ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MULTI-ASPECT SENTIMENT ANALYSIS IN AMAZON REVIEWS USING GMM-ENHANCED TEXTCAPS WITH PROBABILISTIC CAPSULE ROUTING

SRI SOWMIYA.S[1], Dr. L. SANKARI[2]

[1]Research Scholar, Department of Computer Science, Sri Ramakrishna College of Arts & Science For Women, Coimbatore, Tamilnadu, India

[1]Assistant Professor, Department of Computer Technology and Data Science, Sri Krishna Arts and Science College, Coimbatore, Tamilnadu, India

.[2]Professor & Head B.Sc (IT), Department of Computer Science, Sri Ramakrishna College of Arts & Science For Women, Coimbatore, Tamilnadu, India

Email: [1]srisowmiya812@gmail.com, [2]hod-bscit@srcw.ac.in

ABSTRACT

Online shopping enables consumers to purchase goods through e-commerce platforms, where customer reviews significantly influence buying decisions and vendor reputation. Sentiment analysis of these reviews facilitates the decoding of public opinion, offering actionable insights for product improvement, marketing personalization, and an enhanced customer experience. The reviews often contain mixed sentiments, sarcasm, and abrupt polarity changes, which traditional models tend to oversimplify, thereby losing sentiment flow and interpretability. This study proposes GMM-Enhanced TextCaps, an advanced model addressing these challenges by employing Byte-Pair Encoding for subword-level sentiment extraction, capsule networks for granular aspect representation, and Gaussian Mixture Models for probabilistic sentiment routing. The model processes enriched token embeddings through contextual transformer encoders, sarcasm and negation adjustments, and emotion-weighted attention mechanisms, forming interpretable capsules that yield confident sentiment predictions. Experiments utilize the Clothing, Shoes, and Jewelry category from the Amazon Reviews ’23 dataset, comprising 451,478 reviews divided into 80% training, 10% validation, and 10% testing splits. The model attains a balanced accuracy of 72.51% and a G-Mean of 72.45%, demonstrating robust performance in capturing nuanced sentiment dynamics while effectively maintaining class balance.

Keywords: *Aspect-Based Sentiment Analysis, Capsule Networks, Gaussian Mixture Models, Sentiment Interpretation, Multimodal Sentiment Analysis, Transformer Encoding*

1. INTRODUCTION

Amazon product reviews often carry complex emotional signals embedded in informal, aspect-rich narratives that evolve across the review text [1]. A single user review may praise one feature, complain about another, or even express sarcasm that masks actual sentiment. Traditional sentiment classification models struggle with these cases, especially when using static embeddings or single-output architectures [2]. They often reduce user sentiment to oversimplified labels without considering polarity shifts or sub-aspect variability. Words like “light”, “decent”, or “cheap” may flip their emotional meaning depending on surrounding phrases. Recognizing sentiment from such structurally scattered, semantically ambiguous content demands an approach that respects emotional layering,narrative sequence, and token-specific variability without collapsing everything into a single averaged output [3].

To improve representation quality, the model begins with Byte-Pair Encoding (BPE), which allows for token segmentation into sentiment-rich subwords that traditional tokenizers often overlook. Words like “overhyped” or “non-functional” carry strong emotional units that require individual attention [4]. BPE-based subwords are enhanced with local attention and passed through a light transformer that captures contextual interactions. These subword signals not only strengthen token-level encoding but also refine capsule construction, allowing each capsule to better align with the emotional semantics of its contributing tokens. Instead of compressing this into a flat embedding, capsule vectors are created for each sentiment aspect [5]. These capsules preserve semantic, emotional, and positional data, allowing routing based on token importance. Their structure ensures interpretability, with each capsule traceable to contributing tokens and emotional peaks. This interpretable mechanism is supported through activation-based heatmaps, allowing visual analysis of sentiment pathways and emotional reasoning [6], [7].

A core innovation of this model lies in combining capsule-based abstraction with GMM-driven probabilistic sentiment routing, enabling nuanced polarity interpretation and soft decision boundaries [8]. Gaussian Mixture Models (GMMs) introduce probabilistic reasoning to capsule activation, particularly in reviews that lie near sentiment boundaries. Many user statements convey mixed tones, such as "great performance, terrible battery," which traditional models often mishandle by forcing single-label outputs. GMM-based routing enables capsules to be assigned to multiple sentiment clusters with varying degrees of certainty, based on posterior probabilities [9]. This prevents emotional flattening and supports confidence-weighted decisions. Each sentiment capsule's activation strength reflects not only emotional salience but also certainty of belief. This clarity makes the model more reliable for downstream applications such as product ranking, sentiment-aware alerts, or brand monitoring, where sentiment confidence and polarity traceability are essential [10]. The explainable structure supports ethical deployment and diagnostic validation in commercial settings.

Though the model currently focuses on textual sentiment, it integrates domain-specific sentiment priors as a parallel source, simulating multimodal input fusion. These priors—based on aspect-level distributions or historical co-occurrence data act as auxiliary signals to support ambiguous sentiment resolution [11], [12]. The framework is also extendable to include visual reviews, ratings, and reviewer metadata in future implementations. This dual-channel setup maintains interpretive consistency while preparing the architecture for multimodal scalability [13]. Compared to prior systems, such as RoBERTa classifiers or vision-only Sentinet models, this method stands out for its emotional granularity, traceable capsule design, and soft probabilistic routing [14]. It offers a balanced trade-off between interpretability, adaptability, and robustness, making it highly suited for sentiment analysis across structurally complex and emotionally dense user reviews [15].

1.1. Problem Statement

User-generated reviews on platforms like Amazon exhibit a high degree of emotional complexity, stylistic informality, and aspect-level fragmentation. A single review often blends contradictory opinions across product features, expressing both praise and criticism in scattered positions. Sentiment polarity may shift mid-sentence or unfold nonlinearly across the text. Traditional models, including CNNs, BiLSTMs, and even transformer-based systems like RoBERTa, struggle to preserve this sentiment flow, especially in long, sarcasm-rich, or context-dependent narratives. Most existing approaches employ flat classification, single-label outputs, or deterministic attention, which fail to resolve ambiguous phrases or overlapping sentiment tones. Current models lack structural interpretability, as they are unable to trace decisions back to specific subword units or sentiment-bearing clauses. Without a probabilistic mechanism to quantify sentiment uncertainty or visualize decision pathways, such systems often misrepresent user intent. This deficiency creates technical and practical limitations in sentiment systems that must function across real-world consumer review data, where emotional clarity, positional context, and aspect-level decoding are essential. A critical gap remains in building an interpretable, sentiment-aware architecture that handles polarity ambiguity, token traceability, and context shifts without flattening the emotional complexity embedded in structurally rich user opinions.

1.2. Motivation

In the digital economy, consumer reviews form a critical pillar of decision-making, influencing billions in transactions across e-commerce platforms. These reviews not only guide buyers but also serve as data streams for product enhancement, inventory tuning, and promotional targeting. When sentiment systems misinterpret nuanced user opinions—such as sarcasm, multi-aspect contrast, or emotionally ambiguous phrasing the consequences extend beyond poor analytics. Incorrect sentiment labeling can lead to misleading recommendations, algorithmic bias in product rankings, unjustified promotions, and a decline in consumer trust. From a societal perspective, such failures undermine transparency, discourage genuine feedback, and diminish platform credibility. Ethical AI standards emphasize the need for explainable and trustworthy models, particularly in domains where user emotions intersect with commerce. There is also a growing demand for systems that can visualize sentiment decisions and support auditability, particularly in cases where reviews impact vendor reputation or regulatory compliance. Consumers deserve technology that understands the depth of opinions, not just label-level impressions. The motivation stems from this pressing societal need to build sentiment systems that not only perform well but also reason responsibly and visibly in emotionally sensitive domains.

1.3. Objective

The primary objective of this research is to propose a novel architecture, titled GMM-Enhanced TextCaps, specifically developed to address the challenges outlined in the problem statement. The model is designed to decode complex, multi-aspect Amazon reviews by combining subword-level sentiment abstraction with capsule-based routing and probabilistic interpretation. It integrates Byte-Pair Encoding to extract sentiment-bearing subunits, employs capsule networks to construct interpretable aspect-wise vectors, and utilizes Gaussian Mixture Models to handle ambiguous or overlapping sentiment zones through soft probability-based routing. This architecture enables token-to-capsule traceability, emotional heatmapping, and evidence-driven reasoning across sentiment layers. It is also structured for extensibility, simulating early multimodal fusion by integrating domain-specific sentiment priors alongside textual input. The model explicitly supports real-world interpretability, emotion-layer tracking, and confidence-aware classification. The goal is not only to improve sentiment detection accuracy but also to construct a traceable, probabilistically transparent sentiment system capable of adapting to the complexity of real-world reviews, thereby fulfilling both technical robustness and societal trust requirements.

1.4 Research Questions and Hypotheses

This study is guided by clearly defined research questions and corresponding hypotheses. The primary research question examines whether probabilistic capsule routing improves sentiment classification under class imbalance and mixed polarity conditions. A second question investigates whether subword-level representation and aspect-aware capsules enhance interpretability and sentiment traceability. A third question explores whether Gaussian Mixture Modeling contributes to more stable sentiment decisions near polarity boundaries. Based on these questions, the study hypothesizes that probabilistic capsule routing yields higher balanced performance metrics than deterministic baselines, that aspect-aware capsule structures improve sentiment interpretability, and that GMM-based uncertainty modeling strengthens robustness in ambiguous review scenarios.

2. LITERATURE REVIEW

"RoBERTa-Driven Flask Classifier" [16] applied RoBERTa to encode each movie review into contextual embeddings, capturing sentiment dependencies across long text. These embeddings fed into a CNN that extracted localized n-gram features and used max-pooling to highlight salient emotion cues. Fully connected layers produced binary sentiment outputs, and a Flask interface enabled real-time predictions. Pairing RoBERTa's language modeling with CNN's sentiment extraction yielded a minimal, efficient pipeline for practical sentiment analysis. "Keyword-to-Image Sentinet" [17] created a novel pipeline by first extracting dominant keywords from product reviews using Word2Vec with TextRank. These keywords were visualized via a generative AI model into sentiment-expressive images. A CNN-SVM hybrid classified these images, where the CNN analyzed visual cues and the SVM refined the emotional label assignment. This approach transformed textual sentiment analysis into a visual task, enabling the detection of indirect emotions through image-based classification. "Usability-Driven ABSA Network" [18] focused on non-functional aspects of app reviews by combining BERT, BiLSTM, and CNN in a hybrid stack. BERT provided context-aware token outputs, BiLSTM captured sequential sentiment dynamics, and CNN detected short-range emphasis. Aspect-based mapping sorted reviews into effectiveness, efficiency, and satisfaction, each passing through unique sentiment classifiers. This architecture offered usability-focused insights by applying layered neural strategies to dissect qualitative app performance dimensions.

"Saccade-Aware Patch Transformer" [19] drew inspiration from human eye movement to construct an image sentiment framework. Images were split into patches, organized in sequences that mimicked gaze shifts, with sequence attention modules replicating saccades and gated convolutional recurrent units modeling dynamic visual focus. Positional encoding maintained temporal patch memory. A final self-attention fusion layer correlated these features, enabling refined extraction of subtle emotions from visual input by blending mimicry and patch analysis. The "Sentiment-Marketing Analyzer" [3] developed a comparative setup that evaluated SVM, NB, RNN, and LSTM classifiers for customer hotel reviews. Reviews underwent tokenization, normalization, and transformation into feature vectors or embeddings. Each model was trained and assessed for its ability to detect brand emotions, with feedback trends tracked to inform brand strategies. The approach balanced classical and deep learning methods, offering insights relevant to brand management within hospitality sentiment datasets. "STGP-SATA Trading Fusion" [20] designed a genetic programming architecture with distinct branches for sentiment indicators and technical trading signals. Type constraints ensured a clear division during evolutionary operations, optimizing financial metrics such as the Sharpe ratio. Sentiment signals derived from text were integrated with quantitative trading data in evolved genetic trees, enhancing decision diversity and trading relevance through domain-specific bifurcation.

The "Dual-Problem SSLMM Solver" [21] introduces a unified framework that tackles data scarcity and missing modalities via self-supervised pretraining for intra-modal features. Alternating interaction modules disseminate emotional information across modalities through graph-based propagation. Missing data was regenerated through modality reconstruction, and semi-supervised tuning boosted performance with minimal labeled data. This design reflected real-world data challenges, integrating robust feature sharing and flexible label use for emotion classification. "CTRN Robust Fusion Network" [22] integrated a convolution within a transformer encoder to capture both local and global relationships across multiple modalities. Features were extracted and fused via cross-attention, with convolution preserving semantic details. A robust optimization unit reconstructed noisy data, and adversarial training improved model stability. Momentum distillation minimized prediction variance, creating a structured and resilient fusion for practical sentiment analysis, even when the input data was partial or noisy. "Crisis-Phase Emotion Mapper" [23] applied a time-phased sentiment extraction model tracking public opinion throughout an event's lifecycle. Key terms were identified using TF-IDF, and thematic clusters using LDA. Sentiment lexicons tailored for Chinese assigned emotional labels as the public crisis progressed through outbreak, spread, and decline. Comparative emotional tracking revealed shifting moods, directly tying sentiment changes to evolving events and emerging themes.

"Urdu Biaffine AspectNet" [24] enhanced aspect-level sentiment analysis using biaffine attention to map aspect-term dependencies. BiLSTM provided contextual word representations, and the biaffine mechanism modeled interactions between aspect and opinion words. The architecture scored these pairs for polarity, assigning precise sentiment labels. Especially effective for Urdu's complex morphology, this approach accommodated shifting word boundaries and provided fine-grained sentiment analysis at the aspect level. The "Routing-Based Distribution Matcher" [25] proposes a dual-stage approach, first aligning modality feature distributions through regularized matching and then optimizing feature exchange with dynamic routing. Embedded representations from text, audio, and video are passed through capsule-inspired routing layers, which are controlled to facilitate adaptive information flow. Contrastive loss enforced semantic consistency between aligned features. The architecture prioritized both cohesion and flexibility, enhancing emotionally accurate sentiment inference across different modalities. "Arabic Review Senticheck" [26] utilized COVID-19 mobile app user reviews to evaluate sentiment using traditional machine learning algorithms. Preprocessing steps included script normalization and diacritic removal. Feature extraction relied on TF-IDF and n-gram models. SVM, Decision Tree, and Naive Bayes classifiers were used to identify sentiment trends in usability, reliability, and responsiveness. Satisfaction levels were mapped and grouped thematically, with the method tailored for Arabic's linguistic nuances and focused on app-specific user feedback and experience.

"TF-IDF BERT" [27] engineered a hybrid sentiment analysis framework by integrating Term Frequency-Inverse Document Frequency (TF-IDF) with Bidirectional Encoder Representations from Transformers (BERT). TF-IDF captured statistical importance of lexical tokens, encoding sparse syntactic features, while BERT generated dense contextual embeddings via transformer-based self-attention layers. A concatenation mechanism aligned both representations into a unified vector space. This fusion was input into a deep neural classifier composed of fully connected layers activated by ReLU functions. The architecture preserved both global semantic dependencies and local word salience, ensuring sentiment polarity detection across domain-variant sentence structures. "ENLPPR-ICFFO" [28] deployed a hybridized neuro-evolutionary framework by embedding an Elman Neural Network (ENN) within an optimization loop governed by Improved Chaotic Fruit Fly Optimization (ICFFO). The ENN architecture, characterized by self-feedback connections at the context layer, enables the dynamic retention of sequential information across time steps. Feature vectors derived from Continuous Bag-of-Words (CBoW) encoded semantic priors. ICFFO applied logistic chaotic mapping to perturb the odour-concentration fitness landscape, ensuring robust navigation across non-convex error surfaces. The synergistic loop refines weight matrices iteratively, stabilizing temporal modelling in high-variance textual sentiment classification tasks. Bio-inspired optimization plays a major role to achieve the threshold accuracy in sentiment analysis [29]-[59].

Prior research in sentiment analysis has largely concentrated on improving classification accuracy through stronger feature extractors or deeper neural architectures. Transformer-based classifiers commonly employ pooled representations to infer sentiment labels, which limits their ability to preserve aspect-level polarity variation and emotional transitions within long reviews. Hybrid systems combining recurrent networks, convolutional layers, or attention mechanisms have improved contextual sensitivity, yet they continue to rely on deterministic routing and flat decision boundaries. Capsule-based approaches introduced structural grouping of features, though most existing models apply hard routing or fixed coupling coefficients, restricting their flexibility in handling sentiment ambiguity.

The motivation of the present work diverges from accuracy-centric optimization and instead emphasizes sentiment interpretability, uncertainty awareness, and aspect-wise emotional traceability. The proposed GMM-Enhanced TextCaps framework extends prior capsule-based sentiment models by integrating probabilistic routing through Gaussian Mixture Models, allowing sentiment evidence to contribute across multiple polarity hypotheses. This design explicitly addresses mixed and overlapping sentiment expressions that are frequent in Amazon reviews. Empirical findings demonstrate that probabilistic capsule routing improves balanced accuracy and correlation metrics, confirming that modeling uncertainty and sentiment overlap leads to more reliable sentiment inference than deterministic architectures.

2.1. Methodological Gap

The methodological gap in current sentiment analysis research centres on the absence of unified and flexible frameworks that can effectively integrate heterogeneous modalities—including text, images, audio, and domain knowledge while ensuring interpretability, scalability, and robustness. The reviewed literature encompasses a wide range of architectures, including pre-trained language encoders combined with CNNs, generative visual sentiment models, hybrid neural stacks, saccade-inspired transformers, genetic programming fused with financial signals, and self-supervised learning with graph propagation.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Comparative Study of Literature

Model Name	Adaptability to Unexpected Data	Granularity of Sentiment Signal	Human-Centric Design Element	Transparency of Model Decisions	Ethical Sensitivity	Computational Sustainability	Deployment Agility
RoBERTa-Driven Flask Classifier [16]	Adapts rapidly to streaming reviews	Macro/micro pooled cues	User-facing direct web sentiment service	Pipeline audit from embeddings to output	Binary feedback simplifies interpretation	Minimal resource, practical runtime	Web-ready demo for sentiment insight
Keyword-to-Image Sentinet [17]	Shifts from text to generative vision	Converts textual cues to visual sentiment	Offers indirect emotion capture through images	Traceable transformation from keyword to label	Increases inclusivity for abstract emotion	Image gen cost varies with the pipeline	Deployable on visual feedback tools
Usability-Driven ABSA Network [18]	Decomposes qualitative app reviews	Maps to distinct usability axes	Delivers multi-faceted user experience mapping	Aspect and classifier transparency	Quantifies qualitative performance feedback	Focused mapping avoids overgeneralization	Integrates into App Store Insights
Saccade-Aware Patch Transformer [19]	Mimics human visual exploration	Patch sequence, gaze memory	Replicates visual attention dynamics	Discrete patch attention mapping	Respects temporal perception in images	Patch processing environment-variable	Ideal for smart camera and vision apps
Sentiment-Marketing Analyzer [3]	Evaluates trends across diverse classifiers	Detects brand emotion movement	Assists brand managers in consumer strategy	Comparative trend tracking across models	Cross-validates business risk sensitivity	Standard classifiers keep the computational requirements clear	Dashboard reporting for marketing
STGP-SATA Trading Fusion [20]	Blends stochastic and learned signals	Sharpe-optimized branch selection	Fuses quantitative finance with market psychology	Genetic programming evolution trace	Defends against irrational market emotion	Modular evolution controls resource cost	Fits into trading or analytics suites
Dual-Problem SSLMM Solver [21]	Repairs missing, sparse input	Cross-modal, robust correlation	Minimal labeled data benefits underrepresented tasks	Alternating module reconstruction log	Enables inclusion in low-data contexts	Self-supervision limits annotation cost	Hybrid data pipeline integration

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CTRN Robust Fusion Network [22]	Maintains stability amid partial inputs	Local-global synaptic blending	Guards against input quality drops	Robust optimization pathway visualization	Increases reliability for safety-critical settings	Momentum distillation improves resource use	Suits resilient enterprise solutions
Crisis-Phase Emotion Mapper [23]	Detects evolution through distinct crisis stages	Thematic mood tracking	Supports timely management of public response	Time-phased emotional trend transparency	Draws on tailored lexicons for event relevance	Temporal thematics keep data load manageable	Ready for event-centered reporting
Urdu Biaffine Aspect Net [24]	Tackles morphological complexity in language	Aspect-term, fine-grained mapping	Invites inclusivity for non-English speakers	Biaffine attention log for aspect relation	Adapts to overlooked language nuances	Morphology-aware for efficiency and comprehension	Integrates into regional feedback engines
Routing-Based Distribution Matcher [25]	Matches cross-modality fluidly	Ensures cohesion, flexibility	Values semantic parity between modalities	Routing layer traceability	Mitigates misalignment in emotion signals	Flexible capsules optimize resource sharing	Streamlines multimodal pipeline integration
Arabic Review Sentincheck [26]	Calibrates for script and dialect processing	App usability, reliability	Reflects end-user experience in native context	Classifier performance theme mapping	Addresses language-specific feedback	Lighter classifiers adapt to mobile computation	Fits into mobile app analytics
TF-IDF BERT [27]	Moderate adaptation due to BERT; TF-IDF is less flexible	Captures both local (TF-IDF) and global (BERT) sentiment	Highlights necessary tokens and attention	Partial transparency through attention scores	Standard ethical considerations	Efficient inference with combined features	Easy integration into NLP pipelines
ENLPPR-ICFFO [28]	Strong adaptation via chaotic optimization and recurrent memory	Models sequential and semantic sentiment signals	Supports dynamic sentiment tracking	Moderate transparency; complex inner workings	Implicit attention to data fairness	Efficient training through evolutionary optimization	Modular design for flexible deployment

Despite these innovations, challenges persist in managing noisy, incomplete, or imbalanced multimodal inputs and aligning features dynamically across granular semantic and spatial levels. Many methods involve complex fusion mechanisms with high computational demands, which limit efficiency and deployment feasibility. Transparent sentiment traceability remains scarce, obstructing the explainability of model decisions. There is also a lack of standardized benchmarking protocols to enable objective evaluation across modalities, tasks, and languages. Ethical considerations such as fairness, bias mitigation, privacy, and user trustworthiness receive limited attention. Addressing these aspects requires developing interpretable, robust, and resource-efficient systems that are adaptable to diverse supervision levels, domains, and languages, while incorporating rigorous validation and responsible AI principles. This comprehensive advance promises broader applicability and higher reliability in the real-world.

3. GAUSSIAN MIXTURE MODEL-ENHANCED TEXTCAPS FOR SENTIMENT ANALYSIS

The GMM-Enhanced TextCaps framework is designed to decode complex, multi-aspect product reviews with high accuracy and interpretability. It combines Byte-Pair Encoding for preserving subword sentiment, capsule networks for aspect-specific representation, and Gaussian Mixture Models for probabilistic routing. The process begins with subword embedding, contextual enhancement, and adjustments for sarcasm and negation, followed by emotion-weighted attention and aspect identification. Capsules are formed to retain both emotional and positional cues, which are then clustered via GMM to manage sentiment ambiguity. Final sentiment classification is achieved through confidence-normalized capsule aggregation, enabling balanced performance and traceable decision pathways suitable for large-scale, real-world e-commerce sentiment analysis.

This study follows a structured experimental design grounded in supervised learning and comparative performance evaluation. The research protocol begins with data selection from a large-scale, publicly available Amazon Reviews corpus, followed by preprocessing steps that include data cleaning, token normalization, and stratified dataset partitioning into training, validation, and testing subsets. Model development proceeds through staged representation learning, aspect identification, capsule construction, and probabilistic routing. Performance evaluation is conducted using multiple class-imbalance-aware metrics to ensure robustness and fairness across sentiment categories. Baseline models are implemented under identical data splits and evaluation conditions to maintain experimental consistency. All experiments are repeated under controlled settings to ensure reproducibility and objective comparison.

3.1. Multi-Grained Subword Representation with Context-Aware Encoding

Standard token-level representations often fail to capture the morphological richness and product-specific expressions found in Amazon reviews. Customer sentiments are frequently encoded in complex tokens such as “value-for-money” or “battery-draining,” which contain multiple opinion-rich fragments. Each review token is decomposed into variable-length subword units using byte-pair encoding (BPE) to address this issue. These subwords are mapped into a fixed-dimensional embedding space to preserve fine-grained syntactic and orthographic clues.

$$e_{ij}=E_s(s_{ij}) \quad (1)$$

Where, e_{ij} represents the embedding of the j^{th} subword fragment within the i^{th} token. The matrix E_s denotes the trainable embedding lookup table indexed by subword vocabulary. The subword unit s_{ij} is a sequence derived by segmenting a token t_i through optimized BPE thresholds.

Since tokens are now represented by multiple subwords, a weighted fusion is required to reconstruct the token-level meaning from its subparts. An attention scoring function computes the importance of each subword relative to the rest of the token’s context. This ensures higher influence for semantically loaded fragments while minimizing contribution from auxiliary or transitional units.

$$e_i=\sum_{j=1}^{k_i} \alpha_{ij}\cdot e_{ij} \quad (2)$$

The embedding e_i is the final token representation is computed as a weighted sum over its k_i subword fragments. The scalar α_{ij} represents the normalized attention score for the subword s_{ij}, learned through a feedforward neural attention network trained to highlight emotionally significant sub-lexical patterns.

Each fused token embedding is passed through a lightweight transformer encoder that models inter-token dependencies within the review. The output captures both short-term and long-range contextual associations, which are especially crucial in longer Amazon reviews where sentiment may evolve or reverse. The encoder also enables disambiguation of neutral words when surrounded by strong emotional context.

$$c_i = Transformer(e_i) \quad (3)$$

The vector c_i is the contextually enriched output of a token t_i, generated by the transformer encoder. It reflects positional dependencies and semantic interactions between t_i and the surrounding tokens in the sentence or review. The encoder weights are shared across all review samples to ensure domain consistency.

Sentiment tendencies differ across product categories on Amazon. A statistical prior vector is constructed for each token based on frequency-weighted co-occurrence with positive, neutral, or negative sentiment phrases from historical review corpora. This prior vector is linearly projected and fused with the transformer output through a learnable gate, resulting in a hybrid representation that is sensitive to both text and sentiment frequencies.

$$p_i = \sigma(W_p \cdot e_i + b_p) \quad (4)$$

Where, p_i denotes the prior embedding for token t_i. The matrix W_p and vector b_p are trainable parameters mapping the raw embedding e_i into the sentiment prior space. The sigmoid function σ ensures gating values are constrained within [0,1], controlling fusion intensity in the next operation.

The gated fusion of contextual and prior embeddings is performed to capture both contextual dynamics and domain-specific polarity signals. This linear interpolation enables the model to adjust the emphasis between learned contextual flow and historical sentiment trends, depending on the input structure and learned confidence levels.

$$\mathrm{h}_i = \lambda_i \cdot c_i + (1-\lambda_i) \cdot p_i \quad (5)$$

Where, h_i is the final fused representation for token t_i. The scalar λ_i is a learned gate coefficient between 0 and 1, specific to each token and determined from contextual richness and prior sentiment relevance. This vector h_i encodes all upstream lexical, syntactic, contextual, and probabilistic sentiment features.

Each token representation h_i is evaluated using a Gaussian Mixture Model. This probabilistic clustering determines latent semantic membership, which is especially beneficial in handling ambiguous sentiment tokens or overlapping emotional cues. Instead of complex classification, GMM allows the model to operate over soft assignments that reflect shared sentiment behavior.

$$\gamma_{ic} = \frac{\pi_c \cdot N(\mathrm{h}_i; \mu_c, \Sigma_c)}{\sum_{l=1}^{C} \pi_l \cdot N(\mathrm{h}_i; \mu_l, \Sigma_l)} \quad (6)$$

The scalar γ_{ic} is the posterior probability of vector h_i belonging to Gaussian component c within a mixture of C components. A prior weight parameterizes each Gaussian π_c, mean μ_c, and covariance Σ_c. This probabilistic signal helps direct capsule routing more reliably under uncertain conditions.

The GMM-refined vector is linearly projected into a routing-compatible feature space to feed into the capsule formation layer. This projection applies a non-linear transformation, introducing curvature and dimensional expansion to encourage separation across capsule instantiation paths.

$$v_i = ReLU = (W_r \cdot \mathrm{h}_i + b_r) \quad (7)$$

The output vector v_i becomes the primary input for capsule construction in the next step. The matrix W_r and bias b_r define the affine transformation responsible for expanding the sentiment-encoded token vector into capsule space, where routing dynamics can operate effectively.

3.2. Contextual Embedding via Lightweight Transformer

In sentiment analysis involving Amazon product reviews, capturing semantic variations across different product categories is essential. Static embeddings fail to account for this variation, particularly when sentiments are expressed differently in domains such as electronics, fashion, or books. To address this, the token embeddings v_i from the previous step, the data is fed into a lightweight transformer encoder, optimized for efficiency and domain sensitivity. Each token vector retains its subword composition, emotional signals, and probabilistic weights from the GMM output.

$$h_i^{(0)}=v_i \quad (8)$$

Where, $h_i^{(0)}$ denotes the initial input to the transformer encoder for the i^{th} token. The vector v_i was obtained from Step 1 after combining subword fragments and routing through the GMM projection layer. This token now carries a rich internal structure, but lacks contextual awareness from neighboring tokens or broader document sentiment flow.

The transformer encoder applies self-attention across tokens within a sentence or review segment. This mechanism enables the model to compute the strength of each token's relationship with others in its context. For example, phrases like "not worth the price" require the term "not" to significantly influence the interpretation of "worth" and "price." Using scaled dot-product attention, the self-attention mechanism computes weighted interactions among all token pairs.

$$Attention(Q,K,V)=softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (9)$$

Where, Q, K, and V represent the query, key, and value matrices, respectively, derived from the current layer's token embeddings. The dimension d_k is the size of the key vectors, and the softmax function ensures the attention scores are normalized across each token's context window. This attention construct is essential for resolving sentiment dependencies across non-adjacent phrases in the review.

Transformer models inherently lack knowledge of token order, which is problematic for sentiment interpretation in sentence structures like "Initially great, but later terrible." A deterministic positional encoding is added to each token embedding to encode word order. This encoding introduces temporal or positional signals that enhance the understanding of sentiment trajectories across review segments.

$$\tilde{h}_i^{(0)}=h_i^{(0)}+PE(i) \quad (10)$$

Where, $\tilde{h}_i^{(0)}$ is the positionally encoded input vector for the i^{th} token, and $PE(i)$ denotes the positional encoding vector corresponding to its index i within the sequence. This ensures the transformer encoder captures both the token semantics and its temporal role, which is crucial in Amazon reviews that often contain delayed sentiment shifts.

The transformer encoder processes each positionally encoded token across multiple layers. Each layer involves attention, feedforward computation, and normalization operations. Residual connections are used to enhance gradient stability and convergence during training. The output of each transformer layer becomes the input to the next, forming a deep semantic embedding sequence that reflects the global sentiment structure.

$$h_i^{(l)}=LayerNorm\left(h_i^{(l-1)}+FFN(Attention(Q,K,V))\right) \quad (11)$$

The vector $h_i^{(l)}$ represents the embedding of token i after the l^{th} transformer layer. The layer includes residual addition from the previous state $h_i^{(l-1)}$, and applies a feedforward network (FFN) post-attention. This combination preserves gradient paths and stabilizes semantic evolution as the model deepens.

3.3. Sarcasm-Negation Encoder Using Contrastive Cues

Sarcasm and negation are sentiment-altering constructs that distort the polarity of surrounding words. In Amazon reviews, phrases like "great job sending it broken" or "not bad at all" demand re-interpretation of surface-level sentiment. A contrastive polarity vector g_i is introduced for each token t_i, capturing potential reversals triggered by linguistic markers such as "not," "never," "but," or "still." This vector modulates the token's embedding by highlighting contrastive influence zones.

$$g_i=tanh\left(W_g \cdot h_i+b_g\right) \quad (12)$$

Where, h_i is the context-enhanced embedding from the previous step, and W_g, b_g are trainable weights that transform it into a contrast-aware space. The hyperbolic tangent function introduces non-linearity and restricts the range of g_i, preventing extreme polarity distortion while still enabling controlled reversals.

Once the contrastive polarity vector is computed, it is used to recalibrate the original embedding h_i. This recalibration attenuates the influence of misleading cues and amplifies contextually corrected interpretations. The two vectors are combined using an adaptive weighting gate that learns to balance surface sentiment with its negated or sarcastic counterpart.

$$u_i=\beta_i \cdot g_i+\left(1-\beta_i\right) \cdot h_i \quad (13)$$

The vector u_i is the final sarcasm-negation adjusted representation of token t_i. The scalar β_i is a learnable gate specific to each token, reflecting the likelihood of contrastive influence. Tokens near strong negators or sarcasm pivots tend to have higher β_i values, ensuring sentiment reversal is appropriately integrated into the representation.

In addition to token-level polarity adjustment, convolutional filters scan short windows for abrupt contrast cues. This includes detecting patterns where positive and negative words appear adjacent. A one-dimensional convolutional layer processes each token's local context, producing a scalar sentiment reversal signal s_i that is applied as a multiplicative scale on u_i.

$$z_i = s_i \cdot u_i \quad (14)$$

Where, z_i is the updated token representation post-contrast enhancement. The scalar $s_i$is generated by a convolutional filter applied to a window around u_i, which enhances or suppresses the token's final sentiment weight based on the intensity of local polarity flips. This method allows GMM-TextCaps to isolate sarcasm-encoded contradictions and preserve them in the capsule input stage.

3.4. Emotion-Weighted Multi-Head Attention Layer

The emotional depth of Amazon product reviews varies widely across tokens, ranging from mild approval to extreme dissatisfaction. A scalar emotion signal is derived for each token embedding z_i, intended to magnify emotionally charged terms and suppress neutral tokens that do not impact overall sentiment classification.

$$e_i = tanh(w_e z_i + b_e) \quad (15)$$

Where, e_i is the raw emotion intensity for token i, computed via a dot product with trainable weights w_e and bias b_e. The tanh function ensures that intensity values remain bounded, allowing for both positive and negative activations.

Raw emotion signals are normalized using a softmax operation to ensure comparability across all tokens in the review. This converts unscaled emotion intensities into interpretable attention-like weights.

$$\alpha_i = \frac{exp(e_i)}{\sum_{j=1}^{n} xp(e_j)} \quad (16)$$

The scalar α_i serves as the normalized emotion weight for token i, where n is the total number of tokens. Higher α_i values reflect greater emotional significance in shaping the final sentiment prediction.

The attention mechanism requires separate representations for queries, keys, and values. Each is computed using a linear transformation of the emotion-weighted embedding.

$$q_i = W_q \cdot z_i \quad (17)$$

$$k_i = W_k \cdot z_i \quad (18)$$

$$v_i = W_v \cdot z_i \quad (19)$$

Where, W_q,W_k, and W_v are distinct weight matrices for computing the query, key, and value vectors, respectively. These projections allow the model to relate emotion-weighted tokens within a structured attention space.

Multiple attention heads are deployed in parallel to capture diverse emotional dependencies. The outputs from all heads are concatenated and linearly transformed to maintain dimension consistency across the representation.

$$o_i = Concat(\mathrm{head}_1, \ldots, head_{\mathrm{h}}) \cdot W_o \quad (20)$$

The vector o_i represents the final output for token i after applying multi-head attention. Each head captures a different emotional pattern, such as sarcasm, disappointment, or excitement, commonly found in varied Amazon review contexts.

3.8. Gaussian Mixture Modeling with Domain-Adaptive EM Updating

Amazon reviews often carry sentiment ambiguity, especially with mixed or multi-aspect expressions. To probabilistically model such overlaps, each capsule $r_k^{(final)}$ is evaluated under a Gaussian Mixture Model (GMM). This allows the capsule to belong to multiple sentiment-bearing clusters softly.

$$\gamma_{kc} = \frac{\pi_c \cdot N\left(r_k^{(final)}; \mu_c, \Sigma_c\right)}{\sum_{l=1}^{C} \pi_l \cdot N\left(r_k^{(final)}; \mu_l, \Sigma_l\right)} \quad (33)$$

Where, γ_{kc} is the posterior probability that capsule k belongs to Gaussian component c. The values

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

π_c,μ_c, and Σ_c are the prior, mean, and covariance of component c, respectively, and C is the number of GMM components.

The GMM expects capsule likelihoods to reflect domain-structured sentiment variation. During the EM algorithm's expectation (E) step, each capsule's contribution to every component is updated using the current mixture parameters.

$$Q_{kc}=\gamma_{kc}\cdot logN\left(r_k^{(final)};\mu_c,\Sigma_c\right) \quad (34)$$

This likelihood Q_{kc} estimates the expected assignment weight of capsule k under component c, which guides the updating of GMM statistics in the maximization step.

Each Gaussian component's mean is recalculated using a weighted average of all capsules based on their current posterior probabilities. This is essential for aligning clusters with domain-specific review styles.

$$\mu_c^{new}=\frac{\sum_k \gamma_{kc}\cdot r_k^{(final)}}{\sum_k \gamma_{kc}} \quad (35)$$

This ensures that component centroids move towards clusters that reflect product-specific sentiment organization, such as electronics vs. clothing.

To capture elliptical cluster boundaries, the covariance matrix is updated using the weighted variance of the assigned capsule vectors. This defines the spread and orientation of each Gaussian component.

$$\Sigma_c^{new}=\frac{\sum_k \gamma_{kc}\cdot\left(r_k^{(final)}-\mu_c\right)\left(r_k^{(final)}-\mu_c\right)^{\top}}{\sum_k \gamma_{kc}} \quad (36)$$

This matrix helps distinguish sentiment clusters with subtle variations, including moderately negative vs. strongly negative patterns.

The prior weight for each Gaussian is also updated, reflecting the number of capsules it has effectively captured.

$$\pi_c^{new}=\frac{1}{K}\sum_k \gamma_{kc} \quad (37)$$

Where K is the total number of capsules, this step maintains a balance between dominant and rare sentiment expressions.

3.9. Probabilistic Routing Using GMM Posterior Matching

The routing process begins by aligning each capsule with its most likely sentiment trajectory, inferred through GMM posteriors. Rather than using static routing coefficients, GMM-TextCaps employs a probabilistic approach that dynamically adapts to sentiment ambiguity. A routing coefficient r_{kc} is assigned based on the GMM posterior γ_{kc}, allowing soft assignment into higher-level sentiment capsules.

$$r_{kc}=\gamma_{kc} \quad (38)$$

The scalar r_{kc} represents the routing weight between primary capsule k and class capsule c. This value is directly derived from the posterior probability obtained in Step 3.8. Using these probabilistic scores as routing factors, the model facilitates a smooth transition across sentiment types—ideal for reviews that express mixed sentiments.

Each primary capsule proposes a prediction vector for the higher-level class capsule. The final aggregated input to each class capsule is a weighted sum of these predictions, modulated by routing weights.

$$s_c=\sum_{k=1}^{K} r_{kc}\cdot W_{kc}\cdot r_k^{(final)} \quad (39)$$

Where, s_c is the total input to class capsule c, and W_{kc} is the learned transformation matrix from capsule k to c. The multiplication $W_{kc}\cdot r_k^{(final)}$ produces the prediction vector that represents what capsule k believes capsule c should be. The GMM-based routing ensures predictions are aggregated meaningfully even when sentiment class boundaries overlap.

A non-linear squashing function is applied to prevent dominant capsules from overwhelming weaker ones. This normalizes the output while retaining direction, transforming vector magnitude into a confidence score.

$$v_c=\frac{\|s_c\|^2}{1+\|s_c\|^2}\cdot\frac{s_c}{\|s_c\|} \quad (40)$$

The vector v_c is the activated output of class capsule c. It points in the direction of s_c, but its norm falls within the range (0, 1), ensuring a balanced influence from each class-specific capsule.

3.10. Hierarchical Capsule Aggregation with Temporal-Sentiment Anchoring

Capsule modeling requires multi-level abstraction, primarily when Amazon reviews span multiple sentences with evolving sentiment tone. The first stage of hierarchical aggregation involves reconstructing sentence-level capsules by pooling token-level capsule predictions within each sentence boundary.

$$\mathrm{h}_s=\frac{1}{n_s}\sum_{i\in S_s} v_i \quad (41)$$

The vector h_s represents the sentence-level capsule for sentence s, computed as the average of token-level capsule vectors v_i within a sentence token set S_s. The term n_s is the total number of tokens in sentence s.

Since sentiment often shifts from sentence to sentence, each capsule is assigned a temporal encoding. The positional encoding is computed using a modified sinusoidal function to account for sentence-level granularity and narrative progression.

$$t_s=sin\left(\frac{s}{10000^{2j/d}}\right)+cos\left(\frac{s}{10000^{2j+1/d}}\right) \quad (42)$$

The temporal encoding vector t_s introduces relative sentence position. Index s denotes sentence order and j ranges over the embedding dimensions. The result embeds sentence sequencing for downstream aggregation.

The sentence capsule and temporal vector are fused to encode sentiment progression over time. A gated operation selectively controls how much temporal influence affects each sentence capsule's semantic core.

$$\tilde{\mathrm{h}}_s=\lambda_s\cdot\mathrm{h}_s+(1-\lambda_s)\cdot t_s \quad (43)$$

The fusion vector $\tilde{\mathrm{h}}_s$ is the temporally adjusted sentence capsule. The scalar λ_s is a learned weight determining how much the model trusts semantic content versus positional alignment in shaping sentence meaning.

For full-document interpretation, the sentence capsules are aggregated using a sentiment-attention mechanism. Each sentence is scored by its sentiment intensity to prioritize emotionally dominant segments.

$$\alpha_s=\frac{exp\left(W_a^{\intercal}\cdot\tilde{\mathrm{h}}_s\right)}{\sum_{j=1}^{m} exp\left(W_a^{\intercal}\cdot\tilde{\mathrm{h}}_j\right)} \quad (44)$$

The scalar α_s reflects the attention score assigned to sentence s, computed using a trainable vector w_a. These weights are normalized across all m sentences.

The final document-level capsule is the attention-weighted sum of all temporally modulated sentence capsules. This vector summarizes the entire review, emphasizing time-sensitive and sentiment-intense content.

$$d=\sum_{s=1}^{m}\alpha_s\cdot\tilde{\mathrm{h}}_s \quad (45)$$

The vector d becomes the high-level document capsule representing sentiment abstraction across the full review. It integrates both sequential progression and emotional prominence into a single capsule entity.

3.11. Cross-Aspect Capsule Interaction Modeling via Gated Bilinear Fusion

Amazon reviews often convey interdependent sentiments across aspects—poor packaging may diminish perceived value, even if the product quality is high. The model compares each pair of aspect capsules to represent such relational effects. A similarity matrix is computed by measuring cosine proximity between all capsule pairs, determining which aspects influence each other structurally.

$$s_{pq}=\frac{r_p^{\intercal}\cdot r_q}{\left\|r_p\right\|\cdot\left\|r_q\right\|} \quad (46)$$

In this formulation, s_{pq} quantifies the similarity between capsules p and q. Capsules with high cosine similarity are assumed to express related sentiments, such as a product's durability correlating with price satisfaction or brand trust.

Once similarities are determined, a gate controls whether such relational influence can flow between the capsules. Each interaction weight is passed through a gating function, allowing the model to block spurious or low-confidence relational effects.

$$g_{pq}=\sigma\left(w_s\cdot s_{pq}+b_s\right) \quad (47)$$

Where, g_{pq} is the gate scalar between aspect capsules p and q. The trainable parameters w_s and b_s regulate how similarity values are converted into flow control scores. Only positively correlated or semantically linked capsules are allowed to transmit influence.

Each capsule is adjusted using the gated interaction with every other capsule. A bilinear combination fuses the semantic features of both participating aspect capsules into a composite influence term.

$$b_{pq}=g_{pq}\cdot\left(r_p W_b r_q\right) \quad (48)$$

The scalar b_{pq} represents the bilinear interaction score between capsules p and q. The matrix W_b is a learned bilinear mapping that amplifies aligned features while reducing irrelevant traits. These pairwise interactions form the basis for adjusting capsule activations.

The final representation of each aspect capsule is refined by incorporating the weighted interactions from all other capsules. This ensures that each sentiment capsule reflects its internal signal and the influence it receives from correlated capsules.

$$\tilde{r}_p=r_p+\sum_{q\neq p} b_{pq}\cdot r_q \quad (49)$$

The updated capsule $\tilde{r}_p$ contains original features enriched with relational information from other aspect capsules. This construction improves prediction quality when aspect-specific sentiments co-occur or conflict in the review structure.

3.12. Document-Level Capsule Composition Using GMM-Aspect Fusion

After cross-aspect interactions, each aspect capsule $\tilde{r}_p$ contains sentiment content that reflects both individual and relational features. Each capsule is linearly projected into a shared class capsule space to align these capsules for the final sentiment class prediction. This projection standardizes dimension and prepares the vectors for probabilistic fusion across Gaussian components.

$$a_p=W_a+\tilde{r}_p+b_a \quad (50)$$

Where, a_p is the class-space representation of the p^{th} aspect capsule. The matrix W_a and bias b_a are shared across all aspect capsules, maintaining consistent mapping into class sentiment directions such as positive, negative, neutral, or mixed.

Each projected capsule is weighted by its previously calculated GMM posterior γ_{pc}. This ensures that capsule contributions to the document-level representation are scaled based on probabilistic alignment to sentiment clusters. Capsules that are more confidently aligned with a sentiment class are given more substantial influence.

$$f_c=\sum_{p=1}^{P}\gamma_{pc}\cdot a_p \quad (51)$$

The vector f_c represents the fused sentiment signal for class c, where P is the total number of refined aspect capsules. The summation aggregates all aspect-level signals according to GMM-based confidence levels.

The fused class vectors are normalized through a margin-preserving squashing function, which converts them into final sentiment capsule activations. This procedure ensures class capsules express both sentiment intensity and probabilistic precision.

$$d_c=\frac{\|f_c\|^2}{1+\|f_c\|^2}\cdot\frac{f_c}{\|f_c\|} \quad (52)$$

The vector d_c is the activated sentiment capsule for class c. It encapsulates all available information across aspect capsules and GMM signals into a clean final representation for downstream classification.

The final document representation is formed by concatenating all class capsule activations. This yields a document-level vector that preserves class separability and supports high-confidence multi-class classification across diverse review patterns.

$$D=[d_1 \| d_2 \| \ldots \| d_C] \quad (53)$$

The vector D represents the complete document capsule containing all sentiment class vectors. This becomes the final feature used for classification and interpretability.

3.13. Sentiment Capsule Activation with GMM Confidence Normalization

Final sentiment classification benefits from weighing the class capsule output not only by magnitude but also by probabilistic certainty. Each activated class capsule d_c is multiplied by a normalized GMM confidence score. This ensures that sentiment capsules with stronger posterior support exert greater influence on the final sentiment decision, particularly in semantically vague reviews or those containing conflicting expressions.

$$\phi_c = \sum_{p=1}^{P} \gamma_{pc} \quad (54)$$

Where, ϕ_c is the aggregated posterior mass of class c across all P aspect capsules. This scalar represents how strongly the GMM model aligns various aspect capsules with a given sentiment class. Higher values of ϕ_c imply greater confidence in class c being present in the review.

The posterior mass values ϕ_c are normalized across all sentiment classes to ensure fairness and maintain probability constraints. This normalization forms the confidence scaling factor η_c, applied to each class capsule.

$$\eta_c = \frac{\phi_c}{\sum_{j=1}^{C} \phi_j} \quad (55)$$

The scalar η_c reflects the relative GMM confidence of class c, scaled to lie between 0 and 1. When multiple classes are closely aligned with aspect capsules, this prevents any one capsule from dominating due to the enormous raw norm magnitude alone.

Each class capsule vector is then recalibrated by its confidence scaling factor η_c, forming the final prediction-ready capsule z_c. This adjustment enhances interpretability and prevents overfitting to outlier capsule activations.

$$z_c = \eta_c \cdot d_c \quad (56)$$

The vector z_c is the final GMM-normalized capsule for sentiment class c. It incorporates semantic richness and confidence-driven weight refinement, essential for accurate sentiment interpretation in real-world reviews.

The final class prediction is made by selecting the capsule with the highest adjusted norm. This margin-based classification rule selects the most confident and semantically complete sentiment representation across all classes.

$$\hat{y} = \arg\max_c \|z_c\| \quad (57)$$

The predicted label $\hat{y}$ is the class with the highest confidence-weighted capsule magnitude. This decision rule balances sentiment strength with GMM-derived belief in the class's alignment, ensuring reliable predictions even in complex, multi-aspect reviews.

3.14. Composite Loss Function with GMM-Weighted Focal Margin + Contrastive Regularizer

In Amazon product reviews, sentiment imbalance is frequently observed, with extreme classes (such as "very positive" or "very negative") dominating the data. The focal margin loss enhances model sensitivity to complicated or ambiguous reviews, addressing this issue. The capsule norm $\|z_c\|$is interpreted as the probability of sentiment class c, and margin thresholds m^+,m^- guide classification margins.

$$L_c^{margin} = y_c \cdot max(0, m^+ - \|z_c\|)^2 + (1 - y_c) \cdot max(0, \|z_c\| - m^-)^2 \quad (58)$$

Where, y_c is a binary indicator showing whether class c is the correct sentiment label. The term penalizes misclassified capsule magnitudes based on fixed margin thresholds.

To further suppress the impact of well-classified examples and amplify attention on hard instances, a dynamic focal weight α_c is introduced. This scaling is based on capsule confidence and GMM alignment.

$$\alpha_c = (1 - z_c)^{\gamma} \cdot \eta_c \quad (59)$$

Where γ controls the focus on complex examples, while η_c comes from the GMM confidence scores. The final focal-margin loss is computed by weighting the margin penalty:

$$L_c^{focal} = \alpha_c \cdot L_c^{margin} \quad (60)$$

This function prioritizes capsules with low norm and high class ambiguity, enabling refined learning on underrepresented or confusing sentiment categories.

Sentiment capsules often exhibit directional closeness, especially between adjacent categories like "neutral" and "mildly negative." A contrastive regularizer is used to separate class embeddings. It penalizes similarity between incorrect capsule vectors and the actual class vector.

$$L_{ij}^{contrast}=I\left[y_i\neq y_j\right]\cdot exp\left(-\|z_i-z_j\|^2\right) \quad (61)$$

Where, y_i and y_j are labels of samples i and j, and $I[\cdot]$ is an indicator function. This term suppresses alignment between capsules of different classes, particularly when their geometric distance is small.

To control the magnitude of contrastive regularization concerning focal loss, a scaling coefficient λ is applied. This ensures the balance between classification refinement and capsule separation objectives.

$$L^{contrast}=\frac{1}{B^2}\sum_{i,j}L_{ij}^{contrast} \quad (62)$$

Where B is the batch size, this formulation ensures contrast regularization remains consistent across mini-batches and does not overpower classification loss.

The overall loss for the GMM-TextCaps model is constructed by summing the focal margin loss over all sentiment classes and incorporating contrastive feedback with a tunable coefficient.

$$L_{total}=\sum_{c=1}^{C}L_c^{focal}+L^{contrast} \quad (63)$$

This unified formulation optimizes accurate class separation and robust handling of soft sentiment boundaries.

3.15. Aspect-Level Output with Capsule Heatmaps and Component Tracebacks

The final output layer generates activation heatmaps for each aspect token, supporting the interpretability of sentiment predictions, particularly for detailed Amazon reviews. These scores indicate the extent to which each token contributes to the sentiment orientation of its corresponding aspect capsule. The score is derived from the dot product between token vectors and their projected capsule representation.

$$\mathrm{h}_i^p=o_i^{\mathrm{T}}\cdot\tilde{r}_p \quad (64)$$

The scalar h_i^p denotes the activation intensity of token i toward the refined aspect capsule $\tilde{r}_p$. Higher values indicate greater influence on the capsule's sentiment decision, allowing users and analysts to trace sentiment sources at the lexical level.

In addition to local token traceability, each aspect capsule is linked to its most dominant GMM component. This facilitates understanding how the capsule's behavior aligns with sentiment clusters derived from historical training distributions.

$$c_p^*=\arg\max_c \gamma_{pc} \quad (65)$$

The variable c_p^* indicates the most probable GMM component associated with aspect capsule p, based on posterior probabilities γ_{pc}. This allows the model to output sentiment class and the underlying distributional context supporting that classification.

For practical multi-aspect sentiment output, the final sentiment label for each aspect is derived from its confidence-normalized capsule activation. The norm of each confidence-scaled capsule determines the dominant sentiment type among predefined classes (e.g., positive, negative, neutral).

$$\hat{y}_p=arg\max_c\|z_{pc}\| \quad (66)$$

Where, $\hat{y}_p$ is the sentiment label predicted for aspect p, determined by selecting the capsule with the highest adjusted norm across all sentiment classes c.

Algorithm 1: GMM-TextCaps

Input:

- Amazon product review text corpus $R=\{r_1,r_2,\ldots,r_N\}$

Output:

- Predicted sentiment label $\hat{y}$ for each review and aspect-wise sentiment outputs

Procedure:

1. Decompose each review token into subwords and embed them using emotion-rich vector composition.
2. Contextualize token embeddings using a lightweight transformer encoder with positional awareness.
3. Detect polarity reversals and sarcasm using contrastive cues and recalibrate token vectors.
4. Apply emotion-weighted multi-head attention to amplify emotionally

charged tokens.
5. Identify product-specific aspect terms using a CRF decoder with tag sequence constraints.
6. Construct primary capsules by gating aspect-relevant and emotionally weighted token vectors.
7. Inject sentence-position temporal encodings into capsule seeds to capture sentiment drift.
8. Use GMM to perform soft clustering of capsules and update parameters via the EM algorithm.
9. Perform probabilistic routing from primary to class capsules using GMM posterior weights.
10. Aggregate capsules hierarchically from token to sentence to document level using temporal anchoring.
11. Model cross-aspect sentiment influence through gated bilinear fusion between aspect capsules.
12. Fuse refined aspect capsules into class capsules using GMM-aligned projections for document summary.
13. Normalize class capsule activations with GMM confidence scores and derive the final prediction vector.
14. Compute composite loss using GMM-weighted focal margin loss and inter-class contrastive regularization.
15. Generate interpretable outputs with aspect-wise sentiment labels, token heatmaps, and GMM tracebacks.

4. DATASET

The Clothing, Shoes, and Jewelry segment within the Amazon Reviews '23 corpus comprises 22.6 million distinct reviewers who provide feedback on 7.2 million unique products. Over its collection period, this category has amassed 66.0 million ratings, reflecting a wide distribution of satisfaction levels. The review text volume spans approximately 2.6 billion tokens, capturing rich linguistic variation, while the associated metadata contributes an additional 5.9 billion tokens, providing extensive contextual detail. This dual-layer structure enables in-depth exploration of consumer opinion, blending descriptive narratives with structured identifiers. For the present research, a random sample of 4,51,478 reviews was extracted to balance computational feasibility with representational diversity. The sample was cleaned to remove duplicates and incomplete entries, and partitioned into 80% training, 10% validation, and 10% testing sets. This configuration maintains proportionality across sentiment classes, ensuring that analyses remain consistent with real-world sentiment distribution within the fashion and accessories market.

Table 2. Feature Description

Attribute	Description
User	Number of unique reviewers contributing to the category.
Item	Number of distinct products receiving reviews.
Rating	Total count of ratings assigned by users.
R_Token	Aggregate token count derived from all review texts.
M_Token	Aggregate token count derived from metadata records.

5. RESULTS AND DISCUSSION

The results and discussion section presents a detailed evaluation of the proposed GMM-Enhanced TextCaps model against established baselines using multiple performance metrics. The analysis focuses on assessing classification balance, correlation with ground truth, precision–recall trade-offs, and robustness under class imbalance. Each metric offers a distinct perspective on the model's ability to capture nuanced sentiment patterns while maintaining interpretability and stability across diverse review structures. Comparative results with TF-IDF + BERT and ENLPPR-ICFFO highlight improvements in both quantitative measures and qualitative performance attributes. The following subsections report metric-specific outcomes and critically examine the factors influencing performance gains or limitations.

5.1. Balanced Accuracy Evaluation

Balanced accuracy captures the mean of sensitivity and specificity, serving as a robust evaluation metric when class distributions are uneven, as it prevents performance assessment from being dominated by the largest class. Figure 1, where balanced accuracy (%) is pltted on the y-axis and classification algorithms are marked on the x-axis, reveals essential distinctions between the models assessed. TF-IDF + BERT achieves 53.150%. Sparse representation, static term weighting, and the inability to model word order or semantic overlap all contribute to its weak results. Domain shifts or variations in text length further magnify these limitations, making robust classification unfeasible across classes. For ENLPPR-ICFFO, a score of 62.389% is realized. Although recurrent processing allows for some context retention, shallow memory, sensitivity to initializations, and instability in the optimizer's convergence leave persistent weaknesses—especially for sequences exhibiting long-range dependencies or high noise levels. In sharp contrast, GMM-TextCaps secures 72.507%. Here, subword segmentation, dynamic routing, and Gaussian mixture-based sentiment clustering collectively enable superior discrimination, capturing granular structure and mitigating the ambiguity seen in the alternative approaches. This leap in balanced accuracy not only demonstrates structural robustness but also underlines the method's ability to adapt to realistic sentiment composition.

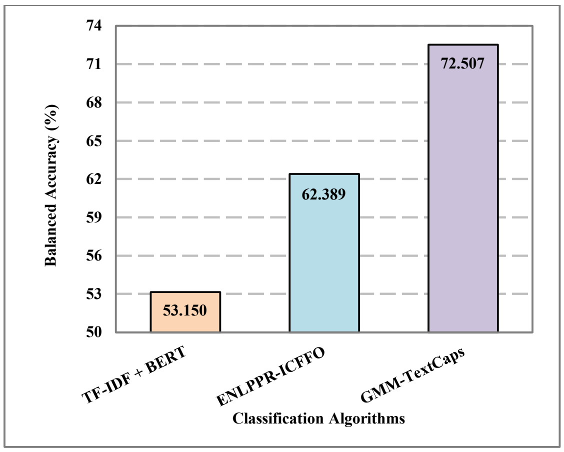

Figure 1. Balanced Accuracy

5.2. Matthews Correlation Coefficient Assessment

The Matthews Correlation Coefficient (MCC) represents the geometric mean of precision and recall for both classes, providing a comprehensive assessment of classifier prediction quality, particularly in cases of class imbalance. Figure 2 presents the MCC (%) on the x-axis and the classification algorithms—TF-IDF + BERT, ENLPPR-ICFFO, and GMM-TextCaps—on the y-axis, providing a comparative view. TF-IDF + BERT achieves a modest 6.299%, revealing profound insensitivity to non-linear decision boundaries, an inability to manage synonymy and polysemy, and susceptibility to overfitting or noise due to high-dimensional sparse features. ENLPPR-ICFFO reaches 24.768%. While it displays improvement, the architecture suffers from the recurrent model's limited memory span and the optimizer's convergence issues, which reduce its ability to reliably decipher intricate sentiment signals—especially in lengthy or noisy sequences. GMM-TextCaps stands out with 45.086%. Its superior MCC can be attributed to its ability to maintain aspect-level distinctions, cluster sentiment boundaries softly, and route evidence dynamically, making its outputs markedly more correlated with proper labels and more robust to class imbalance than the alternatives.

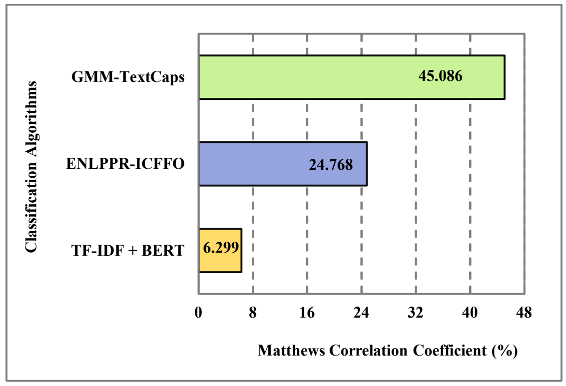

Figure 2. Matthews Correlation Coefficient

5.3. F1 Score Comparison

F1-score, the harmonic mean of precision and recall, encapsulates a classifier's efficacy in balancing false positives and false negatives. Figure 3 visualizes comparative F1-scores on the y-axis, set against the classification algorithms on the x-axis. TF-IDF + BERT yields a value of 49.102%, reflecting limitations stemming from its static term weighting, sparse feature set, and lack of attention to word order or semantic context, which compound the difficulty in optimizing both recall and precision. ENLPPR-ICFFO advances modestly to 49.293%, yet this gain remains constrained by challenges such as vanishing gradients in the Elman recurrent structure and chaos-induced instability that undermine consistent feature utilization. GMM-TextCaps achieves a score of 49.656%, marking the highest score among the models shown. Its advantage arises from BPE segmentation, which preserves sentiment-rich subwords, dynamic capsule routing that aggregates discriminative evidence, and GMM-based clustering near polarity thresholds, which enhances both recall and precision. The progression in values demonstrates that the proposed GMM-TextCaps method addresses critical shortfalls of the alternatives, resulting in measurable improvement.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

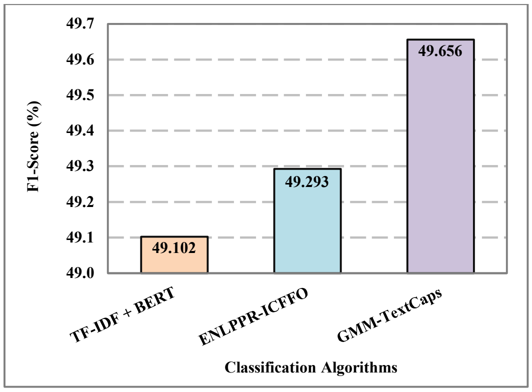

Figure 3. F1 Score

5.4. G-Mean Interpretation

G-Mean evaluates the geometric mean of sensitivity and specificity, providing clear insight into the model's balance on class distributions. Figure 4 illustrates G-Mean (%) along the y-axis, with classification algorithms featured on the x-axis. TF-IDF + BERT attains only 53.119%, which demonstrates that sparse, static features struggle to distinguish between classes, particularly under shifts in data distribution or with short texts. ENLPPR-ICFFO improves to 62.386%, yet recurrent limitations, optimizer convergence instability, and inadequate handling of complex dependencies cap its performance, leaving significant ground uncovered. GMM-TextCaps displays a marked advantage with 72.452%. This increase emerges through the hybrid's BPE-driven subword integrity, aspect-wise capsule routing, and soft probability assignments of Gaussian mixture clustering. GMM-TextCaps not only preserves discriminative features across data regimes but also tempers overconfidence at sentiment boundaries. The data reinforce that this architecture addresses structural and contextual nuances that are missed by TF-IDF + BERT and ENLPPR-ICFFO, producing a robust balance across class outcomes.

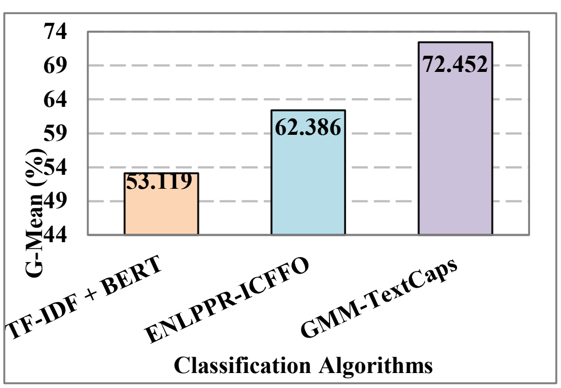

Figure 4. G-Mean

5.5. Critical Success Index Review

Critical Success Index (CSI) calculates the proportion of correctly classified positive cases relative to all cases that were either detected or misclassified as positives. Figure 5 portrays CSI on the x-axis (%), matched against the classification algorithms shown on the y-axis. The TF-IDF + BERT approach achieves a CSI of 36.167%, indicating fundamental shortcomings in capturing nuanced sentiment due to its context-agnostic, sparse representation and static weighting, which limit the detection of meaningful true positives. ENLPPR-ICFFO achieves 44.871%. Despite outperforming TF-IDF + BERT, persistent limitations, including shallow contextual retention and sensitivity to the optimizer, restrict further advancement. GMM-TextCaps achieves a notable CSI of 57.851%. Its progress stems from the ability to retain sentiment-rich subwords and suppress noise through dynamic capsule routing combined with Gaussian mixture modeling, ensuring improved discrimination between relevant and irrelevant signals. The higher CSI indicates that the model more accurately identifies correct cases without an undue increase in false alarms, outperforming its counterparts.

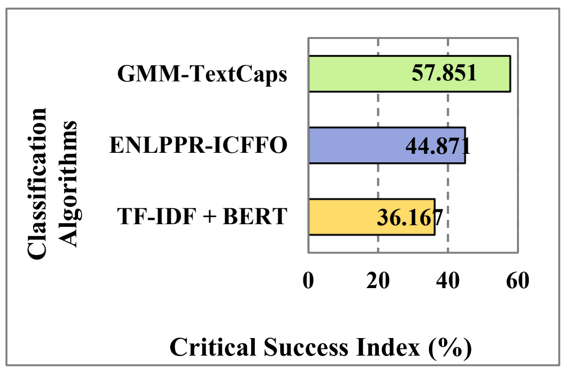

Figure 5. Critical Success Index

5.6 Comparative Analysis and Model Limitations

Several published studies report sentiment analysis results using transformer-based classifiers, recurrent neural networks, and neuro-evolutionary models on Amazon or review-based datasets. These works demonstrate improvements in conventional metrics, yet reported gains often diminish under class imbalance or mixed-sentiment conditions. Comparative evaluation shows that such models struggle to maintain consistent sensitivity across minority sentiment classes, leading to inflated accuracy with limited robustness. In contrast, the proposed model demonstrates stronger balance across sensitivity and specificity, reflected through higher balanced accuracy, G-Mean, and MCC values. This suggests that probabilistic capsule routing contributes measurable stability when sentiment boundaries overlap.

Despite these improvements, the present work has identifiable limitations. The architecture introduces higher computational overhead than standard transformer classifiers, which may constrain scalability in low-resource deployment scenarios. The model also focuses exclusively on textual data, leaving multimodal sentiment signals unexploited in the current evaluation. Performance remains dependent on accurate aspect extraction, meaning errors in aspect tagging can propagate into capsule construction. These limitations indicate that future extensions should explore efficiency optimization, multimodal integration, and robustness to aspect detection noise. Acknowledging these constraints ensures that the reported findings are interpreted as a step toward interpretable and uncertainty-aware sentiment modeling rather than a complete solution.

5.7 Research Challenges and Open Issues

Despite the demonstrated effectiveness of the proposed framework, several research challenges remain open. Probabilistic capsule routing introduces additional computational cost, raising concerns related to efficiency and scalability in large-scale deployments. Aspect detection accuracy remains a critical dependency, as errors at this stage influence downstream capsule formation. The current study focuses exclusively on textual sentiment signals, leaving cross-modal sentiment fusion unexplored. Generalization across domains and languages also requires further validation. Addressing these challenges presents open research opportunities related to efficiency optimization, multimodal integration, adaptive capsule structures, and broader domain transferability.

6. CONCLUSION

The study confirms that sentiment interpretation in user-generated reviews benefits from architectures capable of modelling emotional variation without collapsing multiple perspectives into a single, oversimplified view. The proposed GMM-Enhanced TextCaps framework demonstrates that integrating structured aspect-level representation with probabilistic decision boundaries yields consistent performance gains across imbalanced sentiment classes. Evaluation results show that the model maintains stability and robustness even in the presence of linguistically complex patterns, indicating its adaptability to diverse expression styles. Achieving a balanced accuracy of 72.507% and a G-Mean of 72.452%, GMM-Enhanced TextCaps effectively balances precision and recall, essential for reliable commercial analytics. Beyond numeric improvements, the model produces interpretable outputs that enable tracing sentiment outcomes to specific input elements, reinforcing accountability in automated sentiment analysis. This combination of quantitative effectiveness and qualitative transparency supports its suitability for deployment in platforms where sentiment-driven decisions have significant financial and reputational implications.

Future research may extend the proposed GMM-Enhanced TextCaps framework by incorporating multimodal sentiment cues, including product images, star ratings, and reviewer metadata, enabling richer sentiment reasoning beyond textual evidence. Exploration of efficiency-oriented adaptations, including lightweight capsule routing, parameter sharing, or compression strategies, may support scalability in large-scale or real-time deployment settings. Broadening evaluation to multilingual and cross-domain datasets would allow assessment of robustness across linguistic structures and domain-specific sentiment expressions. Joint optimization of aspect extraction and sentiment capsule construction may further reduce error propagation and strengthen end-to-end reliability. Investigation of adaptive probabilistic clustering strategies, including dynamic mixture component selection, may refine sentiment boundary modeling and improve uncertainty handling in highly ambiguous reviews.

REFERENCES

[1] M. S. M. Suhaimin, M. H. A. Hijazi, E. G. Moung, P. N. E. Nohuddin, S. Chua, and F. Coenen, "Corrigendum to 'Social media sentiment analysis and opinion mining in public security: Taxonomy, trend analysis, issues and future directions' [J. King Saud Univ. – Comput. Inform. Sci. 35(9) (2023) 101776]," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 6, p. 102121, 2024, doi: https://doi.org/10.1016/j.jksuci.2024.102121.
[2] P. Wang, L. Tao, M. Tang, L. Wang, Y. Xu,

and M. Zhao, "Incorporating syntax and semantics with dual graph neural networks for aspect-level sentiment analysis," *Eng. Appl. Artif. Intell.*, vol. 133, p. 108101, 2024, doi: https://doi.org/10.1016/j.engappai.2024.108101.

[3] M. Selimovic, A. A. Almisreb, and S. Amanzholova, "The Role of Sentiment Analysis in Brand Management and Marketing: A Comparative Study," *Procedia Comput. Sci.*, vol. 251, pp. 579–584, 2024, doi: https://doi.org/10.1016/j.procs.2024.11.152.

[4] M. George and R. Murugesan, "Improving sentiment analysis of financial news headlines using hybrid Word2Vec-TFIDF feature extraction technique," *Procedia Comput. Sci.*, vol. 244, pp. 1–8, 2024, doi: https://doi.org/10.1016/j.procs.2024.10.172.

[5] M. Tang, K. Yang, L. Tao, M. Zhao, and W. Zhou, "Multi-granularity enhanced graph convolutional network for aspect sentiment triplet extraction," *Big Data Res.*, vol. 39, p. 100506, 2025, doi: https://doi.org/10.1016/j.bdr.2025.100506.

[6] A. Chauhan, A. Sharma, and R. Mohana, "A Pre-Trained Model for Aspect-based Sentiment Analysis Task: using Online Social Networking," *Procedia Comput. Sci.*, vol. 233, pp. 35–44, 2024, doi: https://doi.org/10.1016/j.procs.2024.03.193.

[7] Z. Li and Z. Zou, "Punctuation and lexicon aid representation: A hybrid model for short text sentiment analysis on social media platform," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 3, p. 102010, 2024, doi: https://doi.org/10.1016/j.jksuci.2024.102010.

[8] K. Jiang, B. Cao, and J. Fan, "A Robust Framework for Multimodal Sentiment Analysis with Noisy Labels Generated from Distributed Data Annotation," *C. - Comput. Model. Eng. Sci.*, vol. 139, no. 3, pp. 2965–2984, 2024, doi: https://doi.org/10.32604/cmes.2023.046348.

[9] C. Ouni, E. Benmohamed, and H. Ltifi, "Deep learning-based Soft word embedding approach for sentiment analysis," *Procedia Comput. Sci.*, vol. 246, pp. 1355–1364, 2024, doi: https://doi.org/10.1016/j.procs.2024.09.720.

[10] M. Suzuki, Y. Ishikawa, M. Teraguchi, and H. Sakaji, "Sentiment works in small-cap stocks: Japanese stock's sentiment with language models," *Int. J. Inf. Manag. Data Insights*, vol. 5, no. 1, p. 100318, 2025, doi: https://doi.org/10.1016/j.jjimei.2024.100318.

[11] S. Panickar, R. Sinha, V. Chawla, O. Singh, and O. Londhe, "Sentiment Analysis of Custom Speech Corpus: A proof of concept for NLP," *Procedia Comput. Sci.*, vol. 244, pp. 220–228, 2024, doi: https://doi.org/10.1016/j.procs.2024.10.195.

[12] L. Deng, B. Liu, and Z. Li, "Multimodal Sentiment Analysis Based on a Cross-Modal Multihead Attention Mechanism," *Comput. Mater. Contin.*, vol. 78, no. 1, pp. 1157–1170, 2024, doi: https://doi.org/10.32604/cmc.2023.042150.

[13] X. Wang and D. Zhao, "A comparative experimental study of citation sentiment identification based on the Athar-Corpus," *Data Sci. Inf.*, 2025, doi: https://doi.org/10.1016/j.dsim.2025.01.008.

[14] A.-G. Ursa and L. S. Dioşan, "Bridging Linguistic Gaps: SENTIROM's Approach to Romanian Sentiment Analysis," *Procedia Comput. Sci.*, vol. 246, pp. 616–625, 2024, doi: https://doi.org/10.1016/j.procs.2024.09.475.

[15] A. Abdedaiem, A. H. Dahou, M. A. Cheragui, and B. Mathiak, "FASSILA: A Corpus for Algerian Dialect Fake News Detection and Sentiment Analysis," *Procedia Comput. Sci.*, vol. 244, pp. 397–407, 2024, doi: https://doi.org/10.1016/j.procs.2024.10.214.

[16] B. Paneru, B. Thapa, and B. Paneru, "Sentiment analysis of movie reviews: A flask application using CNN with RoBERTa embeddings," *Syst. Soft Comput.*, vol. 7, p. 200192, 2025, doi: https://doi.org/10.1016/j.sasc.2025.200192.

[17] J. Li, Y. Huang, Y. Lu, L. Wang, Y. Ren, and R. Chen, "Sentiment Analysis Using E-Commerce Review Keyword-Generated Image with a Hybrid Machine Learning-Based Model," *Comput. Mater. Contin.*, vol. 80, no. 1, pp. 1581–1599, 2024, doi: https://doi.org/10.32604/cmc.2024.052666.

[18] N. Alsaleh, R. Alnanih, and N. Alowidi, "Hybrid Deep Learning Approach for Automating App Review Classification: Advancing Usability Metrics Classification with an Aspect-Based Sentiment Analysis Framework," *Comput. Mater. Contin.*, vol. 82, no. 1, pp. 949–976, 2025, doi: https://doi.org/10.32604/cmc.2024.059351.

[19] J. Zhang, J. Zhu, H. Sun, X. Zhang, and J. Liu, "Saccade inspired Attentive Visual

Patch Transformer for image sentiment analysis," *Appl. Soft Comput.*, vol. 174, p. 112963, 2025, doi: https://doi.org/10.1016/j.asoc.2025.112963.

[20] E. Christodoulaki, M. Kampouridis, and M. Kyropoulou, "A novel strongly-typed Genetic Programming algorithm for combining sentiment and technical analysis for algorithmic trading," *Knowledge-Based Syst.*, vol. 311, p. 113054, 2025, doi: https://doi.org/10.1016/j.knosys.2025.113054.

[21] Y. Wang, H. Jian, J. Zhuang, H. Guo, and Y. Leng, "SSLMM: Semi-Supervised Learning with Missing Modalities for Multimodal Sentiment Analysis," *Inf. Fusion*, vol. 120, p. 103058, 2025, doi: https://doi.org/10.1016/j.inffus.2025.103058.

[22] B. Sun, L. Jia, Y. Cui, N. Wang, and T. Jiang, "Conv-Enhanced Transformer and Robust Optimization Network for robust multimodal sentiment analysis," *Neurocomputing*, vol. 634, p. 129842, 2025, doi: https://doi.org/10.1016/j.neucom.2025.129842.

[23] Y. Zheng *et al.*, "Public Attitudes Toward Violence Against Doctors: Sentiment Analysis of Chinese Users," *JMIR Med. Informatics*, vol. 13, 2025, doi: https://doi.org/10.2196/63772.

[24] K. Aziz *et al.*, "Enhanced UrduAspectNet: Leveraging Biaffine Attention for superior Aspect-Based Sentiment Analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 9, p. 102221, 2024, doi: https://doi.org/10.1016/j.jksuci.2024.102221.

[25] Z. Li *et al.*, "Representation distribution matching and dynamic routing interaction for multimodal sentiment analysis," *Knowledge-Based Syst.*, vol. 316, p. 113376, 2025, doi: https://doi.org/10.1016/j.knosys.2025.113376.

[26] M. Ramzy and B. Ibrahim, "User satisfaction with Arabic COVID-19 apps: Sentiment analysis of users' reviews using machine learning techniques," *Inf. Process. Manag.*, vol. 61, no. 3, p. 103644, 2024, doi: 10.1016/j.ipm.2024.103644.

[27] B. Priya Kamath *et al.*, "Comprehensive Analysis of Word Embedding Models and Design of Effective Feature Vector for Classification of Amazon Product Reviews," *IEEE Access*, vol. 13, no. February, pp. 25239–25255, 2025, doi: 10.1109/ACCESS.2025.3536631.

[28] M. Alghamdi *et al.*, "Improved chaotic fruit fly optimization with elman neural network based data-driven approach on product reviews classification," *Alexandria Eng. J.*, vol. 127, no. February, pp. 628–641, 2025, doi: 10.1016/j.aej.2025.05.042.

[29]. R. Jaganathan, S. Mehta, and R. Krishan, "Preface," Bio-Inspired Intelligence for Smart Decision-Making, pp. xix–xx, 2024.

[30]. R. Jaganathan, S. Mehta, and R. Krishan, "Preface," Intelligent Decision Making Through Bio-Inspired Optimization, pp. xiii–xvi, 2024.

[31]. B. Suchitra, R. Karthikeyan, J. Ramkumar, and V. Valarmathi, "Enhancing Recurrent Neural Network Performance for Latent Autoimmune Diabetes Detection (Lada) Using Exocoetidae Optimization," J Theor Appl Inf Technol, vol. 103, no. 5, pp. 1645–1667, 2025.

[32]. R. Jaganathan, K. Rajendran, and P. S. Ponnukumar, "Peregrine Falcon Optimization Routing Protocol (PFORP) for Achieving Ultra-Low Latency and Boosted Efficiency in 6G Drone Ad-Hoc Networks (DANET)," International Journal of Computing and Digital Systems, vol. 17, no. 1, pp. 1–18, 2025, doi: 10.12785/ijcds/1571111848.

[33]. J. Ramkumar, V. Valarmathi, and R. Karthikeyan, "Optimizing Quality of Service and Energy Efficiency in Hazardous Drone Ad-Hoc Networks (DANET) Using Kingfisher Routing Protocol (KRP)," International Journal of Engineering Trends and Technology, vol. 73, no. 1, pp. 410–430, 2025, doi: 10.14445/22315381/IJETT-V73I1P135.

[34]. J. Ramkumar, B. Varun, V. Valarmathi, D. R. Medhunhashini, and R. Karthikeyan, "Jaguar-Based Routing Protocol (JRP) For Improved Reliability and Reduced Packet Loss in Drone Ad-Hoc Networks (DANET)," J Theor Appl Inf Technol, vol. 103, no. 2, pp. 696–713, 2025.

[35]. B. Suchitra, J. Ramkumar, and R. Karthikeyan, "Frog Leap Inspired Optimization-Based Extreme Learning Machine for Accurate Classification of Latent Autoimmune Diabetes in Adults (LADA)," J Theor Appl Inf Technol, vol. 103, no. 2, pp. 472–494, 2025.

[36]. J. Ramkumar and R. Vadivel, "CSIP—cuckoo search inspired protocol for routing

in cognitive radio ad hoc networks," in Advances in Intelligent Systems and Computing, Springer Verlag, 2017, pp. 145–153. doi: 10.1007/978-981-10-3874-7_14.

[37]. R. Jaganathan and R. Vadivel, "Intelligent Fish Swarm Inspired Protocol (IFSIP) for Dynamic Ideal Routing in Cognitive Radio Ad-Hoc Networks," International Journal of Computing and Digital Systems, vol. 10, no. 1, pp. 1063–1074, 2021, doi: 10.12785/ijcds/100196.

[38]. K. S. J. Marseline, J. Ramkumar, and D. R. Medhunhashini, "Sophisticated Kalman Filtering-Based Neural Network for Analyzing Sentiments in Online Courses," in Smart Innovation, Systems and Technologies, A. K. Somani, A. Mundra, R. K. Gupta, S. Bhattacharya, and A. P. Mazumdar, Eds., Springer Science and Business Media Deutschland GmbH, 2024, pp. 345–358. doi: 10.1007/978-981-97-3690-4_26.

[39]. J. Ramkumar, A. Senthilkumar, M. Lingaraj, R. Karthikeyan, and L. Santhi, "Optimal Approach for Minimizing Delays in IoT-Based Quantum Wireless Sensor Networks Using NM-Leach Routing Protocol," J Theor Appl Inf Technol, vol. 102, no. 3, pp. 1099–1111, 2024.

[40]. J. Ramkumar and R. Vadivel, "Improved frog leap inspired protocol (IFLIP) – for routing in cognitive radio ad hoc networks (CRAHN)," World Journal of Engineering, vol. 15, no. 2, pp. 306–311, 2018, doi: 10.1108/WJE-08-2017-0260.

[41]. J. Ramkumar, S. S. Dinakaran, M. Lingaraj, S. Boopalan, and B. Narasimhan, "IoT-Based Kalman Filtering and Particle Swarm Optimization for Detecting Skin Lesion," in Lecture Notes in Electrical Engineering, K. Murari, S. Kamalasadan, and N. P. Padhy, Eds., Springer Science and Business Media Deutschland GmbH, 2023, pp. 17–27. doi: 10.1007/978-981-19-8353-5_2.

[42]. R. Jaganathan, S. Mehta, and R. Krishan, Intelligent Decision Making Through Bio-Inspired Optimization. IGI Global, 2024. doi: 10.4018/979-8-3693-2073-0.

[43]. R. Jaganathan and V. Ramasamy, "Performance modeling of bio-inspired routing protocols in Cognitive Radio Ad Hoc Network to reduce end-to-end delay," International Journal of Intelligent Engineering and Systems, vol. 12, no. 1, pp. 221–231, 2019, doi: 10.22266/IJIES2019.0228.22.

[44]. P. Menakadevi and J. Ramkumar, "Robust Optimization Based Extreme Learning Machine for Sentiment Analysis in Big Data," in 2022 International Conference on Advanced Computing Technologies and Applications, ICACTA 2022, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICACTA54488.2022.9753203.

[45]. [17] J. Ramkumar, R. Vadivel, and B. Narasimhan, "Constrained Cuckoo Search Optimization Based Protocol for Routing in Cloud Network," International Journal of Computer Networks and Applications, vol. 8, no. 6, pp. 795–803, 2021, doi: 10.22247/ijcna/2021/210727.

[46]. [18] D. Jayaraj, J. Ramkumar, M. Lingaraj, and B. Sureshkumar, "AFSORP: Adaptive Fish Swarm Optimization-Based Routing Protocol for Mobility Enabled Wireless Sensor Network," International Journal of Computer Networks and Applications, vol. 10, no. 1, pp. 119–129, 2023, doi: 10.22247/ijcna/2023/218516.

[47]. J. Ramkumar and R. Vadivel, "Improved Wolf prey inspired protocol for routing in cognitive radio Ad Hoc networks," International Journal of Computer Networks and Applications, vol. 7, no. 5, pp. 126–136, 2020, doi: 10.22247/ijcna/2020/202977.

[48]. R. Vadivel and J. Ramkumar, "QoS-enabled improved cuckoo search-inspired protocol (ICSIP) for IoT-based healthcare applications," in Incorporating the Internet of Things in Healthcare Applications and Wearable Devices, IGI Global, 2019, pp. 109–121. doi: 10.4018/978-1-7998-1090-2.ch006.

[49]. S. P. Geetha, N. M. S. Sundari, J. Ramkumar, and R. Karthikeyan, "Energy Efficient Routing In Quantum Flying Ad Hoc Network (Q-FANET) Using Mamdani Fuzzy Inference Enhanced Dijkstra's Algorithm (MFI-EDA)," J Theor Appl Inf Technol, vol. 102, no. 9, pp. 3708–3724, 2024, [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85197297302&partnerID=40&md5=72d51668bee6239f09a59d2694df67d6

[50]. J. Ramkumar, R. Karthikeyan, and V. Valarmathi, "Alpine Swift Routing Protocol (ASRP) for Strategic Adaptive Connectivity

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Enhancement and Boosted Quality of Service in Drone Ad Hoc Network (DANET),” International Journal of Computer Networks and Applications, vol. 11, no. 5, pp. 726–748, 2024, doi: 10.22247/ijcna/2024/45.

[51]. J. Ramkumar, R. Karthikeyan, and M. Lingaraj, “Optimizing IoT-Based Quantum Wireless Sensor Networks Using NM-TEEN Fusion of Energy Efficiency and Systematic Governance,” in Lecture Notes in Electrical Engineering, V. Shrivastava, J. C. Bansal, and B. K. Panigrahi, Eds., Springer Science and Business Media Deutschland GmbH, 2025, pp. 141–153. doi: 10.1007/978-981-97-6710-6_12.

[52]. J. Ramkumar and R. Vadivel, “Multi-Adaptive Routing Protocol for Internet of Things based Ad-hoc Networks,” Wirel Pers Commun, vol. 120, no. 2, pp. 887–909, Apr. 2021, doi: 10.1007/s11277-021-08495-z.

[53]. R. Jaganathan, S. Mehta, and R. Krishan, Bio-Inspired intelligence for smart decision-making, vol. i. 2024. doi: 10.4018/9798369352762.

[54]. J. Ramkumar, K. S. Jeen Marseline, and D. R. Medhunhashini, “Relentless Firefly Optimization-Based Routing Protocol (RFORP) for Securing Fintech Data in IoT-Based Ad-Hoc Networks,” International Journal of Computer Networks and Applications, vol. 10, no. 4, pp. 668–687, 2023, doi: 10.22247/ijcna/2023/223319.

[55]. M. Lingaraj, T. N. Sugumar, C. S. Felix, and J. Ramkumar, “Query aware routing protocol for mobility enabled wireless sensor network,” International Journal of Computer Networks and Applications, vol. 8, no. 3, pp. 258–267, 2021, doi: 10.22247/ijcna/2021/209192.

[56]. A. Senthilkumar, J. Ramkumar, M. Lingaraj, D. Jayaraj, and B. Sureshkumar, “Minimizing Energy Consumption in Vehicular Sensor Networks Using Relentless Particle Swarm Optimization Routing,” International Journal of Computer Networks and Applications, vol. 10, no. 2, pp. 217–230, 2023, doi: 10.22247/ijcna/2023/220737.

[57]. L. Mani, S. Arumugam, and R. Jaganathan, “Performance Enhancement of Wireless Sensor Network Using Feisty Particle Swarm Optimization Protocol,” ACM International Conference Proceeding Series, pp. 1–5, Dec. 2022, doi: 10.1145/3590837.3590907.

[58]. J. Ramkumar and R. Vadivel, “Whale optimization routing protocol for minimizing energy consumption in cognitive radio wireless sensor network,” International Journal of Computer Networks and Applications, vol. 8, no. 4, pp. 455–464, 2021, doi: 10.22247/ijcna/2021/209711.