Northumbria University NEWCASTLE

Module	PE7043 – AI and Digital Technology
Module Tutor	Yifeng Zeng
Assignment Title	An Overview of Support Vector Machines, Their Applications, and Their Use Cases in Law Enforcement
Student Name	Scott Cumming
Student Number	21056374
Programme	MSC in Computer Science with Data Analytics
Word Count	3141
Date	23rd June 2023

Contents

1. Introduction

The popularity of Artificial Intelligence (AI) has ebbed and flowed since its inception in the middle of the 20th century (Russell and Norvig, 2022). Currently, it appears to be enjoying a period of relative prosperity following the advent of big data, the advancement of deep learning and introduction of generative AI models such as ChatGPT-4. These recent developments could threaten the future of existing methods like Support Vector Machines (SVMs), which have traditionally been widely utilised in supervised learning.

This paper will therefore review SVMs and consider whether they are still relevant. Section 2 looks at the origins of the method, how it works, what it is used for and why it is deemed to be a significant part of machine learning and AI. Section 3 explores some of the research that has been conducted recently on SVMs, including efforts that have been made to address its main flaws. Theoretical and real-world applications of the method are discussed in section 4, particularly within the fields of neurology, cybersecurity, environmental protection and military intelligence. Finally, section 5 considers how law enforcement agencies could take advantage of SVMs whilst addressing the key ethical issues surrounding the use of AI in this sector.

2. What are SVMs and why are they important in AI?

SVMs were first introduced in the early 1990s by Boser, Guyon and Vapnik (1992) for use in binary classification. Vapnik is widely credited with developing the approach (Joshi, 2020), which is based on work he did on statistical learning theory with Alexei Chervonenkis that began in the 1960s (Han, Kamber and Pei, 2012). Since their introduction, the use of SVMs has been extended to multiclass classification and regression analysis, with the latter often referred to as Support Vector Regression (SVR). They can also be used in other areas, such as novelty detection and density estimation (Zhang, 2010).

SVMs are characterised by their use of margins, hyperplanes and kernels. When used for classification, the classes of data are separated by a hyperplane using the largest possible margin, which is known as the maximum marginal hyperplane (Kapoor, 2019). Soft margins are normally utilised with this type of SVM as they allow misclassifications to occur near the separating hyperplane, which means the models produced are less sensitive to outliers and aren't prone to overfitting (Starmer, 2019).

Figure 1 below (Awad and Khanna, 2015, p.46, fig. 3-4) shows the maximum marginal hyperplane bounded by solid black lines, along which lie the support vectors. Misclassified observations that fall on the wrong side of the dotted boundary line are tolerated and marked with a red X in the figure. These misclassifications – along with any correct predictions that lie close to the dotted boundary line – are penalised using the hinge loss cost function (Nelson, 2023).

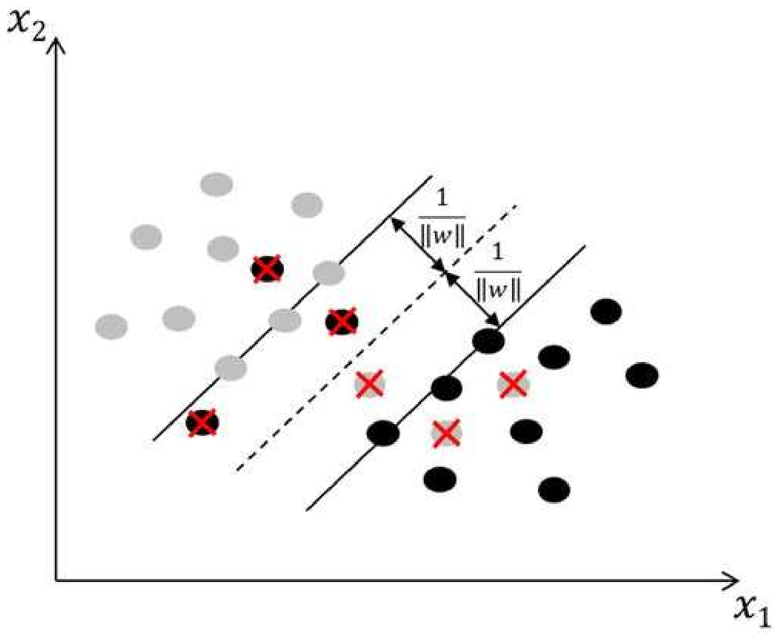

Figure 1 *– A binary classification SVM using soft margins*

When the observations follow a non-linear pattern and cannot be classified, SVMs use a kernel function to separate the data in a higher dimension where linear boundaries can be identified. This function is often called the kernel trick as it calculates these boundaries without transforming the data to the higher dimension, thereby limiting the demand on computational resources (Nelson, 2023).

The methods commonly used to achieve SVM multi-class classification all utilise binary classification and are known as one-vs-one, one-vs-all and Directed Acyclic Graph SVM (DAGSVM). In one-vs-one, a SVM model is generated for each pair of classes, resulting in $K(K-1)/2$ models if there are K classes. In each pair, the positive class is generally assigned +1 and the negative –1; new observations are then tested against all of the models and assigned to the class that receives the most +1 predictions (James *et al.*, 2013). One-vs-all takes a similar approach but compares each class against the combined remaining $K-1$ classes and requires less computation as it generates K models (López, López and Crossa, 2022). DAGSVM mirrors the one-to-one method in the training phase then uses "a rooted binary directed acyclic graph" during testing (Hsu and Lin, 2002, p.416).

When used for regression, SVMs seek to fit the data points into an ε-insensitive band (Bishop, 2006) rather than separate them using a maximum marginal hyperplane. SVR is not as popular as its classification counterpart (Duvvuri and Singhal, 2016) and there is no consensus that it performs better than other forms of regression (López, López and Crossa, 2022), therefore it is not discussed any further in this paper.

SVMs have had a significant impact on the development of AI, specifically the field of machine learning. In the early 2000s, they were the most popular 'off-the-shelf' supervised learning approach for situations where specialised prior knowledge about a domain was lacking (Russell and Norvig, 2022). At that time, the approach was described as being “distinct from the other recent ground-breaking developing [sic] in machine learning” (Andrew, 2000, p.689), with some suggesting it should be considered its own subfield (Cristianini and Shawe-Taylor, 2000). These assertions were supported by early SVMs achieving the highest accuracy when classifying the MNIST benchmark handwriting set (Zhang, 2010) and their superior performance on text categorisation (Joachims, 1998).

SVMs have since been described as “an important pillar of machine learning theory” (Joshi, 2020, p.71). Their popularity has been attributed to a number of factors, including: their high accuracy and effectiveness across multiple applications (Zhang, 2010); their computational efficiency relative to other methods (Steinwart and Christmann, 2008); and “the availability of user-friendly libraries in many languages that are able to implement the SVM method” (López, López and Crossa, 2022, p.338).

3. Recent developments in SVMs

Despite SVMs being considered state-of-the-art in the late noughties (Martens *et al.*, 2008), deep learning networks and random forest models are now preferred for ‘off-the-shelf’ supervised learning (Russell and Norvig, 2022). Equally, generative AI models have recently attracted a lot of attention and business investment following the release of ChatGPT-4 (Garter, 2023), which could have a detrimental impact on the popularity of discriminative AI models like SVMs. Therefore, some may question whether the approach is still relevant.

Due to significant differences in how they work, discriminative AI models will be better suited to some tasks than generative models and vice-versa (Turing, 2023). In addition, there are concerns about the vast amount of energy required to create some generative AI models (Saenko, 2023), so discriminative models may be viewed more favourably if they offer better energy efficiency, particularly given the current climate crisis.

The training time complexity of SVMs is $O(n^2)$, which doesn’t compare well with other supervised learning models (Kumar, 2019) and makes the approach unsuitable for use with large data sets (Cervantes *et al.*, 2008). The resolution of this issue is therefore considered a “major research goal” (Han, Kamber and Pei, 2012, p.415).

A comprehensive review of SVM classification was recently conducted by Cervantes *et al.* (2020), which identified the most common strategies used to improve the handling of large data sets. These strategies include under sampling, the use of parallel algorithms, transforming the Quadratic Programming Problem (QPP) into a simpler problem, adopting a geometric approach, and using alternative optimisation methods. An example of the latter is the Pegasos algorithm, which uses stochastic gradient descent to optimise the SVM objective function. The size of the training set does not directly dictate the run-time of the algorithm, therefore it works well with large data sets (Shalev-Shwartz *et al.*, 2010).

Recently, Pegasos has been adapted by the Stochastic Gradient Twin Support Vector Machine (SGTSVM). As the name suggests, the SGTSVM is based on the

Twin Support Vector Machine (TWSVM), which uses two, nonparallel hyperplanes to classify data points and reduces the QPP to two smaller quadratic problems, rather than a single large one (Tanveer *et al*., 2022). The computational speed of TWSVMs is approximately 4 times faster than a standard SVM and they have increased in popularity in recent years (Huang, Wei and Zhou, 2018), spawning many variants.

Developed by Wang *et al*. (2018), the SGTSVM variant builds on the success of Pegasos while addressing its weaknesses, such as a potential lack of generalisability caused by inadequate sampling of support vectors. It constructs the hyperplanes by sampling two data points at each stochastic iteration, rather than the one used in Pegasos. It also inherits the strengths of the TWSVM, meaning it does not overly depend on particular samples like support vectors and is insensitive to sampling. The data in Appendix B shows that linear and non-linear SGTSVMs learn a lot faster on large data sets than Pegasos and standard SVMs, whilst maintaining a stable performance.

Considerable progress has also been made improving SVMs for online learning, which has proved a challenge for many machine learning algorithms due to the difficulty in handling evolving data streams (Zhou *et al*., 2017). For example, the Online Support Vector Machine (OSVM) created by Wang and Xing (2019) uses an Online Incremental Feature Map (OIM) to maintain representative prototypes for each class. The prototypes constitute a very small part of the original data, so the retraining conducted by the OIM makes OSVMs a lot faster than the best SVMs while achieving comparable accuracy.

Similarly, Sun *et al*. (2018) have advocated the use of Attenuation Incremental SVMs (AISVMs) to classify internet traffic. This approach discards the old training data each time it updates the model with new data, apart from the support vectors which are assigned a weight and retained in the training set until the weighted value falls below a certain level. Table 1 below shows that AISVMs were substantially quicker than standard SVMs when learning the 10 subsets contained in the internet traffic data set provided by the University of Cambridge's Nprobe Project, incurring only a small drop in accuracy.

Model	**Training Time (HH:MM:SS)**	**No of Support Vectors**	**Average Accuracy**
SVM	86:44:51	1071	96.0%
AISVM	6:35:45	901	95.4%

***Table 1** – Performance of SVMs and AISVMs when tested against the internet traffic data set provided by the University of Cambridge's Nprobe Project. Adapted from Sun et al. (2018, p.795).*

Another research area that has shown promise involves the integration of SVMs with deep learning. Some have incorporated them directly into the deep learning architecture, including Tang (2013) who used a linear SVM as the activation function on the output layer of a network. This configuration performed better than the usual softmax function on experiments involving facial recognition and the widely used MNIST and CIFAR-10 data sets. In other cases, features have been extracted from deep neural networks then utilised in SVMs. This approach was employed by Erfani *et al*. (2016) to detect anomalies in high-dimensional spaces, using features drawn from a deep belief network to train a one-class SVM. When compared against a deep autoencoder, the SVM was 3 times faster in training and 1000 times faster in testing, achieving a comparable performance.

Efforts have also been made to enhance the kernel selection and optimisation process, as well as improve the performance of multi-class SVMs. Despite these efforts, further research is required in both areas, with the development of "new heuristic, stochastic or hybrid methods" deemed a desirable aim in the latter (Cervantes *et al.*, 2020, pp.205-206).

4. Applicability of SVMs

SVMs have been employed successfully across a number of domains, ranging from the social to natural sciences (López, López and Crossa, 2022). The table at Appendix C demonstrates their broad applicability and they have performed well in handwritten digit recognition, object recognition, speaker identification and benchmark time-series prediction tests (Han, Kamber and Pei, 2012). Their ability to handle high dimensional data means they have been widely adopted in bioinformatics and natural language processing (Zhang, 2010).

Since their introduction, a lot of research has been conducted on using SVMs as an aid in the diagnosis and prognosis of significant neurological and psychiatric illnesses, such as Alzheimer's disease, schizophrenia and depression (Pisner and Schnyer, 2020; Steardo Jr *et al*., 2020). They are particularly suited to this task as their characterisation at the individual level means they have great potential in a clinical setting, plus their multivariate nature makes them "sensitive to spatially distributed and subtle effects in the brain" (Orrù *et al*., 2012, p.1149). Recently, their popularity in this field has waned as neural networks have become more widely adopted, however they are still recommended for use with general and multimodal data in areas such as dementia classification and amyotrophic lateral sclerosis progression prediction (Myszczynska *et al*., 2020).

Cybersecurity is another area where SVMs have shown considerable potential and applicability, specifically the detection of malicious traffic and network intrusion (Mohammed and Sulaiman, 2012; Kathiresan *et al*., 2023). For example, Ghanem *et al*. (2017) found that a linear two-class SVM achieved 100% accuracy detecting various attack types, although the model was not as good at identifying port scanning attempts and achieved a lower figure of 83.4%. Nevertheless, the results were encouraging, particularly since the model was trained on an unbalanced data set, which often poses problems for SVMs (Batuwita and Palade, 2013). Moreover, their applicability in this area extends to the Internet of Things and modern networks (Koroniotis *et al*., 2019), meaning they are capable of tackling the cybersecurity challenges of the future.

SVMs can also be used to address environmental issues such as water pollution, which affects people all over the world. Studies have shown that they perform reasonably well predicting water quality (Singh, Basant and Gupta, 2011; Kamyab-

Talesh *et al.*, 2019). Furthermore, BBC News (2023) recently reported that AI is being utilised alongside satellite imagery and sensor data to predict water pollution in south-west England, with the ultimate aim of preventing it from occurring. Although details of the algorithm used have not been published, this case, the aforementioned studies and other research on integrating machine learning with remote sensor data and data visualisation techniques (Malek *et al.*, 2019) demonstrate that SVMs could be applied in real-world settings and produce tangible results to help solve environmental problems.

As noted above, SVMs are often employed in natural language processing and several studies have shown they are adept at conducting sentiment analysis (Zainuddin and Selamat, 2014; Devi, Kumar and Prasad, 2016; Bansal, Gupta and Muralidhar, 2019). This type of analysis is typically carried out for the purposes of market research, with social media posts, online reviews and similar content being used to train the algorithms that help businesses better understand how they, their products or their services are perceived. However, it is evident that the technique could also be used by militaries, especially since most of their intelligence is now reported to be derived from publicly available sources (Strobel, 2022). For instance, SVM-enabled sentiment analysis could help military officers ascertain the opinions, emotions and attitudes of a local populace, as well as inform their psychological operations and influence campaigns.

5. Potential SVM use cases in law enforcement

There is huge potential in using SVMs and AI in general to support the work conducted by law enforcement agencies. The individuals working for these agencies are often under-resourced and overworked, therefore adoption of this technology could help them make better decisions and enable more efficient working practices, thus saving time and resources. Equally, it could be argued that the implementation of AI in law enforcement is a necessity given that criminal groups have previously been quick to harness the benefits offered by new technology (United Nations and Interpol, 2019), meaning these agencies risk being outpaced by their adversaries if they do not act.

To date, the use of facial recognition in surveillance applications has attracted a lot of attention, as have predictive policing schemes that aim to identify where crimes may occur or whether an individual will offend (UK Parliament, 2021; NPR, 2023). Understandably, there are numerous concerns around these use cases and AI more widely when employed in a law enforcement context, particularly the potential for bias to become embedded in algorithms and the creation of processes which lack human oversight (Equality and Human Rights Commission, 2022; Reese, 2022). Consequently, if agencies are to reap the benefits of AI and avoid causing injustice, they must address these issues. Interpol (2023) has recognised this and has released an AI toolkit aimed at helping agencies implement the technology in a safe and responsible way.

In practice, the aforementioned issues could be addressed by combining AI with traditional approaches that incorporate adequate checks and balances. For example, some studies have shown that SVMs are effective at identifying propaganda and extremist content online (Miranda *et al*., 2020; Khanday, Khan and Rabani, 2021). The output produced by these models could therefore be used to generate intelligence which is only acted upon if corroborated by other sources. The approach could help agencies identify members of extremist networks, as well as hate crimes and serious threats perpetrated online.

The example above shows how the ability of SVMs in natural language processing could be utilised in law enforcement. This ability could also be put to use in police control rooms to gain insights from call handler data. A recent study found that an AI model was better than human emergency call handlers at identifying stroke cases from transcribed text, an ailment which is often difficult to identify when first reported (European Stroke Organisation Conference, 2023). It is easy to see that SVMs could be used in a similar manner to help police call handlers determine what crimes are potentially occurring and their severity. This could prove particularly useful when the caller is hesitant or cannot speak openly, such as in cases of domestic violence or modern slavery.

Lastly, the proficiency of SVMs at handling high-dimensional data obtained from a relatively small number of samples could be exploited in the investigation of complex offences, like serious fraud and financial crimes. These crimes do not occur as frequently as low-level offences yet when they do, they demand a lot of investigative resources due to their complex nature. As a result, SVMs are better placed than other methods to learn from these limited number of cases and can generate predictive models that maximise the potential of the various dimensions in the data. This use case has been demonstrated in part by Krysovatyy *et al*. (2021), whose SVM model was 99.7% accurate when determining fictitious (and likely fraudulent) businesses from legitimate enterprises.

6. Conclusion

This paper has shown that SVMs form an important part of machine learning and still have a lot to offer since their inception over 30 years ago. They are particularly effective in situations where data is limited but high-dimensional and can achieve excellent results when used in ensemble methods alongside other approaches like deep learning. Moreover, they are getting better at online learning and variants such as the SGTSVM are enabling them to learn a lot quicker when processing big data, which has traditionally been an area of weakness.

It has been demonstrated that SVMs could be employed in a wide variety of applications and have considerable potential in neurological diagnosis, malicious network traffic detection, water pollution identification and military intelligence. Their ability in natural language processing could also be put to good use in law enforcement applications, as long as important ethical concerns are addressed. If so, they could be used to generate intelligence from online sources, help call handlers better interpret incoming crime reports and aid the investigation of complex offences.

Appendix A – Bibliography

Andrew, A.M. (2000) Review of *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, by N. Cristianini and J. Shawe-Taylor. *Robotica*, 18(6), pp.687-689. Available at: https://doi.org/10.1017/S0263574700232827 (Accessed: 20th June 2023).

Awad, M. and Khanna, R. (2015) *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. New York: Apress. Available at: https://doi.org/10.1007/978-1-4302-5990-9 (Accessed: 20th June 2023).

Bansal, A., Gupta, C.L. and Muralidhar, A. (2019) 'A Sentimental Analysis for YouTube Data using Supervised Learning Approach', *International Journal of Engineering and Advanced Technology*, 8(5). Available at: https://www.ijeat.org/portfolio-item/E7756068519/ (Accessed: 16th June 2023).

Batuwita, R. and Palade, V. (2013) 'Class Imbalance Learning Methods for Support Vector Machines', in H. He and Y. Ma (eds) *Imbalanced Learning: Foundations, Algorithms, and Applications*. Hoboken, New Jersey: John Wiley & Sons, pp.83-99. Available at: https://doi.org/10.1002/9781118646106.ch5 (Accessed: 16th June 2023).

BBC News (2023) *AI to Stop Water Pollution Before it Happens*. Available at: https://www.bbc.co.uk/news/science-environment-65913940 (Accessed: 16th June 2023).

Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*. New York: Springer.

Boser, B.E., Guyon, I.M. and Vapnik, V.N. (1992) 'A Training Algorithm for Optimal Margin Classifiers', *COLT92: 5th Annual Workshop on Computational Learning Theory*, Pittsburgh (Pennsylvania, USA), 27-29 June. New York: Association for Computing Machinery, pp.144-152. Available at: https://doi.org/10.1145/130385.130401 (Accessed: 27th May 2023).

Cervantes, J. *et al*. (2008) 'Support Vector Machine Classification for Large Data Sets via Minimum Enclosing Ball Clustering', *Neurocomputing*, 71(4-6), pp.611-619. Available at: https://doi.org/10.1016/j.neucom.2007.07.028 (Accessed: 2nd June 2023).

Cristianini, N. and Shawe-Taylor, J. (2000) *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, UK: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9780511801389 (Accessed: 20th June 2023).

Devi, D.V.N., Kumar, C.K. and Prasad, S. (2016) 'A Feature Based Approach for Sentiment Analysis by Using Support Vector Machine', *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, Bhimavaram (India), 27-28 February. IEEE, pp.3-8. Available at: https://doi.org/10.1109/IACC.2016.11 (Accessed: 16th June 2023).

Duvvuri, S. and Singhal, B. (2016) *Spark for Data Science: Analyze Your Data and Delve Deep into the World of Machine Learning with the Latest Spark Version, 2.0*. Birmingham, UK: Packt Publishing.

Equality and Human Rights Commission (2022) *Artificial Intelligence in Public Services*. Available at: https://www.equalityhumanrights.com/en/advice-and-guidance/artificial-intelligence-public-services (Accessed: 19th June 2023).

Erfani, S.M. *et al*. (2016) High-Dimensional and Large-Scale Anomaly Detection Using a Linear One-Class SVM with Deep Learning, *Pattern Recognition*, 58, pp.121-134. Available at: https://doi.org/10.1016/j.patcog.2016.03.028 (Accessed: 11th June 2023).

European Stroke Organisation Conference (2023) 'AI Tool Outperforms Human Emergency Call Handlers in Identifying Stroke, New Study Shows', 24 May [Press release]. Available at: https://2023.eso-conference.org/wp-content/uploads/2023/05/Press-Release-24.05.2023-1.pdf (Accessed: 20th June 2023).

Garter (2023) 'Gartner Poll Finds 45% of Executives Say ChatGPT Has Prompted an Increase in AI Investment', 3 May [Press release]. Available at: https://www.gartner.com/en/newsroom/press-releases/2023-05-03-gartner-poll-finds-45-percent-of-executives-say-chatgpt-has-prompted-an-increase-in-ai-investment (Accessed: 2nd June 2023).

Han, J., Kamber, M. and Pei, J. (2012) *Data Mining: Concepts and Techniques*. 3rd edn. Boston: Morgan Kaufmann Publishers. Available at: https://doi.org/10.1016/C2009-0-61819-5 (Accessed: 20th June 2023).

Hsu, C-W. and Lin, C-J. (2002) 'A Comparison of Methods for Multiclass Support Vector Machines', *Transactions on Neural Networks*, 13(2), pp. 415-425. Available at: https://ieeexplore.ieee.org/document/991427 (Accessed: 1st June 2023).

Huang, H., Wei, X. and Zhou, Y. (2018) 'Twin Support Vector Machines: A Survey', *Neurocomputing*, 300, pp.34-43. Available at: https://doi.org/10.1016/j.neucom.2018.01.093 (Accessed: 8th June 2023).

Interpol (2023) *Artificial Intelligence Toolkit*. Available at: https://www.interpol.int/en/How-we-work/Innovation/Artificial-Intelligence-Toolkit (Accessed: 19th June 2023).

James, G. *et al*. (2013) *An Introduction to Statistical Learning with Applications in R*. New York: Springer. Available at: https://doi.org/10.1007/978-1-4614-7138-7 (Accessed: 20th June 2023).

Joachims, T. (1998) 'Text categorization with Support Vector Machines: Learning with Many Relevant Features', *ECML-98: 10th European Conference on Machine Learning*, Chemnitz (Germany), 21-23 April. Berlin: Springer, pp.137-142. Available at: https://doi.org/10.1007/BFb0026683 (Accessed: 2nd June 2023).

Joshi, A.V. (2020) *Machine Learning and Artificial Intelligence*. Switzerland: Springer. Available at: https://doi.org/10.1007/978-3-030-26622-6 (Accessed: 20th June 2023).

Kamyab-Talesh, F. *et al*. (2019) 'Prediction of Water Quality Index by Support Vector Machine: A Case Study in the Sefidrud Basin, Northern Iran', *Water Resources*, 46, pp.112-116. Available at: https://doi.org/10.1134/S0097807819010056 (Accessed: 16th June 2023).

Kapoor, A. (2019) *Hands-On Artificial Intelligence for IoT: Expert Machine Learning and Deep Learning Techniques for Developing Smarter IoT Systems*. Birmingham, UK: Packt Publishing.

Kathiresan, V. *et al*. (2023) 'Machine Learning-Based DDoS Attack Detection Using Support Vector Machine', in V. Sarveshwaran, J.L-Z. Chen and D. Pelusi (eds) *Artificial Intelligence and Cyber Security in Industry 4.0*. Singapore: Springer, pp.329-341. Available at: https://doi.org/10.1007/978-981-99-2115-7_15 (Accessed: 2nd June 2023).

Khanday, A.M.U.D., Khan, Q.R. and Rabani, S.T. (2021) 'SVMBPI: Support Vector Machine-Based Propaganda Identification', *International Conference on Cognitive Informatics and Soft Computing*, Balasore (India), 12-13 December 2020. Singapore: Springer, pp.445-455. Available at: https://doi.org/10.1007/978-981-16-1056-1_35 (Accessed: 20th June 2023).

Koroniotis, N. *et al*. (2019) 'Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT dataset', *Future Generation Computer Systems*, 100, pp.779-796. Available at: https://doi.org/10.1016/j.future.2019.05.041 (Accessed: 16th June 2023).

Krysovatyy, A. *et al*. (2021) 'Economic Crime Detection Using Support Vector Machine Classification', *Modern Machine Learning Technology & Data Science Workshop*, Lviv-Shatsk (Ukraine), 5-6 June. CEUR-WS, pp.830-840. Available at: https://ceur-ws.org/Vol-2917/paper46.pdf (Accessed: 20th June 2023).

Kumar, P. (2019) *Computational Complexity of ML Models*. Available at: https://medium.com/analytics-vidhya/time-complexity-of-ml-models-4ec39fad2770 (Accessed: 2nd June 2023).

López, O.A.M., López,A.M. and Crossa, J. (2022) *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Switzerland: Springer. Available at: https://doi.org/10.1007/978-3-030-89010-0 (Accessed: 20th June 2023).

Malek, S. *et al*. (2019) 'Ecosystem Monitoring Through Predictive Modeling', in S. Ranganathan, M. Gribskov, K. Nakai and C. Schönbach (eds) *Encyclopedia of Bioinformatics and Computational Biology: Volume 3*. Amsterdam: Elsevier, pp.1-8. Available at: https://doi.org/10.1016/B978-0-12-809633-8.20060-5 (Accessed: 16th June 2023).

Martens, D. *et al*. (2008) 'Rule Extraction from Support Vector Machines: An Overview of Issues and Application in Credit Scoring', in J. Diederich (ed.) *Rule*

Extraction from Support Vector Machines. Berlin: Springer, pp.33-63. Available at: https://doi.org/10.1007/978-3-540-75390-2_2 (Accessed: 2nd June 2023).

Miranda, E. *et al*. (2020) 'A Study of Radicalism Contents Detection in Twitter: Insights from Support Vector Machine Technique', *2020 International Conference on Information Management and Technology (ICIMTech)*, Bandung (Indonesia), 13-14 August. IEEE. Available at: https://doi.org/10.1109/ICIMTech50083.2020.9211229 (Accessed: 20th June 2023).

Mohammed, M.N. and Sulaiman, N. (2012) 'Intrusion Detection System Based on SVM for WLAN', *Procedia Technology*, 1, pp.313-317. Available at: https://doi.org/10.1016/j.protcy.2012.02.066 (Accessed: 16th June 2023).

Myszczynska, M.A. *et al*. (2020) 'Applications of Machine Learning to Diagnosis and Treatment of Neurodegenerative Diseases', *Nature Reviews Neurology*, 16, pp.440-456. Available at: https://doi.org/10.1038/s41582-020-0377-8 (Accessed: 16th June 2023).

Nayak, J., Naik, B. and Behara, H.S. (2015) 'A Comprehensive Survey on Support Vector Machine in Data Mining Tasks: Applications & Challenges', *International Journal of Database Theory and Application*, 8(1), pp.169-186. Available at: http://dx.doi.org/10.14257/ijdta.2015.8.1.18 (Accessed: 13th June 2023).

Nelson, H. (2023) *Essential Math for AI: Next-Level Mathematics for Efficient and Successful AI Systems*. Sebastopol, California: O'Reilly.

NPR (2023) *Know It All: AI And Police Surveillance*. Available at: https://www.npr.org/2023/02/23/1159084476/know-it-all-ai-and-police-surveillance (Accessed: 19th June 2023).

Orrù, G. *et al*. (2012) 'Using Support Vector Machine to Identify Imaging Biomarkers of Neurological and Psychiatric Disease: A Critical Review', *Neuroscience and Biobehavioral Reviews*, 36(4), pp.1140-1152. Available at: https://doi.org/10.1016/j.neubiorev.2012.01.004 (Accessed: 14th June 2023).

Pisner, D.A. and Schnyer, D.M. (2020) 'Support Vector Machine', in A. Mechelli and S. Vieira (eds) *Machine Learning: Methods and Applications to Brain Disorders*. Academic Press: London, pp.101-121. Available at: https://doi.org/10.1016/B978-0-12-815739-8.00006-7 (Accessed: 14th June 2023).

Reese, H. (2022) *What Happens When Police Use AI to Predict and Prevent Crime?* Available at: https://daily.jstor.org/what-happens-when-police-use-ai-to-predict-and-prevent-crime/ (Accessed: 19th June 2023).

Russell, S.J. and Norvig, P. (2022) *Artificial Intelligence: A Modern Approach*. 4th edn, global edn. Harlow, UK: Pearson.

Saenko, K. (2023) *Is Generative AI Bad for the Environment? A Computer Scientist Explains the Carbon Footprint of ChatGPT and its Cousins*. Available at: https://news.yahoo.com/generative-ai-bad-environment-computer-

122719130.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAACUVlCdwAEHgEBkgsFCTGj39EfBrtU4VyS4zNPg3AWNF2gRgxpxZ3ZYoI4syCaOdGhrjnLSTSHbhye5xZ857nsnZ8LpqMC2795HMNnrPDcSSXEEwgfvg73vmdZtF6B7R-M9aKmYV31GMzfZNHoCG30mx2yAaLWWWALKEa5EafO8J (Accessed: 2nd June 2023).

Shalev-Shwartz, S. *et al*. (2010) 'Pegasos: Primal Estimated Sub-Gradient Solver for SVM', *Mathematical Programming*, 127, pp.3-30. Available at: https://doi.org/10.1007/s10107-010-0420-4 (Accessed: 8th June 2023).

Singh, K.P., Basant, N. and Gupta, S. (2011) 'Support Vector Machines in Water Quality Management', *Analytica Chimica Acta*, 703(2), pp.152-162. Available at: https://doi.org/10.1016/j.aca.2011.07.027 (Accessed: 16th June 2023).

Starmer, J. (2019) *Support Vector Machines Part 1 (of 3): Main Ideas!!!* 30th September 2019. Available at: https://www.youtube.com/watch?v=efR1C6CvhmE (Accessed: 1st June 2023).

Steardo Jr, L. *et al*. (2020) 'Application of Support Vector Machine on fMRI Data as Biomarkers in Schizophrenia Diagnosis: A Systematic Review', *Front Psychiatry*, 11, article number 588. Available at: https://doi.org/10.3389/fpsyt.2020.00588 (Accessed: 16th June 2023).

Steinwart, I. and Christmann, A. (2008) *Support Vector Machines*. New York: Springer. Available at: https://doi.org/10.1007/978-0-387-77242-4 (Accessed: 20th June 2023).

Strobel, W.P. (2022) *Rise of Open-Source Intelligence Tests U.S. Spies*. Available at: https://www.wsj.com/articles/rise-of-open-source-intelligence-tests-u-s-spies-11670710806 (Accessed: 16th June 2023).

Sun, G. *et al*. (2018) 'Internet Traffic Classification Based on Incremental Support Vector Machines', *Mobile Networks and Applications*, 23, pp.789-796. Available at: https://doi.org/10.1007/s11036-018-0999-x (Accessed: 9th June 2023).

Tang, Y. (2013) 'Deep Learning using Linear Support Vector Machines', *International Conference on Machine Learning*, Atlanta (USA), 16-21 June. Available at: https://doi.org/10.48550/arXiv.1306.0239 (Accessed: 11th June 2023).

Tanveer, M. *et al*. (2022) 'Comprehensive Review on Twin Support Vector Machines', *Annals of Operations Research*. Available at: https://doi.org/10.1007/s10479-022-04575-w (Accessed: 8th June 2023).

Turing (2023) *Generative Models vs Discriminative Models: Which One to Choose?* Available at: https://www.turing.com/kb/generative-models-vs-discriminative-models-for-deep-learning (Accessed: 2nd June 2023).

UK Parliament (2021) *AI in Policing and Security*. Available at: https://post.parliament.uk/ai-in-policing-and-security/ (Accessed: 19th June 2023).

United Nations and Interpol (2019) *Artificial Intelligence and Robotics for Law Enforcement*. Available at: https://unicri.it/artificial-intelligence-and-robotics-law-enforcement (Accessed: 19th June 2023).

Wang, X. and Xing, Y. (2019) ‘An Online Support Vector Machine for the Open-Ended Environment’, *Expert Systems with Applications*, 120, pp.72-86. Available at: https://doi.org/10.1016/j.eswa.2018.10.027 (Accessed: 11th June 2023).

Wang, Z. *et al*. (2018) ‘Insensitive Stochastic Gradient Twin Support Vector Machines for Large Scale Problems’, *Information Sciences*, 462, pp.114-131. Available at: https://doi.org/10.1016/j.ins.2018.06.007 (Accessed: 8th June 2023).

Zainuddin, N. and Selamat, A. (2014) ‘Sentiment Analysis Using Support Vector Machine’, *2014 International Conference on Computer, Communications, and Control Technology (I4CT)*, Langkawi Island (Kedah, Malaysia), 2-4 September. IEEE, pp.333-337. Available at: https://doi.org/10.1109/I4CT.2014.6914200 (Accessed: 16th June 2023).

Zhang, X. (2010) ‘Support Vector Machines’, in C. Sammut and G.I. Webb (eds) *Encyclopedia of Machine Learning*. Available at: https://doi.org/10.1007/978-0-387-30164-8 (Accessed: 31st May 2023).

Zhou, L. *et al*. (2017) ‘Machine Learning on Big Data: Opportunities and Challenges’, *Neurocomputing*, 237, pp.350-361. Available at: https://doi.org/10.1016/j.neucom.2017.01.026 (Accessed: 9th June 2023).

Appendix B – SGTSVM learning times and performance

Table B1 – Accuracy of SVM, Pegasos and SGTSVM when tested on large scale data sets, adapted from Wang *et al*. (2018, pp.127-129)

* Ran out of memory

Dataset Name	No of Samples	Dimensions	Ratio (Pos to Neg Class Samples)	Testing Accuracy (%)			
				Liblinear SVM	Pegasos	Linear SGTSVM	Non-Linear SGTSVM
Skin	245,057	3	0.262	84.28	85.39	87.70	85.34
Gashome	928,990	10	0.578	82.57	72.85	76.09	89.13
Susy	5,000,000	18	0.844	78.52	56.44	75.09	68.61
Kddcup	4,898,432	41	0.248	*	96.42	97.45	99.20
Gas	8,386,764	16	0.077	*	50.54	92.45	92.86
Hepmass	10,500,000	28	1.000	*	80.84	81.10	79.59

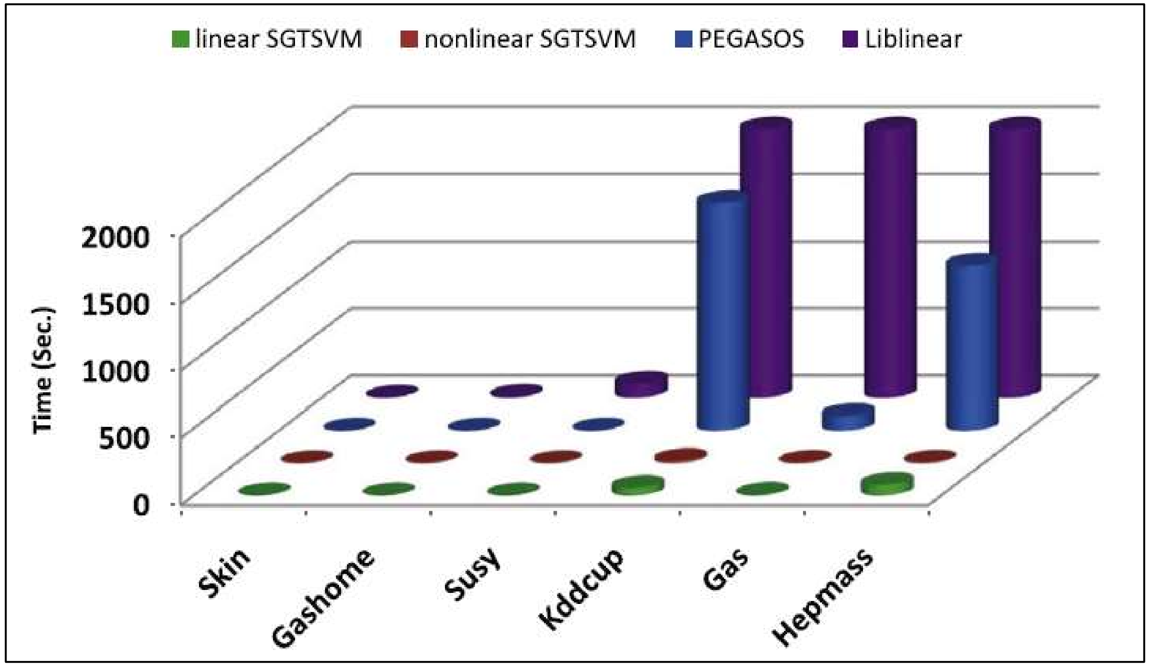

Figure B1 – The learning time of the various algorithms detailed above in Table B1 when used on large scale data sets (Wang *et al*., 2018, p.129)

Appendix C – Various applications of SVM models

Table C1 – List of SVM models and their application in various studies, as identified in the literature review conducted by Nayak, Naik and Behara (2015, p.181)

Model Name	Basis Function Used	Application Area
SVM	Gaussian	Drug Design
SVM	Gaussian	Gene Classification
SVM	Polynomial	Classification
SVM	Gradient Descent	Classification
GSVM	Polynomial	Classification
SVM	Gaussian	Classification
WVSVM	Polynomial	Classification
WC-SVM	Polynomial	Classification
SVM	Polynomial	Image Classification
SVFNN	Polynomial	Pattern Classification
LSSVM	Polynomial	Medical Diagnosis
SVM	Polynomial	Image Classification
SVM	Gaussian	Nuclear Component Classification
OVA SVM	Polynomial	Fingerprint Classification
SVM	Linear, Polynomial, Gaussian	Classification
FSVM	Linear, Polynomial, Gaussian	Product Classification
TSVM	Linear	Pattern Classification
SVM	Linear and Polynomial	Text Classification
SVM	Linear, Polynomial, RBF, Sigmoid	Fault Classification
SVM	Gaussian	TCP Traffic Classification
HGA-SVM	Linear, Polynomial, Gaussian (RBF)	Patent Classification
SS-SVM	Polynomial	Pattern Classification
FSVM	Polynomial	Bankruptcy Prediction
SVM	Linear	Wear degree Classification
SVM	Polynomial	Rule Extraction
LS-SVM	Gaussian	ECG Beat classification
SVM	Gaussian	Power System Classification
SVM	Polynomial	Optimization
CDMTSVM	Quadratic	Classification
SVR-GA/CGA/FA	RBF	Forecasting
SVM	GRPF	Classification
SVM-NN	Polynomial	Text Classification
F2SVM	Sinusoid	Voice Classification
Ɛ-LSSVM	Linear, RBF	Pattern classification
FSVM	Gaussian	Clustering
S-TWSVM	Polynomial	Classification

SVM	Polynomial	Iris Recognition
GASVM	Polynomial, RBF, Sigmoid	Dispute Classification
SVM	Cauchy	Risk Minimization
HS-KCNN-SVM	Polynomial	Text Classification
SVM	Wavelet	System Identification
SVM	Gaussian, Polynomial	Fault Diagnosis
SVR	Gaussian	Reliability Analysis
SVM	Polynomial	Prediction
LSSVM	Least Square	Forecasting
PSVM	Polynomial	Fault Classification
SVM	Polynomial	Signal Analysis