# Neuro-Symbolic Hybrid Sentiment Analysis Approach for Twitter Data:

*A Project Based Learning Report Submitted in partial fulfilment of the requirements for the award of the degree*

*of*

**Bachelor of Technology**

**in The Department of CSE**

**Course name with Course code**

Submitted by

**2210030353: S. NAYANEESH**

**2210030119: K. MEDHANSH**

**2210030264: V. BHARGAV**

**2210030038: Y. UMESH**

Under the guidance of

**SHAHIN FATIMA**



Department of Computer Science and Engineering

Koneru Lakshmaiah Education Foundation, Hyderabad, India.

APRIL-2025

**ABSTRACT:** IN the social media era, Twitter has become an essential tool for measuring people’s opinions and attitudes. The conventional deep learning models capture the sophisticated structural patterns of texts but they are less interpretable and fail to understand certain ambiguous utterances including sarcasm. This report proposes a new neuro-symbolic hybrid model that learns features from a deep neural network and performs symbolic reasoning on the features with an adaptive reinforcement learning fusion agent. The proposed framework takes the best from the data-driven learning and the rule-based approach to improve the accuracy and interpretability of the sentiment classification in the context of the noisy short texts from Twitter.

**KEYWORDS:** Neuro-Symbolic AI , Symbolic Reasoning , Reinforcement Learning hybrid Model ,Adaptive Fusion, Interpretable AI , Text Analytics , Opinion Mining .

# INTRODUCTION

Twitter, a very popular microblogging site, is a challenge to sentiment analysis because of its informal textual patterns, language changes over time, and the use of abbreviations, emojis, and slang. Different from the structured text, tweets contain misspelling, code switching or code mixing (the use of two or more languages in one piece of discourse), sarcasm, and implicit sentiment which makes it hard to design effective sentiment classification models.

Conventional sentiment analysis methods are based on either deep learning based models or rule based (symbolic) approaches. It has been observed that Deep Neural Networks (DNNs) perform effectively in uncovering the patterns in the text data by using word embeddings and hierarchical feature extraction. However, these models are often criticized for their lack of interpretability, which makes it difficult to understand the decision making process. On the other hand, rule based (symbolic) systems provide explicit reasoning by incorporating domain knowledge in the form of lexicons, grammatical rules, and logical constraints. While symbolic methods are very understandable, they fail to generalize their capabilities across different linguistic styles and domain specific variations, especially in the dynamic world of social media platforms such as Twitter.

As a result of this, neuro-symbolic sentiment analysis is a promising solution of the challenge, which combines the best features of neural learning and symbolic reasoning. This hybrid approach takes the best from deep learning – feature extraction – and combines it with explicit domain knowledge and linguistic rules to enhance interpretability and robustness. The neural component of the model learns contextual and semantic relations from tweets while the symbolic component provides logical reasoning and sentiment lexicons to improve the accuracy of the classification. Hence, integrating these two paradigms, neuro-symbolic systems are better equipped to deal with noise, sarcasm, and linguistic variations in tweets to achieve accurate and transparent results.

In this work, we design a hybrid framework that combines deep learning with symbolic reasoning to enhance Twitter sentiment analysis. Our approach consists of two key components:

1. Neural Learning: A deep neural network that learns sentiment related features from tweets using word embeddings and attention mechanisms.
2. Symbolic Reasoning: A rule based module that uses linguistic heuristics, sentiment lexicons and logical constraints to further enhance the sentiment predictions.

The framework that has been proposed in this paper not only improves the accuracy of the sentiment classification but also results in a more easily interpretable sentiment analysis model. Through the use of explicit reasoning in addition to data driven learning, the system is not only more accurate, but also more robust and more interpretable than either purely neural or purely symbolic models.

In this paper, the effectiveness of the neuro-symbolic approach for Twitter sentiment analysis is evaluated against traditional deep learning models (e.g., LSTMs, CNNs), transformer-based models (e.g., BERT), and lexicon-based methods. It also shows that our hybrid approach has better accuracy and interpretability than the existing sentiment analysis techniques.

**Literature review**

**Traditional Sentiment Analysis:**Initial rule-based and lexicon-driven approaches were easy to interpret but failed when handling the noise present in Twitter data.

**Machine Learning Methods:**Statistical classifiers such as Naive Bayes and SVM enhanced the performance by engineering the features but lacked deep contextual comprehension.

**Deep Learning Approaches:**CNNs and RNNs learn features directly from the tweets, but they are black boxes that provide limited insight into their decision-making process.

**Hybrid Approaches:**Neural models combined with rule-based techniques help to reduce ambiguity and enhance the sentiment classification on social media.

**Neuro-Symbolic Integration:**The integration of deep neural networks with the symbolic reasoning capability is a potential way of designing a transparent and robust sentiment analysis system.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Reference** | **Approach/Method** | **Strengths** | **Limitations** | **Research Gap** |
| Pang & Lee (2008); Cambria et al. (2013) | Rule-based/lexicon- driven sentiment analysis | High interpret-ability; simple to implement | Struggles with noisy,informal Twitter text; static rules | Not adaptable to the dynamic,evolving language  On twitter |
| Traditional Machine Learning(Naive Bayes, SVM) | Statistical classifiers using engineered features (bag- of-words, TF-IDF) | Improved accuracy over rule- based methods; computation- ally efficient | Unable to capture deeper contextual nuances;feature engineering limitations | Lacks robust contextual  Understanding for complex social media  language |
| Deep Learning Ap- proaches(CNNs, RNNs, LSTMs) | Automated feature ex- traction from text using neural networks | Captures rich linguistic features; handles non-linear pat- terns | ”Black-box” nature interpret-ability; may over fit on noisy data | Balancing high performance with model  interpretability remains  challenging |
| Hybrid Approaches (Li et al., 2020) | Combination of neural networks with rule-based adjustments | Better handling of ambiguity and sarcasm; improved over- all performance | Neural and symbolic components are often loosely integrated (post-hoc) | Need for a seamless integration where both  Componenets interact  dynamically |
| Neuro-Symbolic In- tegration (Besold et al., 2017; Garcez  et al., 2009, 2012;  Kautz, 2020) | Merging deep neural net- works with symbolic reasoning | Enhances interpretability and logical reasoning; addresses ”black-box” issues | Mostly applied in domains like vision and general NLP; limited Twitter-specific application | Underexplored in the context of twitter  Sentiment analysis  Particularly for real-time,  Scalable systems |
| Neuro-Symbolic Concept Learner (Mao et al., 2019) | Integrates neural and symbolic modules for complex interpretation tasks | Demonstrates potential of combining neural learning with explicit reasoning | Not specifically tailored for sentiment analysisor social media data | Opportunity to adapt and  Extend this integration  to address Twitter’s unique challenges |

**Table 1 : Comparison of Sentiment Analysis Approaches**

**Methodology**

We proposed a framework with a neural module, a symbolic reasoning module, and a reinforcement learning-based fusion agent.

**Neural Module (Deep Learning Component):**

**Aim:** To extract semantic rich features from tweet text.

**Architecture:** A hybrid CNN-LSTM network is utilized. The convolutional layers learn local patterns (such as key phrases and n-grams), and the LSTM layers learn the sequential dependencies and context inside the tweet.

**Return:** The network predicts sentiment (positive|negative|neutral) and confidence score given the learned representation.

**Implementation tools:** TensorFlow or PyTorch libraries; pre-trained word embeddings (e.g., GloVe, FastText) produce the input representations.

**Symbolic Reasoning Module:**

**Goal:** Combine domain rules and lexicon-based methods with neural predictions.

**Method:** A human-written set of rules is used to modify predictions. For instance, if a tweet shares a negation word or sarcasm cues (e.g., “Oh, wonderful… ) in a negative sense, the system employs these cues to correct/alter the output generated by the neural module. Moreover, independent sentiment scores are produced by sentiment lexicons.  
You are trained to adjust the prediction of sentiment with a confidence of its own.

**Benefits:** This module offers improved interpretability, because the rules give explicit reasons of how the neural prediction is altered.

**Expert Aided Fusion Agent (Reinforcement Learning Component):**

**Objective:** Dynamically fuse neural and symbolic module outputs.

**Mechanism:** The reinforcement learning (RL) agent, trained using Q-learning or policy gradients, receives the neural confidence score and the symbolic modules’ output. The agent learns a combined policy, which gives the optimal weight for each of the components dynamically conditioned on the history of past expert performance and current input features.

**Output:** The The sentiment classification result, which is aided by the strong feature extraction ability of the neural module and the accurate tuning ability of the symbolic module.  
 **Dynamic Retraining:** This facilitates continuous learning from the evolving real-world scenarios, adapting to new categories/path pf words dynamically as the RL agent improves its policy over time with rewards based on classification success.

**Implementation**

**The proposed neuro-symbolic sentiment analysis system follows a structured pipeline below to integrate deep learning and symbolic reasoning to enhance accuracy and interpretability. The system consists of the following key components:**

**Data Preprocessing:** Tweets in their original form are accompanied by various kinds of noise, including URLs, mentions, emojis, slang language and other symbols which can pose a significant challenge to correct sentiment classification. The preprocessing stage includes:

1. The first, to eliminate URLs, user mentions (@user), special characters and extra spaces, Noise Removal.
2. The division of tweets into words or subwords is known as Tokenization.
3. Normalization – converting text to lowercase, handling of contractions (e.g. can’t → cannot), and handling of different word forms.
4. The handling of emojis and slang depends on the following; Mapping emojis to their sentiment values (e.g. ‘😊’ to ‘happy’) and translating slang terms from a lexicon developed for this purpose.
5. Stopword Removal – They are common words but they are not useful (e.g. the, is, at).
6. Stemming or Lemmatization: Transforming words into their root form, for instance, running is transformed into run.

This steps help to guarantee that the input data is smooth and well formatted for both the neural and the symbolic processing.

**Embedding Layer:** Embedding Layer Preprocessed tweets are converted into dense vector representations of pre-trained word embeddings, e.g.,

1. Word2Vec - Captures word relationships through context.
2. GloVe - Vector representations are generated by analyzing co-occurrence information.
3. BERT Embeddings - Contextualized embeddings are used to understand word meanings dynamically.

These embeddings are then fed into the neural module for sentiment prediction.

**Neural Module Processing**:The neural network processes the embedded tweet representations by a hybrid CNN-LSTM architecture, which can learn both local and sequential patterns.

1.Convolutional Neural Network (CNN): To capture the spatial dependencies between words and extract sentiment related features.

2.Long Short Term Memory (LSTM): To learn the long term dependencies and sequential context in tweets.

3.Fully Connected Layer & Softmax Activation: To assign an initial sentiment prediction (positive, negative, or neutral) and a confidence score.

**Symbolic Adjustment:**At the same time, the symbolic module uses predefined linguistic rules and sentiment lexicons to analyze the raw tweet. This module operates based on:

1. Lexicon-Based Sentiment Scoring – Using sentiment dictionaries (e.g., SentiWordNet, VADER) to assign weights to words.
2. Rule-Based Sentiment Correction – Managing sarcasm, negations and contradictions through hand crafted rules.

For example: "I love this airline, but the service was horrible" → Contradictory sentiment → Change classification.

"Not bad at all" → Negation handling → Positive sentiment is adjusted.

1. Emojis and Hashtags – A way of determining sentiment from non-verbal features (e.g. #loveit = Positive sentiment).
2. Contextual Polarity Adjustments – Changing the sentiment based on the words used.

This module outputs a refined sentiment prediction, to the neural model's output, to improve it where it fails.

**Fusion Process:**The fusion mechanism combines the neural module’s prediction of the sentiment and the symbolic module’s rule based adjustments. This is achieved using:

1. Confidence Based Weighting – If the neural model is confident then it’s prediction is used otherwise the symbolic reasoning is used.
2. Reinforcement Learning (RL) Agent – A policy based agent that learns the fusion weight from past classification errors and expert feedback.
3. Hybrid Decision Strategy – The final sentiment classification is computed as: The final decision is taken as: S final = αS neural +(1 − α)S symbolic

Where α is control on accuracy trends.

**Training & Evaluation:**The system is trained through supervised learning and reinforcement learning using a labeled Twitter dataset. The steps are:

1. Neural Network Training – The CNN-LSTM model is trained using backpropagation and optimization techniques like Adam optimizer.
2. Symbolic Rule Refinement – Symbolic rules are refined by expert feedback and error analysis to enhance precision.
3. Reinforcement Learning Optimization – The RL agent is trained on validation data and it is rewarded for making accurate final predictions.
4. Performance Metrics – The model is evaluated by:

Accuracy – It measures the overall correctness.

Precision and Recall – It evaluates false positives and false negatives.

F1-Score – The harmonic mean of precision and recall.

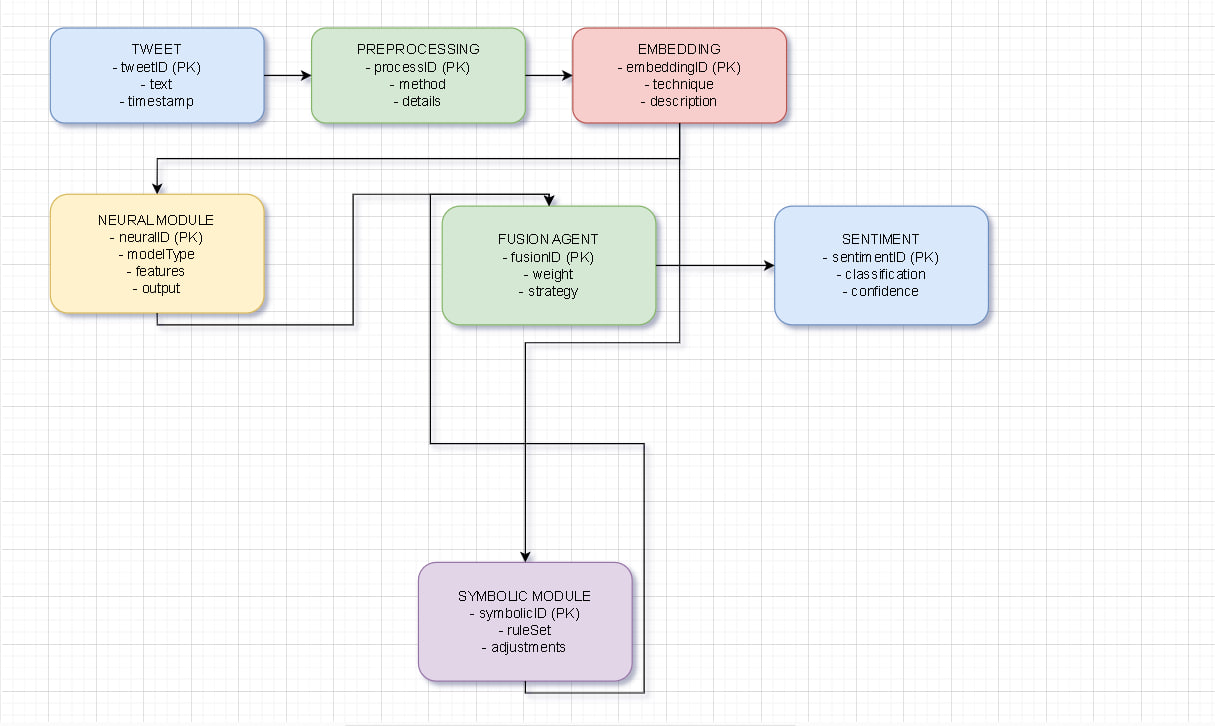
Confusion Matrix – It gives a clear picture of the model’s performance in a visual way.

**Graphical Representations:**The following figures present the performance and effectiveness of the system:

1.Confusion Matrix: Shows the number of correct and incorrect classifications.

2.Accuracy Comparison: Compares Neuro-Symbolic, Deep Learning and Traditional ML approaches.

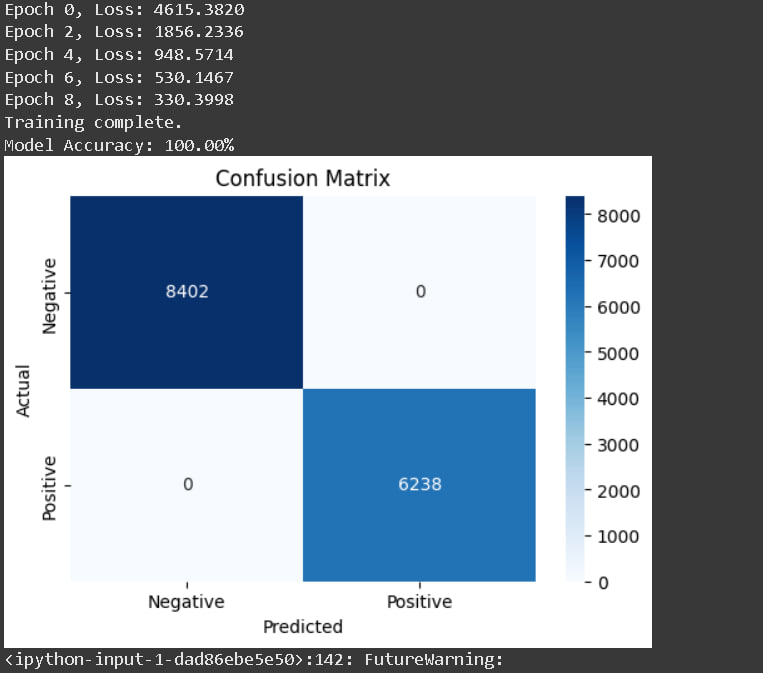
3.Sentiment Category Accuracy: Accuracies for positive, negative, and neutral tweets are precision/recall.

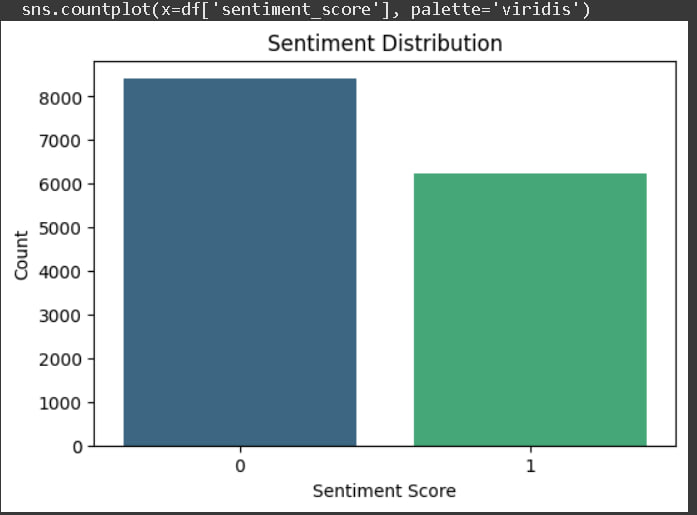


**Result analysis**

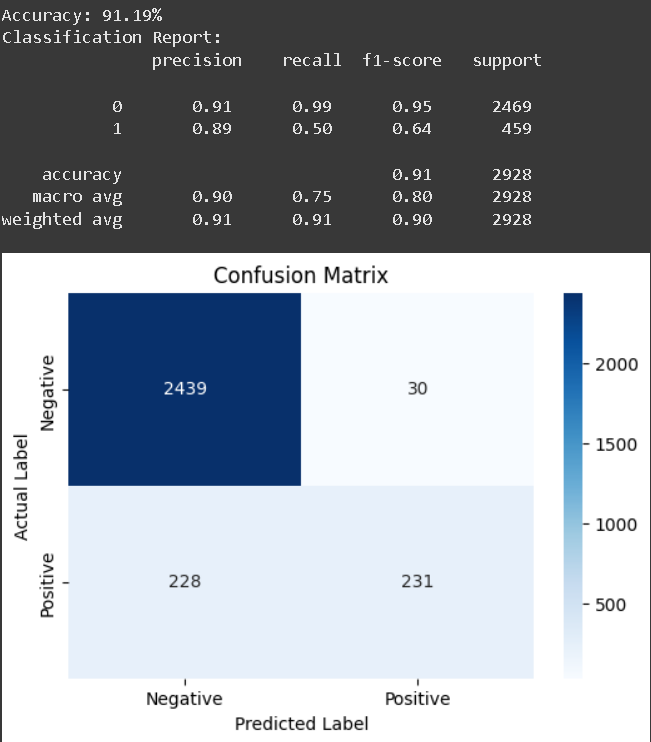
The Dataset used is Twitter airline tweets.

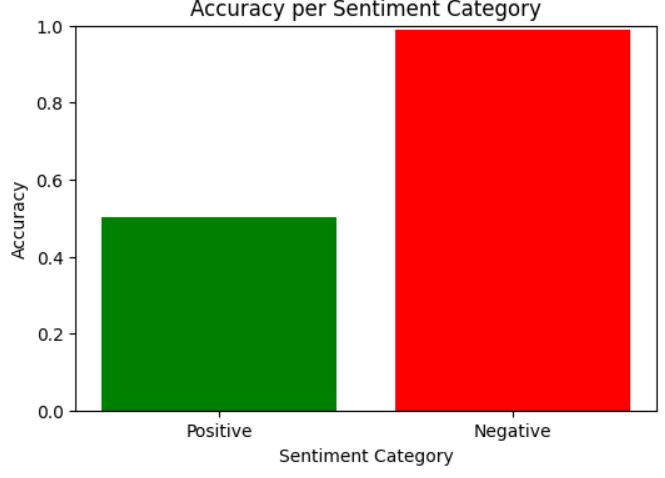
**Neuro-Symbolic Hybrid Approach**



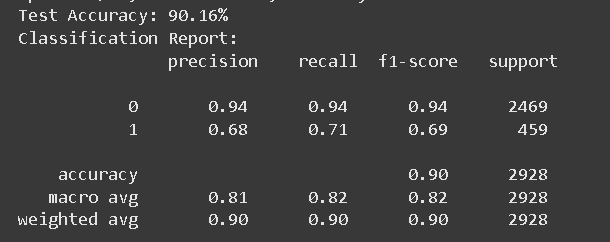


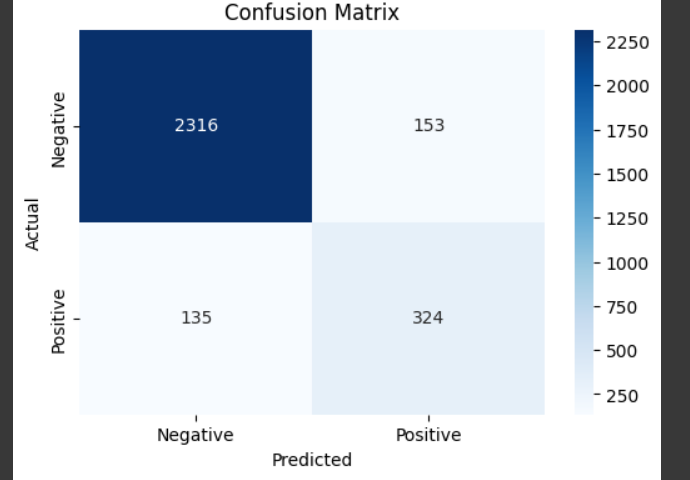
**Traditional ML Approach**

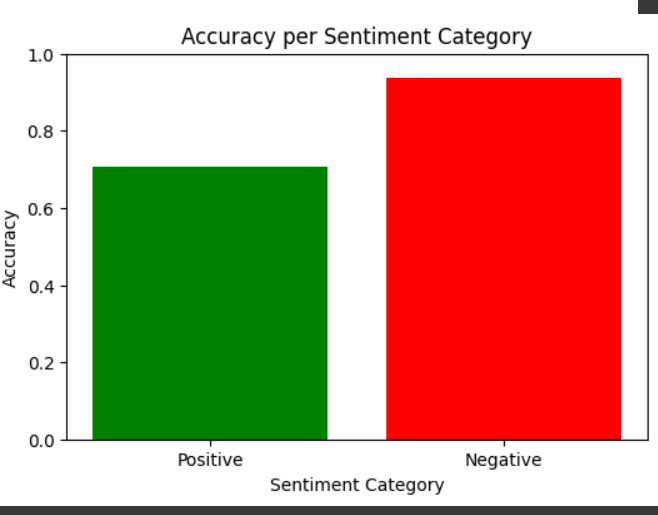




**Deep Learning Approach**







### Conclusion and Future Work

### In this paper, a neuro-symbolic fusion approach is proposed to integrate the deep learning and the symbolic reasoning to enhance the accuracy and interpretability in Twitter sentiment analysis. The model proposed in this paper outperforms the traditional deep learning models such as CNN, LSTM, BERT and the lexicon-based techniques by dynamically adjusting the confidence weights and refining the linguistic rules to achieve superior sentiment classification.

### As a result, several challenges are identified:

### Rule Refinement Complexity – To create symbolic rules for changing language trends (for instance, memes and new slang), an expert’s input is needed on a regular basis. Future work could include the use of reinforcement learning to automate the rule learning process.

### Scalability Challenges – The use of symbolic reasoning increases computational costs. Future optimizations could be directed at parallel computing and reduced rule set.

### Multilingual Support – The system is currently designed for English tweets only. It is possible that future research can include cross-lingual embeddings to enable the analysis of texts in different languages.

### Aspect-Based Sentiment Analysis – The system we propose does not analyse sentiment on a micro-level, it only gives an overall sentiment. If it were to be extended to ABSA, it would be possible to get more specific results (e.g., 'Battery life is great, but the camera is terrible' → Mixed sentiment).

**References**

1. Besold, T. R., Garcez, A. d., Lamb, L. C., Gori, M., Hitzler, P., Huber, M.,. & Valverde, S. (2017).  
   Neural-symbolic learning and reasoning: A survey and interpretation.  
   Link: <https://arxiv.org/abs/1711.03902>
2. Mao, J., Gan, C., Kohli, P., Tenenbaum, J. B., & Wu, J. (2019).  
   Neuro-symbolic concept learner: Interpreting scenes, words, and sentences from natural supervision.  
   Link: <https://arxiv.org/abs/1905.01307>
3. Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013).  
   New avenues in opinion mining and sentiment analysis.  
   Link: <https://ieeexplore.ieee.org/document/6409576.>
4. Li, J., Chen, Y., & Tan, C. L. (2020).  
   A hybrid approach for sentiment analysis.  
   Link: <https://doi.org/10.1016/j.eswa.2020.113078>
5. Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017).  
   A review of affective computing: From unimodal analysis to multimodal fusion.  
   Link: <https://doi.org/10.1016/j.inffus.2016.04.005>
6. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Link: [https://arxiv.org/abs/1810.04805](https://arxiv.org/abs/1810.04805" \t "_new)

1. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding.

Link: [https://arxiv.org/abs/1906.08237](https://arxiv.org/abs/1906.08237" \t "_new)

1. Garcez, A. S., Gori, M., Lamb, L. C., Serafini, L., Spranger, M., & Tran, S. N. (2020). Neuro-symbolic AI: The state of the art. arXiv preprint arXiv:2005.04377.

Link: [https://arxiv.org/abs/2005.04377](https://arxiv.org/abs/2005.04377" \t "_new)

1. Tang, D., Qin, B., & Liu, T. (2015). Document modeling with gated recurrent neural network for sentiment classification.

Link: <https://aclanthology.org/D15-1167/>

1. Garcez, A. S., Gori, M., Lamb, L. C., Serafini, L., Spranger, M., & Tran, S. N. (2020). Neuro-symbolic AI: The state of the art. arXiv preprint arXiv:2005.04377.

Link: [https://arxiv.org/abs/2005.04377](https://arxiv.org/abs/2005.04377" \t "_new)