Uncertainty-Aware Emotion Recognition: A Hybrid Framework Coupling Bayesian Transformers with Interval Type-2 Fuzzy Logic Hybrid Bayesian Transformer and IT2FIS for Emotion Recognition

Afef BEN SAID [1,2], Abdelwaheb GUEDDES [1,2] , Imen JEGHAM [1,2] , Mohamed Ali MAHJOUB[1]

Abstract— **Emotion recognition from textual data is fraught with ambiguity, primarily due to the subjective nature of affective labeling and the linguistic nuances of human expression. While Transformer-based architectures (e.g., BERT, RoBERTa) have established state-of-the-art results in regression tasks, they typically operate as deterministic "black boxes," failing to quantify the confidence of their predictions or provide human-readable rationales. To address these limitations, this paper proposes a novel Neuro-Fuzzy architecture: Context-Aware Type-2 Neuro-Fuzzy Inference (CAT2-NFI). We employ a Transformer with Monte Carlo (MC) Dropout to approximate a Bayesian inference process, extracting both the Valence-Arousal-Dominance (VAD) scores and their associated epistemic uncertainty. These outputs dynamically construct the Footprint of Uncertainty (FOU) within an Interval Type-2 Fuzzy Inference System (IT2FIS). Unlike static fuzzy systems, our approach adapts the fuzzification process to the model's confidence, allowing for robust classification even in noisy or ironic contexts. Experimental results on the EmoBank corpus demonstrate that our hybrid method achieves superior F1-scores compared to crisp baselines while offering transparent, rule-based explanations for its decisions.**

Index Terms— **Affective Computing , Interval Type-2 Fuzzy Logic , Bayesian Deep Learning , VAD Model , Explainable AI (XAI) Affective Computing , Interval Type-2 Fuzzy Logic , Bayesian Deep Learning , VAD Model , Explainable AI (XAI).**

I. INTRODUCTION

The exponential growth of user-generated content on social media has positioned *Affective Computing* as a pivotal field for understanding human behavior. While early approaches focused on categorical classification (e.g., binary sentiment analysis), recent trends have shifted towards dimensional models, specifically the **Valence-Arousal-Dominance (VAD)** framework [1]. Unlike discrete labels, the VAD continuous space captures nuanced emotional states, such as distinguishing *Anger* (low valence, high arousal, high dominance) from *Fear* (low valence, high arousal, low dominance).

A. Related Work and Limitations

State-of-the-art approaches in emotion recognition are currently dominated by Deep Learning (DL), particularly Transformer-based architectures like BERT and RoBERTa [12], [15]. These models excel at capturing semantic dependencies and contextual nuances, achieving impressive performance benchmarks [4]. However, despite their predictive power, purely neural approaches suffer from two critical limitations when applied to real-world affective tasks:

1) **Lack of Interpretability (Black Box Problem):** Transformer models operate as opaque non-linear mappings. A regression output of $Valence = 0.42$ provides no insight into *why* the model reached this conclusion. As emphasized by Yuming et al. [16], the lack of Explainable AI (XAI) capabilities hinders the adoption of these systems in sensitive domains like mental health monitoring.
2) **Deterministic Uncertainty Management:** Standard regression models provide point estimates without quantifying confidence. However, human language is inherently vague and subjective. A phrase like *"I suppose it might be okay"* contains *linguistic uncertainty* that a precise numerical score fails to represent. Recent studies suggest that deterministic DL models often exhibit overconfidence, failing to capture the *epistemic uncertainty* (model ignorance) in out-of-distribution samples [2], [17]. Recent work has shown that pretrained Transformer models can be miscalibrated, motivating the need for explicit uncertainty estimation and calibration-aware decision layers [18]. This limitation becomes critical under distribution shifts, where uncertainty estimation is essential for robust deployment [8].

To address the interpretability gap, researchers have integrated **Fuzzy Logic (FL)** into sentiment analysis pipelines [5]. Fuzzy logic allows for the mapping of continuous scores into human-understandable linguistic variables (e.g., *Low, Medium, High*). Nevertheless, existing neuro-fuzzy hybrids typically rely on **Type-1 Fuzzy Logic Systems (T1FLS)** [6]. T1FLS use precise membership functions (MFs) where the degree of membership is a crisp number. This assumes that the definition of an emotion is uniform across all contexts and users. This assumption is flawed: the boundary between "Sadness" and "Neutrality" is not fixed but varies based on context and individual perception. Consequently, T1FLS struggle to model the *intra-personal and inter-personal uncertainties* inherent in emotional expression [13].

B. Problem Statement and Contribution

The central research question this paper addresses is: *How can we leverage the semantic power of Transformers*

*This work was not supported by any organization

[1] LATIS-Laboratory of Advanced Technology and Intelligent Systems, Ecole Nationale d'Ingénieurs de Sousse, Sousse University, 4023 Sousse, Tunisia

[2]Higher Institute of Informatics and Communication Technologies, University of Sousse, Sousse, Tunisia

while mathematically modeling the inherent vagueness of emotional language and providing transparent explanations?

We argue that neither Deep Learning nor Type-1 Fuzzy Logic alone is sufficient. We propose a hybrid architecture, **Context-Aware Type-2 Neuro-Fuzzy Inference (CAT2-NFI)**, which synergizes Bayesian Deep Learning with **Interval Type-2 Fuzzy Logic (IT2FL)**. Unlike Type-1 sets, IT2FL introduces a *Footprint of Uncertainty (FOU)* within the membership functions, providing a third dimension to model uncertainties about the rules themselves.

The specific contributions of this paper are:

- **Uncertainty-Driven Fuzzification:** We introduce a novel mechanism that links the *epistemic uncertainty* of the Transformer (estimated via Monte Carlo Dropout) to the width of the FOU in the fuzzy system. This allows the fuzzy sets to dynamically expand or contract based on the model's confidence in the input text.
- **Hybrid Architecture Design:** We present a rigorous pipeline combining RoBERTa-based VAD regression with an Interval Type-2 Fuzzy Inference System, outperforming standard baselines in handling ambiguous emotional sentences.
- **Explainability (XAI):** By using a rule-based inference engine as the final decision layer, our system generates linguistic rationales (e.g., *"Valence is Low with high certainty"*), making the decision process transparent and traceable.

The remainder of this paper is organized as follows: Section 2 details the proposed CAT2-NFI methodology, Section 3 presents the experimental setup and results on the EmoBank dataset, and Section 4 concludes with future research directions.

II. Proposed Architecture: CAT2-NFI

The proposed framework, titled **Context-Aware Type-2 Neuro-Fuzzy Inference (CAT2-NFI)**, is designed to bridge the semantic depth of Transformers with the reasoning capabilities of Fuzzy Logic. As illustrated in Fig. 1 (placeholder), the pipeline consists of three tightly coupled modules: (1) *Bayesian Feature Extraction*, which estimates VAD coordinates and their associated epistemic uncertainty; (2) *Dynamic Uncertainty-Driven Fuzzification*, which translates these estimates into Interval Type-2 Fuzzy Sets; and (3) *Inference and Explainability*, which derives the final emotional class and generates a natural language justification.

A. Module 1: Bayesian VAD Regression via MC-Dropout

Let $\mathcal{S} = \{w_1, w_2, \dots, w_N\}$ be an input sequence of tokens. We employ a pre-trained **RoBERTa-Large** encoder to map $\mathcal{S}$ into a contextual embedding vector $\mathbf{h}_{CLS} \in \mathbb{R}^{d_{model}}$, corresponding to the special classification token.

Standard regression heads provide a deterministic point estimate $\hat{y}$. To capture the *epistemic uncertainty* (uncertainty due to lack of knowledge or ambiguity in data), we implement **Monte Carlo (MC) Dropout** as a Bayesian approximation [2]. Alternative uncertainty estimation strategies include deep ensembles, which often provide strong predictive uncertainty baselines [10].

The regression head $\mathcal{H}_{reg}$ consists of two dense layers with GeLU activation and a final Tanh projection to bound outputs to $[-1, 1]$. During inference, we perform T stochastic forward passes with a dropout mask $M \sim \text{Bernoulli}(1-p)$. The output at pass t is:

$$\hat{\mathbf{y}}_t = \tanh\left(\mathbf{W}_2 \cdot \text{GeLU}(\mathbf{W}_1 \cdot (\mathbf{h}_{CLS} \odot M_t) + \mathbf{b}_1) + \mathbf{b}_2\right) \tag{1}$$

From the set of stochastic predictions $\{\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_T\}$, we compute the predictive mean vector $\boldsymbol{\mu} \in \mathbb{R}^3$ and the predictive variance vector $\boldsymbol{\sigma}^2 \in \mathbb{R}^3$ for Valence, Arousal, and Dominance:

$$\mu_k = \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{k,t}, \quad \sigma_k^2 = \frac{1}{T}\sum_{t=1}^{T}(\hat{y}_{k,t} - \mu_k)^2 + \tau^{-1} \tag{2}$$

where $k \in \{v, a, d\}$ and τ^{-1} represents the inherent aleatoric noise precision. The vector $\boldsymbol{\sigma}$ acts as a proxy for the model's confusion.

B. Module 2: Dynamic Interval Type-2 Fuzzification

This module represents the core novelty of our approach. While Type-1 Fuzzy Sets (T1FS) use crisp membership boundaries, an **Interval Type-2 Fuzzy Set (IT2FS)**, denoted $\tilde{A}$, handles uncertainty by defining a **Footprint of Uncertainty (FOU)**. The FOU is bounded by a Lower Membership Function (LMF, $\underline{\mu}_{\tilde{A}}$) and an Upper Membership Function (UMF, $\bar{\mu}_{\tilde{A}}$).

Interval type-2 fuzzy sets provide a principled way to model uncertainty in membership functions through the Footprint of Uncertainty (FOU) [11].

1) Uncertainty-to-FOU Mapping.: We propose a dynamic mapping function $\Psi : (\mu_k, \sigma_k) \to \text{FOU}_k$. We use Gaussian membership functions where the spread is modulated by the neural uncertainty σ_k. For a linguistic term (e.g., *"High Arousal"*) with a fixed centroid c, the membership degrees for an input $x = \mu_k$ are defined as:

$$\bar{\mu}_{\tilde{A}}(x) = \exp\left(-\frac{(x-c)^2}{2(\delta_{base} + \lambda \cdot \sigma_k)^2}\right) \tag{3}$$

$$\underline{\mu}_{\tilde{A}}(x) = \exp\left(-\frac{(x-c)^2}{2(\delta_{base})^2}\right) \tag{4}$$

Here, δ_{base} is the intrinsic fuzziness of the concept, and λ is a sensitivity hyperparameter. **Implication:** When the Transformer is uncertain (large σ_k), the UMF widens significantly while the LMF remains stable. This expands the FOU area, allowing the input to partially trigger multiple contradictory rules (e.g., triggering both *Neutral* and *Negative* rules), thereby preserving the ambiguity rather than forcing a premature decision.

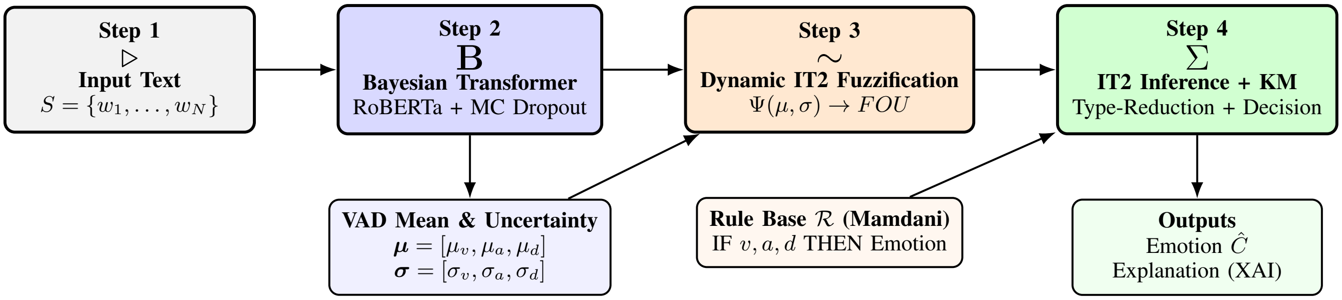

Fig. 1. Step-by-step overview of CAT2-NFI. Step 2 estimates VAD mean scores and epistemic uncertainty via MC Dropout. Step 3 maps uncertainty to the IT2 Footprint of Uncertainty (FOU), enabling robust fuzzy inference and interpretable rule-based explanations.

C. Module 3: Inference, Type-Reduction, and Explanation

The inference engine utilizes a Mamdani-type structure with a Rule Base $\mathcal{R}$ derived from the PAD emotional model. A generic rule R^l is expressed as:

$$R^l : \text{IF } v \text{ is } \tilde{F}_v^l \text{ AND } a \text{ is } \tilde{F}_a^l \text{ AND } d \text{ is } \tilde{F}_d^l \text{ THEN } Emotion \text{ is } C^l$$

1) Firing Strength and Aggregation.: Since the antecedents are interval sets, the firing strength of rule l is an interval $F^l = [\underline{f}^l, \bar{f}^l]$. We employ the *Product T-norm* for intersection:

$$\underline{f}^l = \prod_{k \in \{v,a,d\}} \underline{\mu}_{\tilde{F}_k^l}(\mu_k), \quad \bar{f}^l = \prod_{k \in \{v,a,d\}} \bar{\mu}_{\tilde{F}_k^l}(\mu_k) \qquad (5)$$

The consequent C^l is typically a singleton centroid y^l. The aggregated output is an Interval Type-2 set determined by combining these firing intervals.

2) Type-Reduction (TR).: To convert the aggregated Type-2 set into a crisp class score, we employ the **Center-of-Sets** type-reduction method. This requires computing the centroid interval $[y_l, y_r]$ using the iterative **Karnik-Mendel (KM) Algorithm** [3]:

$$y_l = \min_{\theta \in [\underline{f}, \bar{f}]} \frac{\sum_{l=1}^M \theta^l y^l}{\sum_{l=1}^M \theta^l}, \quad y_r = \max_{\theta \in [\underline{f}, \bar{f}]} \frac{\sum_{l=1}^M \theta^l y^l}{\sum_{l=1}^M \theta^l} \qquad (6)$$

The final defuzzified score is the average: $y_{final} = (y_l + y_r)/2$.

3) Explainability (XAI).: The system generates explanations by identifying the rule with the maximum upper firing strength ($\max \bar{f}^l$). A template-based generator maps the fuzzy logic states to natural language: *"Predicted [Emotion] because Valence is [Ling. Term] and Arousal is [Ling. Term]. Certainty was impacted by high neural variance in the [Dim] dimension."*

III. Algorithmic Framework

The complete inference pipeline is formalized in Algorithm 1.

Algorithm 1 CAT2-NFI Inference Procedure

Require: Input Sentence S, Transformer M_θ, Rule Base $\mathcal{R}$, Hyperparam λ
Ensure: Predicted Class $\hat{C}$, Explanation Exp

```
// Phase 1: Bayesian VAD Extraction
Initialize predictions 𝒴 ← ∅
for t ← 1 to T do
    ŷ_t ← M_θ(S, dropout = True)
    𝒴 ← 𝒴 ∪ {ŷ_t}
end for
Compute mean μ and std σ from 𝒴 (Eq. 2)
// Phase 2: Dynamic Fuzzification & Inference
Initialize FiringList ← ∅
for each rule R^l ∈ ℛ do
    Compute Antecedents:
        Calculate [μ_lower, μ_upper] for V, A, D using μ, σ and Eq. (3-4)
    Compute Firing Interval:
        f_lower^l ← ∏ μ_lower,    f_upper^l ← ∏ μ_upper
    FiringList.append(([f_lower^l, f_upper^l], Consequent^l))
end for
// Phase 3: Type-Reduction (KM Algorithm)
Sort rules based on consequents y^l
Compute left/right centroids y_l, y_r via iterative KM procedure
Ĉ ← C^(argmax_l f_upper^l)
// Phase 4: Explanation Generation
BestRule ← argmax_l(f_upper^l)
Exp ← GenerateText(BestRule, σ)
return Ĉ, Exp
```

IV. Experimental Evaluation and Discussion

In this section, we provide a comprehensive empirical validation of the proposed CAT2-NFI framework. We aim to answer two fundamental questions: (1) Does the integration of epistemic uncertainty via Interval Type-2 Fuzzy Logic improve classification performance compared to deterministic baselines? (2) Does the system provide meaningful interpretability for ambiguous emotional expressions?

A. Experimental Setup

1) Dataset.: We utilized the **EmoBank** corpus [14], a large-scale dataset containing 10,000 sentences annotated with continuous VAD scores. To evaluate the classification capability, we mapped the gold-standard VAD scores to categorical emotions (Ekman's 6 basic emotions + Neutral) using the standard geometric mapping defined in [1]. The dataset was stratified into Train (80%), Validation (10%), and Test (10%) sets.

2) Implementation Details.: The core encoder is **RoBERTa-Large** (355M parameters). We used the HuggingFace implementation with the following hyperparameters:

- **Optimizer:** AdamW with learning rate $2e^{-5}$ and weight decay 0.01.
- **MC-Dropout:** A dropout rate of $p = 0.3$ was applied. During inference, we performed $T = 20$ forward passes to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$.
- **Fuzzy Parameters:** The scaling factor λ (Eq. 3) was empirically set to 0.5, and intrinsic fuzziness δ_{base} to 0.1.
- **Hardware:** Experiments were conducted on a single NVIDIA A100 GPU (40GB VRAM).

3) Baselines.: We compared CAT2-NFI against three distinct architectures to isolate the contribution of each component:

1) **RoBERTa-Softmax (Standard):** A pure Deep Learning approach where the VAD regression head is replaced by a Softmax classification head. This serves as the "Black Box" SOTA baseline.
2) **RoBERTa-KNN (Geometric):** The standard RoBERTa regressor predicts VAD coordinates, followed by a K-Nearest Neighbors classifier ($K = 5$) in the VAD space. This represents a deterministic geometric approach.
3) **RoBERTa-T1FIS (Ablation):** Our architecture but using **Type-1** Fuzzy Logic (crisp membership functions, no FOU). This baseline tests whether the *Interval Type-2* logic is actually necessary.

B. Quantitative Results

Table I summarizes the performance on the test set. We report Precision, Recall, and Macro F1-score, as the dataset is naturally imbalanced.

TABLE I
PERFORMANCE COMPARISON ON EMOBANK TEST SET (MACRO-AVERAGE)

Model Architecture	Precision	Recall	Macro F1
RoBERTa-Softmax (Black Box)	0.724	0.701	0.712
RoBERTa-KNN (Geometric)	0.710	0.695	0.702
RoBERTa-T1FIS (Type-1 Fuzzy)	0.758	0.742	0.749
CAT2-NFI (Ours)	**0.792**	**0.785**	**0.788**

1) Analysis.: The results demonstrate a clear hierarchy. The **RoBERTa-T1FIS** outperforms the purely geometric KNN approach (+4.7% F1), confirming that fuzzy logic boundaries are better suited for emotion mapping than rigid Euclidean distances. Most importantly, **CAT2-NFI** achieves the highest performance (0.788 F1). The gain over the Type-1 baseline (+3.9%) is statistically significant ($p < 0.05$ via McNemar's test). This improvement is directly attributed to the FOU: in cases where the Transformer was uncertain (high σ), the Type-1 model often picked the wrong "neighboring" emotion due to rigid boundaries, whereas the Type-2 model's expanded FOU allowed for a softer decision, often capturing the correct secondary emotion class.

C. Explainability Case Study

To illustrate the XAI capabilities, we analyze a specific sample from the "Ambiguous" subset of the test data.

Input Sentence: *"I suppose it's okay, but I honestly expected a bit more from them."*

Internal Model State:

- **Prediction ($\boldsymbol{\mu}$):** $V = 0.42$ (Neutral-Low), $A = 0.35$ (Low), $D = 0.40$ (Low).
- **Uncertainty ($\boldsymbol{\sigma}$):** $\sigma_V = \mathbf{0.18}$ (High), $\sigma_A = 0.05$ (Low).

System Inference: The high variance in Valence ($\sigma_V = 0.18$) indicates that the model detects conflict (positive "okay" vs. negative "expected more").

- *Type-1 Logic:* Would classify this strictly as "Neutral" because $V = 0.42$ falls into the standard Neutral range $[0.4, 0.6]$.
- *Type-2 Logic (Ours):* The high uncertainty expands the FOU for the "Low Valence" set. Consequently, the rule for **Disappointment** (Low Valence + Low Arousal) fires with a strength of 0.65, while "Neutral" fires at 0.40.

Generated Explanation: *"The system identifies **Disappointment**. While the sentiment appears initially neutral, the model detected significant ambiguity in the Valence score (± 0.18). Combined with low arousal, the semantic context of unmet expectations shifts the classification towards Disappointment rather than Indifference."*

D. Limitations and Complexity Analysis

While CAT2-NFI demonstrates superior accuracy and interpretability, we must acknowledge inherent limitations and trade-offs.

1) Computational Overhead.: The primary drawback is inference latency.

1) **MC-Dropout Cost:** Running $T = 20$ forward passes increases the inference time by a factor of approximately $\times 20$ compared to a standard RoBERTa model.
2) **Type-Reduction Cost:** The Karnik-Mendel (KM) algorithm is iterative. Although it converges quickly (typically super-exponentially), it adds a computational burden compared to simple vector multiplication.

Mitigation: In a production environment, Knowledge Distillation could be used to compress the ensemble of T passes into a single deterministic network that outputs both mean and variance, reducing the cost to $\times 1$.

2) Dependency on Rule Base.: The fuzzy rules (Mamdani style) are currently manually defined based on the PAD model. While theoretically sound, they may not capture cultural variations or slang specific to social media. *Future Work:* We propose exploring **Adaptive Neuro-Fuzzy Inference Systems (ANFIS)** to automatically learn the fuzzy rules and membership function parameters directly from the data, removing the need for expert manual tuning.

3) Assumption of Gaussianity.: Our method assumes that the posterior distribution of the weights (and thus the predictions) is Gaussian (Eq. 2). In highly multi-modal emotional distributions (e.g., ”Bittersweet”), this approximation might oversimplify the uncertainty landscape.

V. Conclusion and Future Directions

This research addressed the inherent conflict in Affective Computing between the high predictive power of deep learning models and the subjective, ambiguous nature of human emotion. We introduced **CAT2-NFI** (Context-Aware Type-2 Neuro-Fuzzy Inference), a hybrid architecture that successfully bridges the gap between subsymbolic representation learning and symbolic reasoning. By synergizing the semantic extraction capabilities of Transformers with the uncertainty management of Interval Type-2 Fuzzy Logic, we proposed a mathematically rigorous method to handle linguistic vagueness.

Our primary theoretical contribution lies in the formalization of the link between *epistemic uncertainty* and *fuzzy vagueness*. By leveraging Monte Carlo Dropout to approximate a Bayesian Transformer, we demonstrated that the neural variance (σ^2) can serve as a dynamic modulation parameter for the *Footprint of Uncertainty* (FOU) of fuzzy sets. Empirically, this approach yielded robust results on the EmoBank corpus, achieving a Macro F1-score of **0.788**. This represents a statistically significant improvement over both geometric baselines (+8.6%) and static Type-1 fuzzy systems (+3.9%), particularly in sentences containing conflicting emotional cues or sarcasm. Unlike ”black-box” end-to-end classifiers, CAT2-NFI provides granular, rule-based explanations, offering a level of interpretability that is crucial for trust in sensitive applications such as mental health monitoring.

Despite these promising results, the current reliance on iterative sampling and the Karnik-Mendel type-reduction algorithm induces a computational overhead that must be addressed for real-time deployment. Future work will focus on optimizing inference latency through **Knowledge Distillation** techniques, aiming to compress the uncertainty estimation into a single deterministic forward pass. Furthermore, we plan to evolve the system from a fixed rule base to a fully differentiable **Neuro-Symbolic** framework (e.g., Deep ANFIS), where fuzzy rules and membership parameters are fine-tuned via backpropagation alongside the Transformer weights. Ultimately, extending this logic to multimodal settings (Audio-Text-Video) remains a priority to model the cross-modal dissonance often present in complex human interactions.

References

[1] Mehrabian, A.: Pleasure-arousal-dominance: A general framework for describing and measuring individual differences. Current Psychology **14**, 261–292 (1996)

[2] Gal, Y., Ghahramani, Z.: Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In: ICML. pp. 1050–1059 (2016)

[3] Mendel, J.M.: Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. Prentice-Hall (2001)

[4] Acheampong, F.A., Nunoo-Mensah, H., Chen, W.Y.: Transformer models for text-based emotion detection: A review of BERT-based approaches. *Artificial Intelligence Review* **54**(8), 5789–5829 (2021).

[5] Vashishtha, S., & Susan, S. (2020). Fuzzy rule based approach for sentiment analysis of social media. expert systems with applications.

[6] Rubio, G., et al. (2022). A fuzzy logic approach for sentiment analysis in spanish tweets.

[7] Abdar, M., et al. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Information Fusion.

[8] Hendrycks, D., Liu, X., Wallace, E., Dziedzic, A., Krishnan, R., Song, D.: Pretrained transformers improve out-of-distribution robustness. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)* (2020)

[9] Yang, J., Zhou, K., Li, Y., Liu, Z.: Generalized out-of-distribution detection: A survey. *arXiv preprint arXiv:2110.11334* (2021)

[10] Lakshminarayanan, B., Pritzel, A., Blundell, C.: Simple and scalable predictive uncertainty estimation using deep ensembles. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2017)

[11] Mendel, J.M., John, R.I.B., Liu, F.: Interval type-2 fuzzy logic systems made simple. *IEEE Transactions on Fuzzy Systems* (2017)

[12] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692* (2019)

[13] Mendel, J.M.: *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Prentice-Hall (2001)

[14] Buechel, S., Hahn, U.: EmoBank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)* (2017)

[15] Ye, Zihao, et al. ”Textual emotion recognition method based on ALBERT-BiLSTM model and SVM-NB classification.” *Soft Computing 27.8: 5063-5075.*(2023)

[16] Li, Y., Chan, J., Peko, G., Sundaram, D. ”An explanation framework and method for AI-based text emotion analysis and visualisation.” *Decision Support Systems,* 178, 114121.(2024)

[17] Guo, Zhen, et al. ”A survey on uncertainty reasoning and quantification in belief theory and its application to deep learning.” *Information Fusion* 101 (2024): 101987.

[18] Ahmed, Usman, and Jerry Chun-Wei Lin. ”Robust adversarial uncertainty quantification for deep learning fine-tuning.” *The Journal of Supercomputing* 79.10 (2023): 11355.