Chapter 12
Recommender Systems in Digital Libraries Using Artificial Intelligence and Machine Learning:
A Proposal to Create Automated Links Between Different Articles Dealing With Similar Topics

Namik Delilovic

https://orcid.org/0000-0002-3955-8816
Graz University of Technology, Austria

ABSTRACT

Searching for contents in present digital libraries is still very primitive; most websites provide a search field where users can enter information such as book title, author name, or terms they expect to be found in the book. Some platforms provide advanced search options, which allow the users to narrow the search results by specific parameters such as year, author name, publisher, and similar. Currently, when users find a book which might be of interest to them, this search process ends; only a full-text search or references at the end of the book may provide some additional pointers. In this chapter, the author is going to give an example of how a user could permanently get recommendations for additional contents even while reading the article, using present machine learning and artificial intelligence techniques.

INTRODUCTION

Natural Language Processing (NLP) is not a term invented recently. As noted by (Liddy, 2001), research already started in the late 40s in the form of machine translation, which was used to encipher enemy messages. Early machine translation used primitive dictionary look-ups and some word reordering

techniques to fit the grammar of the target language. However, after Chomsky's publication *Syntactic Structures,* which made the fields of linguistic and machine translations closer to each other, other application emerged, such as speech recognition. Today when we speak about NLP, we usually refer to it as a discipline of artificial intelligence, and indeed the final goal of NLP is Natural Language Understanding (NLU). The objectives of NLU, as stated by (Liddy, 2001), are:

1. Paraphrase an input text
2. Translate the text into another language
3. Answer questions about the contents of the text
4. Draw inferences from the text

While the increase of memory and processing power led to the rapid development of machine learning (ML) supported Natural Language Processing (NLP) techniques, the ever-increasing amount of data found on the web and technologies which provide faster internet speed (Jelena, 2020) are crucial for the advance of such techniques. This is also the reason the Austria-Forum platform (*Austria-Forum)* which holds over 1.2 million objects in the form of web-books (digital books), documents, images and other multimedia objects is essential for the author's experiments and research. After analysing the usage behaviour of the Austria-Forum users, the author realised that most of the readers (87.34%) come directly from the Google search engine. The different channels and their values are shown in the pie chart of Figure 1. Combining this information with the average number of users from the last 30 days (right plot in Figure 1) the author concludes that out of 7000 daily users, 6000 come from the Google search engine. Even though the platform providers are happy for any channel which leads the user to the content they offer, the fact that the user is utilising the limited Google search option is not satisfying. In section CURRENT STATE the author is going to examine the current state of the Austria-Forum platform and how the users find new contents using essential tools such as navigation, search and browsing, in section FIRST ATTEMPTS the author presents the first version of the linker tool which automatically suggests links for words found in articles. This tool is simple in its implementation and therefore, in section FUTURE IMPLEMENTATIONS the author analyses some current machine learning techniques and proposes a new implementation for the linker tool and finally, in section EVALUATION the author examines the evaluation possibilities for the results which the new linker implementation will provide.

Related Content

An Automated Self-Healing Cloud Computing Framework for Resource Scheduling

Bhupesh Kumar Dewangan, Venkatadri M., Amit Agarwal, Ashutosh Pasrichaand Tanupriya Choudhury (2021). *International Journal of Grid and High Performance Computing (pp. 47-64).*

www.irma-international.org/article/an-automated-self-healing-cloud-computing-framework-for-resource-scheduling/266218

Two Approaches of Workflow Scheduling with QoS in the Grid

Fangpeng Dongand Selim G. Akl (2009). *Quantitative Quality of Service for Grid Computing: Applications for Heterogeneity, Large-Scale Distribution, and Dynamic Environments (pp. 1-27).*

www.irma-international.org/chapter/two-approaches-workflow-scheduling-qos/28268

Cellular Automata and GPGPU: An Application to Lava Flow Modeling

Donato D'Ambrosio, Giuseppe Filippone, Rocco Rongo, William Spataroand Giuseppe A. Trunfio (2012). *International Journal of Grid and High Performance Computing (pp. 30-47).*

www.irma-international.org/article/cellular-automata-gpgpu/69804

Semantic Analysis of Rough Logic

Qing Liu (2010). *Novel Developments in Granular Computing: Applications for Advanced Human Reasoning and Soft Computation (pp. 264-284).*

www.irma-international.org/chapter/semantic-analysis-rough-logic/44707

Dynamic Dependent Tasks Assignment for Grid Computing

Meriem Meddeberand Belabbas Yagoubi (2012). *Grid and Cloud Computing: Concepts, Methodologies, Tools and Applications (pp. 551-565).*

www.irma-international.org/chapter/dynamic-dependent-tasks-assignment-grid/64502