

Polarity Classification of Social Media Comments in Turkish

Authors : Migena Ceyhan, Zeynep Orhan, Dimitrios Karras

Abstract : People in modern societies are continuously sharing their experiences, emotions, and thoughts in different areas of life. The information reaches almost everyone in real-time and can have an important impact in shaping people's way of living. This phenomenon is very well recognized and advantageously used by the market representatives, trying to earn the most from this means. Given the abundance of information, people and organizations are looking for efficient tools that filter the countless data into important information, ready to analyze. This paper is a modest contribution in this field, describing the process of automatically classifying social media comments in the Turkish language into positive or negative. Once data is gathered and preprocessed, feature sets of selected single words or groups of words are build according to the characteristics of language used in the texts. These features are used later to train, and test a system according to different machine learning algorithms (Naïve Bayes, Sequential Minimal Optimization, J48, and Bayesian Linear Regression). The resultant high accuracies can be important feedback for decision-makers to improve the business strategies accordingly.

Keywords : feature selection, machine learning, natural language processing, sentiment analysis, social media reviews

Conference Title : ICACII 2020 : International Conference on Applied Computational Intelligence and Informatics

Conference Location : Paris, France

Conference Dates : September 17-18, 2020