

Sentiment Awareness on Twitter: Relative Analysis and Training a Consensus Image

V.NAGESH DATTA, Ms.BUSHRA TAHSEEN

M.Tech Student, Assistant Professor

DEPT OF CSE

Dr. K.V.SUBBA REDDY INSTITUTE OF TECHNOLOGY, Kurnool

ABSTRACT

Cloud-assisted mobile health (mHealth) monitoring, which applies the prevailing mobile communications and cloud computing technologies to provide feedback decision support, has been considered as a revolutionary approach to improving the quality of healthcare service while lowering the healthcare cost. Unfortunately, it also poses a serious risk on both clients' privacy and intellectual property of monitoring service providers, which could deter the wide adoption of mHealth technology. This paper is to address this important problem and design a cloud assisted privacy preserving mobile health monitoring system to protect the privacy of the involved parties and their data. Moreover, the outsourcing decryption technique and a newly proposed key private proxy re-encryption are adapted to shift the computational complexity of the involved parties to the cloud without compromising clients' privacy and service providers' intellectual property. Finally, our security and performance analysis demonstrates the effectiveness of our proposed design.

I. INTRODUCTION

THE amount of user-generated content on the web grows ever more rapidly, mainly due to the emergence of social networks, blogs, micro-blogging sites and a myriad of other platforms that enable users to share their personal content. Unlike objective and factual professional publishing, user-generated content is richer in opinions, feelings and emotions. These online expressions can have various practical applications. They have been used to predicted stock market fluctuations [1], book sales [2], or movie's financial success [3]. Due to the vast number of texts, manual inspection for emotion classification is infeasible, hence the need for accurate automatic systems. Although in many cases people can easily spot whether the author of a text was angry or happy, the task is quite challenging for a computer — mainly due to the lack of background knowledge that is implicitly considered by humans.

Given some text, emotion recognition algorithms detect which emotions the writer wanted to express when composing it. To treat this problem

as a special case of text classification, we need to define a set of basic emotions. Although emotions have long been studied by psychologists, there is no single, standard set of basic emotions. Therefore, we decided to work with three classifications that are the most popular, and have also been used before by the researchers from computational linguistics and natural language processing (NLP). Paul Ekman defined six basic emotions by studying facial expressions [4]. Robert Plutchik extended Ekman's categorization with two additional emotions and presented his categorization in a wheel of emotions [5]. Finally, Profile of Mood States (POMS) is a psychological instrument that defines a six-dimensional mood state representation [6]. Each dimension is defined by a set of emotional adjectives, like bitter, and the individual's mood is assessed by how strongly (s)he experienced such a feeling in the last month.

Majority of previous studies predict either Ekman's or Plutchik's classifications, while POMS's adjectives had only been used in simple keyword spotting algorithms [1]. We are not aware of any studies that tackle the problem of predicting POMS's categories from the text. Methodologically, they mainly used simple classification algorithms, like logistic regression or support vector machines, on top of word and n-gram counts, and other custom engineered features (capturing the use of punctuation, the presence or absence of negation, and counts of words from various emotion lexicons) .

Deep learning has recently shown a lot of success in the field of NLP: it has been used for improving sentiment analysis, opinion mining and other tasks like part-of-speech tagging, chunking, named entity recognition, and semantic role labelling. Radford et. al even reported on discovering the sentiment unit while training to generate reviews. Despite these successes, deep learning approaches for emotion recognition are almost non-existent. Previous work generally studied only one emotion classification. Working with multiple classifications simultaneously not only enables performance comparisons between different emotion categorizations on the same type of

data, but also allows us to develop a single model for predicting multiple classifications at the same time.

II. EXISTING SYSTEM

The work on POMS is quite rare, as the test is available only to professional psychologists. Most existing studies were led by Johan Bollen. Common to all is the idea of tracking adjectives defined in the POMS questionnaire and using its structure to obtain six-dimensional mood representation. Bollen investigated how Twitter mood predicts the stock market changes. In a similar study, he correlated emotion time series with records of popular events and showed that such events may have a significant effect on various dimensions of public mood. By analyzing emails submitted to futureme.org, Pepe&Bollen also revealed the long-term optimism of its users, but medium-term confusion. Those studies used POMS questionnaire as a tool for obtaining mood representations but did not study the problem of predicting POMS's categories from the text. There are several studies that use other categorizations of emotions. Neviarouskaya and colleagues developed two rule-based systems for detecting nine Izard emotions; one works on blogs, another one on personal stories from experience project11 website. Mishne experimented with detecting 40 different mood states on blog posts from the LiveJournal community. He used features related to ngrams, length, semantic orientation of words, PMI, emphasized words and special symbols to train an SVM classifier.

III. PROPOSED SYSTEM:

We proposed an alternative training strategy that samples training instances based on the difference between train and validation accuracy and showed that it improves over alternating strategy. We confirmed that it is possible to train a single model for predicting all three emotion classifications whose performance is comparable to the three separate models. As a first study working on predicting POMS's categories, we believe they are as predictable as Ekman's and Plutchik's. We also showed that searching for tweets containing POMS adjectives and later grouping them according to POMS factor structure yields a coherent data set whose labels can be predicted with the same accuracy as other classifications. We made our character-based trained RNN models publicly available at <https://github.com/nikicc/twitter-emotion-recognition>. We worked on probably the largest data set for emotion prediction, using tweets from the last seven years. With the aim of developing a universal

emotion detection algorithm, we did not restrict ourselves only to one domain, but rather tested its usefulness for different classifications of emotions. Since the training data was annotated automatically and since we use character-based approaches, our solution is language independent and could easily be adapted for other languages. Past studies of this problem focused on somewhat different goals and used much smaller collections of tweets, which prevented the use of deep learning and resulted in discouragingly low classification performance. Our study, however, shows that, given enough data, emotion prediction may not be such a hard problem after all. We show that recurrent neural networks, especially character-based ones, can improve over bag-of-words and latent semantic indexing models. Although the transfer capabilities of these models are poor, the newly proposed training heuristic produces a unison model with performance comparable to that of the three single models.

IV. LITERATURE SURVEY

Mohamed et al. surveyed enhancement done in the field of emotion mining along with comparative study for various techniques. Also, this paper presented investigation on the technology utilized in this area and also on the implementation of these techniques. These techniques are categorized into two areas: lexical based techniques and ML based techniques. A descriptive process cycle to build such type of systems is also presented in this paper.

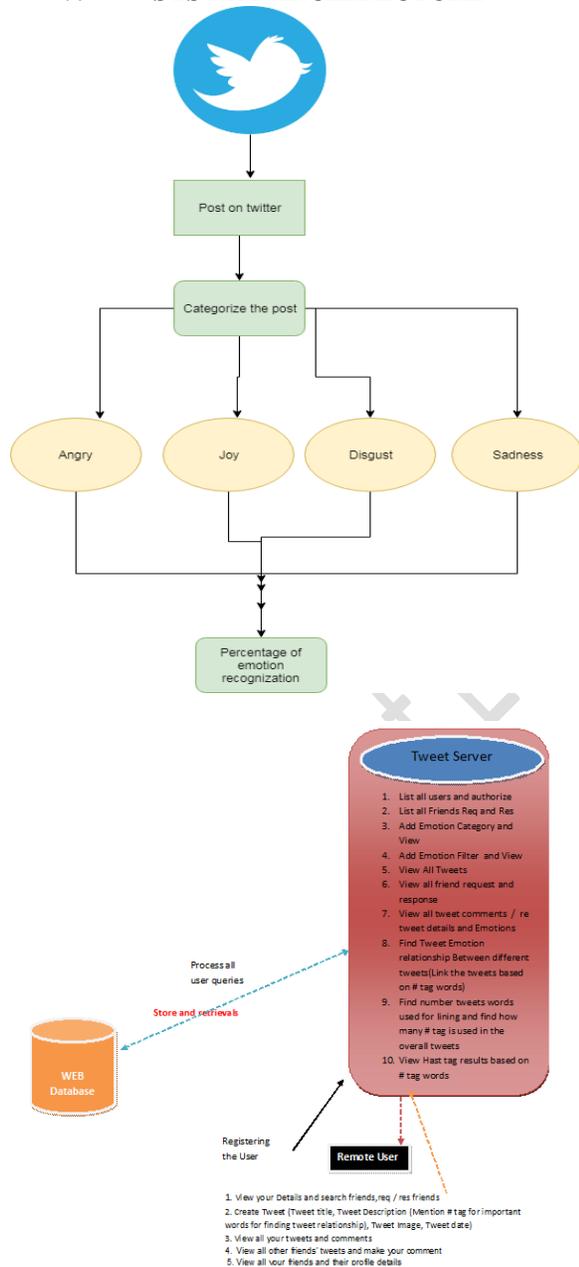
Kumari et al. discuss dissimilar methods and technique to mine emotion from text. Emotion study algorithms are utilize to approximate the public emotion on dissimilar exertion discuss in the common network. Based on text comments author categorized friendship type.

Sharma et al. discuss regarding data mining techniques. This paper offered a survey of the work perform in the region of common network investigation and also focus on the future span in study on common network analysis. This paper presented study related to social networks utilizing Web mining methods.

Dhawan et al. review feeling mining techniques in common networking sites. He stated that latest research shows that it is advisable to deal with various kinds of sentences by various strategies. Also author describe the requirement of some tools to mine certain emotions from different sources of data that gives accurate outcomes.

Wang et al. exploit a technique to automatically annotate a large amount of data. They extracted large amount of tweets (2.5 million) from twitter instead of using already annotated corpus which consists of just thousands of tweets. The main focus of this research is to think about the viability of different element mixes and in addition the impact of the measure of the preparation information on the feeling analysis work.

V. SYSTEM ARCHITECTURE



VI. MODULES DESCRIPTION

Modules:

- **Bag-of-Words & Latent Semantic Indexing Models**
- **Neural Network Models**
- **Transfer Learning**
- **Unison Learning**

Bag-of-Words & Latent Semantic Indexing Models:

To set the baseline performance, we first experimented with common approaches to emotion detection. Within the realm of pure machine learning (as opposed to using, say emotion lexicons), one of the most frequently used approaches is to use simple classifiers on the bag-of-words (BoW) models. We studied two approaches for transforming raw text into BoW model. Vanilla BoW is a model without any normalization of tokens. Normalized BoW reduces the dimensionality of feature space by these transformations. The aim of these normalization techniques is to remove the features that are too specific. For each of these two models, we run experiments on counts of unigrams as well as unigrams and bigrams. Hereafter, we will refer to the combination of unigrams and bigrams simply as bigrams.

Neural Network Models:

Among the most popular neural network (NN) architectures, we decided to use recurrent (RNN) and convolutional (CNN) networks. The former were selected since they can naturally handle texts of variable lengths, and latter since they have already shown to be suitable for text classification [24]. We leave the testing of other neural network architectures, like feed-forward ones, for future work. We experiment with two levels of granularity. In the first approach, we tokenize the tweet's content and then feed a sequence of tokens into the NN. Here the task of the NN is to learn how to combine words to obtain a tweet representation suitable for predicting emotions. Our second setting is an end-to-end learning approach: instead of preprocessing tweets into tokens, we treat the whole tweet as a sequence of characters and pass characters one by one into the NN. The task of the NN is hence to combine characters into a suitable representation and predict emotions. Note that the NN itself has to learn which sequences of characters form words since space is not treated any differently from any other character.

Transfer Learning:

After selecting the best models and their parameters, we test their transfer capabilities and generality. We investigated whether the final hidden state representation — which can be considered as a projection of the tweet’s content into a lower dimensional space — is suitable only for the task for which it was trained or is it sufficient also for predicting other emotion categorizations. We take a model up to the final hidden layer and then re-train the final softmax or sigmoid layer on another data set. In this way, we re-use the embedding from one data set for making predictions on the other. Note that since we are copying weights of one model to the other, we are also forced to use a common model architecture; i.e. the number of neurons, layers, type of layers, number of feature maps, kernel size, etc. The intuition behind these experiments is that if the final hidden state representation can be considered as a general lower dimensional representation suitable for predicting emotions, then the one trained on Ekman might also suffice for predicting POMS’s categories. However, if the performance of such trained model is drastically worse than that of a model initially trained on POMS, this would indicate that final hidden states representations are specifically tuned for particular categorization of emotions.

Unison Learning:

The final set of experiments tests whether it is possible to develop a common model. We define the unison model as a model able of predicting all three emotion classifications while sharing all the parameters that project the input tweet into a final hidden state representation. The utility of such model is at least threefold. First, sharing parameters will hopefully lead to a model whose final hidden state representations are more general. The existence of such hidden state — that could be used to predict various emotional classification — is an indication that there exists a general emotion representation, which could be the starting point for investigating the interdependence between emotional classifications. Second, as is believed for multitask learning approaches, introducing these additional signals during the training of a model could lead to better performance [29]. Finally, when applying such model, we get predictions for all classifications in approximately the same computation time a single model would require for one classification. To build the unison model we propose the following architecture. We have common embedding, followed by a common NN layer. After the final hidden state

representation of the NN, there are three different softmax (for multiclass setting) or sigmoid (for multilabel setting) layers, each predicting one of the three classifications. This architecture, presented in Fig. 3, is learning a low-dimensional embedding that is informative enough for predicting all three categorizations at once.

VII.SYSTEM MODEL

Novel algorithm:

A novel algorithm and architecture are described which have specific application to high performance, digital, adaptive beam forming. It is shown how a simple, linearly constrained adaptive combiner forms the basis for a wide range of adaptive antenna subsystems.

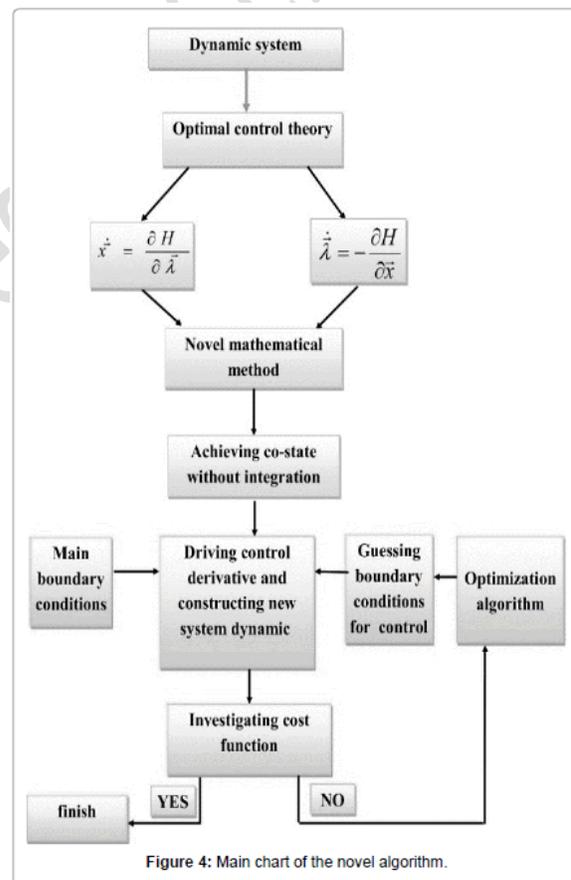
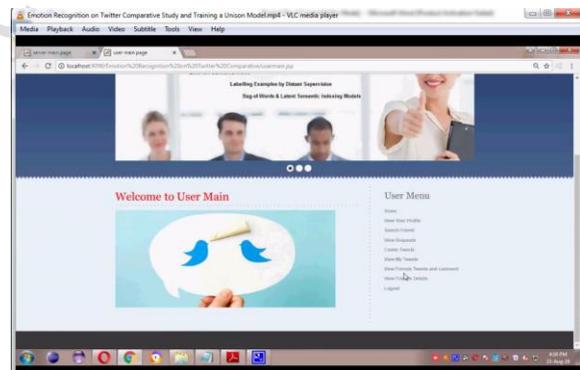
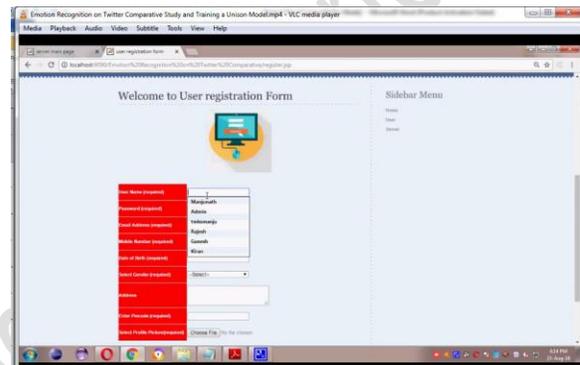
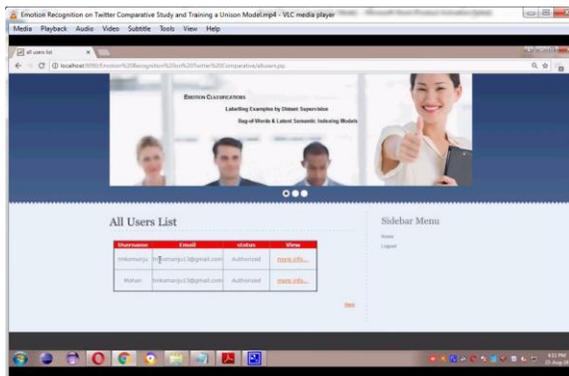
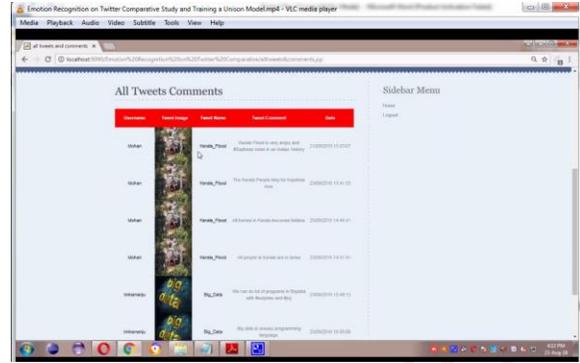
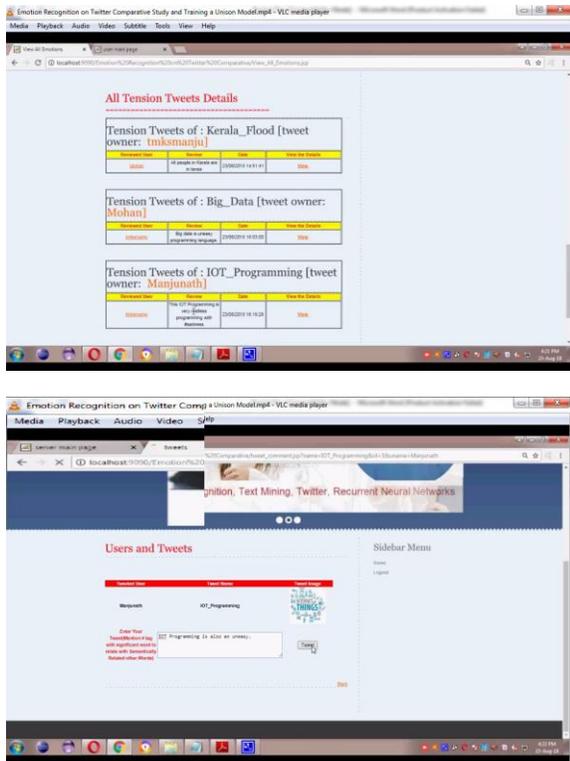


Figure 4: Main chart of the novel algorithm.

VIII. SCREEN SHOTS





IX.CONCLUSION

The central aim of the paper was to explore the use of deep learning for emotion detection. We created three large collections of tweets labeled with Ekman's, Plutchik's and POMS's classifications of emotions. Recurrent neural networks indeed outperform the baseline set by the common bag-of-words models. Our experiments suggest that it is better to train RNNs on sequences of characters than on sequences of words. Beside more accurate results, such approach also requires no preprocessing or tokenization. We discovered that transfer capabilities of our models were poor, which led us to the development of single unison model able to predict all three emotion classifications at once. We showed that when training such model, instead of simply alternating over the data sets it is better to sample training instances weighted by the progress of training. We proposed an alternative training strategy that samples training instances based on the difference between train and validation accuracy and showed that it improves over alternating strategy. We confirmed that it is possible to train a single model for predicting all three emotion classifications whose performance is comparable to the three separate models. As a first study working on predicting POMS's categories, we believe they are as predictable as Ekman's and Plutchik's. We also showed that searching for tweets containing POMS

adjectives and later grouping them according to POMS factor structure yields a coherent data set whose labels can be predicted with the same accuracy as other classifications.

REFERENCES

- [1] J. Bollen, H. Mao, and X.-J. Zeng, "Twitter mood predicts the stock market," *J. of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.
- [2] D. Gruhl, R. Guha, R. Kumar, J. Novak, and A. Tomkins, "The Predictive Power of Online Chatter," *Proc. of the 11th ACM SIGKDD Int. Conf. on Knowledge discovery in data mining*, pp. 78–87, 2005.
- [3] G. Mishne and N. Glance, "Predicting Movie Sales from Blogger Sentiment," *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pp. 155–158, 2005.
- [4] P. Ekman, "An Argument for Basic Emotions," *Cognition & Emotion*, vol. 6, no. 3, pp. 169–200, 1992.
- [5] R. Plutchik, "A General Psychoevolutionary Theory of Emotion," in *Theories of Emotion*. Academic Press, 1980, vol. 1, pp. 3–33.
- [6] J. C. Norcross, E. Guadagnoli, and J. O. Prochaska, "Factor structure of the Profile Of Mood States (POMS): Two Partial Replications," *J. of Clinical Psychology*, vol. 40, no. 5, pp. 1270–1277, 1984.
- [7] S. M. Mohammad and S. Kiritchenko, "Using Hashtags to Capture Fine Emotion Categories from Tweets," *Computational Intelligence*, vol. 31, no. 2, pp. 301–326, 2015.
- [8] S. M. Mohammad, X. Zhu, S. Kiritchenko, and J. Martin, "Sentiment, emotion, purpose, and style in electoral tweets," *Information Processing and Management*, vol. 51, no. 4, pp. 480–499, 2015.
- [9] C. O. Alm, D. Roth, and R. Sproat, "Emotions from text: machine learning for text-based emotion prediction," in *Proc. of Human Language Technology Conf. and Conf. on Empirical Methods in Natural Language Processing*, no. October. ACL, 2005, pp. 579–586.
- [10] B. Plank and D. Hovy, "Personality Traits on Twitter —or— How to Get 1,500 Personality Tests in a Week," in *Proc. of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2015, pp. 92–98.