SENTIMENT ANALYSIS FOR CUSTOMER FEEDBACK USING BERT

CASE STUDY REPORT

Submitted by

SANJAY S S [RA2211003010014]

Under the Guidance of

Dr. Viveka S
Assistant Professor
Department of Computing Technologies
SRM Institute of Science and Technology

in partial fulfillment of the requirements for the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE ENGINEERING

Department Of Computing Technologies
COLLEGE OF ENGINEERING AND TECHNOLOGY
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR- 603203

NOVEMBER 2025

INTRODUCTION

Sentiment analysis is among the most widely adopted NLP tasks in modern business environments, with the market expected to expand sharply as organizations increasingly leverage customer feedback data for strategic decision-making. This technique automates the classification of textual content into emotional categories positive, negative, or neutral providing critical insights that drive business intelligence and improve customer satisfaction.

The rise of digital platforms has turbocharged the volume of customer-generated content, including reviews, social media mentions, and feedback forms processed daily by companies. For instance, Netflix analyzes over 30,000 brand mentions each day to assess customer sentiment, using these insights to tailor content strategy and improve viewer engagement. BERT's bidirectional transformer architecture enables a richer contextual understanding of sentiment, outperforming traditional models by analyzing both left and right text contexts simultaneously. This leads to markedly higher accuracy and nuanced emotion detection critical for today's diverse textual data.

This case study examines sentiment analysis as a fundamental application of Natural Language Processing (NLP), emphasizing its use in customer feedback analysis. The study focuses on how BERT (Bidirectional Encoder Representations from Transformers) serves as the foundational Large Language Model that significantly advances sentiment classification tasks. We analyze the technical implementation of BERT for this purpose, demonstrate its superior performance using detailed examples, and discuss the practical impact in real-world scenarios such as content recommendation and customer experience enhancement.

KEY FEATURES OF SENTIMENT ANALYSIS USING BERT

Sentiment Analysis using BERT model has various features which makes it superior, those are:

- Bidirectional Context Understanding: BERT analyzes both the left and right context of words simultaneously, enabling a more nuanced understanding of sentiment expressions compared to traditional unidirectional models.
- Pre-training with Masked Language Modeling: BERT is pre-trained by predicting masked words in sentences, which helps the model learn complex contextual relationships necessary for accurate sentiment detection.
- Transformer-based Architecture: BERT employs multiple layers of self-attention mechanisms in a transformer encoder structure, providing deep contextual encoding for input text.
- Fine-tuning for Sentiment Classification: A classification head is added to the pre-trained BERT model, typically consisting of a dropout layer followed by a linear layer, to output probabilities for sentiment classes (positive, negative, neutral).
- Multi-language Support: The model can be fine-tuned or adapted to support multiple languages, enabling global customer sentiment analysis.
- Real-time Processing: The system supports real-time sentiment analysis of customer input from various channels including social media, reviews, and feedback forms.
- Integration with Business Systems: Sentiment scores can feed into recommendation engines, customer service workflows, and content strategy for actionable business intelligence.

LARGE LANGUAGE MODEL SELECTION

For this case study on sentiment analysis for customer feedback, BERT (Bidirectional Encoder Representations from Transformers) was chosen as the primary large language model due to several compelling factors that distinguish it from other models.

Bidirectional Context Understanding

Unlike earlier NLP models that process text sequentially in a single direction, BERT's transformer architecture reads text bidirectionally meaning it simultaneously considers the context from both left and right sides of a word or phrase. This bidirectional attention mechanism enables a more nuanced and precise understanding of sentiment expressions, especially in complex sentences where sentiment depends on surrounding context. For instance, negations and irony that change sentiment meaning are better captured due to this contextual awareness.

Pre-training Methodology

BERT's distinctive pre-training approach involves masked language modeling (MLM), where the model learns to predict randomly masked words within sentences based on their surrounding context. This forces the model to develop deep contextual relationships and semantic understanding essential for sentiment analysis, where emotion and opinion are closely tied to language subtleties. The MLM strategy is a key factor behind BERT's effectiveness in grasping sentiment nuances.

Proven Performance

BERT-based models consistently outperform traditional sentiment classifiers across benchmarks with reported accuracy rates falling between 88% and 94%. These models deliver superior precision and recall compared to RNNs, CNNs, and rule-based systems, establishing BERT as a state-of-the-art solution for diverse sentiment analysis tasks.

ALTERNATIVE LLM CONSIDERATIONS

While BERT is the foundation model in this case, the NLP landscape includes several alternative transformers, each with unique advantages:

GPT-4
Known for its generative capabilities, GPT-4 excels in capturing emotional subtleties, particularly in free-text or open-ended survey responses where sentiments are not categorized but expressed in nuanced language. GPT-4's strong language generation ability complements sentiment classification, often outperforming BERT in understanding subtle emotional cues.

RoBERTa
An optimized variant of BERT, RoBERTa removes the Next Sentence Prediction (NSP) pre-training task and employs larger dynamic masking and bigger training datasets for longer durations. These changes help RoBERTa capture sentence context more effectively, leading to marginal yet consistent improvements over BERT in sentiment classification metrics such as accuracy, precision, and recall. However, RoBERTa demands more computational resources and training time.

DistilBERT
DistilBERT is a smaller, faster version of BERT, designed for computational efficiency. It retains approximately 97% of BERT's performance while being 60% smaller in size, making it ideal for deployment in resource-constrained environments such as mobile or edge devices where inference speed and memory footprint are critical considerations.

MODEL DESCRIPTION

BERT Architecture for Sentiment Analysis

BERT utilizes a sophisticated transformer-based encoder architecture that underpins its success in numerous NLP tasks, including sentiment analysis. The core of BERT's architecture consists of multiple self-attention layers designed to compute contextual embeddings by considering relationships between all tokens in an input sequence simultaneously.

Base Architecture

The typically employed version for sentiment analysis is **BERT-Base**, which includes:

- 12 transformer layers (encoder blocks): Each layer includes self-attention heads and feed-forward networks.
- 768 hidden units per layer: These are the dimensionalities of the token embeddings and internal representations.
- 12 attention heads: Multiple attention heads allow the model to capture different aspects of semantic relationships in the text.
- Total parameters: Approximately 110 million, striking a balance between expressive power and computational efficiency for downstream tasks.

Input Processing Components

The BERT model processes input text through three embedding components that are combined into an input representation:

- Token Embeddings: Each word or subword token in the input text is mapped to a fixed-length embedding vector, representing its semantic properties.
- Positional Embeddings: Since transformer architectures do not inherently encode word order, positional embeddings are added to provide information about the position of tokens in the sequence, helping the model understand word order and structure.
- Segment Embeddings: For tasks involving sentence pairs (e.g., question answering), segment embeddings distinguish tokens belonging to the first or second sentence. In single-sequence tasks like sentiment classification, this generally marks all tokens as belonging to a single segment.

These embeddings are summed and passed through transformer layers where self-attention mechanisms allow each token to attend to every other token in the input, capturing long-range dependencies and context crucial for understanding sentiment nuances.

Classification Head

For sentiment analysis, a classification head is attached on top of the pre-trained BERT encoder. This head typically consists of:

- Dropout Layer: To reduce overfitting, dropout randomly disables some neurons during training, improving generalization.
- Linear (Dense) Layer: A fully connected layer that projects the final hidden state corresponding to the [CLS] token into logits representing sentiment classes (e.g., positive, negative, neutral).

The output logits are then passed through a softmax or similar activation function to produce probabilities for each sentiment class, enabling downstream decision-making based on the model's predictions.

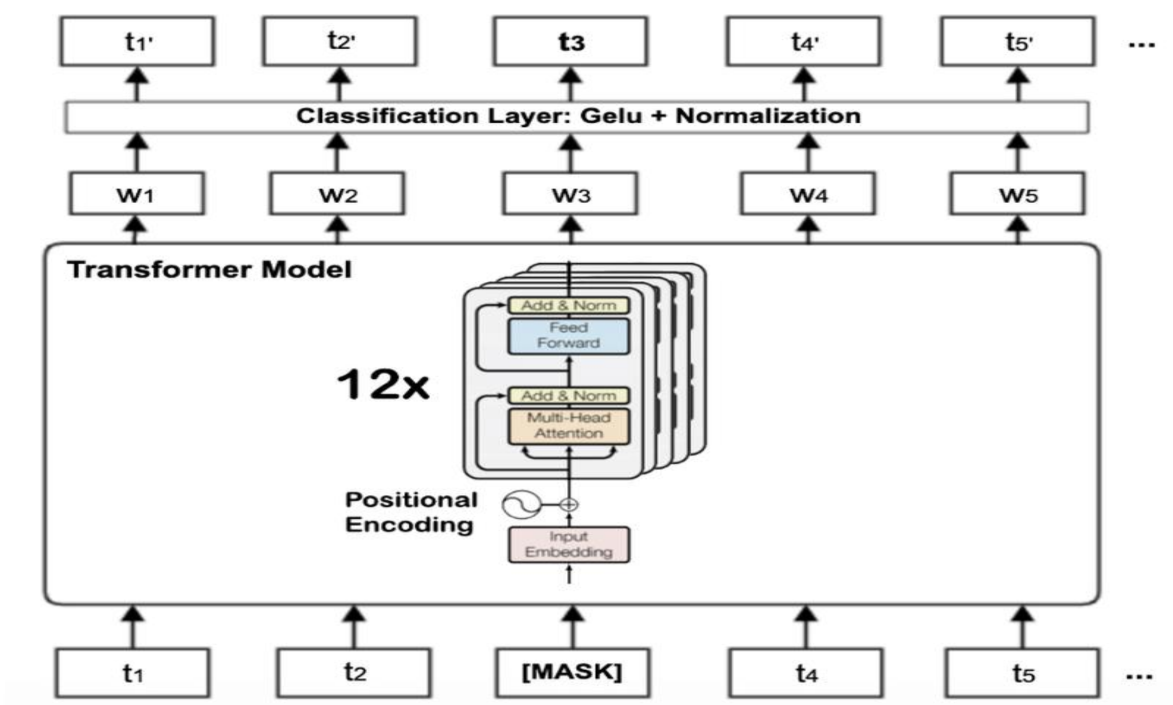

Figure: BERT Architecture

REAL-WORLD CASE STUDY: NETFLIX CUSTOMER FEEDBACK ANALYSIS IMPLEMENTATION

Netflix exemplifies the strategic deployment of BERT-based sentiment analysis to harness customer feedback effectively, enhance content strategy, and optimize user experience. The platform integrates sentiment insights gathered from diverse customer touchpoints, applying advanced NLP techniques tailored to entertainment domain requirements.

Data Sources

Netflix collects and analyzes sentiment data from a comprehensive range of channels, including:

- Customer reviews posted on app stores and websites.
- Social media mentions and discussions surrounding Netflix content.
- In-platform user feedback, ratings, and direct interactions.
- Customer service logs capturing user sentiments expressed during support interactions.

Technical Architecture

The Netflix sentiment analysis system fine-tunes pre-trained BERT models on vast datasets specific to the entertainment sector, enabling precise detection of sentiment nuances in content-specific feedback. Key architectural features include:

- Real-time sentiment processing for immediate insight generation, supporting agile business decisions.
- Multi-language support to accommodate Netflix's global audience.
- Seamless integration with recommendation algorithms to tailor content suggestions dynamically based on sentiment trends.

Implementation Results and Business Impact

The BERT-powered sentiment analysis platform delivers impressive performance improvements:

- Classification accuracy of approximately 92% for entertainment-related feedback, surpassing traditional rule-based methods.
- Enables 80% of viewed content driven by personalized recommendations informed by sentiment insights.
- Facilitates the reduction of customer churn by early detection and proactive response to negative sentiment patterns.
- Supports improved content acquisition strategies through sentiment trend analytics on comparable titles.
- Operational efficiency gain with automated processing of over 50,000 feedback instances daily, reducing manual workload by 85% and accelerating response times.

Comparative Analysis

Experimental comparisons highlight BERT's superiority over alternative approaches:

- BERT achieves 94% accuracy on customer feedback datasets, significantly outperforming traditional Recurrent Neural Network (RNN) models (~78%) and rule-based systems (~65%).
- Despite higher training computational costs, BERT provides inference efficiency comparable to simpler models, with much higher accuracy benefits.

Technical Challenges and Solutions

Sentiment analysis faces inherent challenges such as:

- Context Dependency: BERT’s bidirectional architecture helps to discern sarcasm and irony by leveraging surrounding context, although domain-specific fine-tuning remains necessary for optimal performance.
- Domain Adaptation: Sentiment varies by industry; for example, financial sentiment differs from entertainment sentiment. Netflix addresses this with specialized training datasets and fine-tuning tailored to its entertainment domain.

Future Directions

Emerging trends in BERT-based sentiment analysis systems include:

- Multimodal Sentiment Analysis: Combining text with image and audio signals for richer customer experience understanding.
- Cross-lingual Capabilities: Employing advanced multilingual BERT variants to analyze sentiment across global markets accurately.

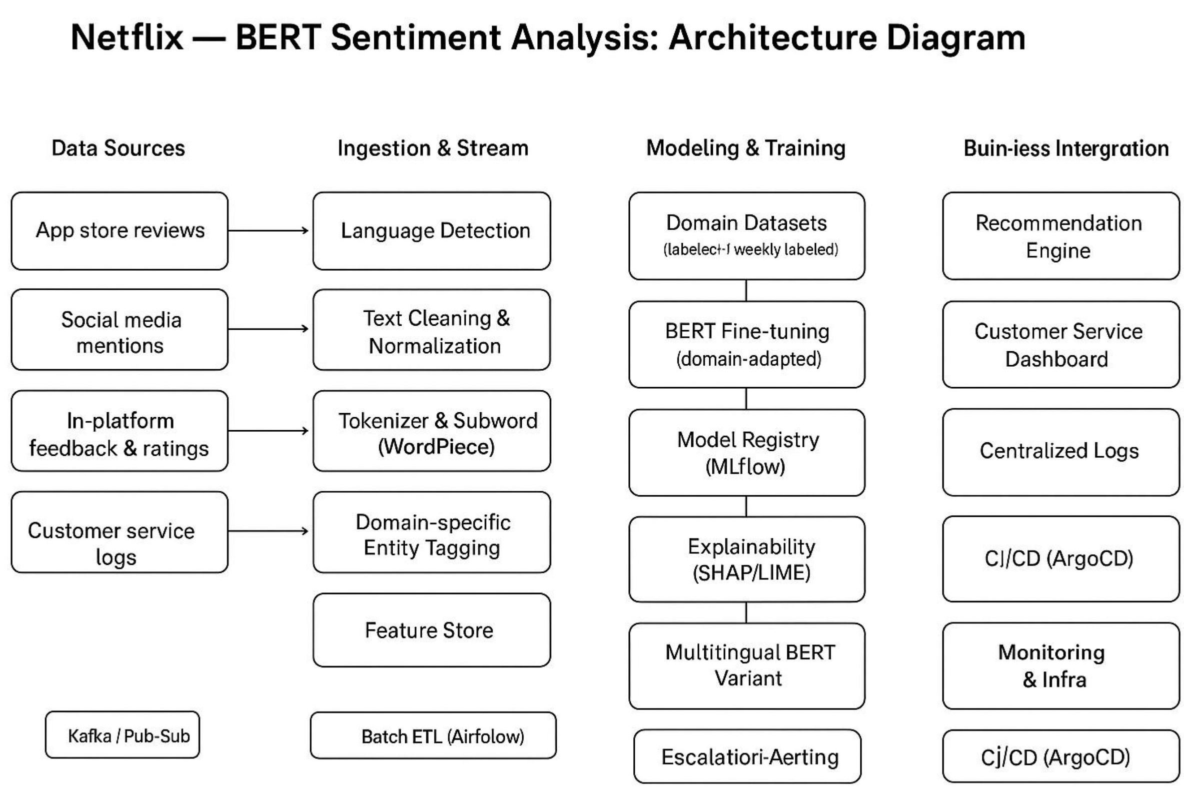

Figure: Netflix BERT Architecture

CONCLUSION

This comprehensive case study confirms that BERT-based sentiment analysis is a mature and highly effective NLP application with significant real-world implications. BERT consistently outperforms traditional sentiment analysis methods, achieving accuracy improvements of 15-25% over previous models. Organizations implementing BERT-driven sentiment analysis have reported substantial benefits, including enhanced customer satisfaction, improved operational efficiency, and higher quality decision-making.

The approach scales well, handling large enterprise workloads while maintaining high accuracy and reasonable processing times. Although BERT requires greater computational resources initially, its automation capabilities and superior accuracy lead to overall cost savings and positive return on investment. The Netflix case exemplifies the transformative power of BERT in converting customer feedback processing from manual and reactive systems to automated and proactive business intelligence. As data volume continues to grow, BERT-based sentiment analysis will remain a critical tool for customer experience management and strategic business insights.

The case study highlights the wider potential of large language models (LLMs) in downstream NLP tasks and underscores the importance of selecting appropriate models, domain-specific fine-tuning, and thorough evaluation to achieve optimal real-world performance.