ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940
www.ijsem.org
Vol 14, Issuse.4 Oct 2023

MECHANICAL TOOLS CLASSIFICATION USING MACHINE LEARNING

DR E.L. NAGESH*, DR RAMESH DARA**

ABSTRACT

It is not reliable to depend on a persons inference on dense data of high dimensionality on a daily basis. A person will grow tired or become distracted and make mistakes over time. Therefore it is desirable to study the feasibility of replacing a persons inference with that of Machine Learning in order to improve reliability. One-Class Support Vector Machines (SVM) with three different kernels (linear, Gaussian and polynomial) are implemented and tested for Anomaly Detection. Principal Component Analysis is used for dimensionality reduction and autoencoders are used with the intention to increase performance. Standard soft-margin SVMs were used for multi-class classification by utilizing the 1vsAll and 1vs1 approaches with the same kernels as for the one-class SVMs. The results for the one-class SVMs and the multi-class SVM methods are compared against each other within their respective applications but also against the performance of Back-Propagation Neural Networks of varying sizes. One-Class SVMs proved very effective in detecting anomalous samples once both Principal Component Analysis and autoencoders had been applied. Standard SVMs with Principal Component Analysis produced promising classification results. Twin SVMs were researched as an alternative to standard SVM.

INTRODUCTION

Machine Learning, in all its diverse forms, has been used to solve a multitude of problems - for example classification[1, 2], detection[3, 4], regression[5] and optimization[6, 7]. At its core Machine Learning is a marriage between statistical theory and signal processing, which gave birth to Support Vector Machines (SVM), Neural Networks (NN), Genetic algorithms and a plethora of other algorithms with various ad-ons. The task in this thesis is detecting flaws and shortcomings in a mechanical product based on the measured moments from rotation of a key component of the product. The moments reveal not only if something is wrong within the product but also clues as to what the source of the problem might be. Currently these measurements are interpreted manually by employees at the production site, however it is anticipated that an autonomous solution will bring several benefits including, but not limited to, improved anomaly detection and better reliability. In general this solution could be useful for a multitude of industries where anomaly detection and classification is needed. There are two main questions regarding this task: 1) How is accurate and reliable anomaly detection as well as classification achieved, and 2) How does it compare to a person performing the task manually? To answer these questions research was conducted broadly to find what methods has been used prior and what their results were. Once a promising

* Principal, MNR College of Engineering and Technology, Sangareddy, Mohd.Shapur, Telangana 502285
** Principal, Visvesvaraya college of engineering and technology, Patel Guda, Mangalpalle, Telangana 501510

ISSN2454-9940
www.ijsem.org
Vol 14, Issuse.4 Oct 2023

direction was found a deeper study was made and extensive testing and analysis was performed. With an initially broad scope which narrows as the research progresses a generally favorable solution is expected to be found. For this particular task a fully autonomous solution is not only favorable to the production but to the workers as well. Over time a worker will grow tired which increases the risk of overlooking measurements and misclassifying samples which will reduce the overall quality of the products - but there is also the risk of repetitive strain-injuries. Such an injury is difficult to recover from and can hurt both the company and the individual. With an autonomous solution it is anticipated that these issues can be overcome and the production process might even become more time efficient. SVMs have been used and developed over several decades and have been proven useful in binary classifications but also for multi-class classification through application of approaches based on cross-class comparisons and Fuzzy Logic [8, 9]. A wide selection of data types and applications have been delt with by SVMs by utilizing various combinations of methods for creating the separating hyperplanes, estimating classification errors and adjusting the hyperplane in an efficient manner. [10, 11, 12]. This thesis will focus mainly on various forms of SVMs for the sake of anomaly detection and classification which will be compared to the capabilities of Back-Propagation NN. To help the algorithms achieve a higher accuracy some pre-processing will be applied to the data sets in advance Figure 1: A model of the testbench. A sensor is attached to a shaft between an electric motor and a mechanism which can not be observed from outside. When the shaft is turned the sensor sends the moment measurements to a computer where a person will analyse the moment plot. In the illustration it is indicated that the mechanism is not properly lubricated, which should be revealed in the plot. 1 S. Bengtsson Machine learning for mechanical analysis and the effects of this will be studied and analysed. Due to the versatility of SVMs it is reasonable to believed that a SVM-based solution can be found that can produce a satisfactory accuracy and reliability for anomaly detection and classification for the mentioned task. The remainder of the report will be structured as; The Problem Formulation will firstly detail the problem which this thesis will attempt to solve as well as stating the hypothesis and research questions. The Background-section will cover the notation and basic concepts used in this report after which the key algorithms and methods used will be introduced more in depth than has been done so far. Related Work will go through the development of SVMs and NNs from conception to the modern era and end with an account of the state of the art. All methods, algorithms and techniques used as well as the reasoning to why they were used will be discussed in the Method section. The stance towards use of personal data and ethical considerations will be stated in Ethical and Societal Considerations, followed by a section dedicated to describing the work process of this thesis. In Results the findings will be presented in detail through tables and analysis. Conclusions will seek to concisely summarize the findings and results of this thesis. In Discussion the results will be examined in regard to the research questions. The report will close with Future Work, suggesting what more might be done to continue the development of a Machine Learning solution with increased capabilities.

LITERATURE REVIEW

In 1943 Warren McCulloch and Walter Pitts sowed the seed for what would become Artificial Neural Networks[23].

ISSN2454-9940
www.ijsem.org
Vol 14, Issuse.4 Oct 2023

They proposed a mathematical model called threshold logic which attempted to mimic some of the functionality of the neurons in a brain. The ”perceptron” was proposed by Rosenblatt in 1958, a further mathematical and computational mimicry of biological cellular functionality. From this perceptron came the first Neural Networks. It was however not until 1974 that the back-propagation algorithm was proposed, stemming from H. Kelleys work in 1960[24, 25]. This invention revitalized the research into Neural Networks to the point that the major limiting factor of the algorithm was the hardware. Due to the size of the network and the large amounts of neurons necessary for it to produce a useful result the amount of time to train the network was often in the range of months, depending on the problem and application of the network. As the speed of processors increased and computer memory became more cheaply accessible the popularity of Neural Networks grew. Wei Zhang et al constructed a multi-layered feed-forward parallel distributed processing model in 1990 which used the Neural Network methodology[26]. This model was capable of classifying letters even when they were tilted, shifted or distorted and would serve as the foundation for Convolutional Neural Networks (CNN or ConvNets). Another branch of Neural Networks is the Recurrent Neural Network (RNN), which is based on the work of Rumelhart in 1986. RNNs incorporate its own output values as inputs to itself to give temporal information. A kind of RNN is Long Short-Term Memory (LSTM) discovered in 1997 by Hochreiter and Schmidhuber which proved excellent in speech-recognition and other contextdependent applications[27, 14]. In 1962 Vapnik and Lerner published an article (translated to English from Russian in 1963) about their idea of a ”Generalized Portrait algorithm” where they also gave an axiomatic definition of patterns based on decomposition of images into subsets[8]. This proposed algorithm was only applicable to linear sets of data and was highly susceptible to noise and outliers. T.M. Cover developed the idea of hyperplanes for pattern separation in 1965, which laid the foundation for large margin hyperplanes[17]. After some development, this algorithm was still fairly limited as it could only be applied to linearly separable binary classes. However, this changed with the introduction of kernels which had previously been researched by Aiserman, Braverman and Rozonoer in 1964[28]. Kernels were realized as a useful tool in SVMs in 1992 by Boser, Guyon and Vapnik and enabled classification of non-linearly separable data by transforming the given data into a feature space where linear separability was possible[16]. In 1995 Cortes and Vapnik introduced the ”soft margin” where each data sample xi is assigned a variable $\zeta i \geq 0$[18]. During training it is attempted to find a solution where these ζi values are minimized as they are indicators of how ill-fitting the current iteration of the classification hyperplane is. Up until this point SVMs utilized what is called a ”hard margin”, meaning that a point of data is either on the right or wrong side of the classification hyperplane with no indication of how right or wrong the samples was classified in terms of distance from the hyperplane. There exists a myriad of adaptations of SVM for various problem solutions[8]. Xi-Zhao Wang and Shu-Xia Lu incorporated Fuzzy Logic into a SVM where a fuzzy membership value was made part of the objective function as a factor to the loss values[9]. By utilizing Fuzzy Logic some of the inherent sensitivity to outliers in ordinary SVMs was overcome. The proposed Improved Fuzzy Multi-category SVM (IFMSVM) achieved a slight but noticeable improvement to the classification scores as compared to a 1vsAll, 1vs1 and Multi-category SVM on various sets of data. Support Vector Machines continue to be developed in the 20th century. Since its conception all SVMs used a single plane with a

ISSN2454-9940
www.ijsem.org
Vol 14, Issuse.4 Oct 2023

surrounding parallel margin to perform classification. In 2006 Mangasarian and Wild introduced the Generalized Eigenvalue Proximal SVM (GEPSVM), in which 10 S. Bengtsson Machine learning for mechanical analysis min w(1),b(1) ,ζ 1 2 (Aw(1) + e1b (1)) T (Aw(1) + e1b (1)) + c1e T 2 ζ (6) subject to − (Bw(1) + e2b (1)) + ζ ≥ e2, ζ ≥ 0 (7) min w(2),b(2) ,ζ 1 2 (Bw(2) + e2b (2)) T (Bw(2) + e2b (2)) + c2e T 1 ζ (8) subject to (Aw(2) + e1b (2)) + ζ ≥ e1, ζ ≥ 0 (9) TSVM as a constrained minimization problem. A and B are matrices containing the training vectors, w is the weight-vector, e is a 1-vector of appropriate dimensions, b is the bias-terms, c is a trade-off constant and ζ is the slack-vector. a plane with a maximized margin towards the given samples are replaced by two non-parallel planes of maximum distance from each other[29]. The planes are eigen-vectors gained from finding a pair of related generalized eigen-value problems with the smallest eigen-values. Each plane attempts to be as close as possible to one class while creating as large distance as possible to the other class, see Figure 8. In 2007 Jayadeva et al proposed another multi-plane SVM which, just like Mangasarians and Wilds GEPSVM, replaces the maximum-margin classifier with two non-parallel planes[30]. The main difference between these two approaches is to do with the formulation of the planes. In GEPSVM the planes are eigen-vectors while in Jayadevas approach, coined ”Twin Support Vector Machine” (TSVM), the planes are very much the same as in normal SVMs. Though they could not display a significant increase in accuracy from a conventional SVM, the TSVM did have a remarkably shorter training time since instead of a single major QP problem, as solved for in ordinary SVMs, two smaller Quadratic Programming (QP) problems are solved for the TSVM. Jayadeva et als formulation for this Twin SVM is given in Equations 6 - 9 as formulated in the notation convention adopted by this thesis.

min w(1) 1 2 ||w(1)||2 + v1 l2 X l1 i=1 L x + i , yi , fi(x + i) (10) min w(2) 1 2 ||w(2)||2 − v2 l1 X l2 j=1 L x − j , yj , fj (x − j) (11) TSVM as an Unconstrained Minimization Problem with a loss-function. w is the weight-vector, v is the difference of the current iteration of w and the previous iteration, l is the number of samples, f() is the classification function. In 2019 Sharma et al reformulated Jayadeva et als TSVM into two unconstrained minimization problems with a loss function. The objective functions are given in Equation 10 and 11. In their report Sharma et al proposed a stochastic solution to this minimization problem using a quasi-Newton method and approximations of the Hessian matrices as these are computationally expensive to calculate. This TSVM is denoted as SQN-PTWSVM (Stochastic Quasi-Newton Pinball Twin Support Vector Machine). Sharma et al also made a comparison between the conventional Hinge loss function and the Pinball loss function with the conclusion that the latter had some prominent advantages over the former[31]. Most importantly, the Pinball loss function produces stability by removing some sensitivity to noisy training data, thus promoting quicker convergence. If τ is set to zero the Pinball loss function would be the exact same as the Hinge loss function. Equation 12 showcases the Pinball S. Bengtsson Machine learning for mechanical analysis Lτ (x, y, f(x)) = (−yf(x), if − yf(x) ≥ 0 (Incorrect classification) τyf(x), if − yf(x) < 0 (Correct classification) (12) The Pinball loss function. x is a data sample, y is the label of x, f(x) is the classification value of x, τ is the penalty rate applied to correctly classified samples. loss function. With extensive testing and comparisons of various SVMs it was shown that, on average, the stochastic quasi-Newton optimization technique coupled with the Pinball loss function for a Twin Support Vector Machine was greater than both

SVMs and conventional TSVMs in terms of both accuracy and training time.

EXISTING SYSTEM:

The existing system for mechanical tools classification utilizing Machine Learning represents a monumental leap forward in automating the categorization of tools based on their distinct characteristics and functionalities. This system harnesses the power of advanced algorithms, particularly convolutional neural networks (CNNs), to process extensive datasets comprising images of various tools. By analyzing pixel-level information, the algorithms can discern crucial features such as shape, texture, and structural attributes. This allows for the precise identification and differentiation of tools into two primary categories: hand tools and power tools.

Hand tools encompass a broad spectrum of manually operated implements designed for specific tasks. Through the application of Machine Learning, these tools can be recognized based on their unique physical attributes. For instance, a wrench can be identified by its adjustable jaws and elongated handle, while a hammer is distinguished by its flat striking surface and short, sturdy handle. Screwdrivers are differentiated by their slender, pointed tips, tailored for turning screws. Within this category, further sub-classifications emerge, including cutting tools like chisels and saws, gripping tools like pliers, and striking tools like mallets and sledgehammers.

In contrast, power tools constitute a subset of mechanical instruments reliant on an external power source, typically electricity or compressed air, to operate. These tools exhibit enhanced efficiency and the capacity to tackle heavier workloads. Machine Learning models excel at distinguishing various power tools based on features such as motor size, handle configuration, and additional attachments. For example, a circular saw is identified by its rotating blade, while a drill is recognized by its rotating bit. Sanders are classified based on their oscillating sanding surface, and angle grinders by their rotating abrasive discs.

Moreover, Machine Learning significantly contributes to the classification of tools according to their specialized functions within industries. Woodworking tools, for instance, comprise saws, planes, and routers, each tailored to execute precise tasks in woodworking projects. Metalworking tools encompass equipment like lathes, milling machines, and welding apparatus, specifically designed for shaping and manipulating metal. Automotive tools, crucial for vehicle maintenance and repair, include wrenches, socket sets, and diagnostic equipment. Plumbing tools, which facilitate plumbing installations and repairs, consist of pipe wrenches and plunger-type drain cleaners.

In addition to categorizing tools based on their physical attributes, the existing system also takes into consideration contextual factors, such as the environment in which they are typically employed and the materials they are intended to work with. This comprehensive approach to tool classification has significantly improved efficiency and accuracy across industries reliant on mechanical instruments. By integrating Machine Learning, the process of identifying, categorizing, and utilizing mechanical tools has been streamlined, resulting in heightened productivity and precision across a diverse range of applications. This existing system stands as a testament to the transformative power of Machine Learning in the field of mechanical tools classification.

PROPOSED SYSTEM

The proposed system for mechanical tools classification using Machine Learning represents a cutting-edge approach to automate the categorization of tools based on their unique attributes and functionalities. This system leverages state-of-the-art deep learning techniques, particularly convolutional neural networks (CNNs), to process extensive datasets comprising images of various tools. By analyzing pixel-level information, the

ISSN2454-9940
www.ijsem.org
Vol 14, Issuse.4 Oct 2023

algorithms can discern crucial features such as shape, texture, and structural attributes. This allows for the precise identification and differentiation of tools into two primary categories: hand tools and power tools.

Hand tools, encompassing a wide array of manually operated implements designed for specific tasks, are accurately recognized based on their distinctive physical attributes. For instance, a wrench is identified by its adjustable jaws and elongated handle, while a hammer is distinguished by its flat striking surface and short, sturdy handle. Screwdrivers are differentiated by their slender, pointed tips, tailored for turning screws. Within this category, further sub-classifications emerge, including cutting tools like chisels and saws, gripping tools like pliers, and striking tools like mallets and sledgehammers.

IMPLEMENTATION

Below will be given a thorough narration of the process, progress and obstacles of the thesis.

Algorithm research and preparation The first step to take was acquiring the data from the company. This proved more difficult than it first seemed as the only available data was from a prototype of the testbench which was fairly limited in resolution and quality - the final version of the hardware had yet to be installed at this point in time. Proper data was promised to be delivered eventually, which was no pressing issue since a reasonable understanding of the data was gained through the prototype outputs. The data was a univariate time-series with a couple of hundred features per sample and each sample was labeled as either ”approved” or ”not approved”. With this knowledge research into potential solutions was initiated. Classification and anomaly detection methods within Machine Learning was researched. Support Vector Machines (SVM) were understood to be a suitable family of classifiers for RQ1, as it was a binary classification problem. After some more research it was learned that SVM has been used with good results for multiclass classification by utilizing the 1vs1 approach (pairwise decomposition) and the 1vsAll approach which opened up for SVM being the topic of RQ3 as well. Though Fuzzy SVM has been shown to have a slightly higher accuracy in the case of multiclass SVM the improvement was so small that the increased complexity as compared to 1vs1 and 1vsAll was not deemed a feasible trade-off at this stage [9]. As a frame of reference for the SVM implementations Back-Propagation Neural Network (BPNN) was chosen based on its current popularity as a classifier. From the multitude of SVMs found during the research three were chosen to be in the scope of this thesis: Standard SVM, One-Class SVM (OC-SVM) and Stochastic Quasi-Newton Pinball Twin SVM (SQN-PTWSVM)[31].

Data processing

research Eventually, data from the final version of the testbench was produced and made available. The data was composed of several thousand features per sample which was a greater number of features than anticipated. This prompted the need to decrease the size of the data samples as the raw data was bulky and cumbersome. Further research was conducted which covered Independent Component Analysis (ICA), subtractive clustering, Fuzzy c-means clustering and more[37, 38, 39]. Eventually it was settled for Principal Component Analysis (PCA) to decrease the dimensionality of samples. It was also learned about the positive effects which autoencoders can have on classification performance, which led to incorporating this into the thesis as well.

Anomaly detection implementation

It was thought best to tackle the research questions chronologically, meaning that the capabilities of OC-SVM on anomaly detection was the first to be implemented and evaluated in conjunction with a BPNN. For the sake of rigorous testing and demonstrating the versatility of the implementations additional data sets were gathered from online sources, such as the UCI Machine Learning Repository and timeseriesclassification.com[35]. Using MatLab, a BPNN was set up and its accuracy was tested on the data sets gathered for the sake of anomaly detection through 10-fold cross-validation. A OC-SVM was also set up and tested in the same fashion with three different kernels: Linear, Gaussian and Polynomial. Once the baseline performance of these two methods was established the data processing methods were applied and the results documented. To gain a solid understanding of the effects of the data processing methods the classifiers were initially tested with the processing methods separately before being tested with both methods applied simultaneously. All results were recorded and can be found in the corresponding subsection in section Results. 18 S. Bengtsson Machine learning for mechanical analysis Figure 10: An illustration of the iterative method used for this thesis.

Classification The final version of the testbench which the company had built features multiclass labeling capabilities. These had yet to be used for the data which had been made available for this thesis. There was hope that multiclass-labeled data would be produced in the time-span of this thesis, however this did not happen. Instead, more data sets were gathered to test the general performance of the implemented classifiers. The testing procedure of the classifiers were very much the same as for anomaly detection: Each solution was tested with 10-fold cross-validation on each data set and each SVM implementation was tested with the linear, Gaussian and polynomial kernel. Once a baseline performance was established the solutions were tested together with PCA. All the results are available in the corresponding subsection in Results. 7.5 SQN-TWSVM Due to its reportedly good performance, Sharma et als Stochastic Quasi-Newton Pinball Twin SVM (SQN-PTWSVM) was implemented based on their descriptions[31]. It was hoped that this new and promising SVM would provide even better results than the other solutions. No conclusive results could be made however, even after weeks of work. The notation and descriptions used in their report was ambiguous and unclear at parts and there were no recommendations to the ranges of several parameters, making it difficult to properly reproduce this algorithm. When it was believed that a functional SQN-PTWSVM had been created matrix search was used to iteratively find functional parameter values. With extensive testing the results were still not very satisfying and this method was eventually abandoned. 19 S. Bengtsson Machine learning for mechanical analysis

CONCLUSIONS

The purpose of this thesis was two-fold: Find a feasible anomaly detection method for the moments as measured from the turning of a key component in a mechanical mechanism, and . Find a classifier which can pin-point the flaws in the mechanical mechanism which caused it to produce anomalous measurements. The results from these then has to be compared to the performance of experienced people on the same data. Research led to the consideration of different kinds of Support Vector Machines (SVM) which were investigated, implemented and tested against the performance of Neural Networks. Two methods were implemented to aid these anomaly detectors and classifiers, namely Principal Component Analysis (PCA) and autoencoders. One-Class Support Vector Machines (OC-SVM) with three different kernels - linear, Gaussian and polynomial - were used for anomaly detection. PCA showed great results in decreasing the dimensionality of samples while retaining the information within these dimensions. Autoencoders in conjunction with PCA gave an overall boost in anomaly detection performance for both the Support Vector

Machines and Neural Networks. For classification ordinary SVMs were used together with methods to enable multiclass classification.

REFERENCES

[1] J. Xie, "Optical character recognition based on least square support vector machine," in 2009 Third International Symposium on Intelligent Information Technology Application, vol. 1, Nov 2009, pp. 626–629.

[2] X. Li, X. Dong, J. Lian, Y. Zhang, and J. Yu, "Knockoff filter-based feature selection for discrimination of non-small cell lung cancer in ct image," IET Image Processing, vol. 13, no. 3, pp. 543–548, 2019.

[3] T. Turki, "An empirical study of machine learning algorithms for cancer identification," in 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), March 2018, pp. 1–5.

[4] P. Shimpi, S. Shah, M. Shroff, and A. Godbole, "A machine learning approach for the classification of cardiac arrhythmia," in 2017 International Conference on Computing Methodologies and Communication (ICCMC), July 2017, pp. 603–607.

[5] A. Datta, M. Augustin, N. Gupta, S. Viswamurthy, K. M. Gaddikeri, and R. Sundaram, "Impact localization and severity estimation on composite structure using fiber bragg grating sensors by least square-support vector regression," IEEE Sensors Journal, 2019.

[6] V. Pliuhin, M. Pan, V. Yesina, and M. Sukhonos, "Using azure maching learning cloud technology for electric machines optimization," in 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S T), Oct 2018, pp. 55–58.

[7] D. Silder, J. Schrittwieser, K. Simonyan, and et al, "Mastering the game of go without human knowledge," Nature, vol. 550, p. 354, October 2017.

[8] V. Vapnik and A. Lerner, "Pattern recognition using generalized portraits (translated from russian)," Avtomatika i Telemekhanika, vol. 24, no. 6, pp. 774–780, Dec 1963.

[9] X. Wang and S. Lu, "Improved fuzzy multicategory support vector machines classifier," in 2006 International Conference on Machine Learning and Cybernetics, Aug 2006, pp. 3585– 3589.

[10] G. Yan and S. Fenzhen, "Study on machine learning classifications based on oli images," in Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC). IEEE, 2013, pp. 1472–1476. [11] H. Liu, X. Wang, X. Bi, X. Wang, and J. Zhao, "A multi-feature svm classification of thangka headdress," in 2015 8th International Symposium on Computational Intelligence and Design (ISCID), vol. 2, Dec 2015, pp. 160–163.

[12] Y. Tan and G. Zhang, "The application of machine learning algorithm in underwriting process," in 2005 International Conference on Machine Learning and Cybernetics, vol. 6, Aug 2005, pp. 3523–3527 Vol. 6.

[13] R. Rojas, Neural networks: a systematic introduction. Springer Science & Business Media, 2013.

[14] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back propagating errors," Nature, vol. 323, pp. 533–536, 10 1986.

[15] Y. Li, Y. Fu, H. Li, and S. Zhang, "The improved training algorithm of back propagation neural network with self-adaptive learning rate," in 2009 International Conference on Computational Intelligence and Natural Computing, vol. 1, June 2009, pp. 73–76.