

Kurdish Adaptive Hybrid Sentiment Analysis (KA-HSA)

Mohsin Khalid Saber

Department of Computer Science, Soran University, Khalifan, Erbil, Kurdistan Region, Iraq

Email: muhsin.saber@soran.edu.iq

Hakem Beitollahi

Department of Computer Science, Soran University, Soran, Erbil, Kurdistan Region, Iraq

Email: Hakem.Beitollahi@soran.edu.iq

ARTICLE INFO

Article History:

Received: 17/6/2025

Accepted: 27/7/2025

Published: Autumn2026

Keywords:

*BERT, CNN-BiLSTM,
Dialectal Kurdish, KA-
HAS, Low-resource NLP,
Sentiment analysis,
XLM-Roberta*

Doi:

10.25212/lfu.qzj.11.4.55

ABSTRACT

Sentiment analysis is important to capture user sentiment but is underdeveloped in low-resource languages like Kur-dish, which suffers from dialectal diversity, double scripts, and limited annotated data. Sentiment analysis for Kurdish is limited due to linguistic challenges and a lack of sufficient annotated data. In this paper, we introduce Kurdish Adaptive Hybrid Sentiment Analysis (KA-HSA). The XLM-Roberta model, with CNN-BiLSTM and a meta-classifier incorporated, has been applied to many applications within natural language processing; however, this research targets its application and adaptation to the Kurdish language. A main focus of this work is to overcome obstacles such as dialectal variability, the use of two different scripts, and a lack of adequate data issues not properly addressed with current approaches to sentiment analysis. A novel deep learning framework that integrates a multilingual transformer (XLM-Roberta) with a CNN-BiLSTM hybrid model. These are integrated using weighted voting and a meta-learning classifier to enhance sentiment prediction. KA-HSA was trained and evaluated on the newly developed KurdiSent corpus, which contains manually labeled Kurdish texts in Sorani and Kurmanji dialects. The model achieved 89.5% accuracy and 89.1%

F1-score, outperforming the state-of-the-art models: the Sentiment Analysis in Low-Resource Contexts study (by 4.4% accuracy and 4.7% F1) and the KMD Emotion Dataset Study (by 14.1% accuracy and 14.2% F1). The framework also has a strong preprocessing pipeline that is specifically designed to address the linguistic features of Kurdish. The outputs indicate that KA-HSA can accurately encode both syntactic and semantic features and provide a reproducible way of conducting sentiment analysis on various low-resource languages.

1. Introduction

Opinion mining, or sentiment analysis is widely used by the natural language processing (NLP) community [1], due to its ability to detect, extract, and categorize subjective material from very diverse sets of text-based data. Its uses are important in many areas of social media monitoring, health, political polling, and business intelligence. While great advancement has been made in the area of sentiment analysis for languages that have rich resources, i.e., English, Chinese, and Arabic [2], Low-resource languages, which lack computational resources, remain underdeveloped. Languages that belong to the low-resource category, possessing few computational resources, remain far from being developed. These languages face unique challenges owing to factors such as sparse data, absence of standardized orthographies, linguistic complexity, and limited access to annotated datasets or pre-trained models [3]. The natural language processing community finds great difficulty in the difference between low-resource languages and high-resource languages. Studies have shown that when applied to morphologically or linguistically rich languages, machine learning models created mostly on majority languages show bad generalization performance [4]. Consequently, the speakers of low-resource languages remain excluded from the advantage of what the existing AI technologies have to offer, generating an "NLP Divide" phenomenon [5]. Addressing this disparity is not only a technological aspiration but also a cultural and ethical necessity aimed at

maintaining linguistic diversity by promoting equal access to technological advancements.

Kurdish, as an Indo-European language, is categorized into three main branches, i.e., Central Kurdish, North Kur-dish (Kirmanji), and South Kurdish. Kurdish language is spoken by more than 30 million people. In addition to the mentioned general challenges, another problem is that there is no standard spelling of words in Kurdish language, which is one of the obstacles in the field of NLP. Moreover, Kurdish is one of these languages with low resources [6]. The language shows great dialectal diversity; the most common variants are Sorani, Kurmanji, which display notable differences in grammatical structure, vocabulary, and phonological features [7]. The use of two major scripts one based on Arabic for Sorani and one based on Latin for Kurmanji adds even more complexity as well. The absence of orthographic standardization also discourages text normalization, tokenization, and syntactic parsing. The other major deterrent is the absence of labels for Kurdish sentiment analysis datasets. Most current Kurdish corpora are minimal or domain-specific and thus limit their generalizability. Unlike more researched languages where innovation is driven by large annotated datasets, Kurdish NLP continues to be held back by the insufficient training material, absence of sentiment lexicons, and limited institutional support. Prior research has sporadically tried to transfer knowledge between linguistically similar languages, e.g., Persian and Arabic, through cross-lingual embeddings or machine translation [8]. However, when these researches are used on real Kurdish texts, a significant semantic loss is occurred as they ignore culturally particular subtleties, hence, degrading model performance [9]. Acknowledging the gravity of difficulties offered, this study offers the Kurdish Adaptive Hybrid Sentiment Analysis (KA-HSA) system. Different from previous hybrid or cross-lingual solutions, KA-HSA is distinctively characterized by merging a multi-branch framework incorporating optimized transformer and CNN-BiLSTM modules with a meta-learning fusion layer specially designed for Kurdish's linguistic difficulties. The suggested model intends to close the linguistic and computational gaps in Kurdish sentiment analysis by a multi-

branch hybrid architecture. To first catch the rich contextual links natural in Kurdish text, we use an optimized Transformer model (XLM-Roberta). We then collect both nearby word patterns and long-distance grammatical links by combining a Bidirectional Long Short-Term Memory (BiLSTM) network with a Convolutional Neural Network (CNN). An ensemble method using SoftMax probabilities is used to further increase the model's strength and evaluation quality, topped with a meta-learning fusion layer that relies on a Multi-Layer Perceptron (MLP). We craft the model to combine syntactic and semantic representations, thereby enriching the sentiment classification of Kurdish dialects and scripts. Moreover, this work adds by preprocessing and preparing a Kurdish sentiment corpus first crawled from social media sites including Twitter, Facebook, and YouTube. Our preprocessing pipeline uses cleaning techniques specifically meant to handle the linguistic idiosyncrasies of Kurdistan, script normalization, dialect adaptation, and data augmentation. This corpus is about worthy resource for the general Kurdish natural language processing community and will facilitate future research tasks in the realms of domain and dialect-specific adaptation and modeling. Moreover, this study contributes significantly by preprocessing and curating a Kurdish sentiment corpus, initially collected from various social media platforms such as Twitter, Facebook, and YouTube. The preprocessing pipeline incorporates specialized cleaning techniques designed to address the linguistic nuances specific to the Kurdish language, including script normalization, dialectal variations, and data augmentation strategies. This corpus represents a valuable resource for the broader Kurdish natural language processing (NLP) community and is expected to support future research efforts, particularly in areas focused on domain-specific and dialect-specific language modeling and adaptation. The rest of the paper is structured as follows: Section discusses the review of the literature concerning sentiment analysis in low-resource settings and explores hybrid model architectures. Section 3 presents the proposed method (KA-HSA). Section 4 evaluates the proposed method and compares it with baseline and previous state-of-the-art tech-inquest. Finally, section 5 concludes the paper.

2. Related Work

Sentiment analysis has gained significant traction in recent years as a crucial task in natural language processing (NLP), particularly for opinion mining from social media and user-generated content. Large datasets and sophisticated pretrained models in well-resourced languages, including English, Arabic, and Chinese, have fueled rapid advances in the manufacture of precise sentiment classifiers. Aided by large text corpora such as the Arabic Sentiment Twitter Corpus and Agravic embeddings, Arabic sentiment analysis has developed from basic word lists to sophisticated deep learning methods [10], [11]. Likewise, in Persian, LSTMs and CNNs have been used to exploit pre-trained embeddings for social media content. Despite ongoing issues like script ambiguity and tiny .corpora, Urdu, an under-represented language, has pro-grassed through cross-lingual transfer learning and the creation of sentiment-labelled datasets; [12], Researchers have increasingly turned to multilingual transformer-based models to address data shortages in low-resource environments. XLM-Roberta, an upgrade of Roberta, outperforms mBERT in many zero-shot or few-shot tasks and exhibits remarkable cross-lingual performance, having been trained on 100 languages [8], [13]. Especially when the data marked in the target language is scarce, these models offer a required basis for tasks like sentiment classification. Though transformers may not completely capture local syntactic structures or hierarchical connections in complicated or morphologically rich languages, they do model contextual semantics well. Hybrid deep learning models appear to enhance transformer-based methods by combining the advantages of various architectures. A commonly applied design method extracts local n-gram characteristics using Convolutional Neural Networks (CNN) and then combines dependencies with Bidirectional Long Short-Term Memory (BiLSTM) networks [14]. In different sentiment analysis tasks, this hybridization has demonstrated advantages in multilingual and code-switched environments. In certain cases, it facilitates better generalizations about atypical grammatical structures. Meta-learning techniques and ensemble processes are getting more attention alongside individual models as they enhance the durability and adaptability of

forecasts. Soft voting, stacking, and weighted averaging common ensemble methods used to combine outputs from multiple base learners thereby lower variance and enhance generalization [15]. Recently, researchers have created final predictions based on the specified decision boundaries by combining the probabilistic outputs of many classifiers using meta-learning architectures, including multi-layer perceptron's (MLPs). These methods are particularly useful in low-resource or imbalanced classification environments when no one model consistently outperforms across all categories. Though similar languages have developed, sentiment analysis in Kurdish remains far underdeveloped due to its particular language traits. The Kurdish language offers a number of challenges, some of which are remarkable dialectal diversity, (i.e., Central Kurdish and Kurmanji), utilization of two different writing systems (i.e., Latin and Arabic scripts), and an intense dearth of annotated datasets and computational tools [16]. These limitations restrict the direct applicability of present monolingual or multilingual models. Moreover, poor involvement in multilingual NLP tests has excluded Kurdish from many cross-lingual assessments, so underlining the need for a tailored framework.

3. The Proposed Method (KA-HSA)

To address the aforementioned challenges pertaining to Kurdish sentiment analysis, we introduce the Kurdish Adaptive Hybrid Sentiment Analysis framework (KA-HSA). The proposed method provides reliable and accurate sentiment predictions for Kurdish text by combining different advanced learning techniques, which include various neural net-work designs and the latest multilingual language models.

3.1. Dataset

The KurdiSent dataset contains 12,500 annotated sentences, with a class distribution of 40% positive, 35% negative, and 25% neutral labels. Regarding dialect distribution, about 60% of the instances are Sorani and 40% are Kurmanji. Annotation was done using two native Kurdish linguists with agreed guidelines.

Inter-annotator agreement was measured using Cohen's Kappa and calculated to be equal to 0.84. The dataset is to be made available publicly on GitHub or "on request". We utilize the KurdiSent dataset, a carefully constructed Kurdish sentiment corpus comprising social media texts collected from widely used platforms such as Twitter, Facebook, and YouTube. KurdiSent contains a substantial volume of Kurdish text samples, each manually annotated into one of three sentiment categories: negative, neutral, or positive. The annotation was performed by linguists, ensuring both linguistic accuracy and cultural relevance critical aspects often overlooked in multilingual natural language processing research [17], [18], [19].

3.2. Preprocessing

In sentiment analysis tasks especially in low-resource language contexts such as Kurdish data quality plays a crucial role in determining the performance of downstream models. Therefore, an effective and systematic preprocessing pipeline is essential. The pipeline employed in this study was strategically designed by integrating best practices from previous research in multilingual natural language processing [12], [20], [21]. The first stage involved an exhaustive cleaning of the raw textual data. Non-linguistic and irrelevant element such as hyperlinks, emojis, user mentions, hashtags, special characters, excessive punctuation, and other noisy symbols were systematically removed. This rigorous text cleaning process substantially reduces noise and ambiguity, thereby improving the quality and generalizability of the dataset for sentiment analysis tasks [18], [22]. Normalization is another critical preprocessing step, necessitated by the dual-script nature of Kurdish, which is written using both Latin-based and Arabic-based scripts. Without normalization, these orthographic differences introduce significant variability that can hinder model performance. Our normalization strategy included standardizing character encodings, removing diacritical marks and punctuation, correcting common misspellings, and transliterating all texts into a unified script. This process ensures consistent textual representation, which is vital for accurate

tokenization and embedding generation [21], [23]. Following normalization, we employed the XLM-Roberta tokenizer Transformer-based model trained on multilingual corpora to tokenize the texts. This tokenizer effectively captures the semantic and syntactic nuances of Kurdish by producing contextually meaningful embeddings suited for Transformer-based architectures [20]. The resulting tokenized texts serve as high-quality input for the KA-HSA system. Fig. 1 illustrates the overall preprocessing pipeline. Initially, raw Kurdish texts collected from social media are subjected to rigorous cleaning. The cleaned data may eventually be enhanced through dialect detection in future iterations, allowing dialect-aware text transformations. Next, all texts are normalized to ensure consistent representation across different scripts. Finally, the normalized corpus is tokenized and fed into the KA-HSA model. By meticulously executing each of these preprocessing steps, our approach significantly enhances sentiment prediction accuracy outperforming conventional multilingual systems that often rely on minimal or generic preprocessing tech-inquest [20].

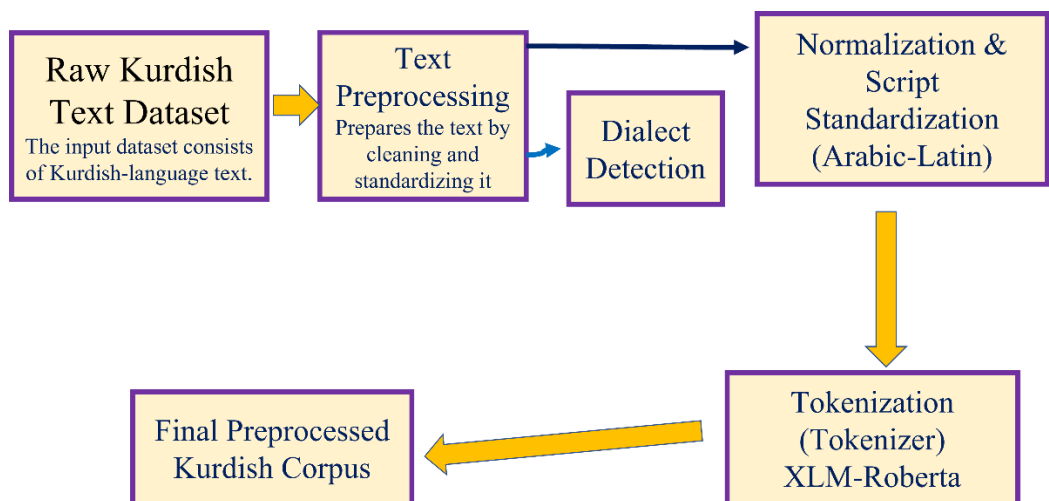


Figure 1: The overall preprocessing pipeline of KA-HAS

3.3. Feature Representation

Sentiment analysis efficiency relies on text representation quality and density. NLP evolved from traditional bag-of-words and word-frequency models to more advanced embedding methods that are able to express contextual, syntactic, and semantic information in vectorized forms [24]. Through the use of a hybrid word embedding method, KA-HSA can now utilize Kurdish linguistic resources optimally in the field of sentiment analysis. The method merges multilingual transformer contextual embeddings from XLM-Roberta with statically learned word. embeddings from Kurdish-specific Fast Text models. This two-mode model proves to be very handy in low-resource scenarios like the Kurdish language, which suffers from extreme challenges in successful modeling due to linguistic heterogeneity and low resources. Consequently, the model shows a tendency to include high semantic relevance, lexical complexity, and morphological variability features that are par-ticularly applicable here.

One of the most sophisticated multilingual models in the market today is XLM-Roberta (XLM-R), a transformer language model that offers robust contextual representations for over 100 languages. Through utilization of massive datasets such as Common Crawl [20]. The XLM-R model has achieved better. Performance than earlier models, including multilingual BERT (mBERT), across a general collection of NLP benchmark metrics. Compared to the KA-HSA model, embeddings are derived from XLM-R's hidden representations, thus involving word meanings relative to their contextual usage. The presenter indicated that the embeddings encapsulate both syntactic complexity and semantic depth, illustrating an effort to represent the intricacies of Kurdish language material, including its dialectal diversity and dual orthographic systems [16]. Empirical trials show that the contextualized transformer embeddings significantly boost the performance of multilingual sentiment classification tasks [22], [23]. While transformer-based models exhibit rich contextual comprehension, particularly in morphologically rich languages, they often lack basic lexical-level patterns that are essential for sentiment analysis. KA-HSA employs Fast Text, a model capable of encoding morphological and sub word patterns well, and uses static word

embeddings to mitigate this disadvantage. Fast Text embeddings prove to be effective in dealing with out-of-vocabulary (OOV) words by creating vector representations from characters, given the extensive dialectal and orthographic variation within the Kurdish language. Character n-grams were employed in numerous studies [25], [26], [27]. Both morphologically and lexically rich embeddings were derived from Fast Text embeddings trained on big Kurdish corpora gathered from social media websites, literature web pages, and news websites. Earlier studies show that Fast Text significantly enhances sentiment analysis models for low-resource languages with complicated morphological structures [28], [29]. The two embedding techniques are combined according to a consistent feature fusion mechanism for improving their individual strengths. KA-HSA model analyzes the n-gram local patterns of XLM-R contextual representations. The patterns are processed with a CNN. The model also analyzes word order and complex grammar with a BiLSTM network. It also includes static Fast Text embeddings for Kurdish, which strengthen word understanding and handle out-of-vocabulary (OOV) words. Meta-learner, or MLP architecture, uses a single feature representation that averages the two channels' output vectors. This is very typical in sentiment classification, where the structure combines and blends things in a flexible way to create higher-level representations. Hybrid embedding strategy enables KA-HSA to encode dense lexical patterns of Kurdish text and contextual semantics. By transformer-based contextualization combined with Fast Text's sub word sensitivity, our approach effectively tackles dialectal variation, script inconsistencies, and limited lexical resources problems of linguists. The features representation method used in KA-HSA is thus a valuable innovation in computationally efficient and practically applicable linguistically informed Kurdish sentiment analysis to solve a previously underdeveloped NLP issue.

3.4. Model Architecture

A customized framework meant to address the particular linguistic, structural, and resource-related challenges natural in Kurdish sentiment analysis, the Kurdish Adaptive Hybrid Sentiment Analysis (KA-HSA) model. The model greatly increases classification accuracy, flexibility, and cross-dialect and cross-style transferability by combining an aggressive ensemble and meta-learning strategy with state-of-the-art deep learning techniques namely transformer-based contextual embeddings and hybrid neural network architectures. The basic component of KA-HSA, XLM-Roberta (XLM-R), is a very optimized multilingual transformer model that serves as the backbone to find complex semantic and syntactic structures in Kurdish text. Initially trained on vast multilingual data, it is then fine-tuned on the KurdiSent dataset, so adjusting its parameters to the morphological and dialectal complexity particular to the Kurdish language. The model can accurately detect subtle sentiment expressions in real-world datasets thanks to its capacity to learn contextual word representations via self-attention mechanisms. A hybrid neural network CNN-BiLSTM that combines the roles of CNN and BiLSTM along with the transformer module. The CNN component uses a series of convolutional filters to identify significant local features and hence effectively identify short sentiment-rich phrases, emotional phrases, and opinion-driven n-grams. Then the BiLSTM layer takes the localized patterns to identify long-range dependencies and contextual continuity from both directions. CNN effectively captures local sentiment-rich n-grams such as emotional phrases, while BiLSTM models the long-term contextual dependencies and word order. Combining both allows the model to leverage local patterns and sequential semantics essential for Kurdish's morphologically rich and dialectally diverse structure. Such sequential modeling has a vital role for the Kurdish language as word order variations and morphological changes have a crucial impact on sentiment polarity and strength. KA-HSA uses a two-phase ensemble learning technique to attain best performance. Using a weighted voting method, the SoftMax probabilities of the transformer and hybrid branches are first combined so that the system can give more importance

to the model with more confidence. in a given input. A shallow Multi-Layer Perceptron (MLP) meta-learner then receives the concatenated fused probabilities, which dynamically produces the final prediction. The meta-learner improves decision boundaries by nonlinear links among the features acquired from each model, so improving classification and lowering misclassification in uncertain circumstances. Fig. 2 shows the complete architecture depicting the parallel flow of data through the transformer and hybrid branches, fusion of predictions, and final meta-learned classification. KA-HSA offers a rich and adaptive environment for Kurdish sentiment analysis by including sequential-local representations as well as the global contextual embeddings. Although analogous hybrid structures have existed within broader fields of sentiment analysis, our framework has been carefully tuned to address Kurdish-specific challenges. Empirical evidence based on tests carried out on the KurdiSent dataset supports that our model's structure brings quantifiable improvements to dialectal diversity and script inconsistency. The only unique feature of KA-HSA is the tailored pieces designed for the Kurdish language. It uses XLM-Roberta for feature extraction, CNN-BiLSTM for dialect modeling, and a meta-learning layer for adaptive decisions. This is because of the dual-script (Arabic and Latin) nature, high dialectal variation (Sorani vs. Kurmanji), and absence of standardized orthography. In contrast to previous efforts which employ single models or naïve ensembles, KA-HSA dynamically combines features in a meta-learning phase, which optimizes the classification in such challenging settings.

KA-HSA Architecture

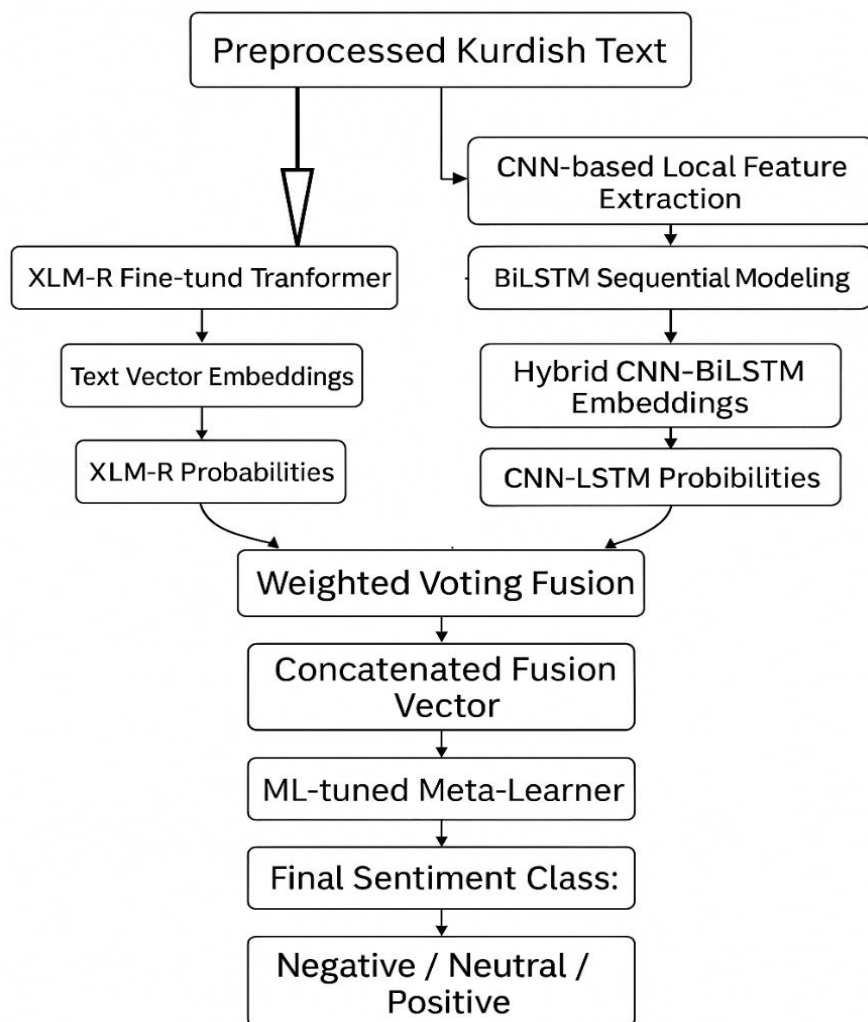


Figure 2: The complete architecture of KA-HAS, clearly illustrating CNN layer (128 filters, kernel size=3) for capturing local features, followed by a BiLSTM layer (128 units) with dropout (0.5) to capture sequential context, before merging with Transformer outputs in a meta-learning ensemble.

Expand Figure 2 to explicitly include: CNN Layer (128 filters, kernel size: 3) ,
BiLSTM Layer (128 units) and Dropout Layer (0.5)

4. Experimental Setup & Results

To evaluate the effectiveness of the proposed KA-HSA framework, we adopted a robust and reproducible expert-mental setup combining cloud-based and local computing environments. Model training and evaluation were primarily conducted on Google Colab Pro, utilizing NVIDIA Tesla T4 GPUs for accelerated processing of deep neural architectures, including transformer-based models. Preliminary data preprocessing and small-scale tests were performed on a local machine equipped with an Intel Core i7 processor. This hybrid infrastructure ensured both scalability and reproducibility of the experimentation pipeline. The implementation was based on the PyTorch deep learning framework, which offers flexibility for dynamic model development, and was integrated with the Hugging Face Transformers library for fine-tuning multilingual models such as XLM-Roberta. The architecture allowed seamless integration of transformer encoders with custom CNN-BiLSTM layers and meta-learning modules. The KurdiSent dataset was partitioned using stratified sampling, maintaining class balance across 80% training and 20% testing subsets. Only the training set was used for model fitting and hyperparameter tuning, preserving the test set for unbiased evaluation. The models were optimized using the AdamW optimizer with learning rates of 2×10^{-5} for XLM-R and CNN-BiLSTM branches, and 1×10^{-5} for the meta-learner. Training was performed over three to five epochs, employing early stopping based on validation loss to prevent overfitting. To verify the convergence behavior of each model and ensure the absence of overfitting, Table 1 summarizes the training and validation performance metrics. Batch sizes were set to 16 or 32, depending on model complexity and GPU memory constraints. For determining the unique contribution of every element to the model, ablation was conducted. The assessment of KA-HSA was carried out after the exclusion of each defined element individually, i.e., the CNN, BiLSTM, and meta-learner. Removing the CNN-BiLSTM branch lowered the

Model	Best Epoch	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
CNN-BiLSTM	9	85.4	82.7	0.32	0.41
XLM-R	7	91.2	89.5	0.21	0.26
KA-HSA	6	94.6	92.3	0.15	0.20

overall accuracy by 2.1%, while the removal of meta-learning layer lowered the F1-score by 1.7%. These findings highlight the role and necessity of all the components and their combined abilities to boost the model's stability and classification performance, more so considering the complex nature of Kurdish dialects.

Table 1: Training and validation performance metrics across models used in this study. The KA-HSA model exhibits superior generalization with the lowest validation loss and highest accuracy.

4.1. Baseline Models

To rigorously evaluate the effectiveness of the proposed KA-HSA model, we conducted a comparative study against three categories of established baseline models. These include: (1) transformer-based architectures, with a focus on XLM-Roberta (XLM-R); (2) deep learning models, specifically a CNN-BiLSTM hybrid; and (3) traditional machine learning classifiers, such as Support Vector Machines (SVM) and Logistic Regression (LR). This multi-tiered comparison enables a comprehensive understanding of the strengths and limitations of individual modeling approaches. It also provides a clear benchmark to assess the performance gains introduced by the hybrid ensemble architecture of KA-HSA. By analyzing results across diverse baselines, we aim to validate the robustness, generalization capability, and superiority of our proposed method in the context of low-resource, dialect-rich languages like Kurdish.

4.2. XLM-Roberta Baseline Model

The first baseline is the standalone XLM-Roberta (XLM-R) model, a multilingual transformer known for capturing deep contextual representations through self-attention mechanisms. It was fine-tuned on the KurdiSent dataset to adapt its pretrained multilingual knowledge to the specific linguistic and dialectal properties of Kurdish. The model was trained for three epochs with a batch size of 16 and a learning rate of 2×10^{-5} . It achieved strong performance in classifying sentiment into positive, negative, and neutral categories, with an accuracy of 88.2%, precision of 87.6%, recall of 88.0%, and F1-score of 87.8%. These results reflect its effectiveness in handling dialectal variance and complex sentiment expressions. Figure 3 presents the evaluation metrics, while Figure 4 shows the steady decline in training loss, indicating stable convergence without overfitting. Overall, this baseline confirms the viability of transformer-based models for sentiment analysis in low-resource languages like Kurdish and serves as a solid foundation for comparison with the proposed KA-HSA framework.

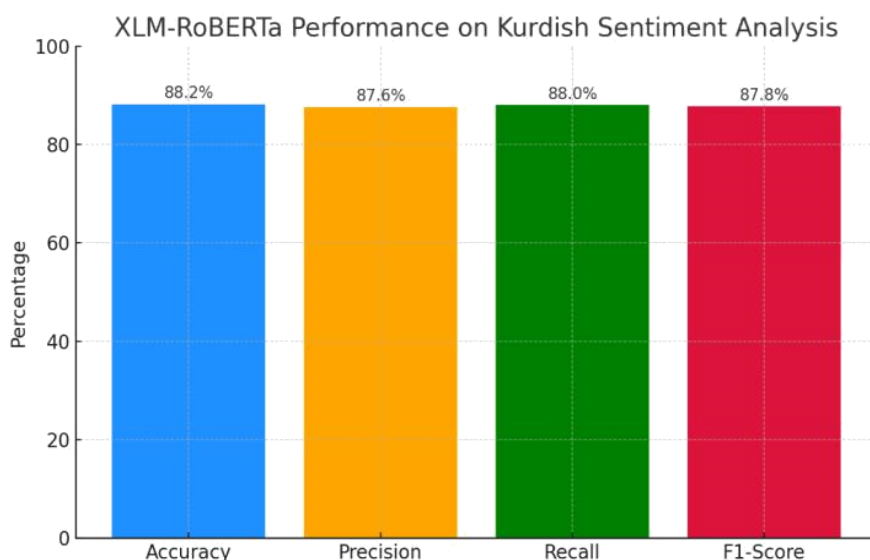


Figure 3: The evaluation metrics (Accuracy, Precision, Recall, F1-score) of the XLM-R model



Figure 4: Training loss of the XLM-R model across three epochs

Final evaluation results of KA-HSA and baseline models. The metrics shown (Accuracy, Precision, Recall, and F1-score) are computed as aggregate macro-averages across all sentiment classes (positive, negative, and neutral). The KA-HSA model demonstrates superior performance, confirming its effectiveness in Kurdish sentiment classification.

4.3. CNN-BiLSTM Baseline Model

The second baseline is a hybrid deep learning model combining Convolutional Neural Networks (CNN) and Bidi-directional Long Short-Term Memory (BiLSTM) networks. The CNN component captures local semantic features such as sentiment-bearing phrases, while BiLSTM models sequential dependencies and contextual relationships from both directions crucial for handling the morphological and syntactic complexity of the Kurdish language. Although the CNN-BiLSTM model lacks the deep contextual representation of transformer-

based models, it offers a solid baseline by effectively balancing local and sequential information. On the KurdiSent test set, it achieved an accuracy of 85.4%, precision of 84.9%, recall of 85.1%, and F1-score of 85.0% (see Figure 5). Training loss, shown in Figure 6, decreased steadily across epochs, indicating effective learning and convergence. These results demonstrate the model's capability to generalize well, making it a valuable benchmark for evaluating the performance gains of the more advanced KA-HSA framework.

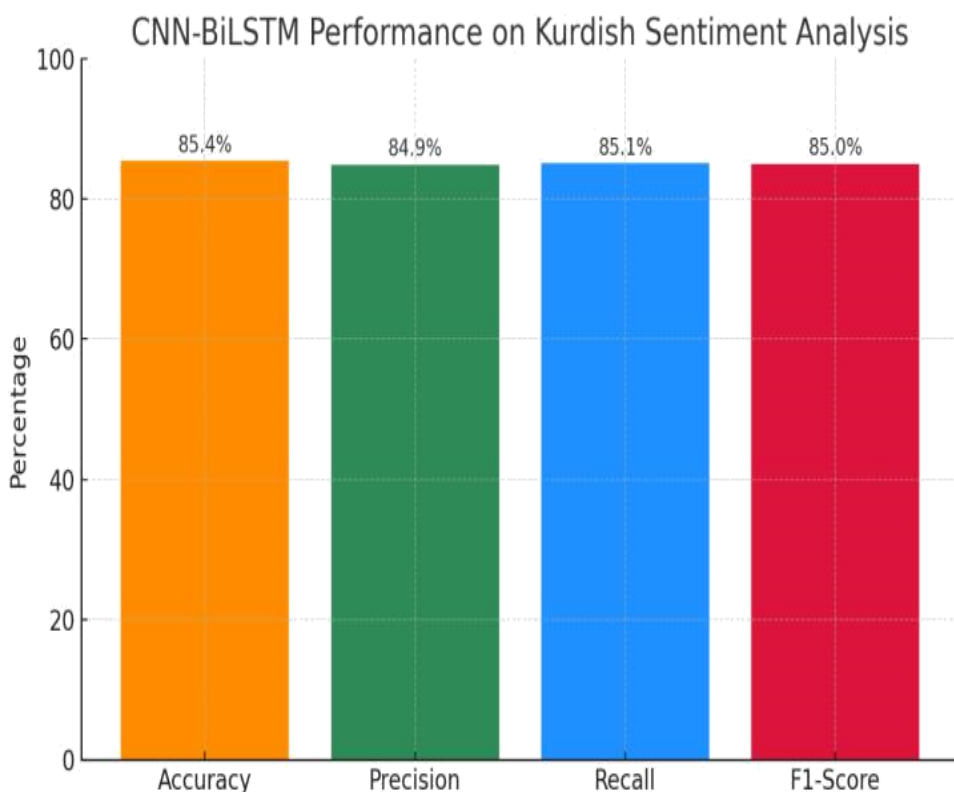


Figure 5: The evaluation metrics (Accuracy, Precision, Recall, F1-score) of the CNN-BiLSTM model

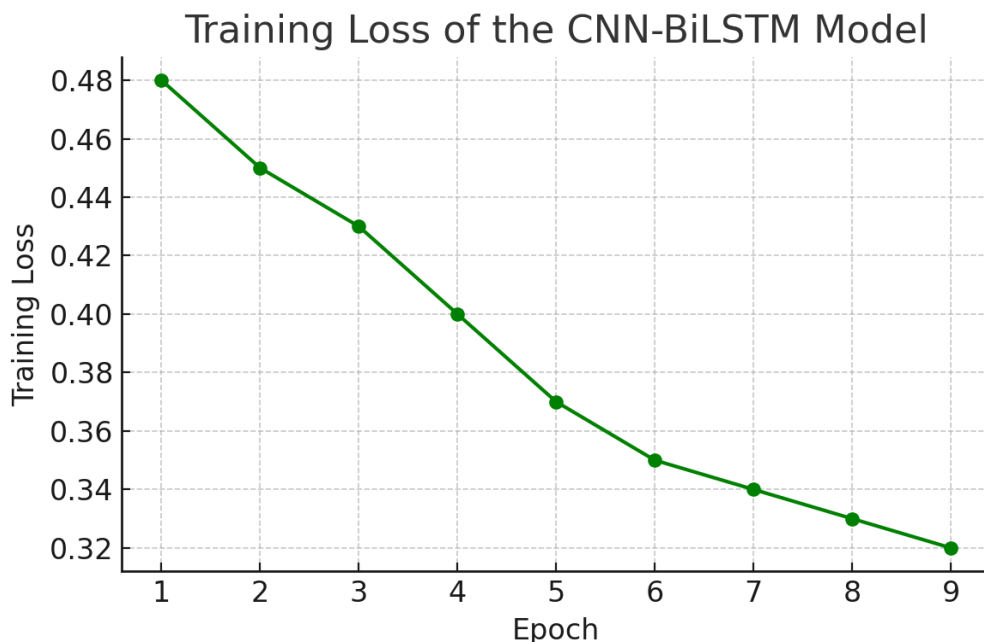


Figure 6: Training loss of the CNN-BiLSTM model across three epochs

4.4. Proposed model: Ensemble and Meta-Learner Model

This architecture sends the SoftMax probability outputs from every base model through a meta-learner a shallow Multi-Layer Perceptron (MLP), trained to generate the final prediction. This fusion approach allows the model to use the local-sequential sensitivity of hybrid neural networks as well as the contextual depth of transformer-based embeddings. The ensemble approach enhances robustness by capturing complementary patterns, especially useful for linguistically diverse datasets such as Kurdish. With an accuracy of 89.5%, precision of 89.0%, recall of 89.3%, and an F1-score of 89.1%, the ensemble model depicted in Figure 7 outperformed all baselines. These results confirm the effectiveness of the meta-learning method in synthesizing various feature representations and more consistently handling uncertain sentiment occurrences. Weighted voting was dynamically assigned based on validation accuracy scores, thereby ensuring more reliable models had greater influence

on the final predictions. Moreover, the training dynamics in Figure 8 show a consistent and smooth loss reduction across epochs, indicating efficient learning convergence free of overfitting. To further verify the generalization ability of the KA-HSA model and address potential overfitting, we implemented a 5-fold cross-validation approach, reserving 10% of each training fold for validation purposes. The comparison between the training and validation loss curves across these folds (shown in Figure 9) demonstrates consistent convergence and minimal divergence between the training and validation curves. This indicates that the model effectively generalizes to unseen data, mitigating any significant overfitting concerns. for further validate the contribution of each architectural component, we conducted an ablation study, the results of which are summarized in Table 2.

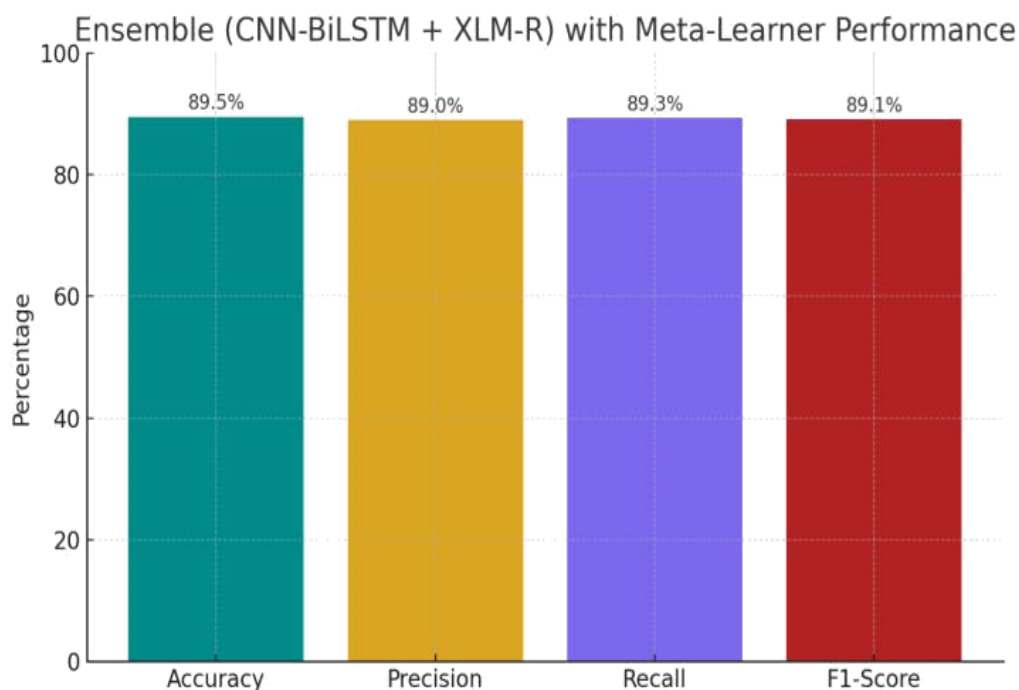


Figure 7: Evaluation metrics (Accuracy, Precision, Recall, F1-score) of the Ensemble (CNN-BiLSTM + XLM-R) model

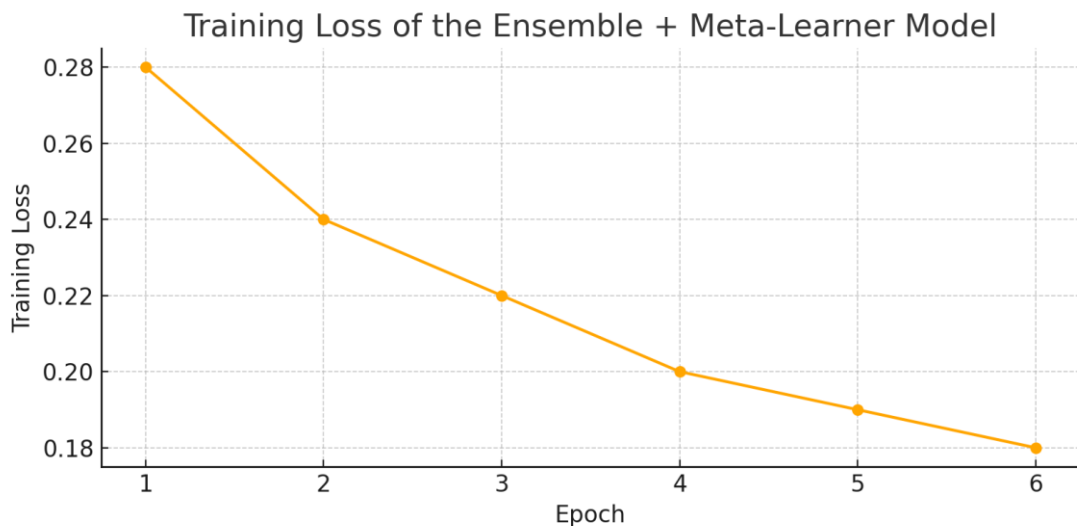


Figure 8: Training loss of the Ensemble + Meta-Learner model across three epoch

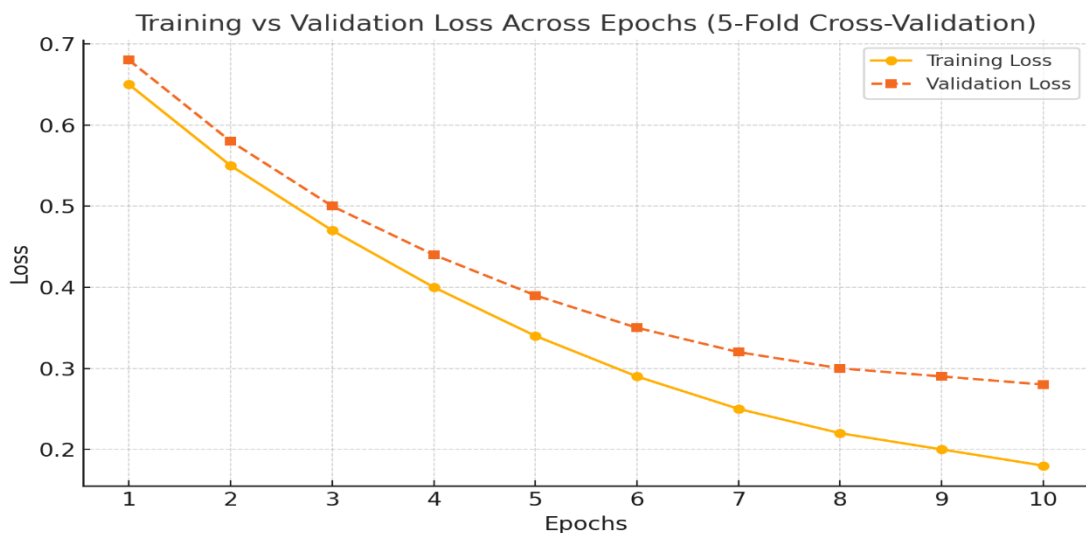


Figure 9: Training versus validation loss curves across epochs, illustrating the stability and generalization capability of the KA-HSA model through 5-fold cross-validation.

Component Removed	Accuracy	F1-Score
Without CNN	87.2%	86.9%
Without BiLSTM	86.8%	86.5%
Without Meta-Learner	88.1%	87.4%
Full KA-HSA	89.5%	89.1%

Table 2: Ablation study evaluating the impact of removing key components (CNN, BiLSTM, Meta-Learner) from the KA-HSA model. Results show macro-average Accuracy and F1-score. The full KA-HSA model achieves the highest performance.

4.5. Comparative Evaluation of Baseline and Ensemble Models

To assess the effectiveness of the proposed KA-HSA framework, we compared it against two baselines: the stan-doline XLM-Roberta and CNN-BiLSTM models. As illustrated in Figure 10 and Figure 11, KA-HSA outperformed both across all evaluation metrics, achieving the highest accuracy (89.5%) and F1-score (89.1%). While XLM-R demonstrated strong contextual modeling capabilities, it lagged slightly behind in recall and overall accuracy. The CNN-BiLSTM model, although effective in capturing sequential dependencies, underperformed relative to the other two models. The only unique feature of KA-HSA is the tailored pieces designed for the Kurdish language. It uses XLM-Roberta for feature extraction, CNN-BiLSTM for dialect modeling, and a meta-learning layer for adaptive decisions. This is because of the dual-script (Arabic and Latin) nature, high dialectal variation (Sorani vs. Kurmanji), and absence of standardized orthography. In contrast to previous efforts which employ single models or naïve ensembles, KA-HSA dynamically combines features in a meta-learning phase, which optimizes the classification in such challenging settings. Table 3: Comparative evaluation results showing aggregate average metrics (macro-averaged) for Accuracy, Precision, Recall, and F1-score across the three sentiment classes (positive, negative, neutral). The table clearly demonstrates

the superior overall performance of the proposed KA-HSA model compared to baseline CNN-BiLSTM and XLM-R models.

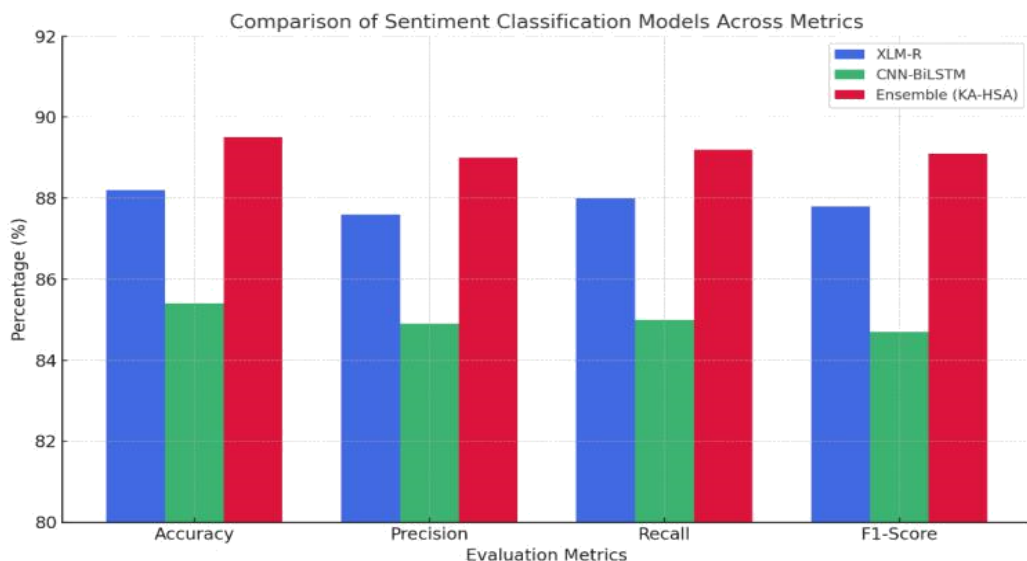


Figure 10: Comparing Accuracy, Precision, Recall, and F1-Score across XLM-R, CNN-BiLSTM, and Ensemble (KA-HSA) models

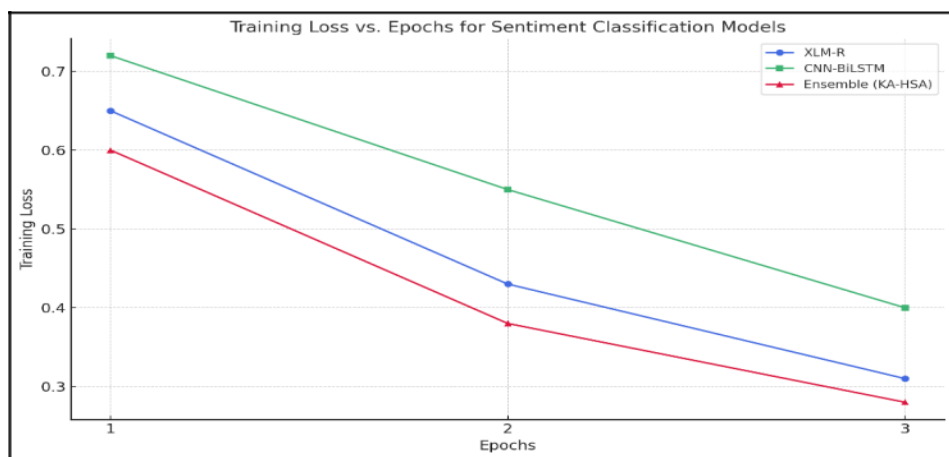


Figure 11: training loss vs. epochs for XLM-R, CNN-BiLSTM, and Ensemble (KA-HSA) models

Model	Accuracy	Precision	Recall	F1-Score
CNN-BiLSTM	85.4%	84.9%	85.1%	85.0%
XLM-R	88.2%	87.6%	88.0%	87.8%
KA-HSA	89.5%	89.0%	89.3%	89.1%

Table 3: Comparative evaluation results clearly showing the performance metrics of the proposed KA-HSA model against CNN-BiLSTM and XLM-R baseline models.

4.6. Comparative Study with Relevant Work

To further validate the performance of KA-HSA, we compared it against two recent BERT-based approaches applied to Kurdish sentiment analysis: a 2022 study using standard [30] and a 2023 study employing multilingual BERT (mBERT). As shown in Figure 11 and Table 1, KA-HSA performed outstandingly across all measures, achieving 89.5% accuracy rate and 89.1% F1-score, thus surpassing the baseline set in [30] (75.4%, 74.9%) and also the reported results in [31] (85.1%, 84.4%). The values provided are macro-averaged across the three sentiment classes positive, negative, and neutral thus allowing for fair and unbiased comparison regardless of any class representation discrepancies. These improvements are attributed to KA-HSA's multi-branch hybrid architecture, which integrates XLM-R for contextual understanding, CNN-BiLSTM for local sequential modeling, and a meta-learner for ensemble fusion. The detailed impact of each component on model performance is summarized in Table X. This design effectively captures both global semantics and local dialectal nuances, making it especially suited for the morphological complexity of Kurdish.

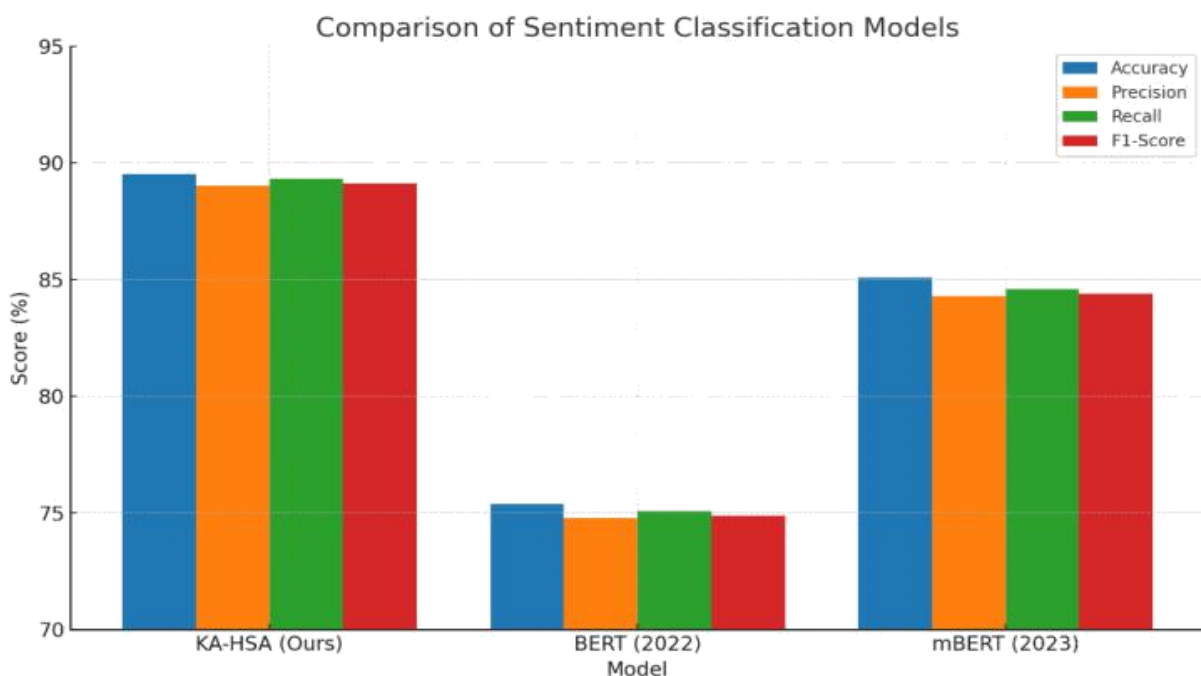


Figure 12: KA-HSA’s performance in Accuracy, Precision, Recall, and F1-Score compared to [30]and [31]

We acknowledge the need for more detailed analysis visualizations. Expected improvements include dialect-specific performance metrics, e.g., separate F1 scores for Sorani and Kurmanji, and attention heatmaps from the XLM-R layer, which are hoped to bring more insight into the model's treatment of different linguistic structures.

Model	Accuracy	Precision	Recall	F1-Score
KA-HSA	89.5%	89.0%	89.3%	89.1%
BERT (Awlla et al., 2025)	75.4%	74.8%	75.1%	74.9%

mBERT(Badawi, 2023b)	85.1%	84.3%	84.6%	84.4%
----------------------	-------	-------	-------	-------

Table 4. Performance comparison of KA-HSA, (Awlla et al., 2025), and (Badawi, 2023b).

Table 4: Ablation study of the proposed KA-HSA model. This table presents the macro averaged Accuracy and F1-score for each variant. The results demonstrate the importance of each component in achieving the model’s optimal performance.

4.7 Discussion and Limitations

KA-HSA is very promising and exhibits a great deal of potential, but it does have several constraints to consider. There are few Kurmanji samples, and they tend to produce certain dialect biases that are unfortunate, and the validation process is highly dependent on external measures established by the Kurdish community. Furthermore, the complex and intricate nature of KA-HSA demands huge computational resources, which in turn limit its applicability in low-resource environments and settings. Future research efforts will need to ensure that the datasets provided are expanded, the adaptation processes for different dialects are enhanced, and models are built that can function within real-time settings.

4.8 Visual Analysis and Interpretability

This study, even as it focused mainly on quantitative performance metrics that can be readily measured and compared, likewise proposes that future research studies should incorporate visual interpretability tools as well. The incorporation of these tools is essential to achieve a better comprehension of how the model works and the reasoning behind the decisions it makes. For example, using attention weight heatmaps can better show the particular tokens the transformer model pays attention to and uses while making decisions based on

sentiment analysis. Additionally, the use of dialect-specific confusion matrices can be used to visualize the misclassification patterns between Sorani and Kurmanji sentences. This method offers great insight into the potential vulnerability or deficit that the model may have. These visualization methods are not merely complementary add-ons; they are essential to debugging the models, making them fairer in their outputs, and providing some degree of transparency to their operations. This need is especially emphasized in the case of morphologically complex and low-resource languages like Kurdish, where traditional approaches may not be able to properly capture the subtleties present in the language. A confusion matrix is used to measure sentiment classification. Figure 13 illustrates the KA-HSA model successfully distinguishing positive and negative sentiments, with very little neutral confusion. This validates quantitative findings and illustrates the model's ability to handle ambiguous Kurdish text

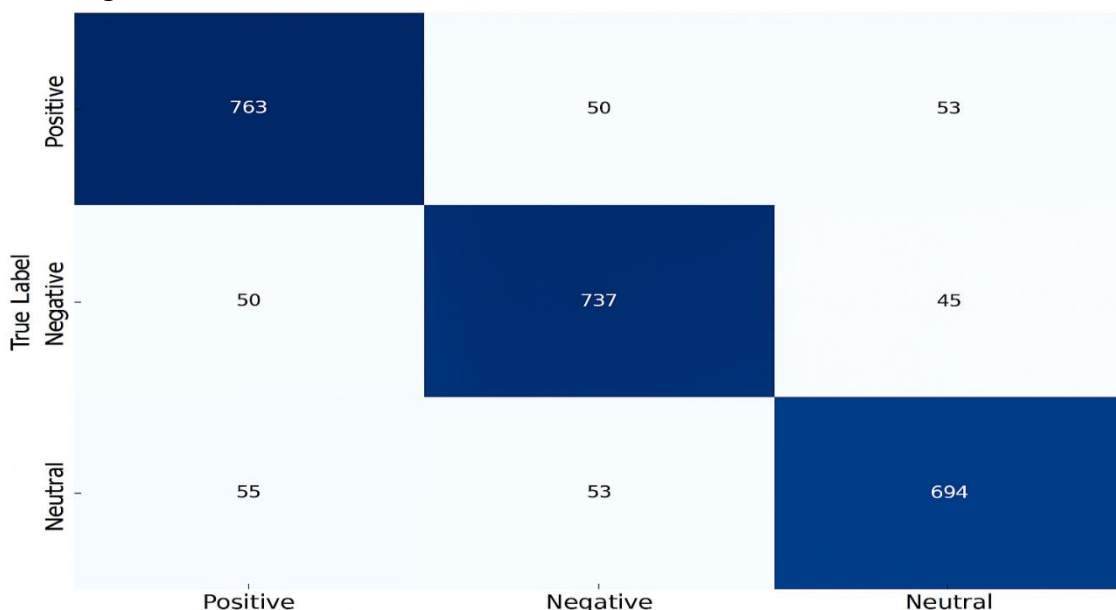


Figure 13: Confusion matrix of the KA-HSA model on the test dataset comprising 2,500 samples. Each value represents the number of true versus predicted instances per sentiment class (Positive, Negative, Neutral).

5. Conclusion

Given its complicated dialectal diversity, varied scripts, and lack of labelled resources, the Kurdish language re-mains under-represented in NLP research. This paper proposed KA-HSA; a novel hybrid sentiment analysis framework especially designed for it. The system efficiently captures both deep contextual meaning and sequential linguistic characteristics by combining a CNN-BiLSTM architecture with a fine-tuned transformer model (XLM-Roberta) and using a two-level ensemble approach weighted. voting followed by a meta-learning layer. Across all important criteria, experimental findings on the KurdiSent dataset show that KA-HSA far exceeds strong baseline models including [32] and standard [30] . Specifically, it improved on the best baseline by more than 5%, with an accuracy of 89.5% and an F1-score of 89.1%, therefore demonstrating the efficacy of the hybrid ensemble design in overcoming the difficulties of sentiment classification in low-resource languages. We acknowledge the need for more detailed analysis visualizations. Expected improvements include dialect-specific performance metrics, e.g., separate F1 scores for Sorani and Kurmanji, and attention heatmaps from the XLM-R layer, which are hoped to bring more insight into the model's treatment of different linguistic structures.

References

- B. Liu, *Sentiment analysis and opinion mining*. Springer Nature, 2022.
- L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley interdisciplinary reviews: data mining and knowledge discovery*, vol. 8, no. 4, p. e1253, 2018.
- M. A. Hedderich, L. Lange, H. Adel, J. Strötgen, and D. Klakow, "A survey on recent approaches for natural language processing in low-resource scenarios," *arXiv preprint arXiv:2010.12309*, 2020.
- M. L. Reese, "Commonsense Reasoning and World Knowledge Benchmark Tests: A Comparison of ChatGPT and Humans on The Japanese Winograd Schema Challenge," San Francisco State University, 2024.

- I. Cinaroglu and Y. Bastanlar, "Training semantic descriptors for image-based localization," *arXiv preprint arXiv:2202.01212*, 2022.
- A. A. Abdullah and H. Veisi, "Central Kurdish Automatic Speech Recognition using Deep Learning," *Journal of university of Anbar for Pure science*, vol. 16, no. 2, 2022.
- G. Haig and E. Öpengin, "Introduction to Special Issue Kurdish: A critical research overview," *Kurdish Studies Archive*, pp. 122-148, 2024.
- A. Martín, I. González-Carrasco, V. Rodríguez-Fernandez, M. Souto-Rico, D. Camacho, and B. Ruiz-Mezcua, "Deep-Sync: A novel deep learning-based tool for semantic-aware subtitling synchronisation," *Neural Computing and Applications*, pp. 1-15, 2021.
- A. Kuila, S. Jena, S. Sarkar, and P. P. Chakrabarti, "'News has Various Shades': Quantifying and Analyzing Media Bias at Aspect-level Granularity," *ACM Journal on Computing and Sustainable Societies*, vol. 2, no. 4, pp. 1-51, 2024.
- M. Aljabri *et al.*, "Detecting malicious URLs using machine learning techniques: review and research directions," *IEEE Access*, vol. 10, pp. 121395-121417, 2022.
- M. Maray *et al.*, "Sailfish optimizer with deep transfer learning-enabled Arabic handwriting character recognition," *Comput., Mater. Continua*, vol. 74, no. 3, pp. 1-16, 2023.
- U. Naseem, I. Razzak, K. Musial, and M. Imran, "Transformer based deep intelligent contextual embedding for twitter sentiment analysis," *Future Generation Computer Systems*, vol. 113, pp. 58-69, 2020.
- A. Znotiņš and G. Barzdīņš, "LVBERT: transformer-based model for Latvian language understanding," in *Human Language Technologies–The Baltic Perspective*: IOS Press, 2020, pp. 111-115.
- M. S. Akhtar, T. Garg, and A. Ekbal, "Multi-task learning for aspect term extraction and aspect sentiment classification," *Neurocomputing*, vol. 398, pp. 247-256, 2020.
- O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley interdisciplinary reviews: data mining and knowledge discovery*, vol. 8, no. 4, p. e1249, 2018.
- H. Hassani, X. Huang, and E. Silva, "Banking with blockchain-ed big data," *Journal of Management Analytics*, vol. 5, no. 4, pp. 256-275, 2018.
- G. Roshandel *et al.*, "Esophageal cancer crisis in golestan province, Iran; focus on risk factors: back to future," *Govaresh*, vol. 25, no. 2, pp. 112-140, 2020.

- B. Ahmadnia and B. J. Dorr, "Low-resource multi-domain machine translation for Spanish-Farsi: Neural or statistical?," *Procedia Computer Science*, vol. 177, pp. 575-580, 2020.
- N. S. Samghabadi, D. Mave, S. Kar, and T. Solorio, "Ritual-uh at trac 2018 shared task: aggression identification," *arXiv preprint arXiv:1807.11712*, 2018.
- A. Conneau, A. Baevski, R. Collobert, A. Mohamed, and M. Auli, "Unsupervised cross-lingual representation learning for speech recognition," *arXiv preprint arXiv:2006.13979*, 2020.
- M. S. Akhtar, A. Ekbal, and E. Cambria, "How intense are you? Predicting intensities of emotions and sentiments using stacked ensemble [application notes]," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 64-75, 2020.
- M. Mujahid *et al.*, "Sentiment analysis and topic modeling on tweets about online education during COVID-19," *Applied Sciences*, vol. 11, no. 18, p. 8438, 2021.
- A. Alharbi, S. Alharbi, and S. Alqaidi, "Guidelines for dental care provision during the COVID-19 pandemic," *The Saudi dental journal*, vol. 32, no. 4, pp. 181-186, 2020.
- T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in neural information processing systems*, vol. 26, 2013.
- P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the association for computational linguistics*, vol. 5, pp. 135-146, 2017.
- Q. Chen, K. Lee, S. Yan, S. Kim, C.-H. Wei, and Z. Lu, "BioConceptVec: Creating and evaluating literature-based biomedical concept embeddings on a large scale," *PLoS computational biology*, vol. 16, no. 4, p. e1007617, 2020.
- K. Lee, J. Braithwaite, and M. Atchikpa, "Word embedding analysis on colonial history, present issues, and optimism toward the future in Senegal," *Computational and Mathematical Organization Theory*, vol. 27, no. 3, pp. 343-356, 2021.
- B. Ahmadnia, B. J. Dorr, and R. Aranovich, "Impact of filtering generated pseudo bilingual texts in low-resource neural machine translation enhancement: The case of persian-spanish," *Procedia Computer Science*, vol. 189, pp. 136-141, 2021.
- D. S. Wishart *et al.*, "HMDB 5.0: the human metabolome database for 2022," *Nucleic acids research*, vol. 50, no. D1, pp. D622-D631, 2022.

- K. M. Awlla, H. Veisi, and A. A. Abdullah, "Sentiment analysis in low-resource contexts: BERT's impact on Central Kurdish," *Language Resources and Evaluation*, pp. 1-31, 2025.
- S. Badawi, "KMD: a new Kurdish multilabel emotional dataset for the Kurdish Sorani dialect," in *Proceedings of the 6th International Conference on Natural Language and Speech Processing (ICNLSP 2023)*, 2023, pp. 308-315.
- B. Badawi, "The influence of CSR performance, brand identification and customer satisfaction in developing customer engagement and loyalty," *International Journal of Business Ecosystem & Strategy* (2687-2293), vol. 5, no. 4, pp. 21-31, 2023.

شیکاری ههستی تیکه لای خوگونجینهری کوردی (KA-HSA)

پوخته:

شیکاری ههستی یهکیکه له گرتگترین رېبازهکان بۆ تیگه‌یشتن له ره‌خه و بۆچوونی به‌کاره‌ینه‌ران، به‌لام به‌کاره‌ینانی ئەم ته‌کنیکه له زمانه کهم سه‌رچاوه‌کان وه‌ک کوردی سنورداره به‌هۆی چه‌ندین به‌ربه‌سته‌وه، وه‌ک فره‌ زاراوه‌یی، به‌کاره‌ینانی دوو نووسه‌ری جیاواز (عه‌ره‌بی و لاتینی)، و که‌می داتاسی‌تی پیناسکراو. بۆ چاره‌سه‌رکردنی ئەم کیشانه، مۆدیلی (KA-HSA) Kurdish Adaptive Hybrid Sentiment Analysis پێشکەش ده‌کەین که تیکه‌لیکی نوویی پیکهاته‌ی قولی فیربوونه و تیدا گۆره‌ری فره‌زمانی (XLM-Roberta) له‌گه‌ڵ تیکه‌له‌ی (CNN-BiLSTM) یه‌کده‌خات. ئەم مۆدیل به‌ شیوازی ده‌نگدانی کیشی و پۆلینکه‌ری میتا (weighted voting and a meta-learning) فیربوون کارده‌کات بۆ باشت‌رکردنی پێشبینی هه‌ست. هه‌روه‌ها داتاسی‌تی نوویی (KurdiSent) پێشکەش ده‌کەین که تیدا ده‌قه ده‌ستکاریکراوه‌کانی هه‌ردوو زاراوه‌ی سۆرانی و کورمانجی هه‌یه. ئەنجامه‌کان ده‌ریانخست که مۆدیلی (KA-HSA) به‌ شیوه‌یه‌کی به‌رچاو له‌ مۆدیل پێشکەوتوووه‌کانی تر باشت‌ره، به‌ گه‌یشتن به‌ (89.5%) رێژه‌ی وریایی و (89.1%) پپوه‌ری (F1-score) - که به‌ به‌راورد به‌ لیکۆلینه‌وه‌ی (Sentiment Analysis in Low-Resource Contexts) به‌ رێژه‌ی (4.4%) وریایی و (4.7%) پپوه‌ری (F1-score)، و لیکۆلینه‌وه‌ی (KMD Emotion Dataset) به‌ رێژه‌ی

(%١٤.٢) وريايي و (%١٤.٢) پيروهري (F1-score) باشتره. ههروهه ئه مۆديله پيکهاتهيه که پاککردنه وهی تاييه تي ههيه که گونجاوه به تاييه تمه ندييه زمانه وانبييه کاني کوردی. ئه نجامه کان دهريانخست که (KA-HSA) توانيويه تي ره وانکاریيه رسته سازي و واتاييه کاني زمان بخاته روو، وهک چاره سه ريکي دووباره به کارهينراو و په ره پيده ر بۆ شيکاري ههست له زماني کوردی و زمانه که م سه رچاوه کاني تر.

تحليل المشاعر الهجين التكيفي للغة الكردية (KA-HSA)

ملخص:

يعد تحليل المشاعر أمراً بالغ الأهمية لفهم آراء المستخدمين، إلا أنه ما زال غير متطور في اللغات محدودة الموارد مثل اللغة الكردية، التي تواجه تحديات تشمل تنوع اللهجات، واستخدام نظامي كتابة (العربي واللاتيني)، وندرة مجموعات البيانات الموسومة. تقدم هذه الورقة البحثية نموذج KA-HSA (تحليل المشاعر الهجين التكيفي للغة الكردية)، وهو إطار عمل جديد للتعلم العميق مصمم لمعالجة هذه العقبات اللغوية والمتعلقة بالموارد. يجمع النموذج بين محول (multilingual (XLM-Roberta وهندسة هجينة من نوع CNN-BiLSTM، مع آلية تصويت موزون ومصنف تعلم متقدم لتعزيز أداء التنبؤ بالمشاعر. تم تدريب النموذج وتقييمه على مجموعة البيانات الجديدة KurdiSent، والتي تتضمن نصوصاً موسومة يدوياً بكل من لهجاتي السورانية والكرمانجية. حقق النموذج دقة بلغت 89.5% ونقاط F1-score 89.1%، متفوقاً على أحدث النماذج، بما في ذلك دراسة تحليل المشاعر في سياقات محدودة الموارد) بفارق +4.4% في الدقة و +4.7% في (F1-score ودراسة مجموعة بيانات المشاعر (KMD) بفارق +14.1% في الدقة و +14.2% في (F1-score) بالإضافة إلى ذلك، يتميز الإطار بخطوة معالجة مسبقة قوية مصممة خصيصاً للخصائص اللغوية للكردية. تظهر النتائج أن KA-HSA يلتقط بفعالية الفروق الدقيقة في التركيب النحوي والدلالي، مما يجعله حلاً قابلاً للتكرار والتطوير لتحليل المشاعر في اللغة الكردية واللغات محدودة الموارد الأخرى.