Chapter XIV

A Machine Learning Approach for One-Stop Learning

Marco A. Alvarez, Utah State University, USA

SeungJin Lim, Utah State University, USA

Abstract

Current search engines impose an overhead to motivated students and Internet users who employ the Web as a valuable resource for education. The user, searching for good educational materials for a technical subject, often spends extra time to filter irrelevant pages or ends up with commercial advertisements. It would be ideal if, given a technical subject by user who is educationally motivated, suitable materials with respect to the given subject are automatically identified by an affordable machine processing of the recommendation set returned by a search engine for the subject. In this scenario, the user can save a significant amount of time in filtering out less useful Web pages, and subsequently the user's learning goal on the subject can be achieved more efficiently without clicking through numerous pages. This type of convenient learning is called one-stop learning (OSL). In this chapter, the contributions made by Lim and Ko (2006) for OSL are redefined and modeled using machine learning algorithms. Four selected supervised learning algorithms:

support vector machine (SVM), AdaBoost, Naive Bayes, and Neural Networks are evaluated using the same data used in Lim et al. (2006). The results presented in this chapter are promising, where the highest precision (98.9%) and overall accuracy (96.7%) obtained by using SVM is superior to the results presented by Lim et al. (2006). Furthermore, the machine learning approach presented here demonstrates that the small set of features used to represent each Web page yields a good solution for the OSL problem.

A Machine Learning Approach for One-Stop Learning

Using the Web, a global repository of information, for educational purposes requires more accurate and automated tools than general-purpose search engines. Innovative tools should support the learning experience and focus the attention of the learner on his or her desired target subject. A typical learner would be interested in going directly to the point and learn without spending time with useless or non-informative pages. In this context, harvesting the Web using current search engines and technologies, however, looking for concepts, subjects, or general information usually imposes significant overhead that is denoted when the user spends time in filtering irrelevant pages or when he or she is simply distracted with advertisements, latest news, or attractive but not suitable Web sites for learning.

Before the advent of the Web, students and occasional learners studied new subjects by reading books or well-known articles in which they could find all the required information. Certainly, these primary sources of information can be considered adequate and sufficient for learning the subject when the learner satisfies his or her aspirations with them. In most cases, there is no need to look for additional resources for the same subject. This conventional learning strategy is called one-stop learning (OSL) in Lim and Ko (2005). On the other hand, when considering the Web as a repository for learning, the learners very often rely on available general-purpose search engines like Google, Yahoo, or Microsoft Live in finding suitable materials for OSL. Here, it must be emphasized that these search engines were not designed with the specific goal of assisting educational activities. The use of such engines for one-stop learning needs to be revisited in order to optimize the time that learners spend searching for self-contained sources of knowledge/information. One clear advantage of existing search engines is the fact that they maintain billions of updated pages already indexed for fast search and retrieval. Previously proposed strategies for OSL using the Web take advantage of the results returned by search engines (Lim et al., 2005, 2006). The major motivation in this chapter is to present a machine learning approach for the OSL problem making use of existing search

23 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/machine-learning-approach-one-stop/7523

Related Content

Testing the SmartMunk's Story.ly App for the Analysis of Online Customer Reviews on Face Masks
Arabela Briciu, Cristian-Laureniu Romanand Victor-Alexandru Briciu (2021). *New Opportunities for Sentiment Analysis and Information Processing (pp. 90-116).*
www.irma-international.org/chapter/testing-the-smartmunks-storyly-app-for-the-analysis-of-online-customer-reviews-on-face-masks/286906

Information Extraction: Methodologies and Applications
Jie Tang, Mingcai Hong, Duo Liang Zhangand Juanzi Li (2008). *Emerging Technologies of Text Mining: Techniques and Applications (pp. 1-33).*
www.irma-international.org/chapter/information-extraction-methodologies-applications/10174

Ranking News Feed Updates on Social Media: A Review and Expertise-Aware Approach
Sami Belkacemand Kamel Boukhalfa (2021). *International Journal of Data Warehousing and Mining (pp. 15-38).*
www.irma-international.org/article/ranking-news-feed-updates-on-social-media/272016

Novel Text Summarization Techniques for Contextual Advertising
Giuliano Armanoand Alessandro Giuliani (2014). *Innovative Document Summarization Techniques: Revolutionizing Knowledge Understanding (pp. 185-204).*
www.irma-international.org/chapter/novel-text-summarization-techniques-for-contextual-advertising/96745

Using Big Data to Improve the Educational Infrastructure and Learning Paradigm
Areej Fatemah Meghjiand Naeem A. Mahoto (2016). *Effective Big Data Management and Opportunities for Implementation (pp. 158-181).*
www.irma-international.org/chapter/using-big-data-to-improve-the-educational-infrastructure-and-learning-paradigm/157691