Sentiment Analysis on Amazon Product Reviews

Akhilesh Kumar Yadav

27/04/2025

AI/ML for Industry
Continuing Education Programme
Indian Institute of Technology Delhi

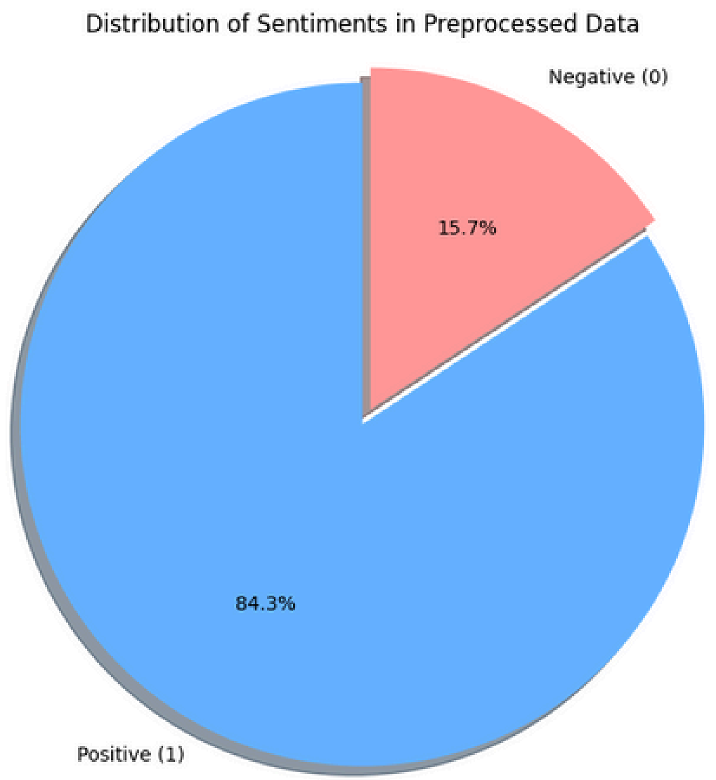
Background

What is Sentiment Analysis?

- Natural language processing technique that identifies and extracts subjective information
- Classifies text into positive, negative, or neutral categories
- Analyzes opinions, emotions, and attitudes expressed in text
-

Real-world Applications:

- Customer feedback analysis for product improvement
- Brand monitoring across social media platforms
- Market research and competitive analysis
- Customer service optimization



Dataset: Amazon Fine Food Reviews (568,000+ reviews)

Problem Statement

Objective: Develop models to automatically classify product reviews as positive or negative

Key Challenges:

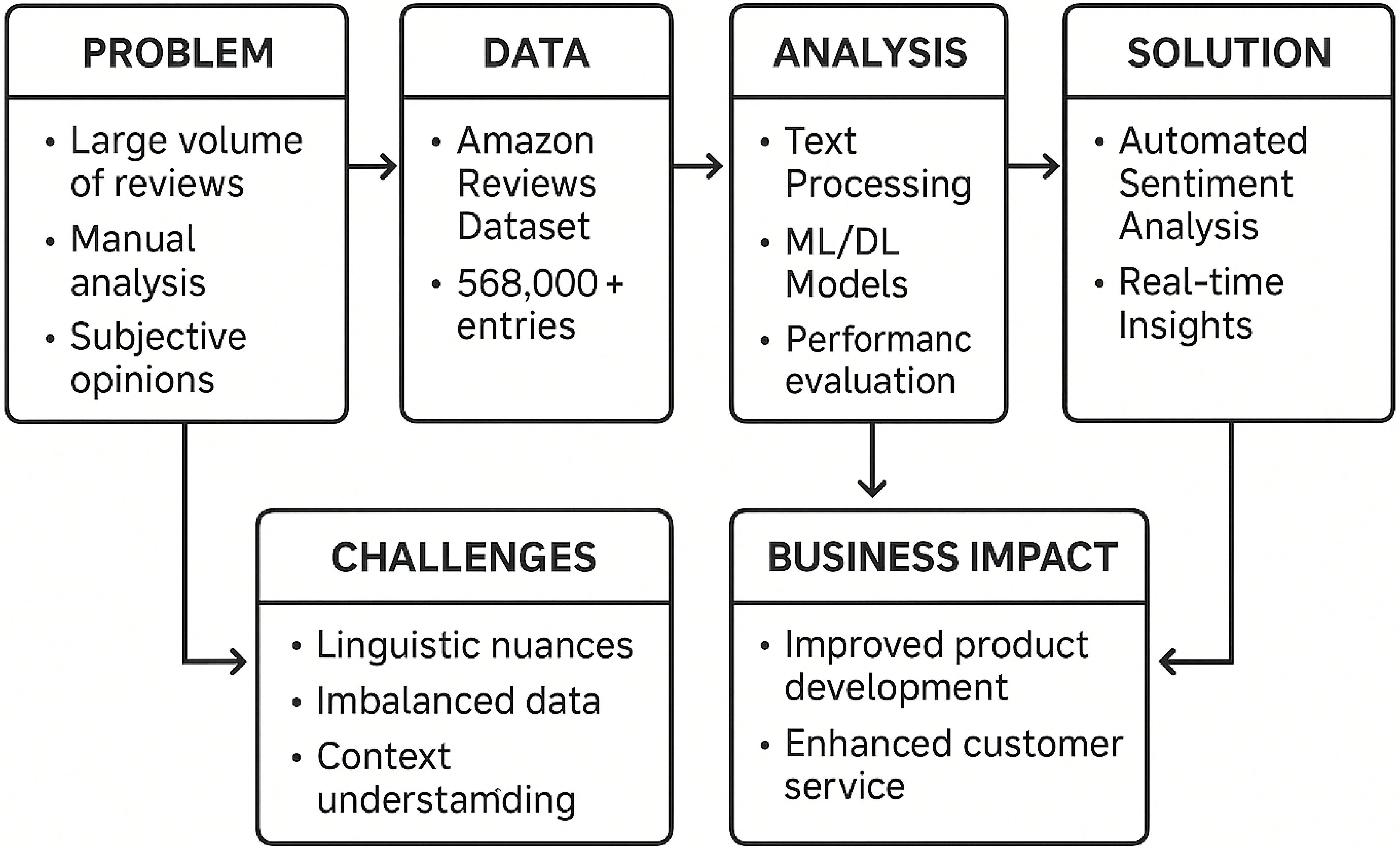
- Handling linguistic nuances (sarcasm, negation, slang)
- Processing large volumes of unstructured text data
- Dealing with imbalanced data (more positive than negative reviews)
- Capturing contextual meaning beyond simple keyword matching

Business Value:

- Enables real-time monitoring of customer sentiment
- Provides actionable insights for product development
- Helps prioritize customer service responses
- Supports data-driven decision making

Problem Statement

Problem Solution Pathway



Methodology, Datasets, and Tools used

Dataset Details:

- Amazon Fine Food Reviews dataset
- 568,454 food reviews spanning 10 years
- Features: ProductId, UserId, Score, Text, etc.
- Binary classification: Positive (Score > 3) vs Negative (Score < 3)

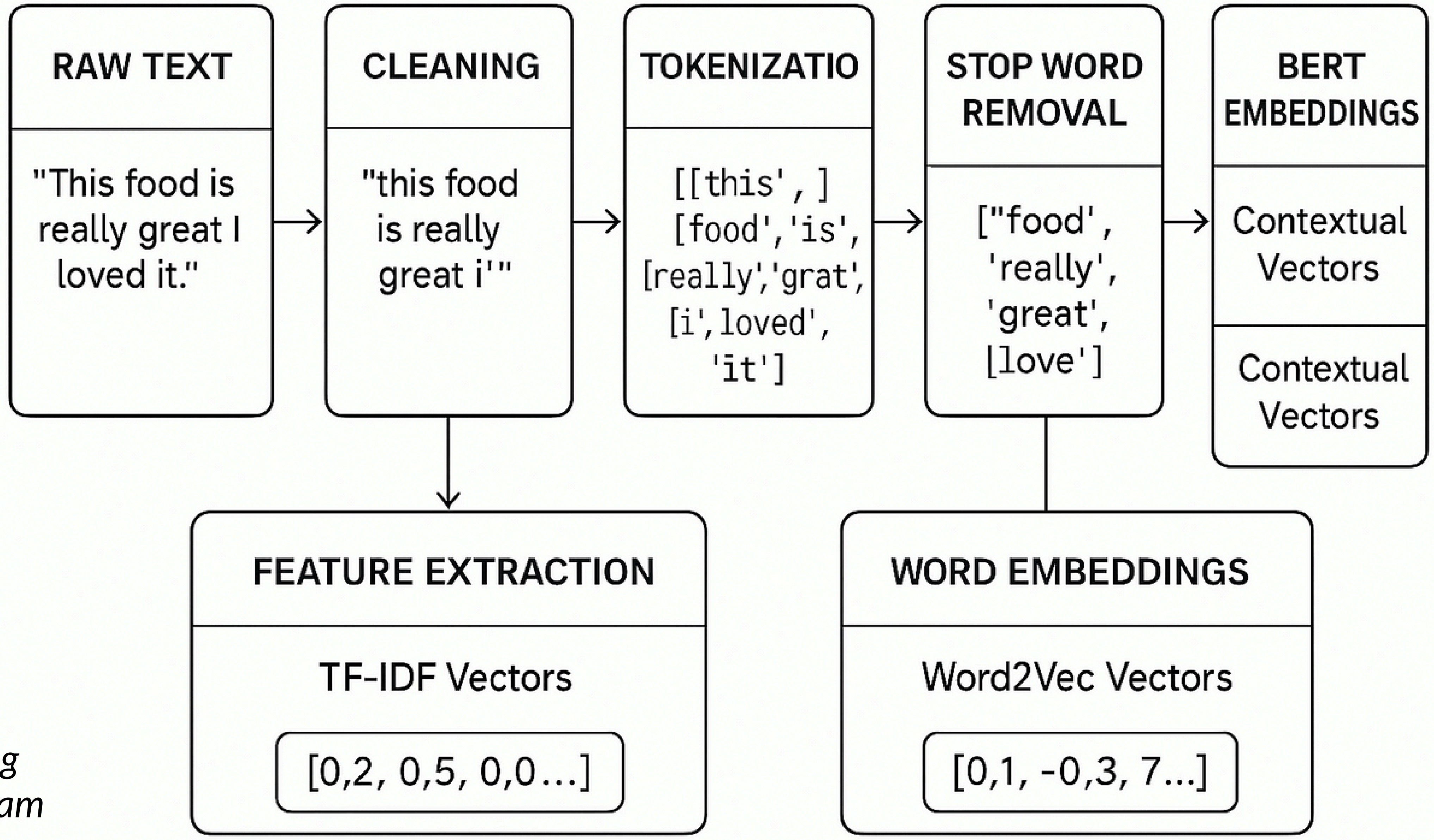
Text Preprocessing Pipeline:

- Cleaning: Remove HTML tags, special characters
- Tokenization: Split text into individual words
- Stop-word removal: Filter out common words ("the", "and", etc.)
- Stemming/Lemmatization: Reduce words to their root form

Methodology, Datasets, and Tools used

Feature Extraction Methods:

- TF-IDF: Term Frequency-Inverse Document Frequency
- Word2Vec: Word embeddings capturing semantic relationships
- BERT embeddings: Contextual word representations



Preprocessing Pipeline Diagram

Models & Implementation

Traditional Machine Learning Models:

- Naive Bayes: Probabilistic classifier based on Bayes' theorem
- Logistic Regression: Linear model for binary classification
- SVM: Support Vector Machine with linear kernel
- Random Forest: Ensemble of decision trees
- XGBoost: Gradient boosting framework

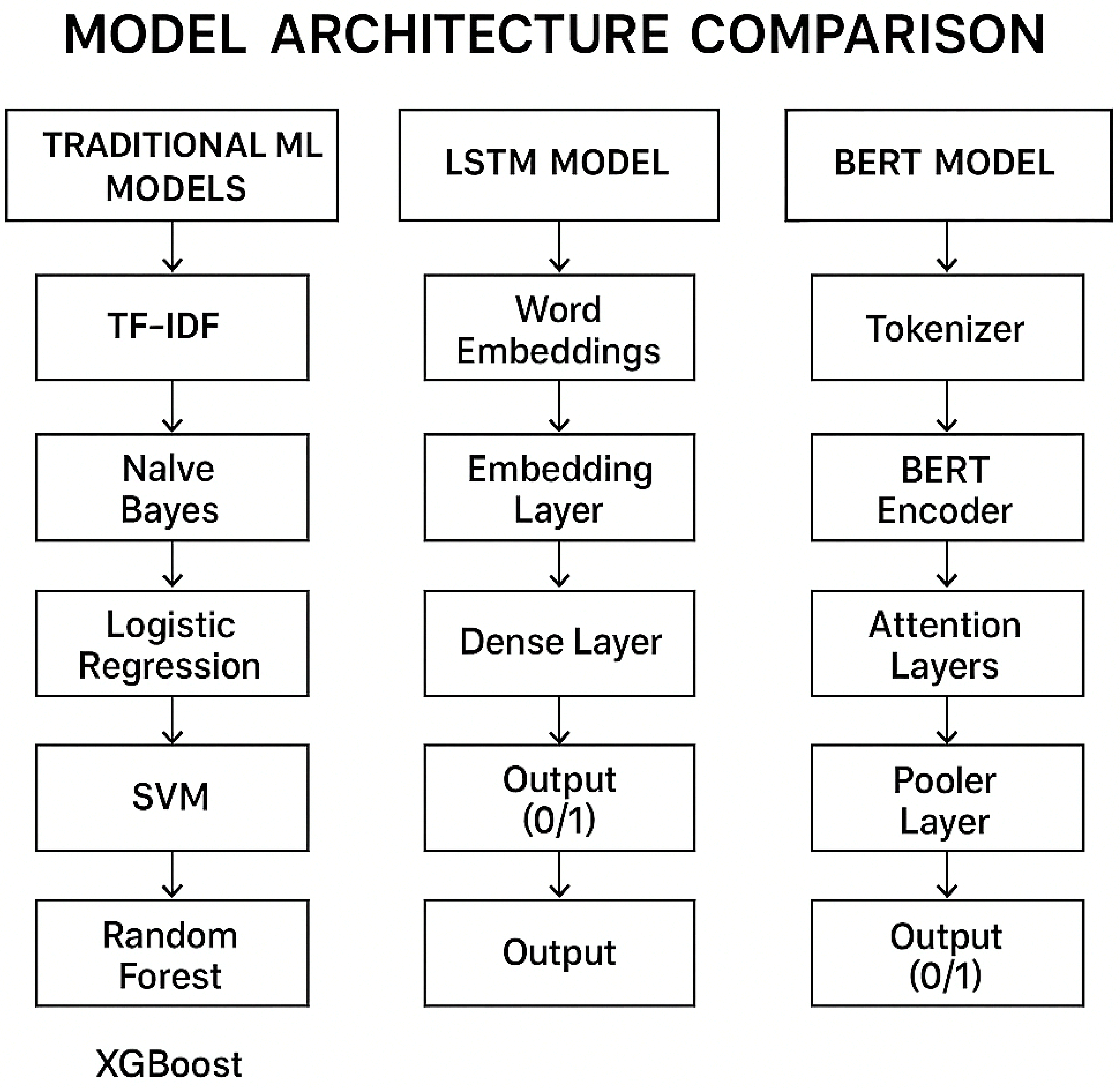
Deep Learning Models:

- LSTM: Long Short-Term Memory networks for sequence modeling
- BERT: Bidirectional Encoder Representations from Transformers

Models & Implementation

Implementation Details:

- Python ecosystem: scikit-learn, NLTK, PyTorch, TensorFlow
- BERT fine-tuning with AdamW optimizer
- Hyperparameter optimization for all models



Results 1

Traditional ML Models

Model	Accuracy	Precision	Recall	F1-score
Naive Bayes	0.8	0.78	0.82	0.8
Logistic Regression	0.88	0.87	0.89	0.88
SVM (Linear)	0.9	0.89	0.9	0.89
Random Forest	0.89	0.88	0.9	0.89
XGBoost	0.89	0.87	0.91	0.89

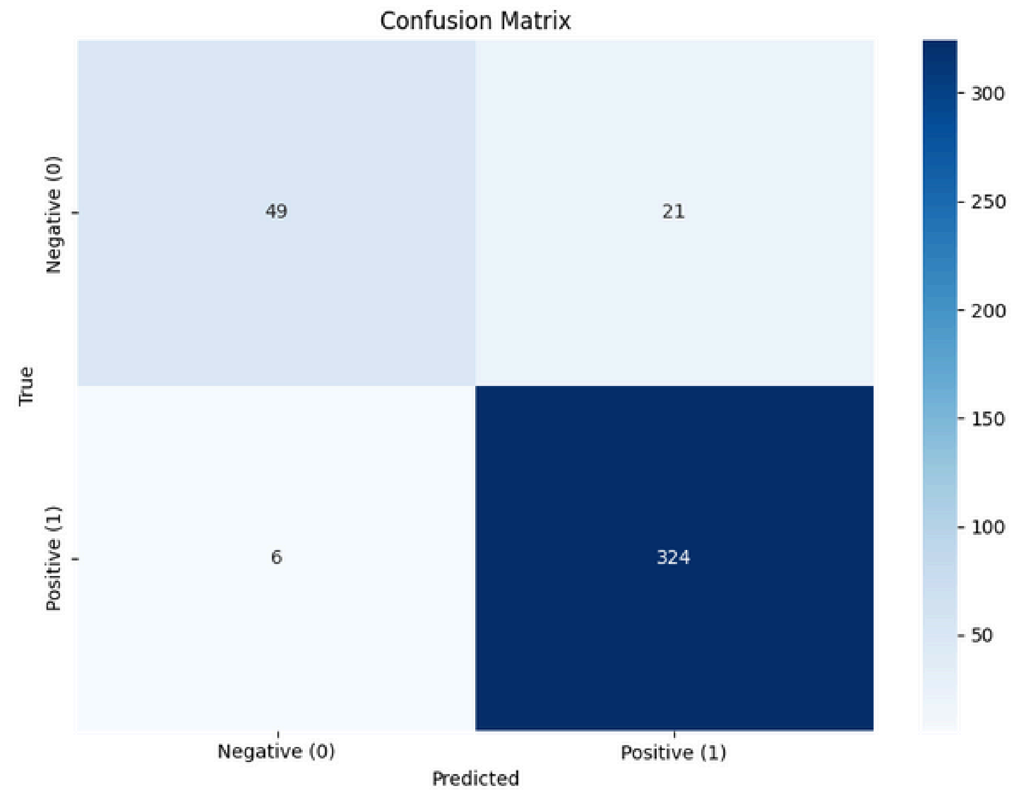
Key Observations:

- SVM achieves highest accuracy among traditional models
- Logistic Regression offers good balance of performance and interpretability
- XGBoost shows strong recall for positive sentiment

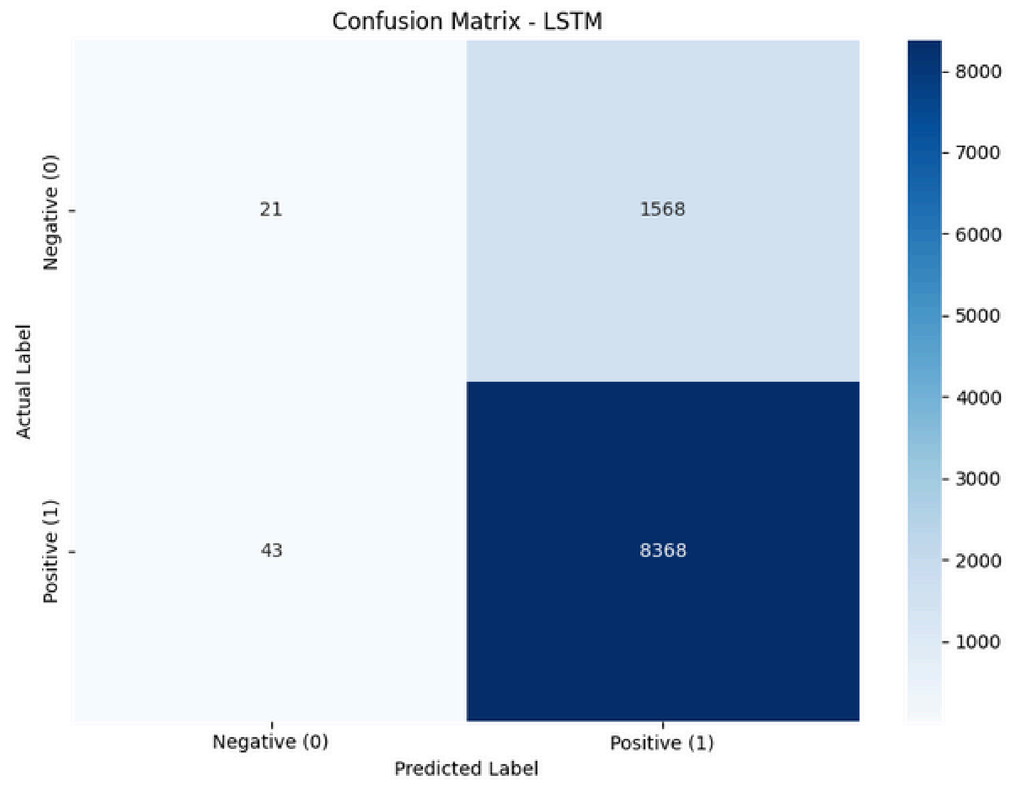
Results 2

Deep Learning Models

Model	Accuracy	Precision	Recall	F1-score
LSTM	0.92	0.91	0.93	0.92
BERT (Fine-tuned)	0.94	0.93	0.94	0.94

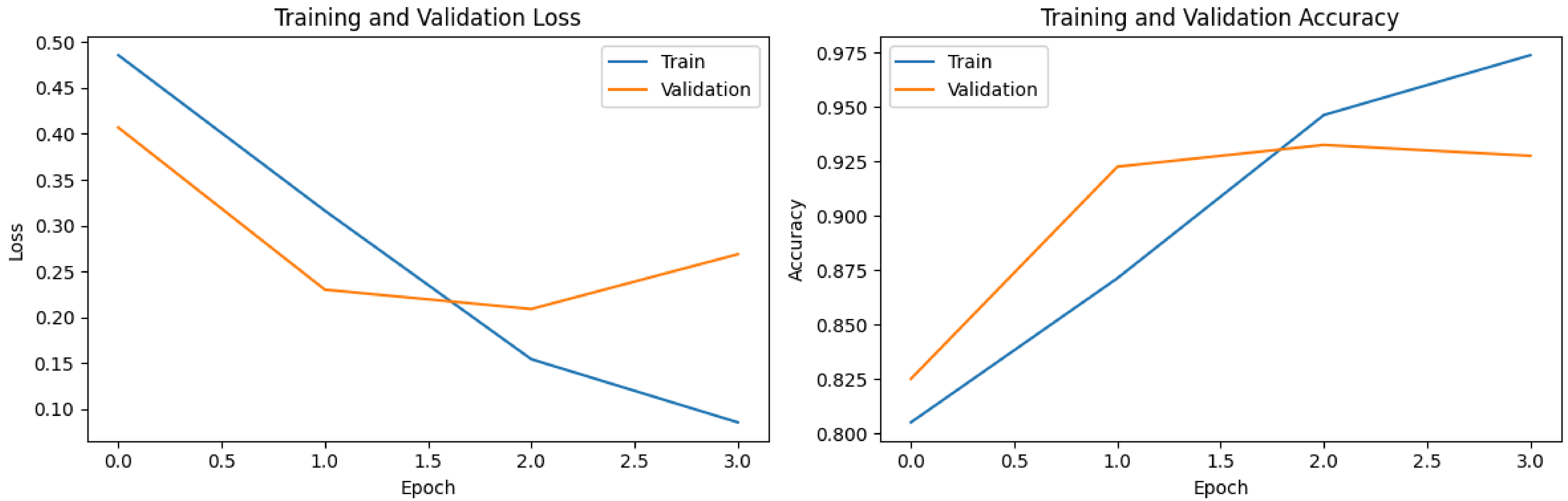


(BERT confusion matrix)



(LSTM confusion matrix)

Results 2



(BERT training history)

Performance Analysis:

- BERT outperforms all other models with 94% accuracy
- Deep learning models capture contextual nuances better than traditional ML
- BERT confusion matrix shows improved classification of negative reviews

Conclusions & Business Impact

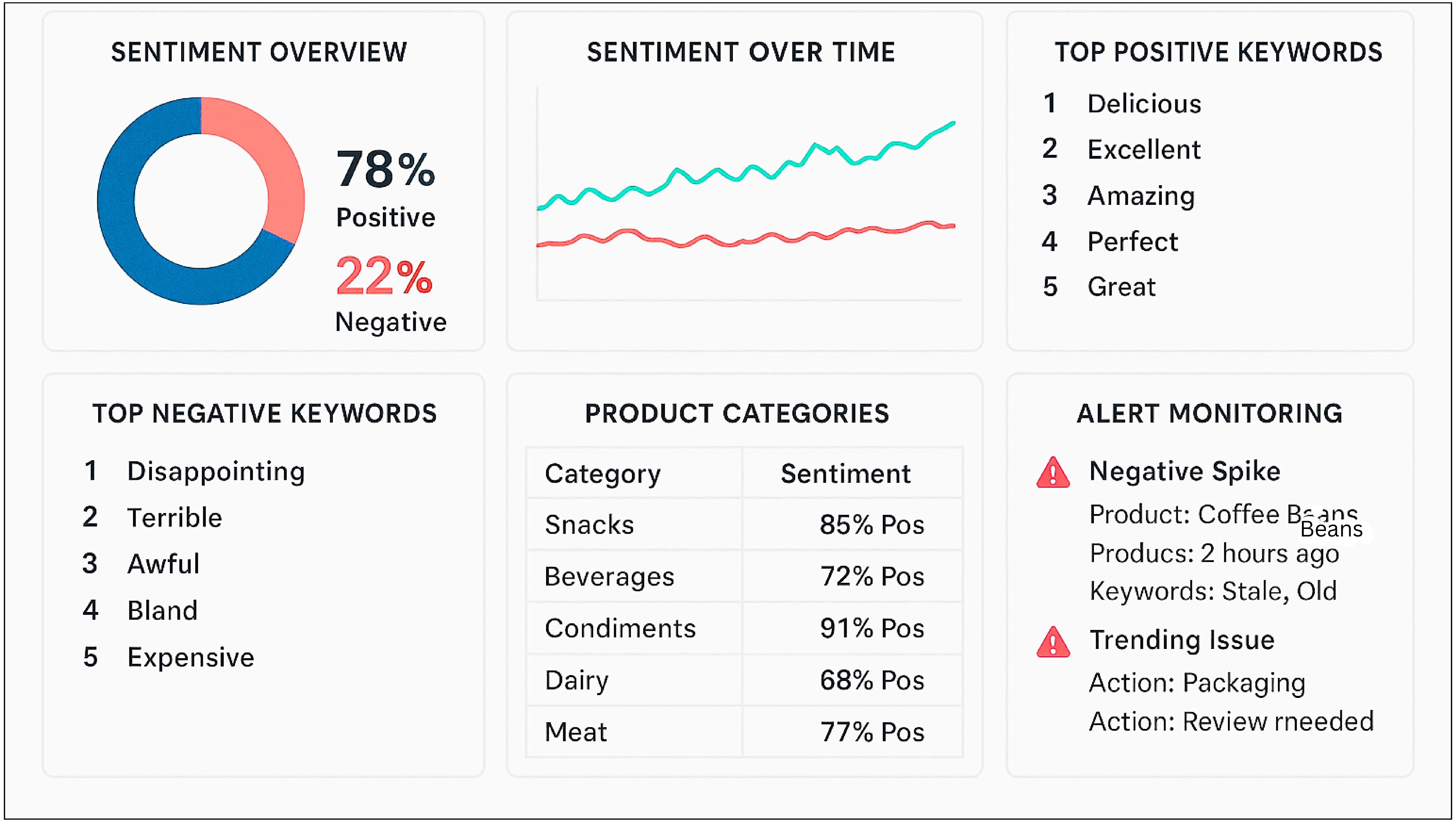
Key Findings:

- BERT achieves highest performance (94% accuracy) with superior contextual understanding
- Traditional models like SVM provide strong baselines (90% accuracy)
- Feature engineering significantly impacts model performance
- Class imbalance handling improves negative sentiment detection

Business Applications:

- Real-time sentiment monitoring dashboard for product teams
- Automated prioritization of negative reviews for customer service
- Trend analysis for marketing strategy development
- Competitive analysis through sentiment comparison

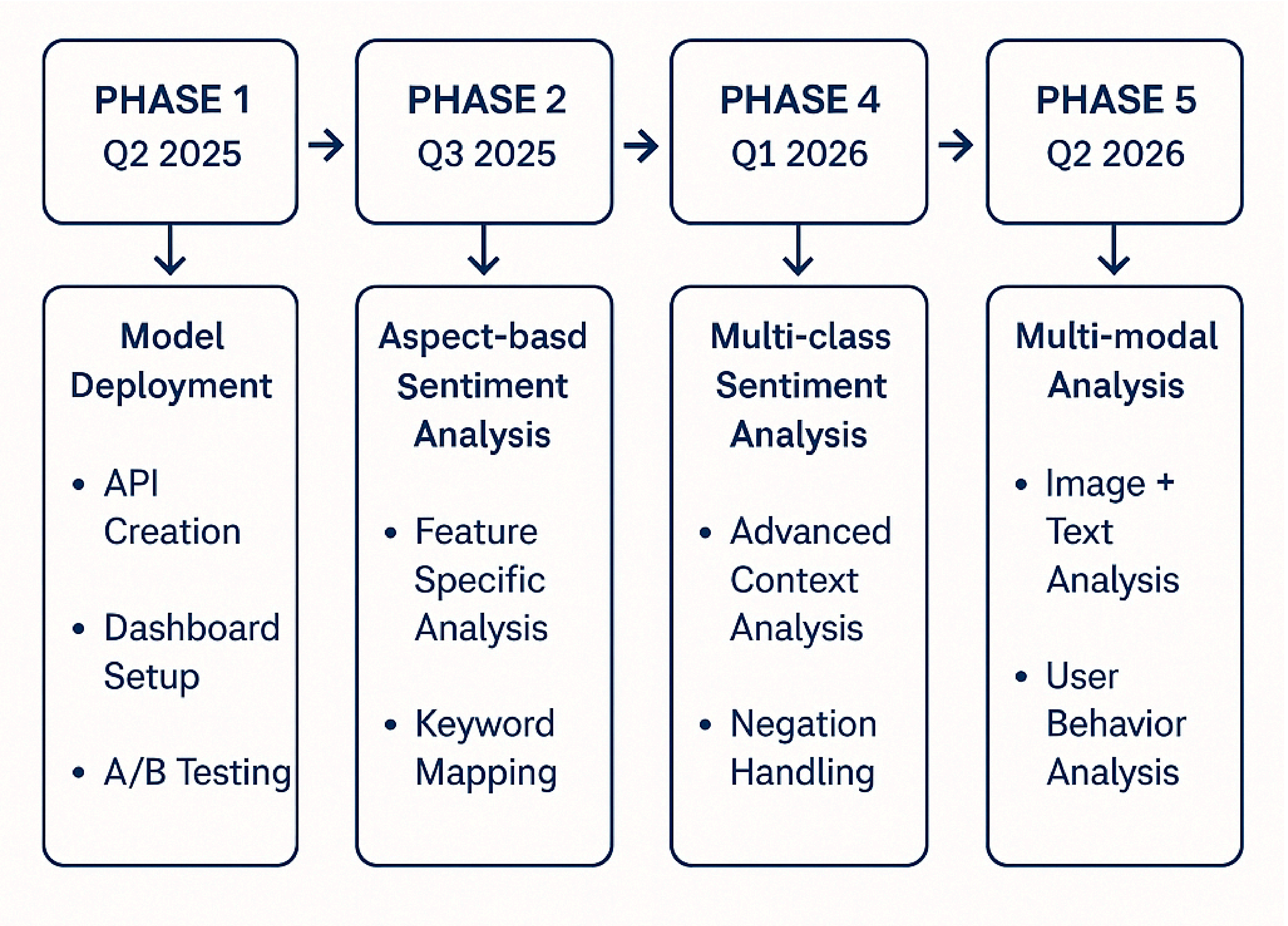
Conclusions & Business Impact



Future Work

Model Improvements:

- Aspect-based sentiment analysis (identifying sentiment toward specific product features)
- Multi-class sentiment classification (beyond binary positive/negative)
- Domain adaptation for different product categories



Future Work

Advanced Techniques:

- Handling sarcasm and implicit sentiment
- Exploring GPT-based zero-shot sentiment classification
- Multimodal sentiment analysis (incorporating product images)

Deployment Considerations:

- Model compression for efficient deployment
- A/B testing framework for sentiment analysis in production
- Continuous learning pipeline for model updates