

PAPER ID: 17-05-0227

A Highly Accurate PDF-to-Text Conversion System for Academic Papers using Natural Language Processing Approach

Tien Fui Yong¹, Saiful Azad^{2,3,*}, Mohammed Mostafizur Rahman⁴, and Kamal Z. Zamli^{2,3}

¹Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, Bandar Barat, 31900 Kampar, Perak, Malaysia.

²Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.

³IBM Centre of Excellence, UMP, Gambang, Malaysia.

⁴American International University Bangladesh (AIUB), Banani, Dhaka, Bangladesh

Extracting text out of PDF documents is never an easy task when a higher degree of accuracy and consistency are the two main criteria to be attained. Although, there exist a considerable number of such systems; however, most of them are falling short of offering desirable performance especially when academic literature is the concern. Researches, those involved heavily in text mining and project analyzing, need an accurate and consistent supporting tool for PDF-To-Text (PTT) conversion. Therefore, in this paper, we propose a Natural Language Processing based PDF-to-text (NLPDF) conversion system, which comprises of two major steps, namely i) reads contents from the PDF and ii) reconstruct the text. The performance of the proposed system is evaluated via four metrics, namely Precision, Recall, F-Measure (AF), and standard deviation, and compared with eight other similar benchmarked systems available in the market. The experimental results evidently demonstrate the effectiveness of the proposed system.

Keywords: PDF-to-Text Conversion, Natural Language Processing, Edit Distance

PAPER ID: 17-05-0175

The Impact of Information Quality and Usability on Intention to Use Student Information System

Mohammed Thamir Atta^{1,2}, Awanis Romli^{1,2}

¹ Faculty of Computer System and Software Engineering, Universiti Malaysia Pahang

² Faculty of Computer System and Software Engineering, Universiti Malaysia Pahang Kuantan, Pahang, Malaysia

Educational institutions become more dependent on information systems because of the essential functions that can benefit from these systems. The aim of this research is to evaluate and update the success model of DeLone and McLean based on their theory of information system in order to be implemented in academic environment. This study introduced usability, information quality, the intention to use SIS, and user's satisfaction in the domain of information systems. The findings from this study shows that information quality of content provided by SIS, the usability of system interface, and the intention to use student information system, all these three variables affect user's satisfaction in higher education institutions. The study recommends to focus on quality measures of content and usability of system which in turn increase the intention of students to use student information system and enhance their satisfaction about the system.

Keywords: Information System, Student Information Systems (SIS), Information Quality (IQ), Usability, Intention to Use SIS, User's Satisfaction.