

Emotion Recognition from Text

Sipra Sahoo¹, Bikram Keshari Ratha², Saroj Kumar Mohanty³

¹Department of Computer Science & Engineering, Siksha 'O' Anusandhan University, India

²Department of Computer Science & Engineering, Utkal University, India

³Department of Computer Science & Engineering, Gandhi Institute for Technological Advancement, India

Abstract: Harvesting knowledge from huge information is a tedious task. Recommender systems play an important role in suggesting to a user. Emotion is a personal feeling of every individual. Various types of emotions include anger, anticipate, disgust, fear, joy, sadness, surprise and trust. This paper details about the machine learning based emotion detection. Various classifiers like Naive Bayes, Support vector machines, Neural networks, and their combinations were used to detect the human emotions. The attained accuracy for each of the method against each of the emotions are provided. In addition to finding emotions, a recommendation approach is proposed based on the emotions from the text.

Index Terms: Classifier, Emotions, Feature extraction, Natural language processing, POS tagging, Recommendation.

I. INTRODUCTION

The growth of information is highly explosive that it makes humans a tedious task to retrieve valuable information. Recommender systems come to handy in this type of situations. Recommender systems can be classified as content based, collaborative based, knowledge based and hybrid. It also has a problem of sparsity or named as cold start problem. Conventional recommendations are performed based on the similarity measures between an experienced customer and a new customer. Recommendations can be entangled with the emotions of an user. Emotion is a strong feeling being derived based on circumstances. The ability which acknowledges emotions is extremely helpful to human-machine communication.

Many sorts of the communication system, such as dialogue system, automatic answering system and human-like automaton, can apply the feeling recognition techniques thus that a user feel as if the system is like human. In addition, the systems can react properly for the human's emotional actions.

The recognition of emotion has been used in several types of media. Some examples are speech, image, signal, facial expressions, textual in formation, and so on. Among this, the textual information is terribly fashionable medium, consisting of books, newspapers, and letters. Text is the simplest and main tool for the human to speak thoughts, convey messages and express aspiration. In addition to the range and complexity of matter information build it potential for folks to exchange ideas, opinions, and emotions using text solely. For these reasons the research for recognizing from the matter information is valuable. Traditional researches tried to acknowledge feeling victimization emotional keywords. If input sentence have emotional keywords, we apply the keyword-based approach to the system since the keyword-based approach shows high recognizing accuracy for emotional keywords. The advancement in the field of computer sciences and information technology has multiplied the probabilities of interaction between humans and computers. The most conventional medium for the world to communicate with computers is still via texts. As the internet emerges, more and more people are maintaining blogs to share their feelings/opinions with the public [1]. Even Though sharing pictures and videos has becoming popular, texts and blog articles still stand as a major medium for expressing opinions [2]. When analyzing text, detecting emotions such as joy, sadness, fear, anger, trust, surprise, disgust and anticipate is useful for variety of functions, including distinctive blogs that categorical specific emotions towards the topic of interest, identifying what feeling a newspaper headline is making an attempt to evoke, and devising dialogue systems that respond fittingly to totally different emotional states of the user. Often different emotions are expressed through different words. For example, delightful and pleasant-tasting indicate the feeling of joy, gloomy and cry are indicative of unhappiness, shout and boiling are indicative of anger, and so on. The main objective of this article is to create an Intelligent system that helps in recognizing emotions from text document or any other textual form that can be used for recommendations. For a given input text, this system would be able to detect the class of emotion of the input text. The emotions expressed through text could be very complex, Since the main intention is to extract emotion from text for recommendations, eight basic classes of emotions that are generally used in any kind of human interactions are deployed here. This paper consists of five sections. Section I gives the introduction of the paper. Section II gives the literature survey. Section III explains the methods and resources we used in collection of datasets. Section IV shows the step by step procedure that has been done to develop an emotion detection system. This section also describes a method for emotion based

recommendation. Experiments conducted and the results are presented in Section V. Finally, the conclusions are given in Section VI.

II. RELATED WORK

Emotions are great values of Mankind. These emotions could be of positive type (like happiness, joy, love etc.) or it could be of negative type (like fear, anger, disgust, etc.). Studying such incredible emotions by machine is always a challenging task. Affective computing is the study of interpretation and simulation of human emotion to computer has been implemented by different kind of media such as speech, image, body gesture. However, recognizing emotion from textual context is consider to complex task till today. The varieties of language and variety grammar make the task of extracting emotion from text tougher than any other modalities.

Natural Language processing (NLP) comes in handy for smooth interaction between computers and human language [3]. NLP could help machine to understand the language issues such as textual syntax, Semantics, pragmatics, discourse. Ambiguities are also resolved using the rule-based or statistical Methods. Though the language issues could be overcome by NLP, there is an intelligence required to extract the abstract information from the text. In this case abstract information is emotion. For recommendation, one of the important features that can be used are human emotions. Extracting the emotion and determining its type is the place where the intelligence have the role to play.

Text mining, sometimes alternately referred to as text data mining, roughly equivalent to text analysis [4], refers to the process of deriving high-quality information (abstract information) from text. Association, Classification, Prediction, clustering, etc., are the different approaches that help to derive high-quality information. In our case emotion in text are the high-quality information and their types are the different classes that need to be determined. The practical feasibility of extracting emotion from text is possible by using NLP and text mining techniques. It is worth noting that these techniques have certain limitations. Accuracy varies depending on application domains. The rules formed for extracting information from one language will not be useful for other languages. Rules and data dependency are dependent on languages.

Optimized class selection (types of emotion, considering too many types of emotion will decrease the accuracy) could be the good solution to stable the accuracy of the extraction system.

Researchers in the field of affective computing have investigated recognition, interpretation and representation of affect [5]. They considered an extensive variety of modalities. For example, influence in discourse, facial showcase, stance and physiological movement [6]. Due to the expansive volume of content information accessible on the internet, online journals, email and visits which are brimming with feelings, as of late there has been a developing enthusiasm for programmed and extraction of slant, suppositions and feelings in content. Moreover, printed information on the web take up minimal physical space and are effectively exchanged, so they have a high potential to be utilized for sharing thoughts, supposition and feelings. It is likewise such a dynamic region of examination since its applications have spread to different spaces, from purchaser item surveys, human services and budgetary administrations to get-togethers and political races In order to recognize and analyze affect in written text seldom explicitly marked for emotions NLP researchers have come up with a variety of techniques, including the use of machine learning, rule-based methods and lexical approach [7]. Detecting emotions, however, is a very popular concept in social network analysis[8]. Emotion/stimulus interaction, an effective relation, potentially yields crucial information in terms of information extraction. For example, we can predict future events or decide on the best reaction if we know the emotion cause [10].Event-based emotion detection has been addressed in earlier research [11], [12]. They explore emotions and their causes, focusing on five primary emotions happiness, sadness, fear, anger and surprise in Chinese texts. They constructed a Chinese emotion corpus, but they focus on explicit emotions presented by emotion keywords. In their dataset, they observe that most emotions appear within the same clause of there presentation of the emotion, so a clause might be the most appropriate unit to detect an emotion. We find such granularity too large to be considered an emotion stimulus in English. Also, clauses were distinguished by punctuation: comma, period, question mark and exclamation mark. Just four punctuation marks are not enough to capture English clauses adequately. There is no significant past work on the usage of emotions as a feature for recommendations.

III. DATA COLLECTION

To predict the emotions, corpus is required. Proper and appropriate corpuses have to be collected for all eight individual classes in the classifier system. Dataset collection have got more priority in terms work proceeding. It is because it reflects the in the accuracy results of classification if the any compromising had happened in dataset collection. Dataset collection should pave path for the classifier principle of high inter class similarity. It was difficult to find English string that express emotions particularly. A bird view

has to hold to capture these target strings. For emotions classes such as joy, fear etc. their wear plethora of English strings available and corpus collection for these classes has been finished without much difficulties. At the same, it was difficult to find string for few emotions class such trust, disgust etc. It is because that an average human does express these emotions generally and even if express it would be combination with other emotions so it was difficult to find corpus for these emotions.

The only rule followed in corpus collection is it should proper and appropriate to subject emotional classes of the emotion classifier. The main focus is on the statements and the hunting on various blogs sites articles, also focused for famous quotes, well-known poems, etc. Then search area included the social networking sites like twitter, Facebook, where chats that shared among the friends are collected, in a new dimension search were benign on emotional news headlines on different Internet sources (like CNN, Google News, WSJ and twitter) [13] [14] and from newspapers. The system of identifying emotional expressions is available in different granularities such as word, phrase, sentence and document. The emotions in these granularities are mostly shown by parts of speech. Noun, Verb, Adverb, Adjective. are some of the parts of speech that denotes the emotional characteristic in relatively good quantity [15] [16]. Thus we target these parts of speech to retrieve the emotions. These keywords act as the data set for the classes which is further processed by classifier to give to emotions in the input corpus. The diagram shows the parts of speech frequency for emotion words. The corpus consists of sentences which are primarily collected from blogs, user status, comments, news headlines, SMS drawn from various websites, newspapers such as Google News, twitter, Facebook, web dictionaries, news websites, SMS websites as well as various blogs pages. This data set has high load of emotional content. The total corpus was divided into 9:1 ratio of datasets for training data set corpus and test dataset corpuses. All the emotions sentences searched from various websites, blog post and daily life used sentences. In total, we collected 2000 sentences including all the 8 broad classes of emotion that we chose to detect from a text. on a whole, 1800 sentences for training and 200 sentences as testing dataset.

This emotion detection system consists of two phases, training phase and testing phase. Training phase extracts the keyword from the emotion training data and translate the keywords (the words of natural language) into the format that is processed by the classifier to generate the predicting model. In the testing phase, it extracts the keywords from the sentence that was input from the keyword and then translates the keyword into the format that can be processed by the classifier and compare the output by the predicting model of the training phase and finally show the result. Fig. 1. shows the model.

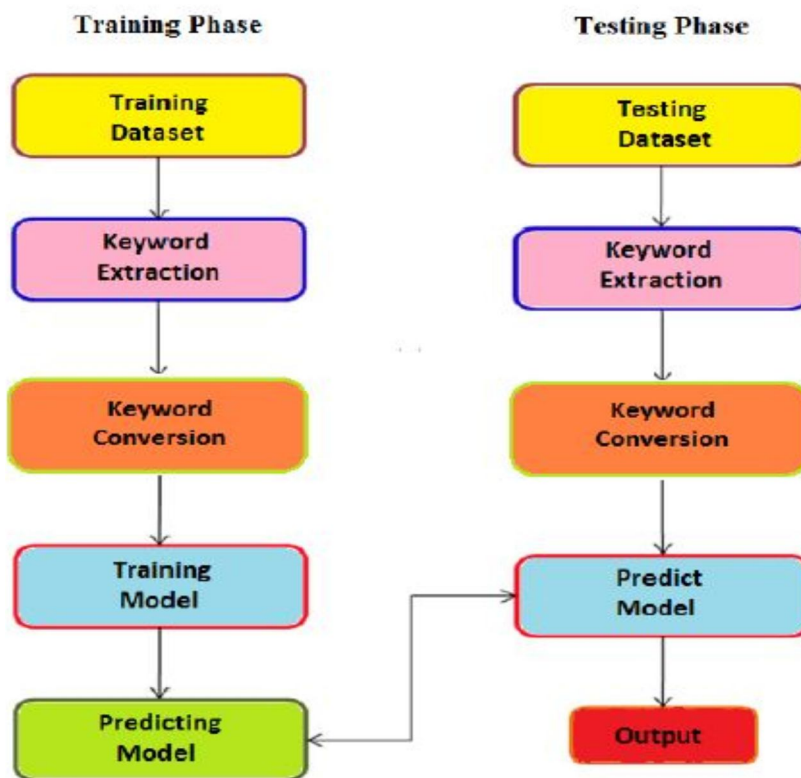


Fig. 1. Predicting Model.

The training phase consists of five important parts: The training dataset, the keyword extraction, the keyword conversion, the training model and the predicting model. These keywords act as the data set for the classes which is further processed by classifiers to give to emotions in the input corpus. The diagram shows the parts of speech frequency for emotion words. A lot of words in a sentences does not help in recognizing the emotions, these are words like pronouns, connectors, etc. If these kind of words are removed before the actual processing of the sentences, then computational time would be reduced to a great extent. This can be done by extracting only the useful words or we prefer to call them as keywords, from the sentences of the document. On our investigation it has been found that noun, verb, adverbs, adjectives are the key items that are more likely to determine emotion than other parts of speech.

In our research study, we have used eight emotion classes (fear, anger, joy, trust, surprise, sadness, disgust and anticipate) that exists in the English literature.

The following table (Table 1.) shows different classes of emotions and their associated POS tags. After the required dataset collection, those sets have to be inputted to training phase of classifier. Training phase consists of two modules: keyword extraction and keyword conversion. Here we cannot input the sentences directly for training the classifier. So, at first, we need to separate the key terms that are useful for classifier from the collected dataset. The keywords here mean the parts of speech which can detect emotion expressed by the text. In order to extract them, the dataset is subjected to a POS tagger(Part-of-speech tagging) and separation from the useless part of the text.

We used POS Tagging concept to tag the words, phrases, sentences and the documents. Since we target only Noun, verb, adjectives and adverb as the keywords to extract the emotion, so tagger play an important role [17]. POS Tagger is marking up the words in a text as corresponding to a particular POS based on its definition as well as its context. Fig. 2. shows different POS taggers that were used.

Once the input sentences or documents get tagged then we need to target the keywords (noun, verb, adjective and adverb) which are useful in retrieving the emotions from the sentences and documents. If there are any other words such as stop words, it will be eliminated from getting extracted. Thus it prevents the system from useless manipulation that saves the time. Subsequently we have the targeted keywords which are stemmed to get the root words from the extracted words.

IV. PROPOSED RECOMMENDATION METHOD

This emotion detection system consists of two phases namely, training phase and testing phase. Training phase extracts the keyword from the emotion training data and translate the keywords (the words of natural language) into the format that is processed by the classifier to generate the predicting model. In the testing phase, it extracts the keywords from the sentence that was input from the keyword and then translates the keyword (the word of natural language) into the format that can be processed by the classifier and compare the output by the predicting model of the training phase and finally shows the result.

The training phase consists of five important parts: the training dataset, the keyword extraction, the keyword conversion, the training model and the predicting emotions [4]. These keywords act as the data set for the classes which is further processed by the classifier to give to emotions in the input corpus. A lot of words in a sentences does not help in recognizing the emotions. Some of these are pronouns, connectors, etc. If these words are removed before the actual processing of the sentences, then the computational time would be reduced to a great extent. This can be done by extracting only the useful words or keywords, from the sentences of the document. We observed that noun, verb, adverbs, adjectives are the key items that are more likely to determine emotion than other parts of speech. There are different emotions exist. In this study, we used only eight emotion classes (fear, anger, joy, trust, surprise, sadness, disgust and anticipate) that exists in the English literature. After collecting the required dataset, they have to be provided to the training phase of the classifier. Training phase consists of two modules: Keyword extraction and Keyword conversion. Here we cannot input the sentences directly for training the classifier. So, we need to separate the key terms that are useful for classifier from the collected dataset. The keywords here take the form of a Noun, Verb, Adverb, Adjective which can detect emotion expressed by the text. In order to extract them, the dataset is subjected to a POS tagger(Part-of-speech tagging) and separate the keywords from the useless part of the text. We used Penn Tree Bank tagging to tag the words. As the target was only Noun, verb, adjectives and adverb as the keywords to extract the emotion, words in a text that were extracted using POS information is based on its context. Existing key tags are shown in the figure 2.

TAGS	REPRESENT	EXAMPLES
/JJ	adjective or numeral, ordinal	multilingual multi-disciplinary
/JJR	adjective, comparative	calmer cheaper choosier cleaner clearer closer colder
/JJS	adjective, superlative	calmest cheapest choicest
/RB	Adverb	occasionally unabatingly adventurously professedly
/VB	verb, base form	ask assemble assess assign assume atone attention avoid bake
/VBD	verb, past tense	wore pleaded swiped soaked
/VBG	verb, present participle	stirring focusing judging
/VBN	verb, past participle	chaired used imitated
/VBP	verb, present tense, not 3rd person singular	wrap resort sue twist spill
/VBZ	verb, present tense, 3rd person singular	bases seals pictures authorizes

Fig. 2. Types of tags with their meanings and examples.

Here, we used a machine learning tagger for which a standard data set shared in a conference on Computational Natural Language Learning which is used for training conditional random fields model for detecting the tags in the sentences provided through test datasets. The precision, recall and f-measure for all the keyword tags are mentioned in the Fig. 3.

Tags	NO. OF TAGS(n)	MODEL (m)	CORRECT (c)	PRECISION (m/c)	RECALL (c/n)	F1-measure (%)
/JJ	2948	2926	2784	95.14	94.43	94.78
/JJR	155	159	143	89.93	92.25	91.07
/JJS	90	83	79	95.18	87.78	91.33
/RB	1463	1471	1373	93.34	93.85	93.59
/VB	1178	1198	1143	95.41	97.03	96.21
/VBD	1518	1518	1455	95.85	95.85	95.85
/VBG	756	772	709	91.84	93.78	92.62

Fig. 3. POS Information.

Once the input sentences are tagged, we need to retrieve the keywords (noun, verb, adjective and adverb) which are useful in retrieving the emotions from the sentences and documents. If there are any other words such as stop words, it will be eliminated from getting extracted. Thus it prevents the system from useless manipulation that saves time [8]. This can be done using the class extractor. After this process, we have the targeted keywords which are stemmed to get the root words from the extracted words. Stemming is the process that is used to reduce the extracted word to its root word [5]. This is useful for query expression o indexing and other natural language processing problems. By stemming, we remove the suffix as well as the prefix from the words to get its root value. Ex: The stemming reduces the words "hashing", "hashed", and "hasher" to the root word, "hash". All the datasets collected and preprocessed to be an input for a classifier. Many researches has proved that Naive Bayes classifier gives preferred option [9].The Naive Bayes algorithm is based on conditional probabilities. It uses Bayes' Theorem, to that calculates a probability by counting the frequency of values and combinations of values in the historical data. The formula is given below.

$$P(c|d) = \frac{P(d|c)P(c)}{P(d)}$$

Bayes' Theorem finds the probability of an event occurring in the testing phase given the probability of another event, i.e. training phase that has already occurred [9]. Naive Bayes calculates a probability by dividing the percentage of pair wise occurrences by the percentage of singleton occurrences. If these percentages are very small for a given predictor, they probably will not contribute to the effectiveness of the model. Occurrences below a certain threshold can usually be ignored. The NB Classifier accepts some specific format from converter and then generates a "Model" file. This file acts as the storage format for each class. As training the data is a time-consuming process, we train the data first and store the result enabling the predict operation to go faster. When we enter the new data for testing, the NB Classifier compares the new dataset with stored model file.

When the user input any emotional sentence, firstly the testing phase extracts the keywords from the sentence and then converts the keyword into the numerical format so that it can be processed by the classifier and generate the model file which is used to compare to the model file generated during the training phase.

V. RESULTS

The testing was conducted on a large set of input sentences to determine the emotion that it holds. These sentences use the tool within the application and finally give the emotions that a sentence exhibits.

The input sentences given during the testing phase under goes various processing steps in various classes and finally gives the output as the type of emotion class that the input sentence exhibits from its keywords. To represent the output, we have created a frame that takes the input from the users and undergo internal processing. During the internal processing the input sentences goes to the Reader, Extractor, Tagger, Stemmer and Convertor classes. It shows the number of keywords extracted from the sentence after tagging and stemming it into the root form. After the internal processing gets over, the system gives us the type of emotion in the sentences. To represent the output in better way, we provide the picture that represent to the resultant emotion in the sentences. Thus every emotion category has a specific picture by which it is represented as the output.

The testing phase is done by providing various input sentences of different classes and getting the result. The individual accuracy for each emotional class is given in Fig. 4.

Emotions	NO. OF TAGS(n)	CORRECT (c)	WRONG	ACCURACY (%)
Anger	25	21	4	84
Anticipate	25	21	4	84
Disgust	25	14	11	56
Fear	25	22	3	88
Joy	25	24	1	96
Sadness	25	20	5	80
Surprise	25	5	20	20
Trust	25	23	2	92

Fig. 4. Results of the developed system using Naive Bayes Classifier.

The classifier takes the input that is given by the Convertor class and generates the Predict model for the input sentence. The generated Predict model in the Testing phase is compared by the Predicting model of the Training phase. This comparison decides that the input sentence belongs to which emotional class. The individual accuracy for each emotional class is given in Fig. 5.

Emotions	SVM	NB	NN	NB+SVM	NB+NN	NB+NN+SVM
Anger	76	84	60	84	76	76
Anticipate	84	84	84	84	84	84
Disgust	36	56	92	56	76	76
Fear	60	88	76	88	88	88
Joy	80	96	92	96	96	96
Sadness	88	80	80	80	80	80
Surprise	40	20	40	20	20	20
Trust	92	92	88	92	84	84

Fig. 5. Results of the developed system using different Classifiers.

VI. CONCLUSION

Extracting emotions form the text has been always an interesting research work. A number of steps are involved in this process. In the testing phase, we deployed a number of classifiers namely, Naïve Bayes, support vector machines, neural networks, ensemble classifiers (NB + SVM, NB + NN, NB + NN + SVM) to evaluate the emotions. Their accuracy measures shows that combined classifier NB + NN shows better accuracy than other individual or combination of classifiers.

REFERENCES

- [1] Strapparava, Carlo, and Rada Mihalcea. "Learning to identify emotions in text." In Proceedings of the 2008 ACM symposium on Applied computing, pp. 1556-1560. ACM, 2008.
- [2] Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis." Foundations and Trends® in Information Retrieval2, no. 1-2 (2008): 1-135
- [3] Collobert, Ronan, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. "Natural language processing (almost) from scratch." Journal of Machine Learning Research 12, no. Aug (2011): 2493-2537
- [4] Tan, Ah-Hwee. "Text mining: The state of the art and the challenges." In Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases, vol. 8, pp. 65-70. sn, 1999

- [5] Tao, Jianhua, and Tieniu Tan. "Affective computing: A review." In International Conference on Affective computing and intelligent interaction, pp. 981-995. Springer, Berlin, Heidelberg, 2005
- [6] Cambria, Erik. "Affective computing and sentiment analysis." IEEE Intelligent Systems 31, no. 2 (2016): 102-107
- [7] Haji binali, vidyasagar potdar. "Emotion detection: state of the art", Proceedings of the CUBE International Information Technology Conference, ACM, New York, NY, USA, 2012 .
- [8] Stieglitz, Stefan, and Linh Dang-Xuan. "Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior." Journal of management information systems 29, no. 4 (2013): 217-248.
- [9] Edward chao-chun kao, ting hao, chang-tai, von-wun. towards text-based emotion detection, 2009 international conference on information management and engineering
- [10] Thalberg, Irving. "Constituents and causes of emotion and action." The Philosophical Quarterly (1950-) 23, no. 90 (1973): 1-14.
- [11] Devillers, Laurence, Laurence Vidrascu, and Lori Lamel. "Challenges in real-life emotion annotation and machine learning based detection." Neural Networks 18, no. 4 (2005): 407-422
- [12] Rellecke, Julian, Marina Palazova, Werner Sommer, and Annkathrin Schacht. "On the automaticity of emotion processing in words and faces: event-related brain potentials evidence from a superficial task." Brain and cognition 77, no. 1 (2011): 23-32.
- [13] Radim BURGET and Jan KARSEK, Zdenk SMKAL, Recognition of Emotions in Czech Newspaper Headlines RADIOENGINEERING, VOL. 20, no. 1, 2013
- [14] Pak, Alexander, and Patrick Paroubek. "Twitter as a corpus for sentiment analysis and opinion mining." In LREc, vol. 10, no. 2010. 2010.
- [15] Olutobi Owoputi, Brendan, Chris, Kevinl, Nathanr, Noah. Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 380-391, 2011
- [16] Wilson, Theresa, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. "OpinionFinder: A system for subjectivity analysis." In Proceedings of hlt/emnlp on interactive demonstrations, pp. 34-35. Association for Computational Linguistics, 2005.
- [17] Cecilia Ovesdotter Alm, Dan Roth, Richard Sproat.: Emotions from text: machine learning for text-based emotion prediction. Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language.
- [18] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011.



Sipra Sahoo is working as an assistant professor in ITER, SOA University. She was born on 6th May 1984. She has completed her B.E (Information Technology) from BPUT in 2005 and M.E.(Computer Science and Engineering) in 2008 from Utkal University. and currently pursuing Ph.D in Utkal University. Her areas on interest include data mining, recommendation systems, soft computing, web personalization and sentiment analysis. She is a member of Institute of Engineers (India).

Sipra is currently working as an assistant professor in SOA university in the department of Computer Science and Engineering since 2009. Prior to this job she has worked in MIET and NIST during 2008 to 2009.



Bikram Kesari Ratha was born on 20th May 1965. He completed his MCA (Tech) from NIT Rourkela in 1988 and subsequently completed his ME (CSE) from the same institute. He completed his Ph.D from Utkal university. He has teaching experience over 25 years in several programmes like B.E., MCA, MSc, ME, M.Tech in different universities and institutes.

Dr. Bikram has worked as a professor in Nepal Engineering College (a constituent college of Pokhara university, Katmandu, Nepal) from 2000 to 2005. Later he has worked as a Professor in SRTM University. He has published more than 50 research papers and delivered many talks on different areas of computer science. His research interests include software engineering, data mining, ICT applications, computer networks and image processing.