SENTIMENTAL ANALYSIS BASED ON SINHALA LANGUAGE SOCIAL MEDIA UPDATES

Shanmuganathan Vasanthapriyan

Department of Computing & Information Systems, Sabaragamuwa University of Sri Lanka.

priyan@appsc.sab.ac.lk

The World Wide Web plays a critical role in collecting public opinion where these opinions play an important role in making business decisions. For factual and subjective information about companies and products, analysts are turning to the Internet to gather information. Extracting the public opinion is the difficult task in a country like Sri Lanka, because most of the time the language spoken is, Sinhala or Tamil rather than English. Sentimental Analysis being a major research topic in computational linguistic community is quite popular and has led building of better products, understanding user's opinion, executing and managing of business decisions. However most of the researches never focused on South Asian languages like Sinhala, often used in Social media websites such as Face book, Twitter and etc. Motivated by Sentimental Analysis researches based on Hindi, another south Asian language, we proposed and developed a system that analyzes social media updates in Sinhala language for the sentiments. Starting with three basic sentiments; Positive, Negative and Neutral we retrieve a set of live updates based on Face book and Twitter. This data set is then deployed in to a cloud service and analyzed and give the proper output. Sinhala is a free order language compared to English which adds complexity while handling user generated content. Our finding focuses on how to build a better platform on sentimental analysis to help bloggers to stop spam, business firms to get feedback, and government firms to get urgent service requests. We hope to do more investigation on implicit factors in Sinhala language and give them as features for the models we described in our work.

Keywords: Sentimental analysis, Social media, Linguistics, Opinion mining