Keywords:

bibliometric software tools, Scopus, VOSviewer, Biblioshiny, innovative economy, education-science-business interaction, K-Means, word embeddings, pretrained models, clustering, clustering quality

Abstract

This paper presents the analysis of scientific publications on the interaction of education, science and business in the innovation economy on the basis of bibliometric software, sources from the Scopus scientometric database, supplemented by data visualization and descriptive analysis. The usage of clustering based on the word semantical similarity as well as clustering quality evaluation has been proposed to extend the data analysis opportunities in the scope of research topic evaluation. Different pretrained word embedding models were tested: GloVe, Word2Vec and transformers models. This allows us to evaluate the effective clustering quantity and extend the topic analysis using both the representation of our methods and known software (VOSViewer, Biblioshiny). It is shown also that performing the dimensionality reduction for this research is more effective before K-Means clustering than after it.