

Natural language inference

Using NLU : SVM vs. Bi-LSTM

Authors Maryam Alkhowildi and Khawla Almarzooqi
The University of Manchester, Department of Computer Science.

In the field of Natural Language Understanding (NLU), Natural Language Inference (NLI) plays a critical role by enabling machines to determine the relationships between parts of the text and infer whether one statement logically follows from another. This capability is essential for understanding complex human language in various applications. Given the advancement of this field, it's important to share knowledge. Research posters play an important role in this process by simplifying ideas about NLI into clear and interesting overviews. This helps experts and those working in the field stay informed about new developments in NLU studies.

Introduction

- NLI challenges AI to discern whether a hypothesis is true (entailment) or false (contradiction) based on a given premise.
- The development and the comparison of Support Vector Machine (SVM) and Bidirectional Long Short-Term Memory (Bi-LSTM) models seeks to uncover which technique more accurately classifies these nuanced textual relationships, thus enhancing machine understanding of complex language dynamics.

Objective

- Develop two supervised classification models for NLI: SVM and Bi-LSTM.
- Compare their performance on NLI task.
- Identify their strengths and weaknesses in handling linguistic nuances.
- Provide insights into the selection of appropriate models for NLI task.

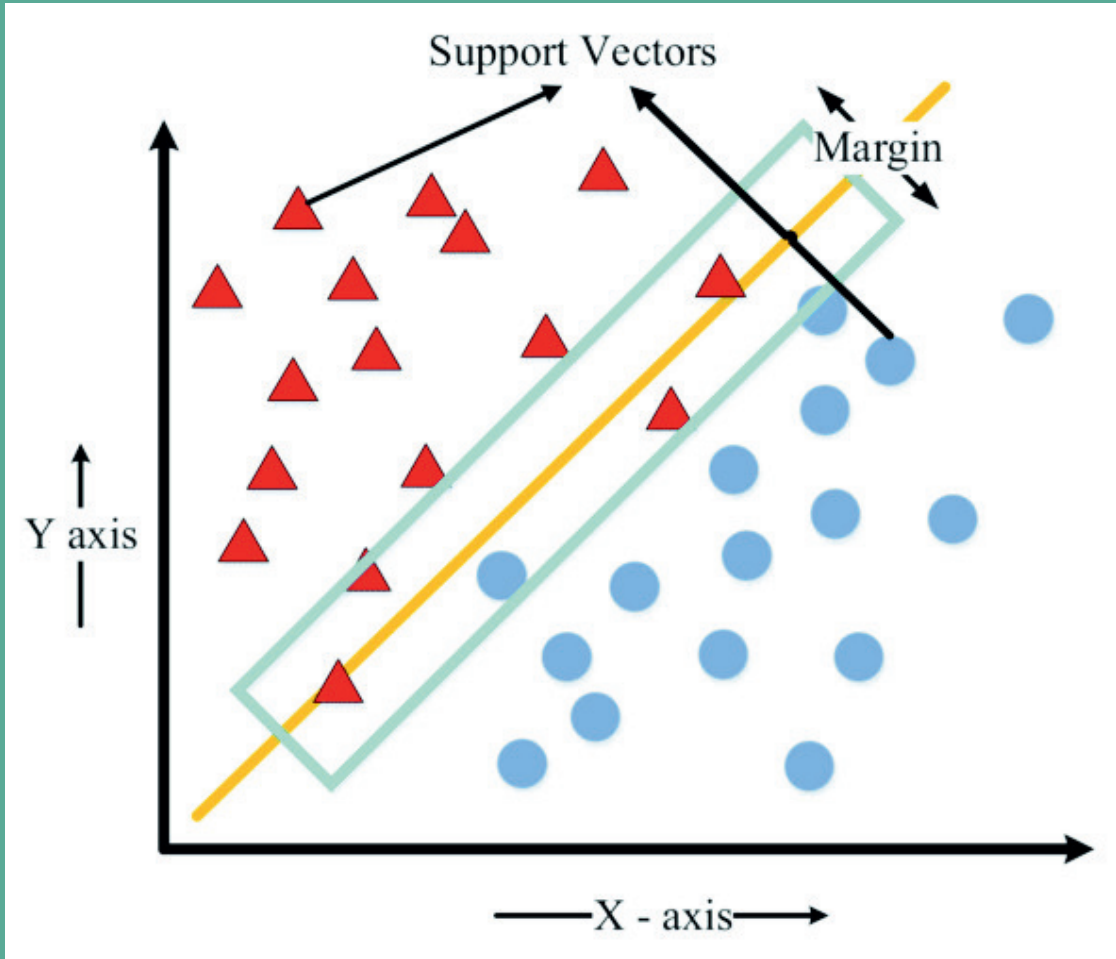
Methodology

SVM Model

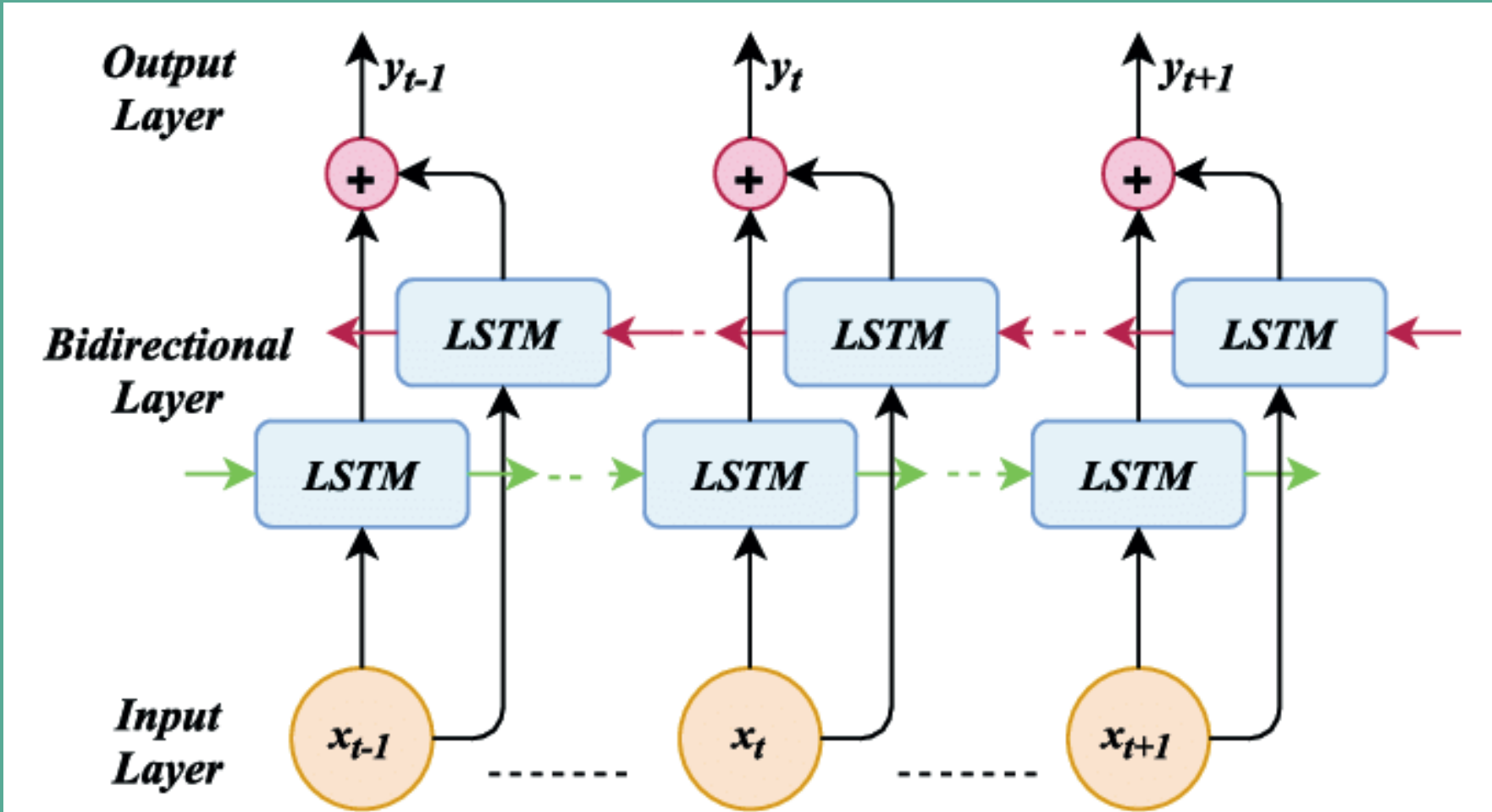
- SVM is used for its ability to create distinct boundaries between data points of different classes, which is essential for clear and robust categorization in text analysis.
- Data was vectorized using TF-IDF to convert text for SVM analysis.
- A linear kernel SVM was implemented, with regularization to prevent overfitting.
- The model was trained and validated against a subset of the NLI dataset.

Bi-LSTM Model

- BiLSTM enhances the model's context understanding by processing data in both forward and backward directions, which is key for tasks where the sequence of the input data is important.
- Text data underwent tokenization and padding to create uniform input sequences.
- The Bi-LSTM network, with 64 units in each direction, was employed to capture bidirectional context.
- Dropout and embedding techniques were applied to enhance model robustness and feature learning.



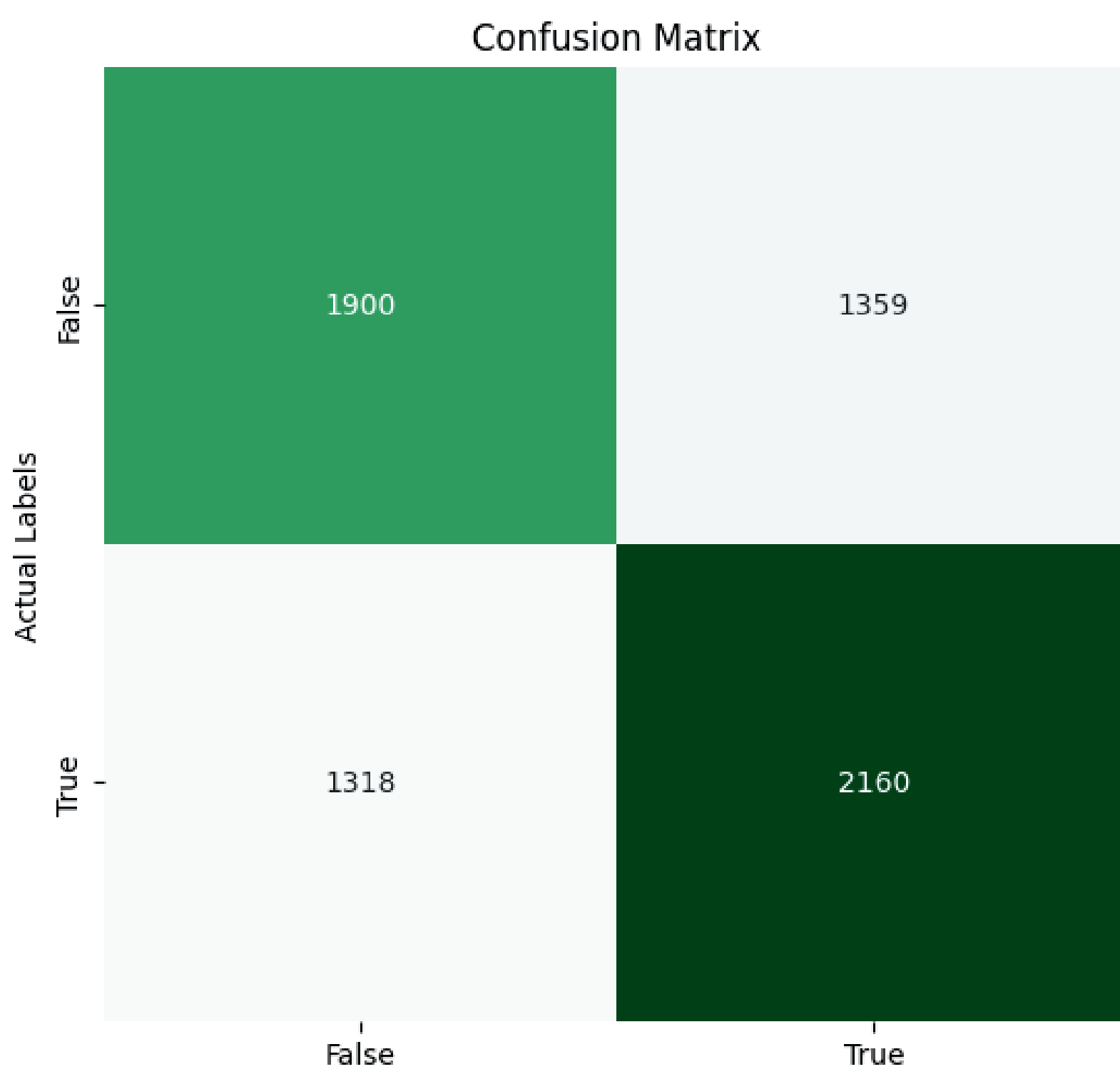
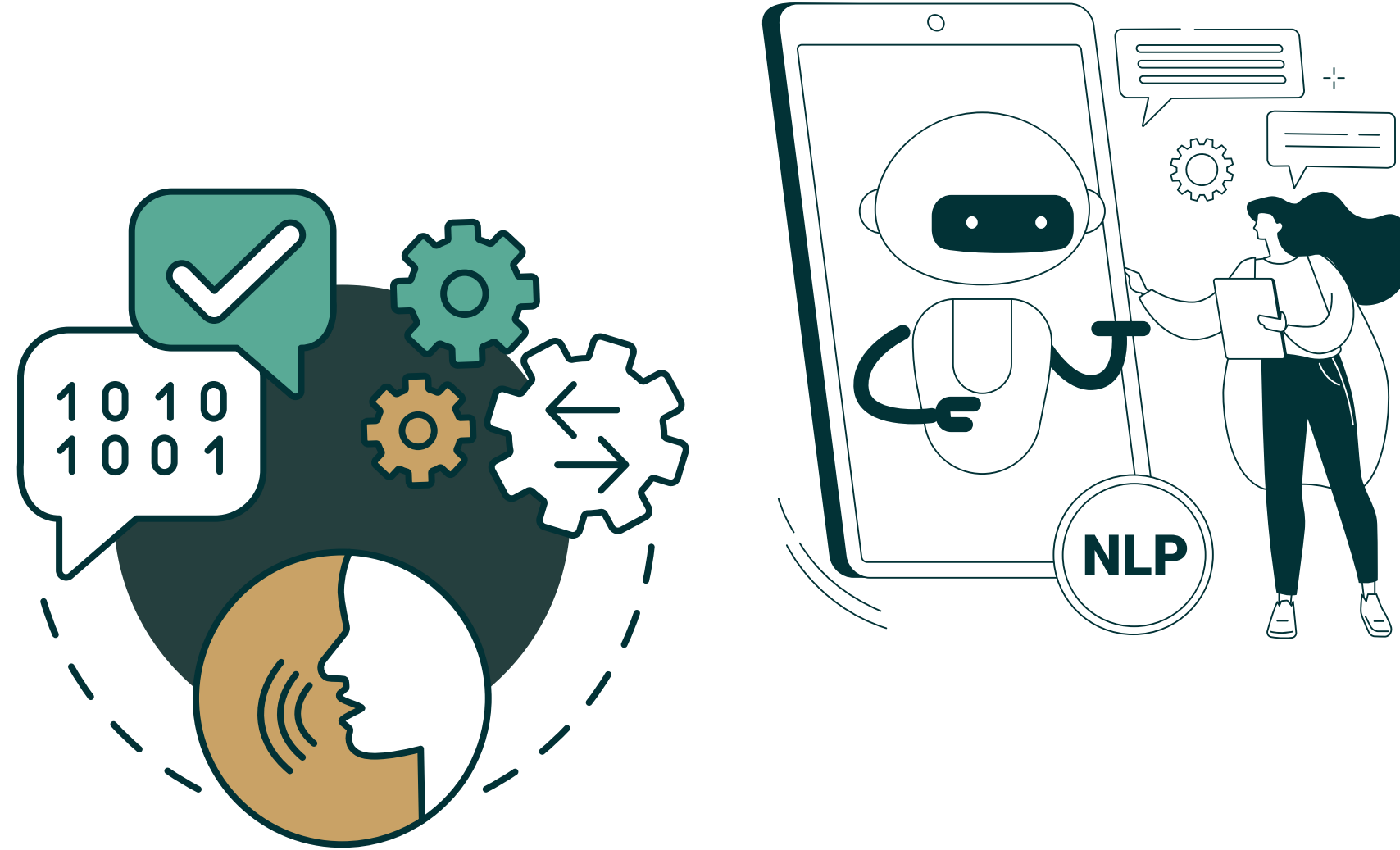
Binary Classification Using SVM [1]



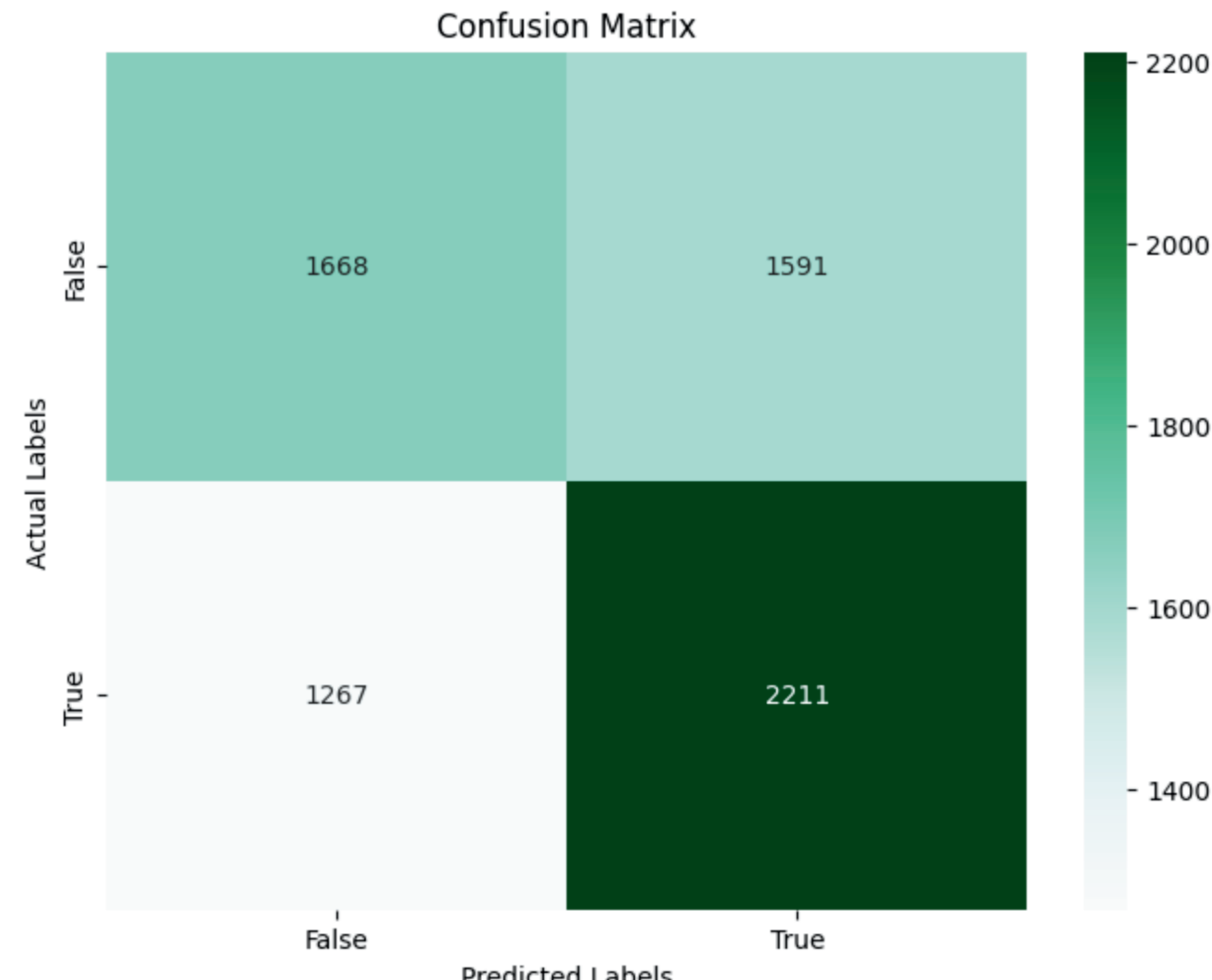
Bi-LSTM network architecture [2]

Results

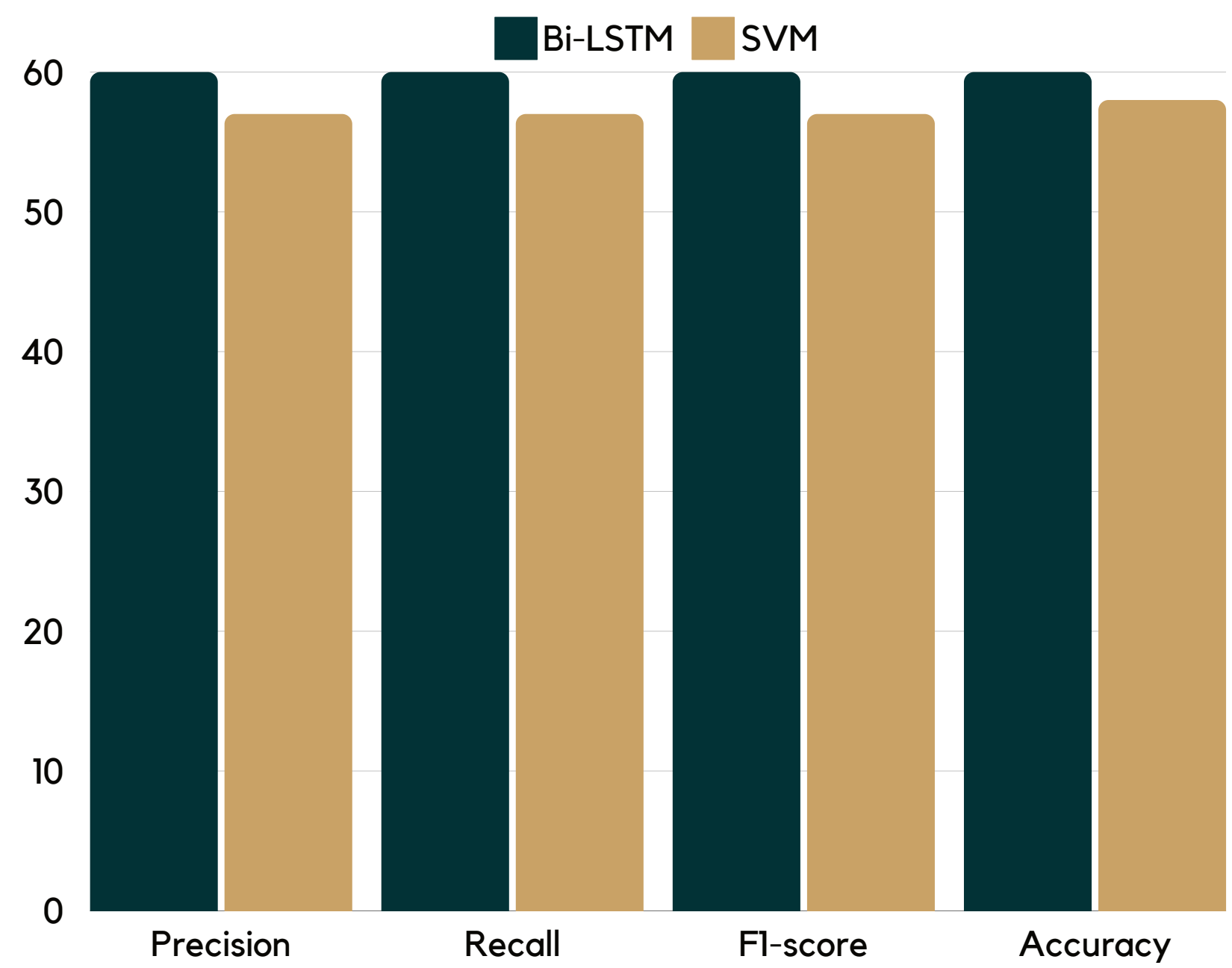
- The Bi-LSTM model achieved a higher accuracy than the SVM in the NLI task.
- The Bi-LSTM model further exhibits strengths in precision and recall.
- Confusion matrices indicated a more balanced prediction distribution for Bi-LSTM, reflecting its enhanced capability to process linguistic context.
- This nuanced processing of language nuances is attributable to the Bi-LSTM's architecture, which considers both preceding and succeeding textual elements, providing a comprehensive understanding that enhances performance on tasks requiring deep semantic analysis.



Bi-LSTM Confusion Matrix



SVM Confusion Matrix



Performance Metrics Bi-LSTM vs. SVM.

Conclusion

- **Bi-LSTM vs. SVM:** Bi-LSTM outperformed SVM with improved accuracy in NLI, underscoring the benefits of such context-aware models.
- **Interesting Results:** The slight advantage in precision and recall for Bi-LSTM suggests a growing edge for deep learning models in language understanding tasks.
- **Relevance to Existing Research:** The results add to the wider understanding in NLU research that models with a strong contextual basis tend to produce more accurate interpretations.

Implications & Future Steps

- The performance of Bi-LSTM models may result in better AI-driven language services.
- Further research could focus on exploring additional supervised classification models for NLI and assessing their effectiveness with more diverse language datasets.

Related Literature

[1] Muzzammel, R., & Raza, A. (2020). A support vector machine learning-based protection technique for MT-HVDC systems. *Energies*, 13(24), 6668.

[2] Ihianle, I. K., Nwajana, A. O., Ekenwa, S. H., Otuka, R. I., Owa, K., & Orisatoki, M. O. (2020). A deep learning approach for human activities recognition from multimodal sensing devices. *IEEE Access*, 8, 179028-179038.