修士論文

Automatic Detection of Repressed Anger from Text Messages

20○○年○月○○日

(Replace with publication date)

学籍番号　16906291

ヘスス·マリア·セスマ·ソランセ

指導教員　湯川　高志

長岡技術科学大学大学院工学研究科

経営情報システム工学専攻

平成　　年　　月　　日

専攻名	経営情報システム工学	学籍番号	16906291
申 請 者 氏 名	ヘスス· マリア· セスマ· ソランセ		
指 導 教 員 氏 名	湯川高志　教授		

審査委員主査　　〇〇　〇〇　教授

審査委員　　　　〇〇　〇〇　准教授

審査委員　　　　〇〇　〇〇　准教授

審査委員

審査委員

専攻主任印

論　文　要　旨

論文題目	Automatic Detection of Repressed Anger from Text Messages

〇〇。

修士論文

Automatic Detection of Repressed Anger from Text Messages

20〇〇年〇月〇〇日

(Replace with publication date)

学籍番号　16906291

ヘスス·マリア·セスマ·ソランセ

指導教員　湯川　高志

長岡技術科学大学大学院工学研究科

経営情報システム工学専攻

Contents

List of Figures

List of Tables

Chapter 1

Algorithms

1.1 Fundamentals of Classification

Classification, which is the task of assigning objects to one of several predefined categories, is a pervasive problem that encompasses many diverse applications [5].

1.2 Support Vector Machine

There are four main advantages: Firstly it has a regularisation parameter, which makes the user think about avoiding over-fitting. Secondly it uses the kernel trick, so you can build in expert knowledge about the problem via engineering the kernel. Thirdly an SVM is defined by a convex optimisation problem (no local minima) for which there are efficient methods (e.g. SMO). Lastly, it is an approximation to a bound on the test error rate, and there is a substantial body of theory behind it which suggests it should be a good idea. The disadvantages are that the theory only really covers the determination of the parameters for a given value of the regularisation and kernel parameters and choice of kernel. In a way the SVM moves the problem of over-fitting from optimising the parameters to model selection. Sadly kernel models can be quite sensitive to over-fitting the model selection criterion [2]

SVMs are a new promising non-linear, non-parametric classification tech- nique, which already showed good results in the medical diagnostics, optical character recognition, elec- tric load forecasting and other fields.

Suitable for binary classification tasks.

The advantages of the SVM technique can be summarised as follows [1]:

1. By introducing the kernel, SVMs gain flexibility in the choice of the form of the threshold separating solvent from insolvent companies, which needs not be linear and even needs not have the same func- tional form for all data, since its function is non-parametric and operates locally. As a consequence they can work with financial ratios, which show a non-monotone relation to the score and to the probability of default, or which are non-linearly dependent, and this without needing any specific work on each non-monotone variable.

2. Since the kernel implicitly contains a non-linear transformation, no assumptions about the functional form of the transformation, which makes data linearly separable, is necessary. The transformation oc- curs implicitly on a robust theoretical basis and human expertise judgement beforehand is not needed.

3. SVMs provide a good out-of-sample generalization, if the parameters C and r (in the case of a Gaussian kernel) are appropriately chosen. This means that, by choosing an appropriate generalization grade, SVMs can be robust, even when the training sample has some bias

4. SVMs deliver a unique solution, since the optimality problem is convex. This is an advantage compared to Neural Networks, which have multiple solutions associated with local minima and for this reason may not be robust over different samples.

5. With the choice of an appropriate kernel, such as the Gaussian kernel, one can put more stress on the similarity between companies, because the more similar the financial structure of two companies is, the higher is the value of the kernel. Thus when classifying a new company, the values of its financial ratios are compared with the ones of the support vectors of the training sample which are more similar to this new company. This company is then classified according to with which group it has the greatest similarity.

Furthermore, $K(\boldsymbol{x}i, \boldsymbol{x}j) \equiv \phi(\boldsymbol{x}_i)^T \phi(\boldsymbol{x}_j)$ is called the kernel function.
four basic kernels [4]:

- **Linear:** $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^T \boldsymbol{x}_j$.
- **Polynomial:** $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\gamma \boldsymbol{x}_i^T \boldsymbol{x}_j + r)^d, \gamma > 0$.
- **Radial Basis Function (RBF):** $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp(-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2), \gamma > 0$.
- **Sigmoid:** $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh(\gamma \boldsymbol{x}_i^T \boldsymbol{x}_j + r)$.

Here, γ, r, and d are kernel parameters.

Support vector machines (SVM) were originally designed for binary classification [3].

solving multi-class SVM in one step: "all-together"methods: [25], [27] and [7]. We then compare their performance with three methods based on binary classifications: "one-against-all,""one-against-one,"and DAGSVM [23]. Our experiments indicate that the "one-against-one"and DAG methods are more suitable for practical use than the other methods.

1.3 K-Nearest Neighbor

1.4 Neural Networks

References

[1]Laura Auria and Rouslan A Moro. “Support vector machines (SVM) as a technique for solvency analysis”. In: (2008).

[2]Gavin C Cawley and Nicola LC Talbot. “On over-fitting in model selection and subsequent selection bias in performance evaluation”. In: *Journal of Machine Learning Research* 11.Jul (2010), pp. 2079–2107.

[3]Chih-Wei Hsu and Chih-Jen Lin. “A comparison of methods for multiclass support vector machines”. In: *IEEE transactions on Neural Networks* 13.2 (2002), pp. 415–425.

[4]Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. “A practical guide to support vector classification”. In: (2003).

[5]Tan Pang-Ning, Michael Steinbach, Vipin Kumar, et al. “Introduction to data mining”. In: *Library of congress*. Vol. 74. 2006.