

Combined Support Vector Machine and Pattern Matching for Hadith Question Classification System

Ali Muttaleb Hasan¹, Taha H. Rassem^{1 2 *}, M.N. Noorhuzaimi@Karimah¹

¹ Faculty of Computer System & Software Engineering
2IBM Centre of Excellence
University Malaysia Pahang, 26300 Gambang, Kuantan,
Pahang DarulMakmur, Malaysia
alimatlab65@yahoo.com, tahahusseini@ump.edu.my*,
nhuzaimi@ump.edu.my

Abstract :

The dimensional phase of question answering (QA) involves an intrinsic form of question classification (QC) that functions to perform an important task in question answering system (QAS). The purpose of QC is to precisely assign labels to questions that are majorly dependent on the form of answer type. Moreover, classification of user's question is a herculean task based on the tractability that natural language (NL) affords with different forms. The information enshrined in a group of words is not sufficient to effectively classify the question in the quote. Until now, few reports have focused on QC for Arabic question answering (QA). The earlier report has employed the technique of handcrafted rules and keyword matching for QC. Nonetheless, these procedures are considered obsolete in terms of applying it to new territories. This study aims on combination of model fixed on SVM and pattern-based techniques for Arabic question classification (AQC). The Hadith purview on QA in the study was focusing on the effect of a feature set on the performance of SVM for QC. About five patterns were employed in the analysis together with classification of three types of questions, namely "Who", "Where" and "What". The dataset employed in this study consisted of 200 questions on Hadith derived from Sahih Al-Bukhari. The performance generated for the Fmeasure values for "Who", "Where" and "What" were 88.39%, 87.66% and 87.93% respectively. The outcome of this study reflected that the proposed prototype of SVM and pattern-based approach is indispensable from the field of QC in the Arabic language.

Keywords: *Question Classification System, Machine learni*