Sentiment Classification in the Financial Domain using ν–SVM and Multi-Objective Optimisation

Fan Sun[1] sun-f@email.ulster.ac.uk
Ammar Belatreche[1] a.belatreche@ulster.ac.uk
Sonya Coleman[1] sa.coleman@ulster.ac.uk
T. M. McGinnity[2] martin.mcginnity@ntu.ac.uk
Yuhua Li[3] yuhua.li@salford.ac.uk

(1) Intelligent Systems Research Centre
Faculty of Computing and Engineering
University of Ulster, Magee, UK
(2) College of Arts and Science, School of Science & Technology
Nottingham Trent University
(3) School of Computing, Science and Engineering
University of Salford, Manchester, UK

***Abstract*— Online financial textual information containing a large amount of investor sentiment is growing rapidly and an effective solution to automate the sentiment classification of such large amounts of text would be extremely beneficial. A novel approach to sentiment classification is the application of multi-objective optimization combined with ν-SVM to improve the overall accuracy and hence we present a Multi-Objective Genetic Algorithm (MOGA) based approach to automatically adjust the free parameters of a ν-SVM classifier to optimise sentiment classification performance. The approach has been implemented and tested using two online financial textual datasets and experimental results show that the overall classification accuracy has improved (4%-7%) compared with other baseline approaches.**

I. INTRODUCTION

Online financial textual information contains a large amount of investor sentiment, i.e. subjective assessment and discussion with respect to financial instruments. Online financial information plays an increasingly important role in financial markets and personal finance for analysis and decision making. Sentiment classification in the financial domain is the process of extracting the emotive content from financial texts and classifying the content as expressing a positive or negative sentiment [1]. The main existing sentiment classification approaches are based on machine learning and semantic-based techniques [2]. Several common classifiers are used in text classification by machine learning techniques, such as k-Nearest Neighbor (k-NN), Artificial Neural Networks (ANNs), Support Vector Machine (SVM), Decision Trees, Naive Bayes (NB) and Maximum Entropy (ME) [3, 4]. Using a classifier for sentiment classification requires fine tuning the classifier parameters to obtain optimal sentiment classification performance.

Parameter estimation is the process of modifying the free parameters in a learning model until the output from the model matches an observed set of data [5]. Automatic parameter estimation using optimization algorithms is relatively easy to implement [5]. In order to obtain optimal sentiment classification performance using automatic parameter estimation, there are three questions that need to be decided: (1) what objective function(s) should be optimized? (2) which machine learning algorithms should be selected for sentiment classification? and (3) which optimization technique should be used?

Single objective optimization applied to machine learning classification has been explored in optimization based research [6, 7]. However, only optimizing a single objective, such as overall accuracy, cannot guarantee selection of optimal parameter(s) of the model, as it may lead to bias and variance in the case of unbalanced classes or asymmetric misclassification costs [7]. As an alternative, using multi-objective optimization (MOO) techniques for model parameter estimation has been used in several works [5,8,9,10]. For example, Ethridge *et al.* [8] designed a multi-objective evolutionary algorithm (MOEA) to maximize sensitivity and specificity of SVM and carried out experiments to solve the over-fitting problem. Clark and Everson [9] proposed a MOEA method to optimize Relevance Vector Machines to overcome the over-fitting problem and Liu and Sun [5] applied two evolutionary algorithms (EA), Non-dominated Sorting Differential Evolution (NSDE) and Non-dominated Sorting Genetic Algorithm (NSGA-II)), on ν-support vector regression (ν-SVR) for parameter estimation of a pressure swing adsorption (PSA) model.

Evolutionary algorithms, such as genetic algorithms and particle swarm optimization, are popular heuristic optimization techniques [8]. Genetic Algorithms (GA) are a popular choice for global optimization applied on parameter estimation, because of their simplicity, global perspective, and inherent parallel processing [5]. In this paper, we present an approach in which a MOGA is combined with ν-SVM in order to automatically select the optimal ν-SVM parameters for

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.134

application to sentiment analysis. A multi-objective optimization method using NSGA-II is applied to financial sentiment classification to achieve enhanced classification accuracy. This is a new application in sentiment analysis where we simultaneously optimize two ν-SVM parameters, ν and λ , , so as to maximize both the correct positive rate and the correct negative rate simultaneously. The goal of introducing MOEA optimization into sentiment classification in the financial domain is to generate an optimal Pareto front that contains a set of optimal trade-off solutions between correct positive rate and correct negative rate. The Pareto front not only contains the set of optimal parameters, which can maximize the overall sentiment classification accuracy, but also offers flexibility for investors to choose a solution based on their subjective preference information. For example, if an investor would like to purchase a stock and is interested in positive sentiment for the stock, s/he will select the optimal parameter on the Pareto front which can maximize the correct positive rate, and vice versa.

The remainder of this paper is structured as follows. We firstly outline the basic theory of ν-SVM and MOEA techniques in Sections II. Section III presents a description of the online financial text used in the experimental evaluation. Section IV presents and analyses the experimental results obtained and compares them with an existing approach based on textual pre-processing for sentiment analysis. Finally, conclusions and future research directions are presented in Section V.

II. BACKGROUND

Sentiment analysis of online text has attracted significant attention from researchers since the early 2000s [11], and research on extracting and classifying investors' sentiment from online financial-domain text has been growing rapidly with the development of sentiment analysis. This paper focuses on machine learning based sentiment classification and a multi-objective optimization method is applied to search for optimal parameters of the classifier algorithm in machine learning sentiment classification. Therefore related research on machine learning for sentiment analysis in the financial domain will be discussed in this section alongside the basic theory of ν-SVM and MOEA techniques.

A. Financial domain machine learning sentiment classification in short informal text

Machine learning based sentiment classification of online financial text has been performed in several research studies [2, 12, 13, 14]. Popular approaches include k-Nearest Neighbour (kNN), Artificial Neural Networks (ANN), Support Vector Machine (SVM), Decision Tree, Naive Bayes (NB) and Maximum Entropy (ME) [11]. Thelwall [12] used ME for sentiment classification and an NB classifier to extract sentiment strength from informal English text. Machine learning approaches have demonstrated the usefulness of sentiment extracted from web based financial information. Chua [13] employed a variation of NB classifiers to classify internet stock message boards associated with term frequency and information gain feature selection methods and achieved an accuracy of 78.72%. Das and Chen [14] proposed a sentiment analysis framework which used both machine learning and corpus-based methods to analyse messages drawn from stock message boards. The Naïve classifier, vector distance classifier and Bayesian classifier were used and the Naïve classifier performed best with 92.25% accuracy when conducting in-sample tests. However, the vector distance classifier achieved the best performance when applied to unseen data with an accuracy of 39.1%, the Naïve classifier only achieved 25.73%.

B. ν-SVM

ANNs and SVMs are the most common classifiers used in sentiment classification [15, 16]. ANN can learn the mapping relationship between the inputs and the outputs sampled from a training set using a supervised learning algorithm. The trained ANN is then used to make a prediction for the test data. SVM uses a Lagrangian formulation to maximise the margin of separation between the training data points of two classes. It finds a kernel function that maps the non-linearly separable training data to a higher dimensional space. Through this mapping, the training data become linearly separable, allowing separating hyperplanes to be found to classify the data into similar groups. SVM performs classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels. Given *l* samples (x_i , y_i), *i*=1,…,*l*, the construction of a model is reduced to the minimisation of the following regularised ε -insensitive loss function:

$$L = \|w\|^2 + \frac{C}{l}\sum_{i=1}^{l}\max\{|y_i - f(x_i) - \varepsilon|\} \quad (1)$$

where ε is the tolerable error, C is a regularisation constant and *f* is the function to be estimated:

$$f(x) = w \bullet x + b \qquad w, x \in R^n, b \in R \quad (2)$$

LibSVM is a simple, easy-to-use and efficient software algorithm for SVM classification and regression. LibSVM, as a library for support vector machine, allows the user to select different variants of SVM (C-SVM, ν-SVM, one-class SVM, ε-SVR and ν-SVR), different types of kernels (linear, polynomial, Gaussian and Sigmoid) and different parameters for each kernel [11]. One commonly used kernel function is the Gaussian Radial Basis Function (RBF) kernel $k(x_1, x_2) = \exp(-\lambda\|x_1 - x_2\|^2)$, for $\lambda > 0$ where the free parameter λ is the kernel parameter, which can be modified to reduce classification error. Through this mapping onto a higher dimensional space, hyperplanes can be constructed in a multidimensional space that linearly separate samples of different class labels. Scholkopf et al. (2000) use the parameter $\nu \in (0,1]$ to control the number of support vectors and training errors. Given training vectors $x_i \in R^n, i = 1, \dots, l$, in two classes, and a vector $y \in R^l$ such that $y_i \in \{1, -1\}$,the primal optimisation problem is formulated as follows:

$$\min_{w,b,\xi,\rho} \frac{1}{2} w^T w - \nu\rho + \frac{1}{l}\sum_{i=1} \xi_i \qquad (3)$$

Subject to

$$y_i(w^T \Phi(x_i) + b) \geq \rho - \xi_i$$
$$\xi_i \geq 0, i = 1, \ldots, l, \rho \geq 0.$$

Here $0 \leq \nu \leq 1$ and training vectors x_i are mapped into a higher- (maybe infinite) dimensional space by the function Φ.

The dual problem is

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha \qquad (4)$$

subject to

$$0 \leq \alpha_i \leq \frac{1}{l}, \quad i = 1, \ldots, l,$$
$$e^T \alpha \geq \nu, \quad y^T \alpha = 0,$$

where e is the vector of all ones, Q is a positive semi-definite matrix, $Q_{ij} = y_i y_j K(x_i, x_j)$, and $K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$ is the kernel.

The decision function is

$$sgn\left(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b\right). \qquad (5)$$

It is shown that $e^T\alpha \geq \nu$ can be replaced by $e^T\alpha = \nu$.

Further detail on the theory and algorithms of the ν-SVM classifier can be found in [14] and [15].

C. *Multi-objective optimisation*

Most real-world optimization problems involve multiple conflicting objectives, from which the user is unable to establish a relative preference. Such considerations give rise to a set of multiple optimal solutions, requiring the process of simultaneous optimization of possibly conflicting multiple objectives, and this is termed multi-objective optimization (6). The multi-objective optimization problem can be stated as:

$$\text{Minimise } F(\theta) = \{ f1(\theta) \ f_2(\theta) \ldots f_m(\theta) \} \qquad (6)$$

where $f_1(\theta), \ldots, f_m(\theta)$ are the m non-commensurable objective functions to be simultaneously minimised with respect to the parameters θ of the model [17].

Multi-objective optimisation is different from that of single-objective optimisation as the latter only searches one optimal solution but the former generates a set of solutions which are superior to the rest of the solutions with respect to all objective criteria but are inferior to other solutions in one or more objectives. These solutions are known as Pareto optimal solutions or non-dominated solutions [5]. In the absence of additional information, it is not possible to distinguish any one of the Pareto solutions as being objectively better than any others with respect to all the objectives concerned (i.e. there is no uniquely "best" solution); therefore, any one of them is an acceptable solution [18]. Decision makers can select one Pareto optimal solution from all generated Pareto fronts based on their experience and prior knowledge and other criteria or constraints. The Pareto optimal front can help the users visualize the trade-offs between different objectives and select an appropriate compromise design [18]. The domination between two solutions can be defined as follows: A solution x_1 is said to dominate another solution x_2, if both of the following conditions are true:

(1) The solution x_1 is no worse than x_2 in all objectives.

(2) The solution x_1 is strictly better than x_2 in at least one objective.

The special sorting used in the Non-dominated Sorting Genetic Algorithm (NSGA-II) is called "Pareto ranking" [5]. The NSGA-II optimization process is defined as follows: (1) Create a random parent population of size N, sort the population based on the non-domination and assign the initial non-dominated individuals as rank 1; (2) Use binary tournament selection, crossover, and mutation operators to create a new offspring population of size N; (3) Combine the offspring and parent population to form an extended population of size 2N and sort the extended population by using the crowding comparison operator; (4) Select the individuals from the sorting fronts starting from the best to create a new population of size N; (5) The non-dominated individuals identified and sorted in the new population are given the rank 2; (6) Repeat step (2) to (5) for a pre-set number of generations [19].

The step-by-step procedure shows that the algorithm automatically changes the two parameters of v-SVM to achieve classification results in each round and compares classification results, which is faster and more comprehensive than a manual operation. Step 2 and step 3 combine the NSGAII with v-SVM to optimise the classification results of v-SVM when applied to our two selected financial datasets.

III. ONLINE FINANCIAL TEXT DATASETS

Two financial textual datasets are used in this paper to evaluate the Multi-Objective Genetic Algorithm (MOGA) approach for v-SVM-based sentiment classification. A summary description of the properties of the three datasets is shown in Table 1. Each dataset is described in more detail below.

Table 1: Two financial sentiment datasets statistics

GKP	No. of Buy-labelled post	512
	No. of Sell-labelled post	512
	No. of total labelled post	1024
IFS	No. of positive labelled articles	357
	No. of negative labelled articles	643
	No. of labelled articles	1000

A. GKP stock forum dataset (GKP)

This collection of financial posts about Gulf Keystone Petroleum stock was extracted from the Interactive Investor (iii.co.uk) stock discussion board. Gulf Keystone Petroleum Ltd is an independent oil and gas exploration and production British company, and it was incorporated in 2001 in Bermuda and listed on the Alternative Investment Market (AIM) of the London Stock Exchange in 2004 (stock quote GKP). GKP is the one of the most active stocks in the discussion boards of Interactive Investor. Author-labelled posts discussing GKP from GKP RSS feeds (http://www.iii.co.uk/rss/cotn:GKP.L.xml) were saved into an XML document for a six months period, from 1st July 2012 to 31st December 2012. The same numbers of posts from two classes— BUY and SELL—were selected.

B. Irish financial sentiment dataset (IFS)

Both raw and preprocessed datasets on IFS were obtained from the University College Dublin Machine Learning Group [20]. The financial news sentiment analysis collection was retrieved from three online news sources (RTE, The Irish Times, the Irish Independent) during a three months period (July to October 2009). A subset of documents was annotated on a daily basis by a group of 33 professional volunteers who labelled the articles as positive, negative, or irrelevant. The first month constituted a “warm-up” period, which provided an initial dataset containing 3858 articles, with 2693 user annotations covering 354 individual articles. This second “main” dataset comprises 12469 documents, with 6910 user annotations resulting in 1306 labelled articles. Positive and negative labelled articles in the “main” dataset were used for the experiments in this paper.

IV. EXPERIMENT SETUP AND RESULTS

When applying the NSGA-II algorithm to optimize ν-SVM parameters, we need to first determine the objective functions, each parameter searching space and the optimization stopping criteria. The experimental process will be described step by step in this section. The results presented use two-fold cross validation on each dataset, 50% for training and 50% for testing.

A. Objective functions

The correct positive rate (CPR) and correct negative rate (CNR) are selected as two objective functions to be maximized simultaneously with the goal of optimizing the overall classification accuracy and the formulations of the two objectives are given as follows:

$$CPR=NCP/TNP *100 \qquad (7)$$

$$CNR=NCN/TNN *100 \qquad (8)$$

where NCP is the number of correct positive, TNP is the total number of positive samples for test, NCN is the number of correct negative and TNN is the total number of negative samples for test.

The overall sentiment correct classification rate (CCR) is then determined by

$$CCR= (NCP+NCN)/(TNN+TNP) *100 \qquad (9)$$

B. Experimental steps

The general flow chart for the parameter estimation process using NSGA-II is presented in Figure 1. As the standard search progresses, the entire population tends to converge to the global Pareto front. This searching process is continued until a maximum number of iterations is reached. The multi-objective optimization process generally requires the four steps below:

(1) Divide the data into training and test data sets (e.g. 50% for training and 50% for testing) or conduct ***n*** fold cross validation

(2) Change the *ν*-SVM model parameter sets and run the *ν*-SVM models for evaluations **(**correct positive rate and correct negative rate)

(3) Run the multi-objective genetic algorithm to get the optimal Pareto front after a pre-specified generation (stopping criteria is given by Clark [9]).

(4) Analyse the optimal Pareto front for decision making

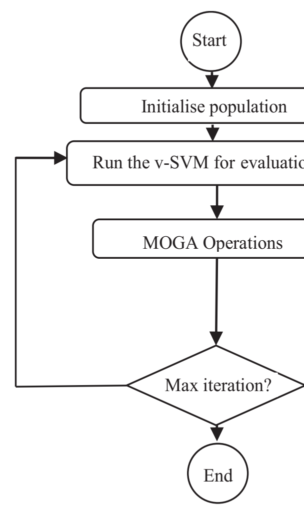

Figure 1: Outline of MOGA optimisation of v-SVM classifier

C. Experimental results

The approach is evaluated using the two online textual datasets introduced in Section III. The RBF kernel λ and the parameters of *ν*-SVM are used in the experiments. The parameter *ν* controls the lower bound of the support vectors and the upper bound of the training error. In this research we use the NSGA-II developed in Matlab [21]. The relevant recommended experimental parameters [8] for NSGA-II for sentiment classification are listed, Table 2 and Table 3 for GKP Dataset, Table 5 and Table 6 for IFS Dataset. For each dataset, there is a figure to illustrate the optimal search results: comparison of Pareto fronts using NSGA-II and a traditional random sampling method [9] with the same number of runs. The default parameters (*ν*= 0.1, λ=0.0001) of the *ν*-SVM are given and the sentiment classification using default parameters with each dataset will be provided for comparison. Traditional random sampling method is chosen for comparison as the method is simple to understand and can help understand the advantage of applying multi-objective optimisation to fine-tune the parameters of *v*-SVM.

a. GKP dataset

Table 2 and Table 3 list the NSGA-II pre-set parameters for the optimization experiment using the GKP dataset and the parameter space of *v*-SVM, respectively. .

Table 2: Experimental parameters of MOGA for GKP dataset

Parameter	Description	Recommended
P	Population size	100
G	Total generations	2000
CR	Crossover rate	0.9
MR	Mutation rate	0.1

Table 3: *v*-SVM model parameter for GKP dataset

Parameter	Description	Lower bound	Upper bound
v	Controls the number of support vectors and training error	0.05	0.5
λ	Gaussian kernel parameter	2^{-15}	2^{3}

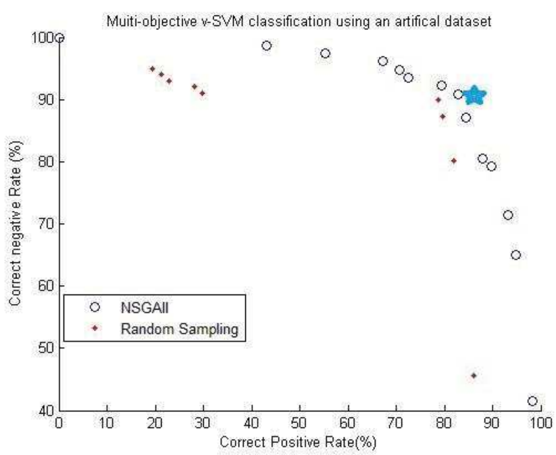

Figure 2 Comparison of Pareto fronts curves using NSGA-II and traditional random sampling method with the same number of runs (GKP dataset)

Table 4: Parameter values and their corresponding objective function values on the optimal Pareto front for GKP dataset

v	λ	CPR (%)	CNR (%)	CCR (%)
1.00e-10	0.08	0	100.00	50.00
1.11e-03	0.17	84.48	87.01	85.75
7.77e-04	**0.35**	**82.76**	**90.91**	**86.83**
7.22e-06	0.23	93.10	71.43	82.27
5.91e-05	0.24	90.87	78.24	84.56
2.24e-05	0.26	43.10	98.70	70.90
6.96e-06	0.20	89.66	79.22	84.44
4.98e-05	0.24	78.75	91.98	85.37
1.30e-05	0.22	72.54	95.31	83.93
3.31e-05	0.23	55.17	97.40	76.29
3.03e-04	0.27	84.48	87.01	85.75
2.12e-05	0.24	95.79	73.28	84.54
2.36e-05	0.21	67.24	96.10	81.67
2.35e-05	0.25	70.04	91.56	80.80
1.20e-05	0.22	43.10	98.70	70.90
1.00e-10	0.08	0	100.00	50.00
2.10e-05	0.26	98.28	41.56	69.92
7.38e-06	0.23	93.10	71.43	82.27

Figure 2 shows that the optimal Pareto front using NSGA-II performs better than traditional random sampling method. As the Pareto front generated by NSGA-II creates a higher curve than that of the random sampling method (Figure 2). Table 4 provides the values of each optimal parameter (v and λ) which corresponds to a circle on the Pareto front curve (Figure 2). Each circle on the Pareto front corresponds to two performance values of CPR (x-axis) and CNR (y-axis). Compared with the pre-processing approach using default parameters with an accuracy of 78.08% [1], the overall accuracies are improved by 7% (as shown in Table 4). The corresponding point of the best CCR, 86.83% (in Table 4) is illustrated in Figure 2 by a blue star.

b. IFS dataset

Table 5 and Table 6 list the parameter values used for *v*-SVM and the NSGA-II pre-set parameters for the optimization experiment using the IFS dataset. Figure 3 displays the optimal Pareto front for IFS dataset. There are 35 optimal circles found by v-SVM using MOGA. Compared with a traditional approach using default parameters with an accuracy of 78.42% [1], the overall accuracies are improved by 4% (as shown in Table 7), and the point corresponding to the best accuracy, 82.02%, in Table 7 is highlighted by a blue star in Figure 3. Table 7 also shows a set of optimal parameters that can maximize CPR and CNR, depending on the decision makers' requirements. For example, when v = 1.34e-03 and λ=0.29, CPR reaches its best performance of 97.33%; when v = 1.34e-03 and λ=0.25, CNR is the highest at 95.45%.

Table 5: Experimental parameters of MOGA for IFS dataset

Parameter	Description	Recommended
P	Population size	100
G	Total generations	30
CR	Crossover rate	0.9
MR	Mutation Rate	0.05

Table 6: v-SVM model parameter for IFS dataset

Parameter	Lower bound	Upper bound
V	1.0e-5	0.01
λ	0.08	0.5

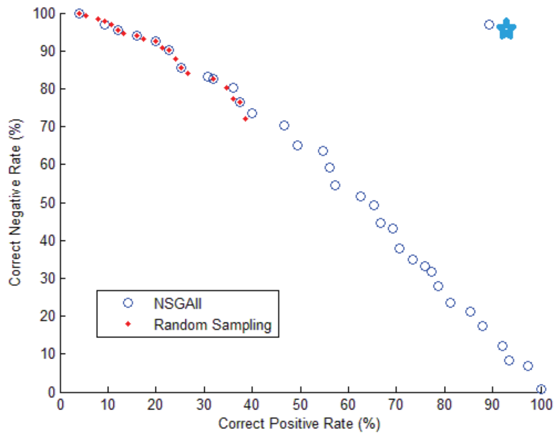

Figure 3 Comparison of Pareto fronts using NSGA-II and traditional random sampling method with the same number of runs (IFS dataset)

Table 7 Parameter values and their corresponding functions' values on optimal Pareto front for IFS dataset

ν	λ	CPR (%)	CNR (%)	CCR (%)
5.47e-3	0.19	04.00	100.00	65.26
1.00e-5	0.33	100.00	00.76	36.67
1.99e-3	0.25	12.00	95.45	65.26
6.12e-4	0.10	32.00	82.58	64.28
3.84 e-4	0.13	37.33	76.52	62.34
1.77 e-3	0.15	16.00	93.94	65.74
1.77e-03	0.13	77.33	31.82	48.29
6.37 e-4	0.13	30.67	83.33	64.28
7.08e-03	0.11	73.33	34.85	48.77
3.12e-04	0.09	69.33	43.18	52.64
4.30 e-2	0.28	36.00	80.30	64.27
1.00e-5	0.33	100.00	00.76	36.67
7.14e-04	0.10	40.00	73.48	61.37
1.63e-04	0.25	93.33	08.33	39.09
7.92 e-4	0.22	25.33	85.61	63.80
2.41e-04	0.09	85.33	21.21	44.41
6.75e-04	0.14	76.00	33.33	48.77
2.34e-03	0.17	81.33	23.48	44.42
2.05e-04	0.15	62.67	51.52	55.55
2.86e-04	0.12	56.00	59.09	57.97
2.28e-0	0.26	88.00	17.42	42.96
1.44 e-3	0.18	20.00	92.42	66.22
2.23 e-3	**0.21**	**89.33**	**96.97**	**82.02**
2.28e-07	0.13	92.00	12.12	41.02
3.97e-04	0.11	70.67	37.88	49.74
1.34e-03	0.29	97.33	06.82	39.57
1.07 e-3	0.14	22.67	90.15	65.73
6.029e-03	0.18	78.67	28.03	46.35
5.00e-04	0.12	66.67	44.70	52.65
2.69e-04	0.11	46.67	70.45	61.85
5.11 e-3	0.15	04.00	100.00	65.26
3.81e-04	0.12	54.67	63.64	60.39
5.84e-04	0.09	49.33	65.15	59.43
6.21e-04	0.15	65.33	49.24	55.06
2.653e-04	0.15	57.33	54.55	55.55

V. CONCLUSIONS AND FUTURE WORKS

This paper investigated the problem of classifying online financial sentiment information from various sources by using machine learning based techniques. The NSGA-II algorithms are applied to estimate ν-SVM classifier optimal parameters. Two objective functions, correct positive rate and correct negative rate, have been chosen for optimization. The experiments show that, by using NSGA-II, sentiment classification achieved a better classification performance by choosing optimal parameter set(s) in the obtained front compared with the baselines of a pre-processing approach based sentiment classification and traditional random sampling method. The application example demonstrated that significant trade-offs between different objectives exist, implying that a single objective function is not able to evaluate all objectives simultaneously. Instead, the parameter estimation analysis is given a set of Pareto ranks for the model parameters.

The novelty in this work is the integration of MOGA with ν-SVM to optimise sentiment classification performance, and in future MOGA techniques can be extended into other classifiers, such as Random Forest and particle swarm optimisation, to enable a full comparison of sentiment classification performance to determine most appropriate optimised classifier for sentiment classification in the financial sector. Another future direction for this research is to investigate domain-specific knowledge in datasets, as sentiment polarity analysis is a domain-dependent task; it would be beneficial to integrate semantic understanding in sentences.

ACKNOWLEDGEMENT

Fan Sun is supported by a Vice-Chancellor's Research Scholarship (VCRS) from the University of Ulster as part of the Capital Markets Engineering (CME) project. The CME project is supported by the companies and organizations involved in the N. Ireland Capital Markets Engineering Research Initiative (InvestNI, Citi; First Derivatives; Kofax; Fidessa; NYSE Technologies; ILEX; ISRC, the University of Ulster and ECIT, the Queen's University of Belfast).

REFERENCES

[1] F. Sun, A. Belatreche, S. Coleman, T. Martin NcGinnity, and Y. Li. "Pre-processing Online Financial Text for Sentiment Classification: A Natural Language Processing Approach." IEEE Computational Intelligence for Financial Engineering & Economics 2014. London, UK, 2014.

[2] B. Pang, , and L. Lee. "Opinion mining and sentiment analysis,." Foundations and Trends in Information Retrieval Vol. 2 No. 1-2 pp. 1-90, 2008.

[3] R. Xia, C. Zong, and S. Li. "Ensemble of feature sets and classification algorithms for sentiment classification." Information Sciences Vol. 181 pp. 1138–1152, 2011.

[4] F. Sun, A. Belatreche, S. Coleman, and Y. Li Martin McGinnity. "Evasluation of LibSVM and Mutual Information Matching Classifiers for Multi-Domain Sentiment Analysis." The 23rd Irish Conference on Artificial Intelligence and Cognitive Science. Dublin, 2012.

[5] Y. Liu, and F. Sun. " Parameter Estimation of a Pressure Swing Adsorption Model for Air Separation Using Multi-objective Optimisation and Support Vector Regression Model." Expert System with Applications Vol. 40, No. 11, pp. 4496-4502, 2013.

[6] C. Wu, Y. G. G. Tzeng, and W. Fang. "A Real-valued Genetic Algorithm to Optimize the Parameters of the Support Vector Machine for Predicting Bankruptcy." Expert Systems with Applications Vol. 32, No. 2, pp. 397-408, 2007.

[7] C. Chatelain, S. Adam, Y. Lecourtier, L. Heutte, and T. Paquet. "Multi-Objective Optimization for SVM Model Selection." 9th International Conference on Document Analysis and Recognition (ICDAR), Curitiba,

Parana, Brazil , 2007.
[8] J. Ethridge, G. Ditzler, and R. Polikar. "Optimal nu-SVM parameter estimation using multi objective evolutionary algorithms." IEEE Congress on Evolutionary Computation , Barcelona Spain, 2010.
[9] A. R. J., Clark, and R. M. Everson. "Multi-objective learning of Relevance Vector Machine classifiers with multi-resolution kernels." .Pattern Recognition Vol. 45, No. 9 pp. 3535-3543, 2012.
[10] M. Sokolova, and G. Lapalme. "A systematic analysis of performance measures for classification tasks." Information Processing and Management, Vol. 45, No. 4, pp. 427-437, 2009.
[11] H. Tang, S. Tan, and X. Cheng, " A survey on sentiment detection of reviews, Expert Systems with Applications", Vol. 36, pp.10760- 10773m 2009.
[12] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment strength detection in short informal text". Journal of the American Society for Information Science and Technology, Vol. 61, No. 12, pp. 2544–2558, 2010.
[13] C. Chua, M. Milosavljevic, and J. Curran, "A sentiment detection engine for internet stock message boards", Proceedings of the Australasian Language Technology Workshop, Sydney, NSW, Australia 2009.
[14] S. Das, and M. Chen, "Yahoo! For Amazon: Sentiment extraction from small talk on the web, Management Science, Vol. 53, No. 9, pp.1375-1388, 2007.
[15] B. Schölkopf, New Support Vector Algorithms. Neural Computation.Vol. 5, No. 15, pp. 1207-1245, 2000.
[16] V. Vapnik The Nature of Statistical Learning Theory. New York : Springer-Verlag, 1995.
[17] P. O. Yapo, H. V. Gupta, and S. Sorooshian. " Multi-objective Global Optimisation for Hydrologic Models ." Journal of Hydrology, Vol. 204, No. 1-4, pp. 83-97, 1998.
[18] T. Goel, R. Vaidyanathan, R. Haftka, W. Shyy, N. Queipo, and K.Tucker. "Response Surface Approximation of Pareto Optimal Front in Multi-objective Optimisation." Computer Methods in Applied Mechanics and Engineering Vol. 196, pp. 879–893, 2007.
[19] K. Deb, A. Pratap, S. Agrawal, and T. Meyarivan. "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimisation: NSGA-II." IEEE Transactions on Evolutionary Computation, Vol. 6, No. 2, pp.182-197, 2002.
[20] A. Brew, D. Greene, and P. Cunningham. "Using crowdsourcing and active learning to track sentiment in online media." Proceedings of the 19th European Conference on Artificial Intelligence. Amsterdam, the Netherlands, 2010.
[21] A. Seshadri, NSGA - II: A multi-objective optimization algorithm. MathWorks. [Online] MATLAB, 19/03/2006. [Cited: 14/08/2014.] http://www.mathworks.com/matlabcentral/fileexchange/10429-nsga-ii--a-multi-objective-optimization-algorithm