MEDINFO23

8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Extracting Symptoms of Agitation in Dementia from Free-text Nursing Notes using Advanced Natural Language Processing

Dinithi Vithanage

PhD Student

University of Wollongong, Australia

AIDH IMIA

Outline

- Introduction
- Problem-Solution
- Methodology
- Results
- Limitations and Future Studies

Introduction

- Nursing staff record observations about older people
- Records are in Free-text nursing notes
- Free-text nursing notes-Unstructured Notes
- Contain: older people care needs, disease symptoms, frequency of symptom occurrence, nursing actions, etc
- Develop a technique to uncover important data from these notes

Introduction

Unstructured **NLP** **Structured**

This resident's agitation symptoms are screaming, shouting, hitting and wandering. Interventions are verbal prompting, reassurance and reorientation.

NLU →

```
<Agitation Symptoms>
   <Symptom> Screaming </Symptom>
   <Symptom> Shouting </Symptom>
   <Symptom> Hitting </Symptom>
   <Symptom> Wandering </Symptom>
</Agitation Symptoms>
<Interventions>
    <Intervention> verbal prompting</Intervention>
    <Intervention> reassurance </Intervention>
    <Intervention> reorientation </Intervention>
</Interventions>
```

Introduction

- Aim: Developed and evaluated a deep learning and transfer learning-based named entity recognition (NER) model for extracting symptoms of agitation in dementia from the nursing notes.
- Apply Model development life cycle

Problem Statement

Sentence 1: Resident lacks insight into her care needs and personal safety and refuses care such as personal hygiene, changing continence aid, dressing and undressing, washing, and sending laundry most of the time and just wanders around lounge room and her room. She is always shouting.

- Agitation Symptom
- Agitation Symptom
- Agitation Symptoms

Design

- Literature Review
  - Dictionary and rule-based approach
  - Feature-based approach
  - Neural network-based approach

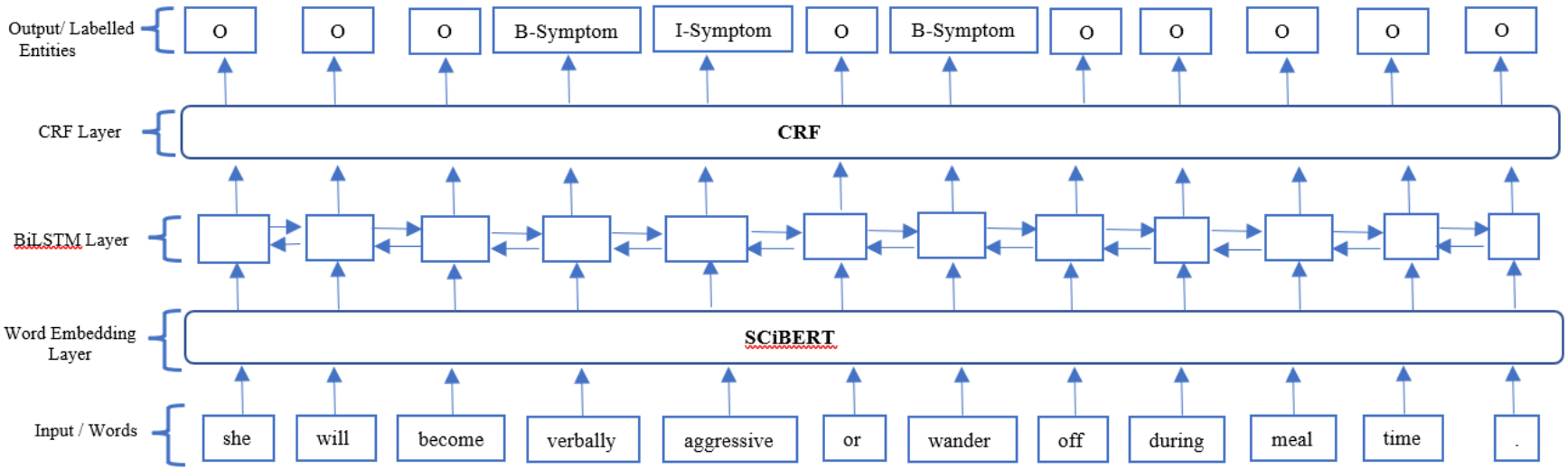
Data Collection

- Granted ethics approval.
- Nursing notes, were collected from 40 RACFs in New South Wales, Australia, in 2019 and 2020.
- The dataset contains 1,000 labelled nursing notes.

Data Collection

- We adopted the Begin-Inside-Outside (BIO) tagging system to label the entities.
- A token is labelled as B-label if it is the first of a named entity, I-label if it is inside a named entity but not the first, and O-label otherwise (e.g., she (O) is (O) verbally (B-Symptom) Aggressive (I-Symptom)).

NER Model Architecture



NER Model Architecture

- Word Embedding-SciBERT- Convert text into numbers
- BiLSTM-Deep learning- Relationship Capturing
  - Task : Capture relationships between the feature vectors generated from the word embedding layer.
  - The BiLSTM model processes each sequence of feature vectors from two directions, both left-to-right and right-to-left.

NER Model Architecture

- Conditional Random Field-Deep Learning-Find the suitable label
  - The optimal predicted label, by learning the order dependency and the transition probability between labels from the output of the relationship capture layer. The CRF model computes the joint.

Results

Number of Notes	NER Model	F1 score	Accuracy
1000 Notes	Clinical BioBERT + BiLSTM + CRF	75%	78%
	Clinical BioBERT	68%	70%
2000 Notes	Clinical BioBERT + BiLSTM + CRF	87%	89%
	SciBERT+ BiLSTM + CRF	90%	93%

Limitations & Future Studies

- Limitation: Model Accuracy.
- Improve the model accuracy.
- Extend the NER model to extract the causal factors and nursing actions to prevent agitation in dementia.
- Combine the variables and values extracted from the free-text nursing notes with the structured data to conduct further machine learning to uncover the patterns or trends that will help nursing care decision-making.


THANK YOU