UNIONIZATION METHOD FOR CHANGING OPINION IN SENTIMENT CLASSIFICATION USING MACHINE LEARNING

ABBAS JALILVAND

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

JANUARY 2020

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time. I would like to dedicate this thesis to my beloved wife SANAZ, for all of her love and support

ACKNOWLEDGEMENT

In the name of Allah, Most Gracious, Most Merciful. I thank Allah s.w.t for granting me perseverance and strength I needed to complete this thesis.

In preparing this thesis, I wish to express my sincere thanks to my main thesis supervisor, Professor Dr.Naomie Salim for friendship, motivation, and encouragement.

I would like to thank the staff of Universiti Teknologi Malaysia, and especially the School of Computing, for their kind cooperation.

Lastly, I wish to thank from all the people who have helped me during conducting my PhD project and the ones who provided such a great academic environment for research and education.

ABSTRACT

Sentiment classification aims to determine whether an opinionated text expresses a positive, negative or neutral opinion. Most existing sentiment classification approaches have focused on supervised text classification techniques. One critical problem of sentiment classification is that a text collection may contain tens or hundreds of thousands of features, i.e. high dimensionality, which can be solved by dimension reduction approach. Nonetheless, although feature selection as a dimension reduction method can reduce feature space to provide a reduced feature subset, the size of the subset commonly requires further reduction. In this research, a novel dimension reduction approach called feature unionization is proposed to construct a more reduced feature subset. This approach works based on the combination of several features to create a more informative single feature. Another challenge of sentiment classification is the handling of concept drift problem in the learning step. Users' opinions are changed due to evolution of target entities over time. However, the existing sentiment classification approaches do not consider the evolution of users' opinions. They assume that instances are independent, identically distributed and generated from a stationary distribution, even though they are generated from a stream distribution. In this study, a stream sentiment classification method is proposed to deal with changing opinion and imbalanced data distribution using ensemble learning and instance selection methods. In relation to the concept drift problem, another important issue is the handling of feature drift in the sentiment classification. To handle feature drift, relevant features need to be detected to update classifiers. Since proposed feature unionization method is very effective to construct more relevant features, it is further used to handle feature drift. Thus, a method to deal with concept and feature drifts for stream sentiment classification was proposed. The effectiveness of the feature unionization method was compared with the feature selection method over fourteen publicly available datasets in sentiment classification domain using three typical classifiers. The experimental results showed the proposed approach is more effective than current feature selection approaches. In addition, the experimental results showed the effectiveness of the proposed stream sentiment classification method in comparison to static sentiment classification. The experiments conducted on four datasets, have successfully shown that the proposed algorithm achieved better results and proving the effectiveness of the proposed method.

ABSTRAK

Klasifikasi sentiment bertujuan untuk menentukan apakah suatu teks mengungkapkan pendapat positif, negatif atau neutral. Kebanyakan pendekatan klasifikasi sentimen yang sedia ada memberi tumpuan kepada teknik pengelasan teks ter penyelia. Satu masalah yang kritikal berkaitan pengelasan sentimen ialah pengumpulan teks mungkin mengandungi puluhan atau ratusan ribu ciri, dengan dimensi tinggi, yang dapat diselesaikan oleh pendekatan pengurangan dimensi. Walaupun kaedah pengurangan dimensi dapat mengurangkan dimensi ciri untuk menghasilkan saiz subset ciri yang lebih kecil, saiz subset biasanya memerlukan pengurangan selanjutnya. Dalam kajian ini, pendekatan pengurangan dimensi yang dipanggil penyatuan ciri telah dicadangkan untuk membina subset ciri yang lebih kecil. Pendekatan ini menggabungkan beberapa ciri untuk mewujudkan ciri yang lebih bermaklumat. Satu lagi kesukaran untuk klasifikasi sentimen yang belum diatasi ialah pengendalian masalah konsep hanyut dalam proses pembelajaran. Pendapat pengguna boleh berubah kerana perubahan entiti sasaran dari masa ke masa. Walau bagaimanapun, pendekatan klasifikasi sentimen sedia ada tidak mengambil kira evolusi pendapat pengguna. Mereka menganggap tika adalah bebas, tersebar secara saksama dan dijana daripada taburan pegun, walaupun ianya dihasilkan dari taburan aliran. Dalam kajian ini, kaedah pengkelasan sentimen aliran dicadangkan untuk menangani perubahan pendapat dan pengagihan data yang tidak seimbang menggunakan kaedah pembelajaran ensembel dan kaedah pemilihan contoh. Sehubungan dengan masalah konsep hanyut, satu lagi isu penting yang belum ditangani secara mendalam ialah pengendalian ciri hanyut dalam klasifikasi sentimen. Untuk mengendalikan ciri hanyut, ciri yang berkaitan perlu dikesan untuk mengemas kini pengelas. Oleh kerana kaedah penyatuan ciri yang dicadangkan sangat berkesan untuk membina ciri-ciri yang lebih relevan, ia terus digunakan untuk menangani ciri hanyut. Oleh itu, kaedah untuk menangani konsep hanyut dan ciri hanyut untuk klasifikasi sentimen aliran telah dicadangkan. Keberkesanan kaedah penyatuan ciri telah dibandingkan dengan kaedah pemilihan ciri dengan lebih empat belas set data awam dalam domain klasifikasi sentimen menggunakan tiga pengelas biasa. Hasil kajian menunjukkan pendekatan yang dicadangkan lebih berkesan daripada pendekatan pemilihan ciri semasa. Di samping itu, keputusan kajian menunjukkan keberkesanan kaedah klasifikasi sentimen aliran yang dicadangkan berbanding dengan klasifikasi sentimen statik. Kajian yang dilakukan pada empat dataset telah berjaya menunjukkan algoritma yang dicadangkan mencapai hasil yang lebih baik dan membuktikan keberkesanan kaedah yang dicadangkan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SC	-	Sentiment Classification
OM	-	Opinion Mining
SA	-	Sentiment Analysis
FS	-	Feature Selection
FU	-	Feature Unionization
IS	-	Instance Selection
RS	-	Random Subspace
VSM	-	Vector Space Model
BOW	-	Bag of Words
IG	-	Information Gain
CHI2	-	CHI Squre
TCM	-	Trigonometric Comparison Measure
MR	-	Movire Review
POS	-	Part of Speech
IDF	-	Inverse Document Frequency
TF	-	Term Frequency
SVM	-	Support Vector Machines
NB	-	Naïve Bayes
KNN	-	K-Nearest Neighbor
ME	-	Maximum Entropy
DT	-	Decision Tree
ANN	-	Artificail Neural Network
CRF	-	Conditional Random Fields
PSO	-	Particle Swarm Optimization
GA		Genetic Algorithm
CNN	-	Condensed Nearest Neighbor
ENN	-	Edited Nearest Neighbor
RNN	-	Reduced Nearest Neighbor

LIST OF SYMBOLS

D	-	Dataset
F	-	Feature
L	-	Label
ACC	-	Accuracy
C	-	Class
P	-	Probability
S	-	Subset
ε	-	Threshold

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In recent years, social media (e.g. social networks, discussion boards, blog and forums) are growing rapidly. Thus, people can easily express their opinions on several topics such as products or services of companies. Thus, a huge numbers of reviews have been increased on the Web. Since information plays an important role in influencing consumer decisions, these reviews containing opinions originated from the user's experiences are useful and helpful for companies and individuals. When people want to buy a product or use a service, they would like to use people's experiences. Furthermore, manufacturers and service providers want to be aware of the opinions of their customers about their products and services.

Awareness of people's opinions and recommendations has always been important. In the past, individuals would ask friends or family, and organizations used surveys and consultants to find people's opinions; while, nowadays, many e-commerce websites, such as Amazon.com, Yahoo's shopping, Epinions.com that made it possible to read many opinions and experiences are posed by users. However, reading and understanding the high amount of reviews is impossible. To address this problem, opinion mining or sentiment analysis approach has recently emerged. Generally, opinions can be expressed on any target entity (e.g., a service a product, an organization, an individual, etc.). For example:

"I bought this camera a few days ago. This camera is very easy to use, but it is very expensive. This camera has a great zoom and captures a nice crisp photo. The battery life is long about three hours of nonstop use."

In this example, customer' opinion is positive about the battery life and negative about the price.

Opinion mining is a sub-discipline of text mining that refers to application of natural language processing to extract, process of the opinions, attitudes, and emotions toward entities and their attributes and to present them friendly to users. Sentiment analysis attempts to detect subjectivity, sentiment, affect, and other emotional states in the opinionated text. In the area of opinion mining, according to respond to different user's request, several fields such as subjectivity classification, sentiment classification and opinion summarization have emerged. The goal of subjectivity classification is determining whether a text (i.e., sentence, paragraph, or document) is objective or subjective. In fact, it distinguishes text containing opinions from text objectively present factual information (Liu, 2012). The goal of sentiment classification as the most active field of opinion mining is to classify the opinion documents into positive or negative class (Chen *et al.*, 2019a). Opinion summarization aims to provide a digest summary of the high numbers of reviews in an easy manner to present users.

There are many applications for sentiment classification such as handling business intelligence tasks (Pang and Lee, 2008), predicting stock market behavior (Bollen *et al.*, 2011), measuring public poll opinion of presidential elections (O'Connor *et al.*, 2010). Sentiment classification can be applied at word-level, sentence-level, document-level and feature/aspect-level using different algorithms ranging from unsupervised to supervised approaches (Wang *et al.*, 2018).

Most sentiment classification works have been applied on document level using supervised classification algorithms due to their predictive power. In the supervised machine learning approach, a classifier is trained to determine the sentiment of reviews using prior training data. The first work of sentiment classification at document-level is by Pang *et al.* (2002), who compared support vector machines (SVMs) , Maximum Entropy (ME) and Naïve Bayes (NB) to classify movie reviews and reached the highest accuracy (82.9 %) using SVM. In recent years, there has been a growing interest in using feature selection and ensemble learning methods to enhance sentiment classification accuracy. The main goal of proposed methods is the improvement of sentiment classification in terms of efficiency and efficacy. Moreover, some researcher work to develop on data preparation to prepare and preprocess the initially available data in learning process (Onan and Korukoğlu, 2016b; Tubishat *et al.*, 2019). Feature selection methods aim to remove irrelevant and redundant features to select an optimal feature subset for classification task (Abbasi *et al.*, 2011; Wang *et al.*, 2011a). The ensemble methods aim to combine the outputs of several base learners to obtain better sentiment classification performance (Wang *et al.*, 2014; Xia *et al.*, 2011; Zhang *et al.*, 2019). An overview of steps and techniques for sentiment classification is shown in figure 1.1.

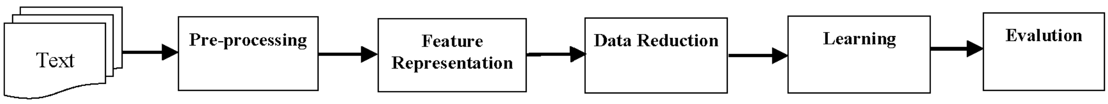

Figure 1.1 The process of supervised sentiment classification at document level

1.2 Problem Background

As we know, text data is unstructured, thus Bag of Words (BOW) model (i.e. term-based Vector Space Model (VSM)) has been most popular for document-feature representation (Baeza-Yates and Ribeiro-Neto, 1999). In VSM model, a document is represented as a feature vector which consists of all words in the document. VSM has been widely applied to text classification due to its simplicity and good performance. However, the main drawback of VSM is that it does not consider the semantic relatedness between words that limits performance of classification task. For example, two words with similar meanings are treated as irrelevant features in VSM. Extracting semantic information by some dictionaries (e.g., WordNet) has a main drawback. Finding the appropriate related concepts for words using WordNet is very difficult. Another problem of VSM is that text collection would result in tens or hundreds of thousands of features. In theory, having more features, should improve the efficiency of classifier but, it is not always true practically. More features may confuse the learning algorithm because most of the features are irrelevant or redundant that may lead classifier to over-fitting. Moreover, a large number of features impose a high computational cost on the learning step.

Accordingly, feature selection (FS) is proposed to remove unnecessary features to improve classifier's generalization ability and computational efficiency, but the number of features need to be more reduced considerably. Therefore developing an effective and efficient method to reduce more features is a vital issue for sentiment classification.

In general, people express their sentiments about a target entity (e.g. product or a service etc.) based on their characteristics which are changed over time. User opinions are changed due to evolution of target entities over time. For example, in the phone product, some features changed (add or remove) at the specific time and some terms (words) associated to the feature may be appeared or disappeared in the phone reviews after this time. The behavior of the customers in an online shop may change over time. One reason for changing opinion may be seasonality, which means that customers' opinion about a product may change seasonally. However, the existing sentiment classification approaches not considered the changing of review document. They assume that instances are independent, identically distributed and generated from a stationary distribution. These techniques generally construct a static learning model from the training dataset and then this model is used to classify a new review. While, labeled reviews used as training set are posted over time and cannot be given to the learner in any order. In fact, each target function inferred at time t can only use the data given before t. The phenomenon of concept changing over time is termed as concept drift in machine learning. In sentiment classification, it can be considered as changing opinion. It is not questionable that the ability to automatically adaptive the classifier over time plays an important role in the real-world application of sentiment classification.

In the concept drift learning, the target function changes over time and need to be adapted (Khamassi *et al.*, 2018; Krempl *et al.*, 2014; Roveri, 2019; Žliobaitė *et al.*, 2016). In stream data mining, For example, the sentiment of texts could change from time to time due to evolving continuously over time. Therefore, in contrast to static concept learning, ordering of the training data is important in concept drift learning. To handle the concept drifts, classifier is adapted to track the changing in data. Most existing concept drift learning algorithms, work based on learning from a window of most recent examples to adapt to concept drift (Klinkenberg and Joachims, 2000; Mena-Torres and Aguilar-Ruiz, 2014). Figure 1.2 illustrates this approach that automatically eliminates older examples which are no longer relevant in learning. Determining the appropriate window size play an important role that is not easy to do. Some researcher developed an adaptive window adjustment heuristic has been to solve the problem that is effective in slow drift rate condition (Gama *et al.*, 2014). Among these methods, the most popular evolving technique for handling concept drift is classifier ensemble (de Mello *et al.*, 2019; Farid *et al.*, 2013; Gomes *et al.*, 2017; Khamassi *et al.*, 2018, 2019; Mirza *et al.*, 2015).

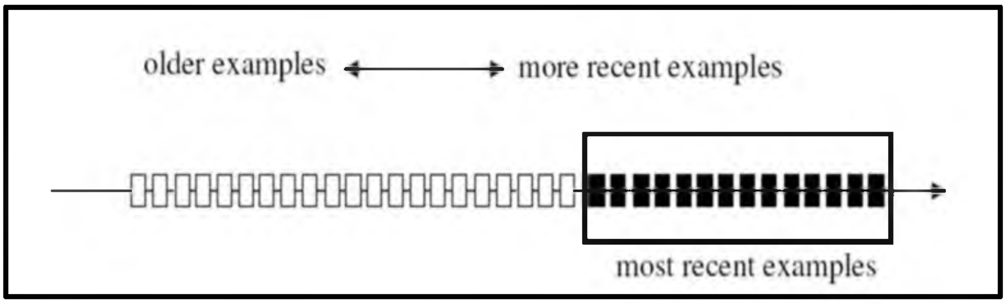

Figure 1.2 A typical approach to concept drifts learning

One important issue, not yet convincingly addressed, is the handling of concept drifts problem in the supervised sentiment classification at the document level. Besides the challenge to adapt learning model, imbalanced data is critical problem needed to be concerned. One of the major challenges in machine learning is that, distribution classes of dataset may be unequal that is called imbalanced data problem. In this situation, learning algorithms are biased towards the majority classes (FernáNdez *et al.*, 2013; Gao *et al.*, 2008; Li *et al.*, 2018; Sun *et al.*, 2017). Previous studies assume sentiment datasets are balanced, while in the real world they are imbalanced, especially when considering in stream learning form. Many approaches were proposed to address the imbalanced data problem both at the data and algorithmic levels. Most approaches for learning from such data are based on under-sampling the majority class or over-sampling the minority class (Sáez *et al.*, 2016). The training dataset is balanced in oversampling by instance generation and in under-sampling methods by instance selection.

Despite of concept drift, another drift is changes in the relevance of features through time, a phenomenon called feature drift (Barddal *et al.*, 2017). The following examples are provided to clarify the concept drift vs. feature drift. Consider the first version of a phone with a particular set of features. The phone producer investigated the customers' feedback and they recognized that their customers prefer a phone with more lightweight and larger screen and higher picture resolution. Therefore, next version is produced. The features of these two different versions remained unmodified but their specifications have evolved (e.g. its weight changed from 250 gr to 150 gr). This illustrates concept drift. However let us now consider that a newer version with a new feature (e.g. Wi-Fi technology). This is a good example of feature drift. Handling feature drift and concept drift simultaneously is very effective to further adapt classifier. Therefore, in this research feature drift is addressed to enhanced stream sentiment classification method.

1.3 Problem Statement

The first step of sentiment classification is data acquisition and data pre-processing. The next step is data reduction that can be applied in two levels: row and column of data. Instance and feature selection can construct a more reduced and discriminative data subset. These steps can be considered as pre-processing before learning. The most important step is learning to train a classifier based on training dataset. In the final, the evaluation of classification process is presented. Dimension reduction and learning steps are two critical factors in supervised classification performance. An important problem of sentiment classification is high dimensionality that is considered in this study. Although feature selection methods can reduce the original feature set and select a reduced feature subset that have more discriminative power for learning task, the size of subset can be more reduced to construct a more compact and discriminative feature subset. Despite of removing all irrelevant and redundant features, there are some features that are correlated to each other implicitly. Similar features tend to occur in documents that are belonging to same class. These relations can be identified if they combined using union operator based on increasing the relevancy to target class. Thus, feature unionization is a new view to reduce feature dimension. This method constructs a more reduced and discriminative feature subset. Taking into account the synonym words that can be considered as a feature, the basic idea of the proposed approach was to reduce dimensionality based on finding and combining features that can construct a more informative single feature based on a feature relevancy criterion. There are implicit and explicit relations between occurring words (features) in the same class. For example, synonyms, or the words of the same group, may tend to occur in the same class. In sentiment classification domain, two words, for example, 'good' and 'great' usually indicate positive class that can be unionized to make an individual feature. Practically, capturing these relations is not easy to do because most of them are latent. The proposed feature unionization approach can capture the relation of features according to their relevancy to the target class in a way to construct an informative feature. Since combination of features is carried out by union operator, the redundancy can also be removed due to inherent characteristic of unionization. In this research, supervised feature unionization is investigated that is suitable for binary datasets.

Supervised sentiment classification is aim to automatically classify an opinion text into the positive or negative class by employing some machine learning techniques (Wang *et al.*, 2014). They usually employs a static supervised learning strategy, in which a classification model is first built using a training set to classify a testing set without considering the time that reviews are posted. However, time may be very useful as an important feature for classification task. In general, people express their opinions about a target entity (e.g. product or a service etc.) based on their characteristics which are changed over time. User's opinions are changed due to evolution of target entities over time. For example, in the phone product, some features changed (add or remove) at the specific time and some terms (words) associated to the features may be appeared or disappeared in the phone reviews. The phenomenon of concept changing over time is termed as opinion changing or concept drift in machine learning. However, the existing sentiment classification approaches incapable to track the changing because they commonly build a static learning model based on training dataset with ignoring time dependency of data. Therefore, there is need to design a sentiment learning model that works based on data continuously flow. The important issue in these models is changing opinion that causes it becomes obsolete and need to be updated (adapted) over time. This research gives a primary focus on sentiment classification model with concept drift. Among evolving different methods for handling concept drift, the most popular technique is ensemble classification that construct different classifiers with diverse windows form stream data. The data selected by a window are imbalanced specially when considering with concept drift. Therefore, imbalanced data distribution is another issue besides changing opinion handling problem that need to be addressed. Some researcher have used ensemble learning to deal with imbalanced data distribution. Therefore, ensemble learning is effective to deal with both imbalanced data distribution and concept drifts problem (Sun *et al.*, 2017). Most approaches for learning from imbalanced data are based on under-sampling the majority class or over-sampling the minority class (Sáez *et al.*, 2016). The training dataset is balanced in Oversampling by instance generation and in under-sampling methods by instance selection. In fact, proposing a stream sentiment classification to deal with concept drift and imbalanced class distribution is an existing problem is addressed in this study.

Another difficulty to stream sentiment classification is feature drift that need to be concerned. To handle feature drift, relevant features need to be detected to adapt classifier. Thus, some work used dynamic feature selection method to address feature drift in stream classification. Since feature unionization method is very effective to construct more relevant features, it can be used to handle feature drift. Therefore a method for dealing with both concept and feature drifts with imbalanced data distribution using ensemble learning based on feature unionization and instant selection for sentiment classification is great demand that is proposed in this study.

1.4 Research Questions

The open issues described in the previous section lead to mentioning some research questions addressed in this research are follows:

1- How to propose a feature unionization method to combine the features to construct a more compact and discriminative features subset?

2- How to propose a stream sentiment classification method to handle concept drift with imbalanced data?

3- How to propose an integrated sentiment classification method to deal with feature drift?

1.5 Research Objectives

Based on the above mentioned research questions the objectives of the research are:

1- To propose a feature unionization method to combine the features to construct a more compact and discriminative features subset.

2- To propose a stream sentiment classification method to handle concept drift with imbalanced data using ensemble learning and instance selection methods.

3- To propose an integrated sentiment classification method to deal with feature drift using feature unionization method.

1.6 Research Scope

Subsequent to the goal and objectives of this study is the research scope. In view of the fact that there is a number of diversity in machine learning to solve such problems, this study is scoped as follows:

1- This study considers supervised sentiment classification at document level and does not include other opinion mining fields.

2- Among different representation data techniques the bag of word (unigrams) technique is used to transfer sentiment review datasets in this study.

3- Three different classification methods (SVM, Naïve Bayes and KNN) are applied as base learners for sentiment classification.

4- Three popular feature relevancy measures (Information gain, Chi square and Trigonometric Comparison Measure) are used to evaluate the relevancy of features in proposed dimension reduction method.

5- Fourteen sentiment classification datasets as widely used by other researchers are investigated in this research. The movie review (MR) crawled from the IMDB movie archive (Pang and Lee, 2004). The multi-domain sentiment (MDS) used by Blitzer *et al.* (2007) crawled fromAmazon.com (Book, DVD, Electronics, and Kitchen). The other nine datasets (Camera, Camp, Doctor, Drug, Laptop, Lawyer, Music, Radio, and TV) were provided by Whitehead and Yaeger (2009).

1.7 Significance of the Research

This research is important and significant from theoretical and practical perspectives for developing sentiment classification. The rationale and motivation for this research are as follow:

1) Reduction of feature dimensionality is very important for classification task to reduce the computational complexity and avoid over-fitting problem, which improves the generalization ability of classifier. For having better generalization performance of the classification, the number of features should be reduced as much as it is required for the number of training samples. These reasons are motivations to propose an effective approach to construct a more compact and discriminative feature subset by features combination idea.

2) The development of stream sentiment classification is important to address the concept drift that is a significant factor to improve the classification performance.

3) Imbalanced data handling which is commonly found in sentiment classification domain especially when considering concept drift. It may decrease the performance of machine learning techniques. Thus it needs to be addressed efficiently.

4) Feature drift handling is an important key to adapt the classifier in stream sentiment classification. Concept and feature drifts with imbalanced data handling may construct a more adaptive classifier in stream sentiment classification.

1.8 Research Contributions

The main contributions of this research from theoretical and practical perspectives are summarized as follow:

1- A new theoretical approach for dimension reduction is proposed which causes feature dimension space is significantly reduced because several features are unionized into a single feature, consequently; the performance of classification is increased due to transformation of feature space to a more discriminative subset. Additionally to sentiment classification, the proposed feature unionization approach can be effective on different machine learning fields.

2- A new theoretical approach is proposed to develop of an adaptive sentiment model that is capable to deal with changing opinion and concept drift to adapt the classifiers for improving the classification accuracy on sentiment datasets. Other supportive contributions lie in handling imbalanced data which is commonly found in stream sentiment classification domain.

3- A method of stream sentiment classification is proposed to handle concept and feature drifts simultaneously using proposed feature unionization method.

1.9 Organization of thesis

The organization of the thesis is as follow:

- Chapter 1: It provides the introduction to the study domain, mainly sentiment classification and machine learning approaches. Then it discussed problem background, problem statement, research objective, scope and contribution.

- Chapter 2: It provides the intensive literature review of the study area, fundamental concepts of relative to this study and background, problems and potential solutions.

- Chapter 3: It provides research methodology used in this research. It discusses problem formulation of sentiment classification based on the literature review.

- Chapter 4: It introduces a novel proposed dimension reduction method based on feature unionization approach and discusses the experimental results.

- Chapter 5: It proposes the adaptive sentiment classification model to deal with concept drift and imbalanced data problem using ensemble learning method.

- Chapter 6: It continues with the improvement on the proposed adaptive sentiment classification model to handle feature drift using proposed feature unionization method.

- Chapter 7: It concludes the thesis with lists of contributions and recommends issues for future studies.

REFERENCES

Abbasi, A., and Chen, H. (2008). CyberGate: a design framework and system for text analysis of computer-mediated communication. *Mis Quarterly*, 811-837.

Abbasi, A., Chen, H., and Salem, A. (2008a). Sentiment analysis in multiple languages. *ACM Transactions on Information Systems, 26*(3), 1-34.

Abbasi, A., Chen, H., Thoms, S., and Fu, T. (2008b). Affect analysis of web forums and blogs using correlation ensembles. *IEEE Transactions on Knowledge and Data Engineering, 20*(9), 1168-1180.

Abbasi, A., France, S., Zhang, Z., and Chen, H. (2011). Selecting attributes for sentiment classification using feature relation networks. *IEEE Transactions on Knowledge and Data Engineering, 23*(3), 447-462.

Abdi, A., Shamsuddin, S. M., Hasan, S., and Piran, J. (2019). Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. *Information Processing & Management, 56*(4), 1245-1259.

Agarwal, B., and Mittal, N. (2016). Prominent feature extraction for review analysis: an empirical study. *Journal of Experimental & Theoretical Artificial Intelligence, 28*(3), 485-498.

Aggarwal, C. C., and Zhai, C. (2012). A survey of text classification algorithms. *Mining text data*, 163-222.

Ahuja, R., Chug, A., Kohli, S., Gupta, S., and Ahuja, P. (2019). The Impact of Features Extraction on the Sentiment Analysis. *Procedia Computer Science, 152*, 341-348.

Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., et al. (2016). *A Comparison between Preprocessing Techniques for Sentiment Analysis in Twitter*. Paper presented at the KDWeb.

Apolloni, J., Leguizamón, G., and Alba, E. (2016). Two hybrid wrapper-filter feature selection algorithms applied to high-dimensional microarray experiments. *Applied Soft Computing, 38*, 922-932.

Argamon, S., Whitelaw, C., Chase, P., Hota, S. R., Garg, N., and Levitan, S. (2007). Stylistic text classification using functional lexical features. *Journal of the Association for Information Science and Technology, 58*(6), 802-822.

Arif, M. H., Li, J., Iqbal, M., and Liu, K. (2017). Sentiment analysis and spam detection in short informal text using learning classifier systems. *Soft Computing*, 1-11.

Asghar, M. Z., Khan, A., Zahra, S. R., Ahmad, S., and Kundi, F. M. (2017). Aspect-based opinion mining framework using heuristic patterns. *Cluster Computing*, 1-19.

Baeza-Yates, R., and Ribeiro-Neto, B. (1999). *Modern information retrieval* (Vol. 463): ACM press New York.

Bai, X. (2011). Predicting consumer sentiments from online text. *Decision Support Systems, 50*(4), 732-742.

Banea, C., Mihalcea, R., and Wiebe, J. (2014). Sense-level subjectivity in a multilingual setting. *Computer Speech & Language, 28*(1), 7-19.

Bao, Y., Quan, C., Wang, L., and Ren, F. (2014). *The role of pre-processing in twitter sentiment analysis*. Paper presented at the International Conference on Intelligent Computing, 615-624.

Barddal, J. P., Gomes, H. M., Enembreck, F., and Pfahringer, B. (2017). A survey on feature drift adaptation: Definition, benchmark, challenges and future directions. *Journal of Systems and Software, 127*, 278-294.

Basari, A. S. H., Hussin, B., Ananta, I. G. P., and Zeniarja, J. (2013). Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization. *Procedia Engineering, 53*, 453-462.

Behera, R. N., Manan, R., and Dash, S. (2016). Ensemble based Hybrid Machine Learning Approach for Sentiment Classification-A Review. *International Journal of Computer Applications, 146*(6).

Bermejo, P., Gámez, J., and Puerta, J. (2008). *On incremental wrapper-based attribute selection: experimental analysis of the relevance criteria*. Paper presented at the Proceedings of IPMU, 638-645.

Bermejo, P., Gámez, J. A., and Puerta, J. M. (2009). *Incremental Wrapper-based subset Selection with replacement: An advantageous alternative to sequential forward selection*. Paper presented at the Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium on, 367-374.

Bezdek, J. C., and Kuncheva, L. I. (2001). Nearest prototype classifier designs: An experimental study. *International journal of Intelligent systems, 16*(12), 1445-1473.

Bickerstaffe, A., and Zukerman, I. (2010). *A hierarchical classifier applied to multi-way sentiment detection*. Paper presented at the Proceedings of the 23rd international conference on computational linguistics, 62-70.

Blitzer, J., Dredze, M., and Pereira, F. (2007). *Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification*. Paper presented at the ACL, 440-447.

Blum, A. L., and Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial intelligence, 97*(1), 245-271.

Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science, 2*(1), 1-8.

Bouazizi, M., and Ohtsuki, T. (2019). Multi-class sentiment analysis on twitter: Classification performance and challenges. *Big Data Mining and Analytics, 2*(3), 181-194.

Bravo-Marquez, F., Mendoza, M., and Poblete, B. (2014). Meta-level sentiment models for big social data analysis. *Knowledge-Based Systems, 69*, 86-99.

Brighton, H., and Mellish, C. (2002). Advances in instance selection for instance-based learning algorithms. *Data mining and knowledge discovery, 6*(2), 153-172.

Burgun, A., and Bodenreider, O. (2001). *Comparing terms, concepts and semantic classes in WordNet and the Unified Medical Language System*. Paper presented at the Proceedings of the NAACL'2001 Workshop,"WordNet and Other Lexical Resources: Applications, Extensions and Customizations, 77-82.

Burns, N., Bi, Y., Wang, H., and Anderson, T. (2011). Sentiment analysis of customer reviews: Balanced versus unbalanced datasets. *Knowledge-Based and Intelligent Information and Engineering Systems*, 161-170.

Caises, Y., González, A., Leyva, E., and Pérez, R. (2009). *SCIS: combining instance selection methods to increase their effectiveness over a wide range of domains*. Paper presented at the International Conference on Intelligent Data Engineering and Automated Learning, 17-24.

Cano, J. R., Herrera, F., and Lozano, M. (2006). On the combination of evolutionary algorithms and stratified strategies for training set selection in data mining. *Applied Soft Computing, 6*(3), 323-332.

Cao, M. D., and Zukerman, I. (2012). *Experimental evaluation of a lexicon-and corpus-based ensemble for multi-way sentiment analysis.* Paper presented at the Australasian Language Technology Association Workshop 2012, 52.

Catal, C., and Nangir, M. (2017). A sentiment classification model based on multiple classifiers. *Applied Soft Computing, 50*, 135-141.

Chandrashekar, G., and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering, 40*(1), 16-28.

Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST), 2*(3), 27.

Chatterjee, M., Watson, E., Perillo, K., and Turan, R. (2018). Method and system of an opinion search engine with an application programming interface for providing an opinion web portal: US Patent App. 15/065,559.

Chen, B., Huang, Q., Chen, Y., Cheng, L., and Chen, R. (2018). *Deep Neural Networks for Multi-class Sentiment Classification.* Paper presented at the 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 854-859.

Chen, C., Zhuo, R., and Ren, J. (2019a). Gated recurrent neural network with sentimental relations for sentiment classification. *Information Sciences, 502*, 268-278.

Chen, J., Huang, H., Tian, S., and Qu, Y. (2009). Feature selection for text classification with Naïve Bayes. *Expert Systems with Applications, 36*(3), 5432-5435.

Chen, K., Koh, Y. S., and Riddle, P. (2016). *Proactive drıft detection: Predicting concept drıfts in data streams using probabilistic networks.* Paper presented at the IJCNN, 780-787.

Chen, T., Su, P., Shang, C., Hill, R., Zhang, H., and Shen, Q. (2019b). *Sentiment Classification of Drug Reviews Using Fuzzy-rough Feature Selection.* Paper presented at the 2019 IEEE International Conference on Fuzzy Systems.

Choi, Y., and Cardie, C. (2008). *Learning with compositional semantics as structural inference for subsentential sentiment analysis.* Paper presented at the Proceedings of the Conference on Empirical Methods in Natural Language Processing, 793-801.

Chou, C.-H., Kuo, B.-H., and Chang, F. (2006). *The generalized condensed nearest neighbor rule as a data reduction method.* Paper presented at the Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, 556-559.

Colas, F., and Brazdil, P. (2006). Comparison of SVM and some older classification algorithms in text classification tasks. *Artificial Intelligence in Theory and Practice*, 169-178.

Colas, F. P. R. (2009). *Data mining scenarios for the discovery of subtypes and the comparison of algorithms*: Leiden Institute of Advanced Computer Science (LIACS), Faculty of Science, Leiden University.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory, 13*(1), 21-27.

Dang, Y., Zhang, Y., and Chen, H. (2010). A lexicon-enhanced method for sentiment classification: An experiment on online product reviews. *IEEE Intelligent Systems, 25*(4), 46-53.

Dasgupta, A., Drineas, P., Harb, B., Josifovski, V., and Mahoney, M. W. (2007). *Feature selection methods for text classification.* Paper presented at the Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 230-239.

de Mello, R. F., Vaz, Y., Grossi, C. H., and Bifet, A. (2019). On learning guarantees to unsupervised concept drift detection on data streams. *Expert Systems with Applications, 117*, 90-102.

Deng, Z.-H., Luo, K.-H., and Yu, H.-L. (2014). A study of supervised term weighting scheme for sentiment analysis. *Expert Systems with Applications, 41*(7), 3506-3513.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*: John Wiley & Sons.

Duric, A., and Song, F. (2012). Feature selection for sentiment analysis based on content and syntax models. *Decision Support Systems, 53*(4), 704-711.

Elwell, R., and Polikar, R. (2011). Incremental learning of concept drift in nonstationary environments. *IEEE Transactions on Neural Networks, 22*(10), 1517-1531.

Esuli, A., and Sebastiani, F. (2007). SentiWordNet: a high-coverage lexical resource for opinion mining. *Evaluation*, 1-26.

Farid, D. M., Zhang, L., Hossain, A., Rahman, C. M., Strachan, R., Sexton, G., et al. (2013). An adaptive ensemble classifier for mining concept drifting data streams. *Expert Systems with Applications, 40*(15), 5895-5906.

Fei, G., Li, H., and Liu, B. (2017). Opinion Spam Detection in Social Networks. In *Sentiment Analysis in Social Networks* (pp. 141-156): Elsevier.

FernáNdez, A., LóPez, V., Galar, M., Del Jesus, M. J., and Herrera, F. (2013). Analysing the classification of imbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches. *Knowledge-based systems, 42*, 97-110.

Fersini, E., Messina, E., and Pozzi, F. A. (2014). Sentiment analysis: Bayesian Ensemble Learning. *Decision Support Systems, 68*, 26-38.

Fleuret, F. (2004). Fast binary feature selection with conditional mutual information. *Journal of Machine Learning Research, 5*(Nov), 1531-1555.

Flores, M. J., and Gámez, J. A. (2005). *Breeding value classification in manchego sheep: a study of attribute selection and construction.* Paper presented at the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 1338-1346.

Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of machine learning research, 3*(Mar), 1289-1305.

Freeman, C., Kulić, D., and Basir, O. (2015). An evaluation of classifier-specific filter measure performance for feature selection. *Pattern Recognition, 48*(5), 1812-1826.

Gama, J. (2010). *Knowledge discovery from data streams*: CRC Press.

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys (CSUR), 46*(4), 44.

Gamon, M. (2004). *Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis.* Paper presented at the Proceedings of the 20th international conference on Computational Linguistics, 841.

Gao, J., Ding, B., Fan, W., Han, J., and Philip, S. Y. (2008). Classifying data streams with skewed class distributions and concept drifts. *IEEE Internet Computing, 12*(6).

Garcia, S., Derrac, J., Cano, J., and Herrera, F. (2012). Prototype selection for nearest neighbor classification: Taxonomy and empirical study. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(3), 417-435.

Gates, G. (1972). The reduced nearest neighbor rule (Corresp.). *IEEE transactions on information theory, 18*(3), 431-433.

Ghag, K. V., and Shah, K. (2015). *Comparative analysis of effect of stopwords removal on sentiment classification.* Paper presented at the Computer, Communication and Control (IC4), 2015 International Conference on, 1-6.

Ghosh, M., and Sanyal, G. (2018). *Document Modeling with Hierarchical Deep Learning Approach for Sentiment Classification.* Paper presented at the Proceedings of the 2nd International Conference on Digital Signal Processing, 181-185.

Gomes, H. M., Barddal, J. P., Enembreck, F., and Bifet, A. (2017). A Survey on Ensemble Learning for Data Stream Classification. *ACM Computing Surveys (CSUR), 50*(2), 23.

Gonçalves Jr, P. M., de Carvalho Santos, S. G., Barros, R. S., and Vieira, D. C. (2014). A comparative study on concept drift detectors. *Expert Systems with Applications, 41*(18), 8144-8156.

Güngör, T. (2010). Part-of-Speech Tagging.

Guyon, I., and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research, 3*(Mar), 1157-1182.

Haddi, E., Liu, X., and Shi, Y. (2013). The role of text pre-processing in sentiment analysis. *Procedia Computer Science, 17*, 26-32.

Hall, M. A. (1999). Correlation-based feature selection for machine learning.

Hardin, D., Tsamardinos, I., and Aliferis, C. F. (2004). *A theoretical characterization of linear SVM-based feature selection.* Paper presented at the Proceedings of the twenty-first international conference on Machine learning, 48.

Hart, P. (1968). The condensed nearest neighbor rule (Corresp.). *IEEE transactions on information theory, 14*(3), 515-516.

Heydari, A., ali Tavakoli, M., Salim, N., and Heydari, Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications, 42*(7), 3634-3642.

Hoang, T., Le, H., and Quan, T. (2019). Towards Autoencoding Variational Inference for Aspect-Based Opinion Summary. *Applied Artificial Intelligence*, 1-21.

Hoens, T. R., Polikar, R., and Chawla, N. V. (2012). Learning from streaming data with concept drift and imbalance: an overview. *Progress in Artificial Intelligence, 1*(1), 89-101.

Hong, H., Xu, D., Wang, G. A., and Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems, 102*, 1-11.

Idris, I., Selamat, A., Nguyen, N. T., Omatu, S., Krejcar, O., Kuca, K., et al. (2015). A combined negative selection algorithm–particle swarm optimization for an email spam detection system. *Engineering Applications of Artificial Intelligence, 39*, 33-44.

Islam, M. T., Sazzad, M. A. U., Ahmed, E., Azad, M., and Islam, S. (2018). *Subjectivity Analysis of Movie Reviews using Machine Learning to Perform Sentence Level Classification*. Paper presented at the 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), 473-478.

Jianqiang, Z., and Xiaolin, G. (2017). Comparison Research on Text Pre-processing Methods on Twitter Sentiment Analysis. *IEEE Access, 5*, 2870-2879.

John, G. H., Kohavi, R., and Pfleger, K. (1994). *Irrelevant features and the subset selection problem*. Paper presented at the Machine learning: proceedings of the eleventh international conference, 121-129.

Joshi, M., and Penstein-Rosé, C. (2009). *Generalizing dependency features for opinion mining*. Paper presented at the Proceedings of the ACL-IJCNLP 2009 conference short papers, 313-316.

Kalaivani, P., and Shunmuganathan, K. (2014). *An improved K-nearest-neighbor algorithm using genetic algorithm for sentiment classification*. Paper presented at the Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference on, 1647-1651.

Kalaivani, P., and Shunmuganathan, K. (2015). Feature reduction based on genetic algorithm and hybrid model for opinion mining. *Scientific Programming, 2015*, 12.

Kang, H., Yoo, S. J., and Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications, 39*(5), 6000-6010.

Kennedy, A., and Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational intelligence, 22*(2), 110-125.

Khamassi, I., Sayed-Mouchaweh, M., Hammami, M., and Ghédira, K. (2013). *Ensemble classifiers for drift detection and monitoring in dynamical environments*. Paper presented at the Annual conference of the Prognostics and Health Management Society. New Orlean.

Khamassi, I., Sayed-Mouchaweh, M., Hammami, M., and Ghédira, K. (2018). Discussion and review on evolving data streams and concept drift adapting. *Evolving systems, 9*(1), 1-23.

Khamassi, I., Sayed-Mouchaweh, M., Hammami, M., and Ghédira, K. (2019). A New Combination of Diversity Techniques in Ensemble Classifiers for Handling Complex Concept Drift. In *Learning from Data Streams in Evolving Environments* (pp. 39-61): Springer.

Khan, J., Alam, A., Hussain, J., and Lee, Y.-K. (2019). EnSWF: effective features extraction and selection in conjunction with ensemble learning methods for document sentiment classification. *Applied Intelligence*, 1-23.

Kim, K., and Zzang, S. Y. (2019). Trigonometric comparison measure: A feature selection method for text categorization. *Data & Knowledge Engineering, 119*, 1-21.

Kim, S.-M., and Hovy, E. (2004). *Determining the sentiment of opinions*. Paper presented at the Proceedings of the 20th international conference on Computational Linguistics, 1367.

Kittler, J. (1980). *On the edited nearest neighbor rule*.

Klinkenberg, R., and Joachims, T. (2000). *Detecting Concept Drift with Support Vector Machines*. Paper presented at the ICML, 487-494.

Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Paper presented at the Ijcai, 1137-1145.

Kolter, J. Z., and Maloof, M. A. (2007). Dynamic weighted majority: An ensemble method for drifting concepts. *Journal of Machine Learning Research, 8*(Dec), 2755-2790.

Krawczyk, B., and Woźniak, M. (2015). One-class classifiers with incremental learning and forgetting for data streams with concept drift. *Soft Computing, 19*(12), 3387-3400.

Krempl, G., Žliobaite, I., Brzeziński, D., Hüllermeier, E., Last, M., Lemaire, V., et al. (2014). Open challenges for data stream mining research. *ACM SIGKDD explorations newsletter, 16*(1), 1-10.

Krishnamoorthy, S. (2015). Linguistic features for review helpfulness prediction. *Expert Systems with Applications, 42*(7), 3751-3759.

Kumar, C. P., and Babu, L. D. (2019). Novel Text Preprocessing Framework for Sentiment Analysis. In *Smart Intelligent Computing and Applications* (pp. 309-317): Springer.

Kuncheva, L. I. (1995). Editing for the k-nearest neighbors rule by a genetic algorithm. *Pattern Recognition Letters, 16*(8), 809-814.

Kuncheva, L. I. (2001). Combining classifiers: Soft computing solutions. *Pattern recognition: From classical to modern approaches*, 427-451.

Kuncheva, L. I. (2004). *Classifier ensembles for changing environments*. Paper presented at the International Workshop on Multiple Classifier Systems, 1-15.

Kuncheva, L. I., and Žliobaitė, I. (2009). On the window size for classification in changing environments. *Intelligent Data Analysis, 13*(6), 861-872.

Lafferty, J., McCallum, A., and Pereira, F. C. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data.

Lee, S., and Choeh, J. Y. (2014). Predicting the helpfulness of online reviews using multilayer perceptron neural networks. *Expert Systems with Applications, 41*(6), 3041-3046.

Li, S., Ju, S., Zhou, G., and Li, X. (2012a). *Active learning for imbalanced sentiment classification*. Paper presented at the Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 139-148.

Li, S., Wang, Z., Zhou, G., and Lee, S. Y. M. (2011). *Semi-supervised learning for imbalanced sentiment classification*. Paper presented at the IJCAI proceedings-international joint conference on artificial intelligence, 1826.

Li, S., Xia, R., Zong, C., and Huang, C.-R. (2009). *A framework of feature selection methods for text categorization*. Paper presented at the Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th

International Joint Conference on Natural Language Processing of the AFNLP: Volume 2-Volume 2, 692-700.

LI, W., WANG, W., and CHEN, Y. (2012b). Heterogeneous ensemble learning for Chinese sentiment classification. *JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE, 9*(15), 4551-4558.

Li, Y., Guo, H., Zhang, Q., Gu, M., and Yang, J. (2018). Imbalanced text sentiment classification using universal and domain-specific knowledge. *Knowledge-Based Systems, 160*, 1-15.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies, 5*(1), 1-167.

Liu, B., and Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463): Springer.

Liu, H., and Motoda, H. (2002). On issues of instance selection. *Data Mining and Knowledge Discovery, 6*(2), 115-130.

Liu, H., and Motoda, H. (2012). *Feature selection for knowledge discovery and data mining* (Vol. 454): Springer Science & Business Media.

Liu, H., and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on knowledge and data engineering, 17*(4), 491-502.

Liu, S. M., and Chen, J.-H. (2015). A multi-label classification based approach for sentiment classification. *Expert Systems with Applications, 42*(3), 1083-1093.

Liu, Y., Bi, J.-W., and Fan, Z.-P. (2017). Multi-class sentiment classification: The experimental comparisons of feature selection and machine learning algorithms. *Expert Systems with Applications, 80*, 323-339.

Liu, Y., Jin, J., Ji, P., Harding, J. A., and Fung, R. Y. (2013). Identifying helpful online reviews: a product designer's perspective. *Computer-Aided Design, 45*(2), 180-194.

Liu, Y., Pang, B., and Wang, X. (2019). Opinion spam detection by incorporating multimodal embedded representation into a probabilistic review graph. *Neurocomputing*.

Lu, B., and Tsou, B. K. (2010). *Combining a large sentiment lexicon and machine learning for subjectivity classification*. Paper presented at the Machine Learning and Cybernetics (ICMLC), 2010 International Conference on, 3311-3316.

Mahajan, S., and Rana, V. (2017). Spam Detection on Social Network through Sentiment Analysis. *Advances in Computational Sciences and Technology, 10*(8), 2225-2231.

Maks, I., and Vossen, P. (2012). A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems, 53*(4), 680-688.

MartíN-Valdivia, M.-T., MartíNez-CáMara, E., Perea-Ortega, J.-M., and UreñA-LóPez, L. A. (2013). Sentiment polarity detection in Spanish reviews combining supervised and unsupervised approaches. *Expert Systems with Applications, 40*(10), 3934-3942.

Martineau, J., and Finin, T. (2009). Delta TFIDF: An Improved Feature Space for Sentiment Analysis. *Icwsm, 9*, 106.

Matsumoto, S., Takamura, H., and Okumura, M. (2005). *Sentiment Classification Using Word Sub-sequences and Dependency Sub-trees*. Paper presented at the PAKDD, 301-311.

McCallum, A., and Nigam, K. (1998). *A comparison of event models for naive bayes text classification*. Paper presented at the AAAI-98 workshop on learning for text categorization, 41-48.

McCallum, A., Pal, C., Druck, G., and Wang, X. (2006). *Multi-conditional learning: Generative/discriminative training for clustering and classification*. Paper presented at the AAAI, 433-439.

McDonald, R., Hannan, K., Neylon, T., Wells, M., and Reynar, J. (2007). *Structured models for fine-to-coarse sentiment analysis*. Paper presented at the Annual meeting-association for computational linguistics, 432.

Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal, 5*(4), 1093-1113.

Mena-Torres, D., and Aguilar-Ruiz, J. S. (2014). A similarity-based approach for data stream classification. *Expert Systems with Applications, 41*(9), 4224-4234.

Mirza, B., Lin, Z., and Liu, N. (2015). Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift. *Neurocomputing, 149*, 316-329.

Mishne, G. (2005). *Experiments with mood classification in blog posts*. Paper presented at the Proceedings of ACM SIGIR 2005 workshop on stylistic analysis of text for information access, 321-327.

Moh'd Mesleh, A. (2011). Feature sub-set selection metrics for Arabic text classification. *Pattern Recognition Letters, 32*(14), 1922-1929.

Mohammadiani, R. P., Mohammadi, S., and Malik, Z. (2017). Understanding the relationship strengths in users' activities, review helpfulness and influence. *Computers in Human Behavior, 75*, 117-129.

Molina-González, M. D., Martínez-Cámara, E., Martín-Valdivia, M.-T., and Perea-Ortega, J. M. (2013). Semantic orientation for polarity classification in Spanish reviews. *Expert Systems with Applications, 40*(18), 7250-7257.

Mollineda, R. A., Ferri, F. J., and Vidal, E. (2002). An efficient prototype merging strategy for the condensed 1-NN rule through class-conditional hierarchical clustering. *Pattern Recognition, 35*(12), 2771-2782.

Moraes, R., Valiati, J. F., and Gavião Neto, W. P. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications, 40*(2), 621-633.

Muhlbaier, M. D., and Polikar, R. (2007). *Multiple classifiers based incremental learning algorithm for learning in nonstationary environments*. Paper presented at the Machine Learning and Cybernetics, 2007 International Conference on, 3618-3623.

Mullen, T., and Collier, N. (2004). *Sentiment Analysis using Support Vector Machines with Diverse Information Sources*. Paper presented at the EMNLP, 412-418.

Ng, V., Dasgupta, S., and Arifin, S. (2006). *Examining the role of linguistic knowledge sources in the automatic identification and classification of reviews*. Paper presented at the Proceedings of the COLING/ACL on Main conference poster sessions, 611-618.

Ngo-Ye, T. L., and Sinha, A. P. (2014). The influence of reviewer engagement characteristics on online review helpfulness: A text regression model. *Decision Support Systems, 61*, 47-58.

Nguyen, H. T., and Le Nguyen, M. (2019). An ensemble method with sentiment features and clustering support. *Neurocomputing*.

O'Connor, B., Balasubramanyan, R., Routledge, B. R., and Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *ICWSM, 11*(122-129), 1-2.

O'Keefe, T., and Koprinska, I. (2009). *Feature selection and weighting methods in sentiment analysis*. Paper presented at the Proceedings of the 14th Australasian document computing symposium, Sydney, 67-74.

Omar, N., Albared, M., Al-Moslmi, T., and Al-Shabi, A. (2014). *A comparative study of feature selection and machine learning algorithms for Arabic sentiment classification*. Paper presented at the Asia information retrieval symposium, 429-443.

Onan, A., and Korukoğlu, S. (2016a). Exploring Performance of Instance Selection Methods in Text Sentiment Classification. In *Artificial Intelligence Perspectives in Intelligent Systems* (pp. 167-179): Springer.

Onan, A., and Korukoğlu, S. (2016b). A feature selection model based on genetic rank aggregation for text sentiment classification. *Journal of Information Science, 43*(1), 25-38.

Onan, A., Korukoğlu, S., and Bulut, H. (2017). A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification. *Information Processing & Management, 53*(4), 814-833.

Paltoglou, G., and Thelwall, M. (2010). *A study of information retrieval weighting schemes for sentiment analysis*. Paper presented at the Proceedings of the 48th annual meeting of the association for computational linguistics, 1386-1395.

Pang, B., and Lee, L. (2004). *A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts*. Paper presented at the Proceedings of the 42nd annual meeting on Association for Computational Linguistics, 271.

Pang, B., and Lee, L. (2005). *Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales*. Paper presented at the Proceedings of the 43rd annual meeting on association for computational linguistics, 115-124.

Pang, B., and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval, 2*(1–2), 1-135.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). *Thumbs up?: sentiment classification using machine learning techniques*. Paper presented at the

Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 79-86.

Piryani, R., Gupta, V., Singh, V. K., and Ghose, U. (2017). A Linguistic Rule-Based Approach for Aspect-Level Sentiment Analysis of Movie Reviews. In *Advances in Computer and Computational Sciences* (pp. 201-209): Springer.

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and systems magazine, 6*(3), 21-45.

Porter, M. E., and Teisberg, E. O. (2006). *Redefining health care: creating value-based competition on results*: Harvard Business Press.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program, 14*(3), 130-137.

Prabowo, R., and Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics, 3*(2), 143-157.

Pratiwi, A. I. (2018). On the feature selection and classification based on information gain for document sentiment analysis. *Applied Computational Intelligence and Soft Computing, 2018*.

Rakitianskaia, A. S., and Engelbrecht, A. P. (2012). Training feedforward neural networks with dynamic particle swarm optimisation. *Swarm Intelligence, 6*(3), 233-270.

Rathan, M., Hulipalled, V. R., Venugopal, K., and Patnaik, L. (2017). Consumer insight mining: Aspect based Twitter opinion mining of mobile phone reviews. *Applied Soft Computing*.

Rehman, A., Javed, K., and Babri, H. A. (2017). Feature selection based on a normalized difference measure for text classification. *Information Processing & Management, 53*(2), 473-489.

Riloff, E., Patwardhan, S., and Wiebe, J. (2006). *Feature subsumption for opinion analysis*. Paper presented at the Proceedings of the 2006 conference on empirical methods in natural language processing, 440-448.

Riloff, E., Wiebe, J., and Wilson, T. (2003). *Learning subjective nouns using extraction pattern bootstrapping*. Paper presented at the Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4, 25-32.

Ritter, G., Woodruff, H., Lowry, S., and Isenhour, T. (1975). An algorithm for a selective nearest neighbor decision rule (Corresp.). *IEEE Transactions on Information Theory, 21*(6), 665-669.

Roveri, M. (2019). Learning Discrete-Time Markov Chains Under Concept Drift. *IEEE transactions on neural networks and learning systems*.

Ruiz, R., Riquelme, J. C., and Aguilar-Ruiz, J. S. (2006). Incremental wrapper-based gene selection from microarray data for cancer classification. *Pattern Recognition, 39*(12), 2383-2392.

Sáez, J. A., Krawczyk, B., and Woźniak, M. (2016). Analyzing the oversampling of different classes and types of examples in multi-class imbalanced datasets. *Pattern Recognition, 57*, 164-178.

Saleh, M. R., Martín-Valdivia, M. T., Montejo-Ráez, A., and Ureña-López, L. (2011). Experiments with SVM to classify opinions in different domains. *Expert Systems with Applications, 38*(12), 14799-14804.

Salvetti, F., Lewis, S., and Reichenbach, C. (2004). Automatic opinion polarity classification of movie. *Colorado research in linguistics, 17*(2).

Schouten, K., and Frasincar, F. (2016). Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering, 28*(3), 813-830.

Schütze, H. (2008). *Introduction to information retrieval.* Paper presented at the Proceedings of the international communication of association for computing machinery conference.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR), 34*(1), 1-47.

Sharma, A., and Dey, S. (2012). *A comparative study of feature selection and machine learning techniques for sentiment analysis*. Paper presented at the Proceedings of the 2012 ACM research in applied computation symposium, 1-7.

Sheikhpour, R., Sarram, M. A., and Sheikhpour, R. (2016). Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation based classifiers in diagnosis of breast cancer. *Applied Soft Computing, 40*, 113-131.

Siahroudi, S. K., Moodi, P. Z., and Beigy, H. (2018). Detection of evolving concepts in non-stationary data streams: A multiple kernel learning approach. *Expert Systems with Applications, 91*, 187-197.

Singh, T., and Kumari, M. (2016). Role of Text Pre-processing in Twitter Sentiment Analysis. *Procedia Computer Science, 89*, 549-554.

Spillmann, B., Neuhaus, M., Bunke, H., Pekalska, E., and Duin, R. P. (2006). *Transforming strings to vector spaces using prototype selection.* Paper presented at the SSPR/SPR, 287-296.

Su, Y., Zhang, Y., Ji, D., Wang, Y., and Wu, H. (2012). *Ensemble learning for sentiment classification.* Paper presented at the Workshop on Chinese Lexical Semantics, 84-93.

Sun, J., and Li, H. (2011). Dynamic financial distress prediction using instance selection for the disposal of concept drift. *Expert Systems with Applications, 38*(3), 2566-2576.

Sun, Y., Wang, Z., Li, H., and Li, Y. (2017). A Novel Ensemble Classification for Data Streams with Class Imbalance and Concept Drift. *International Journal of Performability Engineering, 13*(6), 945.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics, 37*(2), 267-307.

Tan, S., and Zhang, J. (2008). An empirical study of sentiment analysis for chinese documents. *Expert Systems with Applications, 34*(4), 2622-2629.

Tang, D. (2015). *Sentiment-specific representation learning for document-level sentiment analysis.* Paper presented at the Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, 447-452.

Tang, H., Tan, S., and Cheng, X.-q. (2007). Research on sentiment classification of Chinese reviews based on supervised machine learning techniques. *Journal of Chinese information processing, 21*(6), 88-94.

Thelwall, M., Buckley, K., and Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the Association for Information Science and Technology, 63*(1), 163-173.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the Association for Information Science and Technology, 61*(12), 2544-2558.

Tian, F., Wu, F., Fei, X., Shah, N., Zheng, Q., and Wang, Y. (2019). Improving generalization ability of instance transfer-based imbalanced sentiment classification of turn-level interactive Chinese texts. *Service Oriented Computing and Applications*, 1-13.

Tomek, I. (1976). An experiment with the edited nearest-neighbor rule. *IEEE Transactions on systems, Man, and Cybernetics*(6), 448-452.

Tripathy, A., Agrawal, A., and Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications, 57*, 117-126.

Tripathy, A., Anand, A., and Rath, S. K. (2017). Document-level sentiment classification using hybrid machine learning approach. *Knowledge and Information Systems*.

Tripathy, A., and Rath, S. K. (2017). Classification of Sentiment of Reviews using Supervised Machine Learning Techniques. *International Journal of Rough Sets and Data Analysis, 4*(1), 56-74.

Tsutsumi, K., Shimada, K., and Endo, T. (2007). *Movie review classification based on a multiple classifier*. Paper presented at the Proceedings of the 21st Pacific Asia Conference on Language, Information and Computation, 481-488.

Tu, Z., He, Y., Foster, J., van Genabith, J., Liu, Q., and Lin, S. (2012). *Identifying high-impact sub-structures for convolution kernels in document-level sentiment classification*. Paper presented at the Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, 338-343.

Tubishat, M., Abushariah, M. A., Idris, N., and Aljarah, I. (2019). Improved whale optimization algorithm for feature selection in Arabic sentiment analysis. *Applied Intelligence, 49*(5), 1688-1707.

Wang, G., Sun, J., Ma, J., Xu, K., and Gu, J. (2014). Sentiment classification: The contribution of ensemble learning. *Decision Support Systems, 57*, 77-93.

Wang, G., Zhang, Z., Sun, J., Yang, S., and Larson, C. A. (2015). POS-RS: A Random Subspace method for sentiment classification based on part-of-speech analysis. *Information Processing & Management, 51*(4), 458-479.

Wang, M., Ning, Z.-H., Xiao, C., and Li, T. (2018). Sentiment Classification Based on Information Geometry and Deep Belief Networks. *IEEE Access, 6*, 35206-35213.

Wang, S., Li, D., Song, X., Wei, Y., and Li, H. (2011a). A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. *Expert Systems with Applications, 38*(7), 8696-8702.

Wang, S., Li, D., Zhao, L., and Zhang, J. (2013). Sample cutting method for imbalanced text sentiment classification based on BRC. *Knowledge-Based Systems, 37*, 451-461.

Wang, S., Wei, Y., Li, D., Zhang, W., and Li, W. (2007). *A hybrid method of feature selection for Chinese text sentiment classification.* Paper presented at the Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007. Fourth International Conference on, 435-439.

Wang, Z., He, Y., and Jiang, M. (2006). *A comparison among three neural networks for text classification.* Paper presented at the Signal Processing, 2006 8th International Conference on.

Wang, Z., Li, S., Zhou, G., Li, P., and Zhu, Q. (2011b). *Imbalanced sentiment classification with multi-strategy ensemble learning.* Paper presented at the Asian Language Processing (IALP), 2011 International Conference on, 131-134.

Wen, Y., Xu, A., Liu, W., and Chen, L. (2018). *A Wide Residual Network for Sentiment Classification.* Paper presented at the Proceedings of the 2018 2nd International Conference on Deep Learning Technologies, 7-11.

Whitehead, M., and Yaeger, L. (2009). *Building a general purpose cross-domain sentiment mining model.* Paper presented at the Computer Science and Information Engineering, 2009 WRI World Congress on, 472-476.

Whitehead, M., and Yaeger, L. (2010). Sentiment mining using ensemble classification models. In *Innovations and advances in computer sciences and engineering* (pp. 509-514): Springer.

Whitelaw, C., Garg, N., and Argamon, S. (2005). *Using appraisal groups for sentiment analysis.* Paper presented at the Proceedings of the 14th ACM international conference on Information and knowledge management, 625-631.

Wiebe, J., Wilson, T., Bruce, R., Bell, M., and Martin, M. (2006). Learning subjective language. *Learning, 30*(3).

Wilson, D. L. (1972). Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics, 2*(3), 408-421.

Wilson, D. R., and Martinez, T. R. (2000). Reduction techniques for instance-based learning algorithms. *Machine learning, 38*(3), 257-286.

Wilson, T., Wiebe, J., and Hwa, R. (2006). Recognizing strong and weak opinion clauses. *Computational Intelligence, 22*(2), 73-99.

Wu, P., Li, X., Shen, S., and He, D. (2019). Social media opinion summarization using emotion cognition and convolutional neural networks. *International Journal of Information Management*.

Xia, R., Xu, F., Yu, J., Qi, Y., and Cambria, E. (2016). Polarity shift detection, elimination and ensemble: A three-stage model for document-level sentiment analysis. *Information Processing & Management, 52*(1), 36-45.

Xia, R., Zong, C., and Li, S. (2011). Ensemble of feature sets and classification algorithms for sentiment classification. *Information Sciences, 181*(6), 1138-1152.

Yang, D.-H., and Yu, G. (2013). A method of feature selection and sentiment similarity for Chinese micro-blogs. *Journal of Information Science, 39*(4), 429-441.

Yang, Y., and Pedersen, J. O. (1997). *A comparative study on feature selection in text categorization*. Paper presented at the Icml, 412-420.

Yu, H., and Hatzivassiloglou, V. (2003). *Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences*. Paper presented at the Proceedings of the 2003 conference on Empirical methods in natural language processing, 129-136.

Yu, L., and Liu, H. (2004). Efficient feature selection via analysis of relevance and redundancy. *Journal of machine learning research, 5*(Oct), 1205-1224.

Zaidan, O., Eisner, J., and Piatko, C. D. (2007). *Using" Annotator Rationales" to Improve Machine Learning for Text Categorization*. Paper presented at the HLT-NAACL, 260-267.

Zhai, Z., Xu, H., Kang, B., and Jia, P. (2011). Exploiting effective features for chinese sentiment classification. *Expert Systems with Applications, 38*(8), 9139-9146.

Zhang, D., Xu, H., Su, Z., and Xu, Y. (2015). Chinese comments sentiment classification based on word2vec and SVMperf. *Expert Systems with Applications, 42*(4), 1857-1863.

Zhang, Y., Miao, D., Wang, J., and Zhang, Z. (2019). A cost-sensitive three-way combination technique for ensemble learning in sentiment classification. *International Journal of Approximate Reasoning, 105*, 85-97.

Zhang, Z., Ye, Q., Zhang, Z., and Li, Y. (2011). Sentiment classification of Internet restaurant reviews written in Cantonese. *Expert Systems with Applications, 38*(6), 7674-7682.

Zheng, L., Wang, H., and Gao, S. (2015). Sentimental feature selection for sentiment analysis of Chinese online reviews. *International Journal of Machine Learning and Cybernetics*.

Zhou, Z.-H. (2012). *Ensemble methods: foundations and algorithms*: CRC press.

Žliobaitė, I., Pechenizkiy, M., and Gama, J. (2016). An overview of concept drift applications. In *Big Data Analysis: New Algorithms for a New Society* (pp. 91-114): Springer.

Zong, C., Feng, W., Zheng, V. W., and Zhuo, H. H. (2018). *Adaptive Attention Network for Review Sentiment Classification.* Paper presented at the Pacific-Asia Conference on Knowledge Discovery and Data Mining, 668-680.

Zou, B., Zhu, Q., and Zhou, G. (2015). *Negation and Speculation Identification in Chinese Language.* Paper presented at the ACL (1), 656-665.

Zuo, Y., Wu, J., Zhang, H., Wang, D., and Xu, K. (2017). Complementary Aspect-based Opinion Mining. *IEEE Transactions on Knowledge and Data Engineering*.

LIST OF PUBLICATIONS

Journal with Impact Factor

1. **Jalilvand, A.**, and Salim, N. (2017). Feature unionization: A novel approach for dimension reduction. *Applied Soft Computing, 52*, 1253-1261. https://doi.org/10.1016/j.asoc.2016.08.031 **(Q1, IF: 4.873)**

Indexed Journal

1. **Jalilvanda, A.**, and Salima, N. (2016). SENTIMENT CLASSIFICATION WITH CONCEPT DRIFT AND IMBALANCED CLASS DISTRIBUTIONS. *JURNAL TEKNOLOGI, 78*(12-2), 35-42. https://doi.org/10.11113/jt.v78.10120. **(Indexed by SCOPUS)**

Indexed Conference Proceedings

1. **Jalilvand, A.**, & Salim, N. (2012). Sentiment classification using graph based word sense disambigution. *In International Conference on Advanced Machine Learning Technologies and Applications* (pp. 351-358). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35326-0_35. **(Indexed by Springer)**