WELCOME ALL !!

BioLx

A Biomedical Named Entity Recognition App

10^{th} May 2024, Friday

OUR GUIDE

Ms. Anupama M L

Asst Prof of CSE Department

TEAM MEMBERS

Akhilesh M S

S8 CSE Student

Varun Sasikumar

S8 CSE Student

Rintu Sam

S8 CSE Student

Rishal Mohammed V A

S8 CSE Student

TABLE OF CONTENTS

01 Introduction

02 Existing System

03 Objectives

04 Literature Survey

TABLE OF CONTENTS

05 System Diagrams

06 Produced Output

07 Conclusion

08 References

01
Introduction

WHAT IS BioNER ?

- ✓ BioNER is a subfield of natural language processing (NLP) that aims to identify and classify biomedical entities in text.
- ✓ Biomedical entities are specific types of words or phrases that refer to biomedical concepts, such as genes, proteins, diseases, drugs, and chemicals.

WHY IS BioNER IMPORTANT ?

- ✓ BioNER is important because it enables the extraction of biomedical knowledge from unstructured text, which can be used to support a wide range of applications, such as drug discovery, clinical research, and public health surveillance.

FOR EXAMPLE

Biologically Active Substance

Drug

- ✓ The benefits of taking cholesterol lowering statin drugs outweigh the risks even among people who are likely to develop diabetes.

Disorder

Organic chemical

Enzyme

- ✓ Acute exposure to resveratrol inhibits AMPK activity in human skeletal muscle cells.

Cell

02
Existing System

EXISTING SYSTEM

- ✓ Some existing systems employ rule-based approaches for entity recognition, which may lack flexibility in handling complex medical terminology and variations.
- ✓ The existing system may utilize traditional NLP techniques such as tokenization and part-of-speech tagging, which may not adequately capture the nuances of biomedical text.

DISADVANTAGES

- ✓ Inaccuracy
- ✓ Scalability Challenges
- ✓ Expert Dependency
- ✓ Lack of Real-Time Processing
- ✓ Limited Adaptability

PROBLEM DESCRIPTION

- ✓ Manual Effort
- ✓ Limited Handling
- ✓ Language Complexity

PROPOSED SYSTEM

- ✓ The development of a mobile application for Biomedical Named Entity Recognition (BioNER) represents a significant advancement in the domain of biomedical research and healthcare.
- ✓ This specialized mobile app harnesses a state-of-the-art BioNER model hosted on a cloud server, offering a range of features designed to enhance efficiency, accuracy, understanding, and accessibility for biomedical professionals and researchers.

03
Objectives

OBJECTIVES

- ✓ Develop a biomedical entity recognition system using CRF and BERT models.
- ✓ Create an intuitive Flutter mobile app for users to input biomedical text.
- ✓ Implement RESTful APIs with Flask to facilitate communication between the frontend and backend.
- ✓ Integrate CRF and BERT models to enhance entity recognition accuracy.
- ✓ Document the development process and create a reliable, user-friendly application.

04
Literature Survey

PAPER 01

Name	Authors	Abstract	Proposed System
Biomedical Named Entity Recognition via Reference-Set Augmented Bootstrapping	✓ Joel Mathew ✓ Shobeir Fakhraei ✓ Jose Luis Ambite	✓ Train a neural NER model on a small seed of labeled data. ✓ Use a reference set of entity names to identify entity mentions with high precision but low recall on an unlabeled corpus. ✓ Use the NER model to assign weak labels to the corpus and retrain the model iteratively on the augmented training set.	✓ Create an augmented training set using a small labeled seed set and a large unlabeled corpus. ✓ Use automated labeling with a reference set to weakly label the unlabeled corpus. ✓ Apply an iterative label refinement process.

PAPER 02

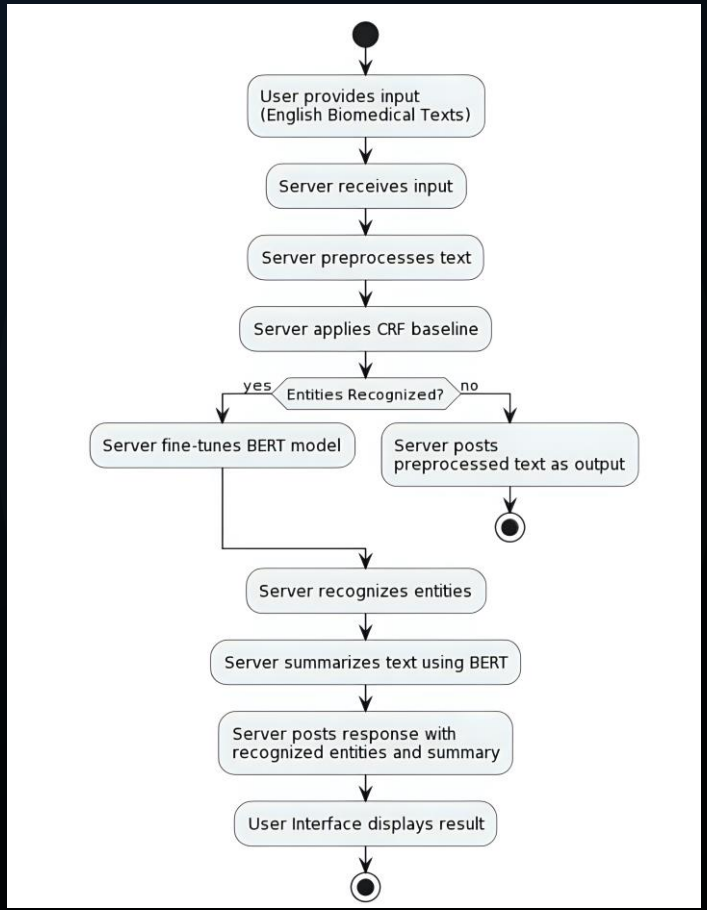
Name	Authors	Abstract	Proposed System
Bio-NER: Biomedical Named Entity Recognition using Rule-Based and Statistical Learners	✓ Pir Dino Soomro ✓ Sanotsh Kumar ✓ Banbhrani ✓ Arsalan Ali Shaikh ✓ Hans Raj	✓ The authors propose a hybrid approach for biomedical NER, employing rule-based and statistical methods. ✓ They introduce a novel method for disease NER that surpasses existing methods. ✓ The effectiveness of their approach is validated using standard evaluation metrics.	✓ The proposed hybrid approach for biomedical NER combines machine learning, rule-based, and decision tree methods. ✓ It uses simple features like affixes and context, as well as more complex features like orthographic, POS tags, and N-grams.

PAPER 03

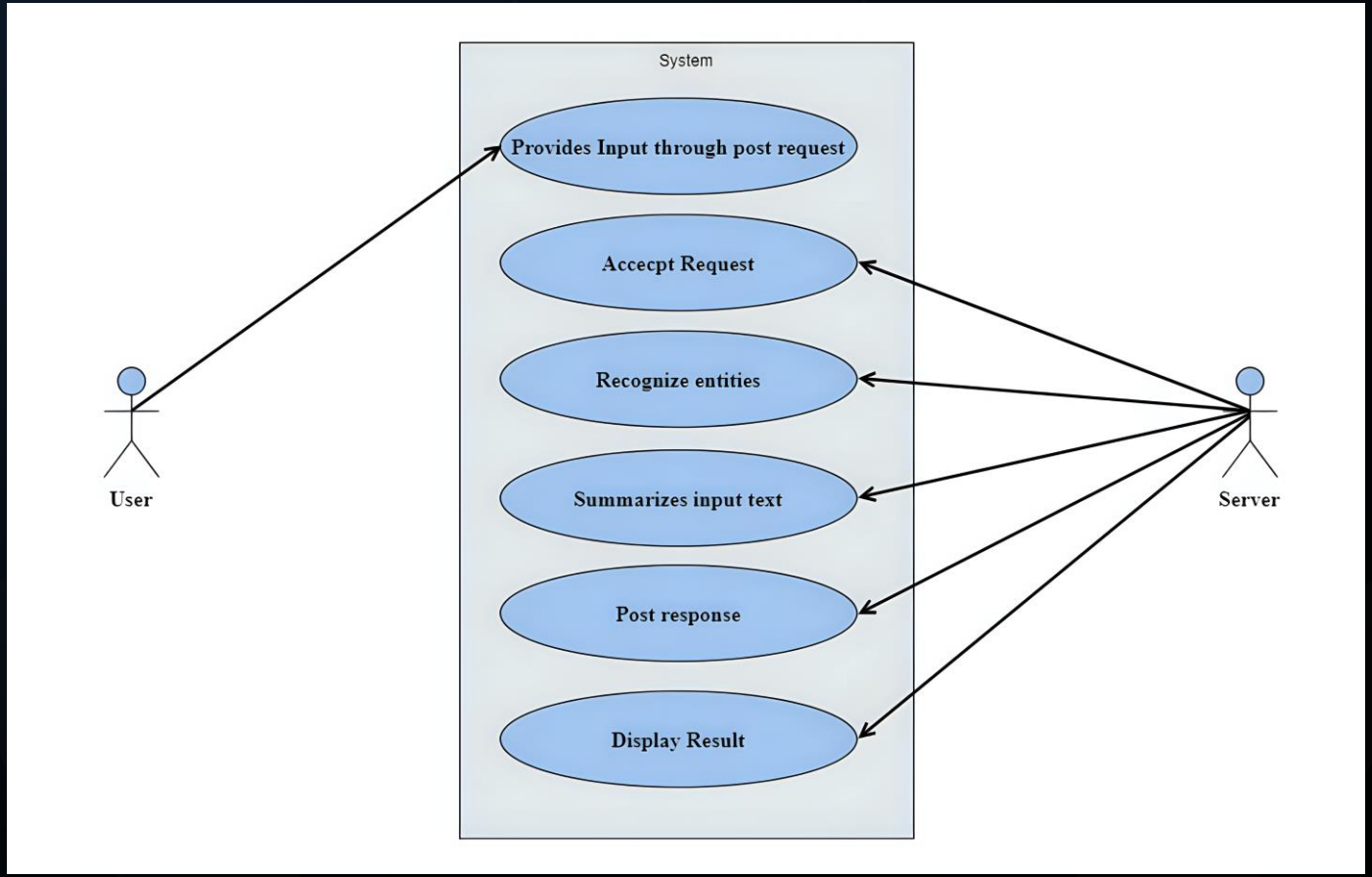
Name	Authors	Abstract	Proposed System
Unsupervised biomedical named entity recognition: Experiments with clinical and biological texts	✓ Shaodian Zhang ✓ Noémie Elhadad	✓ Stepwise approach to entity boundary detection and entity type classification. ✓ Uses noun phrase chunking and distributional semantics principles, without relying on handcrafted rules or annotated data. ✓ Outperforms a baseline dictionary match approach on clinical and biological texts.	✓ Extracts candidate entities from free text using a noun phrase chunker and a filter based on inverse document frequency. ✓ Classifies candidate entities into categories of interest using principles from distributional semantics. ✓ Does not rely on handcrafted rules or annotated data.

05
System Diagrams

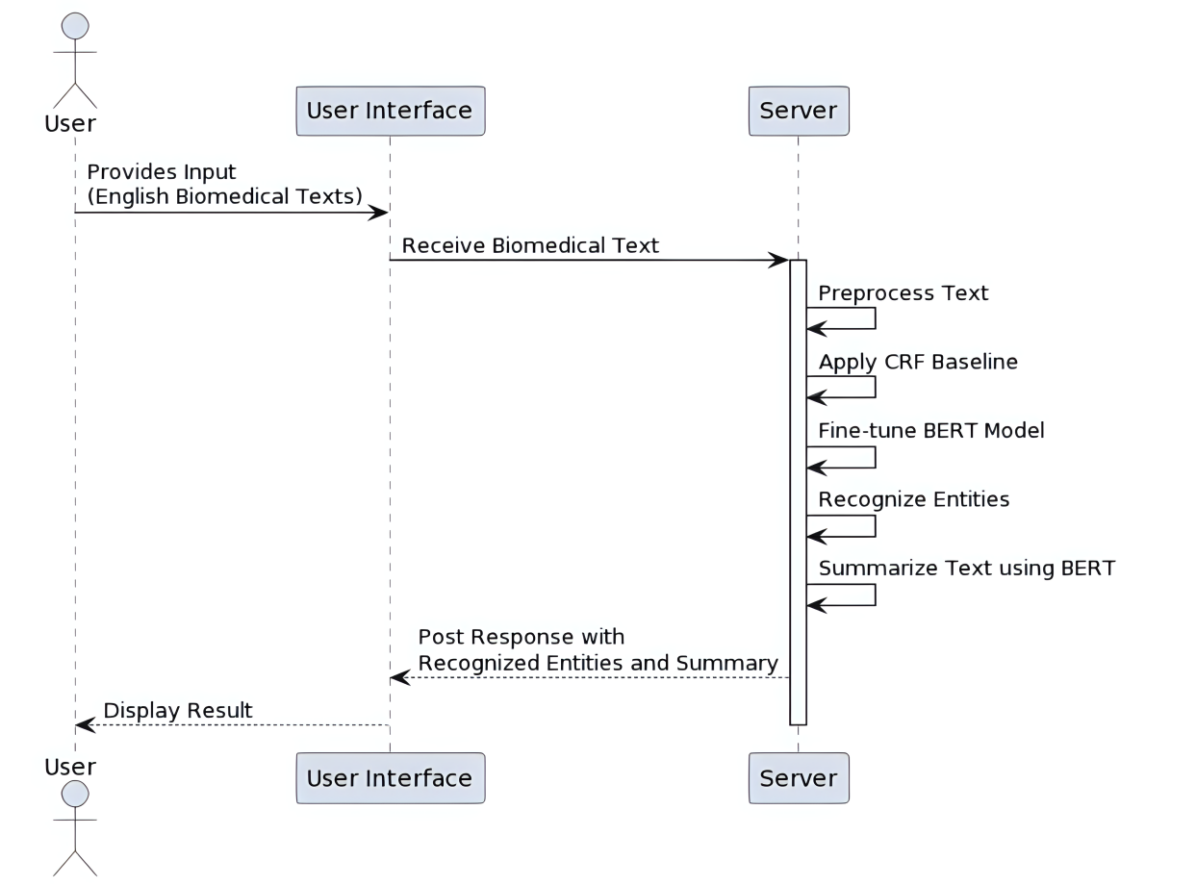
ACTIVITY DIAGRAM



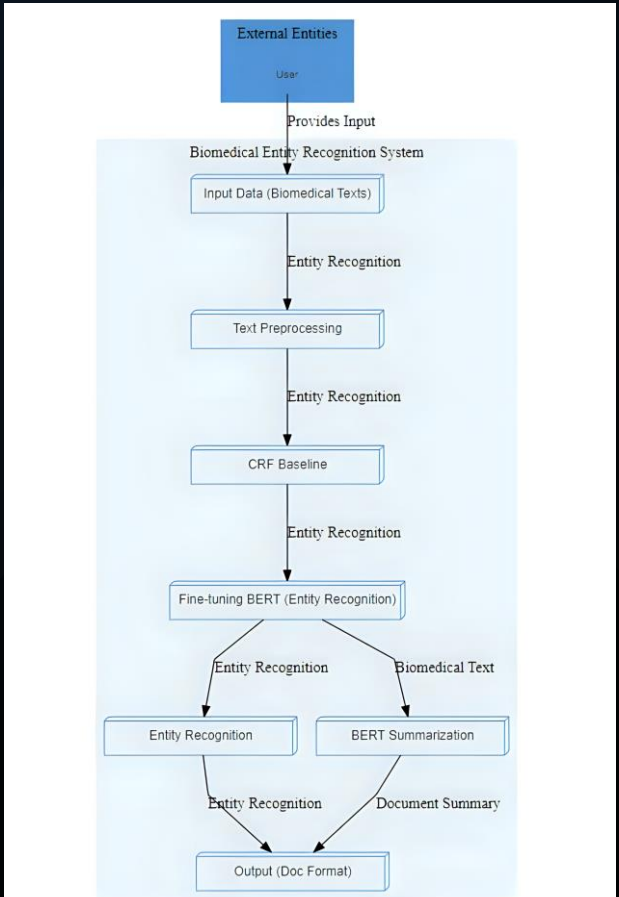
USE CASE DIAGRAM



SEQUENCE DIAGRAM

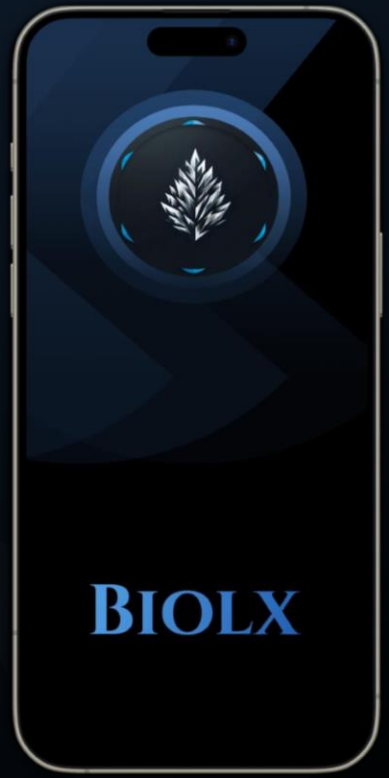


DATAFLOW DIAGRAM

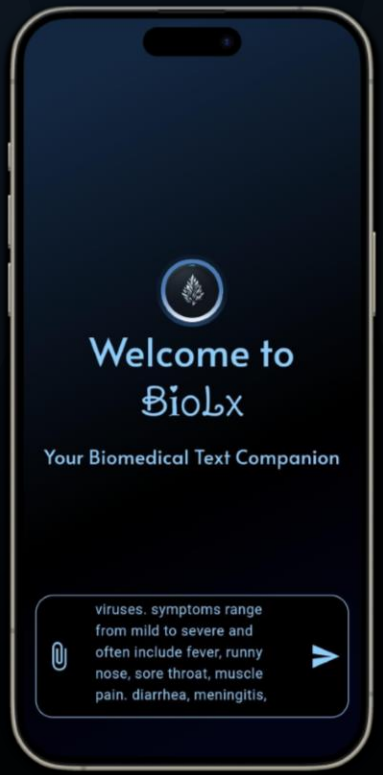


06
Produced Output

SPLASH SCREEN

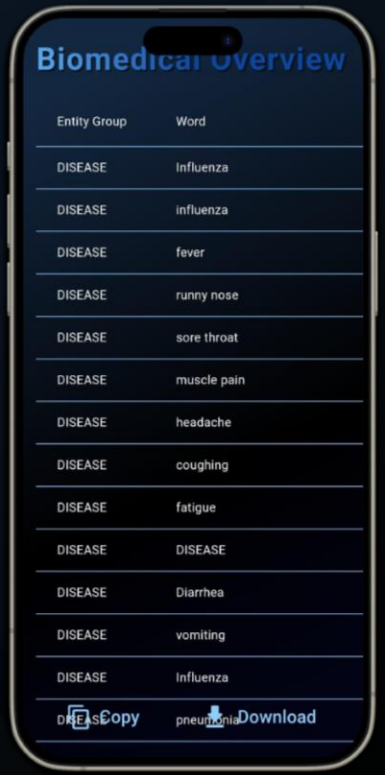


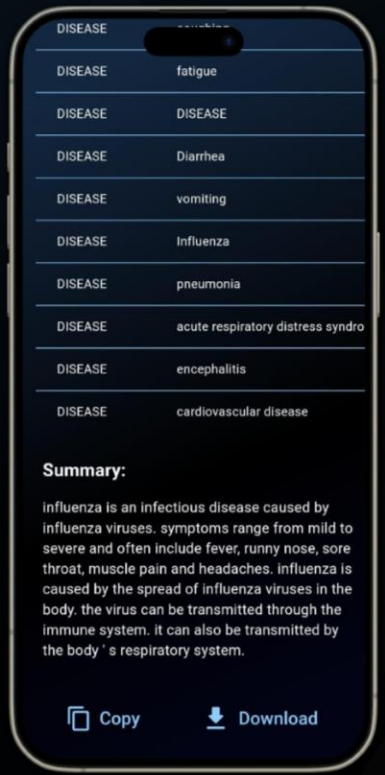
HOME SCREEN



LOADING SCREEN

DISPLAY SCREEN





07
Conclusion

CONCLUSION

- ✓ The biomedical entity recognition project combines CRF and BERT models for accurate identification of entities in biomedical texts.
- ✓ Flask as the primary backend framework, and Python as the programming language supporting the overall backend functionality.
- ✓ The team work between CRF and BERT captures both local and global context, enhancing entity recognition.
- ✓ Avoiding a database simplifies the system, utilizing in-memory knowledge and alternative storage effectively.
- ✓ This system offers a practical, user-friendly solution for biomedical text processing, with potential for scalability and knowledge base refinement in future updates.

08
References

REFERENCES

- ✓ Gamal Crichton, Sampo Pyysalo, Billy Chiu, and Anna Korhonen. 2017. A neural network multi-task learning approach to biomedical named entity recognition. BMC bioinformatics, 18(1):368.
- ✓ Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805
- ✓ Diederik P Kingma and Jimmy Ba. 2014. Adam: Amethod for stochastic optimization. arXiv preprint arXiv:1412.6980.

THANK YOU !!