



Introduction: Diabetes Management in a School Setting

Type 1 Diabetes Mellitus (T1DM) is an autoimmune disease which requires continuous monitoring and management. For students with T1DM, optimally managing T1DM is crucial to avoid severe complications and to improve their quality of life at school. To ensure the safety and proper management of T1DM, school personnel and students should be equipped with relevant medical knowledge which requires access to readable and understandable T1DM content. Existing techniques to provide diabetes education include the use of resource guides, trainings, computer-based education programs, medical websites¹, etc. However, the medical content provided through these means can be unreliable, complex and confusing. This research aims to design and develop new methods using Natural Language Processing and Machine Learning to improve the readability and understandability of gamut of T1DM information through simplification² of medical text.

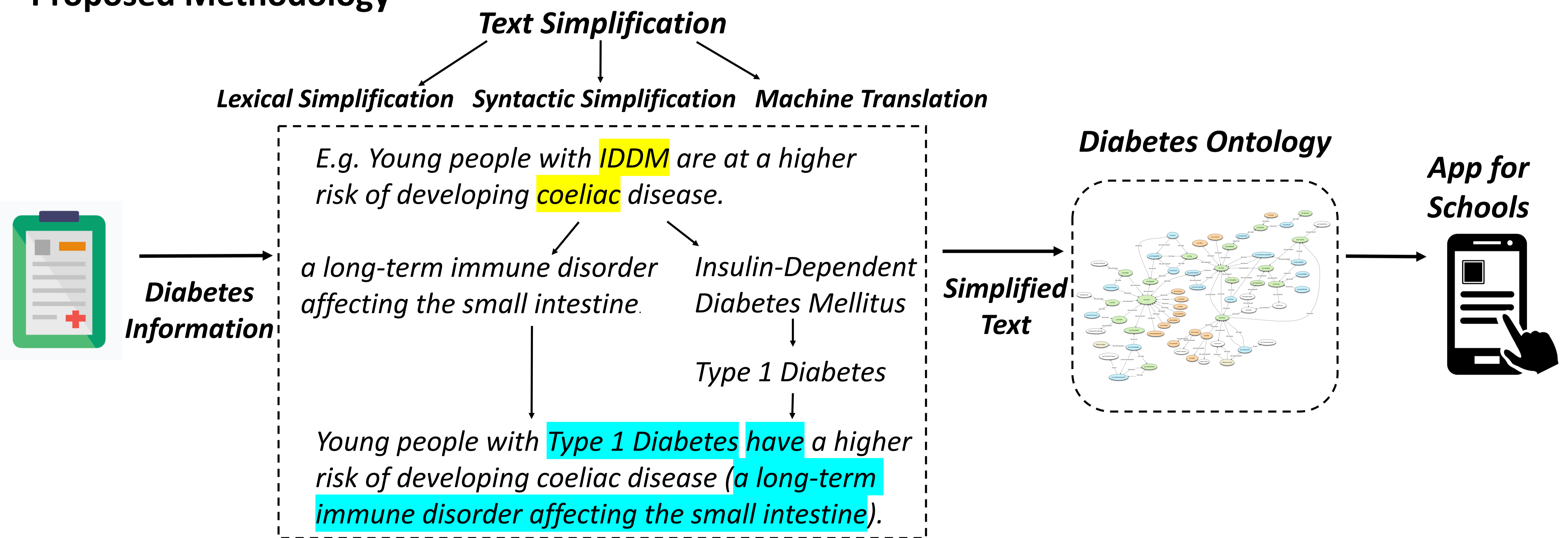
Background

- Understanding medical content with complex words, complex sentence structures and identifying relevant, reliable medical content in an emergency is difficult for people with limited medical background.
- Improving diabetes education can result in better management in a school setting.

Research Objectives

- Medical text simplification to improve the readability and understandability of medical content.
- Ontology learning³ to develop a comprehensive diabetes knowledgebase to provide access to diabetes content.

Proposed Methodology



Current Experiments

Complex Word Identification (first step in lexical simplification)

Data

- Wikipedia
- Wikipedia News
- News

Features

word length, frequency, no. of hypernyms, hyponyms, synonyms, syllables, dependency, relations

→ **Logistic Regression**

Precision:0.84 Recall:0.71 F-score:0.77

Discussion

With the advances in NLP and ML, Machine Translation (MT) techniques have been effectively used for text simplification tasks. However, one of the major constraints of using MT for text simplification is the unavailability of parallel corpora for sequence to sequence modelling. Thus, both lexical and syntactic simplification techniques along with MT should be explored in the medical domain.

References:

[1] Shah, V.; Garg, S.; *Managing Diabetes in the Digital Age. Clinical Diabetes and Endocrinology* 2015.

[2] Paetzold, G.; Specia, L.; A survey of Lexical Simplification. *Journal of Artificial Intelligence Research* 2017.

[3] Mahmoud, N.; Elbeh, H.; Ontology Learning Based on Word Embeddings for Text Big Data Extraction. *14th International Computer Engineering Conference* 2018.

Contact:

Sandaru Seneviratne

sandaru.seneviratne@anu.edu.au

+61 414 986 665

Hanna Neuman Building,

Research School of Computer Science,

The Australian National University

