International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I6P123

Volume 74 Issue 6, 328-342, June 2026

Original Article

Emotion Journey-Aware Vader (EJ-VADER) Approach for Optimized Sentiment Interpretation in Consumer Behaviour Studies

C. Balakrishnan[1], J. Ramkumar[2], R. Karthikeyan[3]

[1]*Department of Commerce with Banking and Insurance, Dr. N. G. P. Arts and Science College, Coimbatore, Tamil Nadu, India.*
[2]*Department of Computer Science, School of Quantum Science and Computing, Rathinam Global (Deemed to be University), Coimbatore, Tamil Nadu, India.*
[3]*Department of Computer Technology, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India.*

[2]*Corresponding Author : jramkumar1986@gmail.com*

Received: 19 January 2026 Revised: 06 May 2026 Accepted: 21 May 2026 Published: 27 June 2026

Abstract *- Emotion-journey consistent expansion of lexicon-based sentiment analysis that could be used to interpret consumer behaviour in the business setting. The consumer-generated text embodies signals of emotion, and their meaning in behavioural terms also changes at different levels of decision making, and interpretation of sentiment at rest cannot be adequate. The proposed approach combines customer journey awareness and sentiment intensity modelling in order to raise the degree of behavioural relevance without compromise in interpretability. EJ-Vader is an extension of Valence Aware Dictionary and Sentiment Reasoner (VADER), which presents journey-aware intensity scaling, transaction-stage negative prioritization, retention-stage positive reinforcement, dynamic threshold calibration, and attribute-level sentiment alignment. These elements allow sentiment indicators to indicate the decision sensitivity, which is associated with evaluation, transaction, and after-sales interaction. Behavioural signal mapping also links optimized sentiment intensity to business analysis, such as risk inclination, formation of loyalty, and advocacy preparation. Validation is designed based on congruence with observable commerce results, as practical. The framework presents a behaviour-focused, scalable, and transparent sentiment analysis framework applicable to consumer analytics, decision support, and commerce-based research applications.*

Keywords *- Sentiment Analysis, Consumer Behaviour, Emotion-Journey, Lexicon-Based Intensity, VADER Optimization, Commerce Analytics.*

1. Introduction

The Sentiment analysis is a computational procedure of identifying emotional significance encoded in written manifestations created on digital platforms. Opinions, appraisals, and responses are always willing to be picked up by the market communication channels as they come in the form of reviews, feedback messages, service interactions, and social conversations [1]. These words contain emotional signs which warn of satisfaction, dissatisfaction, trust, or reluctance. It is out of such raw emotional clues that the sentiment analysis should be an evolution in order to analyze structured analytical evidence in the context of research related to commerce [2].

The study of consumer behaviour is dedicated to the analysis of how people make their choices, choose to purchase, experience, and form prolonged relationships with brands. The emotional perception is incredibly decisive throughout all levels of the consumer decision lifecycle [3]. Emotional tone used in consumer communication has a strong impact on price acceptance, quality perception, service evaluation, and formation of loyalty. The strength of sentiment analysis as a consumer behaviour study lies in the ability to understand the emotional reaction on a large scale that cannot be directly determined by transactional data. Emotional insight helps to supplement behavioural feedback to demonstrate the intensity of motivation, unmet expectations, and after-sales responses [4]. One of the ongoing issues of sentiment analysis has to do with the difficulty of measuring the intensity. Conventional classification of polarity sees positive, negative, and neutral expressions or utterances, but this is a very superficial classification, as it does not provide much interpretation of behaviour. A positive expression displaying some mild positive expression and an expression with a great deal of enthusiasm often have the same label, whereas the difference between them in the implication of behaviour is very high. Repeat purchase, recommendation, escalation of complaints, or even churn do not rely on emotional direction, but rather on the emotional strength of the consumer action. Misinterpretation of intensity results in low correlation between sentiment indicators and actual consumer behaviors, which limits the predictability of commerce environments [5]. Existing sentiment analysis approaches are based on polarity or static estimation of the intensity without variation in the behavioural impact during the customer journey stages. Lexicon-based models guarantee interpretability, but they are not adapted enough to the context, whereas deep learning methods grasp the context with a limited level of transparency. Absence of a framework, combining the interpretable intensity and journey-aware behavioural alignment, creates a great research gap in commerce-oriented sentiment analysis.

The expression of emotions does not occur in the same way throughout the consumer journey, and emotion-journey alignment must be considered. Emotional meaning develops with the consumer between the stages of awareness, evaluation, transaction, and post-purchase experience [6]. Such an expression of concern, when the exploration is in its initial stages, indicates the presence of curiosity or a demand for information, whereas the expression, when the exploration is in its final stages, indicates the perception of risk. Emotion journey alignment is aware that emotional intensity has stage-specific behavioural implications. Sentiment processing prevents behavioural realism and enhances the relevance of a decision by integrating journey context [7]. The term emotion journey is the progression of emotional responses displayed by consumers at different stages of decision making, such as awareness, evaluation, transactions, and post-purchase engagement. The intensity of emotional experience is not constant among these stages, and the behavioural impact varies depending on the context of the decision to be made.

Existing approaches to sentiment analysis consider emotional expressions as context-independent expression signals, relying on polarity or static intensity without taking the stage sensitivity into consideration. Such approaches do not capture variation in behavioural influence throughout the consumer decision process. The proposed EJ-VADER framework proposes emotion-journey awareness by both matching the intensity of sentiments and the relevance of the decision stage. This provides the way to stage-conditioned interpretation through journey-aware scalability, negative prioritization towards transactions, positive reinforcement towards retention, and dynamic calibration. This formulation provides an extension over conventional sentiment models that is behaviour-centric.

Lexicon-based intensity models are a clear-cut way to estimate the strength of the sentiment. The existence of such models is based on pre-established dictionaries in which words and symbols have attached emotional valence values. Due to their reliance on a lexicon, lexicon-based techniques are interpretable, economical to compute, and appropriate to short informal text (when used in consumer communication). Capitalization, punctuation, degree modifiers, and symbols of emotion are intensity cues that are still explicitly modeled. The reason behind these approaches being the best fit in the pipelines of commerce analytics is that lexicon-driven scoring requires little training data, and it can operate with such data, which is not resource-intensive [8].

Valence Aware Dictionary and Sentiment Reasoner (VADER) also stand out as one of the leading sentiment intensity models based on lexicon, which is applicable to social and consumer-oriented text. VADER also uses linguistic heuristics and assigns weighted valence scores on sentiment-bearing words, including emphasis, negation, and amplification of intensity [9] in VADER. The generated compound score of VADER is not a categorical output; it is emotional strength. This feature makes VADER an appropriate reference point for the intensity-sensitive consumer behaviour analysis. The fact that its lexical transparency creates explainable sentiment understanding, which can be linked to the requirements of the business decision, is made possible through rule-based scoring [10]. The proposed EJ-VADER framework adds emotion-journey to lexicon-based sentiment analysis in stage awareness intensity scaling, transaction negative prioritization, retention positive, and dynamic threshold calibration. The interpretability is preserved with the integration of behaviour sensitivity within, and hence the distinct contribution in analysing the consumer behaviour.

In spite of these, the limitations in predicting behaviour by using VADER are based on the fact that there is a limit to the static intensity interpretation. The same sentiment score applied to all the stages of interaction treats variation in sensitivity of decision as it exists across the consumer journey. Emotion journey alignment provides a perfectly straightforward way of maximizing VADER without uprooting lexical principles. Journey-conscious scaling, conditioned prioritization (with stages), and calibration by dynamic threshold bring all the emphasis of baseline intensity interpretation to the demands of behavioural realism.

Loyalty reinforced through emotional expressions is further reinforced at moments that are close to transactions, whereas post-use satisfaction is reinforced with risk fear. VADER is optimized by emotion journey alignment, which is used between sentiment analysis and consumer behaviour modeling. Intensity estimation that occurs via lexicon is preserved, whereas contextual modulation maximizes behavioural relevance. The obtained guide contributes to business-focused reading of emotional cues, making it possible to map the expressed emotion, stage of decision-making, and the willingness to act with the purchases more refined. This correspondence provides a platform for behaviour-sensitive sentiment analysis that should be used in contemporary commerce settings.

1.1. Problem Statement

Sentiment analysis is significant in the interpretation of consumer behaviour in terms of emotional reading of textual responses. The current methods of sentiment analysis mostly utilize polarity-based classification, which restricts the capacity of establishing the deployment of emotional strength. Sentiment models can offer an intensity score of the sentiment, but these scores do not change and cannot be context-dependent. Emotions that buyers exhibit at various steps in the buying process have dissimilar behavioural values. Losing the context of journeys creates poor correlation between sentiment signals and reality consumer behaviours, including churns, repurchase, or advocacy behaviours. There is an obvious issue in how to translate the intensity of sentiment into behaviours indicative of meaning through commerce touchpoints. This is a limitation to the serious application of sentiment analytics to decision support in customer-focused settings.

1.2. Motivation

Contemporary business ecosystems produce significant amounts of consumer-created, text-based expressions of emotions, expectations, and satisfaction. The intensity of the emotion and not the direction of sentiment alone controls behavioural decisions like purchase confirmation, complaint upgrading, or formation of loyalty. The models of sentiment do not present emotional sensitivity at a stage of the consumer journey. This shortcoming decreases forecasting robustness and business feasibility. Lexicon-based methods are transparent and interpretable, but do not contain adaptive mechanisms that are consistent with consumer decision processes. The reasons why it should be motivated are based on the necessity to maintain lexical clarity and make behaviour more realistic. The probability of integrating emotion–journey consciousness into sentiment processing is that it allows enhancing the sentiment-guided consumer behaviour analysis with no complexity of the model or loss of interpretability.

1.3. Objectives

The major aim is to create an emotion–journey resolved optimization framework in the lexicon-based sentiment intensity analysis of consumer behaviour research. The article seeks to improve baseline VADER scoring inheritance with the aid of journey efficient intensity scaling, stage-sensitive prioritization, and scaling of threshold dynamism. The other goal is to be able to match sentiment intensity and commerce attributes that directly affect consumer choices. The framework aims to project optimized sentiment signals to understandable behaviour indicators, which are related to risk, loyalty, and advocacy. Comparison with actual results of commerce is still a vital goal in order to guarantee uniformity of behaviour. In general, the goal focuses on transforming the intensity of sentiment into signals that beget behaviour and which are sustainable in terms of transparency, scalability, and commerce-related analytics.

2. Literature Review

"Peer-Driven EV Uptake" [11] integrates consumer sentiment extraction with peer influence modeling to study new energy vehicle adoption. Large-scale textual data are collected from online platforms and processed through transformer-based language models to derive sentiment polarity and intensity. Network-based clustering captures peer proximity and diffusion strength. "Social AI Perception Mapping" [12] adopts a comparative social media analytics design using Twitter (X) and Reddit datasets. Text preprocessing includes noise removal, lemmatization, and contextual embedding. Sentiment classification employs supervised deep learning aligned with platform-specific discourse traits. "Attention-Driven Ensemble Sentiment" [13] suggests an ensemble deep learning pipeline based on bidirectional long short-term encoders with attention enhancement. There are a number of base classifiers that share the embeddings, as well as attention layers that give precedence to sentiment-bearing tokens. Ensemble Fusion Squeezes the trigger. "Multimodal Review Mining" [14] is a model that combines text, image, and audio features to analyse consumer review sentiment. Image convolutional networks, text recurrent encoders, and audio spectral descriptors are all used in feature extraction. The fusion layer brings features together into a single representation despite the heterogeneity of the features.

"Macroeconomic Sentiment Dynamics" [15] uses a time series econometrics approach to connect consumer sentiment, monetary policy, and inflation dynamics. Long-memory properties are measured using fractional integration methods, and equilibrium relations are measured using cointegration models. "Unsupervised Aspect Attention" [16] builds an unsupervised neural model to extract aspect-based sentiment analysis of consumer reviews. Latent aspect discovery is based on attention mechanisms, which are encoded in neural encoders that remove the labeled dependency. Contextual embeddings predict aspect separation and sentiment alignment. "Graph-Centric Product Insight" [17] executes an enhanced graph convolutional network to program the connections among customers, products, and opinion characteristics. Test texts are converted to a semantic representation and represented as nodes with weighted edges as the degree of interaction. Graph propagation represents contextual effects in terms of related opinions. "Disentangled Multimodal Fusion" [18] presents a hybrid architecture that integrates disentangled representation learning and cross-modal collaborative interaction. All modalities are independently coded in order to isolate common and unique factors of sentiment. The coordination layer is a layer that matches complementary cues in different modalities based on interaction matrices.

"Few-Shot Hierarchical Reasoning" [19] solves few-shot multimodal sentiment tasks using hierarchical reasoning modules. The individual modalities are processed by low-level encoders, and a higher reasoning layer combines semantic relations between modalities and samples. Sentiment generalization can be trained with little labelled data using prototype-based learning. Sentiment inference follows the reasoning paths based on the well-organized levels of abstraction. "Syntax–Semantic Graph Fusion" [20] fuses TextGINConv and KolmogorovArnold networks to boost aspect-based sentiment analysis because of Syntax- Semantic Graph Fusion. Graph convolution is used to learn the semantics of relations by syntactic dependency graphs that organize textual inputs. Semantic feature transformation is enhanced by nonlinear Kolmogorov-Arnold mappings. The aspect-level sentiment learning is a result of joint graph and functional representations. "Dual-Stream Modal Learning" [21], a dual-stream-based collaborative network is used to learn the modality-specific and modality-invariant sentiment representations. Shared affective and distinct modal encoders are obtained in parallel. There is a coordination mechanism that implements the balanced learning between streams. Incorporating joint constraints with modal variation, fusion incorporates invariant stability. "Self-Enhancing Attention Fusion" [22] brings on board attention-guided attributes that are optimized by contrastive goals. The strengthening of salient sentiment cues in fused representations in an iterative process is called self-enhancement—contrastive alignment in the same way that semantic consistency is maintained across modalities. "Robust Multimodal Optimization" [23] involves the use of convolution-enhanced transformers and a robust optimization network to address the variability of multimodal sentiment. The local feature extraction is enhanced by the use of convolution layers, and transformers extract the long-range dependencies. Strong optimization eliminates noise and modality imbalance adaptively with a loss shaping.

"SMSA" [24] is a mass-scale acquisition of social media text on websites where organic food is discussed by consumers. Data processing is performed in the form of orderly cleaning, normalization, and refining tokens to keep expressions of opinion. Word embeddings are built to produce semantic representations with respect to the contextual use of language in food-related language. The polarity and magnitude of sentiments are obtained with the help of supervised and lexicon-based neural classifiers that have been trained on informal language on social media. Topic-sensitive sentiment segmentation associates the tone of emotion with organic, trust, and lifestyle references. Temporal slicing aids in tracking perceptions over the market stages. "OSP-DM" [25] embraces a combined pipeline of deep learning that is customized to work with online shopping platform consumer reviews. On textual inputs, linguistic preprocessing, syntactic structuring, and embedding generation are performed using hybrid convolutional and recurrent encoders. Hierarchies of features obtain not only local opinion cues, but also long-range contextual dependencies. Model integration Model model integration is a layer that uses fusion to stabilize the sentiment representation across product categories between parallel deep learners. The training protocols attribute importance to the class balance, adaptive learning rate, and regularization control to regulate heterogeneity of the review. Aspect-sensitive sentiment alignment relates emotional indicators to features of products.

Emotion-Journey aligned Valence Aware Dictionary and sEntiment Reasoner (EJ-VADER) is a journey-conscious expansion of VADER that optimizes emotion intensity through matching emotional strength and customer journey phases to behaviorally significant consumer choices in a commerce-driven context and analytics.

2.1. Consumer Data Acquisition

Consumer Data Acquisition has been defined as the methodical gathering of the consumer-expressed textual signals produced in commerce touchpoints. This is one of the most crucial steps of EJ-VADER as the sentiment intensity optimisation is fully based on the real consumer voice recorded at the areas of interaction associated with awareness, evaluation, purchase, and post-purchase interaction. Proper acquisition means no modification of emotional signals is done, and that they are context-rich and behaviourally relevant, which supports robust emotion-journey correspondence on subsequent steps.

$$\mathcal{D} = \bigcup_{i=1}^{N} \{ T_i, \tau_i, \kappa_i \} \tag{1}$$

In Equation (1), $\mathcal{D}$ is used to denote the aggregation of consumer information in the form of a dataset. T_i represents the unstructured textual articulation of a consumer. τ_i represents temporal closeness to business communication, like browsing, purchasing, or utilization. κ_i represents contextual commerce data like channel type, product category, and source of interaction. Such formulation makes the sentiment signals contextually dependent on the behavioural context as opposed to independent text pieces.

Structures of commerce touchpoints, arrange incoming data of consumers according to the channel and decision relevance. The text produced during reviews, during service chat, during feedback portals, or during transactional communication varies in terms of intent to emotion, and urgency of behaviour. This coherence of emotion and commerce touchpoints by balancing acquisition and commerce optimises EJ-VADER sensitivity during the decision stage.

$$\omega_i = f(T_i, C_i) \times \lambda_c \tag{2}$$

The score described in Equation (2), which is weighted emotional signal readiness, is denoted by ω_i. C_i refers to the point of commerce that corresponds to the text instance. The semantic relevance between consumer expression and commerce context is represented by the $f(\cdot)$. λ_c It is a commerce alignment coefficient that alters importance with touchpoint criticality, including purchase confirmation or post-use feedback.

The integrity of behavioural signals guarantees that emotional signals used to represent consumer motivation, level of satisfaction, or the degree of dissatisfaction are not distorted. The EJ-VADER is based on the lexical amplification features, so the punctuations, capitalizations, and signs of emotions should be preserved without distortion. This is to protect optimal extraction of sentiments without interfering with the authenticity of raw behavioural data.

$$\Phi_i = \psi(T_i) + \delta_e \tag{3}$$

In equation (3), in which the symbol Φ_i represents the maintained signal of behavioural strength. The transformation function, which is denoted as $\psi(\cdot)$, does not change the lexical and emotive structure but makes it a useful element in calculating intensity by VADER. δ_e denotes capturing residual emphasis on emotion that is added by means of expressive occasions found in consumer text.

2.2. Customer Journey Stage Identification

Customer Journey Stage Identification establishes the procedure of categorization of consumer exchanges into delicate phases of decision making, which denote an evolving sentiment, anticipation, and inclination to act. This is necessary in EJ-VADER as the sentiment intensity only takes behavioural meaning when aligned with the stage of interaction. The emotional expressions obtained in the step above have to be systematically placed in the commerce journey to facilitate the optimization of the intensity modulation and proper interpretation of behaviour.

$$\mathcal{J}_i = \arg\max_{j \in \mathbb{S}} P(j \mid \Theta_i) \tag{4}$$

In Equation (4), the assigning journey stage of the i^{th} Consumer interaction is represented as the value of the character derived from the assigning journey stage. $\mathcal{J}_i$. The set $\mathbb{S}$ denotes predetermined steps of commerce such as awareness, consideration, purchase, post-purchase, retention, and exit. Θ_i It is a representation of behavioural indicators based on metadata acquisition values like time of interaction, depth of engagement, and linkage of transactions. The stage allocation is such that the probability formulation makes sure that the behavioural likelihood is considered and not arbitrary labeling. Behavioural Context Mapping links every expression made by consumers to observable signs of their willingness to make a decision. Consumer behaviour is developed through stages, and emotional interpretation is influenced by the level of motivation and risk sensitivity. This mapping allows EJ-VADER to maintain continuity in behaviour between the raw sentiment signal and optimal journey-related scoring.

$$\mathcal{B}_i = g(\tau_i, \rho_i, \sigma_i) \tag{5}$$

In Equation (5), out of the behavioural context vectors, the behavioural context of the i^{th} An element is denoted by the $\mathcal{B}_i$. Variable i is the time closeness to commerce activity, which is denoted by the variable describing time as being τ_i. Interaction repetition or frequency of engagement is represented by the variable of ρ_i. The association strength of a transaction is denoted by the variable of σ_i. The mapping process $g(\cdot)$ brings the behavioural signals together in a single expression that aids in stage discrimination.

Optimization of the journey boundary calibration optimizes the stage separation by minimizing overlaps between consumer decision phases. Different stages have varying levels of influence through emotion, which needs to be defined in behavioural terms. EJ-VADER also incorporates decision threshold calibration to improve congruence between the stage of sentiment without interfering with lexicon-based sentiment calculation.

$$\Gamma_j = \mu_j + \alpha_j \cdot \nu_i \tag{6}$$

In Equation (6), Γ_j represents the boundary threshold of journey stage j, which has been calibrated. μ_j This is the parameter that is used to indicate the baseline behavioural expectation of the stage. Sensitivity is modified in terms of the priority of commerce by the coefficient. α_j. The variable ν_i, encodes normalized engagement strength through the patterns of consumer interaction.

Stage-Linked Sentiment Readiness Index is a measure of how much consumer emotion is ready to be optimised in terms of intensity. This index allows relating journey positioning to potential sentiment impact, providing EJ-VADER with the ability to distinguish between exploratory emotion and decision-driving emotion.

$$\Omega_i = \mathcal{J}_i \times \mathcal{B}_i \times \kappa_j \tag{7}$$

In Equation (7), Ω_i is the index of sentiment preparedness. The journey stage is allocated, $\mathcal{J}_i$, is in contact with the behavioural context, $\mathcal{B}_i$. The parameter κ_j refers to the stage-specific weight of emotional influence optimization.

2.3. Normalization Target Definition

Text Normalization and Preparation is the transformation of raw consumer text into a uniform and analysis-ready format without destroying emotion-carrying indicators that exist in the consumer behaviour signal. This is an important step following Customer Journey Stage Identification because stage labels need to contain clean text, yet purchase intent, level of satisfaction, urgency of complaint, and trust signals that are necessary to enhance the EJ-VADER intensity processing.

$$\tilde{T}_i = \mathcal{N}(T_i;\, \pi) \tag{8}$$

In Equation (8), $\tilde{T}_i$ has been replaced by the normalized consumer text of interaction i. T_i is the stage-tagged raw text after acquisition, and $\mathcal{N}(\cdot)$ is a normalization operation with configuration π, which influenced metamorphoses like whitespace normalization, Unicode cleaning, and controlled case, without distorting commerce-relevant intent signals.

Removal of disruptive noise and preservation of expressive markers in consumer language are defined by emotion-preserving noise control. This is significant in commerce streams in which brief messages, typing errors, stretches, emojis, and emphasis marks have behavioural significance, which subsequently regulates EJ-VADER optimized scaling amid journey phases.

$$T_i^{\star} = \tilde{T}_i - \mathcal{R}(\tilde{T}_i) + \mathcal{E}(\tilde{T}_i) \tag{9}$$

In Equation (9), $T_i^{\star}$ as the cleaned-yet-expressive text $\mathcal{R}(\cdot)$ eliminates non-informative expressive elements, including broken encodings, redundant separators, and unsafe markup. $\mathcal{E}(\cdot)$ retains emotion carriers, including patterns of emphasis, long-run tokens, emoji indicators, and sentiment modifiers, which are crucial to intensity measurement.

Commerce cue tokenization is a format of segmentation of prepared text into units that retain behavioural cues related to product evaluation, price sensitivity, delivery experience, service quality, and perceived value. This is important to consumer behaviour modelling because such intensity changes can be manifested in modifiers and emphasis tokens and not in nouns.

$$\mathrm{x}_i = \mathcal{T}(T_i^{\star}) \oplus \mathcal{Q}(T_i^{\star}) \tag{10}$$

In Equation (10), x_i represents the token-cue representation. $\mathcal{T}(\cdot)$ yields ordered tokens, which are useful in lexicon lookup. $\mathcal{Q}(\cdot)$ yields quantified indicators of cues like exclamation density, capitalization ratio, repeated-character strength, and negation scope markers. The operator of this kind is denoted by the symbol of addition, which is called the fusion of features.

According to stage-consistent preparation, alignment of features of cleaned text with the previously determined stage of the journey is determined, making it possible to transition to baseline VADER scoring and subsequent journey-aware intensity optimisation. This significance comes due to differences in the stage of sensitivity in behavioural evaluation, commitment to purchase, and post-purchase responses.

$$\mathrm{z}_i = \mathrm{x}_i \odot \mathrm{m}_{\mathcal{J}_i} \tag{11}$$

In Equation (11), the stage-conditioned input to the scoring stage is denoted by the variable. z_i and the stage mask by the variable $\mathrm{m}_{\mathcal{J}_i}$. The operator $\odot$ means element-wise conditioning, which enjoys the optimization of intensity readiness, without change of consumer meaning.

Text normalization consists of several steps: token normalization, noise removal, and intensity-maintaining transformations such as text capitalization, punctuation emphasis, and negations. Feature construction is obtained from lexicon-based sentiment cues wherein the individual tokens have weighted intensity depending on predefined valence scores. Additional features include emphasis, contextual negation impact, and attribute-specific term identification. These features are aggregated to form structured sentiment representations that are aligned to signals for consumer behaviour so that these can be consistently input for further journey-aware scaling and behavioural mapping.

2.4. Baseline VADER Sentiment Scoring

Baseline VADER Sentiment Scoring is the term used for the systematic computation of raw emotional polarity and strength calculations from stage-conditioned consumer text. This step is at the core of EJ-VADER since all optimised and journey-aware sentiment modulation/ is based on the trustworthiness and transparency of the original sentiment signal. Within consumer behaviour analysis, baseline scoring creates an unbiased emotional reference that relates expressed satisfaction, dissatisfaction, trust, or frustration that is embedded in commerce interactions.

$$S_i = \sum_{k=1}^{|\mathbf{z}_i|} v_k \cdot \eta_k \tag{12}$$

In Equation (12), S_i is the aggregate sentiment score for i, i.e., the i^{th} consumer interaction. The term v_k gives the lexicon-assigned valence score of k^{th} token obtained from prepared input $\mathbf{z}_i$. The factor η_k captures local contextual influence, e.g., negation scope, emphasis markers, and modifier strength. Such a formulation maximizes the lexical transparency while it reflects the emotional tone that is present in a consumer's language.

Polarity decomposition is a process in which the sentiment is broken down into positive, negative, and neutral parts to aid in commerce-relevant behavioural interpretation. Consumer behaviour is rarely dependent upon overall emotion without relative dominance of approval over dissatisfaction in continuation of purchase, hesitation, or exit decisions.

$$\{P_i, N_i, U_i\} = \mathcal{D}(S_i) \tag{13}$$

In Equation (13), P_i represents cumulative positive sentiment strength, N_i represents cumulative negative sentiment strength and U_i represents the neutral contribution. The aggregated sentiment is decomposed by the operator $\mathcal{D}(\cdot)$into hypothetically interpretable realimentary components by lexicon polarity classification. This separation aids in later optimised weighing up of stages in the journey.

Compound sentiment normalization consists of compressing the polarity information into a bounded continuous scale suitable for intensity-based optimisation. Consumer behaviour modelling benefits from normalised sentiment values that can be used to ensure comparison can be done consistently between products, services, and stages of the journey.

$$C_i = \frac{s_i}{\sqrt{s_i^2+\beta}} \tag{14}$$

In Equation (14), C_i will be the normalized compound sentiment score. A parameter called β stabilises the normalisation when establishing extreme accumulation of sentiments to ensure a smooth scaling that does not saturate. This score is based on raw emotional intensity relative to the absence of journey bias.

Optimized sentiment readiness encoding of baseline scores for emotion-journey alignment for subsequent emotion journey processing EJ-VADER steps. The baseline score is unalteringly kept at this point to maintain the interpretability and methodological integrity of commerce analytics.

$$\Psi_i = \langle P_i, N_i, C_i \rangle \tag{15}$$

In Equation (15), Ψ_i is an input of the baseline sentiment vector passed to the optimization layer. The vector maintains the separation of polarity and compound intensity, which allows stage-aware scaling, threshold calibration, and behaviour linking prediction in line with the dynamics of consumer decision.

2.5. Journey-Aware Intensity Scaling

Journey-Aware Intensity Scaling refers to the act of controlling the adjustment of baseline sentiment intensity based on the stage of the customer journey identified earlier. This step has high importance to EJ-VADER as the same emotional strength differently affects the consumer behaviour in the different phases of awareness, consideration, purchase, and post-purchase. Baseline VADER outputs give a neutral emotional context against which the emotional responses are being likened (modified) into stage-sensitive intensity, based on commerce-aligned prediction.

$$\hat{C}_i = C_i \cdot s(\mathcal{J}_i) \tag{16}$$

In Equation (16), the symbol symmetrically indicates the journey-scaled compound intensity of interaction. $\hat{C}_i$. C_i is the baseline compound score produced. Function $s(\mathcal{J}_i)$ returns a scale-up coefficient associated with the particular stage $\mathcal{J}_i$ is a label of the label stages. This mapping covers an intensity rise toward commitment-sensitive stages and an intensity ease within exploratory stages, for maintaining the sense of realism in commerce.

Stage-sensitive polarity shaping defines separate scaling for positive and negative components to reflect asymmetric behavioural impact in commerce settings. Consumer dissatisfaction close to purchase or post-purchase typically triggers stronger actions than mild approval, making polarity-aware scaling essential for optimized behavioural fidelity.

$$\hat{P}_i = P_i \cdot p(\mathcal{J}_i), \hat{N}_i = N_i \cdot n(\mathcal{J}_i) \tag{17}$$

In Equation (17), $\hat{P}_i$ and $\hat{N}_i$ denote stage-shaped positive and negative strengths. P_i and N_i originate from baseline polarity decomposition. The functions $p(\mathcal{J}_i)$ and $n(\mathcal{J}_i)$ represent optimized stage coefficients that tune approval and dissatisfaction impact in line with consumer decision sensitivity at the identified journey phase.

Optimized nonlinear intensity stabilization defines a bounded transformation that prevents extreme scaling from overwhelming behavioural interpretation. Commerce conversations contain bursts of emotion, repeated emphasis, and sharp complaint language; stabilization keeps scaled intensity comparable across products, channels, and customer segments.

$$\bar{C}_i = \tanh(\hat{C}_i \cdot \gamma_{\mathcal{J}_i}) \tag{18}$$

In Equation (18), $\bar{C}_i$ denotes stabilized journey-aware intensity. $\hat{C}_i$ is the scaled compound intensity. The parameter $\gamma_{\mathcal{J}_i}$ controls stage-wise sensitivity under stabilization, enabling stronger response near transaction-critical stages without uncontrolled growth.

Commerce-ready intensity packet defines the final scaled sentiment representation forwarded to subsequent EJ-VADER optimisation components. This packaging is important for maintaining clear connectivity between baseline VADER scoring and later stage-driven penalties, reinforcements, and threshold calibration.

$$\mathrm{I}_i = \langle \bar{C}_i, \hat{P}_i, \hat{N}_i, \mathcal{J}_i \rangle \tag{19}$$

In Equation (19), I_i represents the optimized intensity packet. The elements combine to stabilize compound intensity. $\bar{C}_i$, stage-shaped polarity $\hat{P}_i$ and $\hat{N}_i$, and the journey stage $\mathcal{J}_i$. This structure supports direct linkage between emotional intensity and commerce journey position for behaviour-focused prediction.

2.6. Transaction-Stage Negative Intensity Prioritization

Transaction-Stage Negative Intensity Prioritization defines the focused amplification of negative sentiment expressed near purchase, payment, delivery, or immediate usage stages. This step holds strong importance within EJ-VADER since consumer behaviour becomes highly action-oriented at transaction-linked moments. Emotional dissatisfaction expressed at this stage signals cancellation risk, return intention, complaint escalation, or trust erosion. Building upon journey-aware intensity scaling, this step sharpens behavioural sensitivity by prioritizing negative emotion where financial commitment and expectation peaks.

$$\widetilde{N}_i = \widehat{N}_i \cdot \chi(\mathcal{J}_i) \tag{20}$$

In Equation (20), the negative of the prior intensity is denoted by the term. $\widetilde{N}_i$, which is a transaction-prioritized negative intensity. The term, $\widehat{N}_i$, stage-form negative sentiment weighted obtained from journey aware scaling. The functionality $\chi(\mathcal{J}_i)$ Provides transaction-specific amplification coefficient. Triggered strongly from purchase to immediate post-purchase. This mechanism ensures that dissatisfaction approaching commitment points has proportionately more behavioural weight.

Loss-dominant behavioural emphasis brings home the principle of commerce that dissatisfaction has more powerful control over decision reversal than satisfaction has over reinforcement. This principle is in line with observed consumer behaviour, whereby negative emotion motivates faster and stronger action. This dominance is coming to EJ-VADER's intensity prioritization.

$$\Lambda_i = \widetilde{N}_i - \delta \cdot \hat{P}_i \tag{21}$$

In Equation (21), Λ_i denotes the net loss-dominant sentiment force. The variable $\widetilde{N}_i$ captures amplified negative intensity, while $\hat{P}_i$ represents a stage-shaped positive intensity. The coefficient δ regulates positive sentiment offset, maintaining dominance of dissatisfaction without nullifying approval signals. This formulation emphasizes risk over reassurance during transaction-critical interactions.

Transaction sensitivity reinforcement is responsible for emotional volatility found near payment confirming, delivery acknowledging, and service resolving points. Emotional expressions in these moments have a high behavioural immediacy, which has to be optimized as well for the sake of not underestimating.

$$\Theta_i = \Lambda_i \cdot \xi_i \tag{22}$$

In Equation (22), Θ_i is the reinforcement transaction-stage emotional pressure. The factor ξ_i represents transaction proximity intensity obtained by interaction time, relates order status linking their relationship, or resolves the state of service. Increasing proximity gives increased strength of reinforcement, so that urgency-sensitive sentiment is kept in the highest importance.

Risk-oriented intensity consolidation packages added and pushed negative emotion into its commerce-ready signal, that is, a churn-prediction, refund-likelihood-estimation, and complaint-escalation-detection signal in the following EJ-VADER steps.

$$\mathrm{R}_i = \langle \Theta_i, \widetilde{N}_i, \mathcal{J}_i \rangle \tag{23}$$

In Equation (23) R_i is the gathered risk-oriented intensity vector. The elements combine reinforced transaction pressure. Θ_i, prioritized negative sentiment $\widetilde{N}_i$, and the journey stage $\mathcal{J}_i$. This structure maintains an optimized emotional dominance while keeping it in line with consumer decision risk on commerce streams.

2.7. Retention-Stage Positive Intensity Reinforcement

Retention-Stage Positive Intensity Reinforcement is the act of consciously strengthening positive sentiment articulated after product consumption or service experience, in which long-term consumer value formation takes place. For this reason, this step is of strategic importance in EJ-VADER, given that retention-driven behaviour is conditioned on emotional satisfaction, continuity of trust, and value reinforcement faced by the person rather than a transactional reaction to the moment. Building on the arrangement of transaction stage risk prioritization, this stage shifts analytical reception to remaining devoted, repeat interaction, and potential prospects utilizing optimised sentiments handling.

$$\tilde{P}_i = \hat{P}_i \cdot \rho(\mathcal{J}_i) \tag{24}$$

In Equation (24), reinforced positive intensity at the retention stage is represented as $\tilde{P}_i$: The $\hat{P}_i$ represents an earlier obtained stage-shaped positive sentiment attribute. The application of the reinforcement is contained in this function $\rho(\mathcal{J}_i)$, for which do does not apply except for compilers in the journey stage, which is retention-oriented, such as continued usage/feedback sharing/ subscription renewal.

Loyalty-oriented emotional accumulation describes the effect of repeated positive emotional expressions in making attachment stronger and decreasing the probability of switching away in the future. Consumer behaviour at this stage shows the tone of emotional consolidation and not an isolated reaction, which demands cumulative treatment of positive sentiment.

$$L_i = \sum_{t=1}^{T_i} \tilde{P}_{i,t} \cdot \omega_t \tag{25}$$

In Equation (25), L_i is the score of the loyalty intensity. The term $\tilde{P}_{i,t}$ is positive sentiment that is reinforced in the course of multiple retention interactions indexed by t. Coefficient ω_t catches the relevance of the weight of the interaction in terms of recency and strength of interaction. This accumulation model sustains satisfaction as contrasted to single-event approval.

Advocacy readiness enhancement turns reinforced positive emotion into predictive signals for recommendation, positive word-of-mouth, and brand defense behaviour. Strong satisfaction at the levels of retention often converts into proactive consumer behaviour that adds value beyond the time of actual purchase.

$$A_i = L_i \cdot \zeta_i \tag{26}$$

In Equation (26), A_i represents advocacy readiness intensity. The variable L_i denotes accumulated loyalty sentiment. The factor ζ_i reflects engagement amplification derived from review depth, feedback richness, or voluntary expression effort. This formulation emphasizes that emotionally invested consumers are more likely to influence others.

Retention-centric intensity consolidation packages reinforced positive emotion into a structured signal aligned with long-term consumer value assessment. This consolidation preserves optimized intensity while maintaining connectivity with earlier journey-aware scaling and risk prioritization steps.

$$\mathrm{Q}_i = \langle A_i, \tilde{P}_i, \mathcal{J}_i \rangle \tag{27}$$

In Equation (27), Q_i denotes the retention-focused sentiment vector. The elements combine advocacy readiness. A_i, reinforced positive intensity $\tilde{P}_i$, and the identified journey stage $\mathcal{J}_i$. This representation supports loyalty prediction, lifetime value estimation, and sustained relationship modeling within commerce-driven consumer behaviour analysis.

2.8. Dynamic Threshold Calibration

Dynamic Threshold Calibration refers to the stage-conditioned tuning of decision cut-offs that separate positive, neutral, and negative sentiment intensity classes. Dynamic calibration holds high importance in EJ-VADER, since commerce-driven consumer behaviour shows different action sensitivity across journey stages. Journey-aware intensity scaling and transaction or retention reinforcement produce optimized intensity vectors that require calibrated boundaries to prevent mislabeling weak emotions as actionable signals.

$$\theta_{\mathcal{J}} = [\theta_{\mathcal{J}}^{-}, \theta_{\mathcal{J}}^{0}, \theta_{\mathcal{J}}^{+}] \tag{28}$$

In Equation (28), $\theta_{\mathcal{J}}$ denotes the threshold triplet assigned to a journey stage $\mathcal{J}$. The value $\theta_{\mathcal{J}}^{-}$ represents the negative action boundary, $\theta_{\mathcal{J}}^{0}$ represents the neutral band boundary, and $\theta_{\mathcal{J}}^{+}$ represents the positive action boundary. Stage-wise threshold sets enable commerce-consistent interpretation of the optimized intensity values produced in earlier EJ-VADER steps.

Commerce-outcome anchored threshold learning defines the estimation of thresholds using observed behavioural outcomes, enabling alignment between sentiment intensity and measurable consumer actions. Transaction-stage prioritization and retention reinforcement generate risk and loyalty signals that become more useful when thresholds reflect cancellation, return, complaint escalation, repeat purchase, and advocacy patterns.

$$\theta_{\mathcal{J}}^{*} = \arg\min_{\theta_{\mathcal{J}}} \sum_{i \in \mathcal{I}_{\mathcal{J}}} (\mathcal{A}(y_i, \hat{y}_i(\theta_{\mathcal{J}}))) \tag{29}$$

In Equation (29), $\theta_{\mathcal{J}}^{*}$ denotes the optimized threshold set for stage $\mathcal{J}$. The index set $\mathcal{I}_{\mathcal{J}}$ collects interactions tagged with stage $\mathcal{J}$. The symbol y_i denotes an observed commerce outcome label, and $\hat{y}_i(\cdot)$ denotes the predicted label generated by applying thresholds to EJ-VADER intensity signals. The function $\mathcal{A}(\cdot)$ represents a penalty capturing mismatch between predicted and observed behaviour.

Risk–loyalty balanced boundary adjustment defines boundary tuning that respects asymmetric commerce effects: transaction negativity drives rapid behaviour shifts, retention positivity drives sustained value growth. Calibration integrates the risk vector and the retention vector to balance action thresholds.

$$\Delta_{\mathcal{J}} = \lambda_{\mathcal{J}} \cdot \mathbb{E}[\Theta_i \mid \mathcal{J}] - (1 - \lambda_{\mathcal{J}}) \cdot \mathbb{E}[A_i \mid \mathcal{J}] \tag{30}$$

In Equation (30), $\Delta_{\mathcal{J}}$ denotes the stage adjustment signal. The term Θ_i represents reinforced transaction pressure and A_i represents advocacy readiness. The coefficient $\lambda_{\mathcal{J}}$ controls emphasis toward risk or loyalty within a stage, using expected values conditioned on $\mathcal{J}$.

Calibrated decision assignment defines the final mapping from optimized intensity into commerce-interpretable sentiment states used by later attribute alignment and behavioural mapping steps. Calibrated thresholds convert continuous signals into stable action categories aligned with consumer decision dynamics.

$$\hat{s}_i = \begin{cases} -1, & \bar{C}_i \le \theta_{\mathcal{J}_i}^- \\ 0, & \theta_{\mathcal{J}_i}^- < \bar{C}_i < \theta_{\mathcal{J}_i}^+ \\ +1, & \bar{C}_i \ge \theta_{\mathcal{J}_i}^+ \end{cases} \quad (31)$$

In Equation (31), $\hat{s}_i$ denotes the calibrated sentiment state for interaction i. The term $\bar{C}_i$ denotes stabilized journey-aware intensity, and $\mathcal{J}_i$ denotes the journey stage label. The thresholds $\theta_{\mathcal{J}_i}^-$ and $\theta_{\mathcal{J}_i}^+$ provide stage-conditioned boundaries, enabling optimized, commerce-aligned sentiment classification.

2.9. Attribute-Level Sentiment Alignment

Attribute-Level Sentiment Alignment refers to the structured association of calibrated sentiment intensity with specific commerce attributes that shape consumer judgement. This step holds strong importance in EJ-VADER, since consumer behaviour rarely responds to overall emotion alone. Decisions emerge from evaluations of price fairness, product quality, delivery reliability, service responsiveness, and perceived value. Dynamic threshold calibration yields stabilized sentiment states; attribute alignment transforms those states into actionable, attribute-focused behavioural signals.

$$\mathcal{A}_i = \{a_{i1}, a_{i2}, \dots, a_{im}\} \quad (32)$$

In Equation (32), $\mathcal{A}_i$ represents the attribute set that was identified in the consumer interaction i. Each element a_{ij} refers to an attribute relevant to commerce in the text, that is, referenced explicitly or implicitly: a price, a quality, a delivery, support, or consistent experience. This extraction retains the behavioural granularity necessary for optimized interpretation of the sentiments.

Attribute-specific intensity projection assigns a calibrated intensity of sentiment to each identified attribute. Consumer dissatisfaction toward one attribute may be dominating behaviour despite the overall sentiment seemingly being in balance. EJ-VADER has this projection in order to prevent the dilution of important attribute-level emotion.

$$\phi_{ij} = \hat{s}_i \cdot w_{ij} \cdot v_{\mathcal{J}_i} \quad (33)$$

Whereas in Equation (33), ϕ_{ij} is the sentiment intensity as it is projected on the attribute a_{ij}. The caloric variable was conducted as a Beta instance named variable. User variable to construct the calibrated sentiment state as $\hat{s}_i$. Weight w_{ij} These weights reflect the strength in relevance from lexical proximity and emphasis.

The use of the coefficient $v_{\mathcal{J}_i}$ allows various representations of the situation, which are implemented with the help of reflecting on the situation. Normalization for attribute dominance produces normal data conditions of comparability between attributes and is accomplished by stabilizing the intensity magnitude. This process helps against an exaggerated attention for verbose expressions while keeping the behavioural salience of the concise but forceful emotional cues.

$$\bar{\phi}_{ij} = \frac{\phi_{ij}}{\sum_{k=1}^{m} |\phi_{ik}| + \epsilon} \quad (34)$$

In Equation (34), $\bar{\phi}_{ij}$ represents the normalized intensity of attribute sentiments. The denominator is an aggregate of absolute attribute intensities in the same interaction. The constant ϵ guarantees that there is no instability with the sparse presence of attributes. Normalization allows for the fair comparison between attributes that affect consumer choice.

Commerce-ready attribute sentiment vector packages the aligned attribute intensities into a structured attribute vector that is suitable to take as input to behaviour mapping and prediction. This vector maintains the optimized emotion - journey logic set up in previously executed steps in EJ-VADER.

$$\mathrm{M}_i = \langle \bar{\phi}_{i1}, \bar{\phi}_{i2}, \dots, \bar{\phi}_{im}, \mathcal{J}_i \rangle \quad (35)$$

In Equation (35), M_i is the final attribute aligned sentiment structure. Each normalized attribute intensity is combined with the journey stage label. $\mathcal{J}_i$ allowing for an accurate linkage between an emotional response, attribute evaluation, and consumer behavioural tendency in commerce-driven analysis.

2.10. Behavioural Signal Mapping

Behavioural Signal Mapping for how optimized sentiment intensity (in terms of attributes) would translate into explicit forms of behavior, in terms of consumer behavior. This step is crucial in EJ-VADER, as the success of sentiment analysis is only attained after it is converted to interpretable behavioural outcomes. Outputs of sentient information produced by Attribute-Level Sentiment Alignment gives structured signals of emotions; the mapping takes these signals and converts them into behavioural tendencies relevant for commerce, such as continuing or inclined to churn, complain, complaint escalation, loyalty reinforcement, and advocacy-ready.

$$\mathcal{B}_i = h(\mathrm{M}_i) \quad (36)$$

In Equation (36), $\mathcal{B}_i$ is used to denote the behavioural signal vector for the interaction i. The function $h(\cdot)$ is used to transform the attribute-aligned representation of sentiment, M_i into behaviour-oriented representations by aggregating the influence of emotion across the attributes and the journey context. This way of formulating guarantees that emotional input is put directly to use in the modeling of the consumer action. Behaviour-specific intensity allocation spreads out sentiment influence in independent behavioural dimensions. Consumer behaviour is the result of numerous potential responses rather than just a single action. EJ-VADER disassembles the emotional influence into interpretive behavioural pathways so as not to oversimplify the process.

$$b_{ik} = \sum_{j=1}^{m} \bar{\phi}_{ij} \cdot \eta_{jk} \tag{37}$$

In Equation (37), b_{ik} indicates the intensity contribution towards behavioural dimension k, e.g., churn risk, repeat purchase likelihood, and advocacy strength. The term $\bar{\phi}_{ij}$ is the normalized sentiment of the attribute. To do this, the coefficient appears to be η_{jk} that captures learned influence strength of attribute j on behaviour k, ensures commerce-grounded mapping.

Journey-conditioned behaviour adjustment works on behavioural signals with the help of the customer journey stage, indicating probabilistic behaviour variation in the decision phase. Identical sentiment intensity produces different behavioural results depending on the readiness and commitment of the consumer.

$$\tilde{b}_{ik} = b_{ik} \cdot \psi(\mathcal{J}_i) \tag{38}$$

In Equation (38), $\tilde{b}_{ik}$ is the journey-adjusted behavioural intensity. The function including and applying at stage-specific modulation that is derived from the journey label $\psi(\mathcal{J}_i)$. This adjustment makes the predictive relevance stronger in awareness, transaction, and retention phases.

Optimized behavioural readiness index merges mapped behavioural signals into a stable representation for use in decision systems, dashboards, or predictive analytics. This consolidation maintains emotional color and facilitates actionable interpretation of the commerce.

$$\mathbf{Z}_i = \langle \tilde{b}_{i1}, \tilde{b}_{i2}, \ldots, \tilde{b}_{iK}, \mathcal{J}_i \rangle \tag{39}$$

In Equation (39), Z_i is the final behavioural readiness vector. Optimized behavioural intensity according to journey stage is captured in each component. This is established to guarantee consistency between emotion-based sentiment analysis and observable consumer behaviour outcomes in the case of EJ-VADER.

2.11. Commerce Outcome Validation

Commerce Outcome Validation specifies the project of systematically evaluating the outputs of EJ-VADER when compared to actual consumer actions recorded in real commercial settings. This step has a high importance in meeting the purpose of the work, as optimized sentiment and behavioural signals only gain credibility through the match of actual responses from consumers in purchasing, occurrence of churn, repeat engagement, registration of complaint, or expression of advocacy. Behavioural readiness vectors, as generated in the previous stage, provide the predictive basis tested in terms of observable outcomes of commerce.

$$\mathcal{O}_i = \{o_{i1}, o_{i2}, \ldots, o_{ir}\} \tag{40}$$

In Equation (40) $\mathcal{O}_i$ is the set of observed outcomes of commerce related to interaction i. Each element o_{ir} is a consumer action that is realized in the form of transactional logs, service, or engagement history. This formulation puts a factual standard against which optimised behavioural signals are validated.

Behaviour- outcome consistency measurement is the measure of the consistency between predicted signals of behavioural patterns and realized behavioural patterns of consumers. EJ-VADER optimization is to keep the divergence between emotional inference and the actual decision expression as low as possible. This measurement supports the reliability assessment in journey-aware sentiment handling.

$$\kappa_i = \frac{1}{K} \sum_{k=1}^{K} \mathbb{I}(\tilde{b}_{ik} \sim o_{ik}) \tag{41}$$

Equation (41), κ_i is the consistency score for interaction i. The term $\tilde{b}_{ik}$ refers to journey-adjusted behavioural intensity and o_{ik} refers to the corresponding observed outcome. The way the indicator function $\mathbb{I}(\cdot)$ works is that it measures how close the prediction and action are towards each other; the normalized value is called the validation signal.

Stage-aware validation weighting includes customer journey sensitivity into the evaluation of ECG, where prediction accuracy takes on different meanings for different stages of the commerce. Transaction-stage predictions have a greater immediate impact, while retention-stage predictions impact long-term value estimation.

$$V_i = \kappa_i \cdot \omega_{\mathcal{J}_i} \tag{42}$$

In Equation (42) V_i represents the validation weighted score. The variable κ_i is consistency strength and $\omega_{\mathcal{J}_i}$ Is a journey dependent on the importance weight based on the stage label? $\mathcal{J}_i$. This weighting is in accordance with the strength of the validation versus the risk of commerce and value relevance. Optimized performance aggregation aggregates the validation signals across the consumer dataset to assess the performance effectiveness of EJ-VADER, like overall EJ-VADER, in behaviour prediction and decision support.

$$\mathcal{V}_{\mathrm{EJ}} = \frac{1}{N} \sum_{i=1}^{N} V_i \tag{43}$$

In Equation (43), $\mathcal{V}_{EJ}$ is the aggregate score of validation of commerce for EJ-VADER. The variable N represents the number of the evaluated interactions. This aggregation represents alignment strength between optimized sentiment-driven behavioural modeling and actual consumer outcomes within commerce streams.

3. Dataset Description

The Customer_Behaviour dataset runs hand-made structured records apprehending purchasing patterns, preferences, and interaction attributes that encounter the world from a retail-surrounding point of view. The dataset consists of entries that describe the demographics of customers, types of products, frequency of purchases, spending levels, and response tendencies associated with commercial offerings.

Each record represents the uniform signal of observable behaviour that is useful for the analysis of variation in decision tendencies, variation in levels of satisfaction, and the orientation to loyalty. Transaction-related attributes facilitate the identification of recurring purchase activity, whereas categorical features can be used in the segmentation of interest patterns and style of consumption. Numerical fields are fields for expenditure intensity and engagement magnitude, to support quantitative behavioural modelling.

The dataset is appropriate for analysis in consumer behaviour, sentiment-driven analysis, and decision support analysis. Clear organization of attributes facilitates preprocessing tasks, feature alignment, and mapping behaviors. The dataset is generic, balanced, and adaptable across commerce analytics, behavioural prediction, and optimization-oriented sentiment frameworks-application of retail analysis, marketing insight, and strategy. The choice of the dataset is relevant to the analysis of consumer behaviour, which has structured attributes that serve as purchasing patterns, demographic factors, and decision tendencies. The dataset is made publicly available and highly used for academic research purposes to ensure accessibility and reproducibility. No personally identifiable information is included, and all of the records are anonymized. Data usage is in line with ethical research practices with a focus on aggregated behavioural insights, without individual-level inference. The study follows responsible data handling principles that are appropriate for commerce-oriented analytics.

4. Results and Discussions

Results and discussions constitute a structured interpretation of experimental results, which pays attention to the effect of numerical outcomes on the effectiveness of the model in solving the defined research problem. This section describes observed patterns of performance and milestones of comparative performance using accepted evaluation measures. Classification Accuracy (CL-AC) is the ratio of correctly classified observations to observations, and is a measure of total reliability in a decision. The reported values indicate a clear differentiation between the different approaches evaluated. SMSA gives a classification accuracy of 66.655%, indicating a low degree of discrimination ability. OSP-DM improves accuracy to 70.419%, which means that the accuracy of decision-making is moderately enhanced. EJ-VADER achieves a maximum classification accuracy of 79.852%, showing a closer match between the interpretation of sentiments and real consumer behaviour. The improvement verifies the success of emotion-journey aligned optimization in minimizing misclassification between commerce interactions as reflected in Figure. 1.

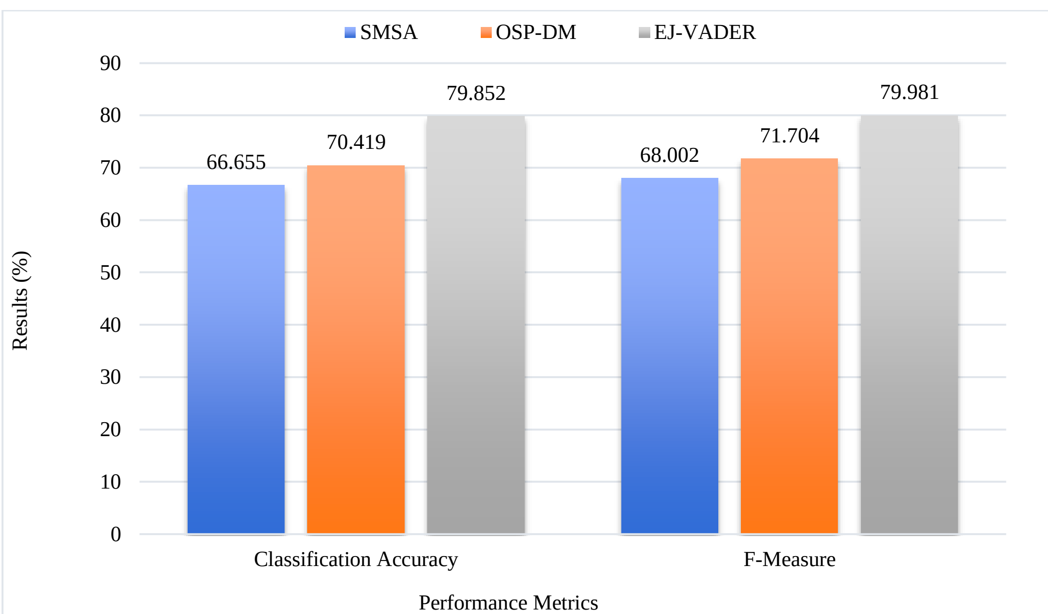

Fig. 1 CL-AC and F-MSR

F-Measure (F-MSR) is the balance of accuracy and recall, which reflects stability in positive identification and error control. SMSA gives an F-Measure of 68.002%, while OSP-DM also achieves 71.704%, so there is incremental improvement. EJ-VADER achieves the highest F-Measure value of 79.9809%, which confirms that it is balanced in identifying behaviourally relevant sentiment signals.

Higher values of the F-Measure score indicate better consistency in handling both positive identification and minimizing errors in the domain of consumer behaviour analysis.

The Fowlkes-Mallows Index (FMI) is used to measure the geometric balance between the precision and recall, indicating the consistency between the predicted and actual class memberships. SMSA has an FMI score of 68.072, indicating poor correlation of accuracy between positive detection and error control. OSP-DM enhances FMI to 71.710, which means the clustering consistency in behavioural classification is good. EJ-VADER achieves the highest FMI value of 79.981, demonstrating high agreement between the predicted sentiment-driven and the actual consumer behaviour pattern as illustrated in Figure 2.

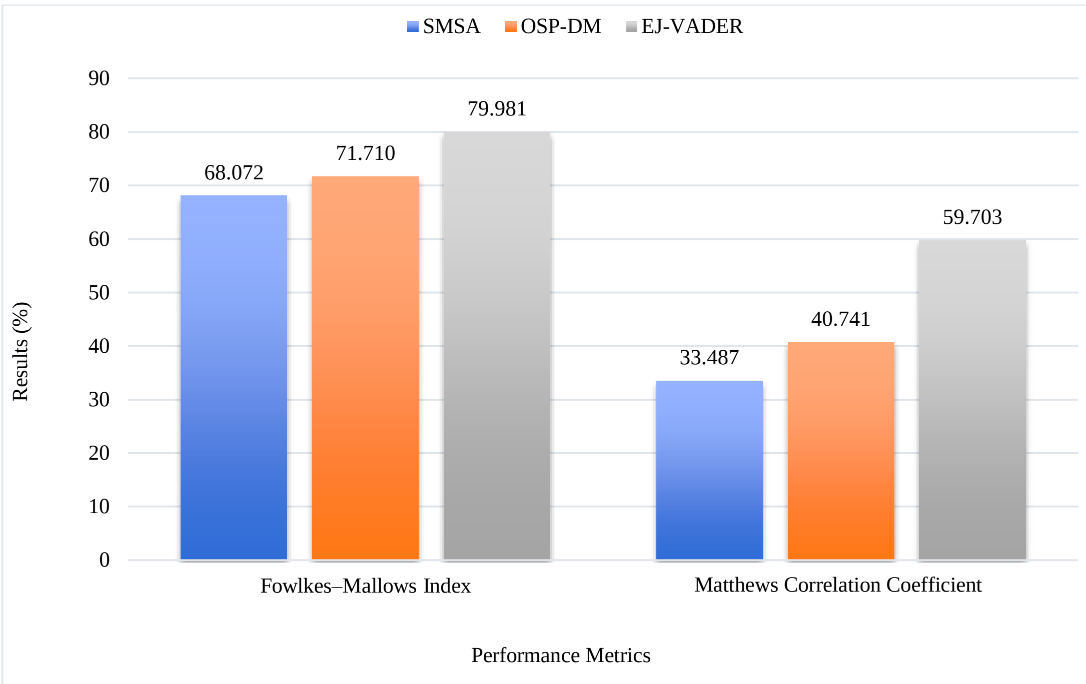

Fig. 2 FMI and MCC

Matthews Correlation Coefficient (MCC): MCC is an evaluation method for the quality of classification, which takes into account all four parts of the confusion matrix and hence, it is a strong measure in case of class imbalance.

SMSA achieves an MCC of 33.487 which implies the weak predictive correlation. OSP-DM 40.741 Moderate Improvement EJ-VADER achieves much higher MCC of 59.703, thereby establishing the much better-balanced prediction strength and reliable behavioural discrimination in commerce-oriented sentiment analysis.

Precision is a percentage of the number of positive instances correctly identified out of the total predicted positives, thus giving an indication of the reliability of positive sentiment detection. SMSA has a precision of 65.070, indicating that it is not very confident in confirming positive classification. OSP-DM increases the precision to 70.806 and indicates improved control of false positives. EJ-VADER shows the highest value of precision, which is 79.787, so it is able to identify behaviourally relevant positive sentiment expressions accurately, as shown in Figure. 3.

Recall is the ratio of the proportion of actual positives that are correctly identified, which means sensitivity to relevant consumer signals.

SMSA achieves a recall value of 71.212 while OSP-DM barely increases recall to 72.626. EJ-VADER shows the highest recall score of 80.175, which shows its increased capability in capturing emotion-critical consumer responses.

Any increased recall and precision satisfies the evidence of improved behavioural sensitivity and reliability in the context of commerce-oriented sentiment analysis.

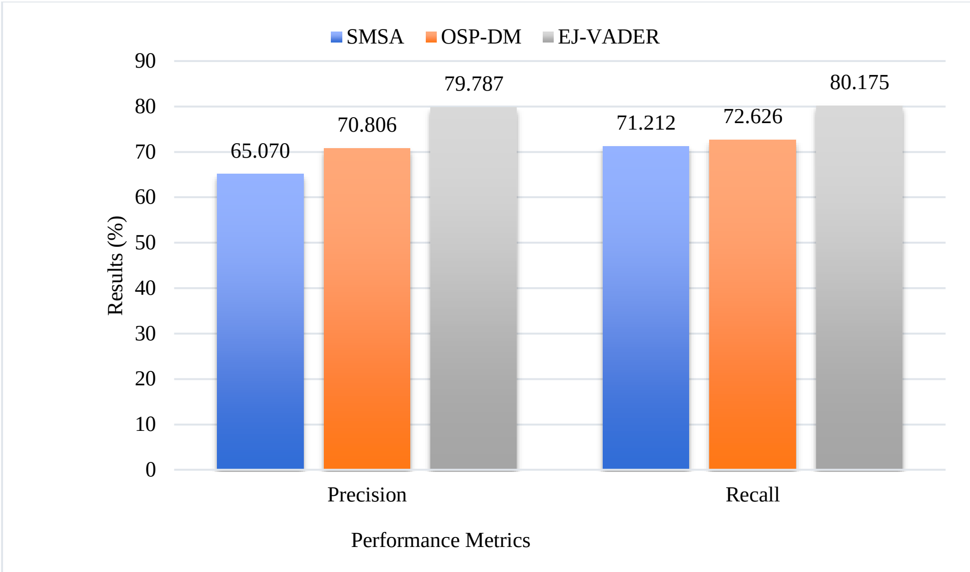

Fig. 3 Precision and Recall

4.1. Behavioural Impact Analysis

Behavioural Impact Analysis is used to assess the impact of optimized sentiment intensity on consumer decision signals. EJ-VADER improves behavioural interpretation by improving emotion journey alignment, especially at transaction and retention stages—negative intensity prioritization, improving identification of risk-oriented behaviour such as churn tendency and complaint escalation. Positive intensity reinforcement increases recognition of loyalty-oriented behavior, such as repeat engagement and advocacy inclination. Attribute-level alignment is used to make behavioural mapping fine-grained by evaluating and correlating the sentiment with decision drivers like price, quality, and service interaction. Behavioural Signal Mapping takes sentiment intensity and translates it into behavioural indicators so that it is easy to identify any association between emotional expression and consumer response patterns in commerce environments.

5. Conclusion

The work proposes a structured framework for sentiment analysis that aims at representing the reality of consumers' decision-making in the context of a commerce environment. Emotional expressions produced by consumers have behavioural meaning that differs across decision stages and requires context-aware interpretation beyond static polarity interpretation. The EJ-VADER framework solves for this need by combining the emotion--journey alignment concept with lexicon-based sentiment intensity estimation. Each stage of the framework facilitates transparent processing from journey identification, intensity scaling, to attribute alignment and behavioural mapping.

The design maintains its interpretability while improving its behavioural relevance, leading to meaningful linkage between emotional expression and readiness for action by the consumer. Focus on journey sensitivity since sentiment interpretation is strengthened in a transaction retention phase where behavioural impact is critical. The framework focuses on clarity, as well as flexibility and commerce alignment, without adding complexity to computation.

EJ-VADER offers a sound methodological basis for sentiment-driven consumer behaviour analysis, applicable to a variety of commerce analytics scenarios. The structured realization of the link between emotional intensity, the context of the journey, and behavioural reasoning contributes towards the realization of more reliable and context-aware sentiment interpretation in consumer-centric research and decision-support systems.

References

[1] Navreen Kaur Boparai, Himanshu Aggarwal, and Rinkle Rani, "Analyzing Fuzzy Semantics of Reviews for Multi-Criteria Recommendations," *Data and Knowledge Engineering*, vol. 152, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Zuhe Li et al., "Multi-Level Correlation Mining Framework with Self-Supervised Label Generation for Multimodal Sentiment Analysis," *Information Fusion*, vol. 99, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Nasir Salari, "Electric Vehicles Adoption Behaviour: Synthesising the Technology Readiness Index with Environmentalism Values and Instrumental Attributes," *Transportation Research Part A: Policy and Practice*, vol. 164, pp. 60-81, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Narong Pleeruxa, and Attawut Nardkulpatb, "Sentiment Analysis of Restaurant Customer Satisfaction During COVID-19 Pandemic in Pattaya, Thailand," *Heliyon*, vol. 9, no. 11, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Hui Li et al., "E-Word of Mouth Sentiment Analysis for user Behavior Studies," *Information Processing and Management*, vol. 59, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yanping Huang et al., "Sentiment Classification using Bidirectional LSTM-SNP Model and Attention Mechanism," *Expert Systems with Applications*, vol. 221, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Pinar Savci, and Bihter Das, "Prediction of the Customers' Interests using Sentiment Analysis in E-Commerce Data for Comparison of Arabic, English, and Turkish Languages," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 3, pp. 227-237, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Handan Cam et al., "Sentiment Analysis of Financial Twitter Posts on Twitter with the Machine Learning Classifiers," *Heliyon*, vol. 10, no. 1, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Muhammad Umer et al., "ETCNN: Extra Tree and Convolutional Neural Network-based Ensemble Model for COVID-19 Tweets Sentiment Classification," *Pattern Recognition Letters*, vol. 164, pp. 224-231, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Serpil Aslan, "A Deep Learning-based Sentiment Analysis Approach (MF-CNN-BILSTM) and Topic Modeling of Tweets Related to the Ukraine-Russia Conflict," *Applied Soft Computing*, vol. 143, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] Tong Fu, Shuyi Yu, and Shiyu Tan, "Peer Effect Matters for the Adoption of New Energy Vehicles: Evidence from Consumer Sentiment Analysis using Chat-GPT," *Energy Economics*, vol. 148, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[12] Christian Graham, and Rusty Stough, "Consumer Perceptions of AI Chatbots on Twitter (X) and Reddit: An Analysis of Social Media Sentiment and Interactive Marketing Strategies," *Journal of Research in Interactive Marketing*, vol. 19, no. 7, pp. 1096-1124, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[13] Ali Jaber Almalki, "Enhanced Sentiment Analysis Framework: Ensemble Attention Enhanced Bidirectional Long-Short-Term Encoder for Accurate Classification of Consumer Reviews," *Alexandria Engineering Journal*, vol. 127, pp. 265-283, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[14] Jing You et al., "Sentiment Analysis Method of Consumer Reviews based on Multi-Modal Feature Mining," *International Journal of Cognitive Computing in Engineering*, vol. 6, pp. 143-151, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[15] Manuel Monge, Ana Lazcano, and Juan Infante, "Monetary Policy and Inflation Rate in the Behavior of Consumer Sentiment in the us. A Fractional Integration and Cointegration Analysis," *Research in Economics*, vol. 78, no. 3, pp. 1-6, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Atanu Dey, and Mamata Jenamani, "Aspect based Sentiment Analysis of Consumer Reviews using Unsupervised Attention Neural Framework," *Applied Soft Computing*, vol. 167, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Ming Xu et al., "Consumer Sentiment Analysis and Product Improvement Strategy based on Improved GCN Model," *Journal of Organizational and End user Computing*, vol. 36, no. 1, pp. 1-38, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Xue Lei, Miao Ma, and Yutong Li, "DRLCCI: A Hybrid Fusion Network Leveraging Disentangled Representation Learning and Cross-Modal Collaborative Interaction for Multi-Modal Sentiment Analysis," *Neurocomputing*, vol. 658, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[19] Jiali You et al., "Hierarchical Reasoning Enhanced Few-Shot Multimodal Sentiment Analysis," *Neurocomputing*, vol. 651, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[20] Xiaoru Li, and Yuxia Lei, "Enhancing Syntactic and Semantic Features Via TextGINConv and Kolmogorov-Arnold Networks for Aspect-based Sentiment Analysis," *Neurocomputing*, vol. 651, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[21] Shengbing Chen, "Dual-Stream Collaborative Network (DSCN): Multimodal Sentiment Analysis Via Modality-Invariant and Modality-Specific Representation Learning," *Neurocomputing*, vol. 652, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yijun Chen et al., "SAFCN: Self-Enhancing Attention Fusion Contrastive Network for Multimodal Sentiment Analysis," *Neurocomputing*, vol. 654, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[23] Bin Sun et al., "Conv-Enhanced Transformer and Robust Optimization Network for Robust Multimodal Sentiment Analysis," *Neurocomputing*, vol. 634, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[24] Yavuz Selim Balcioglu, and Erkut Altindağ, "Social Media Sentiment Analysis: Understanding Consumer Perceptions of Organic Food," *Food and Humanity*, vol. 5, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[25] Yun Yang, "Sentiment Analysis of Consumer Reviews on Online Shopping Platforms using Integrated Deep Learning Models," *ICT Express*, vol. 11, no. 5, pp. 881-887, 2025. [CrossRef] [Google Scholar] [Publisher Link]