Sentiment Analysis for Smart Cities: State of the Art and Opportunities

Kaoutar Ben Ahmed[1], Atanas Radenski[2], Mohammed Bouhorma[1] Mohamed Ben Ahmed[1]
[1]Abdelmalek Essaâdi University, Tangiers, Morocco
[2]Chapman University, Orange, CA 92866, USA

Abstract - *Advances of information and communications technologies in general and of social media platforms in particular have changed the way people communicate and express themselves. Citizens are now using smartphones and other mobile devices to share, on an unprecedented scale, their experiences and views in blogs, micro-blogs, comments, photos, videos, and other postings in social sites. The smart city research community has already recognized that sentiment analysis can contribute to a better understanding of, and timely reactions to, public's needs and concerns by city governments. Yet, relatively little is known about how to best harness the potential benefits for smart cities of opinion mining and sentiment analysis. The objective of this article is to help fill the void by reviewing the state of the art, challenges and opportunities of sentiment analysis platforms, architectures and applications for the smart city application domain.*

Keywords: sentiment analysis, opinion mining, smart cities, big data, text mining

1 Introduction

Smart (or smarter) cities "are urban areas that exploit operational data, such as that arising from traffic congestion, power consumption statistics, and public safety events, to optimize the operation of city services" [1] A smart city's main objective is to increase the quality of life for its citizens and to make the city more attractive, lively and greener. To achieve this goal, smart city technologies (SCTs) are fused with the traditional city's infrastructure. SCTs refer to all the information and communications technologies (ICTs) that enable cities to harness big data gathered and analyzed in order to become connected and sustainable. The smart city concept emerged during the last decade as a fusion of ideas about how ICTs might improve the functioning of cities, enhancing their efficiency, improving their competitiveness, and providing new ways in which problems of poverty, social deprivation, and poor environment can be addressed. Urban communities worldwide are planning, developing and adopting digital systems and technologies to improve efficiency and quality of life for the citizens. According to a 2014 forecast, in 2025 there will be at least 26 global smart cities, about 50 percent of which will be located in North America and Europe [2]. Open data initiatives around the world make public data available online to wider audiences. Such initiatives have provided in recent years support for innovative projects, aiming the design of SCTs that enhance cities' smartness. A number of technological developments illustrate this trend. For example, real time big data analyses are used to boost public safety effectiveness by integrating smart solutions in disaster and emergency management, and in law enforcement systems. Smart mobility solutions based on road sensors and intelligent transportation systems are employed in transport and environment fields aiming to reduce urban congestion and CO2 emission. Websites and cloud services make a number of city government services accessible for anyone over the Internet; they can also become smarter due to tracking and analytics technologies that can discover usage and access patterns and respond to citizen's individual needs and preferences. Recommendation systems, mobile location-based applications, and other advanced ICTs offer city visitors novel and personalized experiences during which they are assisted by smartphone apps acting like electronic tourist guides updated with real time information about accommodation, dining options, weather, currency rates, nearby points of interest based on the visitor's personal preferences or geographical location or other strategies and touch-sensitive maps are found in all stations to help tourist easily find their way. Social media triggered the raise of sentiment analysis which brings new possibilities to city governance in general and decision making in particular.

Sentiment analysis is the examination of people's opinions, sentiments, evaluations, appraisals, attitudes, emotions, and personal preferences towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [3].

Traditionally, sentiment analysis mines information from various text sources such as reviews, news, and blogs then classifies them on the basis of their polarity as positive, negative or neutral. An important preliminary task of sentiment analysis is to evaluate the subjective or objective nature of source texts. Subjectivity indicates that the text bears opinion content whereas objectivity indicates that the text is without opinion content. Recently, sentiment analysis aims to exploit audio, video, location, and other non-traditional data sources.

An essential issue in sentiment analysis is to identify how sentiments are expressed in texts and whether the expressions indicate positive (favorable) or negative (unfavorable) opinions toward the subject. Thus, sentiment analysis involves identification of sentiment expressions, polarity and strength of the expressions, and their relationship to the subject [4].

Historically, sentiment analysis has been exclusively approached as a natural language processing task at many levels of granularity. Starting from being a document level classification task [5], it has been handled at the sentence level ([6]; [7]) and more recently at the phrase level ([8]; [9]).

Over the years, scholars and developers have introduced a number of terms that refer to tasks that are very similar to sentiment analysis, such as opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, and others. Recently, such diverse terms are converging under the umbrella of either sentiment analysis, or opinion mining [3]. To our knowledge, the two terms were first published at the same time, in 2003: the term sentiment analysis appeared first in [4] and the term opinion mining was first published in [10]. Currently, the two terms seem to represent the same field of study. In industry, the term sentiment analysis is more commonly used than opinion mining; in academia both sentiment analysis and opinion mining are frequently employed. In this paper, we choose to use the term sentiment analysis because it is well adopted in both industry and academia. We also acknowledge that sentiment analysis now expands beyond its traditional text sources to aim the analysis of non-text data, such as image and video data from social networks, and human vitals as measured by wearable computers.

In this paper, we focus on sentiment analysis as a promising smart city technology and offer a review of the state of the art, the challenges, and the opportunities in the areas of tools and techniques (section 2) and system architectures (section 3). We limit our considerations, unless otherwise required, to the smart city application domain and do not aim to discuss sentiment analysis in its full generality.

2 Tools and Techniques

Sentiment analysis relies on a variety of tools and techniques; their weaknesses and limitations justifies the necessity of a new generation of tools and techniques that either solve present challenges or offer new opportunities in the smart city domain.

2.1 State of the art

Comprehensive surveys like [5] and [3] reviewed various tools and techniques described in an ever growing pool of sentiment analysis publications. The particular focus of our article is on sentiment analysis tools for the smart city domain; therefore, we only briefly review results that we deem applicable to this domain. Thus, [11] used a sentiment detection tool named LIWC2007 which works with a psychometrically validated dictionary. [12] proposed a fusion between a lexicon-based sentiment classifier and a machine-learning based classifier. [13] investigate and evaluate NTLK, a platform for natural language processing in python, and also SentiWordNet3.0, a lexical resource for sentiment analysis. [14] used a Naïve Bayes model on unigram features and Crowdsourcing via Amazon Mechanical Turk (AMT). [15] devised a new sentiment detection mechanism that determines the keywords related to public reaction and descriptions of terrorism. This is done by feeding a list of potential sentiment-related keywords into their sentiment finder module, which will then tag the messages based on category of the keywords detected in the incoming message. The category defines the sentiment of the message. [16] applied logistic regression, Naive Bayes and Decision Tree (ID3) classifiers to perform sentiment classification of Twitter messages regarding Hurricane Irene. They have concluded that Tweets provide real-time insight into public perceptions of a disaster. [17] utilized Saltlux, which is a proprietary crawling and sentiment mining software, in order to convert the social media streams gathered mainly from Twitter into RDF stream and analyze its sentiment. The R programming packages and Support Vector Machines are applied by [18] to extract sentiment scores from newspaper articles and relate these scores to an economic index. [19] argues that the calculation of a metric named "Gross national happiness" serves as a representation of the overall emotional health of the nation. To calculate this metric, he used the sentiment analysis approach via Text Analysis and Word Count (TAWC) program. [20] used OpinionFinder tool that measures positive vs. negative mood and Google-Profile of Mood States (GPOMS) that measures mood in terms of 6 dimensions (Calm, Alert, Sure, Vital, Kind, and Happy). Their goal was to investigate whether public mood as measured from large-scale collection of tweets is correlated or even predictive of economic indicators.

2.2 Opportunities

Current research in sentiment analysis focuses on the reducing of human effort needed to analyze content and increasing accuracy. Future research is expected to address such issues as visual representation, multilingual audiovisual opinion mining, building usable, peer to peer opinion mining tools for citizens, real-time sentiment analysis, multilingual reference corpora, recommendation algorithms, and automatic irony detection [21]. Additionally, the use of open source software is also expected to increase, most notably Hadoop ecosystem for batch data analysis (with Hive and Hbase), for interactive data analysis (with Presto and Storm) and for recommender systems in particular and statistical analysis in general (with Mahout).

3 Systems

Various system architectures have been proposed for sentiment analysis in the smart city domain. We describe some of them, discuss principle challenges for such systems, and we propose our own system architecture to address some of the difficulties with earlier systems.

3.1 State of the art

The architecture of a restaurant recommendation system [17] consists of three parts: (i) a client segment that interacts with the user and communicates to the back-end sending SPARQL (Protocol and RDF Query Language) queries, (ii) a data initiated segment that continuously analyses the social media streams, and (iii) a query initiated segment that uses the LarKC platform to answer the SPARQL queries of the client by combining several forms of reasoning. A proposed system for terrorism events detection [15] is implemented in Perl using the Net-Twitter Perl module that connects to Twitter with the Twitter (API); it queries the Twitter 'trends' API for names of places and identifies discussions of a flurry of activities at a specific location. Once the location of a possible threat has been identified, the system harvests all related Twitter messages using Twitter's Search API and Streaming API. A spam and noise filtering phase precedes a sentiment detection phase and a demographic exploration of the message pool. Finally, data mining and reporting phase takes place. In a smart learning context, a sentiment evaluation system by [12] extracts information through the Facebook API and records it in a NOSQL database. The system then models, evaluates, and visualizes the users' sentiment polarity (positive, neutral or negative) and users' significant emotional changes. TweetAlert [22] is a citizen's behavior analysis system which extracts tweets via the Twitter API and stores them it in a data warehouse which runs on top of Apache Lucene. The system uses the Textalytics APIs to perform sentiment analysis on the extracted data. Most notably, its user demographics analysis module extracts some important demographics (type, gender, age). End users can rely on the system's visualization component to exploit the stored annotated data. Several widgets have been developed to present the data, either just for query and reporting or for data analytics purposes. A city sensing system [23] goes through two different phases: the offline training phase and the online phase. In the offline training phase, it collects messages from Twitter that contain emotion-word hashtags. The messages are preprocessed and features are extracted from them. The features, together with the emotion derived from the emotion word hashtags are used to train a neural network. In the online phase, the system collects live geo-tagged tweets from the area of interest, e.g., a city. The trained neural network is used to detect emotion in these new tweets. The geotagged emotion data is then aggregated and visualized on a map. A presidential election analysis system is designed by [14]. The system's real-time data processing infrastructure is built on the IBM's InfoSphere streams platform. All relevant tweets are collected in real-time from the entire Twitter traffic via Gnip Power Track, a commercial Twitter data provider. Christopher Potts' basic Twitter tokenizer is used to preprocess the collected tweets. The system relies on manual sentiment annotation by Amazon Mechanical Turk users. The system outputs the number of positive, negative, neutral and unsure tweets in a sliding five-minute window. An Ajax-based HTML dashboard displays volume and sentiment by candidate as well as trending words and system statistics. The dashboard pulls updated data from a web server and refreshes its display every 30 seconds. This election analysis system can very well be applied to cover smart city elections.

3.2 Opportunities and Proposed System

The info-foundation of a smart city is related to the availability of big data, the interconnection of all city components, e.g., pollution sensors, traffic systems, social media and smartphones generates an increasing amount of data of all types, public or private, structured or unstructured and streamed or static. Collecting, analyzing and visualizing this big data come with the promise of empowering and enhancing the smart city governance. Smart cities data is available in variety of formats. Shared content by users on social media sites is often not only text but images as well and with the recent advances in visual sentiment analysis the now popular selfies can form a rich data source for sentiment analysis on social media. A few recent works attempted to predict visual sentiment using features from images ([24]; [25]; [26]; [27]; [28]). There are also studies that analyze speech-based emotion recognition ([29]; [30]; [31]; [32]; [33]). Additionally, a growing number of studies ([34]; [35] ;[36]; [37]; [38]; [39]) are concerned with multimodal sentiment analysis, integrating visual, acoustic and linguistic modalities. Some scholars ([40]; [41]) have claimed that the integration of visual, audio, and textual features can improve the analysis precision significantly over the individual use of one modality at a time leading to error rate reductions of up to 10.5%. Last, but not least, smartwatches are expected [42] to become capable of providing data on emotions: mood valence (positive vs. negative) and arousal (high vs. low). The location of smartwatches permits easy recording of heart rate variability and galvanic skin response (GSR). These two elements can be used to identify physiological arousal. With the advent of smartwatches and effective multisensor data collection, new algorithms (for sensor data fusion) might be developed that can identify valence without the need for processing a facial image. At this time, only the text data format is sufficiently well-understood for machine learning and sentiment analysis. In this paper, we propose a novel architecture of a sentiment analysis system designed for smart city governance purposes that will support a gradual transition sentiment analysis of text data towards sentiment analysis of multi-modal data. We believe that such architecture should be based on the emerging Apache Hadoop big data ecosystem in general and its Spark engine, and also on recently evolving tools that permit the incorporation within Hadoop of powerful statistical analysis languages and libraries, most notably R and Python's pandas. As a result, our proposed system architecture will amalgamate the big data potential of Hadoop with the existing wealth of R and Python libraries.

R is a popular open source software for statistical computing and data visualization. It was initially announced in 1993 as an extension of the S programming language with Scheme-like scoping rules. R is cross-platform (UNIX platforms, Windows and MacOS). R can be used as an interactive console, where users can try out individual statements and observe the output directly. This is useful in creating custom R scripts and exploring the data, where the output of the first statement can inform which step to take next [43].

R has a rich ecosystem of packages that is continuously enhanced by the active R community. It has over 4800 packages in topics like econometrics, data mining, spatial analysis, and bio-informatics. R excels over other open source machine learning libraries for it implements the majority of ML algorithms, including support vector machines, artificial neural network, naïve bayes, bayesian network, maximum entropy, part-of-speech tagging, named entity recognition, n-gram statistics, regressions, k-nearest-neighbors, hidden markov models, extreme learning machine. The rich R ecosystem has already been used in a number of sentiment analysis projects, such as ([44]; [45]; [18]), to mention a few.

Python is dynamic programming language which was first announced in 1991. Initially conceived as an easy top learn educational language, Python has recently gained popularity as a productive general purpose platform - in contrast to R, which is specialized for statistical computing. The adoption of Python in our architecture will support necessary system functionality beyond statistical analysis, such as web services for example. Yet, Python can be used for statistical processing as well, mainly through the emerging pandas framework - an open-source data manipulation and analysis library, which provides high-performance, easy-to-use data structures and data analysis tools. Python pandas has been created with the goal of becoming the most powerful and flexible open source data analysis tool available in any language [46].

The Apache Hadoop framework allows distributed processing of large data sets across commodity clusters by means of relatively simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. Rather than rely on hardware to deliver high-availability, the framework itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures [47]. The Hadoop core consists of four modules: Hadoop Commons, Distributed File System, YARN and MapReduce. A number of additional modules complement the Hadoop core by providing specialized services. For instance, Hive is a data warehouse with an SQL interface and Pig is an interpreter of a high-level programming language; both Hive and Pig compile into MapReduce. Hadoop's ability to process large volumes of data in a fault-tolerant batch mode has already attracted the attention of sentiment analysis researchers and its potential to solve various sentiment analysis tasks has been recognized ([18]; [19]). The MapReduce distributed computing engine, for example, can be pared with the HBase distributed storage to handle opinion lexicons and Mahout's machine learning algorithms can be applied to execute sentiment analysis tasks [48].

Apache Spark is notable big data framework, initially developed at the University of California in Berkeley and later transferred to Apache. Spark's capability to fully use the cluster's available fast memory gives it a significant performance edge to Hadoop's original MapReduce engine. Additionally, Spark offers a more flexible – in comparison to MapReduce - programming model that readily supports iterative processing – an area of weakness of the MapReduce model. Hence, Spark is emerging as a viable MapReduce alternative, which has become a most intensively developed big data framework. (According to OpenHub, during 2015 Spark has had 10194 Commits and 656 total contributors). Spark application can run in standalone mode or on Hadoop and be managed by YARN or Mesos. In the sentiment analysis domain, Spark has a clear edge over MapReduce because of its capability to support iterative machine learning algorithms.

While the sentiment analysis capabilities of Spark and, separately, of R and Python have been acknowledged, little is known about the benefits of their combined potential for sentiment analysis in general, and sentiment analysis for smart city governance in particular. This combined potential can now be explored and realized by means of the SparkR, PySpark, and Sparkling pandas tools. SparkR is a tool that provides a light-weight frontend to use the Spark big data engine from R [49], thus enabling native R programs to scale in distributed setting. Such integration of R and Spark brings a number of benefits including scalability to many cores, and machines within large clusters, optimizations in terms of code generation and memory management and the also ability to connect to a variety of data sources, such as Hive tables, JSON files, Parquet files etc. On the other side, the PySpark tool permits the use of the Spark engine from Python. PySpark, in combination with the SparklingPandas tool supports large scale data analysis with pandas on top of Spark.

We propose a system architecture (Fig. 1) which uses emerging tools, such as SparkR, PySpark, and Sparkling pandas to amalgamate the big data potential of the Hadoop ecosystem in general and its Spark engine in particular with powerful statistical analysis languages and libraries, most notably R and Python. We envision a two-stage implementation of such architecture. The core of this architecture comprises well established technologies - such as

R and Hadoop – that are traditionally used to handle text data for machine learning and sentiment analysis. Extension components for multi-modal data analysis, based on Python, PySpark, and Sparkling pandas, can be designed and developed after gathering some initial smart city sentiment analysis experience with the core components.

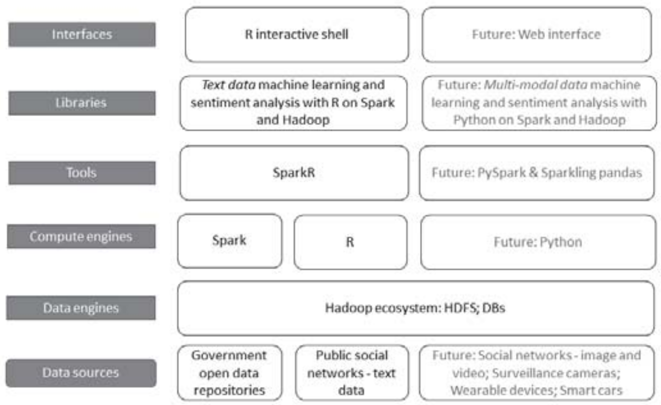

Fig.1. System architecture of sentiment analysis applications for smart cities

A central goal of our architecture is to integrate within Apache Hadoop the leading open-source big-data machine learning libraries, SparkR and Sparkling pandas, and support their use as smart city big data technologies, and to also fuse all the advantages and strengths of these tools via an easy-to-use web interface. The proposed architecture can be extended toward a high level, user-friendly decision support system that aims to assist governments and smart cities' managers by the use of civic engagement in policy development and implementation. Sentiment data can be extracted, in the near future, from several sources in modern cities, such as social media, government open data, surveillance cameras, selfie images, wearable devices, humanoid robots, to mention a few. The system core, together with extensions for analysis of non-text data represents a multimodal sentiment analysis system that can be employed as a sensing tool to gauge citizens' sentiments toward public subjects. A user-friendly interface complemented with interactive graphical visualizations capabilities will permit the use of this system by managers and civic leaders who have no technical programming and statistical background.

4 Conclusion

The development of smart city technologies requires joint efforts by the academia, the industry and the government to provide evolving systems and services. In contrast to the past decade, nowadays "the leading urban centers are not placing their technological futures in the hands of a company or a single university research group". "Instead, they are relying on a combination of academics, civic leaders, businesses, and individual citizens working together to create urban information systems that could benefit all these groups" [50]. In recognition of the need for joint efforts, the US Government has recently announced a new "Smart Cities" Initiative that will invest over $160 million in federal research and leverage more than 25 new technology collaborations to improve city services [51]. We believe that sentiment analysis is a key factor in the development of all smart city domains.

5 References

[1] C. Harrison, B. Eckman, R. Hamilton, P. Hartswick, J. Kalagnanam, J. Paraszczak, and P. Williams, "Foundations for Smarter Cities," *IBM J. Res. Dev.*, vol. 54, no. 4, pp. 1–16, Jul. 2010.

[2] F. & Sullivan, "Frost & Sullivan: Global Smart Cities market to reach US$1.56 trillion by 2020." [Online]. Available: http://www.prnewswire.com/news-releases/frost--sullivan-global-smart-cities-market-to-reach-us156-trillion-by-2020-300001531.html. [Accessed: 30-May-2016].

[3] B. Liu, "Sentiment analysis and opinion mining," *Synth. Lect. Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.

[4] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proceedings of the 2nd international conference on Knowledge capture*, 2003, pp. 70–77.

[5] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*, 2004, p. 271.

[6] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2004, pp. 168–177.

[7] S.-M. Kim and E. Hovy, "Determining the sentiment of opinions," in *Proceedings of the 20th international conference on Computational Linguistics*, 2004, p. 1367.

[8] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proceedings of the conference on human language technology and empirical methods in natural language processing*, 2005, pp. 347–354.

[9] A. Agarwal, F. Biadsy, and K. R. Mckeown, "Contextual Phrase-level Polarity Analysis Using Lexical Affect Scoring and Syntactic N-grams," in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, Stroudsburg, PA, USA, 2009, pp. 24–32.

[10] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews," in *Proceedings of the 12th International Conference on World Wide Web*, New York, NY, USA, 2003, pp. 519–528.

[11] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment.," 2010.

[12] A. Ortigosa, J. M. Martín, and R. M. Carro, “Sentiment analysis in Facebook and its application to e-learning,” *Comput. Hum. Behav.*, vol. 31, pp. 527–541, Feb. 2014.

[13] D. Yang, D. Zhang, Z. Yu, and Z. Wang, “A Sentiment-enhanced Personalized Location Recommendation System,” in *Proceedings of the 24th ACM Conference on Hypertext and Social Media*, New York, NY, USA, 2013, pp. 119–128.

[14] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan, “A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle,” in *Proceedings of the ACL 2012 System Demonstrations*, Stroudsburg, PA, USA, 2012, pp. 115–120.

[15] M. Cheong and V. C. S. Lee, “A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter,” *Inf. Syst. Front.*, vol. 13, no. 1, pp. 45–59, Sep. 2010.

[16] B. Mandel, A. Culotta, J. Boulahanis, D. Stark, B. Lewis, and J. Rodrigue, “A Demographic Analysis of Online Sentiment During Hurricane Irene,” in *Proceedings of the Second Workshop on Language in Social Media*, Stroudsburg, PA, USA, 2012, pp. 27–36.

[17] M. Balduini, I. Celino, D. Dell’Aglio, E. Della Valle, Y. Huang, T. Lee, S.-H. Kim, and V. Tresp, “BOTTARI: An augmented reality mobile application to deliver personalized and location-based recommendations by continuous analysis of social media streams,” *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 16, pp. 33–41, Nov. 2012.

[18] P. Hofmarcher, S. Theu\s sl, and K. Hornik, “Do Media Sentiments Reflect Economic Indices?,” *Chin. Bus. Rev.*, vol. 10, no. 7, 2011.

[19] A. D. I. Kramer, “An Unobtrusive Behavioral Model of ‘Gross National Happiness,’” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2010, pp. 287–290.

[20] J. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” *J. Comput. Sci.*, vol. 2, no. 1, pp. 1–8, 2011.

[21] D. Osimo and F. Mureddu, “Research challenge on opinion mining and sentiment analysis,” *Univ. Paris-Sud Lab. LIMSI-CNRS Bâtim.*, vol. 508, 2012.

[22] J. Villena-Román, “TweetAlert: Semantic Analytics in Social Networks for Citizen Opinion Mining in the City of the Future.”

[23] B. Guthier, R. Alharthi, R. Abaalkhail, and A. El Saddik, “Detection and Visualization of Emotions in an Affect-Aware City,” in *Proceedings of the 1st International Workshop on Emerging Multimedia Applications and Services for Smart Cities*, New York, NY, USA, 2014, pp. 23–28.

[24] S. Siersdorfer, E. Minack, F. Deng, and J. Hare, “Analyzing and Predicting Sentiment of Images on the Social Web,” in *Proceedings of the 18th ACM International Conference on Multimedia*, New York, NY, USA, 2010, pp. 715–718.

[25] D. Borth, R. Ji, T. Chen, T. Breuel, and S.-F. Chang, “Large-scale visual sentiment ontology and detectors using adjective noun pairs,” in *Proceedings of the 21st ACM international conference on Multimedia*, 2013, pp. 223–232.

[26] D. Borth, T. Chen, R. Ji, and S.-F. Chang, “Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content,” in *Proceedings of the 21st ACM international conference on Multimedia*, 2013, pp. 459–460.

[27] J. Yuan, S. Mcdonough, Q. You, and J. Luo, “Sentribute: image sentiment analysis from a mid-level perspective,” in *Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining*, 2013, p. 10.

[28] Q. You, J. Luo, H. Jin, and J. Yang, “Robust image sentiment analysis using progressively trained and domain transferred deep networks,” *ArXiv Prepr. ArXiv150906041*, 2015.

[29] D. Ververidis and C. Kotropoulos, “Emotional speech recognition: Resources, features, and methods,” *Speech Commun.*, vol. 48, no. 9, pp. 1162–1181, 2006.

[30] D. Bitouk, R. Verma, and A. Nenkova, “Class-level spectral features for emotion recognition,” *Speech Commun.*, vol. 52, no. 7, pp. 613–625, 2010.

[31] F. Dellaert, T. Polzin, and A. Waibel, “Recognizing emotion in speech,” in *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, 1996, vol. 3, pp. 1970–1973.

[32] R. Tato, R. Santos, R. Kompe, and J. M. Pardo, “Emotional space improves emotion recognition.,” in *INTERSPEECH*, 2002.

[33] M. El Ayadi, M. S. Kamel, and F. Karray, “Survey on speech emotion recognition: Features, classification schemes, and databases,” *Pattern Recognit.*, vol. 44, no. 3, pp. 572–587, 2011.

[34] L. C. De Silva, T. Miyasato, and R. Nakatsu, “Facial emotion recognition using multi-modal information,” in *Information, Communications and Signal Processing, 1997. ICICS., Proceedings of 1997 International Conference on*, 1997, vol. 1, pp. 397–401.

[35] L. S. Chen, T. S. Huang, T. Miyasato, and R. Nakatsu, “Multimodal human emotion/expression recognition,” in *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, 1998, pp. 366–371.

[36] N. Sebe, I. Cohen, T. Gevers, and T. S. Huang, “Emotion recognition based on joint visual and audio cues,” in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 2006, vol. 1, pp. 1136–1139.

[37] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, “A survey of affect recognition methods: Audio, visual, and spontaneous expressions,” *Pattern Anal. Mach. Intell. IEEE Trans. On*, vol. 31, no. 1, pp. 39–58, 2009.

[38] M. Wöllmer, B. Schuller, F. Eyben, and G. Rigoll, “Combining long short-term memory and dynamic bayesian networks for incremental emotion-sensitive

artificial listening," *Sel. Top. Signal Process. IEEE J. Of*, vol. 4, no. 5, pp. 867–881, 2010.
[39] I. D. Addo, S. I. Ahamed, and W. C. Chu, "Toward collective intelligence for fighting obesity," in *Computer Software and Applications Conference (COMPSAC), 2013 IEEE 37th Annual*, 2013, pp. 690–695.
[40] L.-P. Morency, R. Mihalcea, and P. Doshi, "Towards multimodal sentiment analysis: Harvesting opinions from the web," in *Proceedings of the 13th international conference on multimodal interfaces*, 2011, pp. 169–176.
[41] V. Pérez-Rosas, R. Mihalcea, and L.-P. Morency, "Utterance-Level Multimodal Sentiment Analysis.," in *ACL (1)*, 2013, pp. 973–982.
[42] R. Rawassizadeh, B. A. Price, and M. Petre, "Wearables: Has the age of smartwatches finally arrived?," *Commun. ACM*, vol. 58, no. 1, pp. 45–47, 2015.
[43] M. A. Pathak, *Beginning Data Science with R*. Springer, 2014.
[44] "Mining Twitter for Airline Consumer Sentiment | inside-R | A Community Site for R." [Online]. Available: http://www.inside-r.org/howto/mining-twitter-airline-consumer-sentiment. [Accessed: 30-May-2016].
[45] "Vik's Blog - Writings on machine learning, data science, and other cool stuff." [Online]. Available: http://www.vikparuchuri.com/blog/tracking-us-sentiments-over-time-in/. [Accessed: 30-May-2016].
[46] "Python Data Analysis Library — pandas: Python Data Analysis Library." [Online]. Available: http://pandas.pydata.org/. [Accessed: 30-May-2016].
[47] "Welcome to Apache™ Hadoop®!" [Online]. Available: https://hadoop.apache.org/. [Accessed: 30-May-2016].
[48] V. N. Khuc, C. Shivade, R. Ramnath, and J. Ramanathan, "Towards building large-scale distributed systems for twitter sentiment analysis," in *Proceedings of the 27th annual ACM symposium on applied computing*, 2012, pp. 459–464.
[49] "SparkR (R on Spark) - Spark 1.6.1 Documentation." [Online]. Available: https://spark.apache.org/docs/latest/sparkr.html. [Accessed: 30-May-2016].
[50] G. Mone, "The new smart cities," *Commun. ACM*, vol. 58, no. 7, pp. 20–21, Jun. 2015.
[51] "FACT SHEET: Administration Announces New 'Smart Cities' Initiative to Help Communities Tackle Local Challenges and Improve City Services," *whitehouse.gov*, 2015. [Online]. Available: https://www.whitehouse.gov/the-press-office/2015/09/14/fact-sheet-administration-announces-new-smart-cities-initiative-help. [Accessed: 27-Sep-2015].

ISBN: 1-60132-439-1, CSREA Press ©