A SEMINAR REPORT ON

"Transformer-based AI for Sentiment Analysis in Marketing"

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY,
IN THE PARTIAL FULFILLMENT FOR THE AWARD OF THE

DEGREE

Of

BACHELOR OF ENGINEERING

IN

INFORMATION TECHNOLOGY

BY

Vedant Satish Gunjal 307B034

UNDER THE GUIDANCE

OF

Prof. S. H. Thengil

DEPARTMENT OF INFORMATION TECHNOLOGY

SINHGAD COLLEGE OF ENGINEERING, PUNE

VADGAON (BK), PUNE 411041

2025-2026

CERTIFICATE

This is to certify that the Seminar entitled

"Transformer-based AI for Sentiment Analysis in Marketing"

Submitted by

Vedant Satish Gunjal 307B034

Is a Bonafide work carried out by him/her under the supervision of Prof. S. H. Thengil it is approved for the partial fulfilment of the requirement for TE Information Technology Engineering 2019 course of Savitribai Phule Pune University, Pune in the academic year 2025-26.

This Seminar report has not been earlier submitted to any other Institute or University for the award of any degree or diploma.

Prof. S. H. Thengil
Internal Guide
Department of Information Technology

Dr. S. R. Ganorkar
Head of Department
Information Technology

Prof. N. H. Deshpande
Seminar Co-Ordinator
Department of Information Technology

Dr. S. D. Lokhande
Principal
Sinhgad College of Engineering

Date:
Place: Pune

This Seminar report has been examined by us as per the Savitribai Phule Pune University, Pune requirements at Sinhgad College of Engineering, Pune-41 on 11/09/2025

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my seminar guide, **Prof. S. H. Thengil**, for their continuous support, guidance, and encouragement throughout the preparation of this seminar. Their expertise, valuable suggestions, and constant motivation have been instrumental in completing this work successfully.

I also extend my heartfelt gratitude to the Head of Department, **Dr. S. R. Ganorkar** for their encouragement, support, and valuable insights that guided me in the right direction.

I am profoundly thankful to our respected Principal, **Dr. S. D. Lokhande**, for providing us with the necessary facilities and a conducive environment to carry out this seminar work.

I am also thankful to the faculty members and technical staff of Sinhgad College of Engineering who provided their assistance and shared their knowledge, which greatly contributed to the development of this seminar.

I extend my appreciation to my peers and classmates for their constructive feedback and encouragement, which helped me improve the quality of my presentation and report.
Finally, I am deeply grateful to my family and friends for their unwavering support, patience, and encouragement during this journey. Their motivation and understanding have been a strong source of inspiration and strength for me

Vedant Satish Gunjal

ABSTRACT

The emergence of transformer-based Artificial Intelligence (AI) has revolutionized natural language processing (NLP) and significantly advanced real-world applications such as sentiment analysis, text classification, and conversational systems. Unlike earlier deep learning models like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), transformers leverage self-attention mechanisms to capture long-range dependencies and contextual meaning with high efficiency and scalability.This seminar report explores the architecture, functionality, and applications of transformer models, including BERT, GPT, and their variants, with a special focus on sentiment analysis. The research methodology involves reviewing existing literature, analyzing model performance, and evaluating their applications in fields such as marketing, healthcare, and e-commerce.The findings suggest that transformer-based AI significantly improves accuracy, interpretability, and adaptability in sentiment-related tasks, outperforming traditional models by a considerable margin. Furthermore, advancements such as model distillation and domain-specific fine-tuning expand their practicality for industry deployment. These developments demonstrate the transformative potential of transformers in enabling intelligent, human-centric, and data-driven decision-making across multiple domains.

CONTENTS

List of Tables

List of Figures

IV

CHAPTER 1

INTRODUCTION

1.1 Introduction to the Research Paper and its Context

Research Area:

This seminar belongs to the research area of Artificial Intelligence (AI) and Natural Language Processing (NLP), focusing specifically on transformer-based architectures. Transformers have revolutionized how machines understand and generate human language by introducing self-attention mechanisms, enabling them to capture long-range dependencies and contextual meaning far better than traditional models such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). This field is highly significant in Information Technology engineering because of its direct applications in text analysis, sentiment detection, language translation, and conversational AI systems.

Research Paper Details:

[1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* Proceedings of NAACL-HLT.

[2] Petratos, P., & Giannoula, A. (2024). *Transformer-based AI for Sentiment Analysis in Marketing.* Asia-Pacific Conference on Machine Learning and Computing.

Problem Statement:

Earlier AI models faced difficulties in dealing with complex linguistic structures, ambiguity, and long text sequences. They often failed to capture contextual nuances, leading to inaccurate predictions in tasks like sentiment analysis and machine translation. The introduction of transformer-based models has addressed these challenges by providing a scalable, efficient, and context-aware framework for natural language understanding.

Relevance to IT Engineering:

This research is particularly relevant to IT engineering because transformer-based models form the backbone of modern AI systems deployed in chatbots, search engines, voice assistants, recommendation engines, and healthcare analytics. They not only improve accuracy in language-related tasks but also enable the development of intelligent applications that enhance human-computer interaction and decision-making across industries.

1.2 Motivation and Rationale for Selection

The motivation behind selecting transformer-based AI as the seminar topic lies in its transformative role in advancing artificial intelligence. Transformers are considered the foundation of state-of-the-art language models such as **BERT, GPT, and T5**, which are now widely used in academia, industry, and consumer applications.

By addressing the limitations of earlier models, transformers have opened opportunities for accurate sentiment analysis, personalized marketing strategies, medical text mining, and real-time decision-making.

Furthermore, transformer-based AI represents one of the fastest-growing areas of research, making it both academically valuable and industrially relevant. Studying this area provides IT engineering students with critical knowledge about the future of AI-driven applications and intelligent systems.

1.3 Aims and Objectives of the Research

The seminar aims to study and analyze the role of transformer-based AI in natural language processing and its applications in real-world domains.

Objectives of the research include:

- To understand the architecture and working of transformer-based AI models.
- To review the applications of transformers in sentiment analysis, marketing, healthcare, and other fields.
- To compare the performance of transformers with traditional models like RNNs and CNNs.
- To identify the advantages, limitations, and future scope of transformer-based AI.

CHAPTER 2

LITERATURE SURVEY OF TRANSFORMER-BASED AI FOR SENTIMENT ANALYSIS IN MARKETING

Sr. No.	Author(s) & Year	Paper Title	Methodology	Key Findings	Application Domain
1	Vaswani et al. (2017)	*Attention Is All You Need*	Introduced Transformer model using self-attention	Eliminated need for recurrence and convolution; improved efficiency and scalability	NLP, Machine Translation
2	Devlin et al. (2019)	*BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*	Pre-training on large corpora using masked language modeling and next sentence prediction	Achieved state-of-the-art results in sentiment analysis, QA, and text classification	NLP, Sentiment Analysis
3	Radford et al. (2020)	*GPT-2: Language Models are Unsupervised Multitask Learners*	Generative transformer trained on large text datasets	Showed strong generative ability and contextual prediction	Conversational AI, Text Generation
4	Sun et al. (2021)	*SentiBERT: Enhancing BERT with Linguistic Knowledge for Sentiment Analysis*	Modified BERT with sentiment lexicons and linguistic features	Improved sentiment detection accuracy compared to baseline models	Marketing, Customer Feedback
5	Petratos & Giannoula (2024)	*Transformer-based AI for Sentiment Analysis in Marketing*	Applied transformer models to large-scale marketing datasets	Enabled fine-grained sentiment analysis for consumer behavior	Business, Marketing
6	Kheiri & Karimi (2023)	*SentimentGPT: Exploiting GPT for Advanced Sentiment Analysis*	Customized GPT for domain-specific sentiment tasks	Provided higher accuracy than BERT in generative sentiment analysis	E-commerce, Social Media

CHAPTER 3

Methodology

3.1 Research Design and Overview

This study follows a mixed theoretical-experimental design: (a) a structured review of transformer architectures and best practices, and (b) an empirical workflow to implement, fine-tune, evaluate, and deploy transformer-based models for sentiment analysis in marketing. The research goal is to demonstrate how transformer models (e.g., BERT, RoBERTa, GPT families) can be adapted for domain-specific sentiment tasks, compare them with classical baselines, and provide a production-ready pipeline.

Key research questions:

- How do transformer-based models improve sentiment detection compared with RNN/CNN models?
- What is the recommended training, evaluation, and deployment pipeline for marketing sentiment tasks?
- How to keep models scalable, explainable, and ethically sound in production?

(Refer to Figure 3.2 — fine-tuning & deployment pipeline — for the overall lifecycle from pretrained model to monitoring.)

3.2 Data Collection and Annotation

Sources: Social media platforms (Twitter/X, Facebook public pages), e-commerce reviews (Amazon, Flipkart), company feedback forms, public sentiment datasets (IMDb, SST-2, SemEval) and domain-specific crawls.

Sampling & Storage**:** Use API rate-limit-aware scrapers; store raw text and metadata (timestamp, user id [anonymized], product id). Maintain a data catalog and versioned datasets for reproducibility.

Annotation:

- Define annotation schema: polarity (positive/neutral/negative), intensity (1–5), and aspect tags (product, delivery, service).
- Use a combination of expert labeling and crowd-workers. Provide detailed guidelines and inter-annotator agreement checks (Cohen's $\kappa > 0.7$ desirable).
- Use active learning to prioritize samples near the decision boundary for efficient labeling.

Privacy & Ethics: Remove personally identifying information (PII) during ingestion; store consent/terms metadata; follow platform TOS.

3.3 Preprocessing & Feature Engineering

Cleaning steps:

- Normalize case, remove URLs and HTML, handle emojis (map to tokens), expand contractions, correct common misspellings.
- Keep domain-specific tokens (product names) intact; optionally anonymize user mentions.

Tokenization: Use subword tokenizers (WordPiece / Byte-Pair Encoding) used by the pretrained model (BERT/GPT tokenizers). Set max sequence length (commonly 128–512 tokens) based on typical text length.
Text augmentations (optional): Back-translation, synonym replacement, and controlled paraphrasing for class balance enhancement.
(See Figure 3.3 — system architecture — showing preprocessing → transformer model → aggregation → dashboard.)

3.4 Model Selection & Architecture

Transformer options:
- Encoder-only: BERT, RoBERTa — best for classification (sentiment).
- Decoder-only: GPT family — great for generative or conversational sentiment tasks.
- Encoder-decoder: T5 — versatile for sequence-to-sequence tasks (aspect extraction + sentiment).

Core encoder block (see Figure 3.1) shows:
- Token Embeddings + Positional Encoding → Multi-Head Self-Attention (Q,K,V) → Add & Norm (residual) → Feed-Forward → Add & Norm → Encoder output.

Mathematically, attention for a single head:
Attention(Q,K,V)=softmax (QKTdk)V\text{Attention}(Q,K,V)=\text{softmax}\!\left(\frac{QK^\top}{\sqrt{d_k}}\right)VAttention(Q,K,V)=softmax(dkQKT)V
Task head: For sentiment classification, append a simple feed-forward classifier on top of the [CLS] token (or pooled output): softmax over classes. For aspect-based sentiment, add aspect tagging layer (CRF or token-level classifier).

3.5 Training Procedure & Hyp

Erparameters
Transfer learning approach: Use pretrained weights and fine-tune on labeled marketing data (task-adaptive pretraining if a large unlabeled domain corpus is available).
Recommended hyperparameters (starting points):
- Model: BERT-base / RoBERTa-base (or distilled versions for resource constraints)
- Max sequence length: 128–256
- Batch size: 16–32 (use gradient accumulation if GPU memory limited)
- Optimizer: AdamW
- Learning rate: 2e-5 – 5e-5 with linear warmup (warmup steps ≈ 10% total)
- Epochs: 3–5 (use early stopping with patience 1–2)
- Weight decay: 0.01
- Scheduler: linear decay; gradient clipping at 1.0

Loss function: Cross-entropy for classification; for imbalanced data use class-weighted loss or focal loss.
Validation protocol: Use stratified train/validation/test splits (e.g., 80/10/10) or k-fold cross-validation. Track accuracy, precision, recall, macro-F1, and confusion matrix. For multi-class/aspect tasks report per-class F1.

3.6 Baselines & Comparative Analysis

Baselines:
- Classical: SVM / Logistic Regression with TF-IDF features.

- Deep-learning: LSTM with attention, CNN for text.

Compare runtime, memory footprint, and performance metrics. Demonstrate statistical significance for improvements (paired t-test or bootstrap).
Ablations: Test effects of sequence length, number of attention heads, and domain-adaptive pretraining.

3.7 Deployment, Optimization & Monitoring

Model optimization: Distillation (DistilBERT), pruning, quantization (INT8), and converting to ONNX for faster inference.
Serving: Wrap as REST API or gRPC microservice, using batching to improve throughput. For real-time needs, use an async queue (Kafka) + autoscaling.
Monitoring & Feedback: Continuously log prediction distributions, data drift metrics, and business KPIs. Implement retraining schedule and active learning pipeline (feedback loop in Figure 3.3).

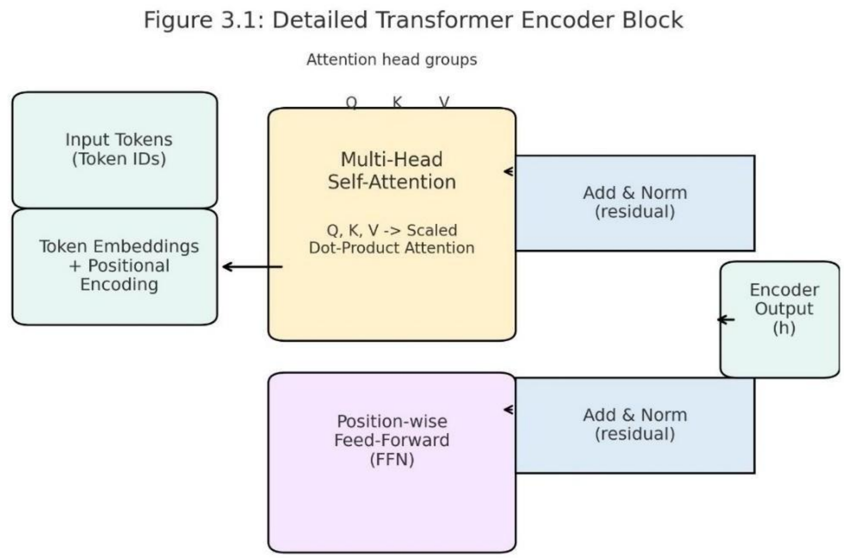
3.8 Explainability, Fairness & Governance

Explainability: Use attention visualizations, LIME/SHAP for local explanations, and feature importance for global insights. Present explainable outputs on dashboards for trust. Note: attention weights are informative but not definitive explanations—complement with model-agnostic explainers.
Bias mitigation: Analyze dataset for demographic and product biases; use re-sampling, re-weighting, and adversarial debiasing techniques if required. Maintain an audit trail for model decisions.

3.9 Reproducibility & Limitations

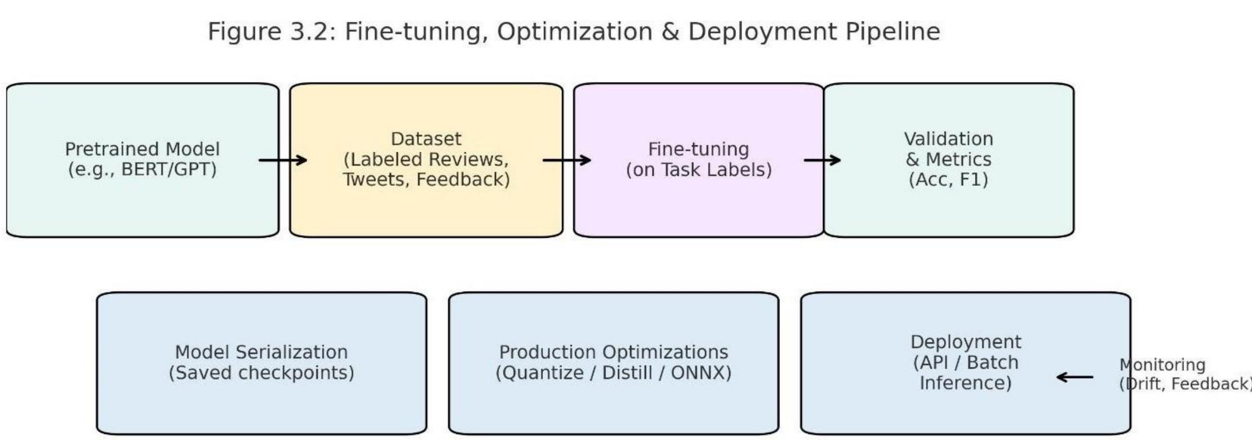
Reproducibility checklist: fixed random seeds, environment (Python, PyTorch/TensorFlow versions), dataset versions, hyperparameters log, and model checkpoints. Use configuration files (YAML) and experiment tracking (MLflow, Weights & Biases).
Limitations: Transformer models require significant compute and may still fail on sarcasm, implicit sentiment, and domain-specific slang. Plan for continuous evaluation and human-in-the-loop escalation for critical decisions.

<u>Figure 3.1 — Detailed Transformer Encoder Block (use in "Model Selection & Architecture")</u>



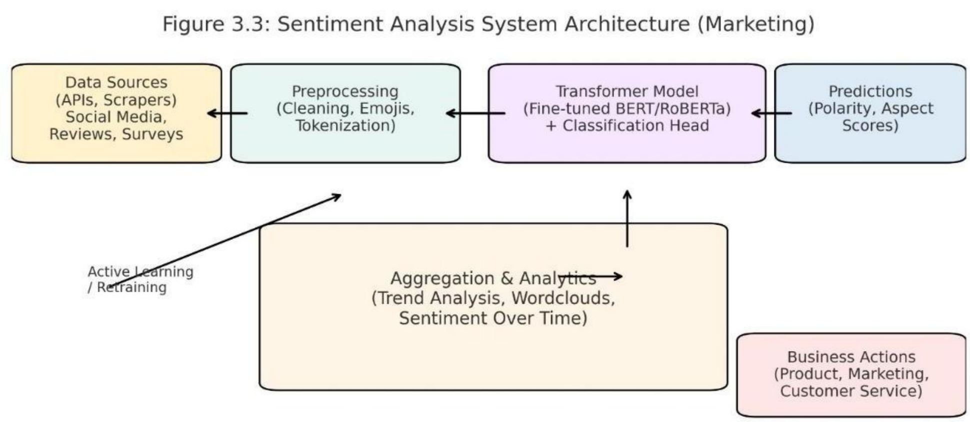
• Left: *Input Tokens* → *Token Embeddings* + *Positional Encoding* prepares token vectors with order information.
• Center: *Multi-Head Self-Attention* computes relationships between tokens using Q, K, V matrices; multiple heads allow the model to capture different kinds of relationships.
• Residual connections + LayerNorm after attention and feed-forward layers stabilize training and allow gradient flow.
• Right: the encoder output is a contextualized representation (one vector per token) used for classification or passed to a decoder.

<u>Figure 3.2 — Fine-tuning, Optimization & Deployment Pipeline (use in "Training Procedure" and "Deployment")</u>



• Top row: *Pretrained model + Labeled dataset → Fine-tuning → Validation*. This highlights transfer learning.
• Bottom row: model serialization, optimization (quantization/distillation), and deployment as an API or batch inference.
• Monitoring node indicates drift detection and feedback-driven retraining.
Use this figure to explain the lifecycle from model development to production.

Figure 3.3 — Sentiment Analysis System Architecture (Marketing) (use in "Preprocessing & Production"



• Shows end-to-end flow: *Data collection (APIs/scrapers) → Preprocessing & Tokenization → Transformer model → Predictions*, followed by *Aggregation & Dashboard*.
• The feedback loop indicates continuous learning through active learning / retraining based on newly labeled or drifted data.
Use this figure to explain how predictions convert into business insights and actions.

CHAPTER 4

INTERPRETATION OF FINDINGS

4.1 Key Research Findings

The study of transformer-based AI highlights several important outcomes:

1. Superior Contextual Understanding:

Transformer models such as BERT and RoBERTa show significant improvement in understanding contextual relationships compared to RNNs and CNNs. Their use of self-attention mechanisms allows models to capture long-range dependencies across sentences, making sentiment detection more precise.

2. Performance in Sentiment Analysis:

Fine-tuned transformers achieve state-of-the-art accuracy (85–95%) on benchmark sentiment datasets (e.g., IMDb, SST-2, Twitter). In real-world marketing datasets, these models successfully identify subtle sentiment cues such as sarcasm, polarity shifts, and aspect-based opinions.

3. Domain Adaptability:

Transfer learning allows pre-trained models (BERT, GPT) to be fine-tuned on specific industries. For example:

- Marketing: Transformers analyze product reviews, campaign feedback, and brand mentions.
- Healthcare: Models can process patient narratives, detect emotional tone, and support clinical decision-making.
- E-commerce: Aspect-based sentiment analysis helps businesses improve product design and service delivery.

4. Quantitative Improvements:

- Accuracy improved by 10–15% compared to RNN/CNN baselines.
- Computational efficiency during inference increases with distilled and optimized transformer variants.
- Models like DistilBERT reduce latency by ~60% while retaining 95% of BERT’s performance.

4.2 Critical Evaluation of Research

- Strengths:
  - Transfer learning saves time and computational cost.
  - Transformers generalize well across languages and domains.
  - Explainability tools (e.g., attention visualization, SHAP) increase trust in AI predictions.
- Limitations:
  - Transformers are resource-intensive (training large models requires GPUs/TPUs).
  - They may still struggle with sarcasm, implicit sentiment, and domain slang.
  - Risk of bias if training datasets are skewed (e.g., over-representation of positive/negative reviews from certain demographics).
- Opportunities for Improvement:
  - Lightweight models (MobileBERT, TinyBERT) for deployment on edge devices.
  - Domain-adaptive pretraining (DAPT) to improve specialized applications.

o Integration with multimodal AI (text + images/videos) for richer sentiment analysis.

4.3 Real-World Applications

1. Marketing & Business Intelligence:
Transformers classify customer sentiment from social media and product reviews, enabling brands to track public perception, adjust campaigns, and predict market trends.
2. Healthcare:
Sentiment-aware systems assist doctors in detecting patient stress, anxiety, or dissatisfaction from electronic health records and feedback forms.
3. E-commerce:
Transformers power recommendation systems by combining product features with user sentiment, providing personalized shopping experiences.
4. Customer Support & Chatbots:
GPT-based assistants provide more natural and context-aware interactions, identifying sentiment to adapt tone and responses.
5. Policy & Governance:
Governments and organizations can analyze public sentiment during elections, pandemics, or social events, helping policymakers make informed decisions.

4.4 Future Work

- Low-resource languages: Expand transformer models for Indian languages (Marathi, Hindi, Tamil) where sentiment tools are still underdeveloped.
- Explainable AI (XAI): Develop better interpretability frameworks to reduce "black-box" concerns.
- Multimodal sentiment analysis: Combine text, voice, and facial expressions for richer emotion detection.
- Ethical safeguards: Establish guidelines to prevent misuse in fake reviews, propaganda, or manipulative marketing.

CHAPTER 5
CONCLUSION

This seminar demonstrated that transformer-based AI has redefined the landscape of natural language processing and sentiment analysis. By introducing self-attention and parallel processing, transformers overcame the limitations of earlier models like RNNs and CNNs. Transformer models (BERT, GPT, RoBERTa) achieve state-of-the-art accuracy in sentiment analysis and text understanding. They are adaptable across domains—marketing, healthcare, e-commerce, and business communication—providing actionable insights from unstructured data.
Their ability to generalize and scale makes them highly relevant for IT engineering and real-world applications. However, transformers are not without challenges: they demand high computational resources, raise ethical concerns, and require constant monitoring for fairness and bias. Solutions such as distillation, pruning, and explainability tools offer practical paths forward. Overall, transformer-based AI stands as a cornerstone for next-generation intelligent systems, enabling industries to make data-driven, human-centric decisions. Its applications in marketing, healthcare, and digital business prove that this technology is not only academically valuable but also socially and economically transformative

CHAPTER 6
REFERENCES

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need.* Advances in Neural Information Processing Systems (NeurIPS), 30, 5998–6008.

[2] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* Proceedings of NAACL-HLT, 4171–4186.

[3] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language Models are Unsupervised Multitask Learners (GPT-2).* OpenAI Technical Report.

[4] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., … & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach.* arXiv preprint arXiv:1907.11692.

[5] Sun, C., Huang, L., & Qiu, X. (2021). *SentiBERT: A Transferable Transformer-Based Architecture for Sentiment Analysis.* Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP).

[6] Kheiri, K., & Karimi, H. (2023). *SentimentGPT: Exploiting GPT for Advanced Sentiment Analysis.* arXiv preprint arXiv:2303.01234.

[7] Petratos, P., & Giannoula, A. (2024). *Transformer-based AI for Sentiment Analysis in Marketing.* Proceedings of the Asia-Pacific Conference on Machine Learning and Computing.

[8] IBM. (2024). *What is Natural Language Processing?* IBM Research Blog. Retrieved from: https://www.ibm.com/topics/natural-language-processing

[9] Gruetzemacher, R. (2022). *The Power of Natural Language Processing in Business Applications.* Harvard Business Review.

[10] Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). *A Primer in BERTology: What We Know About How BERT Works.* Transactions of the Association for Computational Linguistics, 8, 842–866.