Abstract

Text, either online or offline can be presented in two different writing styles: formal and informal writing style. A Piece of text may contain a lot of emotion, ideas or feelings. Various techniques and methods are present in the field of Opinion Mining and Sentiment Analysis to extract the emotions from text. This paper presents a differential analysis of Formal and Informal text pieces in the field of Sentiment Classification. This paper presents a study and analysis of differences of approaches used for Emotion Detection and Sentiment Analysis for both cases. In this study, 10 formal text pieces in form of poetry, proverbs, essay and document are analyzed. These text pieces are present in 7 different International languages (i.e. Persian, Spanish, Chinese, Arabic, Malaysian, English, Ottoman). Informal text, in form of chats, emails, review sites and micro blogs, written in different International languages (Korean, Persian and English) are considered in this study. Various machine learning based methods – Support Vector Machine(SVM), Naive Bayes (NB), Decision Tree are more often used in classification of literary arts especially poetry. Statistical Machine learning approach Support Vector Machine outperforms all other methods in case of poetry and NB performed well in case of Informal Writing Style.

Index Terms

Computer Science
Information Sciences
Keywords
Emotion Detection  Formal Text  Informal Text  Opinion Mining  Sentiment Analysis.