

# Paper: *Contextual Word Representations: A Contextual Introduction*

**Razwan Ahmed Tanvir**

Date: 04/12/2022

Quote	<i>“...After reading this document, you should have a general understanding of word vectors (also known as word embeddings): why they exist, what problems they solve, where they come from, how they have changed over time, and what some of the open questions about them are. ”</i>
Overview	<p>The paper in the discussion discusses about one of the most basic concepts of natural language processing which is Word Vectors. Inherently, computers cannot process raw text data. Hence, numerical representation of textual data is inevitable to work with any natural language processing task. In the initial stage of natural language processing, the words were mapped to integers and every word in a corpus would get a unique integer representation. The mentioned benefits of this ‘integerized’ approach are- every word can be stored in same amount of memory, the words can be accessed in constant time by using array-like data structure. This integer-based word representation is known as discrete representation. However, this integer-based representation is no longer used and currently vector representation of words are widely used in NLP systems. These vectors are generated from a certain corpus by numerous of ways like count-based stats and advanced statistical methods. One-hot-encoding is one such approach to generate word vectors. Furthermore, to get the context of any word in a corpus, full distribution of context is obtained from nearby words or the position of a word occurring in the data set. In this way, the words which have similar meaning gets clustered together and thus we get a vector space semantic of words. Various arithmetic operations can also be done to these vectors to produce analogies in word meaning.</p> <p>ELMo can be immensely helpful to improve NLP systems such as question answering from a certain paragraph, semantic arguments of the verbs, name-entity relationship etc. Authors also pointed the limitation of word embeddings or vectors. They showed that the word vectors can be biased based on the data set used to produce those vectors. Moreover, language models are not just words placed sequentially, it is more that just a collection of words.</p>
Intellectual Merit	<p>This paper explains the concepts of word vectors and their usability in different NLP systems. Though this paper did not introduce any novel algorithm and processes, it sheds light to one of the important concepts of natural language processing. The author Noah A. Smith is a well known NLP researcher having almost 12 years of experience in this field.</p>
Broader Impact	<p>This paper imparts knowledge about the word vectors and their story of origin. It catches the readers attention by showing intuitive examples of several complex problems in NLP. The author of this paper is a scientist in natural language processing research at the University of Washington.</p>

Keywords    Natural Language Processing, Word Vectors, Context, Evaluation

- Discussion        • The authors discussed about the contextual word vectors but they did not explicitly  
Questions        showed how the context of a word changes in terms of data set.
- The author did not talk about the efficiency of the models described in the paper.  
Comparison in between models could be helpful for the readers to understand more  
about the complexity of these NLP models.

Table 1: Grade deductions by section

Overview	Intellectual M.	B. Impact	Keywords	Questions	Is Online?