Abstract

Online news became one of the favorite information sources for most of the people nowadays because of its update rate and availability over the 24 hours rather than the traditional newspapers. Measuring the readability of the news articles gives a clear view for both the readers and the writers about how easily people can read and understand these articles. In this paper, we present MLAR, a new machine learning based system for Arabic text readability, and use it in measuring the readability of the Arabic online news articles from different outlets. The proposed system is able to determine the topic of each article efficiently and calculates its readability score level. The results show that readability of the online Arabic news is affected by the nature of its topic and the source outlet. The writing style of news articles in each topic differs from one outlet to another.

References


Index Terms

Computer Science

Information Sciences

Keywords

MLAR, Machine Learning, Text Mining, Readability, Arabic Online News, Natural Language Processing.