CONTENT-BASED RECOMMENDER SYSTEM FOR SCIENTIFIC PUBLICATIONS

Milovan Milivojevic, Ph.D.[1], Srdjan Obradovic B.Sc.[1], Miklos Pot, M.Sc.[2], Dragoljub Drndarevic Ph.D.[1]

[1]College of Applied Sciences, Užice, Serbia, milovan.milivojevic@vtps.edu.rs
[2]Polytechnical Engineering College, Subotica, Serbia, pmiki@vts.su.ac.rs

***Abstract**: Content-based recommender systems (CBRS) in the last decades had been facing dynamic development in various domains from e-commerce and marketing to the film industry, the games industry, music, education, medicine and other fields. In the paper the CBRS for scientific publications focused on abstracts and keywords is presented. The developed model covers a wide range of processes and methods from natural language processing, feature selection, text-centric document vectorization, clustering methods and classification techniques. In addition to numerous paradigms of content similarity, the concept of diversification has been incorporated in order to prevent overspecialization. Developed solution provides: a scientific domain analyzer based on a database of scientific papers, recommendations base for existing papers and recommendations for paper classification into appropriate scientific fields. Software solutions were implemented using R and Python platforms. The developed model is validated through the case study database of scientific papers from the SED scientific conference. The obtained results show that the CBRS based on abstracts and keywords shows better performance compared to document search only by keywords.*

***Keywords:** content-based recommender system, content similarity, TF-IDF, clustering, classification, scientific publications*