Sentiment Analysis: From Traditional Methods to Advanced Deep Learning Approaches

Jatin Sharma

April 2025

Abstract

This technical report provides a comprehensive overview of sentiment analysis techniques, tracing their evolution from traditional machine learning approaches to state-of-the-art deep learning models. We explore the foundations of sentiment analysis, examine aspect-based sentiment analysis for fine-grained opinion mining, and investigate the transformative impact of BERT and other transformer-based architectures. Through experimental implementation on benchmark datasets, we demonstrate the superior performance of transformer-based models, achieving accuracy rates of up to 92% on movie review sentiment classification tasks. Our research reveals that specific attention heads contribute disproportionately to model performance, suggesting opportunities for targeted model optimization. The report concludes with promising future research directions in this rapidly evolving field.

1 Introduction

Sentiment analysis, a key application of natural language processing (NLP), involves determining the emotional tone or attitude expressed in text. As digital communication platforms proliferate, understanding sentiment has become crucial for businesses, researchers, and organizations seeking to extract meaningful insights from user-generated content. This field has seen remarkable evolution, transitioning from simple rule-based systems to sophisticated deep learning architectures.

1.1 Definition of Sentiment Analysis

Sentiment analysis is the process of computationally identifying and categorizing opinions expressed in text to determine the writer's attitude toward a particular topic, product, or service [**?**]. It is fundamentally a classification problem where text is labeled as positive, negative, or neutral, with challenges in recognizing contextual nuances, sarcasm, and cultural references.

1.2 Types of Sentiment Analysis

The field encompasses several specialized approaches:

- **Fine-grained Sentiment Analysis**: Assigns precise sentiment scores, such as very positive or very negative, valuable for detailed feedback in product reviews [**?**].
- **Aspect-based Sentiment Analysis (ABSA)**: Identifies sentiment for specific aspects, e.g., praising a camera but criticizing battery life in a product review [Sen, 2025].
- **Emotion Detection**: Identifies emotions like happiness or anger, used in social media monitoring and mental health analysis [**?**].

1.3 Applications and Importance

Sentiment analysis is widely applied in:

- Business Intelligence: Analyzing customer feedback to improve products.
- Market Research: Gauging public reception to campaigns or brands.
- Social Media Monitoring: Tracking brand sentiment to address issues.
- Political Analysis: Refining campaign messaging based on voter sentiment.
- Financial Markets: Predicting market movements from news and social media sentiment [**?**].

2 Evolution of Approaches

Sentiment analysis methodologies reflect advancements in AI and machine learning:

- **Lexicon-based Methods**: Used predefined sentiment dictionaries.
- **Traditional Machine Learning**: Employed algorithms like SVM and Naive Bayes.
- **Deep Learning**: Utilized RNNs and LSTMs to capture sequential patterns.
- **Transformer-based Models**: Leveraged BERT and attention mechanisms for bidirectional context [Sen, 2025, **?**].

3 Traditional Approaches

Before deep learning, sentiment analysis relied on lexicon-based and machine learning methods.

3.1 Lexicon-Based Methods

These methods use dictionaries of words with sentiment scores to compute overall text sentiment.

3.1.1 Dictionary-Based Approaches

Resources like SentiWordNet and VADER provide sentiment scores for words, used to calculate document sentiment.

3.1.2 Corpus-Based Approaches

These derive sentiment lexicons from domain-specific corpora, addressing context-dependent sentiment variations.

3.1.3 Limitations

Challenges include inability to capture context, negation, sarcasm, and domain-specific terms.

3.2 Machine Learning Methods

These treat sentiment analysis as a classification problem, relying on labeled data.

3.2.1 Feature Engineering

Common features include Bag-of-Words, N-grams, TF-IDF, POS tags, and syntactic relationships.

3.2.2 Commonly Used Algorithms

Algorithms like Naive Bayes, SVM, KNN, and Random Forests were widely used.

3.2.3 Performance Analysis

Table 1 shows SVM outperforming other methods [**?**].

Table 1: Performance of Traditional ML Methods on Sentiment Analysis

Method	Accuracy (%)	F1-Score (%)
KNN	38.85	34.33
SVM	57.30	53.66

3.2.4 Limitations

These methods struggle with semantic understanding, context sensitivity, and require extensive feature engineering.

4 Aspect-Based Sentiment Analysis

ABSA provides granular sentiment analysis by identifying sentiments for specific aspects.

4.1 Definition and Components

ABSA identifies aspects and their sentiments, e.g., positive for food but negative for service in a review [Sen, 2025]. It involves:

- Aspect Term Extraction: Identifying aspect terms, similar to NER.
- Aspect Sentiment Classification: Determining sentiment for each aspect.

4.2 Approaches to ABSA

ABSA has evolved from rule-based to deep learning methods:

- **Traditional**: Rule-based, lexicon-based, and topic modeling.
- **Machine Learning**: CRFs and SVMs for extraction and classification.
- **Deep Learning**: ATT-LSTM, ASGCN, and BERT-based models [ove, 2025].

4.3 Proposed Syntax-Guided Methodology

Our novel approach combines syntactic and semantic representations:

- Syntax Representation Encoder: Processes POS tags and token distances.
- Semantic Representation Encoder: Uses BERT for contextual embeddings.
- Syntax-Guided Transformer: Combines syntax and semantics for attention.
- Joint Learning Framework: Integrates representations for classification [Sen, 2025].

4.4 Experimental Results

Table 2 shows our method outperforming baselines [Sen, 2025].

Table 2: Performance Comparison of ABSA Methods

	Restaurant Dataset		Laptop Dataset	
Method	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)
ATT-CNN	68.19	60.35	62.11	58.64
ATT-LSTM	76.60	69.34	68.90	65.26
ATAE-LSTM	77.20	69.86	68.70	65.43
ASGCN	80.77	72.02	75.55	71.05
BERT-BASE	85.97	81.72	79.72	76.91
Proposed Method	86.21	83.22	82.29	78.68

4.5 Applications

ABSA supports product development, competitive analysis, customer support, recommendation systems, and brand monitoring.

5 BERT in Sentiment Analysis

BERT, introduced by **?**, revolutionized NLP with bidirectional context.

5.1 Understanding BERT Architecture

BERT uses transformer layers (12 for BERT-base, 24 for BERT-large) with self-attention mechanisms.

5.2 Self-Attention Mechanism

Attention is computed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

where Q, K, and V are query, key, and value projections [**?**].

5.3 Pre-training and Fine-tuning

BERT is pre-trained on Masked Language Model and Next Sentence Prediction tasks, then fine-tuned for sentiment tasks [Fin, 2025].

5.4 BERT vs. Traditional Models

BERT excels in contextual understanding, bidirectional processing, and reduced feature engineering.

5.5 Performance Analysis

Table 3 shows BERT outperforming traditional models [**?**].

Table 3: Performance Comparison of Deep Learning and Transformer Models

Category	Method	Accuracy (%)
Deep Learning	FNN	56.62
	RNN	57.13
	LSTM	58.38
Transformer-based	BERT	78.83
	BERT-SVM	67.74
	RoBERTa	79.40

6 Implementation and Experiments

We validated transformer models on benchmark datasets.

6.1 Dataset Description

- **Stanford Movie Review Dataset**: 50,000 reviews, split 80% training, 20% test [Fin, 2025].
- **IMDB Dataset Sample**: 99 reviews for attention analysis [imd].

6.2 Model Architecture and Implementation

We fine-tuned BERT-base and implemented GPT-Neo for few-shot learning [Fin, 2025, imd].

6.3 Experimental Results

BERT achieved 92% accuracy on the Stanford dataset (Table 4) [Fin, 2025].

Table 4: Performance Metrics on Stanford Movie Review Dataset

Metric	Value (%)
Accuracy	92.00
Precision	92.10
Recall	91.90
F1-Score	92.00

6.3.1 Attention Head Analysis

Heads 5, 14, and 3 were critical for sentiment classification (Table 5) [imd].

Table 5: Operational Contribution of Top Attention Heads

Attention Head	Contribution Score
Head 5	0.1515
Head 14	0.1313
Head 3	0.1010
Others	< 0.0900

6.3.2 Error Analysis

Challenges included sarcasm, subtle sentiment, mixed sentiment, and domain-specific terms [Fin, 2025].

7 Future Directions

Promising areas include:

- **ABSA Enhancements**: Implicit aspect extraction, cross-domain ABSA [ove, 2025].
- **Multimodal Analysis**: Integrating text, visual, and audio cues.
- **Temporal Analysis**: Tracking sentiment trends over time.
- **Explainable Analysis**: Visualizing attention and feature attribution [Fin, 2025].
- **Robustness and Fairness**: Addressing biases and adversarial robustness [arx, 2025].
- **Efficient Deployment**: Model distillation and attention head pruning [imd].

8 Conclusion

Transformer-based models like BERT achieve up to 92% accuracy, with our syntax-guided ABSA approach outperforming baselines. Specific attention heads drive performance, suggesting optimization opportunities. Future work will enhance multimodal, explainable, and robust sentiment analysis systems.

References

imdb-1.ipynb. `https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/29587170/087f176e-fc12-410c-b6fc-d3d28a72ea4c/imdb-1.ipynb`.

Finetuning-pretrained-trasformer-for-absa. `https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/29587170/344de19c-e6944c58-b558-c860210315da/mireport-1.pdf`, 2025.

Sentiment-analysis. `https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/29587170/ad2e96bf-7e0245d0-833e-1997e1a0e8e4/MS-Thesis-Stage-I-Mid-Sem-Evaluation.pdf`, 2025.

arxiv submission guidelines. `https://info.arxiv.org/help/submit/index.html`, 2025.

Overleaf gallery: arxiv. `https://www.overleaf.com/gallery/tagged/arxiv`, 2025.