MultiCraft

International Journal of Engineering, Science and Technology
Vol. 3, No. 6, 2011, pp. 71-77

INTERNATIONAL JOURNAL OF ENGINEERING, SCIENCE AND TECHNOLOGY
www.ijest-ng.com

Infinite ensemble of support vector machines for prediction of musculoskeletal disorders risk

Pankaj Chandna[1]*, Mahesh Pal[2]

[1] Department of Mechanical Engineering, National Institute of Technology Kurukshetra, Haryana, INDIA*
[2] Department of Civil Engineering, National Institute of Technology Kurukshetra, Haryana, INDIA
[*] Corresponding Author: e-mail: pchandna08@gmail.com

Abstract

Several modeling techniques have been used to classify the low-back disorders (LBDs) risk associated with the industrial jobs. Many researchers have demonstrated the use of artificial neural networks (ANNs) to predict musculoskeletal disorders risk associated with occupational exposures. In order to improve the accuracy of LBDs risk classification, this paper proposes to use the support vector machines (SVMs), a machine learning algorithm used extensively in the last decade. The results of SVMs based ensemble classification approach to classify the low-back disorders (LBDs) risk associated with the industrial jobs are presented. Four different kernels (i.e. the stump kernel, perceptron kernel, Laplacian kernel and exponential kernel) were used to create infinite ensemble of SVMs and their performance have been compared with the SVMs, AdaBoost SVMs, and a backpropagation neural network. The results suggest an increased performance by stump and Laplacian kernel in comparison to the radial basis function and polynomial kernel based SVMs, AdaBoost SVMs, and the back propagation neural network. Highest classification accuracy of 77.01% was achieved by Laplacian kernel based SVMs in comparison to 71.3% and 74.7% by radial basis function kernel based SVMs and back propagation neural network respectively.

Keywords: Low-back disorders (LBDs), Support Vector Machines (SVMs), Ensemble learning, Back propagation neural network

DOI: http://dx.doi.org/10.4314/ijest.v3i6.6

1. Introduction

The term musculoskeletal disorders (MSDs) refer to the conditions that involve the nerves, tendons, muscles, and supporting structures of the body. Work related Musculoskeletal disorders (WMSDs) refer to musculoskeletal disorders to which the work environment and the performance of work contribute significantly or are being made worse or long lasting by the working conditions. WMSDs are amongst the most prevalent lost-time injuries and illnesses in almost every industry (BLS, 1995; BLS, 1996; Itasca, 1995; Tanaka et al., 1995) and those involving the back, are amongst the most costly occupational problems (Frymoyer & Cats, 1991; Webster & Snook, 1994; Guo et al., 1995). WMSDs may cause a great deal of pain and suffering among afflicted workers and may decrease productivity and the quality of products and services. Workers experiencing aches and pains on the job may not be able to do quality work.

Ortiz-Hernandez et al. (2003) analyzed the relationship between musculoskeletal disorders and use of the personal computer (PC) in Mexico City and showed that continuous use the PC increased risk of developing musculoskeletal disorders. Such an increase is mediated by ergonomic factors e.g. mouse use, remaining seated for prolonged periods, adoption of inadequate or uncomfortable postures, performing certain PC tasks, and psychosocial factors etc. Bernard (1997) examined the epidemiologic evidence of the relationship between selected MSDs of the upper extremity & the low back and exposure to physical factors at work. Specific attention is being given to analyzing the weight of the evidence for the strength of the association between these disorders and the work factors. Because the relationship between exposure to physical work factors and the development and prognosis of particular disorders may be modified by psychosocial factors, the literature about psychosocial factors and the presence of musculoskeletal symptoms or disorders is also reviewed. Understanding these associations and relating them to the cause of disease is critical for identifying exposures amenable to preventive and therapeutic interventions. Also a substantial body of credible epidemiologic research provides strong evidence of an association between MSDs and certain work-related physical factors when there are high levels of exposure and especially in combination with exposure to more than one physical factor e.g. repetitive lifting of heavy objects in extreme or awkward postures.

On the basis of the survey conducted over the past 25 years and with a random sample of about 2,50,000 private sector establishments, Bureau of Labour Statistics (1994) of the U.S. Department of Labour reported that approximately 705,800 cases (32%) were the result of overexertion or repetitive motion, 367,424 injuries due to overexertion in lifting (65% affected the back); 93,325 injuries due to overexertion in pushing or pulling objects (52% affected the back); 68,992 injuries due to overexertion in holding, carrying, or turning objects (58% affected the back); total across these three categories, 47,861 disorders affected the shoulder. 83,483 injuries or illnesses were caused by unspecified overexertion events, 92,576 injuries or illnesses due to repetitive motion, including typing or key entry, repetitive use of tools, and repetitive placing, grasping, or moving of objects other than tools. Of these injuries or illnesses, 55% affected the wrist, 7% affected the shoulder, and 6% affected the back. It has been reported by National Research Council (USA) that the estimated costs associated with the lost days and compensation claims related to musculoskeletal disorders-including back pains and repetitive motion injuries-range from $13 billion to $20 billion annually. This is a serious problem that has spurred considerable debate about the causal links between such disorders and risk factors at the workplace. Musculoskeletal disorders (MSDs) are a major occupational health problem in Europe also, affecting over 40 million workers. Even current EU legislation has included some ergonomic provisions relating to MSDs prevention.

The modeling techniques may be used for the development of models that explicitly describe the risk associated with various work designs so that specific, quantitative workplace assessments can be made (Dempsey, 1997). For example, although there are established risk factors for LBDs, the manner in which these factors interact to promote the risk of LBDs in industry, and ultimately disability, is not well understood (Zurada et al., 1997). Many researchers have demonstrated the use of, for example, advanced statistical methods including logistic regression and generalized additive models (Dempsey, 1995) and (Marras et al., 1993) and artificial neural networks (ANNs) (Zurada et al., 1997; Karwowski, 1994; Killough, 1995 and Chen et al. 2000) and semi-supervised learning approach (Chandna et al., 2010) to predict musculoskeletal disorders risk associated with occupational exposures. It is found that neural networks provide superior predictive capability in comparison to multiple linear regression models (Zurada et al., 1997 and Killough, 1995). Some of the problems with a back-propagation neural network based modeling algorithm includes setting up of different learning parameters (like learning rate, momentum), the optimal number of nodes in the hidden layer and the number of hidden layers so as to have a less complex network with a relatively better generalization capability. A large number of training iterations may force ANN to over train, which may affect the predicting capabilities of the model. Keeping these issues in view, the main objective of this study is to develop a support vector machines based classification system that could classify industrial jobs according to the potential risk for low-back disorders (LBDs). Support vector machines based classification algorithms are found to work well in comparison to back-propagation neural network (Pal and Mather, 2006), thus suggesting that such an algorithm could be very useful in modeling potential risk for low-back disorders in an industrial environment.

In recent years, a lot of work proposing the combination of multiple classifiers, referred to as an ensemble classifier, have been reported in literature and found to be more accurate than any of the individual classifiers making up the ensemble. An ensemble is defined as a set of individually trained classifiers whose predictions are combined when classifying a new data. In the last decade a number of ways to create ensemble of classifiers are suggested in the literature (Bauer and Kohavi, 1999 and Dietterich, 2000).

Recently, some studies suggest that it is beneficial to use infinite ensemble learning (Demiriz et al., 2002) in comparison to other ensemble approaches. The learning algorithms such as SVMs are found to be working well by combining infinite hypotheses (Vapnik, 1998). An infinite ensemble approach with SVMs can be obtained by embedding the base learners (hypothesis) into a SVMs kernel, as SVMs is capable of handling infinite number of features through the use of kernels. This approach can embed an infinite number of base learning models altogether, which is not possible by traditional ensemble learning algorithms. This allows constructing new kernels; Lin (2005) derived four new kernels, the stump kernel, the perceptron kernel, Laplacian kernel and exponential kernel to create an infinite ensemble of SVMs.

Keeping in view the potential of ensemble approaches for classification (Pal, 2007; 2008), this paper discusses the results of an infinite ensemble approach as proposed by Lin (2005) with SVMs as base classifier to classify industrial jobs according to the potential risk for low-back disorders (LBDs). Results obtained with this ensemble approach are compared with AdaBoost SVMs, SVMs and back-propagation neural network in terms of overall classification accuracy.

2. Support Vector Machines

The support vector machine classifier is based on statistical learning theory (Vapnik, 1995) and tries to find an optimal hyperplane as a decision function in high dimensional space (Boser et al., 1992) and (Cristianini & Shawe-Taylor, 2000). The SVMs use structural risk minimisation, rather than using empirical risk minimisation. Empirical risk minimises the misclassification error on the training set, whereas structural risk minimises the probability of misclassifying a previously unseen data point drawn randomly from a fixed but unknown probability distribution (Lin, 2005). Linearly separable classes are the simplest case on which to train a support vector machine. Let the training data with k number of samples be represented by $\{\mathbf{x_i}, y_i\}$, i = 1… k, where $x \in \mathbf{R^N}$ is an N-dimensional space and $y \in \{-1, +1\}$ is the class label. These training patterns are said to be linearly separable if there exists a vector **w** (determining the orientation of a discriminating plane) and a scalar b (determine the offset of the discriminating plane from the origin) such that

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \tag{1}$$

The hypothesis space can be defined by the set of functions given by

$$f_{\mathbf{w},b} = sign(\mathbf{w} \cdot \mathbf{x} + b) \tag{2}$$

If the set of examples is linearly separable, the goal of the SVMs is to minimise the value $\|\mathbf{w}\|^2$, which is equivalent to find the separating hyperplanes for which the distance between the classes of training data, measured along a line perpendicular to the hyperplane, is maximised. This distance is called the margin (Vapnik, 1995) .The problem of minimising $\|\mathbf{w}\|^2$ can be solved by using standard Quadratic Programming (QP) optimisation techniques and Lagrangian multipliers by forming a Lagrangian. For a two-class problem the decision rule that separates the two classes can be written as:

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{k} \lambda_i y_i (\mathbf{x} \cdot \mathbf{x}_i) + b\right) \tag{3}$$

Where λ_i, $i = 1,\ldots,k$ are positive Lagrange multipliers. In the case of linearly non-separable data Cortes and [24] suggested that the restriction that every training vector of a given class lie on the same side of the optimal hyperplane be relaxed by introducing a positive "slack variable" ξ_i, that takes the value $\xi_i \geq 0$. In this case, SVMs algorithm searches for the hyperplane that maximises the margin and that, at the same time, minimizes a quantity proportional to the number of misclassification errors. This trade-off between margin and misclassification error is controlled by introducing a positive constant C such that $\infty > C > 0$.

For non-linear decision surfaces (Boser et al., 1992) proposed that a feature vector is mapped into a higher dimensional Euclidean space (feature space), via a non-linear vector function. The optimal margin problem in the Euclidean space can be written by replacing $\mathbf{x}_i \cdot \mathbf{x}_j$ with $\mathbf{\Phi}(\mathbf{x}_i) \cdot \mathbf{\Phi}(\mathbf{x}_j)$ and the equation (3) will have the form:

$$f(\mathbf{x}) = sign\left(\sum_i \lambda_i y_i \mathbf{\Phi}(\mathbf{x}) \cdot \mathbf{\Phi}(\mathbf{x_i}) + b\right) \tag{4}$$

As suggested by (7) the only quantities that one needs to compute are the scalar products, of the form $\mathbf{\Phi}(\mathbf{x})$. $\mathbf{\Phi}(\mathbf{y})$ which is computationally quite expensive. To cope up with problem, Vapnik (1995) introduced the concept of the *kernel function* K in the design of non-linear SVMs. A kernel function is defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{\Phi}(\mathbf{x}_i) \cdot \mathbf{\Phi}(\mathbf{x}_j) \tag{5}$$

Thus, based on equation (5), equation (4) will have the form:

$$f(\mathbf{x}) = sign\left(\sum_i \lambda_i y_i K(\mathbf{x_i}, \mathbf{x_j}) + b\right) \tag{6}$$

In this optimization problem the kernel function is computed rather than $\mathbf{\Phi}(\mathbf{x})$. This approach consists in moving the training dataset into a higher-dimensional feature space where the training data may be spread further apart and a larger margin may be found for the optimal hyperplane. A number of kernel functions can be used with SVMs. Vapnik (1995) discussed kernels and the parameters of the kernel function in detail. SVMs were initially designed for binary (two-class) problems. When dealing with multiple classes, a "one against one" approach (Knerr et al, 1990) is used in this study.

3. Ensemble of SVMs

Following section discusses the effect of infinite ensemble approach proposed by Cortes and Vapnik (1995) with support vector machines as a base classifier to classify low-back musculoskeletal disorders. it also discusses the infinite approach to create ensemble of SVMs in details.

3.1 Infinite Ensemble of SVMs

In place of using a powerful learning model, ensemble learning algorithms such as AdaBoost (Freund and Schapire,1997) deals with a simple base learning model H whereas the classifier $h \in H$ is called a base hypothesis. Ensemble learning algorithms construct a classifier that averages over some base hypotheses in a set of hypotheses H. AdaBoost iteratively selects K hypotheses $h_i > H$ to construct an ensemble classifier

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{M} w_i\ h_i\ (x)\right) \quad \text{for} \quad w_i > 0 \ \text{and} \ i = 1, 2, \ldots, M. \tag{7}$$

A classifier *f*, expressed by equation (7) is called an ensemble classifier, and w_i is called the hypothesis weights. Another property of using finite ensemble learning algorithms is that they select each hypothesis h_i by calling a base learning algorithm. Ratsch et al., (2001) suggested that under $M \to \infty$ AdaBoost asymptotically approximates an infinite ensemble classifier $sign\left(\sum_{i=1}^{\infty} w_i\ h_i\ (x)\right)$ such that (*w*, *h*) is an optimal solution to

$$\min_{w_i \in R,\ h_i \in H} \|w\|_1$$

$$\text{subject to} \quad y_j\ \left(\sum_{i=1}^{M} w_i\ h_i\ (x)\right) \geq 1, \ \text{for} \ w_i > 0 \tag{8}$$

Lin (2005) suggested that SVMs and AdaBoost can easily be related as same role is played by the elements of $\Phi(x)$ and hypotheses $h_i(x)$ in SVMs and AdaBoost respectively and both works on linear combination of these elements. SVMs minimizes the l_2-norm of weights as well as having an additional term *b*, where as AdaBoost minimizes the l_1-norm of weights and requires $w_i > 0$ for ensemble learning.

This connection between SVMs and ensemble learning is utilized by Lin (2005) to formulate kernels those embodies infinite number of hypotheses in H and equation (6) with these kernels is a linear combination of infinite number of hypotheses (with an intercept term). Lin (2005) defined a kernel based on feature mapping as:

Assume that $H = \{h_\alpha : \alpha \in C'\}$, where C' is a measure space. The kernel that embodies H is defined as:

$$K_{H,r}(x, x') = \int_{C'} \Phi_x(\alpha)\, \Phi_{x'}(\alpha)\, d\alpha \tag{9}$$

Where $\Phi_x(\alpha) = r(\alpha) h_\alpha(x)$, and $r : C' \to R^+$ is chosen such that the integral exists for all $x, x' \in \chi$ and α is the parameter of the hypothesis h_α.

The algorithm proposed by Lin (2005) for infinite ensemble of SVMs has the following steps:

1. Start with a hypothesis space *H* and training data represented by $\{(x_i, y_i)\}_{i=1}^{k}$. The hypothesis set is negation complete (i.e. $h \in H \ iff \ (-h) \in H$) (Freund and Schapire, 1997) and contains a constant hypothesis (Vapnik, 1998).
2. A kernel (K_H) is constructed using equation (9) and a proper value of *r*.
3. Choose parameters such as regularization parameter C of SVMs classifier.
4. Obtain the value of Lagrange multipliers (λ_i) and the value of intercept *b* by solving the optimization problem with K_H.
5. Output a classifier $sign\left(\sum_{i=1}^{k} y_i\ \lambda_i\ K_H(\mathbf{x_i}, \mathbf{x}_j) + b\right)$, which is equivalent to some ensemble classifier over *H*.

Four different kernels used to create the infinite ensemble of SVMs are:

- The stump kernel ($\left(\|x - x'\|_1\right) + coefficient$, which contains an infinite number of decision stumps.
- Stump kernel can be extended to the perceptron kernel $\left(\|x - x'\|_2\right) + coefficient$, which consists of infinite number of perceptrons.
- Laplacian kernel $\left(\exp\left(-\gamma \|x - x'\|_1\right)\right)$ is derived from the infinite tree kernel. SVMs with the Laplacian kernel create an infinite ensemble classifier over decision trees of any level.

- Similarly, exponential kernel $\left(\exp\left(-\gamma \left\| x - x' \right\|_2\right)\right)$ also embeds infinite number of decision regions with perceptron boundaries.

Both Laplacian and exponential kernel are RBF kernels with γ and *coefficient* is user-define parameter. For further details about these kernels and their derivation readers are referred to Lin (2005).

4. Dataset and methodology

The dataset collected by Marras et al. (1993) in a field study of 235 industrial jobs with low and high-risk values of low-back disorders was used in this study. The dataset consists of independent variables namely lift rate in number of lifts per hour (LIFTR), peak twist velocity average (PTVAVG), peak moment (PMOMENT), peak sagittal angle (PSUB), and peak lateral velocity maximum (PLVMAX) and a single dependent variable (i.e. low-back disorders) having two discrete levels of 'low risk' and 'high risk'. Zurada et al. (1997) used a backpropagation neural network (NN) to classify industrial jobs according to the potential for low-back disorders using this dataset. A detailed table of both training and test data is provided by Zurada et al. (1997). They used a total of 148 randomly selected data to train the neural network. Training dataset consists of 74 low-risk and 74 high-risk jobs. The remaining 87 jobs (50 low risk and 37 high-risk) were used to test the neural network. In order to compare the performance of the proposed approach with that of Zurada et al. (1997), same dataset for training and testing was used in this study. In the present study class 1 refers low-risk jobs whereas class 2 represents high-risk jobs.

The concept of the kernel was introduced to extend the capability of the SVMs to deal with non-linear decision surfaces. There is a little guidance in the literature on the criteria to be used in selecting a kernel, the kernel-specific parameters and the value of regularization parameter (*C*). A grid search method together with k-fold cross-validation error and trial and error is one of the most frequently used methods for parameter selection with SVMs with averaged size of datasets (Ismael et al. 2008). Grid search method was found to be computationally very expensive because the model must be evaluated at many points within the grid for each parameter and cross-validation method does not consider test dataset during evaluation of the model. Keeping in view the problems with the grid search method together with k-fold cross-validation error, trial and error method was used in this study. A number of trials were carried out using polynomial $\left((x.y+1)^{degree}\right)$ and radial basis kernel function $\left(e\left(-\gamma \left\| x_i - x_j \right\|^2\right)\right)$ as well as with different kernels proposed by Lin (2005) with different values of kernel-specific parameter and *C*, considering classification accuracy on test dataset as the measure of quality. The value of γ was varied between 0.1 and 1 whereas value of C was varied between 1 and 10. In case of polynomial kernel, degree of polynomial value was varied between 1 and 3. The value of kernel width parameter γ = 0.167 (with radial basis function kernel), *degree* of polynomial=1(with polynomial kernel) and regularization parameter *C* = 2 was found to be performing well with different kernels. In case of AdaBoost, ten iterations were found to provide highest accuracy with the dataset used in this study.

Classification accuracy was used to compare the performance of different classification algorithms used in this study. In addition, confusion matrices are also provided for different ensemble approaches used in this study.

5. Results

The aim of the present study is to evaluate the usefulness of infinite ensemble of SVMs as base classifier for LBDs classification and comparing its performance with SVMs. To test the performance, four different kernels are considered to create infinite ensemble as mentioned in section 3. A test dataset consisting of 87 jobs, consisting of 50 low risk and 37 high-risk jobs, has been used to judge the performance of the trained classifiers. Table 1 provides the classification accuracy derived by using SVMs with different kernels. Results suggest a comparable performance by polynomial and radial basis function kernels with this dataset. In comparison, result suggests that Laplacian kernel based ensemble approach classified a total of 67 dataset correctly (77.01%). Out of 50 low-risk and 37 high-risk jobs, a total of 42 low-risk and 25 high-risk jobs were correctly classified by Laplacian kernel based SVMs. Similarly, Stump kernel based ensemble of SVMs provides an improved classification performance in comparison to SVMs (Table 1). Results from table 2 suggest a poor performance by Perceptron and exponential kernel based ensemble approach in comparison to the SVMs. A comparison of results from Table 1 suggests that Laplacian and stump kernel based ensemble of SVMs provides an increase of about 5% in classification accuracy in comparison to Polynomial/ RBF kernel based SVMs. In comparison to infinite ensemble of SVMs, AdaBoost based ensemble approach achieve an accuracy of 72.4% suggesting inferior performance by AdaBoost SVMs with this dataset.

Table 1: Classification accuracies with different kernel using SVMs.

	Kernel used					
	Polynomial	RBF	Stump	Perceptron	Laplacian	Exponential
Overall classification (%)	71.30	71.30	75.90	66.70	77.01	63.20

Results from this study also suggest an improved performance by the infinite ensemble based SVMs approach to classify the industrial jobs (with both low and high risk jobs) in comparison to the neural network approach proposed by Zurada et al. (1997). In comparison of classification accuracy of 74.7% achieved by neural network approach, stump and Laplacian kernel based ensemble of SVMs achieves an accuracy of 75.9% and 77.01% respectively. A comparison of table 2 suggests that out of 50 low risk jobs Laplacian kernel based infinite ensemble of SVMs classify 42 correctly and out of 37 high risk jobs this kernel classify 25 correctly.

Table 2: Confusion matrices with different ensemble approaches with SVMs

	Stump		Laplacian		Perceptron		Exponential		AdaBoost	
	Class1	Class2	Class1	Class2	Class1	Class2	Class1	Class2	Class1	Class2
Class 1	41	12	42	12	40	19	36	18	42	16
Class 2	9	25	8	25	10	18	14	19	8	21

6. Conclusions

This study discusses SVMs based infinite ensemble approach to classify low-back disorders (LBDs) risk associated with the industrial jobs. Four different kernels are used to create infinite ensemble of SVMs and their performance has been compared with AdaBoost SVMs, SVMs and a backpropagation neural network. Results suggests that the infinite ensemble approach works well in classifying the low-back disorder (LBDs) risk associated with the industrial jobs in comparison to the neural network approach proposed by Zurada et al. (1997) and AdaBoost SVMs in term of overall classification accuracy. A better performance by Laplacian kernel based infinite ensemble of SVMs in comparison to other kernels is also concluded from the results. Finally, performance degradation of ensemble of SVMs in classifying the high risk jobs may be due to possible misclassification of data in high and low risk jobs (Zurada et al., 1997). Despite the encouraging performance of the proposed approach with used dataset, a major problem with artificial intelligence-based modeling approaches is their data-dependent nature. Their outputs may change depending on the dataset, the scale at which the experiments are conducted or the number of data available for training.

Acknowledgements

The authors would like to thank all the anonymous reviewers for their insightful comments, which helped us to improve the quality of this manuscript.

References

Bauer E. and Kohavi R., 1999. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants, *Machine Learning*, Vol. 36, pp.105–142.

BLS, 1995. Workplace injuries and illnesses by selected characteristics, 1993 Washington, DC: US Department of Labor, Bureau of Labor Statistics, USDL 95–142.

BLS, 1996. Characteristics of injuries and illnesses resulting in absences from work. 1994 Washington, DC: US Department of Labor, Bureau of Labor Statistics, USDL 96–163.

Bernard B., 1997. Musculoskeletal disorders and workplace factors: A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back Second Printing Cincinnati, OH: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention" *National Institute for Occupational Safety and Health DHHS (NIOSH)*, Publication Number 97–141.

Boser B., Guyon I. and Vapnik V.N., 1992. A training algorithm for optimal margin classifiers, *Proceedings of 5th Annual Workshop on Computer Learning Theory, Pittsburgh; PA: ACM*, pp.144–152.

Chen, C-L, Kaber, D. B. and Dempsey, P. G., 2000, A new approach to applying feedforward neural networks to the prediction of musculoskeletal disorder risk, Applied Ergonomics 31, 269-282.

Cortes C. and Vapnik V.N., 1995. Support vector networks, *Machine Learning*, Vol.20, pp.273–297.

Cristianini N. and Shawe-Taylor J., 2000. *An Introduction to Support Vector Machines*, Cambridge University Press, London.

Chandna, P., Deswal, S. and Pal, M., 2010. Semi-Supervised Learning Based Prediction of Musculoskeletal Disorder Risk, *Journal of Industrial and Systems Engineering*, Vol. 3, No. 4, pp. 291-295.

Dempsey, P.G. and Ayoub, M.M., 1995. Westfall PH. The NIOSH lifting equations: a closer look, *Advances in Industrial Ergonomics and Safety VII, Taylor and Francis, Bristol, PA*, pp. 705–712.

Dempsey P. G. and Westfall P.H., 1997. Developing explicit risk models for predicting low - back disability: a statistical perspective, *Indian Journal of Industrial Ergonomics*, Vol.19, pp. 483–497.

Demiriz A., Bennett K.P. and Shawe-Taylor J., 2002. Linear programming boosting via column generation, *Machine Learning*, Vol. 46, pp. 225–254.

Dietterich T.G., 2000. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization, *Machine Learning*, Vol.40, pp.139–157.

Freund Y. and Schapire R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences*, Vol. 55, pp.119–139.

Frymoyer J.W. and Cats-Baril W.L., 1991. An overview of the incidence and costs of low-back pain, *Orthopedic Clinics of North America*, Vol. 22, pp. 262–271.

Guo H., Tanaka S., Cameron L., Seligman P., Behrens V. and Ger J., 1995. Back pain among workers in the United States: national estimates and workers at high risks, *American Journal of Industrial Medicine*, Vol. 28, No.5, pp.591–602.

Itasca I.L., 1995. *Accident facts*, National Safety Council (NSC).

Ismael, K., Salleh, S.H., Najeb J.M. and Jahangir Bakhteri, R.B., 2008. Efficient Parameter Selection of Support Vector Machines, Biomed 2008, N.A. Abu Osman, F. Ibrahim, W.A.B. Wan Abas, H.S. Abd Rahman, H.N. Ting (Eds.), Proceedings 21, pp. 183–186.

Karwowski W., Zurada J., Marras W.S. and Gaddie P., 1994. A prototype of the artificial neural network-based system for classification of industrial jobs with respect to risk of low back disorders, *Advances in Industrial Ergonomics and Safety,* Vol.6, pp.19–22.

Killough M.K., Crumpton L., Calvert A. and Bowden R., 1995. An investigation of using neural networks to identify the presence of carpal tunnel syndrome, *Proceedings of 4th Industrial Engineering Research Conference*, pp. 659–667.

Knerr S., Personnaz L. and Dreyfus G., 1990. Single-layer learning revisited: a stepwise procedure for building and training neural network, *Neurocomputing: Algorithms, Architectures and Applications*, NATO ASI. Berlin: Springer.

Marras W.S., Lavender S.A., Leurgans S., Sudhakar L.R., Allread W.G., Fathallah F. and Ferguson S., 1993. The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders. The effects of workplace factors, trunk position, and trunk motion characteristics on risk of injury, *Spine*, Vol. 18, pp. 617–628.

Ortiz-Hernandez L., Tamez-Gonzalez S., Matýnez-Alcantara S. and Mendez-Ramýrez I., 2003. Computer use increases the risk of musculoskeletal disorders among newspaper office workers, *Archives of Medical Research*, Vol. 34, pp.331–342.

Pal, M., 2008, Ensemble of support vector machines for land cover classification. International Journal of Remote Sensing, 29(10), 3043 -3049.

Pal, M., 2007, Ensemble Learning with Decision Tree for Remote Sensing Classification. Proceedings of World Academy of Science, Engineering and Technology, December 14-16, Bangkok, Thailand, 26, 735-737.

Pal, M. and Mather, P. M., 2006, Some Issues in the Classification of Remote Sensing Data: A Case Study with DIAS Hyperspectral Data. International Journal of Remote Sensing, Vol. 27, No. 14, 2895–2916.

Ratsch G., Onoda T. and Muller K., 2001. Soft margins for AdaBoost, *Machine Learning*, Vol. 42, pp. 287–320.

Tanaka S., Wild D., Seligman P., Halperin W., Behrens V. and Putz-Anderson V., 1995. Prevalence and work-relatedness of self-reported carpal tunnel syndrome among US workers: analysis of the occupational health supplement data to the 1988 National Health Interview Survey, *American Journal of Industrial Medicine,* Vol. 27, pp.451–470.

Vapnik, V.N., 1998. Statistical Learning Theory, *Wiley:* New York.

Lin H.T., 2005. Infinite ensemble learning with support vector machines, *Master's thesis,* California Institute of Technology.

Vapnik V.N., 1995. The Nature of Statistical Learning Theory, *Springer-Verlag*: New York.

Webster B.S. and Snook S.H., 1994. The cost of compensable upper extremity cumulative trauma disorders, *Journal of Occupational Medicine*, Vol. 36, No. 7, pp.713–717.

Zurada J., Karwowski W. and Marras W.S., 1997. A neural network-based system for classification of industrial jobs with respect to risk of low back disorders due to workplace design, *Applied Ergonomics*, Vol.28, No.1, pp. 49–58.

Biographical notes

Pankaj Chandna is presently serving as an Associate Professor in the Department of Mechanical Engineering at the National Institute of Technology Kurukshetra Haryana, India. He has more than forty five research papers in international/national journals and conferences to his credit. His current research interest areas include sequencing and scheduling, ergonomics, supply chain management, inventory management, machine learning, etc.

Mahesh Pal is an Associate Professor in the Department of Civil Engineering at NIT Kurukshetra. He received his Ph.D. degree in year 2002 from university of Nottingham UK. His research interests are remote sensing and soft computing techniques. He has published more than fifty research papers in International journals and conferences.

Received February 2011
Accepted June 2011
Final acceptance in revised form September 2011