

Introduction to the first Workshop on Context-aware NLP in eHealth

Mohammed Hasanuzzaman¹, Jyoti Prakash Singh², Gaël Dias³,
Cristina Soguero-Ruiz⁴, Terje Solvoll⁵ and Phuong Dinh Ngo⁵

¹ADAPT Centre, MTU, Ireland

²NIT Patna, India

³University of Caen Normandy, GREYC CNRS, France

⁴Universidad Rey Juan carlos, Spain

⁵Norwegian Centre for eHealth Research, Norway

Abstract

Artificial intelligence (AI) consists of a variety of technologies that combine data, algorithms and computer technology. As advances in computing power and the availability of data have increased, so too have expectations regarding AI's possibilities in various areas of society. It is currently recognised that as much as 30% of the world's stored data is produced by the healthcare sector.¹ These data include biobanks, medical health records and observational studies. However, this 'data-rich' sector does not currently explore and analyse data to the full potential. Machine Learning (ML) and automatic text processing technologies can support the operations and efficiency of our healthcare sector in numerous ways: a) by improving the efficiencies and operations within clinical laboratories, thus improving patient management and clinical outcomes; b) by offering solutions for more efficient data management to develop better screening mechanisms to analyse clinical data and diagnose disease; c) by performing complex simulations that can assess the effectiveness of new medicines based on clinical research data. WNLPe-Health 2022 - the first Workshop on Context-aware NLP in eHealth was held at IIT Delhi, India on December 18th, 2022 in conjunction with 19th International Conference on Natural Language Processing (ICON 2022).

Keywords

e-health, NLP, context-awareness

1. Introduction

There has as yet been very little research into the use of AI applications for patients or healthy citizens, although the technology may allow for a much more individual and person-centred approach. For example, by combining ubiquitous data with user-generated and publicly available data, AI algorithms can guide and inform citizens about risk modifying behaviors in an appropriate context. The goal of this workshop is to provide a unique platform to bring together researchers and practitioners in healthcare informatics working with health-related data especially textual data, and facilitate close interaction among students, scholars, and industry professionals on eHealth language processing tasks. In particular, we are interested in works

¹<https://data.oecd.org/>

The first Workshop on Context-aware NLP in eHealth, 18 December 2022, Delhi, India

✉ mohammed.hasanuzzaman@adaptcentre.ie (M. Hasanuzzaman)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

that advance state-of-the-art NLP and ML techniques for eHealth domains by incorporating more contextual knowledge in order to make models explainable, trustable and robust in changing situations. The main topics of the proposed workshop are: Modelling of healthcare text in classical NLP tasks (tagging, chunking, parsing, entity identification, relation extraction, coreference, summarization, etc.) for under-resourced languages; Person-centred NLP applications for eHealth including early risk prediction; Algorithm for Context Data reasoning; Context sensitive recommendations to individual citizens and patients; Integration of structured and unstructured resources for health applications; Domain adaptation techniques for clinical data; Medical terminologies and ontologies; Interpretability and analysis of NLP models for healthcare applications; Processing clinical literature and trial reports; Bayesian modelling and feature selection techniques for high-dimensional healthcare data; and Multimodal learning for decision support systems: Ubiquitous data, public databases, user generated content (in combination with wearable sensor technology).

2. Program

WNLPe-Health 2022 was a half-day workshop. The program chairs accepted eight papers; however, one paper was withdrawn. The papers are briefly described below. The paper entitled “MINDS: A Multi-label Emotion and Sentiment Classification Dataset Related to COVID-19” introduces a dataset specifically designed to capture emotions and sentiments in user-generated content during the COVID-19 pandemic. The authors of the paper “Depression detection in Thai language posts based on attentive network models” develop a framework for early detection of individuals at risk of depression from Thai social media platform. The paper “Interestingness from COVID-19 Data: Ontology and Transformer-Based Methods” identifies interesting patterns using ontology-based mining techniques and process them with BioClinicalBERT and CovidBERT. The paper “CME2 Net: Contextual Medical Event Extraction Network for clinical notes” presents an end-to-end model for automatic extracting and classifying the medication change events from a clinical note. The authors of paper “Stress Detection System using Natural Language Processing and Machine Learning Techniques” developed models for stress detection from social media data. The paper “Automatic Annotation of Training Data for Deep Learning Based De-identification of Narrative Clinical Text” utilized dictionaries constructed from publicly available lists of identifiers to automatically annotate a training dataset for a named entity recognition model to de-identify names, streets, and locations in Danish narrative clinical text. Finally, “Recognizing Question Entailment in Consumer Health Using a Query Formulation Approach” develops a query-Based system for recognizing question entailment.

3. Acknowledgements

We would like to thank the program committee members for their effort in reviewing and recommending papers.