International Journal of Pure and Applied Mathematics
Volume 119 No. 12 2018, 13297-13308
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

Hybrid Sentiment Classification of Reviews Using Synonym Lexicon and Word embedding

Sankar. H, Subramaniyaswamy. V*
School of Computing, SASTRA Deemed University, Tanjore, Tamil Nadu, India
***Corresponding Author**

Abstract

Sentiment analysis is used in extract some useful information from the given set of documents by using Natural Language Processing (NLP) techniques. These techniques have wide scope in various fields which are dealing with huge amount of data link e-commerce, business and market analysis, social media and review impact of products and movies. Sentiment analysis can be applied over these data for finding the polarity of the data like positive, neutral or negative automatically or many complex sentiments like happiness, sad, anger, joy, etc. for a particular product and services based on user reviews. Sentiment analysis not only able to find the polarity of the reviews. Sentiment analysis utilizes machine learning algorithms with vectorization techniques based on textual documents to train the classifier models. These models are later used to perform sentiment analysis on the given dataset of particular domain on which the classifier model is trained. Vectorization is done for text document by using word embedding based and hybrid vectorization. The proposed methodology focus on fast and accurate sentiment prediction with higher confidence value over the dataset in both Tamil and English.

Keywords: Sentiment analysis, Natural Language Processing, HWW2V

I. Introduction

Every day, millions of people share their view, opinions and reviews about movies and products on various social media site like Facebook and Twitter, e-commerce sites like Amazon and movie reviewing sites like rottentomato or IMDb.

These reviews and opinions may hold some of uses expectation which is important to business and marketing professionals and researchers.

Automated sentiment analysis and opinion mining is extraction of useful information from the text data on social media, forums, wikis, etc. contacting informations about people’s opinions. This type of analysis can be performed in various levels on the input source materials ranging from separate words, sentences or entire document at stretch. The sentiment analysis use to categorization these data can be characterized in to either positive, negative or neutral.

Sentiment analysis processes topically uses Natural Language Processing (NLP) processing like POS tagging, stemming, etc. and some other resource such as dictionary, thesauri and sentiment lexicons to model the documents. Extraction of important features from the document are said to be successfully sentiment detection. Then, these sentiments are characterized based in their polarity as positive, negative or neutral. Various methods can be used for identifying the sentiment of the document using either by using supervised or unsupervised learning. In supervised method use some machine learning algorithms where the sentiment detection is considered as a classification problem. SVM (Support Vector Machine), Neural Networks or Naive Bayes (NB) [20]are some algorithms which are commonly used for sentiment classification. On the other hand the unsupervised approach uses lexical resources are employed to provide the sentiment scores to words to detect the overall sentiment of a document[1].

The recent and upcoming method of sentiment detection approaches uses word representation on semantic vector spaces which makes use of neural networks or probabilistic models over huge text corpora [21-30]. This approach of sentiment analysis captures the sentiment much accurately along with the semantic and context relation between the documents [31-40].

Lexicon based approach are used for extracting the features which contain information about overall polarity of the entire document [41-45]. This type of approach some times suffers low coverage because in some cases documents may not contain any words present in the lexicon. We can find this type of situation while analyzing documents contain short review. This approach fails to capture the sentiment much efficiently. On the other hand, the word embedding based approach represents documents in the form of vectors which is used to successfully capture the semantic and syntactic similarities of a document in its written language. Unlike lexicon based approach this approach does not use any sentiment score or emotional score for words present in the documents.

In this work we presented a generic sentiment analysis approach used for multiple languages because the single language based methods limits the possibilities of using sentiment analysis in industrial applications. Thus, this approach satisfy performance in multiple language, hence this method can be applied in various language in feature.

Lexicon based approach suffers from low coverage in features that convey information on overall document. This phenomenon can be witnessed which performing sentiment analysis on short text snippets such as tweets and user reviews. This may result in failure of identifying the emotions and sentiments. On other hand, supervised word embedding based approach successfully capture the semantic and syntactic regularities which are generally found in the normal written language but fail to make use of sentiment and emotion values of each and every words present in that document.

In this work, the method is used to determine the polarity of documents, under the point of view put forward that increases the strength of lexicons along with word embedding models which will improve the performance of the classifier model. This methodology derives the features from the documents by using both lexicon based (unsupervised) and word embedding based (supervised) approach which is combined to gather to form a hybrid vectors for concise and clearly expressed form of document representation.

Sentiment analysis algorithms employs various NLP processing techniques like stemming, POS tagging, chinking, chunking, etc. along with some additional feature such as thesauri, sophisticated dictionaries, ontologies and emotion-based lexicons to train the classifier model using obtained data or documents. Once the model is created the sentiment analysis can be perforated on the text document based on their polarity as positive, negative or neutral. There are two type of techniques which are available to sentiment analysis which may be unsupervised or supervised sentiment analysis. Unsupervised sentiment analysis employs lexicon based approach. Supervised sentiment analysis employs the machine learning algorithms such as SVM (Support Vector Machines), Deep Learning, Naive Bayes, KNN (k-Nearest Neighbor) or Neural Networks.

2. Problem statement

In sentiment analysis the vector representation of document is said to be efficient, when such algorithm satisfies any one of the two properties:

a) The first property is to capture the semantic relation between documents or paragraphs in terms of Euclidean space.
b) Second property is to identify the sentiment polarity in terms of positive, negative or neutral or by assigning a sentiment score based on the sentiment

polarity of the document.

But unfortunately, the approach which satisfies the first property fail to capture the sentiment. On the other hand the approach which satisfies the second properties fail to capture the semantic similarity of the between two documents or paragraphs. The main idea of our approach is to combine both the beneficial features properties for sentiment analysis.

The first property improves the capability of any machine learning algorithm model to classify the documents in terms of polarity or emotions based on its contents. For instance the Bag-of-Words approach2, represents each document based on the frequency of each words, irrespective of the order and often consider vocabularies which is present in the dictionaries. By comparing term frequency profile the semantic similarity between documents can be determined. However, the Bag-of-Words (Bow) fails to capturer the synonyms. For example, the sentence: “He came late to party” and “He delayed coming to party” although it has different representation both the sentience convey the same meaning. The Word2Vec[8] is another model, this model tries to satisfy the limitation by vector representation of words such that the words which are frequency appearing in similar contexts are represented closely in the Euclidean space.

The second property, the property of identifying the sentiment can be achieved, with help of the lexicon where the scores are given according the to their sentiment orientation. In earlier days the lexicon based approach is used for sentiment analysis, called lexicon based sentiment analysis[7], but this method suffers lack of coverage, because many sentiences don't content the words which are present in the lexicon and thus this sentiment approach is can’t be evaluated. To resolve the disadvantage of both the approach, we are going to combine the representation.

Based on the above discussion we have proposed a new Hybrid Weighted Word2Vec (HWW2V) document representation approach. The HWW2V is an integration of:

a) Bag-of-Words representation
b) Weighted Word2Vec representation, and
c) Sentiment lexicon based representation

3. Methodology and approach

Sentiment Classification techniques are classified in to lexicon based, machine language approach, lexicon based and hybrid approach. The Machine learning(ML) approach uses some common ML algorithms and linguistic features. Lexicon based approach makes use sentiment lexicons. This technique is sub-classified corpus based and dictionary based method that uses statistical approach to determine the sentiment of the sentence. Hybrid approach is a combination of both lexicon based and ML based approach.

The Machine Learning based approach can be classified into supervised and unsupervised learning approach. Supervised learning approach use a labeled training documents for training the model. In unsupervised learning approach it is difficult to identify label in training documents to train the model.

The lexicon based approach is subdivided into two approach dictionaries based and corpus based approach. Dictionary based approach find the keyword and then search for synonyms and antonyms for that keyword. The lexicon based approach begins with tokenization of the document and extracting the sentence which holds some sentiment value by neglecting the stop word from the document.

3.1. Machine learning approach

Machine learning is the one of the approach of the sentiment analysis which relies on some famous machine learning algorithm to classify the text based on its polarity with help of syntactic and(or) linguistic features.

Text classification problem: Given a suitably annotated collection of document. The classification model relate the feature in to give document to any of the classification label. Such labels can be either nominal (like positive, negative, happy, anger or neutral) or real valued (like 0.5 sad and 0.8 anger). Then given a suitably unannotated collection of document, the model predict a classification label for it.

3.1.1. Supervised Learning algorithm

The supervised learning algorithm depends on the labeled training data. The training data contain a collection training examples. Each example in supervised learning to consist of an input object and output value. In the following subsection we are presenting a brief explain some of the most commonly used classification algorithms for Sentiment Analysis.

3.1.1.1. Probabilistic classification algorithm

Probabilistic classifier[18] is a classifier that predict, given an input, a probability distribution over a set of classes and output the class that's most likely to the given input observation. This classifier type of classifier is also called as generative classifiers.

$$y = f(x)$$

3.1.1.2 Decision tree classifier

Decision tree classifier a method which is commonly used in data mining. The main idea behind the decision tree is to predict the outcome based on input variables. Decision tree classifiers[17] train itself by decomposing the data with help of condition on the attribute values. The decomposition of the data will be continued until the leaf nodes of the tree contain certain minimum number of records which can be later used for classification.

3.1.1.3 Rule based classifiers

In rule based classification the data model is trained based on rules. The left-hand side holds the condition of the rule and right-hand side holds the class label.

$$(Condition) \rightarrow (y)$$

Where y is the class label, LHS rule antecedent or condition and RHS: rule consequence. There are various criteria availability to generate rules, The two most common criteria are confidence and support. The support represents the number of instance which are relevant to rules present in the training datasets. The confidence refers to probability that the RHS of the rule satisfies the LHS. DT has strict hierarchical partitioning, but in rule based classifies allow overlapping the decision space. Quinlan[3] has studied the decision rules and decision tree within a single framework; some branches of DT can be considered as a rule for classifying the text instance.

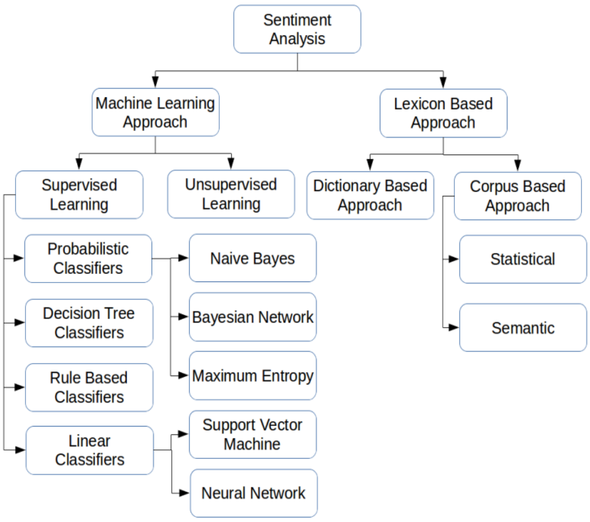

Fig 3.1. Sentiment Classification Techniques

3.1.1.4 Liner classification algorithm

Linear classifier are statistical classifiers, which use an object characteristics to predict the which group or class it belongs to. It performs classification decision base on combination of linear characteristics. Linear predictor$p = \overline{A} \cdot \overline{X} + b$, where $\overline{X}$is the document frequency of the words, $\overline{A}$is the vectors of linear coefficients. SVM[15] is one of the most preferred algorithm for document classification or SA due to it accuracy and ease of training to model. In the below subsection we discussed two line classification algorithms which are famous.

3.1.2 Unsupervised Learning Algorithms

Unsupervised learning method use a different approach for sentiment analysis. They don't use the labeled documents unlike supervised learning. Instead, they make use of the document's statistical properties such as word co-occurrence, Natural Language Processing, existing lexicons with emotional (or) polarized words. Lin and He[12] have proposed a method that makes use of latent dirichlet allocation[5] to identify the sentiment of a document simultaneously. The proposed method has an accuracy close to supervised methods.

There are also methods in unsupervised learning that depend only on lexicons for finding the average sentiment (or) emotion present in a document, These methods starts with a set of opinion words with known orientation and expand that set by using a known thesaurus for synonyms. These methods also make use of structural elements and syntactic patterns present in the text by using NLP Processes like lemmatization and POS tagging[14]. Turney[13] has proposed an unsupervised method that classifies reviews as negative (or) positive by calculating the semantic orientation of phrases by linking them with only two words namely poor and excellent. Heerschop et al [4]. have studied aspect based knowledge extracted which to increase the performance of sentiment analysis. Oiu et al.[16] has used a rule base method combining syntactic parsing and lexicon to extract opinion sentences with negative sentiment and also identify sentence topics. Saif, He, fernandez and Alani[6] have proposed a new lexicon based approach known as senticirles. It captures latent semantics from their co-occurrence patterns to update the sentiment orientation of words.

3.2. Lexicon based classification

The lexicon based sentiment analysis[19] is based on the available sentiment lexicon. Generally sentiment lexicon are document which consist of set of teams which is carrying some emotional weights along with some dimensional information. The dimension can be scored by either a specific rating scales or in binary manner (e.g either positive or negative). There are three key techniques in order to extract feature words list. Manual approaches is very time-consuming and it can't be used separately. There two automated approaches are availability they are Dictionary based approach and corpus based approach. The corpus based approach as be further classified into statistical approach and semantic approach.

4. Proposed architecture

The Figure 4.1 shows the overall architecture for the proposed hybrid sentiment analysis technique which combines both word embedding based and lexicon-based approach.

4.1. Preprocessing of input documents

Before feeding the data input the sentiment analysis algorithm for sentiment prediction there are a number of preprocessing step as mentioned below:

1. **Tokenization:** This is the method of splitting the document in to words or sentence based on requirement.
2. **Contractions:** Replace replacing the contents with two tokens, e.g., don't is replaced by do and not.
3. **Stop-word Removal:** In this step common words such as "a", "an", "the", etc. are removed which contain no sentiment or semantic.

4.2. Bag-of-Words representation

The BoW representation is a collection of all words present in the document after removing the stop words. All words are turned in to lowercase but no lemmatization or stemming was applied. We use TF-IDF[9] representation of documents where the term frequency are forced to be binary which give the word is present or not in the document and only the document frequency will vary.

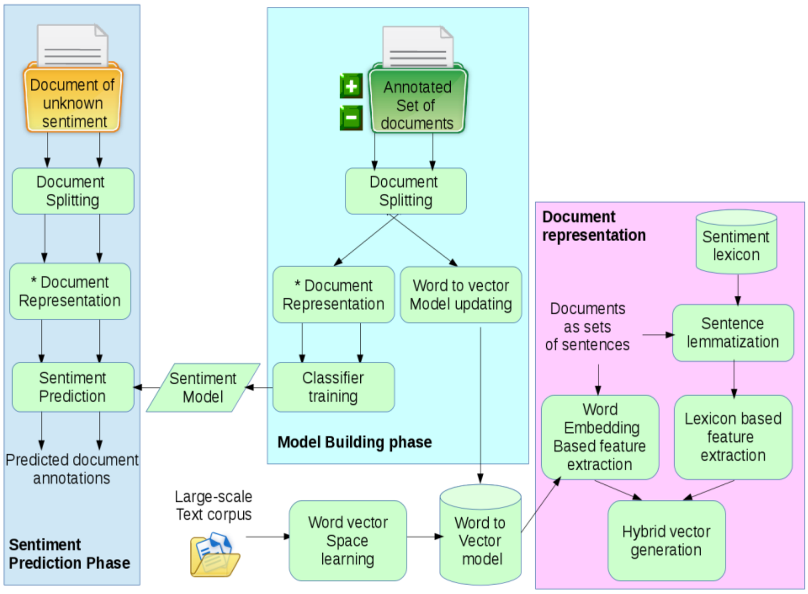

Fig 4.1. Proposed Architecture

4.3. WW2V representation

The Word2Vec model embeds semantic information in a vector space representation of words. The main idea of this model is to capture the words with help of some machine learning algorithm such as neural networks. Word2Vec has the ability to predict the occurrence of a word when the window of previous and next words. Word2Vec follows two different approaches for representation of words for large datasets:

a) **Continuous Bag-of-Words model (CBOW):** when the surrounding words are give we can predict the center word. This approach of word2vec is much faster and accurate than Skip-gram model.
b) **Skip-Gram model (SG):** when the single word is given we can predict surround window in a document. This model works well with small number of training dataset and represent accurately even a word which is rarely used or a phrase.

The learned vector use have many linguistic patterns, for example the vector of the two similar words will occupy the space with minimum Euclidean space between each other. These patterns can be represented as linear transactions. The following these successful techniques, research interested to develop a model that goes beyond word level to achieve sentence level representations. Doc2Vec[10] is a more complex approach, which modifies the word2vec algorithm to a unsupervised learning of large blocks of text, such as sentences, phrases, or entire documents.

Proposed approach: The Weighted Word2Vec is a modifies version of Mean Word2Vec, here we represent the model in the terms of weighted average of the representation of all the words obtained from Word2Vec model. For each word ‘i’ in the document we compute V_i, where V_i is a vector representation by using Continuous Bag of words model, and we construct a dictionary matrix

$$V \in \mathbb{R}^{N \times B}$$

Where N is the word representation dimension and B is the dictionary size. Each term weight depict the significance of a word in a document, thus Term Frequency(TF) is normalized with Inverse Document Frequency(IDF). We are gonna calculate TF-IDF for a word ‘i’ in the document ‘j’ as

$$w_{ij} = f_{ij} \log\left(\frac{D}{D_i}\right)$$

Where the frequency 'f_{ij}' is the number of occurrence of a i^{th} word in the document j, D is the number of documents in the corpus and D_i is the document with the i^{th} word in the documents. The vector representation of the j^{th} document the sum of words appear in it weighted by their TD-IDF values:

$$S_j = \sum_i w_{ij}\, v_i$$

Example:

Text	Sentiment Vector	Positive	Objective	Negative	Unknown
-	1	0	0	0	1
.	1	0	0	0	1
bad	1	0	0.2	0.8	0
Between	1	0	1	0	0
Dead	1	0.1	0.3	0.6	0
Man	1	0	1	0	0
Room	1	0	1	0	0
Smell	1	0.2	0.6	0.2	0
Smt	1	0	0	0	1
So	2	0	1	0	0
Towels	1	0	1	0	0
Wardrobe	1	0.1	0.9	0	0
Wet	1	0.1	0.8	0.1	0

Table 4.1 BoW representation of sentence and respective sentiments polarity and objective values from SentiWordNet

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0.2 & 0.8 & 0 \\ 0 & 1 & 0 & 0 \\ 0.1 & 0.3 & 0.6 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.2 & 0.6 & 0.2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.1 & 0.9 & 0 & 0 \\ 0.1 & 0.8 & 0.1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.036 & 0.671 & 0.121 & 0.171 \end{pmatrix}$$

Positive	Objective	Negative	Unknown
0.036	0.671	0.121	0.171

Table 4.2 The vector representation of the text

4.4. Lexicon based representation

The Word2Vec based representation described in the previous section doesn't contain any sentiment information about the give input document. We can study the represented by adding the semantic structure of words based on a lexicon that associated a set of terms with sentiment polarity with numeric values. This approach also consider the semantic orientation of the words based on lexicon that associate a set of values with their semantic polarity expressed in terms of numeric values. Actually this approach consider the semantic structure of the document independent of its context. These lexicons can be created either manually or by using semi-automatically by untiling resources like WordNet. In our approacher we use a semi-automatic sentiment lexicon called SentiWordNet[11] which is widely used in various sentiment analysis applications.

SentiWordNet is an extension of WordNet lexicon of synonyms adding sentiment vales for each syn set in the collection by means of combining linguistic and statistical classifiers. SentiWordNet set three sentiment scores for each synset of WordNet, based on its polarity and objectivity, accordingly, the sum of these scores will be always 1.

In the model we implemented uses SentiWordNet, to get the values of polarity and objective scores for each word present in the corpus, by averaging the values of synsets in SentiWordNet that has same text representation. To indicate the word which are not present in the dictionary, a fourth feature was added, represent unknown sentiment. The sentiment scores for all words are represented in a sentiment matrix.

Since a token generally does not pass its sentiment in another sentence, we split the reviews input sentences. Vector of each sentence is generated using BoW model, and we can find the overall sentiment of the document by performing dot product between the sentence vector and the sentiment matrix. To find the sentiment for entire document we calculate the mean sentiment representation across all the sentence of the document. For each document four features are extracted:

1. Positive sentiment value,
2. Negative sentiment value,
3. Object sentiment value, and
4. A flag to represent word which is not present or not included in the SentiWordNet.

5. Conclution

In project, We propose a hybrid method that combines the method TF-IDF WW2V representation and Bag-of-Words representation with sentiment lexicon based sentiment values for unstructured text. The word embedding representation, based on Word2Vec, which provides the syntactic and semantic representation, this approach can be refined in further by combining the weighted Word2Vec with TF-IDF weight which gives us the frequent words. We used the sentiment Lexicon based representation that provides emotional and sentiment information for inspection of document. The hybrid approach combines the merits of both the approaches, and this approach can be modified for any language as long as sentiment lexicons are available.

6. Acknowledgment

Authors thank the Science and Engineering Research Board for their financial support (YSS/2014/000718/ES). Authors also express their gratitude to SASTRA University for the infrastructure facilities and support provided to conduct the research.

7. References

1. Chatzakou, D. , & Vakali, A. (2015). Harvesting opinions and emotions from social media textual resources. IEEE Internet Computing, 19 (4), 46–50 .
2. G. Salton and M. J. McGill, Introduction to modern information retrieval (McGraw-Hill, Inc., 1986).

3. Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, *1*(1), 81-106.
4. Heerschop, B., Goossen, F., Hogenboom, A., Frasincar, F., Kaymak, U., & de Jong, F. (2011, October). Polarity analysis of texts using discourse structure. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 1061-1070). ACM.
5. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.
6. Saif, H., Fernandez, M., He, Y., & Alani, H. (2013). Evaluation datasets for Twitter sentiment analysis: a survey and a new dataset, the STS-Gold.
7. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, *37*(2), 267-307.
8. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
9. Joachims, T. (1996). *A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization* (No. CMU-CS-96-118). Carnegie-mellon univ pittsburgh pa dept of computer science.
10. Le, Q., & Mikolov, T. (2014, January). Distributed representations of sentences and documents. In *International Conference on Machine Learning* (pp. 1188-1196).
11. Baccianella, S., Esuli, A., & Sebastiani, F. (2010, May). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *LREC* (Vol. 10, No. 2010, pp. 2200-2204).
12. Lin, C., & He, Y. (2009, November). Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 375-384). ACM.
13. Turney, P. D. (2002, July). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics* (pp. 417-424). Association for Computational Linguistics.
14. Bird, S., & Loper, E. (2004, July). NLTK: the natural language toolkit. In *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions* (p. 31). Association for Computational Linguistics.
15. Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10* (pp. 79-86). Association for Computational Linguistics.
16. Van de Voorde, I., Martin, C. M., Vandewege, I., & Oiu, X. Z. (2000). The superPON demonstrator: an exploration of possible evolution paths for optical access networks. *IEEE Communications magazine*, *38*(2), 74-82.
17. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113.
18. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1-135.
19. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, *37*(2), 267-307.
20. Tan, S., Cheng, X., Wang, Y., & Xu, H. (2009, April). Adapting naive bayes to domain adaptation for sentiment analysis. In *European Conference on Information Retrieval* (pp. 337-349). Springer, Berlin, Heidelberg.
21. Logesh, R., Subramaniyaswamy, V., Vijayakumar, V., Gao, X. Z., & Indragandhi, V. (2017). A hybrid quantum-induced swarm intelligence clustering for the urban trip recommendation in smart city. Future Generation Computer Systems, 83, 653-673.
22. Subramaniyaswamy, V., & Logesh, R. (2017). Adaptive KNN based Recommender System through Mining of User Preferences. Wireless Personal Communications, 97(2), 2229-2247.
23. Logesh, R., & Subramaniyaswamy, V. (2017). A Reliable Point of Interest Recommendation based on Trust Relevancy

between Users. Wireless Personal Communications, 97(2), 2751-2780.

24. Logesh, R., & Subramaniyaswamy, V. (2017). Learning Recency and Inferring Associations in Location Based Social Network for Emotion Induced Point-of-Interest Recommendation. Journal of Information Science & Engineering, 33(6), 1629–1647.
25. Subramaniyaswamy, V., Logesh, R., Abejith, M., Umasankar, S., & Umamakeswari, A. (2017). Sentiment Analysis of Tweets for Estimating Criticality and Security of Events. Journal of Organizational and End User Computing (JOEUC), 29(4), 51-71.
26. Indragandhi, V., Logesh, R., Subramaniyaswamy, V., Vijayakumar, V., Siarry, P., & Uden, L. (2018). Multi-objective optimization and energy management in renewable based AC/DC microgrid. Computers & Electrical Engineering.
27. Subramaniyaswamy, V., Manogaran, G., Logesh, R., Vijayakumar, V., Chilamkurti, N., Malathi, D., & Senthilselvan, N. (2018). An ontology-driven personalized food recommendation in IoT-based healthcare system. The Journal of Supercomputing, 1-33.
28. Arunkumar, S., Subramaniyaswamy, V., & Logesh, R. (2018). Hybrid Transform based Adaptive Steganography Scheme using Support Vector Machine for Cloud Storage. Cluster Computing.
29. Indragandhi, V., Subramaniyaswamy, V., & Logesh, R. (2017). Resources, configurations, and soft computing techniques for power management and control of PV/wind hybrid system. Renewable and Sustainable Energy Reviews, 69, 129-143.
30. Ravi, L., & Vairavasundaram, S. (2016). A collaborative location based travel recommendation system through enhanced rating prediction for the group of users. Computational intelligence and neuroscience, 2016, Article ID: 1291358.
31. Logesh, R., Subramaniyaswamy, V., Malathi, D., Senthilselvan, N., Sasikumar, A., & Saravanan, P. (2017). Dynamic particle swarm optimization for personalized recommender system based on electroencephalography feedback. Biomedical Research, 28(13), 5646-5650.
32. Arunkumar, S., Subramaniyaswamy, V., Karthikeyan, B., Saravanan, P., & Logesh, R. (2018). Meta-data based secret image sharing application for different sized biomedical images. Biomedical Research,29.
33. Vairavasundaram, S., Varadharajan, V., Vairavasundaram, I., & Ravi, L. (2015). Data mining-based tag recommendation system: an overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 5(3), 87-112.
34. Logesh, R., Subramaniyaswamy, V., & Vijayakumar, V. (2018). A personalised travel recommender system utilising social network profile and accurate GPS data. Electronic Government, an International Journal, 14(1), 90-113.
35. Vijayakumar, V., Subramaniyaswamy, V., Logesh, R., & Sivapathi, A. (2018). Effective Knowledge Based Recommeder System for Tailored Multiple Point of Interest Recommendation. International Journal of Web Portals.
36. Subramaniyaswamy, V., Logesh, R., & Indragandhi, V. (2018). Intelligent sports commentary recommendation system for individual cricket players. International Journal of Advanced Intelligence Paradigms, 10(1-2), 103-117.
37. Indragandhi, V., Subramaniyaswamy, V., & Logesh, R. (2017). Topological review and analysis of DC-DC boost converters. Journal of Engineering Science and Technology, 12 (6), 1541–1567.
38. Saravanan, P., Arunkumar, S., Subramaniyaswamy, V., & Logesh, R. (2017). Enhanced web caching using bloom filter for local area networks. International Journal of Mechanical Engineering and Technology, 8(8), 211-217.
39. Arunkumar, S., Subramaniyaswamy, V., Devika, R., & Logesh, R. (2017). Generating visually meaningful encrypted image using image splitting technique. International Journal of Mechanical Engineering and Technology, 8(8), 361–368.
40. Subramaniyaswamy, V., Logesh, R., Chandrashekhar, M., Challa, A., & Vijayakumar, V. (2017). A personalised movie recommendation system based on collaborative filtering. International Journal

of High Performance Computing and Networking, 10(1-2), 54-63.

41. Senthilselvan, N., Udaya Sree, N., Medini, T., Subhakari Mounika, G., Subramaniyaswamy, V., Sivaramakrishnan, N., & Logesh, R. (2017). Keyword-aware recommender system based on user demographic attributes. International Journal of Mechanical Engineering and Technology, 8(8), 1466-1476.
42. Subramaniyaswamy, V., Logesh, R., Vijayakumar, V., & Indragandhi, V. (2015). Automated Message Filtering System in Online Social Network. Procedia Computer Science, 50, 466-475.
43. Subramaniyaswamy, V., Vijayakumar, V., Logesh, R., & Indragandhi, V. (2015). Unstructured data analysis on big data using map reduce. Procedia Computer Science, 50, 456-465.
44. Subramaniyaswamy, V., Vijayakumar, V., Logesh, R., & Indragandhi, V. (2015). Intelligent travel recommendation system by mining attributes from community contributed photos. Procedia Computer Science, 50, 447-455.
45. Vairavasundaram, S., & Logesh, R. (2017). Applying Semantic Relations for Automatic Topic Ontology Construction. Developments and Trends in Intelligent Technologies and Smart Systems, 48.