

1. Introduction & Motivation

Sarcasm detection enhances sentiment analysis and content moderation in multilingual social media.

Challenges: Cross-language semantics, limited dialectal resources, contextual nuance.

Contributions: Unified BiLSTM+Attention architecture; bilingual preprocessing; extensive evaluation.

2. Data & Preprocessing

Datasets: ArSarcasm-v2 (Arabic, 27 479 tweets) [1]; News Headlines Dataset for Sarcasm Detection (English, 5 889 headlines) [2].

Pipeline: Clean text (remove HTML, URLs, mentions), normalize diacritics, lowercase, tag language, tokenize (vocab=25 002), pad/truncate to 24 tokens, balance classes.

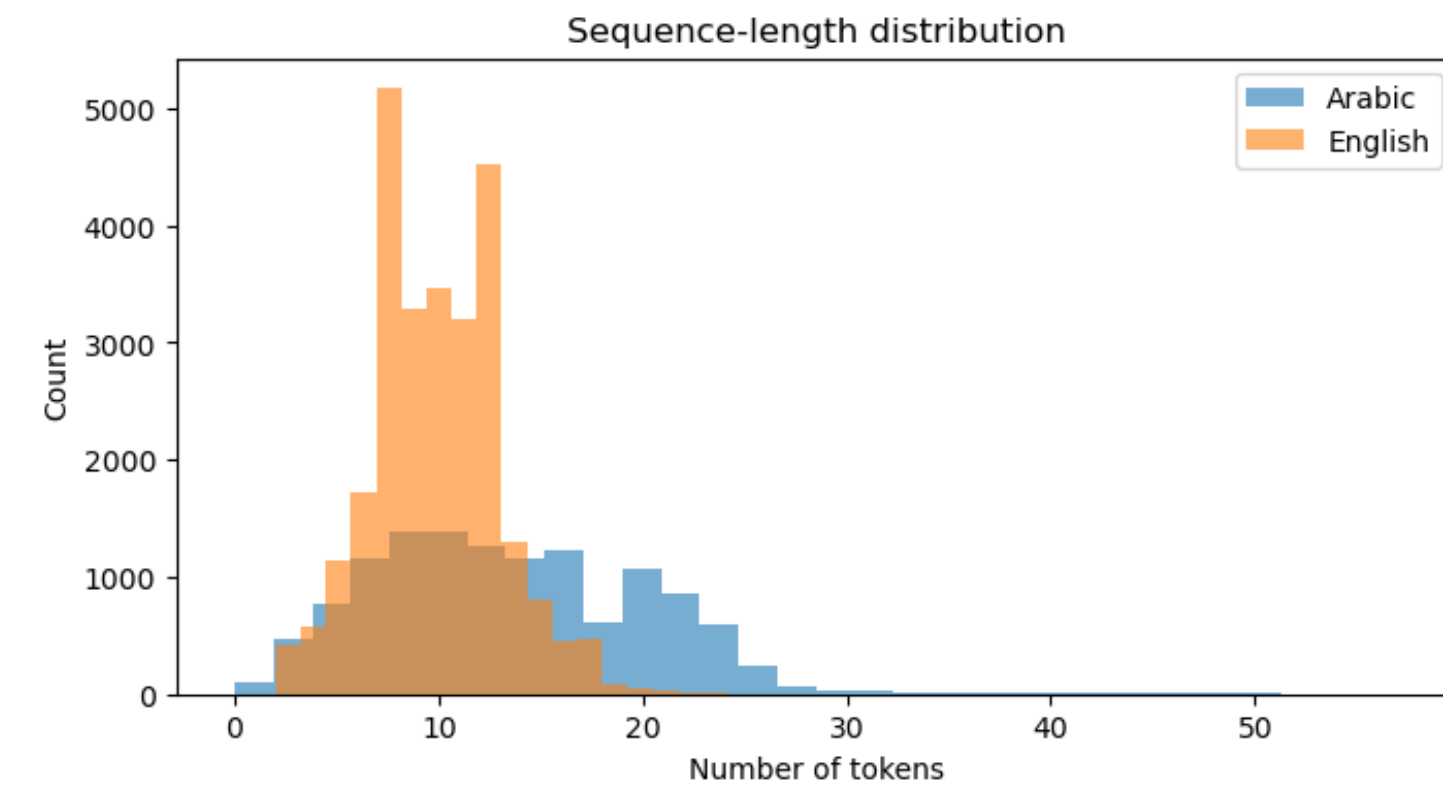


Figure 1. Token-length distribution for Arabic vs. English samples.

3. Objectives

- Build a robust binary classifier for Arabic & English sarcasm.
- Maximize F1-score ROC AUC through dropout, LR scheduling, and L2 regularization.
- Visualize learned embeddings interpret attention distributions.

4. Sample Predictions

Illustrative examples:

Sarcasm: “Great, another Monday morning...” → Sarcastic (p=0.85)

Non-sarcasm: “I love this weather.” → Not Sarcastic (p=0.10)

5. Model Architecture

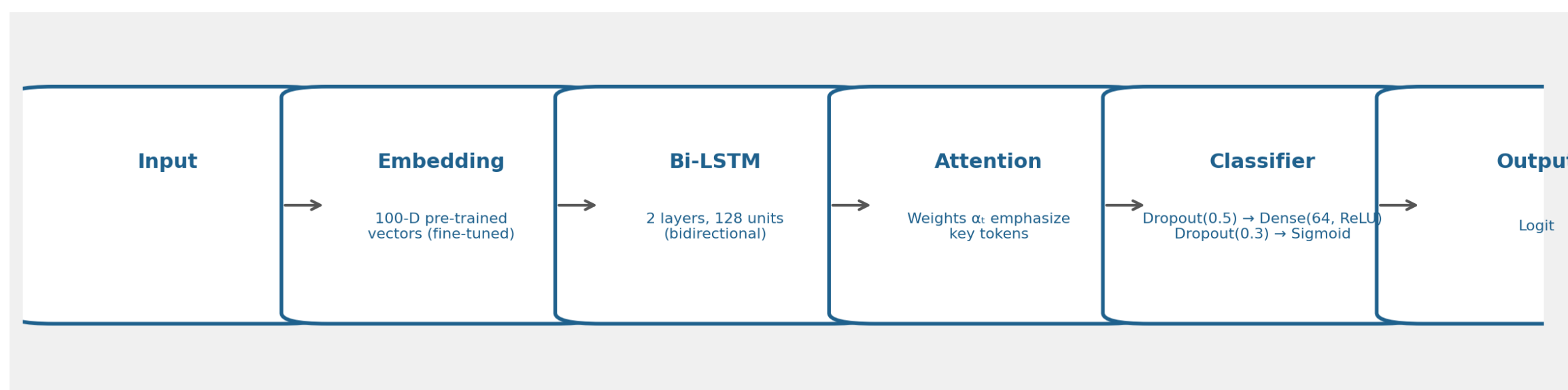


Figure 2. Embedding → BiLSTM → Attention → Classification head.

Key Components:

- Embedding: 100-D pretrained vectors, fine-tuned during training.
- BiLSTM: 2 layers, 128 units in each direction for contextual encoding.
- Attention: Computes token-level weights α_t over hidden states.
- Classification: Dropout(0.5) → Dense(64, ReLU) → Dropout(0.3) → Dense(1, Sigmoid).

6. Training Validation

Setup: 70/15/15 split; weighted BCE loss; Adam (LR=5e-4, weight decay=1e-4); CosineAnnealingLR; early stopping (patience=5).

Metrics: Accuracy, F1-score, ROC AUC per epoch.

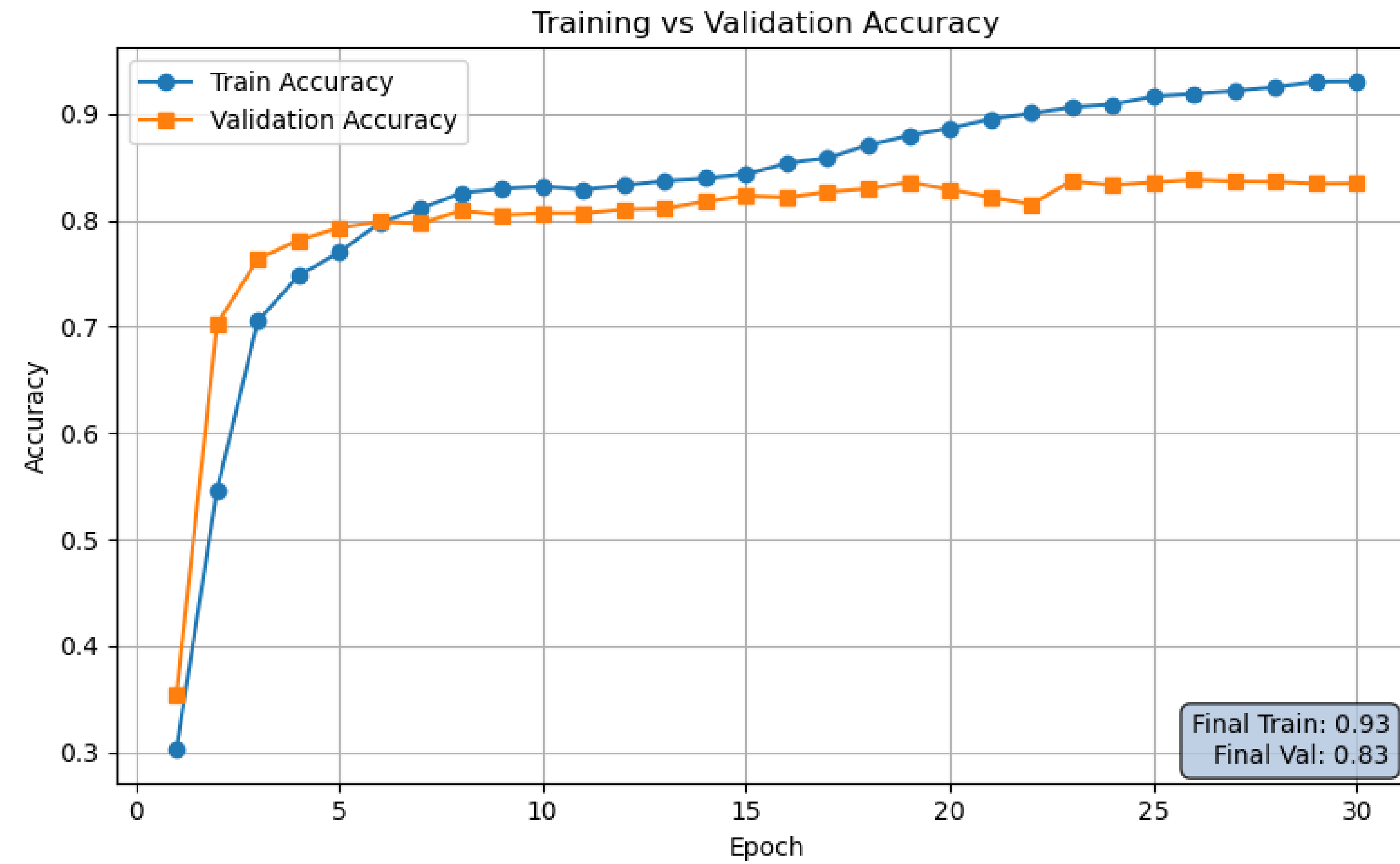


Figure 3. Training vs. Validation Accuracy over epochs.

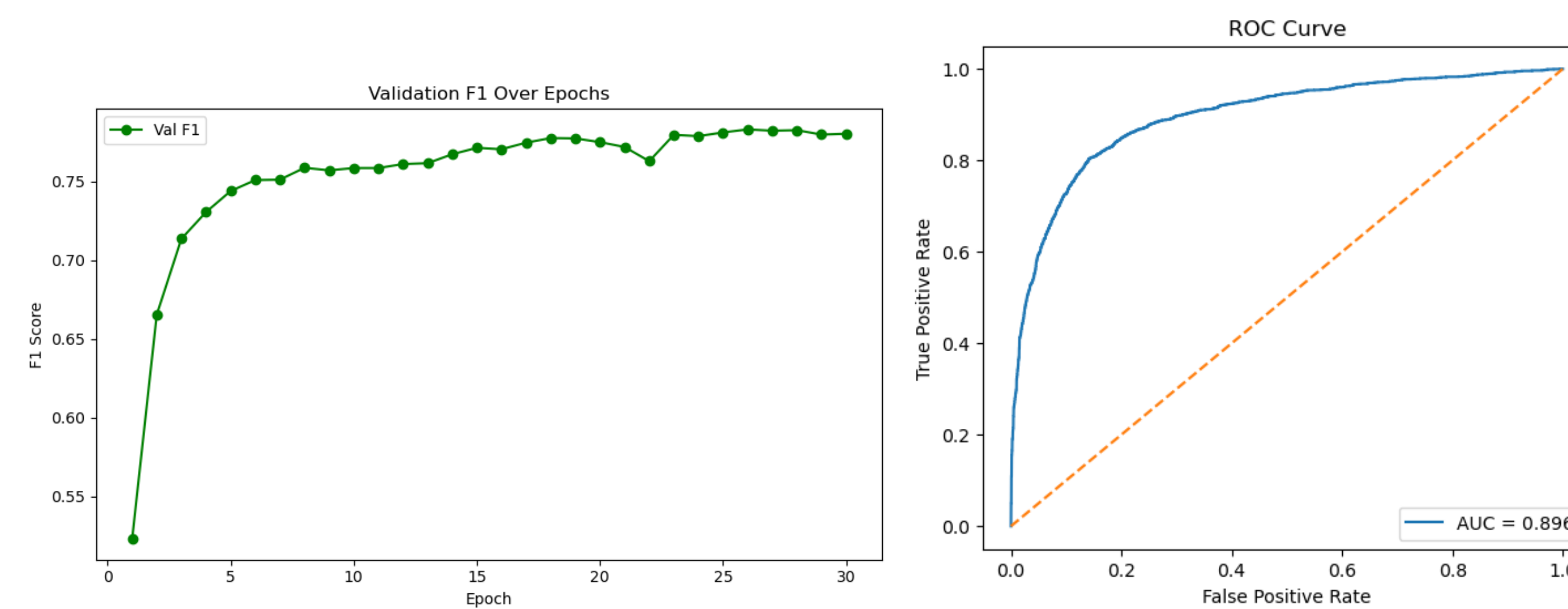


Figure 4. (left) Validation F1 vs. epochs; (right) ROC Curve (AUC = 0.896).

7. Test Results

Test performance: Acc = 84.5%, F1 = 0.7805, AUC = 0.896.

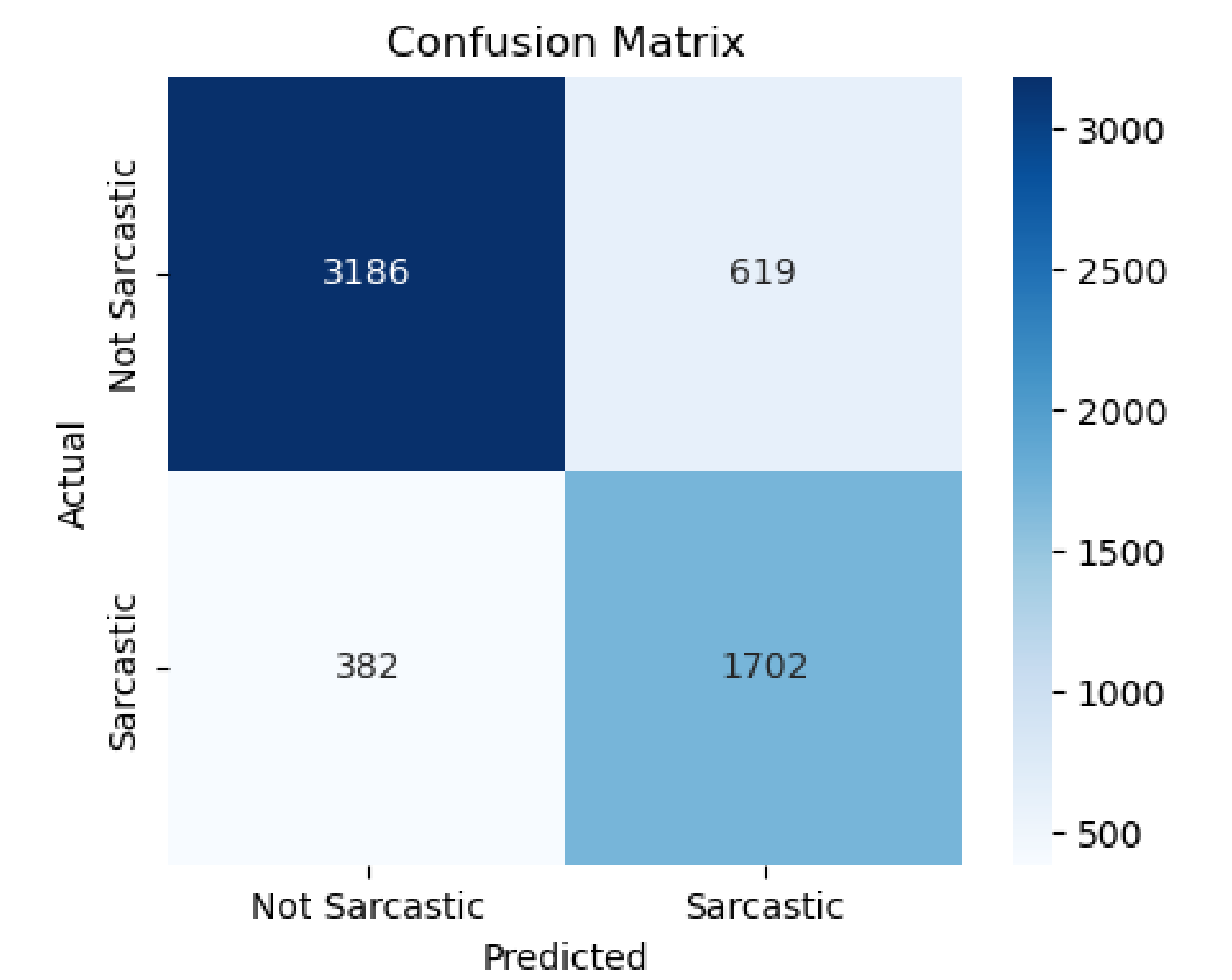


Figure 5. Confusion matrix on test set.

8. Embedding PCA Visualization

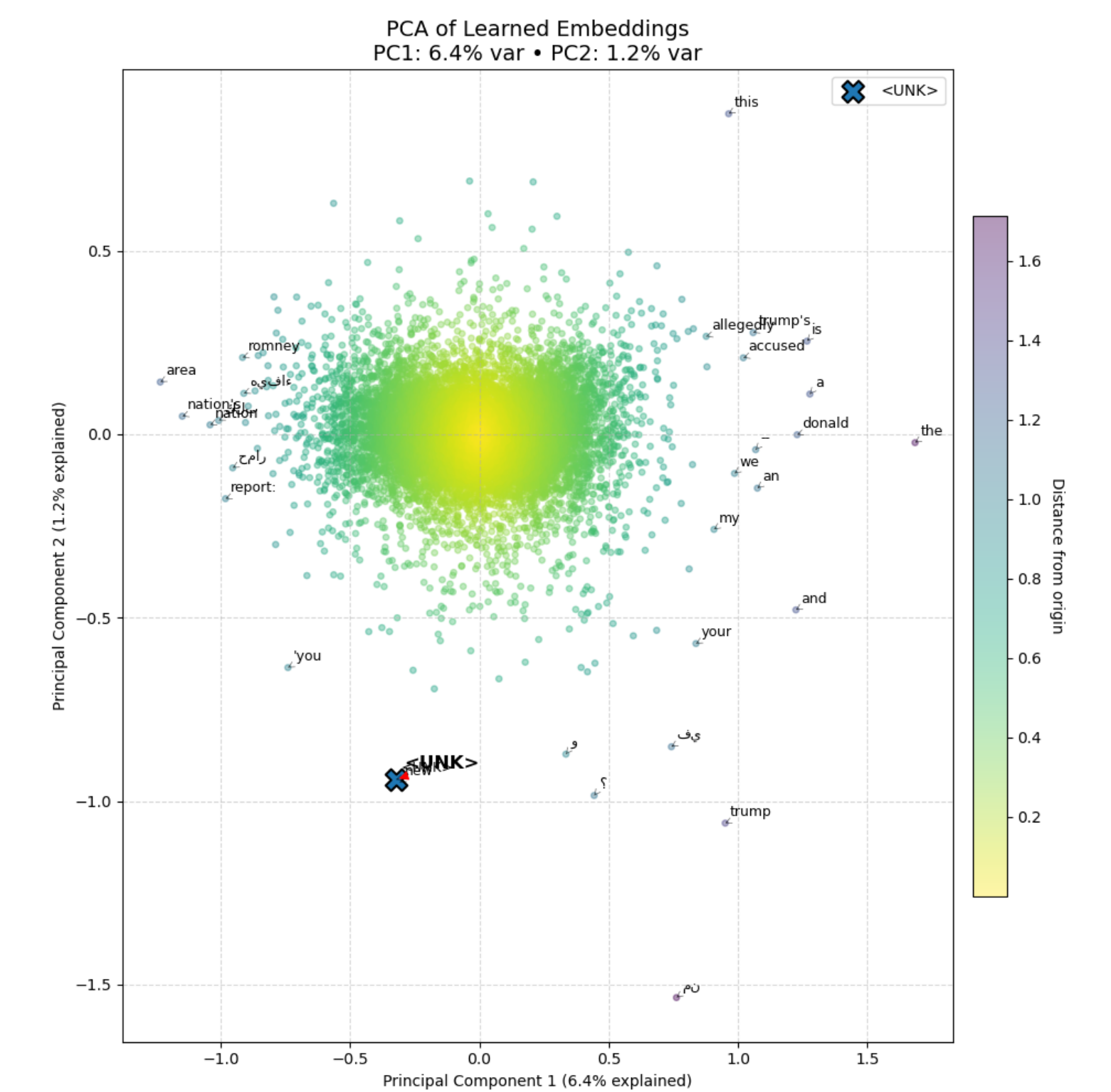


Figure 6. PCA projection of learned token embeddings (PC1: 6.4%, PC2: 1.2%).

9. Discussion Future Work

Attention captures salient sarcasm cues; 10% train-val gap indicates good generalization. Future: integrate transformer backbones; enlarge dialectal corpora; deploy via REST API.

10. References

- Abrar Alotaibi and Hend Al-Khalifa. ArSarcasm-v2: Arabic Sarcasm Detection Dataset. Kaggle, 2022. [Online]. Available: <https://www.kaggle.com/datasets/abraralotaibi00/arsarcasm-v2>
- Rishabh Misra and Prahal Arora. News Headlines Dataset for Sarcasm Detection. Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/rmisra/news-headlines-dataset-for-sarcasm-detection>