Partial Differential Equations and Applied Mathematics Seminar

Title Mathematical challenges on Korean Text Clustering and Applications

Speaker   장봉수 교수

Affiliation   울산과학기술대학 수학과

Date   May 30th, Wed., 5:00~6:00 Pm

Location   Science Building #225 Yonsei University

Abstract

As the name of fourth industrial revolution, a machine learning technology with neural network is one of the most important technology to analyze a large amount of data. In this talk, we introduce a basic understanding of text clustering based on word2vec, which is a shallow, two-layer neural network model that produces a word embedding vector. From the fact that many asian characters have their own meanings, we propose a new syllable based feature extraction method for korean text data and show the effectiveness of its performance. Moreover, we apply it to cluster precedents of korean supreme court.