

Persian Multi-Documents Summarization by Deep Learning

Shima Mehrabi Hamid Reza Ahmadifar Seyed Abolghasem Mirroshandel

Faculty of Engineering, University of Guilan, Rasht, Iran

ABSTRACT

With the increasing amount of the accessible textual information via internet, it seems necessary to have summarization system which is able to generate summary of information for user demands. Since long time ago, summarization has been considered by natural language processing researchers. Today, with improvement in processing power and development of computational tools, efforts to improve the performance of summarization system are continued. In this paper, a novel Persian multi-document summarization system is proposed that works based on a machine learning method called Deep Learning. Mostly Deep Learning uses artificial neural networks in learning process. The result of using deep learning in speech recognition and picture processing are sound promising which convinced natural language processing researchers to apply Deep Learning in NLP tasks. The proposed system ranks the sentences based on some predefined features and by using a deep artificial neural network called Autoencoder. The performance of the system is evaluated in Persian and the result of evaluations demonstrates the effectiveness and success of proposed summarization system.

Keywords: Artificial Neural Networks, Deep Learning, Multi-Documents Summarization, Persian Language Processing.