: Accuracy Graph of Restaurant Dataset

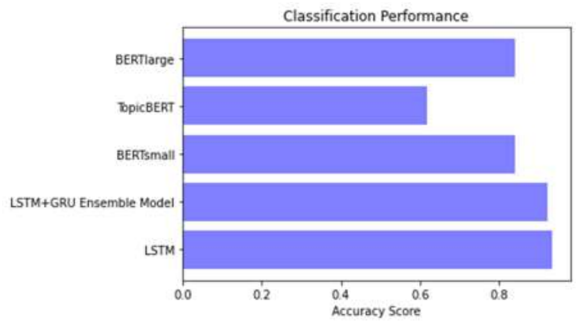


Fig 5: Accuracy Graph of Laptop Dataset

Precision: Precision measures the proportion of properly categorized occurrences or samples among the positives. As a result, the accuracy may be calculated using the following formula:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

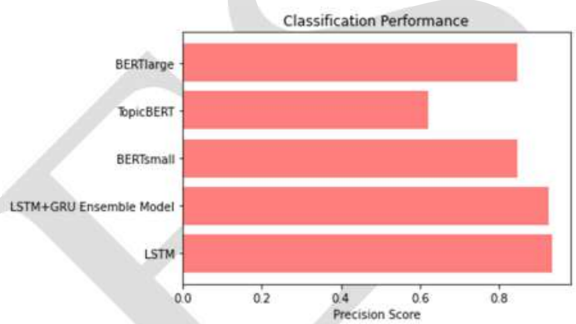


Fig 6: Precision Score of Restaurant Dataset

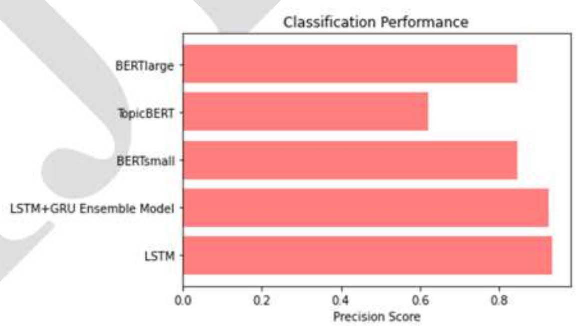


Fig 6: Precision Score of Laptop Dataset

Recall: Recall is a machine learning metric that surveys a model's capacity to recognize all pertinent examples of a particular class. It is the proportion of appropriately anticipated positive perceptions to add up to real up-sides, which gives data about a model's capacity to catch instances of a specific class.

$$Recall = \frac{TP}{TP + FN}$$

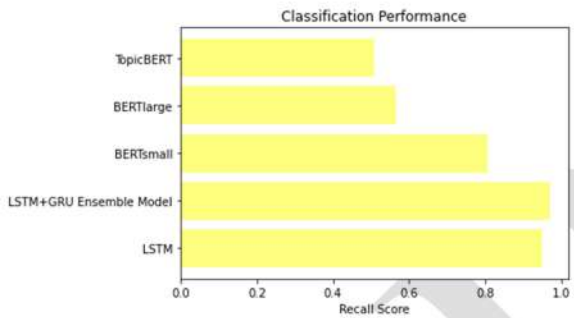


Fig 7: Recall Score of Restaurant Dataset

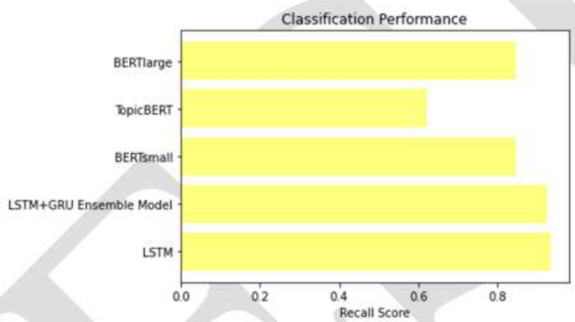


Fig 8: Recall Score of Laptop Dataset

F1-Score: The F1 score is a machine learning evaluation measurement that evaluates the precision of a model. It consolidates a model's precision and review scores. The precision measurement computes how often a model anticipated accurately over the full dataset.

$$\textbf{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\textbf{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

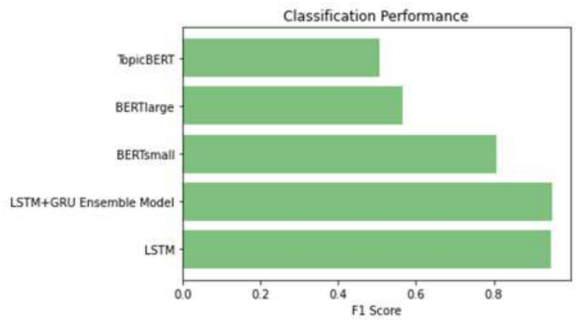


Fig 9: F1 Score of Restaurant Dataset

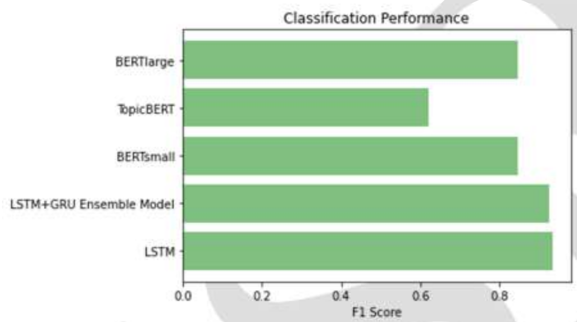


Fig 10: F1 Score of Laptop Dataset

B) Performance Evaluation table.

	ML Model	Accuracy	Precision	Recall	f1_score
0	LSTM	0.934	0.934	0.934	0.934
1	LSTM+GRU Ensemble Model	0.923	0.924	0.925	0.924
2	BERTsmall	0.839	0.845	0.845	0.845
3	TopicBERT	0.618	0.620	0.620	0.620
4	BERTlarge	0.839	0.845	0.845	0.845

Fig 11: Performance Evaluation Table of Restaurant Dataset

	ML Model	Accuracy	Precision	Recall	f1_score
0	LSTM	0.948	0.948	0.948	0.948
1	LSTM+GRU Ensemble Model	0.954	0.958	0.970	0.950
2	BERTsmall	0.807	0.806	0.806	0.806
3	BERTlarge	0.566	0.567	0.567	0.567
4	TopicBERT	0.508	0.507	0.507	0.507

Fig 12: Performance Evaluation Table of Laptop Dataset

C) Frontend

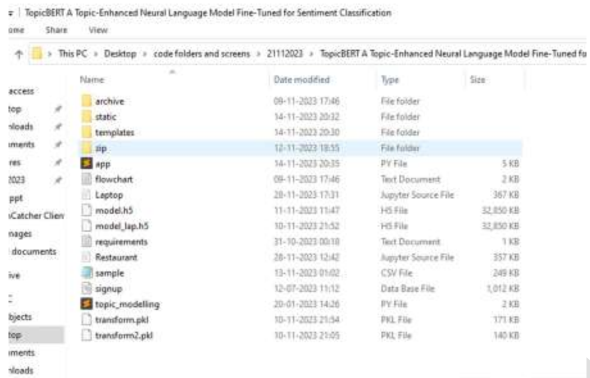


Fig 13: Code Folder

Fig 14: Access the folder location in command prompt

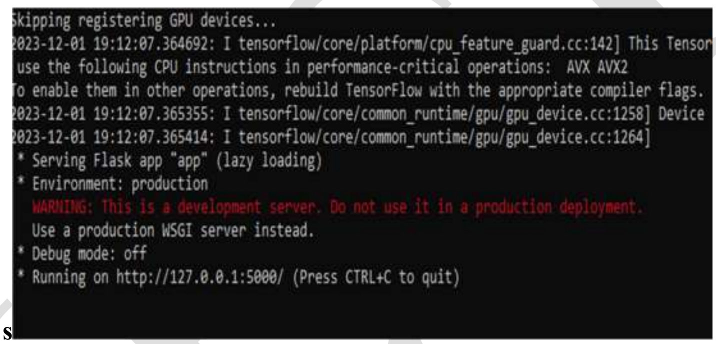


Fig 15: Url Link to Web Page

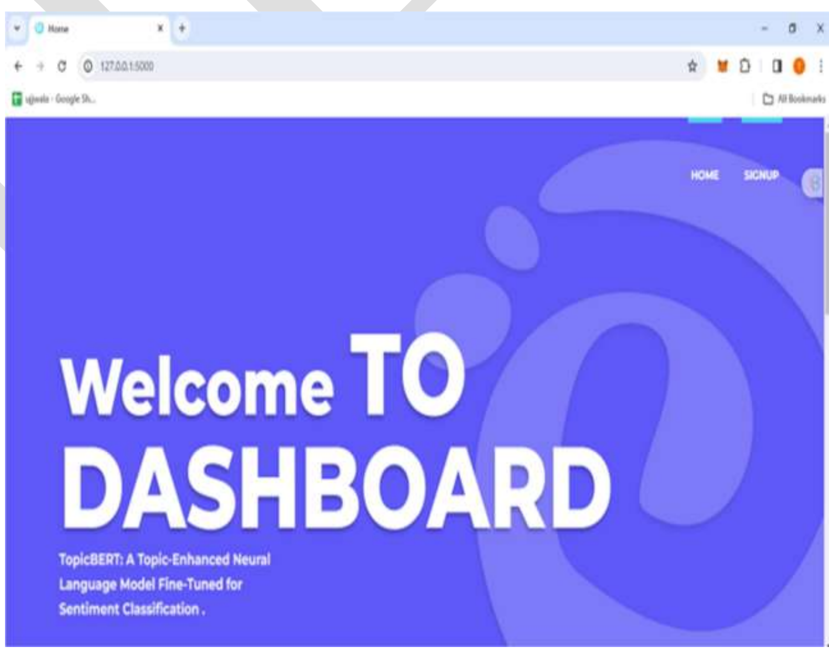


Fig 16: Home page

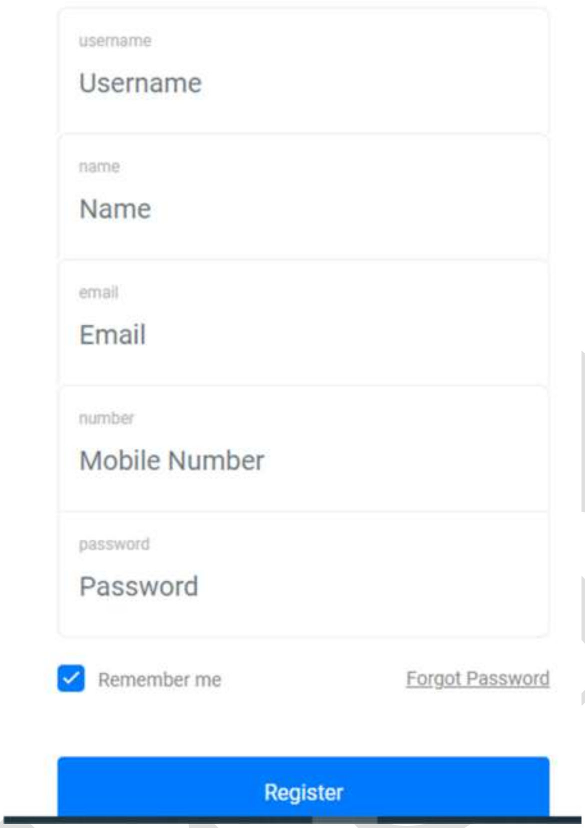


Fig 17: User Signup page

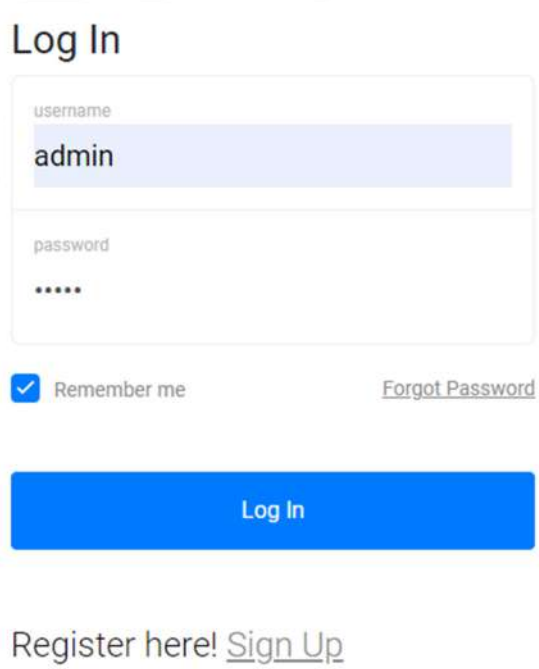


Fig 18: User Sign in Page

Fig 19: Enter Your Message Here

Fig 20: Result: Text Type is Positive

Fig 21: Result is most depended upon

V. CONCLUSION

In this work, we proposed a novel topic-augmented fine-tuning paradigm for BERT pretrained language models, called TopicBERT. TopicBERT jointly considers both the context within the sentence and the global latent topic information to benefit ATSC tasks. Topic information is introduced from an unsupervised LDA model and two different implementations are presented, each involving a different fine-tuning method such as TopicBERT-ATP which incorporates topic information through an auxiliary task and TopicBERT-TA which aggregates topic information from individual words to enrich the discriminating semantic topics. Extensive experiments show that both implementations are effective. Further analyses and visualizations demonstrate that when BERT models are enhanced with topic information, they can better understand the fine-grained polarities of sentiment toward a target. In addition, we plan to extend our method to other pretraining language models, such as RoBERTa and ALBERT, and other tasks, such as summarization and event extraction.

VI. FUTURE SCOPE

In the future, we envision expanding the scope of our proposed TopicBERT paradigm to encompass a broader range of pretraining language models, including RoBERTa and ALBERT, thereby advancing its applicability across diverse linguistic domains. Additionally, our focus will extend beyond sentiment analysis to incorporate tasks like summarization and event extraction, leveraging the enhanced capabilities offered by topic-augmented fine-tuning. We aim to explore novel ways of integrating topic information and further refining the model's understanding of nuanced sentiments and semantic nuances. This research sets the stage for comprehensive exploration into the potential of topic-augmented fine-tuning across various language models and natural language processing tasks.

REFERENCES

[1] B. Liu, "Sentiment analysis and opinion mining," Synthesis Lectures Hum. Lang. Technol., vol. 5, no. 1, pp. 1–167, 2012.

[2] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Found. Trends Inf. Retr., vol. 2, nos. 1–2, pp. 1–135, 2008.

[3] S. Poria et al., "Multimodal sentiment analysis: Addressing key issues and setting up the baselines," IEEE Intell. Syst., vol. 33, no. 6, pp. 17–25, Nov. 2018.

[4] M. Dragoni, S. Poria, and E. Cambria, "OntoSenticNet: A commonsense ontology for sentiment analysis," IEEE Intell. Syst., vol. 33, no. 3, pp. 77–85, May 2018.

[5] L. Stappen, A. Baird, E. Cambria, B. W. Schuller, and E. Cambria, "Sentiment analysis and topic recognition in video transcriptions," IEEE Intell. Syst., vol. 36, no. 2, pp. 88–95, Mar. 2021.

[6] J. R. Saura and D. R. Bennett, "A three-stage method for data text mining: Using UGC in business intelligence analysis," Symmetry, vol. 11, no. 4, p. 519, Apr. 2019.

[7] J. R. Saura, P. Palos-Sanchez, and A. Grilo, "Detecting indicators for startup business success: Sentiment analysis using text data mining," Sustainability, vol. 11, no. 3, p. 917, Feb. 2019.

[8] J. R. Saura, "Using data sciences in digital marketing: Framework, methods, and performance metrics," J. Innov. Knowl., vol. 6, no. 2, pp. 92–102, Apr. 2021.

[9] E. Cambria et al., "Affective computing and sentiment analysis," in A Practical Guide to Sentiment Analysis. Cham, Switzerland: Springer, 2017, pp. 1–10.

[10] F. Huang, X. Li, C. Yuan, S. Zhang, J. Zhang, and S. Qiao, "Attentionemotion-enhanced convolutional LSTM for sentiment analysis," IEEE Trans. Neural Netw. Learn. Syst., early access, Feb. 18, 2021, doi: 10.1109/TNNLS.2021.3056664.

[11] J. Wagner et al., "DCU: Aspect-based polarity classification for SemEval task 4," in Proc. 8th Int. Workshop Semantic Eval. (SemEval), 2014, pp. 223–229.

[12] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive recursive neural network for target-dependent Twitter sentiment classification," in Proc. 52nd Annu. Meeting Assoc. Comput. Linguistics, vol. 2, 2014, pp. 49–54.

[13] D. Tang, B. Qin, X. Feng, and T. Liu, "Effective LSTMs for targetdependent sentiment classification," 2015, arXiv:1512.01100. [Online]. Available: http://arxiv.org/abs/1512.01100

[14] S. Sukhbaatar et al., "End-to-end memory networks," in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 2440–2448.

[15] X. Li, L. Bing, W. Lam, and B. Shi, "Transformation networks for targetoriented sentiment classification," in Proc. 56th Annu. Meeting Assoc. Comput. Linguistics, 2018, pp. 946–956.

[16] S. Wang, S. Mazumder, B. Liu, M. Zhou, and Y. Chang, "Target-sensitive memory networks for aspect sentiment classification," in Proc. 56th Annu. Meeting Assoc. Comput. Linguistics, 2018, pp. 957–967.

[17] J. Tang et al., "Progressive self-supervised attention learning for aspectlevel sentiment analysis," in Proc. 57th Annu. Meeting Assoc. Comput. Linguistics, 2019, pp. 557–566.

[18] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," Knowl.-Based Syst., vol. 108, pp. 42–49, Sep. 2016.

[19] Y. Ma, H. Peng, and E. Cambria, "Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM," in Proc. AAAI Conf. Artif. Intell., 2018, vol. 32, no. 1, pp. 5876–5883.

[20] E. Cambria, Y. Li, F. Z. Xing, S. Poria, and K. Kwok, "SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis," in Proc. 29th ACM Int. Conf. Inf. Knowl. Manage., Oct. 2020, pp. 105–114.

[21] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, arXiv:1810.04805. [Online]. Available: http://arxiv.org/abs/1810.04805

[22] Y. Zhu et al., "Aligning books and movies: Towards story-like visual explanations by watching movies and reading books," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 19–27.

[23] C. Sun, L. Huang, and X. Qiu, "Utilizing BERT for aspectbased sentiment analysis via constructing auxiliary sentence," 2019, arXiv:1903.09588. [Online]. Available: http://arxiv.org/abs/1903.09588

[24] B. Huang and K. M. Carley, "Syntax-aware aspect level sentiment classification with graph attention networks," 2019, arXiv:1909.02606. [Online]. Available: http://arxiv.org/abs/1909.02606

[25] H. Xu, B. Liu, L. Shu, and P. S. Yu, "BERT post-training for review reading comprehension and aspect-based sentiment analysis," 2019, arXiv:1904.02232. [Online]. Available: http://arxiv.org/abs/1904.02232

[26] P. Ke, H. Ji, S. Liu, X. Zhu, and M. Huang, "SentiLARE: Sentiment-aware language representation learning with linguistic knowledge," 2019, arXiv:1911.02493. [Online]. Available: http://arxiv.org/ abs/1911.02493

[27] W. Liu et al., "K-BERT: Enabling language representation with knowledge graph," in Proc. AAAI, 2020, pp. 2901–2908.

[28] P. Wang, N. Peng, F. Ilievski, P. Szekely, and X. Ren, "Connecting the dots: A knowledgeable path generator for commonsense question answering," 2020, arXiv:2005.00691. [Online]. Available: http: //arxiv.org/abs/2005.00691

[29] X. Li, A. Taheri, L. Tu, and K. Gimpel, "Commonsense knowledge base completion," in Proc. 54th Annu. Meeting Assoc. Comput. Linguistics, vol. 1, 2016, pp. 1445–1455.

[30] D. Maier et al., "Applying LDA topic modeling in communication research: Toward a valid and reliable methodology," Commun. Methods Measures, vol. 12, nos. 2–3, pp. 93–118, Apr. 2018.