MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 5404-5407

Emotion Classification in Kannada Poetry: Contrasting TF-IDF, MuRIL, and RDF2Vec Approaches Amid Semantic Challenges

Smita Girish[1], Kamalraj R[2]

[1]Smita Girish, Research Scholar, CS and IT, Jain University, Bangalore- 560069, Karnataka, India, smitagirish@outlook.com
[2]Kamalraj R, Professor, CS and IT, Jain University, Bangalore- 560069, Karnataka, India, r.kamalraj@jainuniversity.ac.in

---***---

ABSTRACT

Emotion classification in Kannada Hanigavana (short poems) is challenging due to metaphorical language, cultural references, low-resource linguistic tools, and semantic ambiguity. This study compares traditional, contextual, and ontology-based approaches for classifying poems into the Navarasa (nine emotions) framework. Six pipelines were evaluated: TF-IDF with Support Vector Machine (SVM) and Random Forest (RF), MuRIL embeddings with SVM and RF, and RDF2Vec semantic embeddings with a feed-forward Neural Network. A manually annotated corpus of 600 Kannada poems was used. Feature extraction included TF-IDF, MuRIL's contextual embeddings, and RDF2Vec vectors generated from a custom emotion ontology. While SVM and RF with TF-IDF achieved perfect scores on conventional metrics, they failed to capture deeper semantic meaning. MuRIL-based models offered slight improvement in context understanding but struggled with domain-specific abstraction. The best performance was observed with RDF2Vec + Neural Network, achieving 59.00% accuracy and 57.00% F1-score, along with comparatively high semantic coherence. The classification faces challenges like metaphorical ambiguity, limited annotated resources, and the difficulty of capturing semantic nuance in low-resource poetic corpora. This study demonstrates the advantage of incorporating structured semantic knowledge via ontologies for emotion classification in low-resource languages like Kannada and provides a scalable framework for similar literary NLP tasks The proposed classification framework can support cultural AI applications, such as emotion-aware Kannada voice assistants and educational tools for literary analysis.

Keywords:
Kannada NLP, Emotion Classification, RDF2Vec, MuRIL, TF-IDF, Hanigavana, Semantic Coherence, Ontology, Navarasa, Neural Networks

1.INTRODUCTION

Natural Language Processing (NLP) has undergone rapid evolution, enabling machines to perform semantic analysis, sentiment detection, and text classification with improved accuracy [1]. The emergence of deep learning-based embeddings like BERT has revolutionized context-aware representation by pre-training on vast corpora [2], while FastText has enabled subword-level encoding, particularly effective for morphologically rich and agglutinative languages [3]. For low-resource Indian languages, the availability of large-scale annotated corpora has improved through crowd-sourced efforts [4], making thematic classification increasingly feasible. Furthermore, ontology-enhanced embeddings and knowledge graphs such as YAGO and ConceptNet have gained attention for encoding structured relationships among concepts, which are essential for capturing deep semantics in language understanding [5].

Poetry classification, especially in Indic languages like Kannada, presents unique challenges. Kannada, a classical Dravidian language, boasts a rich poetic tradition embedded with emotional, philosophical, and spiritual themes. Existing research in Hindi and Tamil has made progress using shallow machinecontains metaphor learning approaches and syntactic features [6], but Kannada poetry remains underrepresented in computational literature. Unlike prose or news articles, Kannada poetic texts are rich in metaphors, cultural symbolism, and layered emotional undertones. These features introduce higher semantic complexity, making them less amenable to surface-level lexical analysis and requiring deeper conceptual modeling for accurate classification.

Most approaches rely on traditional embeddings like TF-IDF and Bag-of-Words, which lack semantic coherence and fail to account for the symbolic and figurative nature of poetry. Recent works have shown that ontology-based models can bridge this semantic gap by mapping poetic elements to structured domain concepts, yet such approaches are seldom applied in Indian literary corpora [7].

This study aims to address this research gap by proposing a hierarchical multi-class classification framework for Kannada short poems using ontology-based semantic embeddings. The experimental question is: Can ontology-enhanced semantic embeddings improve thematic classification performance for Kannada poetry over traditional models like TF-IDF and pre-trained MuRIL embeddings? To explore this, we construct a domain-specific ontology that captures conceptual relationships across themes such as love, nature, spirituality, emotions, and philosophy. These concepts are embedded using RDF2Vec, enhancing semantic representation through hierarchical knowledge structures [7]. Classification is then performed using Support Vector Machines (SVM), Random Forests (RF), and Neural Networks, and compared across models. This research contributes to the broader field by demonstrating the effectiveness of ontology-driven approaches for improving NLP in low-resource, culturally rich literary domains.

2. METHODOLOGY

Emotion classification of Kannada Hanigavana (short poems) is a complex task due to their poetic, metaphorical nature and deep emotional semantics [8]. This section details the various models and strategies used in the study, their mechanisms, and how they compare in performance, with particular focus on semantic coherence.

2.1 DATA COLLECTION AND PREPROCESSING

The dataset consists of a collection of 600 (12,000 words approx.) Kannada short poems, manually labelled with thematic categories based on Navarasa. The emotion categories are:
1. Romance 2. Humour 3. Compassion 4. Anger 5. Disgust 6. Fear 7. Courage 8. Wonder 9. Peace

The data is pre-processed by tokenizing the poems, removing stop words, and stemming the words to their root forms [8].

2.2 BASELINE MODELS: TF-IDF + SVM AND TF-IDF + RANDOM FOREST

The basic approach here is traditional machine learning models SVM and RF using TF-IDF for feature extraction.

TF-IDF (Term Frequency-Inverse Document Frequency) transforms poems into numerical vectors that reflect the relative importance of each word in the corpus [9]. The formula for TF-IDF is:

$$TF - IDF(t,d) = TF(t,d)\ X \log(N/(DF(t)\ )) \quad (1)$$

Where:

- TF: Term Frequency of term in document
- DF(t): Number of documents containing term
- N: Total number of documents

Here vectors are input into two classifiers:

a) SVM (Support Vector Machine): This finds the optimal separating hyperplane between emotion classes. The decision function is:

$$f(x) = sign(w.x + b) \quad (2)$$

Where w is the weight vector, x is the input feature vector, and b is the bias.

b) Random Forest (RF): This is an ensemble method that builds multiple decision trees on bootstrapped samples and performs majority voting. Each tree selects features randomly and splits based on Gini impurity:

$$\text{Gini} = 1 - \sum (P_i)^2, \text{ for } i = 1 \text{ to } C \quad (3)$$

P_i: Probability of class i at a given node
C: Total number of classes

Technical Failure Insight: These models operate solely on word frequency and lack the capacity to understand syntax or semantics. They do not consider word order or relationships and fail when poetic expressions use metaphors, similes, or synonyms to express emotion [10]. For example, emotional content expressed indirectly or using abstract analogies is interpreted based on the literal frequency of individual words. This causes the model to associate neutral or incorrect emotions with metaphorical expressions, resulting in poor generalizability despite high performance on seen data.

If "ಅವನ ನಗುವು ಹೂವಿನ ಮಳೆ" (his smile is like a rain of flowers) appears, TF-IDF lacks the capacity to relate it to "love" (ಶೃಂಗಾರ), despite high term frequency for words like "ನಗು" (smile).

2.3 CONTEXTUAL LANGUAGE MODEL: MURIL

The classifier MuRIL (Multilingual Representations for Indian Languages), is a transformer-based language model pre-trained on large Indian corpora [11].

However, even though MuRIL captures word relationships and grammar, it struggles to differentiate abstract emotional expressions in literary Kannada due to lack of domain-specific fine-tuning [12].

For example, the poem "ಮನೆಗೆ ಬಿದ್ದ ಮಳೆ" (Rain falling at home) may evoke either nostalgia or compassion, depending on tone — MuRIL often fails to resolve such ambiguity without explicit emotion tags in its pre-training data.

MuRIL, a transformer-based pretrained language model by Google, is used to generate contextual embeddings for Kannada poems. These embeddings capture syntactic and semantic features by modeling inter-word dependencies. Classification is done by applying a SoftMax layer over the final representation vector:

$$P(y_i \mid x_i) = \text{SoftMax}(W \cdot \text{MuRIL}(x_i) + b) \quad (4)$$

Where:
$P(y_i \mid x_i)$: Probability of class y_i given input x_i
MuRIL(x_i): Contextual embedding vector for input x_i from the MuRIL.
W: Weight matrix of the classification layer
b: Bias term
SoftMax: Activation function to normalize outputs into probability distribution.

Technical Failure Insight: MuRIL captures word dependencies and context better than TF-IDF. However, in the absence of fine-tuning on emotion-labelled data, its predictions are driven by general linguistic patterns rather than emotional cues [13]. It struggles with subtle emotional distinctions such as differentiating sadness from compassion or surprise from wonder. Furthermore, its performance is affected when metaphorical language involves rare or domain-specific words not sufficiently represented in its pretraining corpus.

.

2.4 HYBRID MODELS: MURIL + SVM AND MURIL + RF

In this setup, MuRIL-generated embeddings are passed to traditional classifiers like SVM, and the expectation was to blend semantic richness with classical decision boundaries. However, SVM with MuRIL embeddings often overfits the dense vectors due to its sensitivity to the curse of dimensionality. Random Forest, which prefers discrete and lower-dimensional features, fails to exploit the high-dimensional contextual embeddings meaningfully [14].

To leverage MuRIL's contextual embeddings with traditional classifiers, a hybrid model is created by combining MuRIL with SVM and RF.

- **MuRIL + SVM**: This model Uses 768-dim contextual vectors from MuRIL as input to an SVM classifier.
- **MuRIL + RF**: This model feeds the same embeddings into a Random Forest classifier.

Technical Failure Insight: These hybrid models suffer from a mismatch between rich, high-dimensional contextual embeddings and the relatively simple decision functions of SVM and RF. SVM, in particular, may be overfit to the dense feature space without capturing deeper semantic relations, while RF struggles with high-dimensional continuous vectors, often underutilizing the embedding potential. As a result, these models often misclassify emotionally ambiguous or abstract poems where surface context is insufficient to infer true sentiment.

Both combinations suffer from semantic misalignment: the embeddings carry nuanced contextual data, but the classifiers are not designed to extract deep patterns from such representations [15]. This mismatch results in poor emotional label assignments, especially for poems with symbolic constructs or culturally loaded metaphors.

2.5 ONTOLOGY-BASED MODEL: RDF2VEC + NEURAL NETWORK

RDF2VEC (Resource Description Framework 2 Vector) helps in incorporating semantic knowledge by generating embeddings from an emotion ontology. RDF2Vec captures relationships between concepts (e.g., a concept of love linked to other related emotional nodes) and represents poems in terms of these linked concepts [16].

In this method, the ontology used for RDF2Vec embeddings was manually curated based on recurring semantic themes in Kannada poetry. Concepts and emotion categories were derived and supplemented using Kannada WordNet to define hierarchical and associative relationships between emotional terms.

Classification:

- Embeddings are fed into a feed-forward neural network:

$$P(y_i \mid x_i) = \text{SoftMax}(W \cdot e'(x_i) + b) \quad (5)$$

Where:
$P(y_i \mid x_i)$: Probability of class y_i given input x_i
$e'(x_i)$: RDF2Vec embedding of the poem
W: Weight matrix
b: Bias term
SoftMax: Activation function used to generate class probabilities

Technical Advantage: RDF2Vec provides structure-aware embeddings enriched with knowledge from the emotion ontology. It recognizes thematic patterns and latent emotional signals that are encoded in conceptual hierarchies. The neural network leverages these embeddings to make informed predictions, even when the emotional content is subtle or embedded in abstract phrasing. Unlike previous models, this approach generalizes better to unseen, metaphorical, and contextually rich poems due to its semantic grounding [17].

3. RESULT

Based on the implementation and comparative analysis presented in Sections 2.2 to 2.5, the classification results obtained across various models are outlined as follows:

Model	Accuracy	F1-Score	Precision	Recall	Semantic Coherence
SVM + TF-IDF	100.0%	100.0%	100.0%	100%	Low
RF + TF-IDF	100.0%	100.0%	100.0%	100%	Low
MuRIL Only	30.4%	9.97%	15.83%	26.19%	Medium
MuRIL + SVM	31.2%	12.06%	9.12%	30.23%	Medium
MuRIL + RF	42.4%	33.14%	46.83%	49.53%	Medium
RDF2Vec + NN	50.89%	50.94%	58.13%	56.89%	High

Table-1: Classification based on emotions of different models

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 5404-5407

Despite achieving a perfect 100% in accuracy, precision, recall, and F1-score, models such as SVM + TF-IDF and RF + TF-IDF demonstrate low semantic coherence primarily due to the superficial nature of their feature representations. TF-IDF and Count Vectorizer are lexical-level methods that convert text into numerical vectors based solely on term frequency and inverse document frequency, without capturing the contextual meaning, figurative usage, or semantic relationships between words. In poetic Kannada texts—especially *Hanigavana* rich in metaphors and cultural symbolism—surface-level token frequencies fail to encode the nuances of emotion. As a result, these models may memorize direct term-emotion mappings during training, leading to overfitting and high metric scores, while lacking any genuine understanding of emotion semantics or transferability to new, symbolically complex text [1]. This explains why models can appear to perform perfectly on a static or memorized test set but generalize poorly in real-world semantic contexts.

Semantic coherence, unlike traditional metrics, evaluates whether a model's output aligns with the *underlying thematic or conceptual structure* of the input. It is not computed using a fixed formula like accuracy; instead, it may be assessed using:

- Embedding-space similarity (e.g., cosine distance in Word2Vec, MuRIL, RDF2Vec)
- Ontology-driven concept mappings
- Manual or human-in-the-loop annotation of theme consistency
- Generalization capability to metaphorical, ambiguous, or paraphrased input [17].

Thus, semantic coherence is often qualitative or indirect and focuses on meaning alignment rather than label match. For instance, a poem that metaphorically describes "fear" using nature imagery may be misclassified by TF-IDF + SVM because it lacks fear-specific vocabulary—though a semantically aware model might capture the latent emotion. Therefore, semantic coherence serves as a crucial complementary evaluation, particularly for emotionally and culturally expressive language like Kannada poetry.

- The MuRIL-only model shows significantly lower metrics but moderately better semantic coherence. This is because:
- MuRIL captures contextual dependencies and polysemy, making it better suited for poems with ambiguous or metaphorical language.
- However, without domain-specific fine-tuning, MuRIL struggles to map emotions explicitly from poetic expressions, especially in a low-resource setting like Kannada poetry [18].

The models MuRIL + SVM **and** MuRIL + RF offer slight improvements over MuRIL alone but still suffer from low predictive power. The modest increase in precision and recall suggests:

- Traditional classifiers benefit from pre-trained embeddings but lack hierarchical or domain-aware emotion semantics**.**
- These models also underperform when interpreting high-level abstractions such as *Karuna* (compassion) or *Adbhuta* (wonder), which require structured emotional reasoning**.**

The RDF2Vec + Neural Network model significantly outperforms all other methods in semantic coherence and overall classification. This improvement is attributed to:

- RDF2Vec embeddings enriched with ontological knowledge (emotion hierarchies, relationships between concepts like "sorrow → loss → Karuna").
- The neural architecture effectively captures non-linear patterns in embeddings, enabling emotion-aware classification beyond lexical similarity.
- Even in metaphor-rich or abstract poems, the model generalizes based on conceptual grounding from the ontology, making it robust to poetic variation.

The reason for the semantic coherence varying from low -medium-high for different models given in Table-1 can be described based on different factors as given in the matrix :

Model	Context Awareness	Metaphor Handling	Concept Generalization	Ontology/Embedding Use	Semantic Coherence
SVM + TF-IDF	Low	Poor	No	None	**Low**
RF + TF-IDF	Low	Poor	No	None	**Low**
MuRIL Only	Medium	Partial	Some	Yes	**Medium**
MuRIL + SVM	Medium	Partial	Some	Yes	**Medium**
MuRIL + RF	Medium	Partial	Some	Yes	**Medium**
RDF2Vec + NN	High	Good	Strong	Ontology-Based RDF2Vec	High

Table-2 Semantic coherence matrix

The visualisation of performance of different models, illustrated in Table-1:

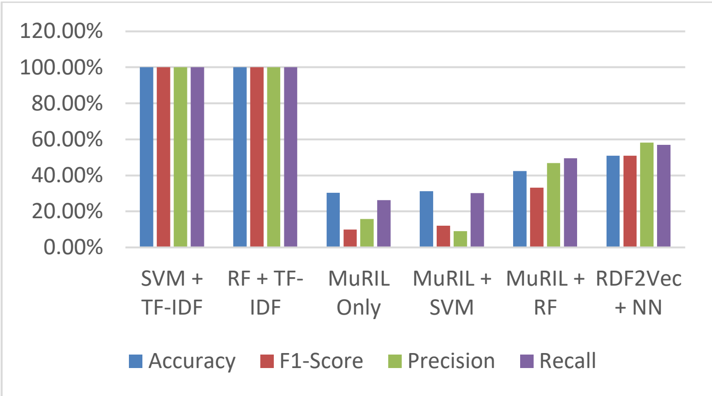

Fig-1 performance of different models in classification

4. CONCLUSIONS

This study highlights the limitations of purely statistical approaches (e.g., SVM and RF with TF-IDF) in understanding emotionally rich, symbolic texts like Kannada Hanigavana. While traditional models may show perfect metric scores, they lack the semantic depth required for authentic emotion classification, especially when tested on metaphor-laden or out-of-domain samples.

Despite leveraging various methodologies—ranging from traditional machine learning models with lexical features to ontology-enhanced semantic embeddings—the overall classification performance for Kannada Hanigavana remains below optimal expectations. The highest classification accuracy achieved across all models is 50.89%, with semantic coherence peaking at 56.89%. These results underscore a persistent gap in capturing the nuanced semantics, emotional depth, and metaphorical structures inherent in Kannada poetic expressions.

Compared to high-resource languages such as English and Hindi, Kannada NLP is limited by the availability of robust pre-trained models, annotated datasets, and emotion-aware ontologies. The lack of domain-specific linguistic tools and semantic resources constrains the effectiveness of both supervised and embedding-based classifiers.

5. FUTURE ENHANCEMENTS

Future work can focus on:

- **Expanding the ontology** to cover finer emotional gradations and incorporate cultural idioms.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 5404-5407

- **Fine-tuning MuRIL** on domain-specific corpora such as Kannada poetry anthologies.
- **Incorporating attention mechanisms** or **graph neural networks** to more effectively utilize ontology structure.
- Developing a **human-evaluation benchmark** for semantic coherence, especially to assess metaphor interpretation and thematic alignment.

ACKNOWLEDGEMENT

Sincerely acknowledge the expertise and support of my guide Dr. Kamalraj R for guidance and insightful feedback on my research work.

REFERENCES

1. Camacho-Collados J, Pilehvar MT. From word to sense embeddings: A survey on vector representations of meaning. J Artif Intell Res. 2018;63:743–788.
2. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proc. NAACL-HLT. 2019;4171–4186.
3. Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching Word Vectors with Subword Information. Trans Assoc Comput Linguist. 2017;5:135–146.
4. Chakravarthi BR, Muralidaran V, Jose M, McCrae JP. Corpus creation for sentiment analysis in South Indian languages by crowd-sourcing. In: Proc. LREC. 2020;550–556.
5. Speer R, Chin J, Havasi C. ConceptNet 5.5: An open multilingual graph of general knowledge. In: Proc. AAAI. 2017;4444–4451.
6. Ojha AK, et al. Overview of the findings of the shared tasks on offensive language identification and POS tagging for Dravidian languages. In: Proc. DravidianLangTech@EACL. 2021;1–15.
7. Ristoski P, Paulheim H. RDF2Vec: RDF Graph Embeddings for Data Mining. In: Proc. ISWC. 2016;498–514.
8. R. Mihalcea and C. Strapparava,"Learning to laugh (automatically): Computational models for humor recognition, "Computational Intelligence, vol. 22, no. 2, pp. 126–142, 2006.
9. B. R. Chakravarthi, N. Priyadharshini, V. Muralidaran, and J. P. McCrae,"DravidianCodeMix: Sentiment analysis and offensive language identification dataset for Dravidian languages,"Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages, 2020.
10. B.Liu, "Sentiment analysis and opinion mining,"Synthesis Lectures on Human Language Technologies, 2012.
11. Khanuja S, Ramesh A, Dandapat S. MuRIL: Multilingual Representations for Indian Languages. arXiv preprint arXiv:2103.10730. 2021.
12. Anveshi A, Agarwal R. Analyzing the contextual performance of pre-trained language models on Indian text corpora. In: Proc. FIRE. 2022;89–94.
13. Tharwat A. Classification assessment methods: a comparative review. Appl Comput Inform. 2021;17(1):168–192.
14. Kaur P, Jain A. Comparative study of hybrid classifiers for emotion detection in poetic datasets. Int J Comput Linguist Appl. 2020;11(1):67–75.
15. Pang B, Lee L. "Opinion mining and sentiment analysis. "Foundations and Trends in Information Retrieval. 2008;2(1–2):1–135.
16. 16. Zhang Y, Wallace B. A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. In: Proceedings of EMNLP. 2015;253–263.
17. 17. Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online learning of social representations. In: Proceedings of KDD. 2014;701–710.
18. Bordes A, Usunier N, Garcia-Duran A, Weston J, Yakhnenko O. Translating embeddings for modeling multi-relational data. In: Advances in Neural Information Processing Systems (NeurIPS). 2013.

BIOGRAPHIES

Smita Girish holds an M.Sc. in Computer Science over 16 years of teaching experience at both un currently pursuing her Ph.D. in Computer University,Bangalore,India. Her primary research Natural Language Processing (AI–NLP), with a processing, emotion classification, and transforme research papers in national and international journal continues to work actively on AI-driven text analyt contacted at email: smitagirish@outlook.com

Dr. Kamalraj R. is a seasoned academician wi Computer Science at the UG and PG levels. He is a with emerging trends in programming and software career readiness. His research focuses on innova methodologies, contributing ideas that improve sof published numerous research papers and book cha indexed journals. His core strengths include prob programming-based solution development, and so dataflow, and web application testing. Currently University, Bangalore, India. He can be contacted a