A Comparative Study on Various Vocabulary Knowledge Scales for Predicting Vocabulary Pre-Knowledge

Di Zou, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
Haoran Xie, The Education University of Hong Kong, Tai Po, Hong Kong
Yanghui Rao, Sun Yat-sen University, Guangzhou, China
Tak-Lam Wong, The Education University of Hong Kong, Tai Po, Hong Kong
Fu Lee Wang, Caritas Institute of Education, Tseung Kwan O, Hong Kong
Qingyuan Wu, Beijing Normal University Zhuhai Campus, Zhuhai, China

ABSTRACT

The world has encountered and witnessed the great popularity of various emerging e-learning resources such as massive open online courses (MOOCs), textbooks and videos with the development of the big data era. It is critical to understand the characteristics of users to assist them to find desired and relevant learning resources in such a large volume of resources. For example, understanding the pre-knowledge on vocabulary of learners is very prominent and useful for language learning systems. The language learning effectiveness can be significantly improved if the pre-knowledge levels of learners on vocabulary can be accurately predicted. In this research, the authors model the vocabulary of learners by extracting their history of learning documents and identify the suitable vocabulary knowledge scales (VKS) for pre-knowledge prediction. The experimental results on real participants verify that the optimal VKS and the proposed predicting model are powerful and effective.

KEYWORDS
E-Learning Systems, Learner Profile, Pre-Knowledge Prediction, User Modeling, Vocabulary Learning

1. INTRODUCTION

We have encountered and witnessed the great popularity of various emerging e-learning resources such as massive open online courses (MOOCs), textbooks and videos with the development of the big data era. On the other hand, the emerging social media and ubiquitous techniques have played an important role in our daily lives. These evolving techniques and e-learning applications have not only provided a large volume of learning resources on the web but also significantly changed the ways people learn new knowledge and access information.

It is critical to understand the characteristics of users to assist them in finding desired and relevant learning resources in such a large volume of resources. For example, understanding the pre-knowledge on vocabulary of learners is very prominent and useful for language learning systems (Chen & Li, 2010; Zou, Xie, Li, Wang & Chen, 2014). That is, the language learning effectiveness can be significantly improved if the pre-knowledge levels of learners on vocabulary can be accurately predicted. In this research, the authors model the vocabulary of learners by extracting their history of learning documents and identify the suitable vocabulary knowledge scales (VKS) for pre-knowledge prediction. The experimental results on real participants verify that the optimal VKS and the proposed predicting model are powerful and effective.

DOI: 10.4018/IJDET.2017010105

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
improved if the pre-knowledge levels of learners on vocabulary can be accurately predicted. Another typical example of modeling the characteristics of users is the tag-based user profiles which have widely adopted to facilitate personalized search or recommendation in the social media domain (Xie, Li & Mao 2012; Xie, Li, Mao, Li, Cai & Rao 2014).

In this research, we focus on the research problem of how to model language learners in the e-learning systems. As the fast development and popularity of e-learning systems, a collection of user learning documents such as writing assignments, reading essays and accessed lecture notes are available in log database of such systems (Li, Lau, Wah, Ashman, Leung, Li & Lee, 2009). To understand the language learners, we try to predict the pre-knowledge levels on vocabulary of learners according to their history of learning documents. In particular, we focus on two kinds of learning documents, which are writing assignments and reading essays. The reason of selecting writing assignments and reading essays is that they can reflect both receptive and productive knowledge of words (Webb, 2005). The main contributions of this research are summarized as follows.

1. We propose three rating functions according to different granularity of the vocabulary knowledge scales (VKS) for identifying pre-knowledge levels on vocabulary of learners;
2. We model the language learner by adopting vector-based learner profile;
3. We present a framework which can integrate with different rating functions for writing assignments and reading essays;
4. We compare accuracy of the proposed framework with baselines in real-world study circumstances;
5. We examine the impact of various parameters and document sources in the experiment.

The rest of this article is organized as follows. In Section 2, we review the related research on language learning and user modeling. In Section 3, we introduce the proposed rating functions, framework and models. Section 4 reports the experimental settings, metrics and results. The implications of this research are discussed in Section 5. In Section 6, the findings of this research are summarized and future directions are discussed.

2. RELATED WORK

In this section, we review the related research in two areas, which are user modeling techniques, and language learning systems.

2.1. User Modeling Techniques

As more and more data generated by users (e.g., the popularity of user-generated data in social media), user modeling becomes an essential and important topic in the research communities. Kobsa (2001) reviewed the development of generic user modeling systems over the last twenty years in the 20th century, and summarized that shell systems form part of the application, central server systems communicate with several applications and possible future user modeling agents physically follow the user. Shen, Tan & Zhai (2005) investigated a major deficiency of existing retrieval systems is that they generally lack user modeling and are not adaptive to individual users, resulting in inherently non-optimal retrieval performance. A new approach was proposed to construct DLs that satisfy a user’s necessity for information: Adaptive DLs, libraries that automatically learn user preferences and goals and personalize their interaction using this information (Frias-Martinez, Magoulas, Chen, & Macredie, 2006).

With the popularity of recommender and social computing systems, user modeling techniques were adopted to provide the personalized recommendations (Abel, Gao, Houben & Tao, 2011) and motivate user participation (Vassileva, 2012). More recently, user-generated data such as Twitters, collaborative tags are adopted for modeling users to support context-aware applications (Xie et al.,
User Interface Design Pedagogy: A Constructionist Approach
www.igi-global.com/article/user-interface-design-pedagogy/49705?camid=4v1a

Open and Distance Learning Administration: The Organizational Models, Cultures, and Structure and the Barriers and Trends in ODL Administration
Gürhan Durak (2018). *Administrative Leadership in Open and Distance Learning Programs* (pp. 84-116).
www.igi-global.com/chapter/open-and-distance-learning-administration/182904?camid=4v1a