MixedEmotions: Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics MarketsInnovation Action

Gabriela Vulcu and Paul Buitelaar

Insight Centre for Data Analitics at National University of Ireland, Galway,
{gabriela.vulcu,paul.buitelaar}@insight-centre.org

1   Introduction

Emotion analysis is central to tracking user behavior and satisfaction, which can be observed from user interaction in the form of explicit feedback through email, call center interaction, social media comments as well as implicit acknowledgment of approval or rejection through facial expressions, speech or other nonverbal feedback. Making sense of accumulated user interaction from different ("mixed") data sources, modalities and languages is challenging and has not yet been explored in fulness in an industrial context. MixedEmotions aims at developing an integrated Big Linked Data platform for emotion analysis across heterogeneous data sources, languages and modalities, building on existing state of the art tools, services and approaches that will enable the tracking of emotional aspects of user interaction and feedback on an entity level.

2   MixedEmotions approach

MixedEmotions aims at providing an integrated data platform solution for: i) large-scale emotion analysis and fusion on heterogeneous, multilingual, text, speech, video and social media data streams, leveraging open access and proprietary data sources, exploiting also social context by leveraging social network graphs, ii) semantic-level emotion information aggregation and integration through robust extraction of social semantic knowledge graphs for emotion analysis along multidimensional clusters

The platform will be developed and evaluated in the context of three cross-domain Pilot Projects that are representative of a variety of data analytics markets: Social TV, Brand Reputation Management, Call Centre Operations. Fig. 1 shows the overview of the MixedEmotions architecture focusing on the data-value chain from the data acquisition of the available "mixed" resources, to their analysis and integration, and finally to exploitation in terms of the three pilots.

Pilot Project I (Social TV): For social TV, MixedEmotions will provide solutions for entity-level emotion and sentiment analysis (program, actor, sportsperson), related contents analytics (liked content, mentioned entities), multimedia media analytics (emotion and sentiment analysis on online videos), and Social TV (related social networks activity while watching TV programs).
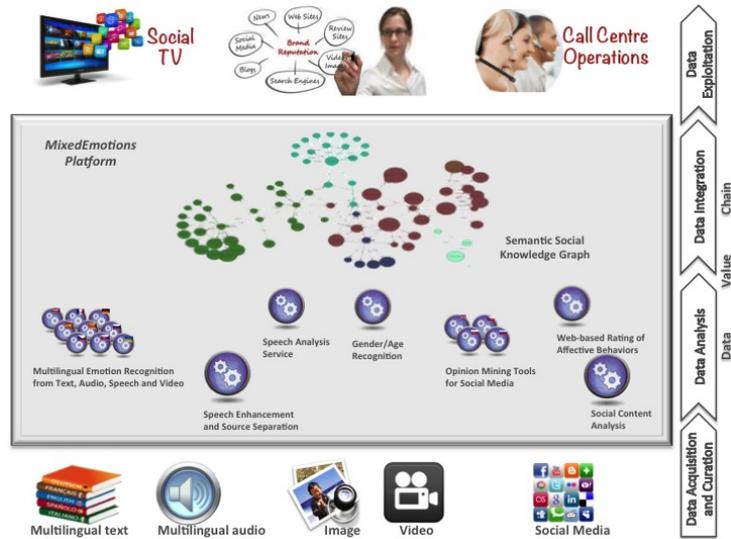
Fig. 1. Basic Architecture of the MixedEmotions Platform

Pilot Project II (Brand Reputation Management): MixedEmotions aims at contributing on several directions: enhance real-time capabilities, adding multi-modal input would dramatically enhance the quality and accurateness of generated reputation reports, multilingual data analysis and report generation to increase internationalisation of the product, to deal with growing data to be analysed faster and more reliably, advanced extraction of emotions and opinions, automated content-based filtering and emotion analysis of speech and video, enhanced multilingual processing, linked data anchoring in order to better deal with data variety and volatility by providing the necessary unified ground.

Pilot Project III (Call Centre Operations): For call centre operations MxedEmotions will: enable the implementation and use of optimization algorithms for speech recognition in the context of big data(huge streams of speech recordings in mid-size and large call centres), perform advanced extraction of emotions from speech and enhanced multilingual processing.

3   Connection with other EU projects

The MixedEmotions project intersects in different aspects with various EU projects. Among these we mention "Big Data Europe" in terms of big data platforms; "Pheme" and "SSIX" for Social Media analysis; "LIDER" in terms of linguistic linked data; "MultiSensor" for multilingual sentiment analysis; "XLiMe" for multi-modal data fusion and "FREME" for semantic enrichemnt.