GROUP 12 - PROJECT PROGRESS

Domain-Specific Sentiment Analysis in Textual Feedback

AI-Powered Aspect Detection and Sentiment Analysis for Textual Feedback and Reviews

Dr. Nadeesha Sandamali
(Supervisor)

Dr. Kushan Sudheera
(Co-Supervisor)

Dr. Nisansa de Silva
(Co-Supervisor)

Mr. Sachintha Rajith
(Co-Supervisor)

Isuranda M.A.K. EG/2019/3607
Jayakody D.S. EG/2019/3609
Malkith A.V.A. EG/2019/3664

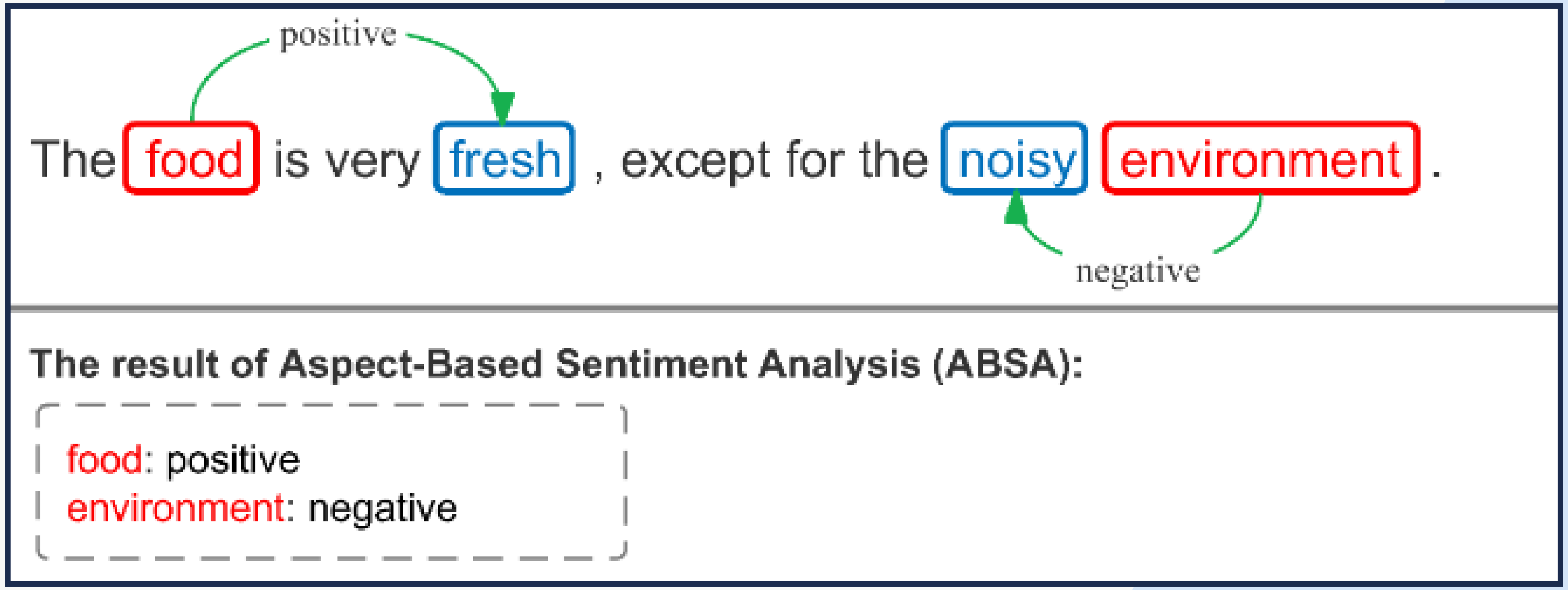
Problem Statement

The Emojot platform struggles with high costs and low accurate sentiment analysis due to using the Open AI, AWS Comprehend , and Auto ML API calls.

Our project aims to develop an **AI solution** to accurately identify sentiments in domain-specific text reviews. Through our **cost-effective** solution, we aim to **reduce expenses while improving accuracy** and enabling better decision-making and user engagement within the Emojot platform.

What is Aspect Based Sentiment Analysis ?

Aspect-Based Sentiment Analysis (ABSA) is a text analysis technique that **breaks down** textual data and **determines** its **sentiment** based on **specific aspects**. This method analyzes consumer feedback by connecting sentiments to various aspects of a product or service.

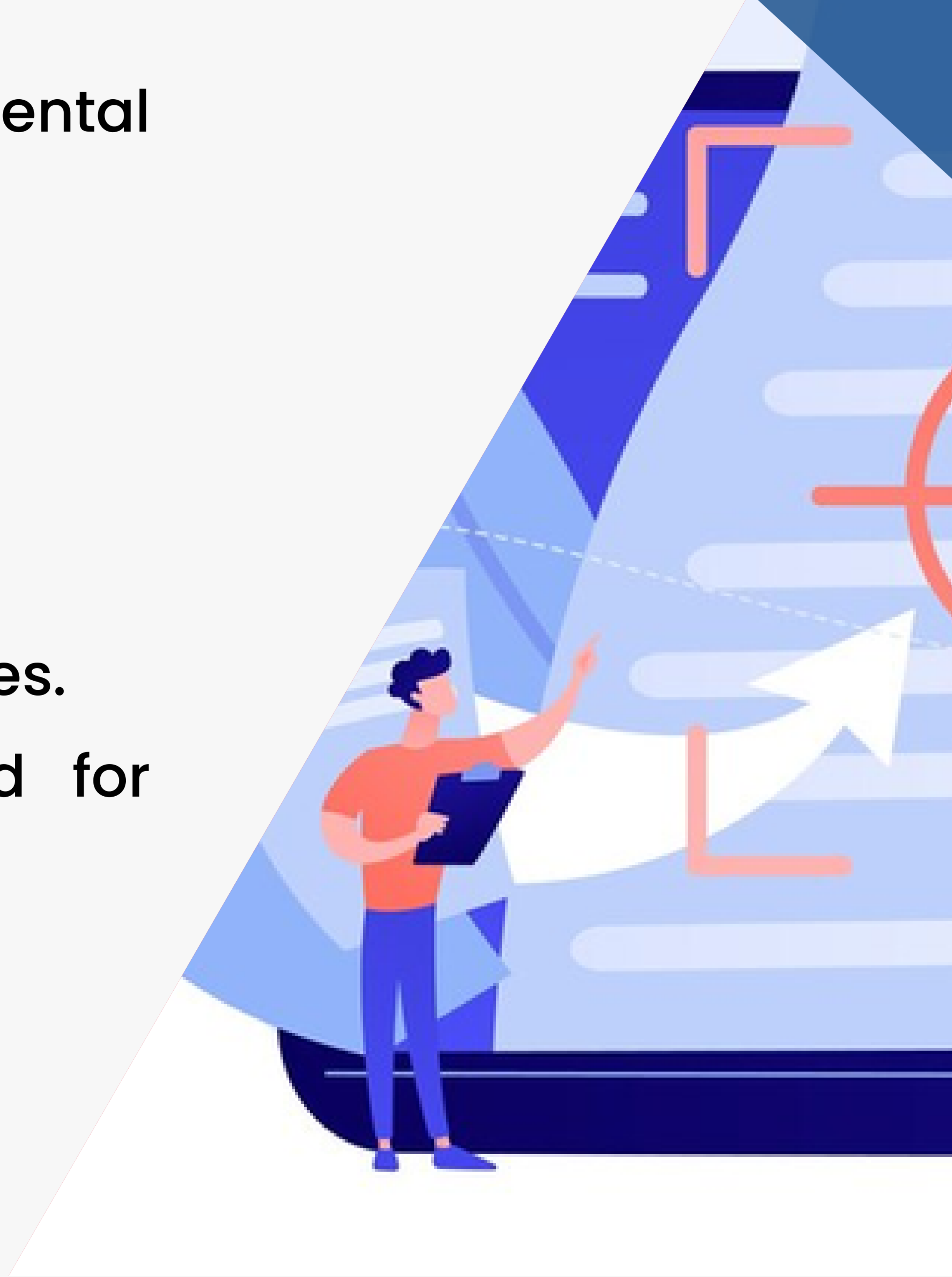


Project Objectives

- Investigate and assess the key differences, advantages, and disadvantages of sentiment analysis techniques.
- Develop a model that can achieve a high accuracy in recognizing aspects from sentences related to a specific domain within textual reviews
- Build a machine learning model with a high accuracy to identify the polarity of sentences in sentiment analysis within the Emojot platform.
- Develop a user-friendly interface.

Project Scope

- Research will be limited to “aspect-based” sentimental analysis methods.
- Reviews are considered under a specific domain.
- Aspects are decided by the clients.
- Only English reviews are analyzed.
- positive, neutral, and negative are the sentiment polarities.
- The interface for the application will be developed for demonstration purposes.



Progress Evolution

APRIL

JUNE

AUGUST

MERCon 2024
Under Review

ICTer 2024
Under Review

Literature Review

Our Main Goals

High Accuracy Model Generation

Cost Effective Implementation

Research Gap

Less research is done in the joint(ATE+ASC) task hybrid model development for aspect based sentiment analysis.

First Objective

- Investigate and assess the key differences, advantages, and disadvantages of sentiment analysis techniques.

Literature Review

Overview of Tasks for Aspect-Based Sentiment Analysis

Task	Input	Output
Aspect Term Extraction (ATE)	S_i	a^1, a^2
Aspect Sentiment Classification (ASC)	$S_i + a^1, S_i + a^2$	sp^1, sp^2
Joint Task (ATE + ASC)	S_i	$(a^1, sp^1), (a^2, sp^2)$

S_i: The price was high, but the restaurant was breathtaking.

Task	Input	Output
Aspect Term Extraction (ATE)	S_i	$price, restaurant$
Aspect Sentiment Classification (ASC)	$S_i + price, S_i + restaurant$	$Negative, Positive$
Joint Task (ATE + ASC)	S_i	$(price, Negative), (restaurant, Positive)$

Literature Review

More than 70 single task models were found.

- **LSTM**
- **BERT**
- **DeBERTa**
- **RoBERTa**
- **Other Models**

F1 Scores of Models Evaluated on the SemEval 2014 Benchmark.

Model	F1 Score (%)							
	Res-14		Lap-14		Res-15		Res-16	
	AE	SP	AE	SP	AE	SP	AE	SP
InstructABSA [7]	**92.10**	85.17	**92.30**	81.56	**76.64**	84.50	**80.32**	89.43
GPT2(med) [98]	75.94	–	82.04	–	–	–	–	–
GRACE [96]	85.45	–	87.93	–	–	–	–	–
LCF-ATEPC-CDM [94]	89.78	85.88	85.29	79.84	–	–	–	–
ASCNN [99]	–	73.10	–	66.72	–	58.90	–	70.43
ASGCN [99]	–	70.48	–	68.06	–	60.78	–	70.29
MCRF-SA [100]	–	73.78	–	74.23	–	61.59	–	75.92
KaGRMN-DSG [101]	–	81.98	–	79.42	–	–	–	–
RAM [70]	–	68.54	–	68.43	–	–	–	–
TN (BiLSTM+C2A+Pas) [53]	–	65.75	–	65.34	–	–	–	–
MGAN [53]	–	71.48	–	71.42	–	–	–	–
CNN-ASP [103]	–	65.11	–	65.31	–	–	–	–
PRET+MULT [102]	–	69.73	–	67.46	–	68.74	–	–
MemNet [68]	–	69.94	–	65.17	–	58.28	–	65.99
IAN [58]	–	70.09	–	67.38	–	52.65	–	55.21
BART-ABSA [97]	87.07	–	83.52	–	75.48	–	–	–
LSAT [95]	–	90.86	–	86.31	–	–	–	–
DeBERTaV3 [6] [90]	–	83.06	–	79.45	–	73.76	–	73.59
ABSA-DeBERTa[93]	–	89.46	–	82.76	–	–	–	–
DeBERTa-V3-base-absa-v1.1* [92, 104]	–	**90.94**	–	**90.32**	–	**89.55**	–	**84.91**
LSA-X-DeBERTa [91]	–	87.02	–	84.41	–	81.29	–	84.87

Note ∗ : The F1-scores for the DeBERTa-v3-base-absa-v1.1 model was calculated by us separately.

BERT [5]	–	71.91	79.28	71.94	–	–	74.10	–
BERT-DK [75]	77.02	75.45	83.55	73.72	–	–	–	–
BERT-SPC [76] [80]	–	76.98	–	75.03	–	–	–	–
BERT-MRC [77]	74.21	74.97	81.06	74.10	–	–	–	–
BERT-PT [78] [79]	–	76.96	84.26	75.08	–	–	81.57	–
BAT [79]	–	76.50	85.57	79.24	–	–	81.50	–
P-SUM [80]	–	79.68	85.94	76.81	–	–	81.99	–
H-SUM [80]	–	79.67	86.09	76.52	–	–	81.34	–
SK-GCN-BERT [81]	–	75.19	–	75.57	–	66.78	–	72.02
SDGCN-BERT [82]	–	76.47	–	78.34	–	–	–	–
RGAT-BERT [39]	–	80.92	–	78.20	–	–	–	–
DGEDT-BERT [105]	–	80.00	–	75.60	–	71.00	–	79.00
DualGCN-BERT [83]	–	81.16	–	78.10	–	–	–	–
BERT-ADA [106]	–	80.05	–	74.09	–	–	–	–
TF-BERT [84]	–	81.15	–	78.46	–	–	–	–
Dual-MRC [107]	–	82.04	–	75.97	–	–	–	–
dotGCN-BERT [85]	–	80.49	–	78.10	–	–	–	–
DPL-BERT [86]	–	84.86	–	78.58	–	–	–	–
SSEGCN-BERT [87]	–	81.09	–	77.96	–	–	–	–
TGCN-BERT [83]	–	79.95	–	77.03	–	82.77	–	72.81
RoBERTa [52]	–	82.10	–	79.73	–	62.41	–	80.88
ASGCN-RoBERTa [89]	–	80.59	–	80.32	–	–	–	–
RGAT-RoBERTa [89]	–	81.29	–	79.95	–	–	–	–
PWCN-RoBERTa [89]	–	80.85	–	81.08	–	–	–	–
SARL-RoBERTa [88]	–	82.44	–	82.97	–	73.83	–	81.92
RoBERTa+MLP [89]	–	80.96	–	80.73	–	–	–	–

MN [108]	–	64.34	–	62.89	–	–	–	–
MN(+AS) [109]	–	69.15	–	65.24	–	–	–	–
TNet [4]	–	71.27	–	71.75	–	–	–	–
TNet-LF [4]	–	71.03	–	70.14	–	59.47	–	70.43
TNet-ATT [109]	–	69.44	–	71.51	–	–	–	–
TNet-AS [4]	–	71.27	–	71.75	–	–	–	–
TNet-ATT(+AS) [109]	–	72.90	–	73.84	–	–	–	–
AE-LSTM [3]	–	64.32	–	62.50	–	–	–	–
ATAE-LSTM [3]	–	64.95	–	62.45	–	–	–	–
TD-LSTM [2]	–	66.73	–	68.43	–	-	–	–
BILSTM-ATT-G [66]	–	70.78	–	69.90	–	–	–	–
LSTM-ATT-CNN [4]	–	68.71	–	68.03	–	–	–	–
LSTM-FC-CNN-LF [4]	–	70.23	–	70.60	–	–	–	–
LSTM+SynATT+TarRep [110]	–	71.32	–	69.23	–	66.05	–	–
LSTM-FC-CNN-AS [4]	–	70.06	–	70.72	–	–	–	–
AEN-BERT [76]	–	73.76	–	76.31	–	–	–	–
AEN-GloVe [76]	–	72.14	–	69.04	–	–	–	–

Literature Review

F1 Scores of Different Models which perform the joint task

Model	Laptop-14	Restaurant-14
InstructABSA [7]	79.34	79.47
GRACE [96]	70.71	77.26
SPAN [112]	68.06	74.92
RACL-BERT [111]	63.40	75.42
BERT-E2E-ABSA [53]	61.12	74.72
DOER [114]	60.35	72.78
IMN [115]	58.37	69.54
E2E-TBSA [113]	57.90	69.80

Second and Third Objectives

- Develop a model that can achieve a high accuracy in recognizing aspects from sentences related to a specific domain within textual reviews

- Build a machine learning model with a high accuracy to identify the polarity of sentences in sentiment analysis

Current Progress

Current Progress

Implemented different approaches to get the most accurate and robust solution.

1. Fine-Tuning LLaMA 2-7B with Quantized Low Rank Adaptation (QLoRA)
2. Fine-Tuning Mistral-7B with Quantized Low Rank Adaptation (QLoRA)
3. SETFIT for efficient few-shot fine-tuning of Sentence Transformers
4. InstructATE-DeBERTaASC (Novel Model)

What is a Large Language Model

A LLM is an **advanced artificial intelligence** designed to **understand** and **generate human-like text** by imposing **massive datasets** and extensive **neural network architectures.**

LLaMA 2 -7B by Meta and **Mistral 7B** by Mistral AI are both 7 billion parameter models designed for general-purpose natural language understanding and generation. Both models are specifically optimized for **high performance** and **efficiency** across various **NLP tasks**.

Fine-Tuning LLMs (LLaMa 2 , Mistral)

1.Parameter Efficient Fine-Tuning (PEFT) with QLoRA on NVIDIA L4 GPU

- **PEFT**
  - Reduce the resources needed for fine-tuning LLMs.
- **QLoRA (Quantized Low-Rank Adaptation)**
  - Uses **decomposed weight matrices** for fewer parameters, maintaining performance with **lower resource demand**

2. Fine-Tuning on Benchmark Datasets and used relevant prompt in the pipeline.

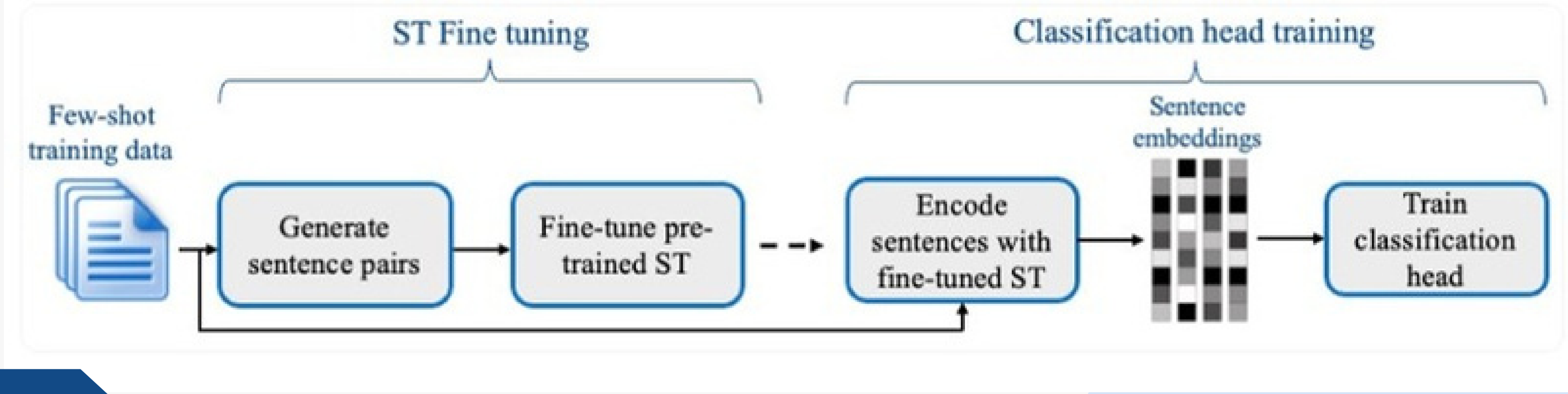
3. Save the fine-tuned models in a model hub separately.

4. Loading and Evaluating Fine-Tuned Models on environment which created in GCP VM.

SetFit

SETFIT(Sentence Transformer Fine-tuning) presents an innovative framework for efficient and prompt-free few-shot fine-tuning of Sentence Transformers (ST).
The SETFIT approach consists of two main steps:

1. Fine-tuning a pretrained ST on a small number of text pairs in a contrastive Siamese manner.
2. Training a classification head using the resulting ST to generate rich text embeddings.



InstructATE-DeBERTaASC

- We developed an aspect-based sentiment analysis pipeline utilizing transformer-based models to automatically extract aspects and analyze sentiments in textual data

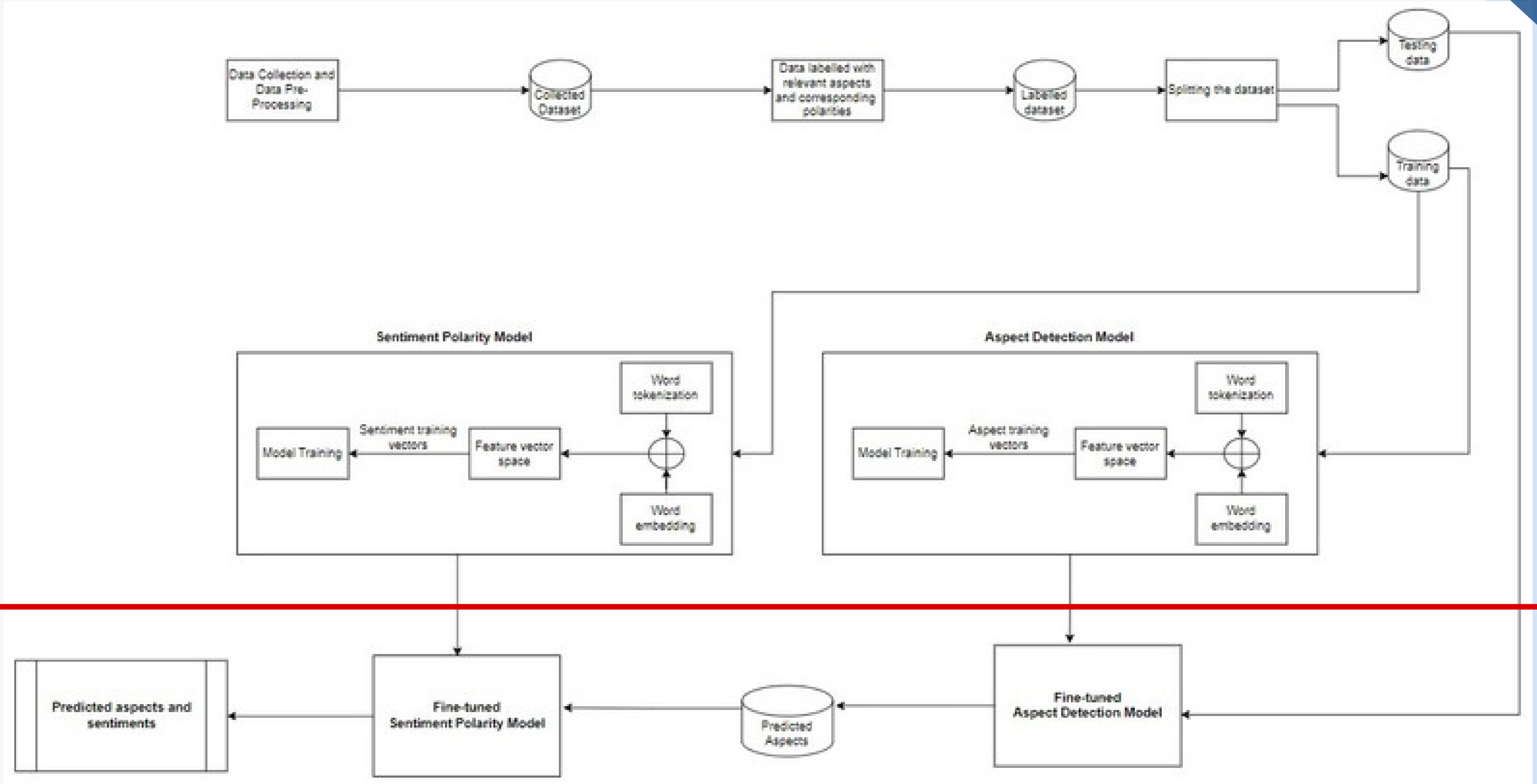
- Our goal was to implement a model that can perform the joint task (ATE + ASC)

- Through our rigorous literature review we identified the best performing models for the single tasks and built our **novel hybrid model**.

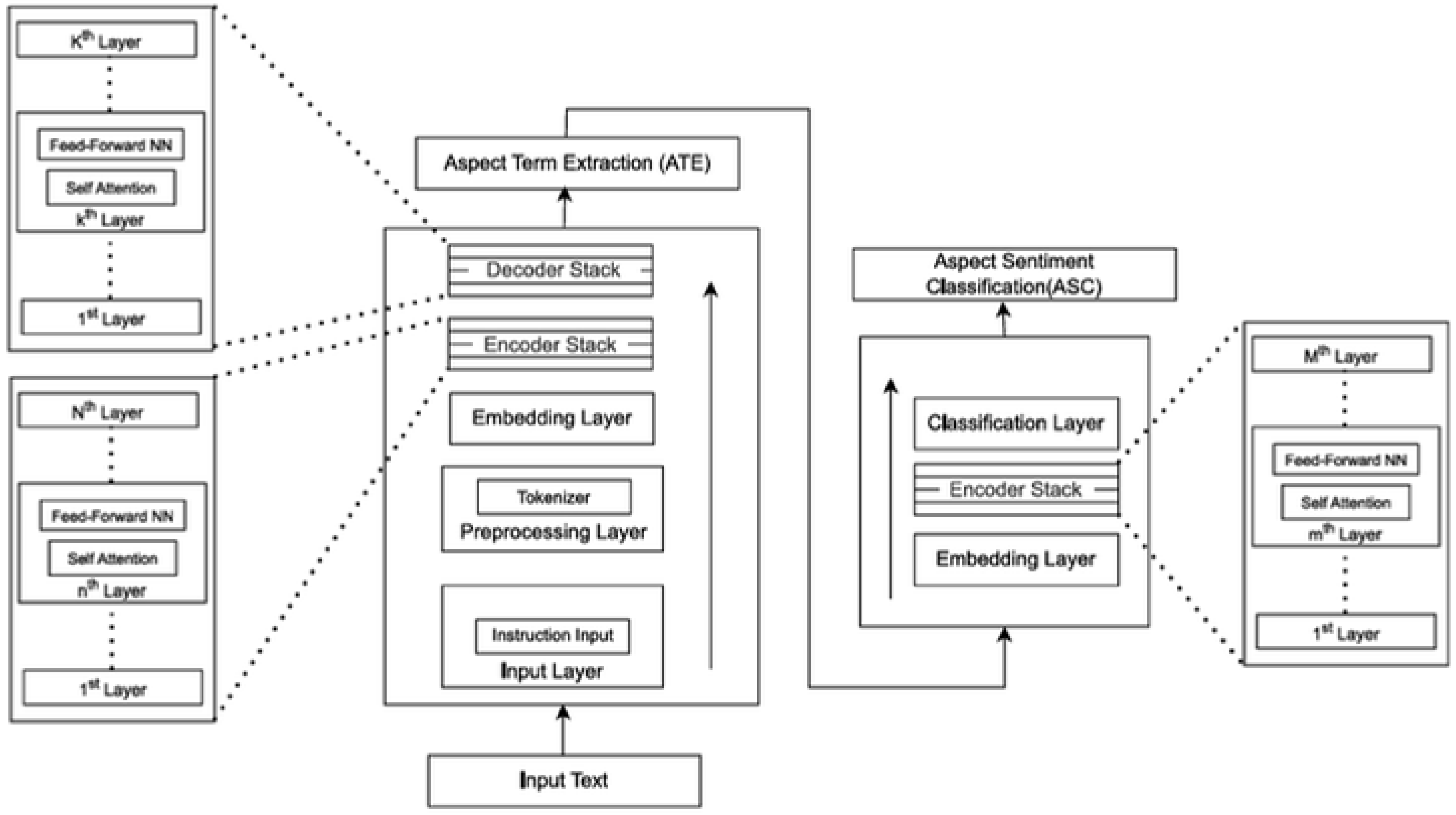
- The sentiment polarity model utilized is **DeBERTa-V3-baseabsa-V1** and the aspect detection model utilized is **InstructABSA.**

- Exploited the strengths of these models to develop our own novel hybrid model named **InstructATE-DeBERTaASC.**

Generalized Novel Model



Layered Architecture of the Novel Model



Model Algorithm

Algorithm 1: Aspect-Based Sentiment Analysis (ABSA)

Require: Review X

Aspect Term Extraction (ATE):

$A_c = f_{ATE}(X)$

Target Aspect Filtering:

$A = f_{filter}(A_c)$

Aspect Sentiment Classification (ASC):

$S = \{\}$

for each aspect term a in A **do**

$s = f_{ASC}(X, a)$

Add (a, s) to S: $S = S \cup \{(a, s)\}$

end for

return Final aspect terms A with sentiment labels S: $\{(a, s) \mid a \in A, s = f_{ASC}(X, a)\}$

Sets:

• X: Review represented as a word sequence (X = {x1, x2, . . . , xn})

• A: Set of extracted aspect terms (A X)

• S: Set of sentiment labels for aspect terms (S {positive, negative, neutral})

Functions:

• f_ATE(X): Function for Aspect Term Extraction. Takes review X and returns candidate terms (Ac). (Ac X)

• f_filter(Ac): A filtering function. Takes candidate terms and returns refined aspects (A). (A Ac)

• f_ASC(X, a): Function for Aspect Sentiment Classification. Takes review X and an aspect term a, returns sentiment labels (S {positive, negative, neutral})

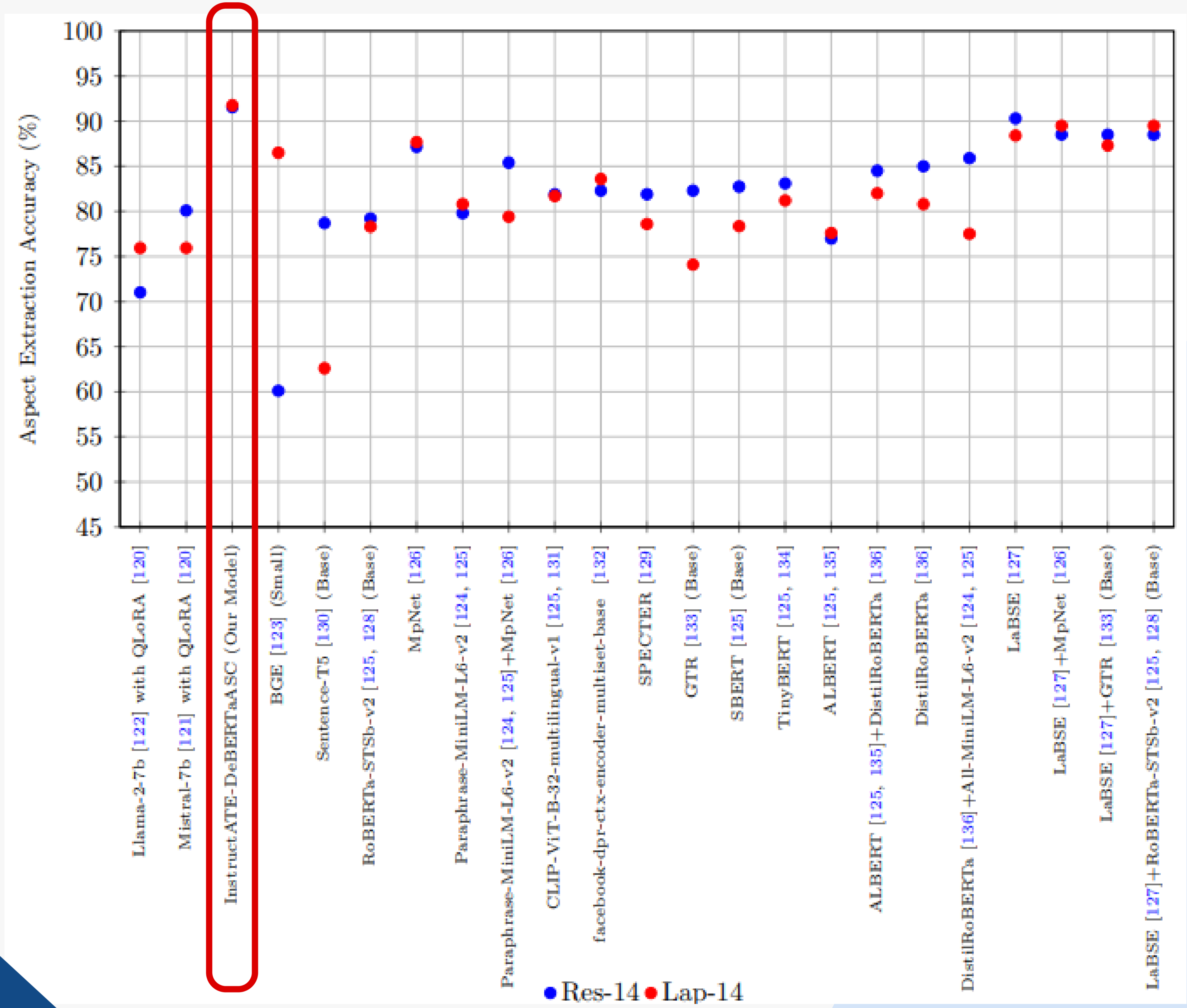
Progress Results

Progress Results

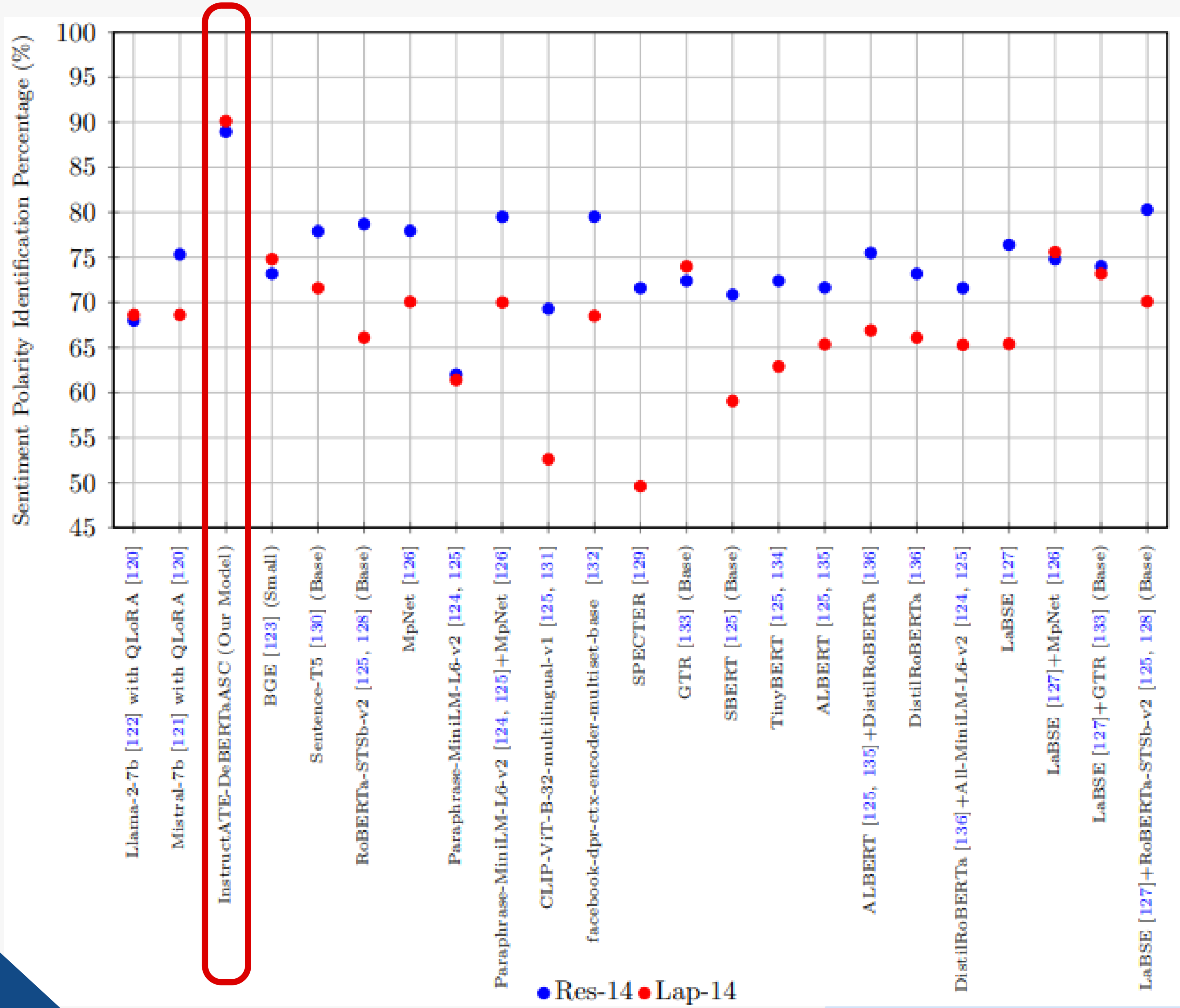
F1 Scores of Joint Task Implemented Models on the SemEval 2014 Benchmark

	Model	F1 Score (%)			
		Res-14		Lap-14	
		Aspect Extraction	Sentiment Polarity	Aspect Extraction	Sentiment Polarity
	Llama-2-7b [122] with QLoRA [120]	71.94	69.29	71.66	66.53
	Mistral-7b [121] with QLoRA [120]	81.33	76.46	77.65	72.40
	InstructATE-DeBERTaASC (Our Model)	**91.39**	**88.63**	**91.56**	**89.65**
SETFIT [8]	BGE [123] (Small)	72.24	75.59	64.79	75.59
	Sentence-T5 [130] (Base)	56.82	78.74	63.29	74.01
	RoBERTa-STSb-v2 [125, 128] (Base)	82.37	77.95	84.26	67.71
	Paraphrase-MiniLM-L6-v2 [124, 125]	84.58	71.65	83.14	64.56
	+MpNet [126]	84.58	78.74	79.40	70.07
	CLIP-ViT-B-32-multilingual-v1 [125, 131]	81.49	59.05	73.03	53.54
	facebook-dpr-ctx-encoder-multiset-base [132]	76.54	78.74	75.52	77.45
	SPECTER [129]	81.93	71.65	77.52	55.11
	GTR [133] (Base)	81.85	74.80	84.70	74.80
	SBERT [125] (Base)	83.18	70.86	84.32	61.41
	TinyBERT [125, 134]	78.76	73.22	82.83	63.77
	ALBERT [125, 135]	80.08	74.80	80.22	66.92
	+DistilRoBERTa [136]	81.49	74.81	79.40	68.50
	DistilRoBERTa [136]	84.95	75.59	80.97	69.29
	+All-MiniLM-L6-v2 [124, 125]	85.46	71.65	81.27	63.77
	MpNet [126]	86.28	77.95	88.80	73.22
	LaBSE [127]	89.38	73.23	90.30	64.57
	+MpNet [126]	88.55	74.80	89.51	75.59
	+GTR [133] (Base)	88.55	74.02	87.27	73.22
	+RoBERTa-STSb-v2 [125, 128] (Base)	90.30	77.17	89.51	70.08

Aspect Extraction Accuracy of Joint Task Implemented Models

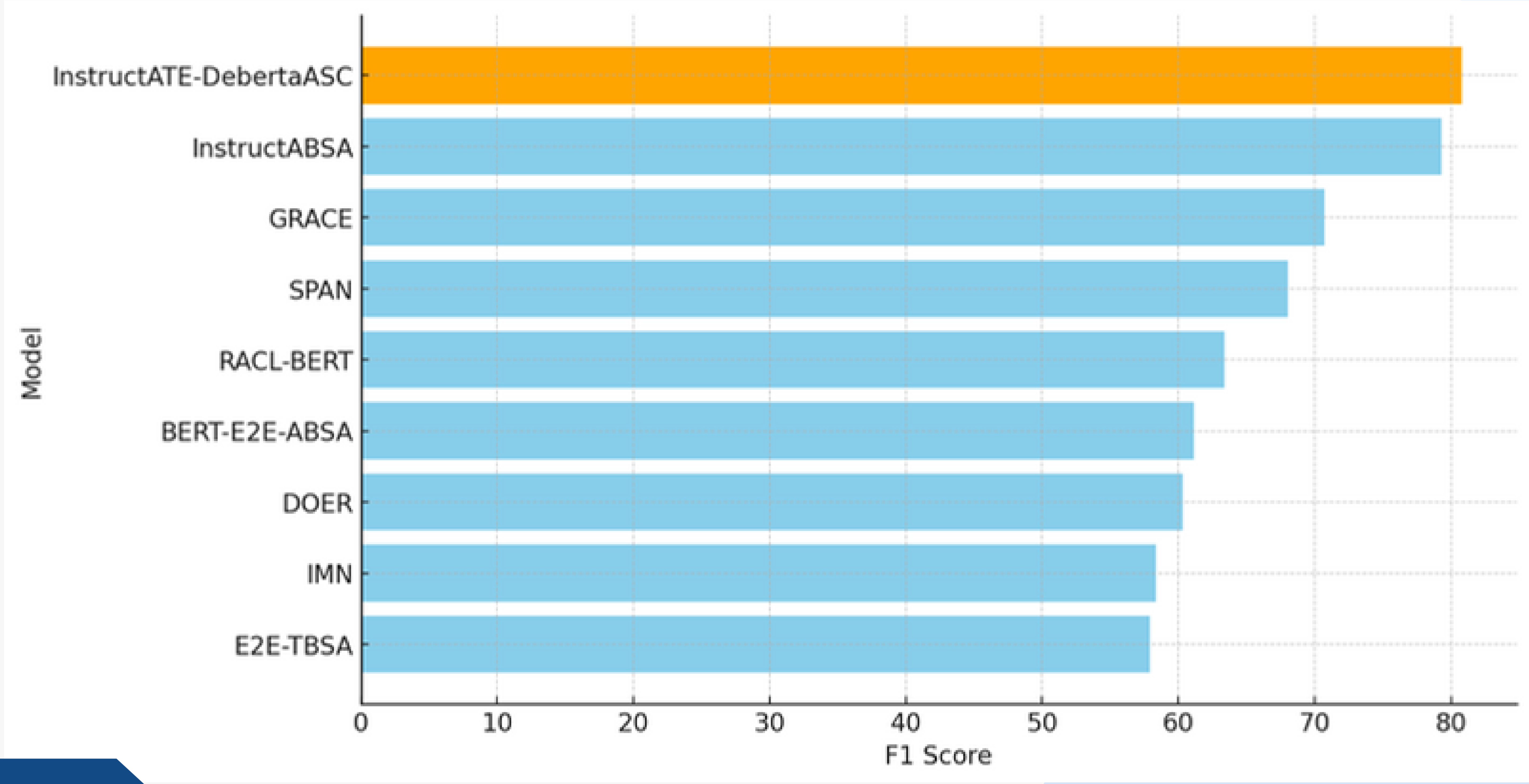


Sentiment Polarity Percentage of Single Task Implemented Models



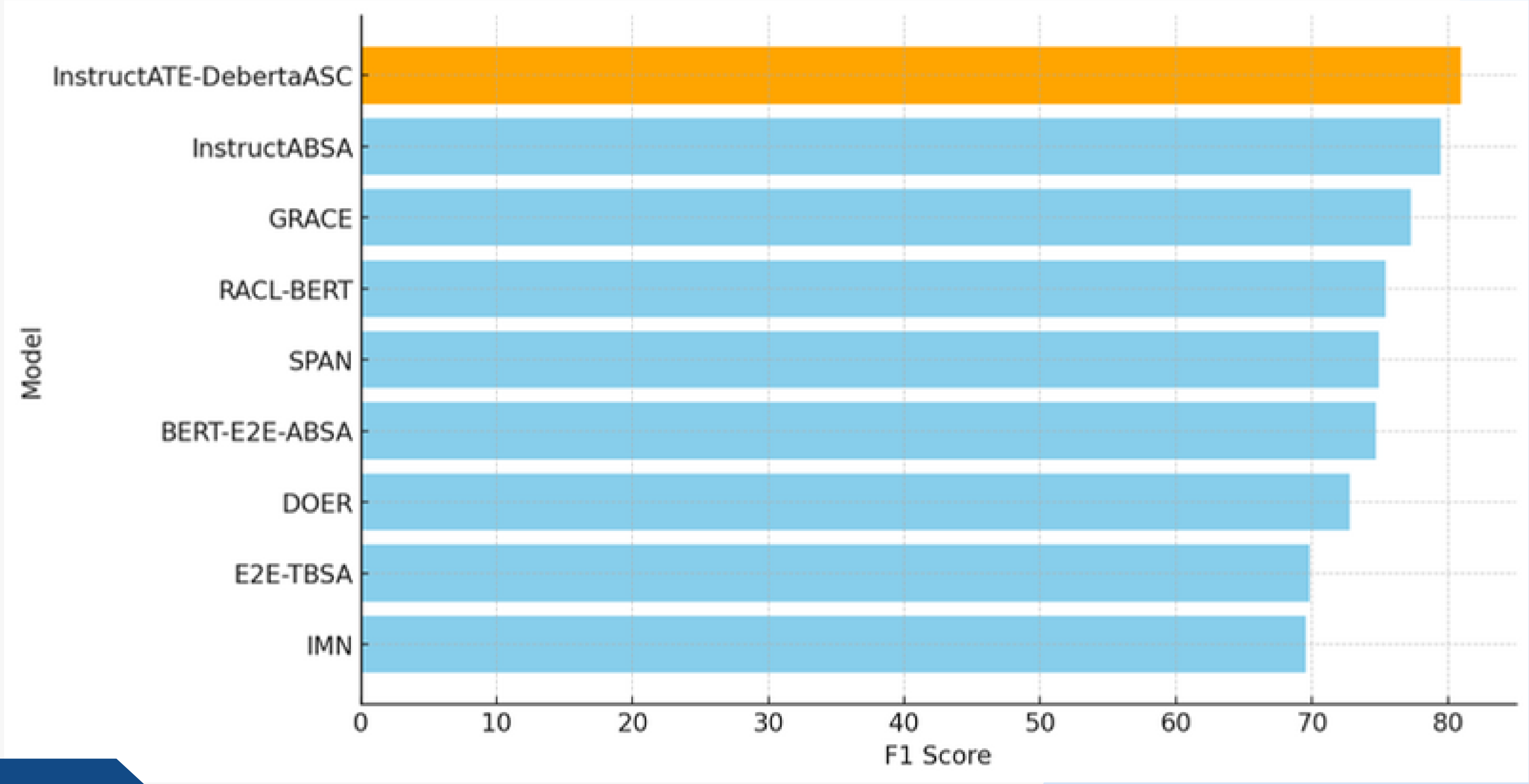
Progress Results

F1 scores of Models for the joint task of ASC and ATE for laptop-14



Progress Results

F1 scores of Models for the joint task of ASC and ATE for restaurant-14



Cost Effectiveness and Benefits for Both Parties

Total Monthly Cost for our proposed model

Approximate Cost of our Project =
+ $26 (Cloud GPU)
+ $20 (Deploy Costs)
+ $6 (Other)
= **$52**
= **LKR 15,500**

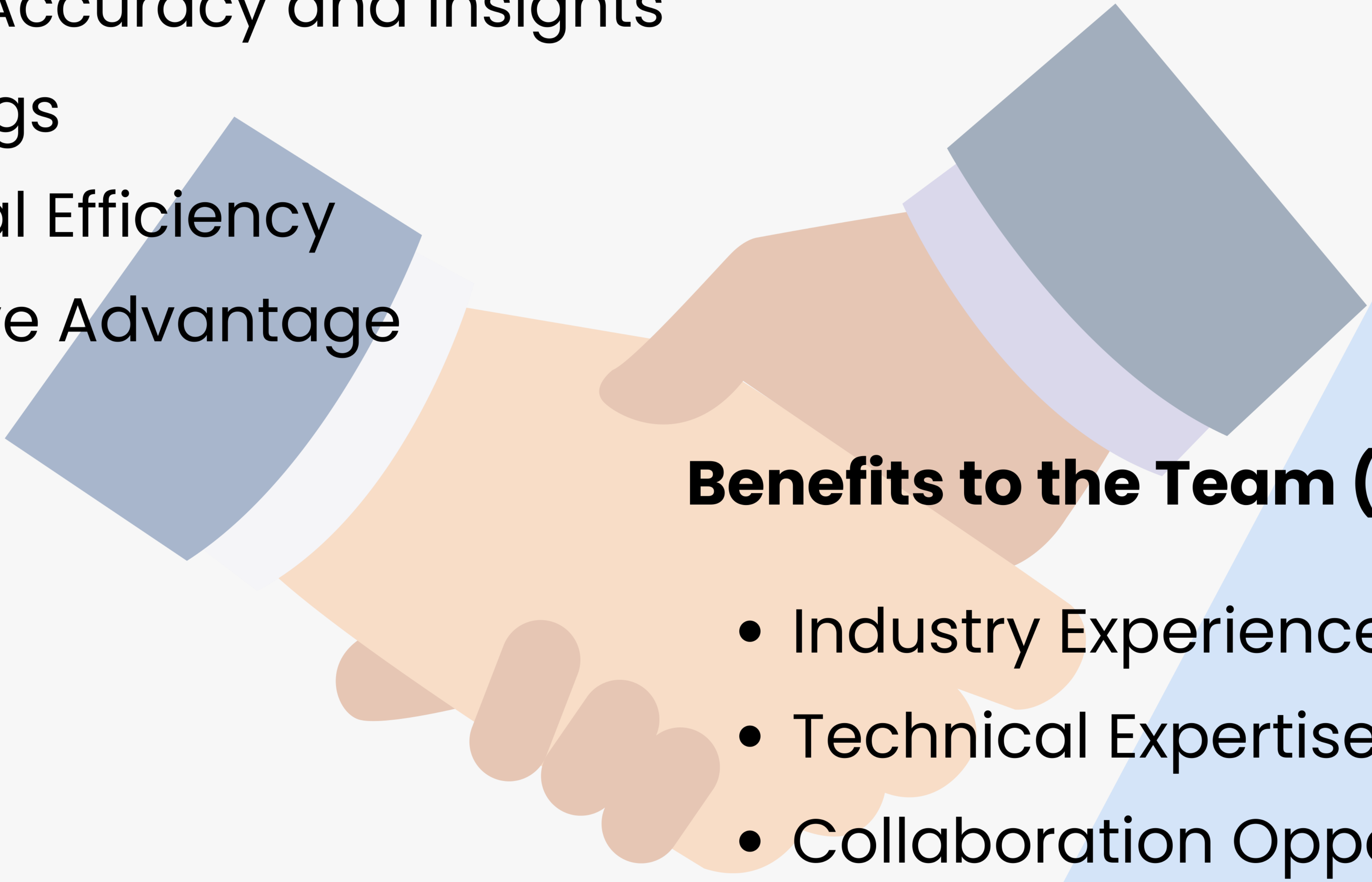
Total Monthly for the company for the current employed system

Total Cost of Company =
+ $3 (Training)
+ $450 (Inference)
+ $0.023 (Storage)
+ $0.00675 (Data Transfer)
= **$453.03**
= **LKR 136,200**

Benefits for both parties

Benefits to Emojot (Pvt) Ltd.

- Improved Accuracy and Insights
- Cost Savings
- Operational Efficiency
- Competitive Advantage

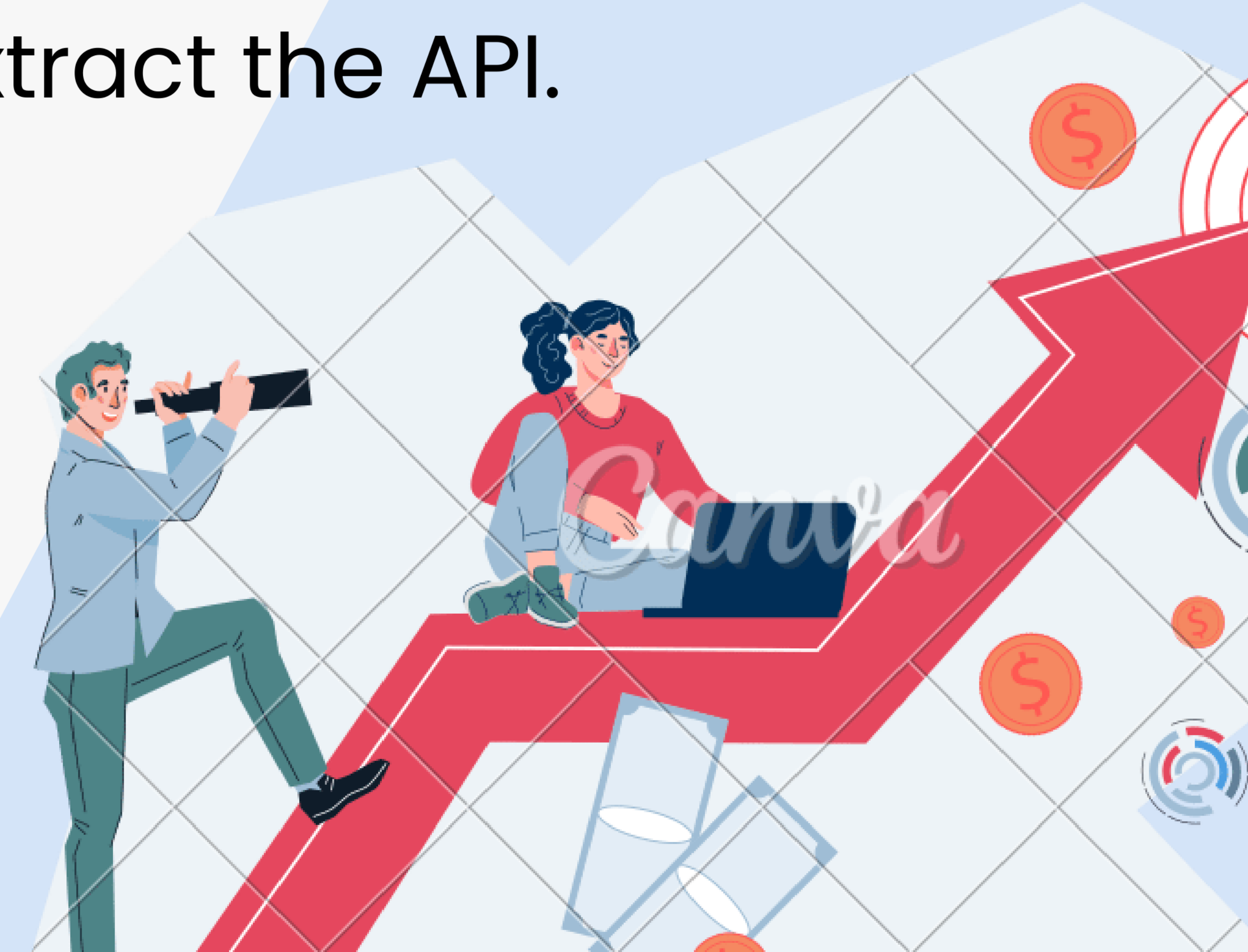


Benefits to the Team (Us)

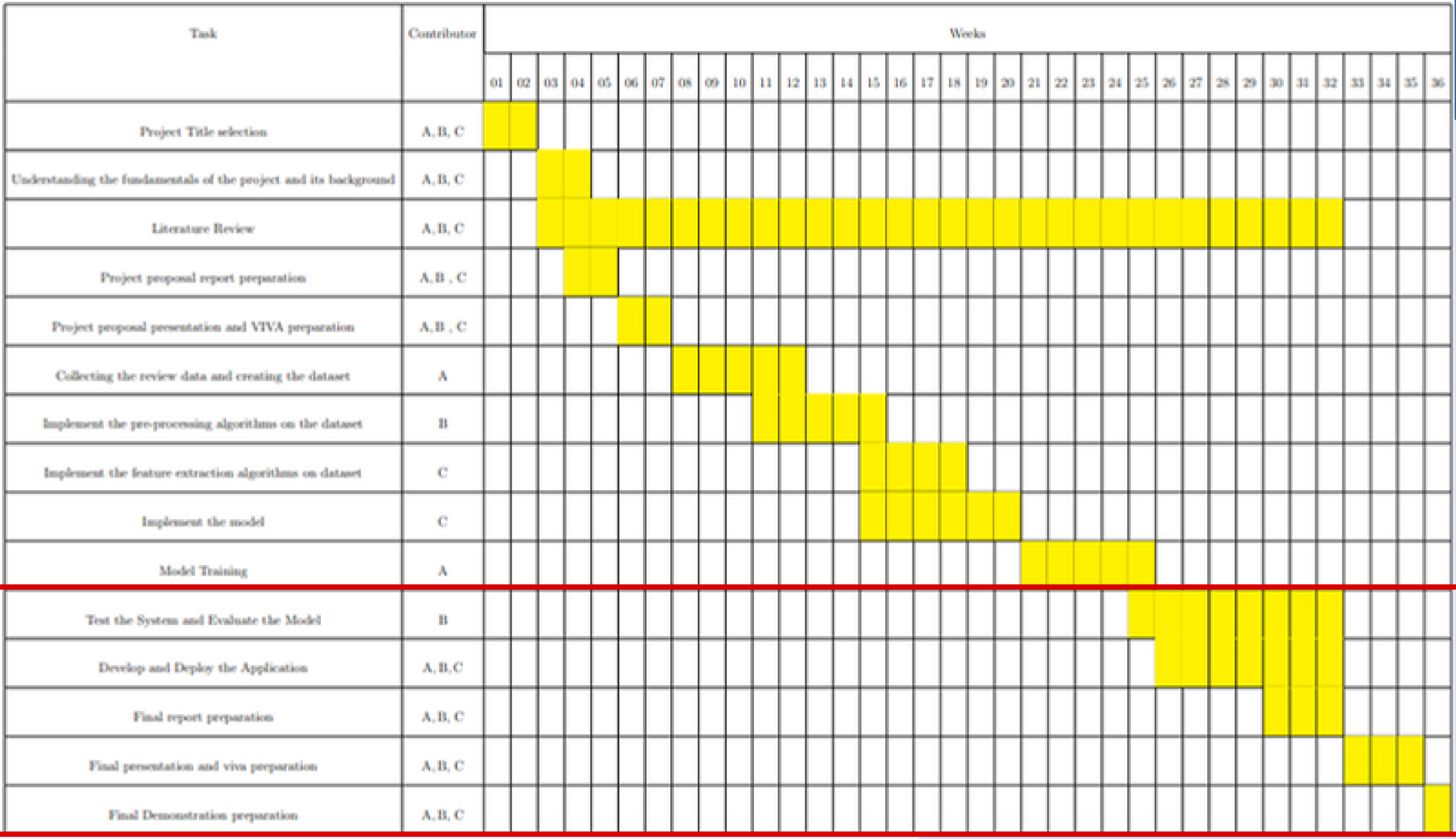
- Industry Experience
- Technical Expertise
- Collaboration Opportunities

Future Works

- Increase the accuracy of the developed novel model further.
- Categorize the aspects according to company requirements
- Develop a user-friendly application.
- Deploy the model in a Cloud Environment and extract the API.
- Extend the research further.
- New implementation for a complete user review experience system



Project Timeline



References

[1] M. Kentour and J. Lu, “An investigation into the deep learning approach in sentimental analysis using graph-based theories,” PLoS One, vol. 16, no. 12, p. e0260761, 2021.
[2] D. Tang, B. Qin, X. Feng, and T. Liu, “Effective LSTMs for target-dependent sentiment classification,” arXiv preprint, arXiv:1512.01100, 2015.
[3] Y. Wang, M. Huang, X. Zhu, and L. Zhao, “Attention-based LSTM for aspect-level sentiment classification,” in Proc. of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016, pp. 606–615.
[4] X. Li, L. Bing, W. Lam, and B. Shi, “Transformation networks for target-oriented sentiment classification,” arXiv preprint, arXiv:1805.01086, 2018.
[5] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint, arXiv:1810.04805, 2018.
[6] P. He, X. Liu, J. Gao, and W. Chen, “DeBERTa: Decoding-enhanced BERT with disentangled attention,” in ICLR, 2020.
[7] K. Scaria et al., “InstructABSA: Instruction learning for aspect-based sentiment analysis,” arXiv preprint, arXiv:2302.08624, 2023.
[8] L. Tunstall et al., “Efficient few-shot learning without prompts,” arXiv preprint, arXiv:2209.11055, 2022.
[9] M. Pontiki et al., “SemEval-2014 Task 4: Aspect-based sentiment analysis,” in SemEval 2014, 2014, pp. 27–35.
[10] A. Ullah, S. N. Khan, and N. M. Nawi, “Review on sentiment analysis for text classification techniques from 2010 to 2021,” Multimedia Tools and Applications, vol. 82, no. 6, pp. 8137–8193, 2023.
[11] A. Ligthart, C. Catal, and B. Tekinerdogan, “Systematic reviews in sentiment analysis: A tertiary study,” Artificial Intelligence Review, pp. 1–57, 2021.
[12] P. Chakriswaran et al., “Emotion AI-driven sentiment analysis: A survey, future research directions, and open issues,” Applied Sciences, vol. 9, no. 24, p. 5462, 2019.
[13] O. Alqaryouti et al., “Aspect-based sentiment analysis using smart government review data,” Applied Computing and Informatics, vol. 20, no. 1/2, pp. 142–161, 2024.

References

[14] P. Bhatia, Y. Ji, and J. Eisenstein, "Better document-level sentiment analysis from RST discourse parsing," arXiv preprint, arXiv:1509.01599, 2015.

[15] B. Yang and C. Cardie, "Context-aware learning for sentence-level sentiment analysis with posterior regularization," in Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2014, pp. 325–335.

[16] S. Poria et al., "Dependency-based rules for concept-level sentiment analysis," Knowledge-Based Systems, pp. 1–32, 2014.

[17] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 3, pp. 813–830, 2015.

[18] J. P. Mellinas, S.-M. M. María-Dolores, and J. J. B. García, "Booking.com: The unexpected scoring system," Tourism Management, vol. 49, pp. 72–74, 2015.

[19] D. M. E.-D. M. Hussein, "A survey on sentiment analysis challenges," Journal of King Saud University-Engineering Sciences, vol. 30, no. 4, pp. 330–338, 2018.

[20] A. Kumar and T. M. Sebastian, "Sentiment analysis: A perspective on its past, present and future," International Journal of Intelligent Systems and Applications, vol. 4, no. 10, pp. 1–14, 2012.

[21] P. K. Jain et al., "Airline recommendation prediction using customer-generated feedback data," in Proc. 2019 4th International Conference on Information Systems and Computer Networks (ISCON), IEEE, 2019, pp. 376–379.

[22] A. Ritter et al., "Open domain event extraction from Twitter," in Proc. of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2012, pp. 1104–1112.

[23] A. Razon and J. Barnden, "A new approach to automated text readability classification based on concept indexing with integrated part-of-speech n-gram features," in Proc. of the International Conference Recent Advances in Natural Language Processing, 2015, pp. 521–528.

[24] A. Sharma et al., "A feature extraction technique using bi-gram probabilities of position specific scoring matrix for protein fold recognition," Journal of Theoretical Biology, vol. 320, pp. 41–46, 2013.

References

[25] M. Straka, J. Hajic, and J. Straková, “UDPipe: Trainable pipeline for processing CONLL-U files performing tokenization, morphological analysis, POS tagging and parsing,” in Proc. of the Tenth International Conference on Language Resources and Evaluation (LREC’16), 2016, pp. 4290–4297.

[26] T. Mikolov et al., “Efficient estimation of word representations in vector space,” arXiv preprint, arXiv:1301.3781, 2013.

[27] Y. Al Amrani, M. Lazaar, and K. E. El Kadiri, “Random forest and support vector machine based hybrid approach to sentiment analysis,” Procedia Computer Science, vol. 127, pp. 511–520, 2018.

[28] P. Bojanowski et al., “Enriching word vectors with subword information,” Transactions of the Association for Computational Linguistics, vol. 5, pp. 135–146, 2017.

[29] M. Ling et al., “Hybrid neural network for Sina Weibo sentiment analysis,” IEEE Transactions on Computational Social Systems, vol. 7, no. 4, pp. 983–990, 2020.

[30] A. Duric and F. Song, “Feature selection for sentiment analysis based on content and syntax models,” Decision Support Systems, vol. 53, no. 4, pp. 704–711, 2012.

[31] G. Rizos, K. Hemker, and B. Schuller, “Augment to prevent: Short-text data augmentation in deep learning for hate-speech classification,” in Proc. of the 28th ACM International Conference on Information and Knowledge Management, 2019, pp. 991–1000.

[32] S. Poria et al., “Sentic patterns: Dependency-based rules for concept-level sentiment analysis,” Knowledge-Based Systems, vol. 69, pp. 45–63, 2014.

[33] G. Adomavicius and Y. Kwon, “Improving aggregate recommendation diversity using ranking-based techniques,” IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 5, pp. 896–911, 2011.

[34] Y. Mezquita et al., “A review of k-NN algorithm based on classical and quantum machine learning,” in Distributed Computing and Artificial Intelligence, Special Sessions, 17th International Conference, Springer, 2021, pp. 189–198.

References

[35] K. L. Chiew et al., “A new hybrid ensemble feature selection framework for machine learning-based phishing detection system,” Information Sciences, vol. 484, pp. 153–166, 2019.
[36] E. Kasmuri and H. Basiron, “Subjectivity analysis in opinion mining—a systematic literature review,” Int. J. Adv. Soft Comput. Appl., vol. 9, no. 3, pp. 132–159, 2017.
[37] M. Crawford et al., “Survey of review spam detection using machine learning techniques,” Journal of Big Data, vol. 2, no. 1, pp. 1–24, 2015.
[38] Z. Fang et al., “An implicit opinion analysis model based on feature-based implicit opinion patterns,” Artificial Intelligence Review, vol. 53, pp. 4547–4574, 2020.
[39] R. Ferreira et al., “A semantic-based method for sentiment analysis in large datasets,” Applied Intelligence, vol. 47, no. 4, pp. 1001–1011, 2017.
[40] S. Montserrat et al., “User behavior modeling for in-store shopping in retailing,” Decision Support Systems, vol. 74, pp. 15–25, 2015.
[41] M. Bandhakavi et al., “Generating and ranking aspect-based summarizations,” in Proc. of the 2017 ACM on Conference on Information and Knowledge Management, 2017, pp. 747–756.
[42] J. Chen et al., “Global optimization of fine-grained opinion summary by dynamic programming,” in Proc. of the 25th International Conference on World Wide Web, 2016, pp. 297–307.
[43] J. Y. Jang and S. H. Bae, “Sentiment analysis of live streaming viewers’ comments based on self-attention mechanism,” IEEE Access, vol. 9, pp. 125799–125809, 2021.
[44] P. Liu, X. Zheng, and X. Li, “Aspect-based sentiment analysis with multi-hop self-attention network,” Neurocomputing, vol. 425, pp. 234–243, 2021.
[45] M. R. Islam, M. Hasan, and A. A. Hashem, “A deep learning-based approach for aspect-based sentiment analysis,” IEEE Access, vol. 7, pp. 148515–148525, 2019.

References

[46] R. Dey and F. M. Salem, “Gate-variants of gated recurrent unit (GRU) neural networks,” in Proc. of the IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), 2017, pp. 1597–1600.
[47] A. V. d. Oord, S. Dieleman, and B. Schrauwen, “Deep content-based music recommendation,” in Proc. of the 26th International Conference on Neural Information Processing Systems (NIPS), 2013, pp. 2643–2651.
[48] X. Zhang and Y. LeCun, “Text understanding from scratch,” arXiv preprint, arXiv:1502.01710, 2015.
[49] X. Zhang, J. Zhao, and Y. LeCun, “Character-level convolutional networks for text classification,” in Proc. of the 28th International Conference on Neural Information Processing Systems (NIPS), 2015, pp. 649–657.
[50] L. Yang, Z. Yang, and C. Dyer, “Hierarchical attention networks for document classification,” in Proc. of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, pp. 1480–1489.
[51] P. Bojanowski et al., “Enriching word vectors with subword information,” Transactions of the Association for Computational Linguistics, vol. 5, pp. 135–146, 2017.
[52] Y. Kim, “Convolutional neural networks for sentence classification,” in Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1746–1751.
[53] A. Sun, S. Lo, and J. Tang, “A hierarchical transformer model for document classification,” in Proc. of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 3281–3287.
[54] A. Agarwal et al., “Learning sentiment-specific word embeddings for twitter sentiment classification,” in Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2014, pp. 30–35.
[55] C. B. Jones et al., “GIScience: Understanding and addressing the grand challenges of spatial data science,” International Journal of Geographical Information Science, vol. 33, no. 12, pp. 2499–2514, 2019.
[56] P. D. Turney, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” in Proc. of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), 2002, pp. 417–424.

References

[57] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? Sentiment classification using machine learning techniques,” in Proc. of the ACL-02 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2002, pp. 79–86.

[58] M. T. P. Adamopoulou et al., “Mining opinions on the go: Summarizing Yelp reviews using a lexicon-based approach,” in Proc. of the IEEE 16th International Conference on Data Mining Workshops (ICDMW), 2016, pp. 626–633.

[59] Y. Zhang et al., “Deep learning for sentiment analysis: A survey,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 9, no. 3, p. e1253, 2019.

[60] H. He and D. Kifer, “Adversarial active learning for sequence labeling and generation,” arXiv preprint, arXiv:2001.05665, 2020.

[61] W. Tan et al., “Sentiment analysis for stock price prediction using deep learning models with Bayesian optimization,” Applied Soft Computing, vol. 96, p. 106734, 2020.

[62] L. Balakrishnan and N. Narayanaswamy, “Aspect-level sentiment analysis using unsupervised models,” Journal of Intelligent Information Systems, vol. 56, no. 3, pp. 527–550, 2021.

[63] S. Li et al., “Mutual learning of sentiment and semantic representations for aspect-level sentiment analysis,” arXiv preprint, arXiv:1911.08375, 2019.

[64] H. Xu et al., “Targeted aspect-based sentiment analysis for drug reviews,” BMC Medical Informatics and Decision Making, vol. 22, no. 1, pp. 1–14, 2022.

[65] Y. Sun, H. Yin, and J. Chen, “Deep convolutional network for sentiment analysis of short texts,” IEEE Transactions on Computational Social Systems, vol. 6, no. 4, pp. 907–917, 2019.

[66] Z. Zhang, “Deep learning based sentiment analysis: A survey,” arXiv preprint, arXiv:1801.07883, 2018.

[67] F. Seyyed-Kalantari and M. Rahgozar, “Fuzzy-based aspect sentiment analysis for Persian reviews,” Computers and Electrical Engineering, vol. 87, p. 106787, 2020.

[68] M. K. M. Rahim et al., “Improved aspect-based sentiment analysis with double embeddings and multi-task learning,” IEEE Access, vol. 8, pp. 100031–100043, 2020.

References

[68] M. K. M. Rahim et al., “Improved aspect-based sentiment analysis with double embeddings and multi-task learning,” IEEE Access, vol. 8, pp. 100031–100043, 2020.
[69] S. Saha et al., “Sentence-level sentiment analysis in Bangla using BiLSTM-CNN,” IEEE Access, vol. 7, pp. 67335–67344, 2019.
[70] J. Devlin et al., “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pp. 4171–4186.
[71] P. Bojanowski et al., “Enriching word vectors with subword information,” Transactions of the Association for Computational Linguistics, vol. 5, pp. 135–146, 2017.
[72] M. Ulinski, “Survey on aspect-based sentiment analysis,” arXiv preprint, arXiv:2106.15381, 2021.
[73] J. A. Hartigan and M. A. Wong, “Algorithm AS 136: A k-means clustering algorithm,” Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 28, no. 1, pp. 100–108, 1979.
[74] T. Mikolov et al., “Efficient estimation of word representations in vector space,” in Proc. of the International Conference on Learning Representations (ICLR), 2013, pp. 1–12.
[75] W. Yin et al., “Comparative study of CNN and RNN for natural language processing,” arXiv preprint, arXiv:1702.01923, 2017.
[76] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint, arXiv:1412.6980, 2014.
[77] M. Zhang et al., “Aspect-based sentiment analysis using deep learning: A survey,” IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 3, pp. 738–755, 2020.
[78] A. Peixoto et al., “A review of machine learning techniques for sentiment analysis: Challenges and solutions,” Pattern Recognition Letters, vol. 124, pp. 1–12, 2019.
[79] T. Mikolov et al., “Distributed representations of words and phrases and their compositionality,” in Proc. of the 26th International Conference on Neural Information Processing Systems (NIPS), 2013, pp. 3111–3119.

References

[80] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532–1543.

[81] Y. Kim et al., "Character-aware neural language models," in Proc. of the 30th AAAI Conference on Artificial Intelligence, 2016, pp. 2741–2749.

[82] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," in Proc. of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 328–339.

[83] A. Joulin et al., "Bag of tricks for efficient text classification," arXiv preprint, arXiv:1607.01759, 2016.

[84] Z. Yang et al., "Hierarchical attention networks for document classification," in Proc. of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, pp. 1480–1489.

[85] X. Zhang and Y. LeCun, "Text understanding from scratch," arXiv preprint, arXiv:1502.01710, 2015.

[86] R. Collobert et al., "Natural language processing (almost) from scratch," Journal of Machine Learning Research, vol. 12, no. 70, pp. 2493–2537, 2011.

[87] A. Vaswani et al., "Attention is all you need," in Proc. of the 31st International Conference on Neural Information Processing Systems (NIPS), 2017, pp. 5998–6008.

[88] T. Mikolov et al., "Recurrent neural network based language model," in Proc. of the 11th International Conference on Artificial Intelligence and Statistics (AISTATS), 2010, pp. 587–595.

[89] T. Wolf et al., "Transformers: State-of-the-art natural language processing," in Proc. of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 2020, pp. 38–45.

[90] D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed., 2021.

[91] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," in Proc. of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 328–339.

References

[92] A. Vaswani et al., “Attention is all you need,” arXiv preprint, arXiv:1706.03762, 2017.

[93] S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python, 1st ed., O'Reilly Media, 2009.

[94] M. Tan et al., “EfficientNet: Rethinking model scaling for convolutional neural networks,” in Proc. of the 36th International Conference on Machine Learning (ICML), 2019, pp. 6105–6114.

[95] H. He, K. Zhang, and J. D. Villalobos, “Toward practical visual sentiment analysis: A transductive learning method,” IEEE Transactions on Affective Computing, vol. 10, no. 4, pp. 639–652, 2019.

[96] M. T. Ribeiro, S. Singh, and C. Guestrin, “"Why should I trust you?": Explaining the predictions of any classifier,” in Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2016, pp. 1135–1144.

[97] D. D. Lewis, “Naive (Bayes) at forty: The independence assumption in information retrieval,” in Proc. of the 10th European Conference on Machine Learning (ECML), 1998, pp. 4–15.

[98] D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed., 2021.

[99] J. J. O'Connor and E. F. Robertson, “Satoru Iwata,” MacTutor History of Mathematics Archive, 2015.

[100] H. He, K. Zhang, and J. D. Villalobos, “Toward practical visual sentiment analysis: A transductive learning method,” IEEE Transactions on Affective Computing, vol. 10, no. 4, pp. 639–652, 2019.

[101] M. A. Hearst, “Direction-based text interpretation as an information access refinement,” arXiv preprint, arXiv:1911.08034, 2019.

[102] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” Annals of Eugenics, vol. 7, no. 2, pp. 179–188, 1936.

[103] M. R. Islam, M. Hasan, and A. A. Hashem, “A deep learning-based approach for aspect-based sentiment analysis,” IEEE Access, vol. 7, pp. 148515–148525, 2019.

[104] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2016, pp. 785–794.

[105] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

References

[106] A. Graves, N. Jaitly, and A.-r. Mohamed, “Hybrid speech recognition with deep bidirectional LSTM,” in Proc. of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), 2013, pp. 273–278.

[107] H. Bai, “Recommender system based on user behavior analysis,” Pattern Recognition Letters, vol. 123, pp. 32–37, 2019.

[108] A. Goldstein et al., “An information-theoretic approach to understanding deep neural networks,” in Proc. of the 34th International Conference on Machine Learning (ICML), 2017, pp. 984–993.

[109] S. Wang et al., “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 82–97, 2012.

[110] R. T. Smith, “An overview of the Tesseract OCR engine,” in Proc. of the Ninth International Conference on Document Analysis and Recognition (ICDAR), vol. 2, 2007, pp. 629–633.

Thank You

For Your Attention

Supplementary Slides

For Literature Review
Single Task Models

Detailed Literature Review

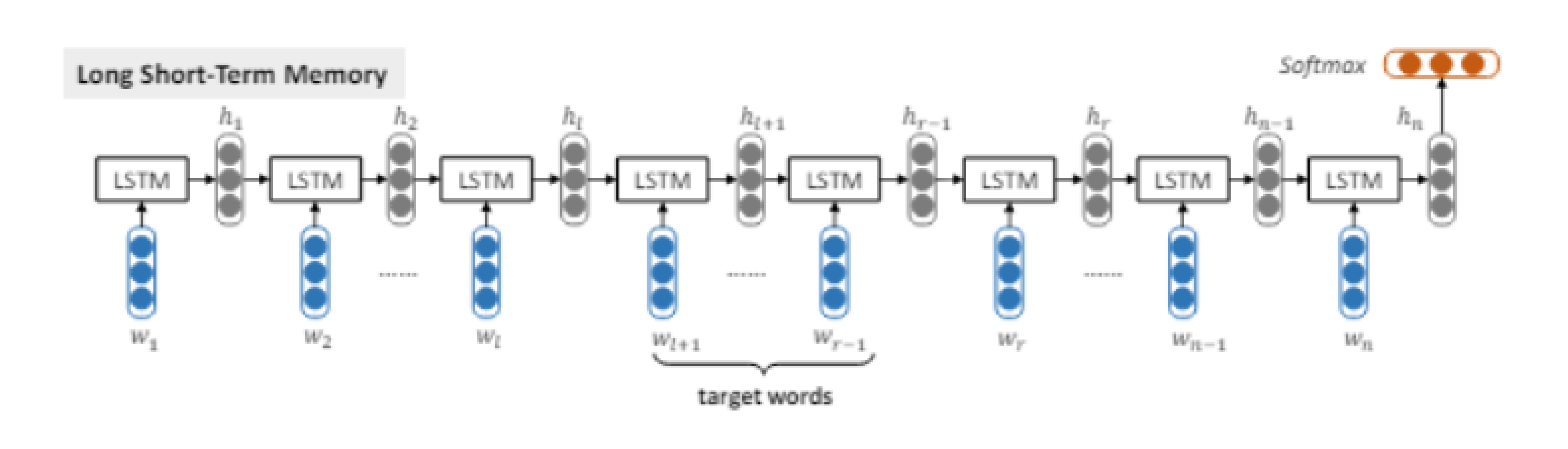
LSTM Based

- Recurrent Neural Networks (RNNs) are an extension of conventional feed-forward neural networks designed to handle sequential data by integrating loops within the network.

- RNNs got problems as, primarily the vanishing gradient and exploding gradient problems.

- In order to overcome these challenges, Long Short-Term Memory (LSTM) network was developed.

Detailed Literature Review



Detailed Literature Review

LSTM Based

- To effectively extract aspect information, Wang et al. proposed a model called **LSTM with Aspect Embedding (AE-LSTM)**
- By learning an embedding vector for each aspect, the AE-LSTM can capture the context and sentiment associated with different aspects within a sentence properly.
- .In the AE-LSTM model, aspect information is included by enabling the aspect embedding which affect the calculation of attention weights

Detailed Literature Review

LSTM Based

- To further leverage aspect information, Wang et al. developed an enhanced model called **Attention-based LSTM with Aspect Embedding (ATAE-LSTM)**

- **Target-Dependent LSTM (TD-LSTM)** model uses two LSTM networks.

- Building on the vanilla LSTM model, a new model named **BILSTM-ATT-G** was developed by Liu and Zhang. This model employs a bidirectional LSTM to represent the input word sequence, enhancing the model's ability to capture dependencies from both directions

Detailed Literature Review

LSTM Based

- Li et al. proposed **Target-Specific Transformation Networks (TNet)**, a new architecture designed to improve target sentiment classification by effectively handling multiple targets and extracting relevant features without introducing noise.

- TNet introduces a novel Target-Specific Transformation (TST) component for generating target-specific word representations. TNet aims to enhance the strengths of both RNNs and CNNs while mitigating their respective limitations.

- To enhance context information, which has been shown to be useful by Lai et al. , two strategies were investigated: Lossless Forwarding (LF) [4] and Adaptive Scaling (AS) .The model variants are named **TNet-LF and TNet-AS.**

Detailed Literature Review

LSTM Based

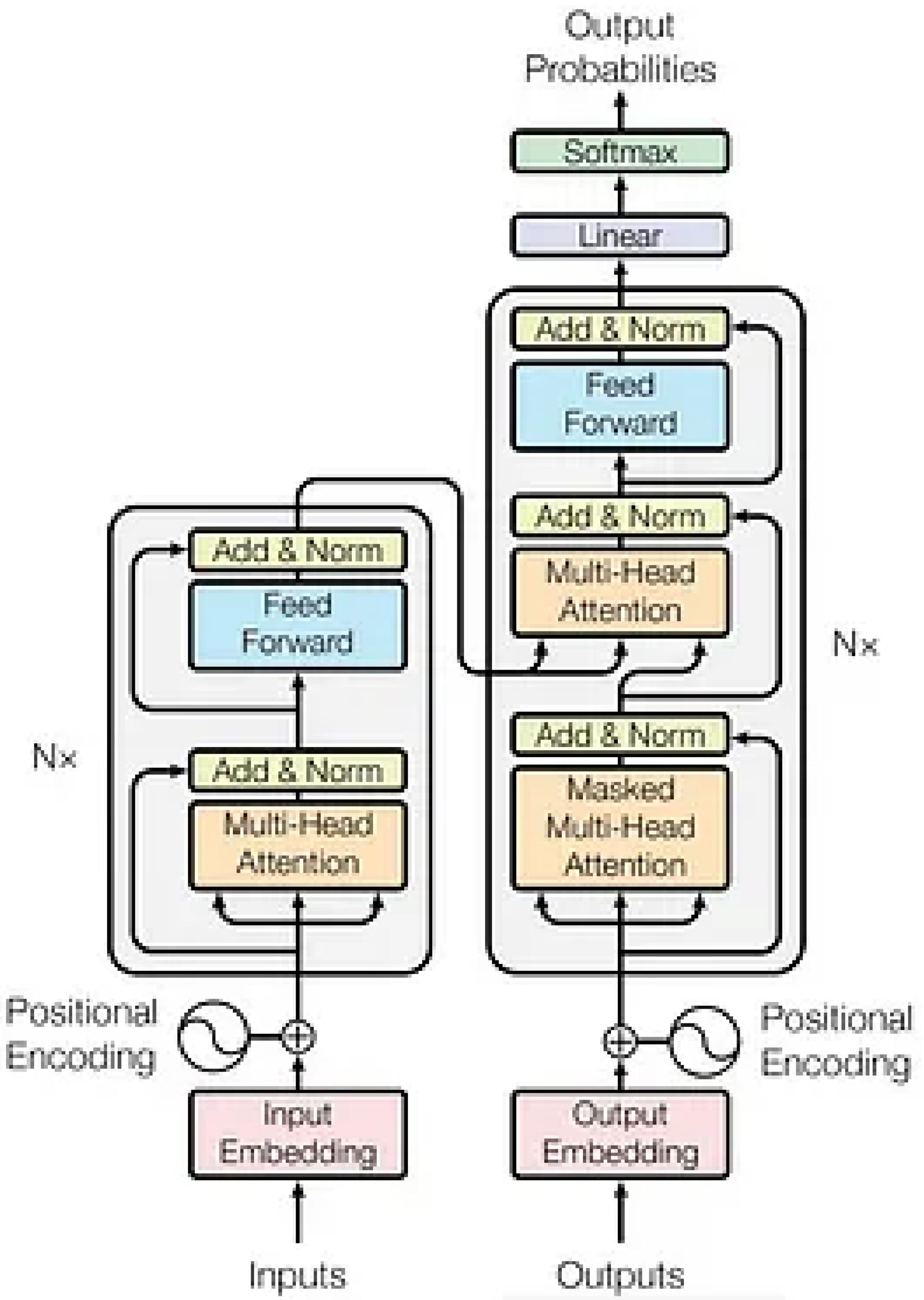
- **LSTM-ATT-CNN** which is mentioned by Li et al. applies attention as an alternative and does not require a context-preserving mechanism, but it performs significantly worse than the TNet variants

- The models **LSTM-FC-CNN-LF** and **LSTM-FC-CNN-AS** are built by Li et al. by applying a fully connected (FC) layer to replace the TST and retaining the context-preserving mechanism

Detailed Literature Review

BERT

Encoder



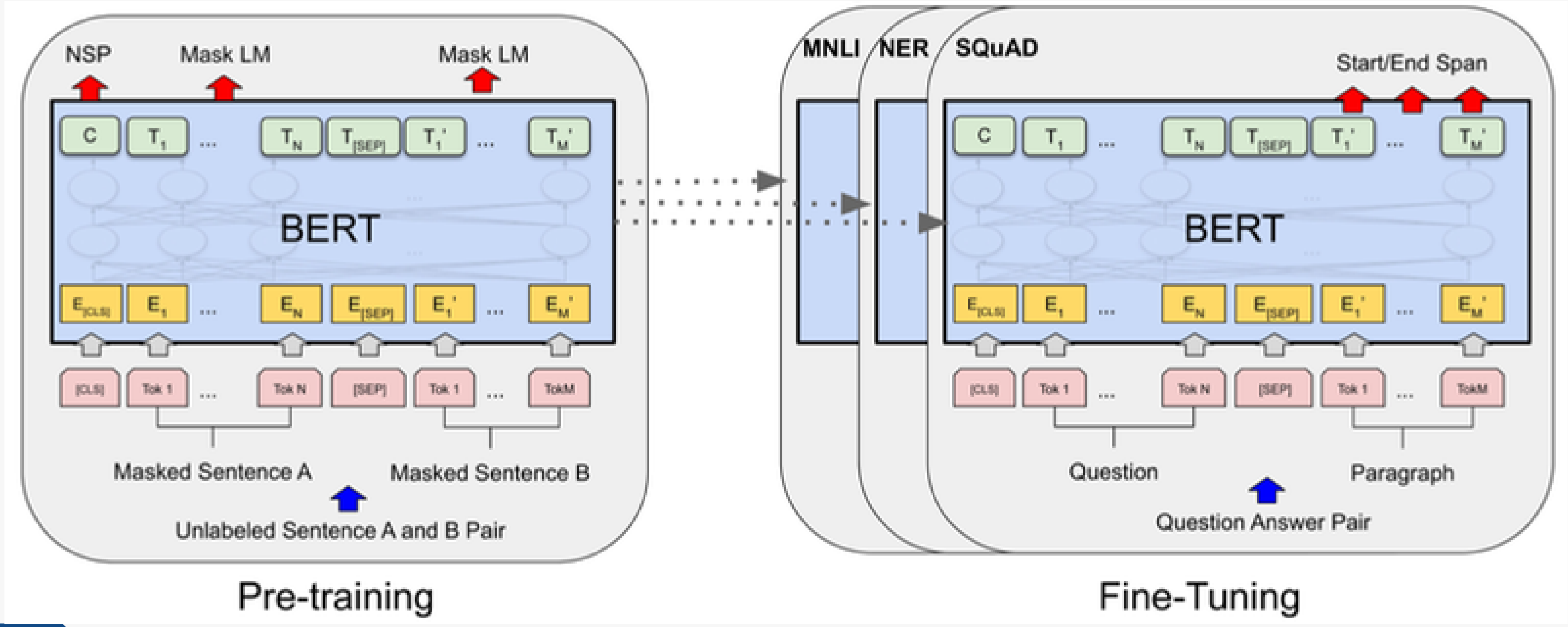
GPT

Decoder

Detailed Literature Review

BERT Based

- **BERT,** which stands for Bidirectional Encoder Representations from Transformers, enhances traditional Transformers by eliminating the unidirectional constraint.



Detailed Literature Review

BERT Based

- BERT’s training process involves two main steps. Pre training and fine tuning.
- During the pre-training phase, BERT is trained on a large corpus of unlabeled text data through various pre-training tasks, including the MLM and next sentence prediction.
- Once pre-trained, BERT is fine-tuned for specific downstream tasks using labeled data. During fine-tuning, all the parameters are adjusted based on the specific requirements of the task.
- Although different tasks use the same pre-trained BERT model as a starting point, each task results in a separately fine-tuned model optimized for that particular application.

Detailed Literature Review

BERT Based

- Researchers have developed enhanced versions of BERT.

- **BERT-DK**, introduced by Zhao, integrates domain-specific knowledge to improve ABSA performance

- **BERT-SPC**, developed by Song et al., employs a Sentence Pair Classification framework to better understand the context of aspect-specific sentences. The sentence pair framework helps in understanding relationships between sentences, thereby improving sentiment classification accuracy.

- Although different tasks use the same pre-trained BERT model as a starting point, each task results in a separately fine-tuned model optimized for that particular application.

Detailed Literature Review

BERT Based

- Innovative approaches such as **BERT-MRC**, proposed by Zhao et al. , frame ABSA tasks as machine reading comprehension problems.

- **BERT-PT** which involves pre-training BERT on domain-specific data followed by fine-tuning.

- **BAT** which stands for **BERT with Adversarial Training**, introduced by Karimi et al., enhances ABSA by generating adversarial examples during training .

- Another approach taken by Karimi et al. utilizes summarization-based approaches to enhance ABSA performance. **P-SUM** and **H-SUM** are such proposed methods.

Detailed Literature Review

BERT Based

- Innovative approaches such as **BERT-MRC**, proposed by Zhao et al. , frame ABSA tasks as machine reading comprehension problems.

- Zhou et al. developed **SK-GCN-BERT** which integrates structural information from sentences and uses graph convolutional networks to capture complex relationships between aspects and sentiments.

- **SDGCN-BERT**, introduced by Zhao et al, combines semantic dependency graphs with BERT. These models use structural and dependency information to enhance their understanding of the text, which eventually improves the performance of ABSA tasks.

Detailed Literature Review

BERT Based

- **RGAT-BERT**, proposed by Bai et al., uses relational graph attention networks to improve aspect extraction and sentiment classification abilities. The use of relational information significantly improves the model's ability to understand and classify relationships within text.

- **DualGCN-BERT** introduced by Li et al., uses dual graph convolutional networks to handle both aspect extraction and sentiment classification. The dual graph approach helps the model to balance and improve both tasks effectively.

- **TF-BERT,** developed by Zhang et al., uses task-specific fine-tuning strategies to improve ABSA performance. Fine-tuning strategies enable the model to adapt better to the complexities of ABSA tasks

Detailed Literature Review

BERT Based

- **DotGCN-BERT,** proposed by Chen et al., uses dot-product based **graph convolutional networks** to improve ABSA performance. By using dot-product operations, this model is able to effectively understand and classify relationships between aspects and sentiments.

- **DPL-BERT** proposed by Zhang et al. has deep prompt learning, which makes the prompt update rapid when new ABSA tasks come through. Prompt learning effectively guides the learning of the model itself to have a sharp focus on one task.

- **SSEGCN-BERT** proposed in Zhang et al. introduces semantically structural embeddings into the BERT pre-trained model.

Detailed Literature Review

RoBERTa Based

- **RoBERTa** (Robustly optimized BERT approach) is an advanced language model that builds upon the foundational work of BERT.

- The authors of RoBERTa carefully examined the effects of important hyperparameters and training data size through a replication study of BERT pretraining, and they found that BERT was significantly undertrained.

- RoBERTa’s development involved several critical modifications to BERT’s original design: **training the model for a longer period with larger batch sizes** and **over more data**, **removing the next sentence prediction objective** which was part of BERT’s original training regimen, **training on longer sequences to better capture context**, and **employing a dynamically changing masking pattern applied to the training data.**

Detailed Literature Review

RoBERTa Based

- **SARL-RoBERTa** model, which was introduced by Wang et al., uses span-based dependency modeling to align opinion candidates with aspects and uses an adversarial learning strategy to reduce sentiment bias in aspect embeddings.

- Strong performance is achieved by the **ASGCN-RoBERTa** model, which combines an aspect-specific graph convolutional network with dependency tree syntactic information.

- With a relational graph attention network integrated to collect relational information between words, **RGAT-RoBERTa** performs admirably.

- **RoBERTa+MLP** integrates a multi-layer perceptron with RoBERTa, it highlights the flexibility of combining RoBERTa's embeddings with simple classifiers.

Detailed Literature Review

RoBERTa Based

- **SARL-RoBERTa** model, which was introduced by Wang et al., uses span-based dependency modeling to align opinion candidates with aspects and uses an adversarial learning strategy to reduce sentiment bias in aspect embeddings.

- Strong performance is achieved by the **ASGCN-RoBERTa** model, which combines an aspect-specific graph convolutional network with dependency tree syntactic information.

- With a relational graph attention network integrated to collect relational information between words, **RGAT-RoBERTa** performs admirably.

- **RoBERTa+MLP** integrates a multi-layer perceptron with RoBERTa, it highlights the flexibility of combining RoBERTa's embeddings with simple classifiers.

Detailed Literature Review

RoBERTa Based

- **PWCN-RoBERTa** combines a position-weighted convolutional network with RoBERTa to emphasize positional information.

Detailed Literature Review

DeBERTa Based

- **DeBERTa (Decoding-enhanced BERT with disentangled attention)** presents a significant advancement in the field of natural language processing by enhancing the attention mechanisms used in earlier models like BERT and RoBERTa.

- DeBERTa introduces a disentangled attention mechanism. Unlike BERT, where each word is represented by a single vector that combines content and positional embeddings, DeBERTa utilizes two separate vectors for each word to represent its content and position independently.

- While traditional MLM tasks predict masked words based only on the surrounding context, DeBERTa improves this by using absolute positional information along with the con-textual embeddings of word contents and positions before the softmax layer

Detailed Literature Review

DeBERTa Based

- Further, DeBERTa during fine-tuning also uses virtual adversarial training, which is a procedure to boost the generalization power over a wide variety of linguistic tasks through perturbing the input data with small perturbations and retraining the model against those perturbations.

- Improving from the vanilla DeBERTa model a new model named **DeBERTaV3** by He et al. It improves the original DeBERTa model by replacing masked language modeling (MLM) with replaced token detection (RTD), a more sample-efficient pre-training task.

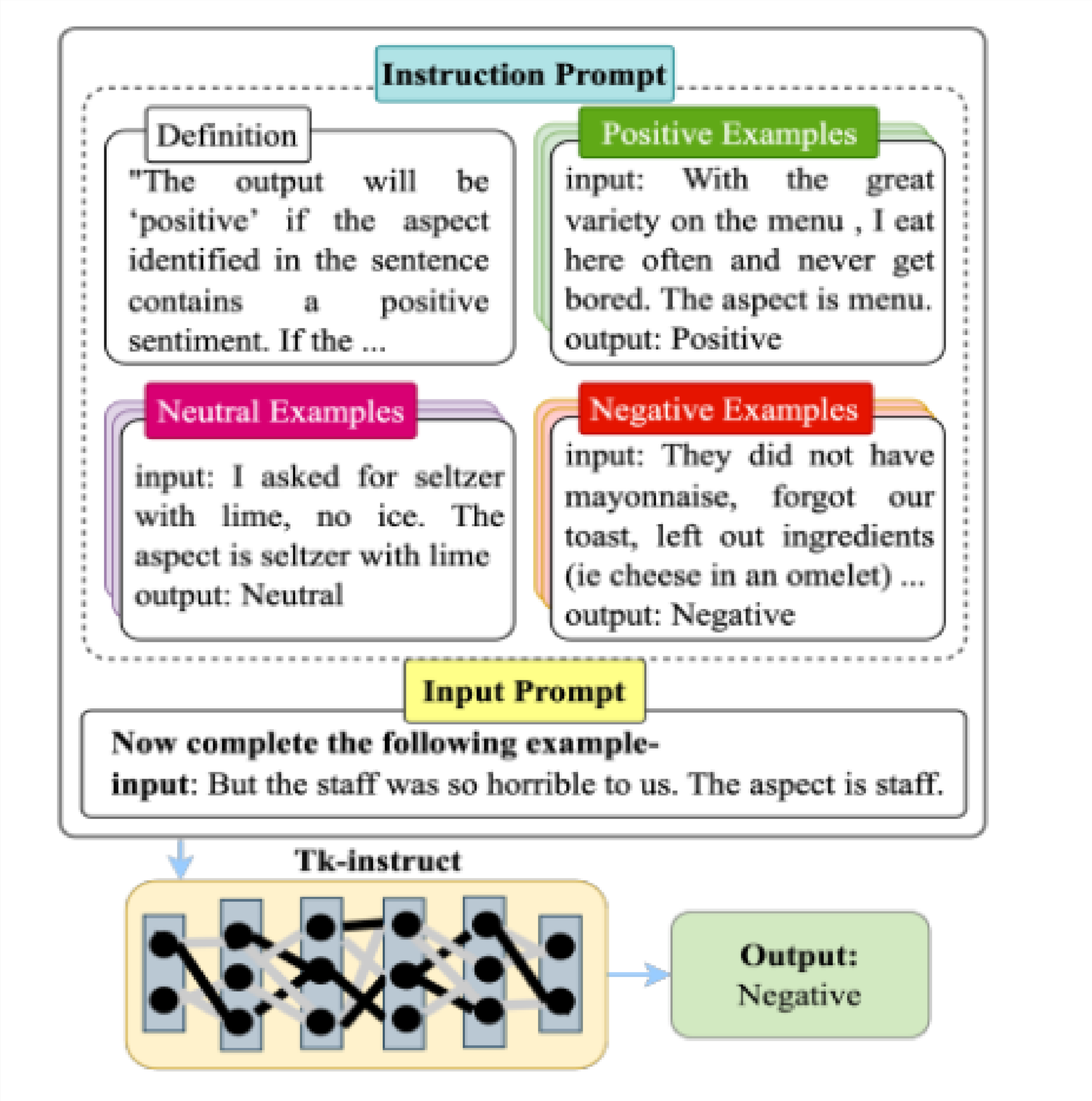
- Neighbouring aspects usually share similar sentiments, which is known as "aspect sentiment coherency." To address this, they proposed a local sentiment aggregation paradigm (LSA) to effectively model fine-grained sentiment coherency. In respect to that, the **LSA-X-DeBERTa** model was introduced.

Detailed Literature Review

Other Models

- **LCF-ATEPC-CDM**, employs a Local Context Focus technique and performs well.

- **InstructABSA , based on T5**



Detailed Literature Review

Other Models

- **BART-ABSA** model in a comprehensive approach to ABSA has also been demonstrated by Yan et al. , which combines all ABSA subtasks into a single generative formulation**.**

- BART is a transformer-based machine learning model that combines both bidirectional and autoregressive techniques. It's primarily designed for sequence-to-sequence tasks and has been pretrained on a large corpus of text data. BART is unique in that it is trained by corrupting text with an arbitrary noising function and learning to reconstruct the original text. This pretraining helps it understand context, grammar, and the nuances of language effectively.

- **RAM,** leveraging multiple attention mechanisms, showed lower performance in sentiment polarity classification, suggesting that while attention mechanisms can capture long-range dependencies, they might not be as effective in sentiment polarity classification compared to other models

Detailed Literature Review

Other Models

- **ASCNN and ASGCN**, proposed by Zhang et al., integrate attention mechanisms and graph convolutional networks to address syntactical constraints and long-range dependencies, showing moderate performance in sentiment polarity classification

- **MCRF-SA**, presented by Xu et al., employs multiple CRFs (Conditional Random Fields)-based structured attention to capture variable-length opinion spans effectively.

- **KaGRMN-DSG**, introduced by Xing and Tsang, combines local syntactic and global relational information to enhance aspect sentiment classification, addressing challenges in aspect representation.

Detailed Literature Review

Other Models

- **MGAN,** proposed by Li et al., utilizes multiple attention networks to differentiate between aspects, showing competitive performance in aspect sentiment classification.

- **PRET+MULT** focuses on transferring knowledge from document-level data, while **MemNet** explicitly captures the importance of each context word for sentiment polarity.

- **IAN,** proposed by Ma et al., uses interactive attention networks to separately model targets and their contexts, demonstrating effectiveness in aspect sentiment classification.

For Literature Review
Joint Task Models

Detailed Literature Review

- **GRACE,** which was proposed by Luo et al. and addresses the imbalance issue in aspect term extraction, obtained reasonable performance in aspect extraction.
- Through a combination of multitasking and relation propagation, **RACL-BERT** considers the connections between these elements, leading to more accurate sentiment analysis.
- **SPAN** introduced by Hu et al. uses another novel approach. Instead of tagging each word in a sentence, it focuses on picking out key opinion points directly.
- **E2E-TBSA** which stands for End-to-End Target Aspect-Based Sentiment Analysis, is a model proposed by Li et al. in 2019 to address two tasks at once: ATE and ASC.

Detailed Literature Review

- **GRACE,** which was proposed by Luo et al. and addresses the imbalance issue in aspect term extraction, obtained reasonable performance in aspect extraction.

- Through a combination of multitasking and relation propagation, **RACL-BERT** considers the connections between these elements, leading to more accurate sentiment analysis.

- **SPAN** introduced by Hu et al. uses another novel approach. Instead of tagging each word in a sentence, it focuses on picking out key opinion points directly.

- **E2E-TBSA** which stands for End-to-End Target Aspect-Based Sentiment Analysis, is a model proposed by Li et al. in 2019 to address two tasks at once: ATE and ASC. Here the model uses a collapsed approach where the model combines ATE and ASC into a single, unified task. The model learns the relationship between identifying aspects and their sentiment simultaneously

Detailed Literature Review

- **BERT-E2E-ABSA** introduced by Li et al. relies on the same concepts as E2E-TBSA and is purely based on BERT models

- **DOER** introduced by Luo et al. uses a cross-shared RNN framework to generate all aspect term-polarity pairs of the input sentence simultaneously.

- **IMN** which was introduced by He et al. uses an interactive architecture with multi-task learning for end-to-end ABSA tasks. It contains aspect term and opinion term extraction besides aspect-level sentiment classification

Our Models

SETFIT Models

- The **BGE** (small) model is a variant of the **Bilingual Google Embeddings** model, optimized for efficiency and reduced computational resources while still providing effective bilingual word embeddings.

- Google/**sentence-t5** (Text-To-Text Transfer Transformer) family of models encode text into high-dimensional vectors that can be used for text classification, semantic similarity, clustering.

- **MPNet** (Masked and Permuted Network) is a transformer-based model introduced by Microsoft. It combines the strengths of Masked Language Modeling (MLM) and Permuted Language Modeling (PLM) to capture both local and global context, leading to better performance in understanding the relationships between words and the overall meaning of sentences.

SETFIT Models

- **SBERT** (Sentence-BERT) is an adaptation of BERT designed to generate efficient and high-quality sentence embeddings. It modifies BERT by adding a pooling layer to produce fixed-size sentence vectors.

- **TinyBERT** is a distillation-based model designed to create a smaller, faster version of BERT while preserving much of its performance, and less layers.
- **ALBERT** (A Lite BERT) is a variant of BERT.

- **SPECTER** (Scalable Embedding for Content-Based Retrieval) is a neural network architecture designed for generating document embeddings.

- **clip-ViT-B-32-multilingual-v1** is a multi-lingual version of the OpenAI CLIP-ViT-B32 model. (VT: Vision Transformer, CLIP: Contrastive Language-Image Pre-training)

SETFIT Models

- **LaBSE** (language-agnostic BERT sentence embedding) encodes text into high dimensional vectors.

- **GTR** (Generalizable T5-based dense Retrievers) models are dual encoders that encode two pieces of text into two dense vectors respectively. This is typically used to encode a query and a document to compute their similarity for dense retrieval.

- **MiniLM** (Miniature Language Model) is a family of small and efficient transformer models.

- **facebook-dpr-ctx_encoder-multiset-base** model is based on the Facebook DPR (Dense Passage Retrieval) framework, specifically using the context encoder component.

STSB stands for the "**Semantic Textual Similarity Benchmark**."

Layered Architecture

Layered Architecture

Inside ATE Model

- **Instruction Input-** instructions or queries regarding the aspect extraction.
- **Tokenization -**The input text and instruction are tokenized into sequences of tokens.
- **Embedding Layer-** Converts tokens into dense vectors (embeddings).
- **Encoder Stack-** Processes the input sequence to generate contextualized embeddings.
- **Decoder Stack-** Generates the output sequence with identified aspects.

Layered Architecture

Inside Encoder and Decoder Stack

- **Self-Attention Mechanism-** This mechanism allows the model to focus on different parts of the input text to capture relationships between words, regardless of their position in the sequence. It helps the model understand the context in which each word appears.
- **Feed-Forward Neural Network-** After the self-attention mechanism, the output is passed through a feed-forward neural network, which applies additional transformations to the data.

Layered Architecture

Inside ASC model

- **DeBERTa Encoder Stack-** Processes the input sequence to generate contextualized embeddings for sentiment classification.
- **Classification Layer-** Assigns sentiment labels (positive, negative, neutral) to each aspect term.

What is a GCP


Cost Effective
Highly Scalable
Custom Machine Types
Internet of Things
Google Cloud Platform
Serverless
Cloud AI
Big Data Analytics
API Platform and Ecosystem
IoT

Fine-Tunning LLM

Fine-Tuning on Benchmark Datasets

- Fine-tuned on **SemEval 14 Laptop** and **Restaurant** data sets
- **Configure parameters** (learning rate, batch size, epochs).
- **Save** fine-tuned models separately in to open-source hubs

Loading and Evaluating Fine-Tuned Models

- **Load** the saved models execute on prepare environment in GCP VM .
- Evaluate on SemEval 14 test data sets