A MODIFIED WEIGHTED SUPPORT VECTOR MACHINE (WSVM) TO REDUCE NOISE DATA IN CLASSIFICATION PROBLEM

SYARIZUL AMRI MOHD DZULKIFLI

A thesis submitted in
fulfillment of the requirement for the award of the
Doctor of Philosophy in Information Technology

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

DECEMBER, 2021

ACKNOWLEDGEMENTS

All praise to the Almighty Allah, for the good health and wellbeing that were necessary to complete this study. I would like to express my sincere gratitude to Assoc. Prof. Dr. Mohd Najib bin Mohd Salleh, my supervisor, for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

I take this opportunity to express gratitude to Faculty of Computer Science and Information Technology (FSKTM) and Centre for Graduate Studies (CGS) of Universiti Tun Hussein Onn Malaysia (UTHM) for providing good facilities and inspiring environment for me to complete this study.

I also place on record, my sense of gratitude to one and all, who directly or indirectly, have lent their helping hands in this study.

Last but not least, deepest appreciation to my parents, Mohd Dzulkifli bin A. Jalil and Shaikhah binti Ahmad, and my siblings, Syahrun Neizam bin Mohd Dzulkifli, Dzunnur Zaily binti Mohd Dzulkifli and Syafiee Syukrie bin Mohd Dzulkifli for their support and give me a lot of encouragement, guidance and motivation.

ABSTRACT

Classification refers to a predictive modeling problem where a class label is predicted for a given example of input data. Data is everywhere and the amount of digital data that exists is growing exponentially. However, data is rarely perfect and there are many inconsistencies that affect data quality such as noise data. Nowadays, the use of SVM is very perspective for the big data classification. SVM provides a global solution for data classification but SVM is highly sensitive to noise data and may not be effective when the level of noise data is high. When noise exists in training data, the decision boundary of SVM would deviate from the optimal hyperplane severely. To overcome SVM drawback for noise data problem, WSVM using KPCM algorithm was used but WSVM using kernel-based learning algorithm such as KPCM algorithm suffer from training complexity, expensive computation time and storage memory when noise data contaminate training data. Thus, through a simple pruning and speed-up method such as clustering method, WKM-SVM has been proposed. However, WKM-SVM has several limitations that are related to *k*-Means Clustering. One of the limitations of WKM-SVM is the clustering centers may not suitably represent original data structures which can potentially cause poor prediction results. Therefore, this research work proposes a modified WSVM utilized with instance selection method and weighted learning to improve WSVM training and classification accuracy. The modification of WSVM will reduce noise data by producing multiple hyperplanes and selecting the optimal hyperplane based on the lowest noise data. The overall result shows that the proposed method outperforms WSVM, OWSVM and WKM-SVM in all datasets in terms of classification accuracy. Specifically, the proposed method produces classification accuracy equal to or higher than 85% for three datasets and lower than 85% for six datasets. However, the performance of the proposed method for test data may not be as good as anticipated since most of the datasets produced classification accuracy lower than 85%.

ABSTRAK

Pengkelasan merujuk kepada masalah pemodelan ramalan yang mana label kelas diramalkan untuk contoh tertentu bagi kemasukan data. Data ada di mana-mana dan jumlah data digital yang wujud telah berkembang dengan pesat. Walau bagaimanapun, data jarang sempurna dan terdapat banyak ketidakseragaman yang mempengaruhi kualiti data seperti hingar data. Masa kini, penggunaan SVM adalah sangat perspektif bagi pengkelasan data besar. SVM menyediakan penyelesaian umum untuk pengkelasan data tetapi SVM amat sensitif terhadap hingar data dan mungkin tidak efektif jika hingar data adalah tinggi. Apabila hingar wujud dalam data latihan, sempadan keputusan SVM akan tersasar jauh dari sempadan optimum. Bagi mengatasi kekurangan SVM dalam masalah hingar data, WSVM yang menggunakan algoritma KPCM telah digunakan tetapi WSVM yang menggunakan algoritma pembelajaran berasaskan *kernel* seperti algoritma KPCM mempunyai masalah dalam latihan, masa pengiraan yang tinggi dan ruang memori apabila hingar data mengubah data latihan. Dengan demikian, melalui kaedah mempercepat dan pengurangan mudah seperti kaedah pengelompokan, WKM-SVM telah dicadangkan. Namun begitu, WKM-SVM mempunyai beberapa kekangan berkaitan dengan *k-Means Clustering*. Salah satu kekangan tersebut adalah pusat pengelompokan tidak sesuai mewakili struktur data asal yang berpotensi menyebabkan keputusan ramalan yang rendah. Lantaran itu, kerja penyelidikan ini mencadangkan agar WSVM diubah suai menggunakan kaedah *instance selection* dan *weighted learning* untuk meningkatkan latihan WSVM dan ketepatan pengkelasan. Pengubahsuaian WSVM akan mengurangkan hingar data melalui penghasilan sempadan keputusan yang pelbagai dan pemilihan sempadan keputusan berdasarkan jumlah hingar data yang rendah. Hasil keseluruhan keputusan menunjukkan bahawa kaedah yang dicadangkan mengatasi WSVM, OWSVM dan WKM-SVM berdasarkan pada ketepatan pengkelasan dalam semua set data. Secara khususnya, kaedah yang dicadangkan memperoleh ketepatan pengkelasan sama atau lebih tinggi dari 85% untuk tiga set data dan lebih rendah dari 85% untuk enam set data. Namun begitu, pencapaian bagi kaedah yang dicadangkan untuk data ujian tidak sebaik yang dijangkakan kerana kebanyakan set data memperoleh ketepatan pengkelasan lebih rendah dari 85%.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ML	-	Machine Learning
AI	-	Artificial Intelligence
w	-	Weight vector
ξ_i	-	Slack variables
b	-	Bias
x	-	Input vector
ρ	-	Margin for a hyperplane
T	-	Transpose
α	-	Lagrange multiplier
S	-	i where $\alpha_i > 0$
C	-	Regularization parameter
h_0/h_1	-	hyperplane
m	-	Distance between hyperplane h_0 and h_1
ω_{ij}	-	Weight function
QP	-	Quadratic Programming
k	-	Number of hyperplanes that will produced depends on the subsets
V	-	Difference slope value between two hyperplanes
R	-	Ranking of the subset
m_i	-	Classification loss
p	-	Orthogonal projection
RBF	-	Radial Basis Function
KKT	-	Karush-Kuhn-Tucker
KPCM	-	Kernel-based Possibilistic C-Means
EP	-	Emerging Patterns
PCA	-	Principal Component Analysis
CSA	-	Cuckoo Search algorithm

BA	-	Bat algorithm
FPA	-	Flower Pollination algorithm
FFA	-	Firefly algorithm
CSISA	-	Cuckoo Search Instance Selection Algorithm
BISA	-	Bat Instance Selection Algorithm
FPISA	-	Flower Pollination Instance Selection Algorithm
FFISA	-	Firefly Instance Selection Algorithm
DT	-	Decision Tree
ANN	-	Artificial Neural Network
KNN	-	*K*-Nearest Neighbors
LUPI	-	Learning Using Privileged Information
NCAR	-	Noise Completely at Random
NAR	-	Noise at Random
NNAR	-	Noise Not at Random
SRM	-	Structural Risk Minimization
SMO	-	Sequential Minimal Optimization
SVM	-	Support Vector Machine
SV	-	Support vector
WSVM	-	Weighted Support Vector Machine
KEEL	-	Knowledge Extraction based on Evolutionary Learning
DASH	-	Difference Average Slope Hyperplane
FPR	-	false positive rate
FNR	-	false negative rate
ROC	-	receiver operating characteristic
AUC	-	area under the ROC curve
TP	-	true positives
TN	-	true negatives
FP	-	false positives
FN	-	false negatives
PPV	-	positive predictive value
TPR	-	true positive rate
PC	-	personal computer
CPU	-	Central Processing Unit

RAM	-	Random Access Memory
OWSVM	-	One-step Weighted Support Vector Machine
IWSVM	-	Iteratively Weighted Support Vector Machine
KM-SVM	-	Support Vector Machine using *k*-Means Clustering
WKM-SVM	-	Weighted Support Vector Machine using *k*-Means Clustering
MHI	-	Multiple Hyperplanes and Instance-weighted
MWSVM-MHI	-	Modified WSVM using Multiple Hyperplanes and Instance-weighted

LIST OF PUBLICATIONS

Journals:

(i) Syarizul Amri Mohd Dzulkifli and Mohd Najib Mohd Salleh (2020). Modified Weighted Support Vector Machine (WSVM) algorithm using Multiple Hyperplanes and Instance-Weighted for class noise. International Conference on Interdisciplinary Computer Science and Engineering 2020 (ICICSE2020).

Proceedings:

(i) Syarizul Amri Mohd Dzulkifli, Mohd Najib Mohd Salleh and Abdul Mutalib Leman (2017). Customer and performance rating in QFD using SVM classification. 3rd Electronic and Green Materials International Conference 2017 (EGM 2017).

(ii) Syarizul Amri Mohd Dzulkifli, Mohd Najib Mohd Salleh and Kashif Hussain Talpur (2019). Improved Weighted Learning Support Vector Machine (SVM) for High Accuracy. 2nd International Conference on Computational Intelligence and Intelligent Systems (CIIS 2019).

(iii) Syarizul Amri Mohd Dzulkifli, Mohd Najib Mohd Salleh and Ida Aryanie Bahrudin (2020). A Comparison of Weighted Support Vector Machine (WSVM), One-Step WSVM (OWSVM) and Iteratively WSVM (IWSVM) for Mislabeled Data. 4th edition of Soft Computing and Data Mining International Conference (SCDM 2020).

CHAPTER 1

INTRODUCTION

1.1 Background of Research

Machine learning (ML) is a continuously developing field, given that some considerations need to be taken into account in working with machine learning methodologies or analysing the impact of machine learning processes (Tagliaferri, 2017). ML applications are highly automated and self-modifying and continue to improve over time with minimal human intervention as they learn with more data. According to a recent study, ML algorithms are expected to replace 25% of global jobs in the next decade (Mathews & Aasim, 2021). The basic objective of ML is to allow computers to automatically learn to recognise complex patterns, make intelligent decisions, and improve performance over time based on the input data. Most often, this involves using a set of historical outcomes to make predictions about future outcomes. ML is also seen as a discipline in artificial intelligence (AI) that consists of designing and developing algorithms. Generally, ML aims to find patterns in data and subsequently use a model that recognizes those patterns in making predictions on new data.

However, ML has its own unique challenges compared to other approaches (Nair, 2017); the first challenge is understanding which processes need automation. Intelligent process automation is about robotic process automation fundamentals combined with ML capabilities to robotize the tasks and learn to perform a job even better (Joshi, 2019). The most straightforward processes to automate are the ones that are performed each day manually with no variable output. The complicated processes require further self-analysis before automation. Though, while ML can help automate some processes, not all automation problems need ML. The second challenge is the lack of good data. Noise in data is a significant concern for many ML techniques used in modeling data (Atla *et al.*, 2011). The solution is to evaluate data through data acquisition, data integration, and data exploration until generating a good dataset. The third challenge is the inadequate infrastructure. For most organisations, managing the various aspects of the infrastructure surrounding ML activities can become a significant challenge (Dean, 2017). The solution to handle ML is to upgrade storage accompanied by hardware acceleration and distributed computing.

The fourth challenge is implementation. Various data driven organizations have spent many years developing successful analytics platforms for implementing ML. Implementing a ML algorithm will provide a deep and practical appreciation for how the algorithm works. This knowledge can also help to internalize the mathematical description of the algorithm (Novikov, 2020). Moreover, integrating newer ML methodologies into existing methodologies is a complicated task. Though maintaining proper interpretation and documentation is an excellent solution to ease the implementation of new methodologies, the last challenge results from the limited skilled resources. With the rapid growth of big data and the availability of programming tools, ML is becoming increasingly popular for data scientists (Mathews & Aasim, 2021). Data scientists often need a combination of domain experience and in-depth knowledge of science, technology, and mathematics. Consequently, the recruitment for data scientists requires companies to pay large salaries since these jobs are often in high demand. This is due to the emergence of big data and how data is being generated and consumed by companies (Das, 2020).

Given these challenges, the second challenge is related to this research as, over the last few decades, noise data has attracted a considerable amount of interest and attention from researchers. The research community has developed several techniques and algorithms to address this issue (Prati *et al.*, 2019). ML can assist people who are frequently susceptible to making mistakes during analyses and trying to establish relationships between multiple features and improve the efficiency of systems and the designs of machines. ML also provides knowledge on making more informed, data driven decisions faster than traditional approaches. Having said that, there are three types of ML: supervised learning, unsupervised learning, and reinforcement learning (Atul, 2019).

In supervised learning, the algorithms are designed to learn by example. When training a supervised learning algorithm, the training data consists of inputs paired with the expected outputs. The training data can accept any type of data as an input, such as values of a database row, the pixels of an image, or an audio frequency histogram. During training, the algorithm searches for patterns in the data that correlate with the expected outputs; then, after training, the algorithm will take in new unseen inputs and determine the label for the new inputs classified based on prior training data. The supervised learning model aims to predict the correct label for newly presented input data (Wilson, 2019a).

Unlike supervised learning, unsupervised learning does not use labeled data but focuses on the data's attributes. Unsupervised learning will frequently find subgroups or detect hidden patterns based on the typical characteristics of the input data within the dataset. In unsupervised learning, the targeted outputs are not subjects of concern as making predictions is not the desired outcome of unsupervised learning algorithms (Wilson, 2019b).

On the other hand, reinforcement learning is considered a hit and trial method of learning. This type of ML is the training of ML models to make a sequence of decisions. To get the machine to do what the programmer wishes, the AI gets either rewards or penalties for the actions. The goal is to maximise the total reward. Moreover, the model has to determine how to perform the task to maximize the reward, beginning from random trials and finishing with advanced tactics (Błażej & Konrad, 2018).

The majority of practical ML uses supervised learning (Brownlee, 2016). There is also no single learning algorithm that works best on all supervised learning problems. A broad range of supervised learning algorithms is available, each with strengths and weaknesses. Supervised learning has been successful in real world applications, divided into two categories: classification and regression (Jaiswal, 2018). Classification predicts discrete values such as true or false and male or female, while regression predicts continuous values such as price, salary, or age. This research focuses on classification because classification is an important technique used in data mining and data analysis applications (Pruengkarn *et al.*, 2015).

In classification, reliability depends on correctly detecting the class label (Sarangam, 2021). Classification refers to a predictive modeling problem where a class label is predicted for a given input data example (Brownlee, 2020). The success of prediction values for the class label aims to measure the overall accuracy, percentage of data for which the class label is correctly predicted. Moreover, classification algorithms have been designed to achieve the maximum possible number of correct class label predictions. In addition, classification seek to predict the target class by analyzing the training data (Priyadarshiny, 2019) and make good predictions on unseen data (Pérez-Ortiz *et al.*, 2016). Attributes and class labels typically characterize the quality of training data, in which the quality of attributes represents how well attributes describe the data for the training purposes.

In contrast, the quality of the class label indicates whether the label of each data is correctly assigned (Nazari *et al.*, 2018). However, having said that, data is rarely perfect, as many inconsistencies affect data quality, such as noise data (Garcia, 2016). Noise data is also considered one of the most challenging classification problems (Farid *et al.*, 2013). Even with extreme efforts to avoid noise, it is challenging to ensure a data acquisition process without errors. Noise data tend to increase the complexity of the classification problem (Napolitano, 2009) within a wide range of research areas. Several studies have concluded that, even in controlled environments, there are at least 5% of noise and errors in a dataset (Maletic & Marcus, 2000). Even though there are various strategies and techniques to manage and deal with noise data, it is often difficult to determine if a given data is indeed noisy or not.

Support Vector Machine (SVM) is a promising and powerful tool for solving practical binary classification problems (Cervantes *et al.*, 2020) and provides a global solution for data classification (Abdiansah & Wardoyo, 2015). One way to learn classification algorithms in the presence of noise data is to correct the labels on the noise data and subsequently to learn the classification algorithm. SVM treats all training data of a given class equally and relies on convex quadratic programming (QP), whose computational complexity is commonly subject to data size. Various studies have indicated that noise data have several consequences, such as significantly reducing the classification accuracy of the classifier (Li *et al.*, 2019), increase in the numbers of necessary training data, increase the classification model building time, alterations in the observed frequencies of the possible classes (Frénay & Kabán, 2014) and increasing the size and interpretability of the classifier (Rani & Rao, 2019). Therefore, this research emphasizes dealing with noise data by using SVM and reducing the high computational complexity of SVM training.

1.2 Problem Statement

Nowadays, the use of SVM is very perspective for the big data classification (Demidova *et al.*, 2016). Training SVM from extremely large and difficult datasets has become an issue given the high training time and memory complexity of SVM training (Nalepa & Kawulok, 2018). SVM requires all training data to be stored in memory during the training when the model's parameters are learned. Once the model parameters are identified, SVM only depends on a subset of training data commonly referred to as support vectors (SV) that lie near the margin. Here, the complexity of the classification task with SVM depends on the number of SV rather than the dimensionality of the input space (Awad & Khanna, 2015). The number of SV retained from the original dataset is data-dependent and varies depending on the complexity of the data, which is captured by the data dimensionality and class. When the data have noise, it is possible that these SV could be construed as noise as well.

Noise data causes decreased performance on SVM given SVM is highly sensitive to noise data (Almasi & Rouhani, 2016) and may not be effective when the level of noise data is high (Li *et al.*, 2013). The performance of SVM can also dramatically decrease with a relatively small number of noise data, which will make the decision boundary deviate from the optimal hyperplane severely (Zhu *et al.*, 2016). Figure 1.1 shows that the noise data influences the decision boundary severely. The thin solid line is the decision plane with no noises, while the bold dotted line is the decision plane with some noises. Circles denote noise data.

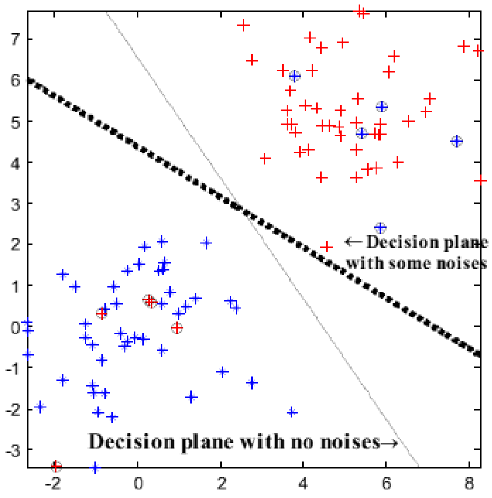

Figure 1.1: Noise data influence the decision boundary severely (Zhu *et al.*, 2016)

In addressing SVM drawback for noise data problem, Yang *et al.* (2007) discovered one of the considered solutions by proposing Weighted SVM (WSVM) using a Kernel-based Possibilistic C-Means (KPCM) algorithm. The KPCM algorithm generates the weights used in WSVM, and these weights will be given to noise data to reduce the effect of noise data as if they do not exist in training data. Indeed, different data have different impacts on the learning of the decision boundary, and the function of weight can make noise data contribute differently. If the data are already associated with the weights, the information can be directly utilized to train the data. As a result, the effect of noise data on the decision boundary is reduced during the training. However, WSVM using kernel-based learning algorithms such as the KPCM algorithm suffer from training complexity, expensive computation time and storage memory when noise data contaminate training data. Nevertheless, it can be reduced through a simple pruning and speed-up method.

Thus, through a simple pruning and speed-up method such as the clustering method, WSVM using *k*-Means Clustering (WKM-SVM) was proposed by Bang & Jhun (2014) and Kim (2016) to reduce noise data. However, WKM-SVM has several limitations related to *k*-Means Clustering. Considering the limitations of WKM-SVM, the intention of scaling down the training data by selecting support vector candidates using a small subset to reduce SVM training time while assigned weight of each noise data for a different penalty of misclassification is considered in this work. The instance selection method is a set of techniques that reduce the quantity of data by selecting a

REFERENCES

Abbasi, Z., & Rahmani, M. (2019). An Instance Selection Algorithm Based on ReliefF. *International Journal on Artificial Intelligence Tools*, *28*(1), 14.

Abdel Maksoud, E. A., Barakat, S., & Elmogy, M. (2019). Medical Images Analysis Based on Multilabel Classification. In *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging* (pp. 209–245). Elsevier Inc.

Abdiansah, A., & Wardoyo, R. (2015). Time Complexity Analysis of Support Vector Machines (SVM) in LibSVM. *International Journal of Computer Applications*, *128*(3), 0975–8887.

Aich, A. (2019). Overfitting and Underfitting With Algorithms in Machine Learning. Retrieved March 20, 2021, from https://www.knowledgehut.com/blog/data-science/overfitting-and-underfitting-in-machine-learning

Akinyelu, A. A., & Adewumi, A. O. (2017). Improved instance selection methods for support vector machine speed optimization. *Security and Communication Networks*, *1*(1), 11.

Alcalá-Fdez, J., Sánchez, L., García, S., Jesus, M. J. del, Ventura, S., Garrell, J. M., … Herrera, F. (2009). KEEL: A software tool to assess evolutionary algorithms for data mining problems. In *Soft Computing* (pp. 307–318).

Almasi, O. N., & Rouhani, M. (2016). Fast and de-noise support vector machine training method based on fuzzy clustering method for large real world datasets. *Turkish Journal of Electrical Engineering and Computer Sciences*, *24*(1), 219–233.

Altidor, W., Khoshgoftaar, T. M., & Van Hulse, J. (2011). Robustness of filter-based feature ranking: A case study. In *Proceedings of the 24th International Florida Artificial Intelligence Research Society, FLAIRS - 24* (pp. 453–458).

Andronicus, A. A. (2017). *Intelligent Instance Selection Techniques for Support Vector Machine Speed Optimization with Application to e-Fraud Detection*. University of KwaZulu-Natal, Durban, South Africa.

Arefi, M. (2018). *Data-Driven Diagnostics of Issues Related to Power System Dynamics Using PMU Measurement*. University of North Carolina.

Arnaiz González, Á. (2018). *Estudio de métodos de selección de instancias*. Universidad De Burgos.

Asiri, S. (2018). Machine Learning Classifiers. Retrieved September 9, 2020, from https://towardsdatascience.com/machine-learning-classifiers-a5cc4e1b0623

Atla, A., Tada, R., Sheng, V., & Singireddy, N. (2011). Sensitivity of different machine learning algorithms to noise. *Journal of Computing Sciences in Colleges*, *26*(5), 96–103.

Atul, H. (2019). Machine Learning Tutorial for Beginners. Retrieved July 22, 2020, from https://www.edureka.co/blog/machine-learning-tutorial/

Awad, M., & Khanna, R. (2015). Support Vector Machines for Classification. In *Efficient Learning Machines* (pp. 39–66). Apress, Berkeley, CA.

Bagchi, T. P. (2014). SVM Classifiers Based On Imperfect Training Data. In *POMS Conference* (pp. 1–7).

Bang, S., & Jhun, M. (2014). Weighted Support Vector Machine Using k-Means Clustering. *Communications in Statistics - Simulation and Computation*, *43*(10), 2307–2324.

Barros De Almeida, M., De Pádua Braga, A., & Braga, J. P. (2000). SVM-KM: Speeding SVMs learning with a priori cluster selection and k-means. In *Brazilian Symposium on Neural Networks, SBRN* (pp. 162–167).

Ben-Hur, A., Ong, C. S., Sonnenburg, S., Schölkopf, B., & Rätsch, G. (2008). Support vector machines and kernels for computational biology. *PLoS Computational Biology*, *4*(10), 1–8.

Bersimis, F. G., & Varlamis, I. (2019). Use of health-related indices and classification methods in medical data. In *Classification Techniques for Medical Image Analysis and Computer Aided Diagnosis* (pp. 31–66). Academic Press.

Bhandari, A. (2020). AUC-ROC Curve in Machine Learning Clearly Explained. Retrieved August 4, 2020, from https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/

Blachnik, M. (2015). Reducing time complexity of SVM model by LVQ data compression. In *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)* (pp. 1–9).

Błażej, O., & Konrad, B. (2018). What is reinforcement learning? The complete guide. Retrieved July 25, 2020, from https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/#:~:text=Reinforcement learning is the training,faces a game-like situation.

Brodley, C. E., & Friedl, M. A. (1996). Improving automated land cover mapping by identifying and eliminating mislabeled observations from training data. In *International Geoscience and Remote Sensing Symposium (IGARSS)* (Vol. 2, pp. 1379–1381).

Brodley, Carla E., & Friedl, M. A. (1996). Identifying and eliminating mislabeled training instances. *Proceedings of the National Conference on Artificial Intelligence*, *1*(1), 799–805.

Brownlee, J. (2014). Feature Selection to Improve Accuracy and Decrease Training Time. Retrieved May 20, 2018, from https://machinelearningmastery.com/feature-selection-to-improve-accuracy-and-decrease-training-time/

Brownlee, J. (2016). Supervised and Unsupervised Machine Learning Algorithms. Retrieved January 21, 2019, from https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

Brownlee, J. (2019). Information Gain and Mutual Information for Machine Learning. Retrieved December 20, 2019, from https://machinelearningmastery.com/information-gain-and-mutual-information/

Brownlee, J. (2020). 4 Types of Classification Tasks in Machine Learning. Retrieved June 7, 2020, from https://machinelearningmastery.com/types-of-classification-in-machine-learning/

Burgos, S. A., & Lorite, F. J. C. (2001). Noisy Data in Data Mining. Retrieved February 10, 2020, from https://sci2s.ugr.es/noisydata

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, *408*(10), 189–215.

Chaudhary, M. (2020). Silhouette Analysis in K-means Clustering. Retrieved September 19, 2021, from https://medium.com/@cmukesh8688/silhouette-analysis-in-k-means-clustering-cefa9a7ad111

Chelliah, I. (2020). An Introduction to Support Vector Machine. Retrieved December 24, 2020, from https://towardsdatascience.com/an-introduction-to-support-vector-machine-3f353241303b

Chen, J., Zhang, C., Xue, X., & Liu, C. L. (2013). Fast instance selection for speeding up support vector machines. *Knowledge-Based Systems*, *45*(8), 1–7.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *20*(3), 273–297.

Das, S. (2020). Reasons, Why There Is A Shortage Of Data Scientists In The Industry. Retrieved August 18, 2020, from https://analyticsindiamag.com/reasons-why-there-is-a-shortage-of-data-scientists-in-the-industry/

De Haro-García, A., & García-Pedrajas, N. (2009). A divide-and-conquer recursive approach for scaling up instance selection algorithms. *Data Mining and Knowledge Discovery*, *18*(3), 392–418.

Dean, J. (2017). 5 Machine Learning Mistakes – and How To Avoid Them. Retrieved November 3, 2020, from https://www.sas.com/en_us/insights/articles/big-data/5-machine-learning-mistakes.html

Delavar, A. G. N., & Jafari, Z. (2016). One Method to Reduce Data Classification Using Weighting Technique in SVM+. *Modern Applied Science*, *10*(9), 1913–1844.

Demidova, L., Nikulchev, E., & Sokolova, Y. (2016). Big Data Classification Using the SVM Classifiers with the Modified Particle Swarm Optimization and the

SVM Ensembles. *International Journal of Advanced Computer Science and Applications*, *7*(5), 294–312.

Ding, H., & Xu, J. (2015). Random Gradient Descent Tree: A Combinatorial Approach for SVM with Outliers. In *Twenty-Ninth AAAI Conference on Artificial Intelligence* (pp. 2561–2567).

Duong, H. T., & Truong Hoang, V. (2019). A Survey on the Multiple Classifier for New Benchmark Dataset of Vietnamese News Classification. In *2019 11th International Conference on Knowledge and Smart Technology, KST 2019* (pp. 23–28).

Eichelberger, R. K., & Sheng, V. S. (2013). An empirical study of reducing multi-class classification methodologies. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 505–519).

Fan, H., & Ramamohanarao, K. (2005). A weighting scheme based on emerging patterns for weighted support vector machines. *IEEE International Conference on Granular Computing*, *2*(2), 435–440.

Fan, R. E., Chen, P. H., & Lin, C. J. (2005). Working set selection using second order information for training support vector machines. *Journal of Machine Learning Research*, *6*(12), 1889–1918.

Farid, D. M., Maruf, G. M., & Rahman, C. M. (2013). A new approach of Boosting using decision tree classifier for classifying noisy data. In *International Conference on Informatics, Electronics and Vision, ICIEV* (pp. 1–4).

Frénay, B, & Kabán, A. (2014). A Comprehensive Introduction to Label Noise. *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 23–25.

Frénay, Benoît, & Verleysen, M. (2014). Classification in the presence of label noise: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(5), 845–869.

Ganegedara, T. (2018). Intuitive Guide to Understanding Decision Trees. Retrieved May 5, 2020, from https://towardsdatascience.com/light-on-math-machine-learning-intuitive-guide-to-understanding-decision-trees-adb2165ccab7

Garcia, L. P. F. (2016). *Noise detection in classification problems*. University of Sao Paulo.

Glasmachers, T., & Igel, C. (2006). Maximum-Gain Working Set Selection for SVMs. *Journal of Machine Learning Research*, *7*, 1437–1466.

Glen, S. (2019). Weighting Factor, Statistical Weight: Definition, Uses. Retrieved April 4, 2020, from https://www.statisticshowto.com/weighting-factor/#:~:text=1.,Weight and the Weighting Factor.&text=It is usually used for,medicine for calculating effective doses.

Gonzalez, W. (2020). Today's AI Solutions Require Quality, Unbiased Training Data. Retrieved October 20, 2020, from https://www.forbes.com/sites/forbesbusinesscouncil/2020/08/26/todays-ai-solutions-require-quality-unbiased-training-data/

Gupta, S., & Gupta, A. (2019). Dealing with noise problem in machine learning data-sets: A systematic review. In *Procedia Computer Science* (pp. 466–474).

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning Data Mining, Inference, and Prediction - Second Edition*. Springer, New York.

Helmenstine, A. M. (2020). What Is the Difference Between Accuracy and Precision? Retrieved December 22, 2020, from https://www.thoughtco.com/difference-between-accuracy-and-precision-609328

Huang, X., Shi, L., & Suykens, J. A. K. (2015). Sequential minimal optimization for SVM with pinball loss. *Neurocomputing*, *149*(3), 1596–1603.

Jadari, S. (2019). *Finding mislabeled data in datasets: A study on finding mislabeled data in datasetsby studying loss function*. Uppsala Universitet.

Jaiswal, S. (2018). Regression vs. Classification in Machine Learning. Retrieved November 13, 2020, from https://www.javatpoint.com/regression-vs-classification-in-machine-learning

Jha, A., Dave, M., & Madan, S. (2019). Comparison of Binary Class and Multi-Class Classifier Using Different Data Mining Classification Techniques. In *ICACM* (pp. 1–10).

Jiang, L., Luo, S., & Li, J. (2012). An approach of household power appliance monitoring based on machine learning. In *Proceedings - 2012 5th International Conference on Intelligent Computation Technology and Automation, ICICTA 2012* (pp. 577–580).

Joshi, N. (2019). Robotic Process Automation Just Got “Intelligent” Thanks to Machine Learning. Retrieved February 2, 2019, from https://www.forbes.com/sites/cognitiveworld/2019/01/29/robotic-process-automation-just-got-intelligent-thanks-to-machine-learning/?sh=1ee7156b52d8

Kao, W., Chung, K., Sun, C., & Lin, C. (2004). Decomposition methods for linear support vector machines. *Neural Computation*, *16*, 1689–1704.

Karasuyama, M., Harada, N., Sugiyama, M., & Takeuchi, I. (2012). Multi-parametric solution-path algorithm for instance-weighted support vector machines. *Machine Learning*, *88*(3), 297–330.

Karthikeyan, T., & Revathy, N. P. (2014). Improved Edge Detection Method by using Weighted Support Vector Machines. *International Journal of Advanced Research in Computer Science*, *5*(6), 255–260.

Kim, S. (2016). Weighted K-means support vector machine for cancer prediction. *SpringerPlus*, *5*(1), 1162.

Lapin, M., Hein, M., & Schiele, B. (2014). Learning using privileged information: SVM+ and weighted SVM. *Neural Networks*, *53*, 95–108.

Lee, H. G., Kim, Y. S., Jeong, C. Y., Han, S. W., & Nam, T. Y. (2005). Multi-Level Objectionable Text Classification Using SVM and Non-Harmful Document Screen. In *The 4th International Conference on Asian Language Processing and Information Technology* (pp. 1–8).

Leyva, E., González, A., & Pérez, R. (2015). Three new instance selection methods based on local sets: A comparative study with several approaches from a bi-objective perspective. *Pattern Recognition*, *48*(4), 1523–1537.

Li, H. X., Yang, J. L., Zhang, G., & Fan, B. (2013). Probabilistic support vector machines for classification of noise affected data. *Information Sciences*, *221*, 60–71.

Li, J., Wong, Y., Zhao, Q., & Kankanhalli, M. S. (2019). Learning to learn from noisy labeled data. In *Proceedings of the IEEE Computer Society Conference on*

Computer Vision and Pattern Recognition (pp. 1–9).

Li, Y., Hu, Z., Cai, Y., & Zhang, W. (2005). Support vector based prototype selection method for nearest neighbor rules. In *Lecture Notes in Computer Science* (pp. 528–535).

Liu, C., Wang, W., Wang, M., Lv, F., & Konan, M. (2017). An efficient instance selection algorithm to reconstruct training set for support vector machine. *Knowledge-Based Systems*, *116*, 58–73.

Liu, H., & Motoda, H. (2001). Data Reduction via Instance Selection. In *Instance Selection and Construction for Data Mining* (pp. 3–20). Springer US.

Loukas, S. (2020). Is your model overfitting? Or maybe underfitting? An example using a neural network. Retrieved September 2, 2020, from https://towardsdatascience.com/is-your-model-overfitting-or-maybe-underfitting-an-example-using-a-neural-network-in-python-4faf155398d2

Low, J. Y. (2020). The Importance of Good Quality Training Data. Retrieved March 10, 2020, from https://blog.supahands.com/2020/01/31/the-importance-of-good-quality-training-data/

Luxburg, U. von, & Schölkopf, B. (2011). *Statistical Learning Theory: Models, Concepts, and Results*. *Handbook of the History of Logic*. Elsevier North Holland.

Lyhyaoui, A., Martínez, M., Mora, I., Vázquez, M., Sancho, J. L., & Figueiras-Vidal, A. R. (1999). Sample selection via clustering to construct support vector-like classifiers. *IEEE Transactions on Neural Networks*, *10*(6), 1474–1481.

Maletic, J., & Marcus, A. (2000). Data Cleansing: Beyond Integrity Analysis. (pp. 1–10).

Martinez, T. R., & Zeng, X. (2014). *US 8, 788, 439 B2*.

Mathews, B., & Aasim, O. (2021). Common Machine Learning Algorithms for Beginners. Retrieved August 16, 2021, from https://www.dezyre.com/article/common-machine-learning-algorithms-for-beginners/202

Matić, N., Guyon, I., Bottou, L., Denker, J., & Vapnik, V. (1992). Computer aided cleaning of large databases for character recognition. In *Proceedings* -

International Conference on Pattern Recognition (Vol. 2, pp. 330–333).

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817.

Mohd Dzulkifli, S. A., Mohd Salleh, M. N., & Leman, A. M. (2017). Customer and performance rating in QFD using SVM classification. In *AIP Conference Proceedings* (pp. 1–8).

Monnappa, A. (2021). Data Science vs. Big Data vs. Data Analytics. Retrieved July 24, 2021, from https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article

Morales, P., Luengo, J., Garcia, L. P. F., Lorena, A. C., de Carvalho, A. C. P. L. F., & Herrera, F. (2017). The Noise Filters R package: Label noise preprocessing in R. *The R Journal*, *9*(1), 219–228.

Mourad, S., Tewfik, A., & Vikalo, H. (2017). Data subset selection for efficient SVM training. In *25th European Signal Processing Conference, EUSIPCO 2017* (pp. 833–837).

Mushtaq, M.-S., & Mellouk, A. (2017). Methodologies for Subjective Video Streaming QoE Assessment. *Quality of Experience Paradigm in Multimedia Services*, 27–57.

Nair, A. (2017). 5 Common Machine Learning Problem and How to Solve Them. Retrieved December 12, 2017, from https://www.provintl.com/blog/5-common-machine-learning-problems-how-to-beat-them

Nalepa, J., & Kawulok, M. (2018). Selecting training sets for support vector machines: a review. In *Artificial Intelligence Review* (pp. 857–900).

Napolitano, A. (2009). *Classification Techniques for Noisy and Imbalanced Data*. Florida Atlantic University.

Nazari, Z., & Kang, D. (2015). Density Based Support Vector Machines for Classification, *4*(4), 69–76.

Nazari, Z., Nazari, M., Danish, M. S. S., & Kang, D. (2018). Evaluation of Class Noise Impact on Performance of Machine Learning Algorithms. *International Journal* *of Computer Science and Network Security*, *18*(8), 148–153.

Nettleton, D. F., Orriols-Puig, A., & Fornells, A. (2010). A study of the effect of different types of noise on the precision of supervised learning techniques. *Artificial Intelligence Review*, *33*(4), 275–306.

Neville, J. (2000). *Iterative Classification*.

Nguyen, M. H., & de la Torre, F. (2010). Optimal feature selection for support vector machines. In *Pattern Recognition* (Vol. 43, pp. 584–591).

Nkoana, R. (2011). *Artificial Neural Network Modelling of Flood Prediction and Early Warning*. University of The Free State Bloemfontein.

Novikov, D. (2020). How to Implement a Machine Learning Algorithm in Code. Retrieved November 2, 2020, from https://resources.experfy.com/ai-ml/how-to-implement-a-machine-learning-algorithm-in-code/

Olson, D. L., & Delen, D. (2008). *Advanced Data Mining Techniques*. Springer Berlin.

Olvera-López, J. A., Carrasco-Ochoa, J. A., Martínez-Trinidad, J. F., & Kittler, J. (2010). A review of instance selection methods. In *Artificial Intelligence Review* (pp. 133–143).

Ortner, A. (2020). Top 10 Binary Classification Algorithms. Retrieved July 23, 2020, from https://medium.com/@alex.ortner.1982/top-10-binary-classification-algorithms-a-beginners-guide-feeacbd7a3e2

Panda, N., Chang, E. Y., & Wu, G. (2006). Concept boundary detection for speeding up SVMs. In *ACM International Conference Proceeding Series* (pp. 681–688).

Parikh, K. S., & Shah, T. P. (2016). Support Vector Machine – A Large Margin Classifier to Diagnose Skin Illnesses. In *Procedia Technology* (Vol. 23, pp. 369–375).

Paulsen, V. I., & Raghupathi, M. (2016). *An introduction to the theory of reproducing kernel Hilbert spaces*. *Cambridge Studies in Advanced Mathematics*. Cambridge University Press.

Pearlman, S. (2019). What is Data Preparation? Retrieved April 4, 2021, from https://www.talend.com/resources/what-is-data-preparation/

Pelletier, C., Valero, S., Inglada, J., Champion, N., Sicre, C. M., & Dedieu, G. (2016).

Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series. *Remote Sensing*, 1–23.

Pérez-Ortiz, M., Jiménez-Fernández, S., Gutiérrez, P. A., Alexandre, E., Hervás-Martínez, C., & Salcedo-Sanz, S. (2016). A review of classification problems and algorithms in renewable energy applications. *Energies*, *9*(8), 1–27.

Platt, J. C. (1999). Fast training of support vector machines using sequential minimal optimization. In *Advances in kernel methods* (pp. 185–208). MIT Press.

Prati, R. C., Luengo, J., & Herrera, F. (2019). Emerging topics and challenges of learning from noisy data in nonstandard classification: a survey beyond binary class noise. In *Knowledge and Information Systems* (Vol. 60, pp. 63–97).

Prem. (2021). A Simple Introduction to the k-Nearest Neighbour (kNN) Algorithm. Retrieved January 12, 2021, from https://www.iunera.com/kraken/fabric/k-nearest-neighbour-knn-algorithm/

Priyadarshiny, U. (2019). Introduction to Classification Algorithms. Retrieved November 26, 2019, from https://dzone.com/articles/introduction-to-classification-algorithms

Pruengkarn, R., Fung, C. C., & Wong, K. W. (2015). Using misclassification data to improve classification performance. In *12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (pp. 1–5).

Qi, B., Zhao, C., Youn, E., & Nansen, C. (2011). Use of weighting algorithms to improve traditional support vector machine based classifications of reflectance data. *Optics Express*, *19*(27), 26816.

Qin, T., Zhang, X. D., Wang, D. S., Liu, T. Y., Lai, W., & Li, H. (2007). Ranking with multiple hyperplanes. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'07* (pp. 279–286).

Rani, S. N., & Rao, S. (2019). Study and Analysis of Noise Effect on Big Data Analytics. *International Journal of Management, Technology and Engineering*, *8*(12), 5841–5850.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. In *Encyclopedia of*

Database Systems (pp. 532–538).

Riquelme, J. C., Aguilar-Ruiz, J. S., & Toro, M. (2003). Finding representative patterns with ordered projections. *Pattern Recognition*, *36*(4), 1009–1018.

Sabzevari, M. (2015). *Ensemble Learning in the Presence of Noise*. Universidad Autonoma de Madrid.

Sáez, J. A., Galar, M., Luengo, J., & Herrera, F. (2014). Analyzing the presence of noise in multi-class problems: Alleviating its influence with the One-vs-One decomposition. *Knowledge and Information Systems*, *38*(1), 179–206.

Sáez, J. A., Luengo, J., & Herrera, F. (2016). Evaluating the classifier behavior with noisy data considering performance and robustness: The Equalized Loss of Accuracy measure. *Neurocomputing*, *176*, 26–35.

Samanta, D. (2018). Support Vector Machine. *IIT Kharagpur*.

Sanz, H., Reverter, F., & Valim, C. (2020). Enhancing SVM for survival data using local invariances and weighting. *BMC Bioinformatics*, *21*(193), 1–20.

Sarangam, A. (2021). Difference between Classification and Prediction in Data Mining - An Easy Guide in Just 3 Points. Retrieved April 11, 2021, from https://www.jigsawacademy.com/blogs/data-science/classification-and-prediction-in-data-mining/

Sarkar, T. (2019). Clustering metrics better than the elbow-method. Retrieved November 7, 2019, from https://towardsdatascience.com/clustering-metrics-better-than-the-elbow-method-6926e1f723a6

Saseendran, A. T., Setia, L., Chhabria, V., Chakraborty, D., & Roy, A. B. (2019). Impact of Noise in Dataset on Machine Learning Algorithms. In *Machine Learning Research* (pp. 0–8).

Segata, N., Blanzieri, E., Delany, S. J., & Cunningham, P. (2010). *Noise reduction for instance-based learning with a local maximal margin approach*. *Journal of Intelligent Information Systems* (Vol. 35).

Sen, P. C., Hajra, M., & Ghosh, M. (2018). Supervised Classification Algorithms in Machine Learning: A Survey and Review. In *Advances in Intelligent Systems and Computing* (pp. 99–111).

Shanab, A. A., Khoshgoftaar, T. M., & Wald, R. (2011). Impact of noise and data sampling on stability of feature selection. In *Proceedings - 10th International Conference on Machine Learning and Applications, ICMLA 2011* (Vol. 1, pp. 172–177).

Sharma, M. (2019). Generalization in Machine Learning for better performance. Retrieved June 27, 2019, from https://mathanrajsharma.medium.com/generalization-in-machine-learning-for-better-performance-51bed74a3820

Sirohi, K. (2019). Support Vector Machine (Detailed Explanation). Retrieved September 11, 2019, from https://towardsdatascience.com/support-vector-machine-support-vector-classifier-maximal-margin-classifier-22648a38ad9c

Solomatine, D. P., & Dulal, K. N. (2003). Model trees as an alternative to neural networks in rainfall—runoff modelling. *Hydrological Sciences Journal*, *48*(3), 399–411.

Stanevski, N., & Tsvetkov, D. (2005). Using Support Vector Machine as a Binary Classifier. In *International Conference on Computer Systems and Technologies* (pp. 1–5).

Sun, R., Luo, Z.-Q., & Ye, Y. (2020). On the Efficiency of Random Permutation for ADMM and Coordinate Descent. *Mathematics of Operations Research*, *45*(1), 233–271.

Tagliaferri, L. (2017). An Introduction to Machine Learning. Retrieved January 10, 2018, from https://www.digitalocean.com/community/tutorials/an-introduction-to-machine-learning

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). Classification: Basic Concepts, and Techniques. In *Introduction to Data Mining* (p. 839).

Tavara, S. (2018). *High-performance computing for support vector machines*. University of Skovde.

Tian, J., Gu, H., Liu, W., & Gao, C. (2011). Robust prediction of protein subcellular localization combining PCA and WSVMs. *Computers in Biology and Medicine*, *41*(8), 648–652.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. *Statistics for*

Engineering and Information Science. Springer New York.

Vapnik, V., & Vashist, A. (2009). A new learning paradigm: Learning using privileged information. *Neural Networks*, *22*(6), 544–557.

Wang, Q., Li, B., & Hu, J. (2009). Human resource selection based on performance classification using weighted support vector machine. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *13*(4), 407–415.

Welch, D. (2017). Applied Data Mining and Statistical Learning - When Data is Linearly Separable. Retrieved August 22, 2019, from https://online.stat.psu.edu/stat508/lesson/10/10.1

Wen, C., Guyer, D. E., & Li, W. (2009). Local feature-based identification and classification for orchard insects. *Biosystems Engineering*, *104*(3), 299–307.

Wilson, A. (2019a). A Brief Introduction to Supervised Learning. Retrieved December 18, 2020, from https://towardsdatascience.com/a-brief-introduction-to-supervised-learning-54a3e3932590

Wilson, A. (2019b). A Brief Introduction to Unsupervised Learning. Retrieved December 18, 2020, from https://towardsdatascience.com/a-brief-introduction-to-unsupervised-learning-20db46445283

Wu, Y., & Liu, Y. (2013). Adaptively Weighted Large Margin Classifiers. *Journal of Computational and Graphical Statistics : A Joint Publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America*, *22*(2), 37–41.

Yang, X. S. (2010). A new metaheuristic Bat-inspired Algorithm. In *Studies in Computational Intelligence* (pp. 65–74).

Yang, X. S. (2012). Flower pollination algorithm for global optimization. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 240–249).

Yang, X. S., & Deb, S. (2009). Cuckoo search via Lévy flights. In *World Congress on Nature and Biologically Inspired Computing, Proceedings* (pp. 210–214).

Yang, Xin-she. (2010). *Nature-Inspired Metaheuristic Algorithms*. *University of Cambridge, United Kingdom*. Luniver Press.

Yang, Xulei, Song, Q., & Wang, Y. (2007). A Weighted Support Vector Machine for Data Classification. *International Journal of Pattern Recognition and Artificial Intelligence*, *21*(5), 961–976.

Yi, K., & Wu, J. (2019). Probabilistic end-to-end noise correction for learning with noisy labels. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 7017–7025).

Zhang, H., & Sun, G. (2002). Optimal reference subset selection for nearest neighbor classification by tabu search. *Pattern Recognition*, *35*(7), 1481–1490.

Zhu, F., Yang, J., Gao, J., & Xu, C. (2016). Extended nearest neighbor chain induced instance-weights for SVMs. *Pattern Recognition*, *60*, 863–874.

Zhu, X., & Wu, X. (2004). Class Noise vs. Attribute Noise: A Quantitative Study. *Artificial Intelligence Review*, *22*(3), 177–210.

VITA

The author was born in March 4, 1981, in Johor, Malaysia. He went to Sekolah Menengah Kebangsaan Tun Sardon, Rengit, Batu Pahat, Johor, Malaysia for his secondary school. He pursued his degree at the Universiti Teknologi Malaysia, Johor, Malaysia, and graduated with the Bachelor of Science in Computer in 2004. Upon graduation, he worked as project coordinator at Auspac Corporation Sdn. Bhd., Johor, Malaysia. He also worked as research assistant and part time lecturer at Universiti Tun Hussein Onn Malaysia, Johor, Malaysia. He then enrolled at the Universiti Tun Hussein Onn Malaysia, Johor, Malaysia, in 2005, where he was awarded the Master of Science in Information Technology (Management) in 2007. Thereafter, he taught Database and Information Technology Skills as contract lecturer in the Information Technology Department, Centre for Diploma Studies at Universiti Tun Hussein Onn Malaysia, Johor, Malaysia. In March 2016, Mr. Syarizul Amri bin Mohd Dzulkifli enrolled for a Ph.D program in Information Technology at Universiti Tun Hussein Onn Malaysia, Johor, Malaysia. During this time, he was the main author of four papers related to his Ph.D research.