An Innovative Deep Neural Network for Stress Classification in Workplace

Conference Paper · February 2023

DOI: 10.1109/ICSCA57840.2023.10087794

3 authors:

Nikhil Patel
15 PUBLICATIONS 132 CITATIONS
SEE PROFILE

Sandeep Trivedi
25 PUBLICATIONS 169 CITATIONS
SEE PROFILE

Nuruzzaman Faruqui
Daffodil International University
28 PUBLICATIONS 282 CITATIONS
SEE PROFILE

An Innovative Deep Neural Network for Stress Classification in Workplace

Nikhil Patel
University of Dubuque
Iowa, United States of America
Email: patelnikhilr88@gmail.com,
ORCID: 0000-0001-6221-3843

Sandeep Trivedi
Deloitte Consulting LLP
Houston, United States of America
Email: sandeeptrivedi@ieee.org
ORCID: 0000-0002-1709-247X

Nuruzzaman Faruqui
Department of Software Engineering
Daffodil International University
Daffodil Smart City, Dhaka, Bangladesh
Email: faruqui.swe@diu.edu.bd
ORCID: 0000-0001-9306-9637

***Abstract*—Human Resource & Management (HRM) plays a vital role in organizational operations. The HRM tries to produce optimal output from human resources through workload balance. One of the core factors of workload balance is stress management. Although Deep Learning technology has introduced revolutionary applications in different sectors, its application in HRM is still nominal. This paper proposes an innovative application of Deep Learning to classify stressed and satisfied employees automatically. This generalized adaptive method utilizes quantitative measures which ensure unbiased classification with 88.40% accuracy and 0.8728 F1-score. The proposed network outperforms similar approaches, paving the path to applying Deep Learning based solutions to ensure a better workplace and proper workload balance through an effortless automatic but reliable stress classifier.**

***Index Terms*—HRM, Deep Learning, Stress Classification, Quantitative, Neural Network.**

I. Introduction

Stress directly influences the overall productivity of organizations [1]. It is essential to maintain proper workload balance to minimize the stage and maximize productivity. And the responsibility of stress management falls upon the HRM. Training, recreational opportunities, annual recognition, bonuses on performance, and many other strategies are there the HRM applies to keep the employees motivated and reduce stress [2]. The first challenge of this endeavor is to identify the stressed-out employees. Self-reporting, questionnaires, and observation are three standard methods to identify stressed employees [3]. These approaches have multiple drawbacks and limitations. Deep Learning-based automatic stressed employee classifier overcomes the limitations and removes the drawbacks with the assurance of accurate classification proposed in this paper.

First, the observational approach is entirely subjective and varies from observer to observer [4]. Moreover, human psychology is complex, influenced by many factors, and never assures unbiased judgment [5]. That is why subjective human observation never guarantees that the stressed-out employees are detected correctly. Instead of subjective observation, a quantitative method performs much more reliably in this regard designed in this research. On the other hand, the questionnaire-based approach leads to another complicated challenge. It imposes a deadlock situation. Studies suggest that when the identity is revealed, participants don't respond truthfully in questionnaire-based approaches [6]–[8]. The HRM must know the participants' identities to discover and help stressed employees. An automatic computerized system that keeps the employee identity anonymous during the data processing is a solution to this problem. Such an automatic system has been designed, experimented with, and compared in this paper.

The self-reporting is considered an effective way to identify and help stressed employees [9]. However, it leaves scope for system misuse. Unethical employees take advantage of such a system to reduce their workload. One of the essentials in stress management is quick response [10]. The HRM of a large organization can easily be overloaded with stress-related self-reports. Instant responses become a challenge and cause of stress for the HRM team. Studies suggest that introverted people are stressed out [11], [12]. It is also evident that introverted people are less likely to express their problems [13]. Altogether, it leads to the conclusion that self-reporting imposes additional challenges and is not practical for certain people who are more likely to be stressed. The deep neural network proposed in this paper overcomes these challenges and dissolves the complication the HRM faces in identifying and helping stressed employees.

The main focus of this research is to solve an existing HRM problem regarding stress management using deep learning. The experiment and the endeavors of the researchers involved in this paper contribute in:

- **Novel Approach:** Design, implementation, analysis, and optimization of a deep neural network to automatically classify stressed and non-stressed employees.
- **Establishing Credibility:** Discovering the proposed solution's credibility through statistical analysis and comparison with similar approaches.
- **Problem Solving:** Providing a solution to an existing HRM-related problem and paving the path to further research to develop the scientific branch of applied Deep Learning in HRM.

The rest of the paper has been organized into five following sections. The second section highlights and correlates our approach with other existing research. The methodology used in this research has been discussed in the third section. The fourth section is about experimental results and performance evaluation. After that, this paper's limitations and future scope have been discussed in section five. Finally, the paper has been concluded in section six by summarizing the entire paper.

II. Literature Review

An innovative and practical approach to detecting human stress using wearable sensors using a convolutional neural network developed by Manuel Gil-Martin at el. classifies stressed and non-stressed classes with 96.6% accuracy. Their three fully connected networks receive bio-signals and make predictions [14]. Although it is a practical application of deep learning in stress classification, it is not a proper solution to workplace stress classification. It is dependent on wearable sensors, which cannot be imposed in most workplaces. Moreover, stress induced by personal life is not the subject of interest of the HRM. However, the approach of [14] fails to cluster different types of stress. The proposed methodology uses deep neural networks to solve a real-world problem the HRM faces regularly. It does not require the employees to wear anything. As a result, it becomes more practical and effective from an application perspective. The study by Anna-Maria Hultén et al. points right on the problem that this paper deals with. They used a Work Stress Questionnaire (WSQ) based approach to discover the correlation between work stress and future sick leave of the employees. This study shows that 36% of the employees were on sick leave for more than 15 days a year. And more than 72% of employees perceiving high stress were on leave for more than 15 days a year. The experimental result indicates a leave application caused by [15]. Although the findings of this research are impactful, the WSQ method suffers from honest responses [16]. There are instances when the participants won't be truthful, raising questions about the overall integrity of the WSQ method [17], [18]. Moreover, this is a manual and time-consuming process. The proposed methodology is entirely automatic and thus faster than any manual process. At the same time, there is no integrity issue here because the participants have no discomfort or fear of identity disclosure to others.

A deep learning-based stress prediction approach for improving and bettering workplace conditions has been studied by Yu Zhang et al. Their purpose and domain of interest align with the research concentration of the proposed paper. It introduces remarkably better results. However, the proposed methodology of this paper achieved 88.40% accuracy, which is 17.20% higher than Yu Zhang et al. Moreover, they did not apply any data preprocessing. There are two variables in the ESI (Employee Satisfaction Index) [19] dataset, which ranges do not maintain coherence with the rest of the dataset. It is essential to normalize these variables, which has been done in the proposed paper.

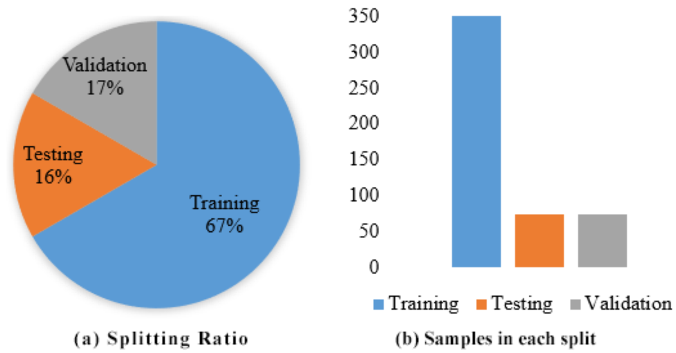

Fig. 1. (a) Data splitting and (b) instance distribution

III. Methodology

A. Dataset

The Employee Satisfaction Index (ESI) dataset has been used in this experiment. This dataset has 14 variables and 500 observations. Two variables are serial number and unique employee ID. These two variables have no impact on employee satisfaction prediction. That is why they have been removed from the experimenting dataset. In the modified dataset, four variables contain string datatype, and the rest of the eight variables contain numeric data. The last variable, the target variable, contains two classes identified by 0 and 1. The four-string variables are department, employee location, education status, and type of recruitment. Two variables, job level and rating, are scaled from 1 to 5. The onsite and certifications variables are categorical. There are only two categories. The age and salary are numeric. However, these two variables contain the most imbalanced data range. The mean normalization [20] has been applied for age and salary variables to maintain the data range balance using equation 1.

$$O_{norm} = \frac{v - v_{min}}{v_{max} - v_{min}} \tag{1}$$

After normalizing the age and salary variable, the dataset maintains a balanced range scale. Not every variable takes part in predicting the target variable. That is why variables without significant correlation with the target variable have been removed. The table I lists a subset of the dataset before and after normalization. It also gives the overall idea about the variables of the dataset.

1) Data Splitting: The dataset has been split into training, validation, and testing datasets by maintaining a 14:3:3 ratio. There are total 500 instances in the dataset. Among these, 350 instances have been used for training, 75 for testing and remaining 75 for validation. The dataset splitting and number of instances in each group have been illustrated in figure 1.

The k-fold cross validation has been used to validate the learning process. The experimental approach shows that the learning process demonstrates optimal performance at $k = 5$. A separate dataset for testing could not be managed because of the scarcity of matured datasets. However, the test segment

TABLE I
A SUBSET OF THE DATASET BEFORE AND AFTER NORMALIZATION

Before							After						
Age	Job_Level	Rating	Online	Awards	Certifications	Salary	Age	Job_Level	Rating	Online	Awards	Certifications	Salary
28	5	2	0	1	0	86750	0.2	5	2	0	1	0	1.00
50	3	5	1	2	1	42419	0.9	3	5	1	2	1	0.29
43	4	1	0	2	0	65715	0.6	4	1	0	2	0	0.66
44	2	3	1	0	0	29805	0.7	2	3	1	0	0	0.09
33	2	1	0	5	0	29805	0.3	2	1	0	5	0	0.09

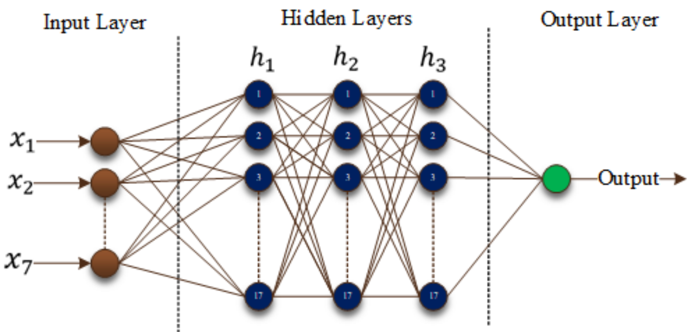

Fig. 2. Caption

TABLE II
THE VALUES OF THE EVALUATION MATRICES

Evaluation Matrix	Value	Percentage
Accuracy	0.884	88.40%
F-measure	0.8728	87.28%
Precision	0.9087	90.87%
Recall	0.8379	83.79%
Phi coefficient	0.7685	76.85%

remained untouched and had not been used in training and validation. It has been used for testing the network. This approach assures the overall integrity of the proposed performance of the experimenting network.

B. Network Architecture

This experiment has designed and optimized a fully connected feed-forward deep neural network with 3 hidden layers to predict employee satisfaction. The input layer has eight nodes, and each hidden layer has 17 nodes. The Sigmoid function [21] has been used as the activation function of hidden and output nodes. The network is defined by equation 2.

$$Output = f_s(\sum_{L=1}^{4}\sum_{P=1}^{15} F_s((N_P^L \times W_P^L) + b_L)) \quad (2)$$

Here in equation 2, the f_S stands for Sigmoid function. The N_P^L means the node of layer L located at position P. The W_P^L represents the weight of layer L and position P. The b_L expresses the bias of layer L. The network has been illustrated in the figure 2.

C. Optimizer Selection

We experimented with four different optimization algorithm to train the network accurately and efficiently. These algorithms are Nesterov Accelerated Gradient, Adaptive Gradient Algorithm (Adagrad), Adaptive Delta (Adadelta), and Adaptive Moment Estimation (ADAM). Among these four, the ADAM gives the best result. To apply the ADAM optimizer, we calcuated the exponentially weighted average gradient descent using equation 3.

$$w_{t+1} = w_t - \alpha m_t \quad (3)$$

Here, the m_t is defined by equation 3.

$$m_t = \beta m_{t-1} + (1-\beta)[\frac{\delta L}{\delta w_t}] \quad (4)$$

The Adaptive Moment Estimation (ADAM) optimizer [22] we used is defined by equation 4.

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. Evaluation Metrics

In this experiment, the accuracy (equation 5), F-measure (equation 6), precision (equation 7), and recall (equation 8) have been used as the evaluation matrices.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

B. Experimental Result

The experimental results have been evaluated using a confusion matrix illustrated in figure 3. The True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) have been calculated from it. Later, these values have been used to calculate the accuracy, F1 score, precision, recall and ϕ-coefficient and listed in table II. Here on the confusion matrix of figure 3, we have 174 true positives, 21 false negatives, 20 false positives, and 143 true negative values.

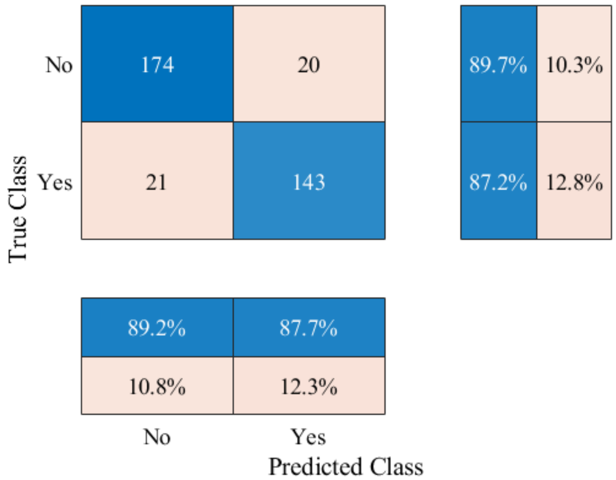

Fig. 3. The confusion matrix analysis

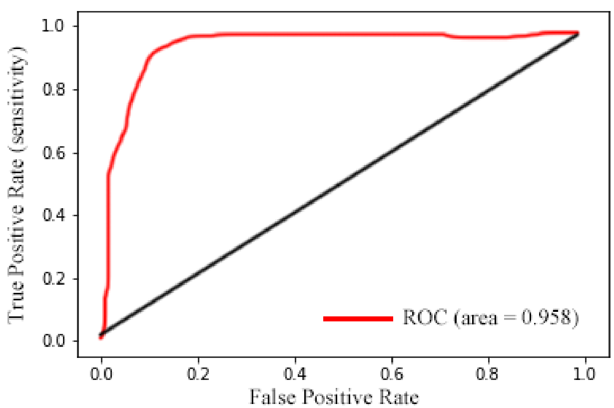

Fig. 4. Performance evaluation through Receiver operating Characteristic (ROC) curve

C. Performance Evaluation

The Receiver operating Characteristic (ROC) curve illustrated in figure 4 demonstrate the ability of the proposed network to perform binary classification accurately. It represents the True Positive Rate (TPR) against the False Positive Rate (FPR). The Area Under the Curve (AUC) of the classifier is 0.958. That means if we randomly pick a set of instances, the probability of getting TPR is 95.8% and the FPR is 4.2%. It proves that the classification by the proposed classifier is reliable.

D. Performance Comparison

Although facial recognition-based attendance systems have become common nowadays, it is a fraction of the diverse operations performed by HRM professionals. The application of deep learning in human resource management is still an underexplored domain. To the best knowledge of the researchers working on this paper, only one well-written comparable paper

TABLE III
PERFORMANCE COMPARISON

Sequence	Algorithm	Accuracy	F1 Score
1	Zhang et al. [17]	71.20%	68.60%
2	SVM	60.90%	57.90%
3	DT	54.20%	50.06%
4	NB	68.44%	63.70%
5	Proposed	88.40%	87.28%

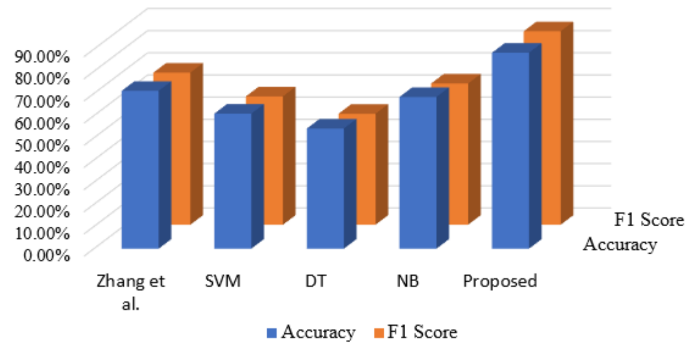

Fig. 5. Caption

by Zhang et al. [19] is available in this field. The performance of our proposed method has been compared with this paper. It has been compared with Support Vector Machines (SVM) [23], Decision Tree (DT) [24], and Naïve Bayes (NB) [25] as well. The result of the comparison has been tabulated in table III.

The performance comparison demonstrates that the proposed network's accuracy and F1 score outperform similar approaches. Figure 5 illustrates this comparison. It demonstrates the significant improvement in the accuracy and F1 score of the proposed method over existing algorithms.

V. Limitation and Future Scope

The proposed method demonstrates remarkable achievement in improving the accuracy and F1-score in employee satisfaction classification using deep learning technology. However, like any system in this universe, it is not immune to limitations.

1) Dataset Invariant Approach: The first limitation is the dataset invariant approach. The HRM datasets are companies' assets,, and usually,, they are not disclosed. As a result, no other applicable dataset was used to evaluate the network performance. Although the performance was evaluated based on a test dataset, which has not been used during the training or validation, testing the network with another dataset would increase the credibility of the proposed method. However, an initiative has been taken to collect relevant data from some organizations with the identity non-disclosure agreement. These datasets will be used in the future to evaluate the performance of this proposed methodology further.

2) Comparison Limitation: The performance of the proposed network has been compared with four different classifiers. Among these four classifiers, only Zhang et al. [19]. was published research. The rest of the three were tested during the experiment. No relevant research has been discovered during the performance comparison. That is why the performance has been compared with a limited number of algorithms. However, initiatives have been taken to replicate this proposed method using other classification algorithms, which will be compared with this proposed network in future.

3) Binary Classification: The proposed network is limited to binary classification. There are employees who are not completely dissatisfied nor fully satisfied with their job. However, it has not been addressed in this paper. The purpose behind not extending the network to multiclass classification is the scarcity of relevant datasets. The research team of this paper is actively working on developing a dataset for multiclass classification, which will be published in subsequent papers.

The limitations of the proposed methodology are caused by the scarcity of relevant datasets and a limited number of published research on the application of deep learning in HRM. However, the outstanding performance dims the limitations of the approach. Despite limitations, the research team of this experiment considers these as an opportunity for further research.

VI. Conclusion

Employee satisfaction is a crucial factor in organizational productivity. However, assuring it is a challenging task for Human Resource Management. Computer-assisted, deep learning-based automatic systems have become state-of-the-art solutions for challenging tasks where human intelligence and sentiment are essential. A successful application of such technology has been uncovered in this paper. The innovative design of the four-layer deep neural network with fifteen nodes in each layer has demonstrated a good accuracy of 88.40% in stressed and non-stressed employee classification. The performance has been evaluated using the ROC curve, PR AUC curve, gain curve, and K-S Statistics to ensure the approach's credibility. It has been compared with the existing research as well. The performance, evaluation, and comparison demonstrate the network proposed in this paper as the best performing workplace stress classifier. It paves the path for adapting deep learning technology to HRM applications.

References

[1] J. He, X. Guo, H. Cui, K. Lei, Y. Lei, L. Zhou, Q. Liu, Y. Zhu, and L. Liu, "Stress sensitivity of tight sand reservoir and its influence on productivity of well," *Petroleum Science and Technology*, vol. 40, no. 14, pp. 1665–1680, 2022.

[2] J. Kapoor and P. Chhabra, "Work stress management in an organization with the role of hrm," *Journal of Positive School Psychology*, vol. 6, no. 2, pp. 233–240, 2022.

[3] C. Gilbert, "Non-verbal pride expressions as a predictor of lgbtq health," 2022.

[4] T.-C. Chou and H.-P. Lu, "How to observe business operations: An empirical study of family business," *Plos one*, vol. 17, no. 4, p. e0267223, 2022.

[5] P. B. Sharp, I. Fradkin, and E. Eldar, "Hierarchical inference as a source of human biases," *Cognitive, Affective, & Behavioral Neuroscience*, pp. 1–15, 2022.

[6] S. M. Hulke, S. L. Wakode, A. E. Thakare, R. Parashar, R. N. Bharshnakar, A. Joshi, and Y. P. Vaidya, "Perception of e-learning in medical students and faculty during covid time: A study based on a questionnaire-based survey," *Journal of Education and Health Promotion*, vol. 11, 2022.

[7] L. Kelly, J. J. Kurinczuk, R. Fitzpatrick, F. Alderdice, *et al.*, "Refinement of the well-being in pregnancy (wip) questionnaire: cognitive interviews with women and healthcare professionals and a validation survey," *BMC pregnancy and childbirth*, vol. 22, no. 1, pp. 1–14, 2022.

[8] R. Lipinski, D. Rogger, and C. Schuster, "Designing survey questionnaires (ii): What types of survey questions do public officials not respond to?," 2022.

[9] M. Balconi, G. Fronda, F. Cassioli, and D. Crivelli, "Face-to-face vs. remote digital settings in job assessment interviews: A multilevel hyperscanning protocol for the investigation of interpersonal attunement," *Plos one*, vol. 17, no. 2, p. e0263668, 2022.

[10] S. L. Martindale, R. D. Shura, M. A. Cooper, S. F. Womack, R. A. Hurley, C. L. Vair, and J. A. Rowland, "Operational stress control service: An organizational program to support health care worker well-being," *Journal of Occupational and Environmental Medicine*, vol. 64, no. 1, p. 64, 2022.

[11] A. Benabed, "Exploring the impacts of the introversion personality traits on algerian efl students' oral proficiency: First year ba students at ibnkhaldoun university-tiaret as a sample," *Technium Social Sciences Journal*, vol. 32, pp. 223–240, 2022.

[12] R. Spencer, "'hell is other people': Contemplating seligman on sartre and recognising the different learning styles of extroverts and introverts in legal education," *Rachel Spencer (2022)"Hell is other people": rethinking the Socratic method for quiet law students, The Law Teacher, DOI*, vol. 10, no. 03069400.2021, p. 2005305, 2022.

[13] H. Senuri, "Impact of extroversion and introversion on the use of communication strategies at different proficiency levels,"

[14] M. Gil-Martin, R. San-Segundo, A. Mateos, and J. Ferreiros-Lopez, "Human stress detection with wearable sensors using convolutional neural networks," *IEEE Aerospace and Electronic Systems Magazine*, vol. 37, no. 1, pp. 60–70, 2022.

[15] A.-M. Hultén, P. Bjerkeli, and K. Holmgren, "Work-related stress and future sick leave in a working population seeking care at primary health care centres: a prospective longitudinal study using the wsq," *BMC public health*, vol. 22, no. 1, pp. 1–12, 2022.

[16] K. M. Alharthi, *The impact of writing strategies on the written product of EFL Saudi male students at King Abdul-Aziz University*. PhD thesis, Newcastle University, 2012.

[17] W.-g. Kim and H. Lee, "Multimodal biometric image watermarking using two-stage integrity verification," *Signal Processing*, vol. 89, no. 12, pp. 2385–2399, 2009.

[18] B. S. Swann, J. M. Libert, and M. A. Lepley, "Tools for quality control of fingerprint databases," in *Biometric Technology for Human Identification VII*, vol. 7667, pp. 82–93, SPIE, 2010.

[19] Y. Zhang and E. Qi, "Happy work: Improving enterprise human resource management by predicting workers' stress using deep learning," *Plos one*, vol. 17, no. 4, p. e0266373, 2022.

[20] G. Yang, J. Pennington, V. Rao, J. Sohl-Dickstein, and S. S. Schoenholz, "A mean field theory of batch normalization," *arXiv preprint arXiv:1902.08129*, 2019.

[21] X. Yin, J. Goudriaan, E. A. Lantinga, J. Vos, and H. J. Spiertz, "A flexible sigmoid function of determinate growth," *Annals of botany*, vol. 91, no. 3, pp. 361–371, 2003.

[22] W. K. Newey, "Adaptive estimation of regression models via moment restrictions," *Journal of Econometrics*, vol. 38, no. 3, pp. 301–339, 1988.

[23] W. S. Noble, "What is a support vector machine?," *Nature biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.

[24] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 18, no. 6, pp. 275–285, 2004.

[25] G. I. Webb, E. Keogh, and R. Miikkulainen, "Naïve bayes.," *Encyclopedia of machine learning*, vol. 15, pp. 713–714, 2010.