**NLP and Deep Learning-based Analysis of Building Regulations to support Automated Rule Checking System**

The objective of this research is to apply deep learning-based NLP for translating building regulations into a computer-readable format. This paper focuses on how computers can be able to understand semantic meaning of building regulations to automate rule interpretation process. Automated rule checking has been developed in various ways and enhanced the efficiency of building design review process. Converting human-readable building regulations to computer-readable format is, however, still time-consuming and error-prone due to the nature of human languages. Several domain independent efforts have been made for NLP.

The proposed process is composed of following steps:

1) learning semantics of words and sentences,

2) utilization of semantic analysis. For semantic analysis, we use word embedding technique which converts meaning of words in numerical values. By using those values, computers can extract related words and classify the topic of sentences. The results of the semantic analysis can elaborate the interpretation with domain-specific knowledge.

Deep learning-based NLP enables the computers to learn the semantics of natural language from raw text data. Unlike conventional method, deep learning model extracts training features from raw text and adjusts the weight without human’s intervention. It can alleviate the manual efforts to interpret regulation sentences which required for automated rule checking.

This paper proposes a semantic analysis process and its utilization to support rule interpretation. In this paper, Word embedding model is used for learning the semantics of text. Through the training, the semantic of word and sentences are represented in vector values. The results of the semantic analysis are utilized for extracting related words and classifying the topic of sentences. Sentence classification based on deep learning model enables the computer to classify the regulatory sentence according to its content. Extracting related words helps both human and computer to find semantic information from raw text.