Assignment 2

Detecting Adverse Drug Events using Sequence Models

Prashant Tandon
B21AI053

- Literature Review
  - Adverse drug event detection using natural language processing
  - MADEx: A System for Detecting Medications, Adverse Drug Events, and their Relations from Clinical Notes
  - Bidirectional LSTM-CRF for Adverse Drug Event Tagging in Electronic Health Records
- Literature Review for Model Architecture
  - Paper - Named Entity Recognition with Bidirectional LSTM-CNNs
  - Paper - Mining Adverse Drug Reactions from Unstructured Mediums at Scale
- Implementation
  - Dataset Description
  - Preprocessing Pipeline
    - Text Processing
  - Embeddings
    - SpaCy
    - GloVe
  - Models
    - BLSTM
    - GRU
  - Performance Evaluation
    - SpaCy Embeddings
    - Glove Embeddings
  - Recommendations for Future Research

Literature Review

Adverse drug event detection using natural language processing

Paper Link - https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0279842

This paper synthesizes findings from studies focusing on the detection of adverse drug events (ADEs) utilizing natural language processing (NLP) techniques, with a particular emphasis on supervised learning methods.

Data preparation emerges as a significant challenge, with complexities stemming from domain-specific clinical text, frequent use of abbreviations, and unconventional formatting. Custom tools and collaborative efforts are recommended to address these hurdles, especially for non-English languages. Annotated gold standard corpora creation demands substantial effort, with semi-automated methods showing promise in reducing annotation time. Involving clinicians in annotation scheme design is crucial for ensuring data accuracy.

Modeling approaches primarily feature Long Short Term Memory (LSTM) and Conditional Random Field (CRF) variations, adept at handling sequential information and contextual dependencies. However, fair comparisons between methods are hindered by differences in dataset size, annotation quality, and performance metric variations. Performance evaluation predominantly relies on F1 scores, with a focus on assessing ADE detection capabilities amidst contextual nuances. Various sources report their performance in similar metrics but differ in their methods of calculation. None of the articles indicated a cut-off value for determining a good performance in advance of performing the analysis.

Despite reported high ADE entity numbers, caution is warranted due to potential dataset biases and lack of formal causality assessment. Class imbalance correction and validation strategies such as k-fold cross-validation are recommended to enhance model robustness. Deployment of NLP applications in clinical practice remains limited, with future implementation plans emphasizing the need for translating research findings into practical healthcare solutions.

MADEx: A System for Detecting Medications, Adverse Drug Events, and their Relations from Clinical Notes

Paper Link - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6402874/

Identifying ADEs from Electronic Health Records (EHRs) is challenging due to the narrative nature of clinical text. Clinical NLP is crucial for extracting relevant information from such unstructured data. The article introduces MADEx, a system designed for detecting medications, adverse drug events (ADEs), and their relationships within clinical notes. This system was developed specifically for the MADE1.0 challenge.

Dataset

MADEx was trained and tested on a dataset comprising 1,089 de-identified clinical notes. These notes were annotated with medication names and attributes, ADEs, indications, and other relevant information, totaling 79,114 entities and 27,175 relations. The annotations were represented using the BioC format, with the corpus divided into training and test sets.

Methods

MADEx employs a Recurrent Neural Network (RNN) model using Long Short-Term Memory (LSTM) for clinical Named Entity Recognition (NER), surpassing a baseline Conditional Random Fields (CRFs) model. A modified training strategy for RNN is utilized, which outperforms traditional early stop strategies. For relation extraction, Support Vector Machines (SVMs) and Random Forests are compared for single-sentence and cross-sentence relations. An integrated pipeline is developed to extract entities and relations simultaneously, combining RNN and SVMs.

Results and Conclusion

MADEx achieved the best performance on both the validation and test sets, with an F1-score of 0.8897 on the validation set and 0.8134 on the test set. This outperformed the CRFs model, which scored 0.8377 on the validation set and 0.7250 on the test set.

The study demonstrates the effectiveness of deep learning methods, such as RNNs, for automatically extracting medications, ADEs, and their relations from clinical text, facilitating pharmacovigilance and drug safety surveillance efforts.

Bidirectional LSTM-CRF for Adverse Drug Event Tagging in Electronic Health Records

Paper Link - https://proceedings.mlr.press/v90/wunnava18a/wunnava18a.pdf

Introduction

The paper addresses the critical issue of detecting Adverse Drug Events (ADEs) in Electronic Health Records (EHRs) to enhance pharmacovigilance and patient safety. They highlight the significance of early ADE detection in preventing severe incidents and discuss how Natural Language Processing (NLP) techniques can aid in this task.

System Architecture:
The proposed system architecture comprises a three-layered deep learning model. It includes a bidirectional LSTM for character-level word representation, another bidirectional LSTM for context representation, and a Conditional Random Field (CRF) for final output prediction. This design aims to leverage the strengths of each component in capturing relevant information from the EHR text.

Preprocessing Techniques:
To handle the noise in EHR text, the authors develop rule-based tokenization techniques. These techniques address challenges such as medical and non-medical abbreviations, acronyms, misspelled words, and ambiguous named entities, ensuring better recognition of critical information.

Experimental Setup and Results:
They train and evaluate DLADE using the MADE1.0 dataset, comprising annotated EHR notes from cancer patients. The system achieves high precision, recall, and F1-score, demonstrating its effectiveness in ADE detection. They also compare DLADE's performance with a baseline system, showing significant improvement with dual-level embeddings.

Error Analysis:
The authors conduct an error analysis to understand the sources of misclassification by DLADE. They identify challenges such as entity span across multiple words, mixture of medical and non-medical text, rare occurrence of medical abbreviations, and ambiguity in entity labeling.

Conclusion:
In conclusion, the paper highlights the successful integration of deep learning techniques and rule-based preprocessing methods for ADE detection in EHRs. They emphasize the importance of their findings in advancing pharmacovigilance efforts and ensuring patient safety.

Literature Review for Model Architecture

Paper - Named Entity Recognition with Bidirectional LSTM-CNNs

Paper - Mining Adverse Drug Reactions from Unstructured Mediums at Scale

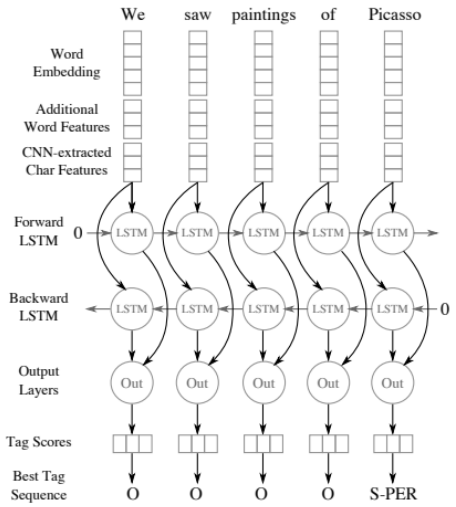


Figure 1: The (unrolled) BLSTM for tagging named entities. Multiple tables look up word-level feature vectors. The CNN (Figure 2) extracts a fixed length feature vector from character-level features. For each word, these vectors are concatenated and fed to the BLSTM network and then to the output layers (Figure 3).

The first paper, "Named Entity Recognition with Bidirectional LSTM-CNNs," introduces a novel neural network architecture for named entity recognition (NER) that eliminates the need for extensive feature engineering and lexicons. By combining bidirectional LSTM and CNN architectures, the model automatically detects word- and character-level features. Additionally, the paper proposes a method for encoding partial lexicon matches in neural networks and demonstrates competitive performance on the CoNLL-2003 dataset and state-of-the-art performance on the OntoNotes 5.0 dataset.

The second paper, "Mining Adverse Drug Reactions from Unstructured Mediums at Scale," addresses the task of mining adverse drug reactions (ADRs) from unstructured data. The authors divide the problem into three main tasks: Document Classification, Named Entity Recognition (NER), and Relation Extraction (RE). They emphasize the importance of NER in identifying entity spans and propose an end-to-end solution by placing all components in a single pipeline. The NER model's output serves as input for the RE model, highlighting the interdependence between the two tasks. The authors experiment with different types of embeddings, such as GLoVe and BERT, to enhance the performance of NER and RE. They find that BERT embeddings, particularly BioBERT, provide more useful information due to their context-awareness and better handling of out-of-vocabulary tokens.

Implementation

Dataset Description

ADE Corpus Charactersitics
The ADE corpus is distributed within two files i.e. DRUG-AE.rel and DRUG-DOSE.rel. DRUG-AE.rel describes relations between drugs and adverse effects whereas DRUG-DOSE.rel describes relations between drugs and dosages. For this project we have only used DRUG-AE.rel.

The format of DRUG-AE.rel is as follows with pipe delimiters:
Column-1: PubMed-ID
Column-2: Sentence
Column-3: Adverse-Effect
Column-4: Begin offset of Adverse-Effect at 'document level'
Column-5: End offset of Adverse-Effect at 'document level'
Column-6: Drug
Column-7: Begin offset of Drug at 'document level'
Column-8: End offset of Drug at 'document level'

Preprocessing Pipeline

Text Processing

- Read the text file into a dataframe for ease.

```
# Read the text file into a DataFrame
df = pd.read_csv('/content/DRUG-AE (copy).txt', sep='|', header=None, name
```

- Remove punctuation marks, convert all capital letters to lowercase, tokenize the sentences by importing necessary libraries from NLTK for text preprocessing, including stopwords removal, tokenization, and lemmatization.
- A function **clean_text** is defined to clean the text by removing punctuation marks, converting uppercase letters to lowercase, and tokenizing the text.
  The code iterates over each row of the DataFrame to process sentences, adverse effects, and drugs separately.
  For each sentence, it identifies and tags adverse effects and drugs using the '**IOB**' tagging scheme.
  - **IOB** tagging, denoting **Inside-Outside-Beginning**, is a tagging scheme crucial for tasks like named entity recognition. It precisely marks tokens in text as beginning, inside, or outside an entity, preserving sequence structure and aiding in handling multiple entities of the same type. Its granularity and simplicity make it widely used for accurately identifying and delineating entities within text.

- After tagging, it appends the processed sentence and corresponding tags to separate lists. It encodes the tags using a label encoding dictionary to convert them into numerical labels.

Embeddings

SpaCy

- Under the hood, spaCy loads these pre-trained embeddings along with its language models. These embeddings are essentially high-dimensional vectors, where each dimension represents some aspect of the word's meaning or usage.
- With spacy we have used "en_core_web_md" that generates an embedding of length 300 for each word.

GloVe

- GloVe, a pre-trained model trained on a vast corpus, to obtain embeddings by directly accessing pre-calculated vectors for each word.
- With GloVe, we have utitlized the "glove.6B.50d" file which generates embeddings of length 50 for words that belong to the corpora of 6B words in the text file.

SpaCy performs better due to its contextualized embeddings, capturing nuanced semantic meanings. GloVe, although widely applicable, often assigns zero vectors to out-of-vocabulary medical terms, limiting its effectiveness in this domain-specific tasks.

To improve embeddings, leveraging pre-trained models specialized in the biomedical domain, such as BioWordVec, could enhance performance by capturing domain-specific nuances.

State-of-the-art models like BERT have shown significant improvements by contextualizing word embeddings, offering potential enhancements for NLP tasks by capturing intricate semantic relationships.

Models

The provided code defines two models: BiLSTM and GRU, both designed for adverse drug event (ADE) detection. Both models use bidirectional recurrent neural networks (RNNs) and consist of an RNN layer followed by a fully connected layer.

BLSTM

- BiLSTM utilizes a bidirectional LSTM layer followed by a fully connected layer.
- It captures bidirectional contextual information, enabling it to understand both past and future dependencies in the input sequence.

```
import torch.nn as nn
# Define BiLSTM model
class BiLSTM(nn.Module):
    def __init__(self, input_dim, hidden_dim, output_dim):
        super(BiLSTM, self).__init__()
        self.lstm = nn.LSTM(input_dim, hidden_dim, bidirectional=True)
        self.fc = nn.Linear(hidden_dim*2, output_dim)  # *2 for bidirectio

    def forward(self, x):
        lstm_out, _ = self.lstm(x.unsqueeze(1))
        out = self.fc(lstm_out.view(len(x), -1))  # Get the last output of
        return torch.sigmoid(out).float()
```

GRU

- GRU follows a similar architecture as BiLSTM but utilizes gated recurrent units (GRUs) instead of LSTMs.
- GRU simplifies the gating mechanism compared to LSTM, potentially leading to faster training and convergence.

```
class GRU(nn.Module):
    def __init__(self, input_dim, hidden_dim, output_dim):
        super(GRU, self).__init__()
        self.gru = nn.GRU(input_dim, hidden_dim, bidirectional=True)
        self.fc = nn.Linear(hidden_dim*2, output_dim)  # *2 for bidirectio

    def forward(self, x):
        gru_out, _ = self.gru(x.unsqueeze(1))
        out = self.fc(gru_out.view(len(x), -1))  # Get the last output of
        return torch.sigmoid(out).float()
```

Despite their architectural differences, both models achieve near-similar scores for ADE detection on medicinal datasets due to their inherent ability to capture sequential dependencies and contextual information from input sequences.

CRF (Conditional Random Field):
CRF is a probabilistic graphical model commonly used for sequence labeling tasks, such as named entity recognition. It models the dependencies between neighboring labels in a sequence, enabling it to consider the global structure of the sequence during prediction. The addition of CRF to the BiLSTM model, creating a BiLSTM-CRF model, can vastly improve ADE

detection scores by incorporating sequential dependencies and jointly optimizing the labeling sequence.

Character-Level Embeddings:

Incorporating character-level embeddings alongside word-level embeddings can enhance the model's ability to capture morphological and orthographic features of words, especially in scenarios with out-of-vocabulary words or domain-specific terms.

BiLSTM-CRF Model:
By combining the BiLSTM architecture with CRF, the model can leverage both local contextual information (captured by BiLSTM) and global sequence dependencies (modeled by CRF). This integration allows the model to make more informed predictions by considering the entire sequence of labels jointly, leading to improved performance in ADE detection tasks.

This improvements are stated in Named Entity Recognition with Bidirectional LSTM-CNNs and adopted for an ADE task in Mining Adverse Drug Reactions from Unstructured Mediums at Scale.

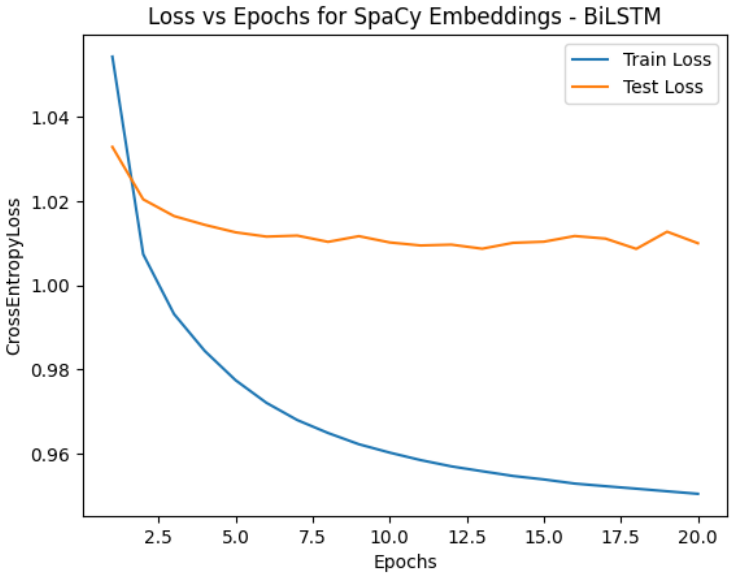
Performance Evaluation

SpaCy Embeddings

Comparing Model Performance

Model	Accuracy	Precision	Recall	F-1 Score	ROC-AUC
GRU	0.8823	0.7275	0.5503	0.6098	0.7413
BiLSTM	0.8801	0.7174	0.5531	0.6033	0.7473

Tag-Wise Metric Scores for BLSTM Model

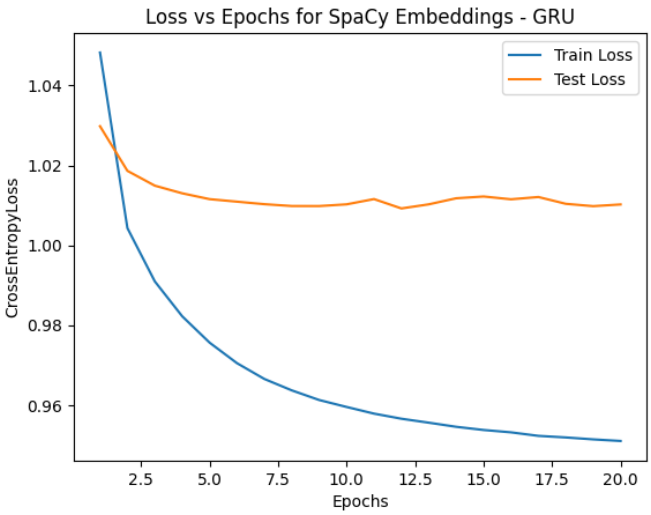


Tags	Precision	Recall	F1-Score
O	0.9101	0.9654	0.9369
B-AE	0.6110	0.4358	0.5087
I-AE	0.6276	0.3354	0.4372
B-Drug	0.6903	0.7354	0.7121
I-Drug	0.7479	0.2937	0.4218

```
Number of sentences with accuracy less than 50%: 7
Index: 5732
Sentence: ['severe', 'abdominal', 'pain', 'in', 'low', 'dosage', 'clofazir
True Tags: ['O', 'B-AE', 'I-AE', 'O', 'O', 'O', 'B-Drug']
Predicted Tags: ['B-AE', 'I-AE', 'I-AE', 'O', 'O', 'B-Drug', 'O']
Accuracy: 42.857142857142854

Index: 5775
Sentence: ['interference', 'with', 'the', 'cortisol', 'axis', 'by', 'the',
True Tags: ['B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'O', 'I-AE', 'O', 'O',
Predicted Tags: ['B-Drug', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-Drug
Accuracy: 40.0

Index: 5828
Sentence: ['cholesterol', 'crystal', 'embolization', 'associated', 'renal'
True Tags: ['B-AE', 'I-AE', 'I-AE', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-
Predicted Tags: ['O', 'O', 'O', 'O', 'B-AE', 'I-AE', 'O', 'O', 'O', 'B-Dru
Accuracy: 28.57142857142857
```

```
Index: 5960
Sentence: ['does', 'acyclovir', 'increase', 'serum', 'lithium', 'levels']
True Tags: ['O', 'B-Drug', 'B-AE', 'I-AE', 'I-AE', 'I-AE']
Predicted Tags: ['O', 'B-Drug', 'O', 'O', 'I-Drug', 'B-AE']
Accuracy: 33.33333333333333

Index: 6323
Sentence: ['relapse', 'in', 'the', 'external', 'auditory', 'canal', 'of',
True Tags: ['B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE
Predicted Tags: ['B-AE', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O',
Accuracy: 47.05882352941176

Index: 6517
Sentence: ['clostridium', 'difficile', 'toxin', 'induced', 'colitis', 'aft
True Tags: ['B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'O', 'O', 'O', 'B-Drug
Predicted Tags: ['O', 'O', 'O', 'O', 'B-AE', 'O', 'O', 'O', 'O', 'B-Drug',
Accuracy: 41.66666666666667

Index: 6738
Sentence: ['we', 'suspect', 'that', 'nefazodone', 'inhibits', 'metabolism
True Tags: ['O', 'O', 'O', 'B-Drug', 'B-AE', 'I-AE', 'I-AE', 'I-AE']
Predicted Tags: ['O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-Drug']
Accuracy: 37.5
```

Tag-Wise Metric Scores for GRU Model



Tags	Precision	Recall	F1-Score
O	0.9076	0.9695	0.9375
B-AE	0.6469	0.4366	0.5213
I-AE	0.6332	0.3364	0.4394
B-Drug	0.7357	0.6953	0.7149
I-Drug	0.7143	0.3135	0.4358

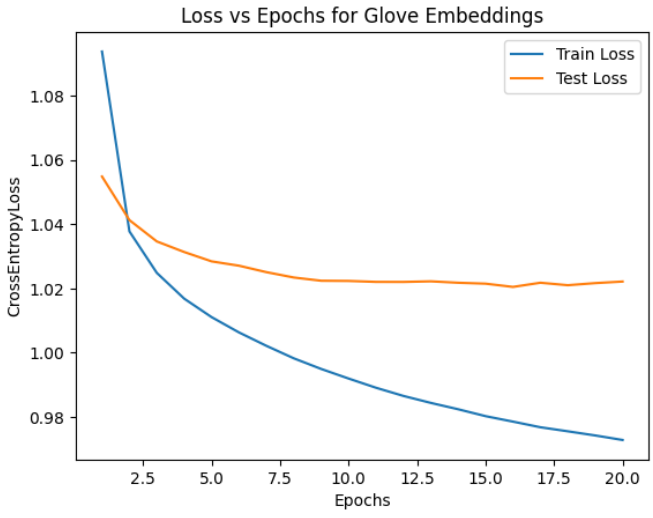
```
Number of sentences with accuracy less than 50%: 8
Index: 5614
Sentence: ['celiprolol', 'pneumonitis']
True Tags: ['B-Drug', 'B-AE']
Predicted Tags: ['B-AE', 'I-AE']
Accuracy: 0.0

Index: 5828
Sentence: ['cholesterol', 'crystal', 'embolization', 'associated', 'renal'
True Tags: ['B-AE', 'I-AE', 'I-AE', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-
Predicted Tags: ['O', 'O', 'O', 'O', 'B-AE', 'I-AE', 'O', 'O', 'O', 'O',
Accuracy: 35.714285714285715

Index: 5960
Sentence: ['does', 'acyclovir', 'increase', 'serum', 'lithium', 'levels']
True Tags: ['O', 'B-Drug', 'B-AE', 'I-AE', 'I-AE', 'I-AE']
Predicted Tags: ['O', 'B-Drug', 'O', 'O', 'O', 'O']
Accuracy: 33.33333333333333

Index: 5967
Sentence: ['when', 'measured', 'the', 'serum', 'lithium', 'level', 'had',
True Tags: ['O', 'O', 'O', 'B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'O', 'O
Predicted Tags: ['O', 'O', 'O', 'O', 'B-Drug', 'O', 'O', 'O', 'I-AE', 'I-A
Accuracy: 46.15384615384615

Index: 5968
Sentence: ['when', 'measured', 'the', 'serum', 'lithium', 'level', 'had',
True Tags: ['O', 'O', 'O', 'B-AE', 'B-Drug', 'I-AE', 'I-AE', 'I-AE', 'O',
Predicted Tags: ['O', 'O', 'O', 'O', 'B-Drug', 'O', 'O', 'O', 'I-AE', 'I-A
Accuracy: 46.15384615384615

Index: 6323
Sentence: ['relapse', 'in', 'the', 'external', 'auditory', 'canal', 'of',
```

```
True Tags: ['B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE'
Predicted Tags: ['B-AE', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O',
Accuracy: 47.05882352941176

Index: 6517
Sentence: ['clostridium', 'difficile', 'toxin', 'induced', 'colitis', 'aft
True Tags: ['B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'O', 'O', 'O', 'B-Drug
Predicted Tags: ['O', 'O', 'O', 'O', 'B-AE', 'O', 'O', 'O', 'O', 'B-Drug',
Accuracy: 41.66666666666667

Index: 6738
Sentence: ['we', 'suspect', 'that', 'nefazodone', 'inhibits', 'metabolism'
True Tags: ['O', 'O', 'O', 'B-Drug', 'B-AE', 'I-AE', 'I-AE', 'I-AE']
Predicted Tags: ['O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-Drug']
Accuracy: 37.5
```

Glove Embeddings

Comparing Embeddings Performance BLSTM model

Embeddings	Accuracy	Precision	Recall	F-1 Score	ROC-AUC
Glove	0.8702	0.5520	0.4885	0.5095	0.7258
SpaCy	0.8801	0.7174	0.5531	0.6033	0.7473



Tags	Precision	Recall	F1-Score
O	0.9028	0.9636	0.9322

Tags	Precision	Recall	F1-Score
B-AE	0.5971	0.4181	0.4919
I-AE	0.5893	0.3318	0.4245
B-Drug	0.6710	0.7287	0.6986
I-Drug	0.0000	0.0000	0.0000

```
Number of sentences with accuracy less than 50%: 5
Index: 5742
Sentence: ['sudden', 'death', 'in', 'an', 'infant', 'from', 'methemoglobi
True Tags: ['B-AE', 'I-AE', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-Drug',
Predicted Tags: ['O', 'O', 'O', 'O', 'O', 'O', 'B-Drug', 'O', 'O', 'O', 'O
Accuracy: 42.857142857142854

Index: 6000
Sentence: ['systemic', 'corticosteroids', 'in', 'the', 'phenytoin', 'hype
True Tags: ['O', 'O', 'O', 'O', 'B-Drug', 'I-AE', 'I-AE']
Predicted Tags: ['B-AE', 'I-AE', 'O', 'O', 'B-Drug', 'O', 'O']
Accuracy: 42.857142857142854

Index: 6323
Sentence: ['relapse', 'in', 'the', 'external', 'auditory', 'canal', 'of',
True Tags: ['B-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE', 'I-AE
Predicted Tags: ['O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'I-AE', 'O',
Accuracy: 23.52941176470588

Index: 6631
Sentence: ['a', 'drug', 'interaction', 'between', 'zafirlukast', 'and', 't
True Tags: ['O', 'B-AE', 'I-AE', 'O', 'B-Drug', 'O', 'O']
Predicted Tags: ['O', 'O', 'O', 'O', 'O', 'O', 'B-Drug']
Accuracy: 42.857142857142854

Index: 6738
Sentence: ['we', 'suspect', 'that', 'nefazodone', 'inhibits', 'metabolism
True Tags: ['O', 'O', 'O', 'B-Drug', 'B-AE', 'I-AE', 'I-AE', 'I-AE']
Predicted Tags: ['O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-Drug']
Accuracy: 37.5
```

Recommendations for Future Research

For future research in adverse drug event (ADE) detection, several recommendations can be made to advance the field:

- **Integration of Transformers and Large Language Models:**
  Explore the use of transformer-based architectures, such as BERT, GPT, or XLNet, which have demonstrated exceptional performance in various natural language processing tasks. These models can capture complex contextual relationships and semantic nuances, potentially improving ADE detection accuracy.
- **Development of Gold Standard Datasets:**
  Create standardized datasets with uniform annotation guidelines, ensuring consistency in labeling ADEs and their associated treatments. Emphasize the representation of ADE treatments as relations between drug and symptom entities rather than standalone entities, facilitating more accurate and comprehensive detection.
- **Refinement of Evaluation Metrics:**
  Define and standardize evaluation metrics tailored specifically for ADE detection tasks. Metrics should account for the nuances of detecting ADEs and their treatments, considering factors like entity span, relation extraction accuracy, and semantic relevance. Ensure transparency and reproducibility by providing clear guidelines on metric computation.
- **Integration of Domain-Specific Knowledge:**
  Incorporate domain-specific knowledge, such as pharmacological databases and medical ontologies, into model training and evaluation.

  Leveraging domain expertise can enhance model understanding of drug-symptom relationships and improve detection performance in real-world healthcare settings.
- **Collaboration with Healthcare Professionals:**
  Collaboration between researchers and healthcare professionals to ensure the practical relevance and clinical applicability of ADE detection systems.

  Engage clinicians in dataset curation, model development, and evaluation processes to address real-world challenges and validate system performance in clinical settings.

By addressing these recommendations, future research endeavors in ADE detection can advance the state-of-the-art, leading to more accurate, efficient, and clinically impactful solutions for medication safety and pharmacovigilance.