

Systematic review of event prediction systems

Fakhrzy Shakirin bin Jamaludin, R Kanesaraj Ramasamy

Faculty of Computing and Informatics (FCI), Multimedia University (MMU), Malaysia

Article Info

Article history:

Received Dec 19, 2019

Revised Feb 2, 2020

Accepted Mar 25, 2020

Keywords:

Algorithm

Event

Hybrid

Prediction

Single

ABSTRACT

In this paper, matters that were discussed were the algorithms used for the event prediction systems, the type of algorithms and the comparison between different types of algorithm. Two main that was discussed and reviewed in the paper were the single machine learning algorithm and the hybrid algorithm, a combination of two or more algorithms, including but not limited to optimization algorithms and machine learning. One of the main observations from the reviewed papers were that hybrid algorithms shows more robustness in handling a large dataset compared to a single algorithm.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

R Kanesaraj Ramasamy,
Faculty of Computing and Informatics,
Multimedia University, Malaysia.
Email: kanes87@gmail.com, r.kanesaraj@mmu.edu.my

1. INTRODUCTION

Predicting an event is something that would be most appreciated by many industries. The ability to know beforehand about the possibility of an event occurring can be an advantage to them. That is one of the reasons that why many companies are starting to pursue this event prediction systems. For example, an airline company can boost their revenue if they knew the likelihood of an event that would cause certain countries to have more traffic at a certain time of the year [1-2]. A prediction system needs to be robust and reliable in most situations or it would not hold much significance.

Generally, a prediction system can be produced using algorithms such as naïve Bayes, decision tree and neural network. Those are some of the more popular algorithms used for these systems. Each with their own strength and weakness. They are utilized based on the purpose of their predictions. This is because there is no absolute algorithm where it is useful in every scenario. Performance varies with the situation of the data available to it We will cover the common algorithms used in the events prediction domain. There are many similarities that can be seen in all those papers. One of it is the choice of algorithms used in the prediction phase.

The motivation for this paper is due to the lack of papers in the area of event prediction area and the increase of different type of data available in recent times [2-5]. In this paper, matters to be discussed and reviewed are the works done by others in the aspect of event prediction system. What motivates the research and the results that they achieved using their methods. By the end of the review, we will take a look whether currently a hybrid algorithm is better than a single algorithm for an event prediction system.

2. ALGORITHMS COMMONLY USED IN PREDICTION MODELS

In this first section, we will look into the commonly used algorithms in prediction models. There are a variety of algorithms has been used by others and this part is only to give a brief introduction to the commonly used ones. First one is the decision tree algorithm, a simple but a very popular algorithm because of its ability to

give out good results and enabling transparency in the process. The algorithm can be interpreted easily and to some that is a huge factor to be considered. There are some disadvantages of the algorithm, which is the memory usage and overfitting problem. A more recent decision tree algorithm has been developed and shows an improvement from the previous version namely in efficiency in memory usage and reduced error pruning [6].

The next algorithm is the neural network algorithm or artificial neural network. It is an algorithm that is very reliable and can produce good results even when handling complex type of data. There are a number of versions of neural network [7], but the usually they are consisting of three layers which is the input layer, hidden layer and output layer. The outcome of the algorithm depends on the weight and nodes of the neural network. Although it is very reliable, some might not choose this method as it is considered as a black-box approach due to it being not interpretable.

Naïve Bayes is the other algorithm used in event prediction systems. Based on the Bayes theorem, it is used when the input has a high dimension [8]. It is labeled as naïve because the algorithm assumes all variables contributes towards classification and are mutually correlated. The advantage of using this method is that it only requires short computational time for training it and has good performances but it is less accurate compared to other methods on certain datasets as the Bayes theorem does not work in every case. Besides those three algorithms, there are other notable ones such as the Support Vector Machine (SVM) and Hidden Markov Model (HMM) [9-11].

3. REVIEW OF RELATED WORKS

The type of events discussed in the paper are broad in category and does not limited to mega-events [6]. There are various works that discussed about the event recognition system, how it categorizes the events and the possible definition of events [13-15]. This paper will be discussing the events by the definition of something took place.

In this section, the related works are discussed and summarized to gain insight on the methods used in producing the event prediction system. It would be divided into two parts where the first part is the review of papers that uses only one type of algorithm in their prediction system and the second part would be where the researchers are using more than one type of algorithms.

3.1. Single algorithms

A paper speaks about a fundamental flaw in classic forecasting models that uses time series analysis [16]. The flaw is that it requires a complete observation for a full period of a time series making it unable to give out a good prediction on an early basis. They seek to create a model that enables them to predict an event at an early stage based on online social networks data which has a lot of noise and complex. The model proposed by them has to be able to give out a high accuracy prediction with limited and high noise data at an early stage and the time complexity should be within an acceptable range. They came up with a model based on the Bayesian Network called Bayesian perspective Early stage Event Prediction model (BEEP). It is a different approach than a normal Bayesian network in the sense that it is an improved version of naïve Bayes classifier but looser constraints. The dataset used is from tweets. The results shown that their method is better at an early stage than other baseline methods, SVM (Support Vector Machine) and GNB (Gaussian Naïve Bayes) but as time passes by the accuracy drops but still better than other methods that was tested with. This paper touches the ability of Bayesian classifier in prediction ability when source of data is limited. Besides the flaw of having the need of complete observation, time series analysis has other factors that hinders its ability to forecast. The time series data offers some challenges that needs to be considered when using it [17]. The work mentions about the time complexity, privacy of data, noise in data and the dimension of time series data. There are steps that can be taken to handle these challenges in order to create a good forecasting model based on time series analysis. The work has also listed the other works done in various fields that utilizes the time series analysis and it covers a wide range of scope [18].

In [19] researches the way of helping farmers in managing their agriculture. Global warming has caused many unpredictable weathers to occur and that has troubled the farmers in managing their farms. This research seeks to develop a decision support system for agriculture management where they use time series analysis. This method however has the flaw as [16] had mentioned in their paper. It requires a complete set of data in a period to make a good prediction. Although their model can predict with the lack of data, it would be considered incomplete and inaccurate as the algorithm depends on that update from the user. In [20-21] and [8] applies decision tree in their method for the forecasting part. Among them, [20] shows the least successful result among them this is because there has not been a significant increase of accuracy despite using many factors and a good amount of data size. However, this is not seen in the study by [21] and [8]. One of the reasons is that the factors chosen by the researchers may not matter much thus does not contribute to the accuracy of the results or because of the domain knowledge of the problem. The paper

by [21] and [8] clearly shows an expertise in those area as they were able to determine the important factors to include in the tree and in addition to that they also employ techniques such as finding the information gain to help decide which factor to consider. A precaution measure to avoid overfitting was also used, pruning. This helps the researchers to maintain the accuracy of the model and avoiding the model from overfitting the training data.

In a paper [22], they are creating a forecasting model based on wavelet transform and back propagation neural network but the authors has also mention about the comparison of algorithms in their paper. To sum it up, a single algorithm is dependent on the situation of the problem and usually they can handle a small sample set data if it were to have a lower time complexity. For example, a neural network algorithm can handle a large and complex dataset but the time complexity would be high and for most cases that is not the ideal solution. [16] has mention the weakness of time series analysis and naïve Bayes classifier has the limitation of needing a large number of data to obtain a good result [23].

A paper predicting the life events of a person based on the user's tweet history uses SVM as their prediction model [24]. They mined the data needed from Twitter's API and they sorted out the data based on the life event that they have defined. They look for keywords from the users to identify the life events that the users might be going through. They did not utilize time series feature in their research and they believe that may have cause inaccurate results for a shorter period of time as their results have the best accuracy in six months duration.

We have looked into the models by others that only uses an algorithm in their research. Based on those findings we can infer that it is plausible to use a single algorithm as long as the algorithm chosen for the problem is suitable because some problems are only suitable to be solved using certain algorithms. For example, a neural network algorithm may find a longer time to compute in computing a problem as compared to Bayesian network which performs faster but may give lower accuracy result. The user needs to be able to discern the main problem to be solved.

3.2. Hybrid algorithms

The following part touches about the models that utilized hybrid algorithms which is a combination of two or more algorithms in the prediction model. A paper tested the power of two algorithms [25] with the same dataset. In addition to simply testing the prediction capabilities, the authors have also tested the algorithm when they are used together with K-means clustering algorithm. The K-means are applied to the dataset and the result is fed to the prediction algorithms. The result was that while without the help of K-means, both of the algorithms perform on an average basis but with K-means, the results were increased significantly. This however does not prove that K-means will work on all other algorithm such as it works on the decision tree and neural network as the author has not try this in their experiment.

K-means is one of the clustering algorithms that can help to cluster the data into similar patterns or class and it is also used in numerous cases as the combination of algorithms in prediction model [26, 27]. In [28] is one of the examples that does so. Apart from K-means, genetic algorithm is one other frequently used algorithm which was used for parameter optimization [29-33]. Genetic algorithm optimizes parameters that will be input to the prediction algorithms. By optimizing the parameters, the prediction models are able to perform better and faster [34-37].

In [29] demonstrates this example where they are trying to create a model to predict the stress of an Anchor bolt. They created a model, a back propagation neural network with genetic algorithm as the parameter optimizer. To prove that their model performs better, they have also tested it with a similar neural network but without the addition of genetic algorithm to optimize the parameters. It is shown that with the help of genetic algorithm, it obtains a better accuracy. In [31] and [32] also combines genetic algorithm with neural network for their prediction model. The former predicts train arrival time and the latter predicts aircraft engine wear. These are two different event scenarios but the main purpose is still the same, which is to predict the event beforehand. As the work mentioned before, the genetic algorithm plays an important role in these cases where it optimizes the parameters for the neural network algorithm [29]. In a previous work done, they are able to solve the problem of low speed in local optimization and convergence of backpropagation neural network [23] and another one was able to achieve high accuracy for their prediction with the deployment of genetic algorithm [31].

There have also been studies where they combine with 2 different algorithms that does not use the help of k-means or genetic algorithm. One uses the hybrid algorithm that combines regression analysis with time series analysis [38]. Both are which an important prediction model. They combine them to gain the advantages of each model for their work. Another study combines decision tree with neural network [39]. They also believe in the merit of combining these 2 algorithms together. They wished to cover up the weakness of a single algorithm for prediction, which they use neural network first and discover that although error was small, it had high time complexity making it less efficient. The combination of those two makes the model to become more efficient and faster even for a larger dataset which the neural network has high time

complexity. The accuracy obtained was slightly lower than a single algorithm, only differs by 1%-2% but the average time taken for the combination was improved drastically. Taking that into consideration, the combination algorithm is a better choice in that case.

A recommendation system for Facebook public events uses the combination of Random Forest and Logistic Regression [40]. The result shows effectiveness but still unclear on the application on a wide scale basis as the experiment was done only in a selected city but the main reason for using the method of the two algorithms is to reduce overfitting in the model. The research also talks about the possibility of an individual model to have a poorer performance compared to a model with a combination of algorithms. They believe that it may be caused by the data distribution or type of the problem being addressed. Another research was done previously to create a method to form an ensemble [41]. A decision tree that was used to predict a flight delay with the help of gradient boosting algorithm also gives out a good accuracy [17]. They decided to use that method as it has been proven that the decision tree with gradient boosting are able to handle regression task quite well [42]. The gradient boosting algorithm reduces the error made by the prediction model allowing a better overall accuracy. It was tested with 70 airports and shown a good result.

A combination of neural network and genetic algorithm has been shown effectiveness from various works. Another work uses the genetic algorithm in conjunction with SVM [43]. The work seeks to predict the event of bankruptcy using those algorithms. Similar to hybrid model that uses genetic algorithm, this work also utilizes it to optimize the parameter for the use of SVM. The result shows that when compared to SVM only model, it performs better showing the effectiveness of genetic algorithm in boosting the model's accuracy.

A hybrid approach in predicting crimes also have been made [44]. The approach is unlike the previous methods discussed as it involves semantic analysis. The main idea is the same though, the idea is to combine two different approaches in order to get a much better result as more and more type of data is available compared to 15 years back.

A hybrid algorithm is not just the combinations between two different algorithms. It should seek to maximize the abilities of both algorithms to its full potential. As such, [39] has talked about the importance of choosing which algorithms for the combinations. Some combination may prove to be more useful in certain cases than the other. The metrics used in assessing the machine learning model should also follow the correct standard, statistical approaches [45-50]. There can be multiple metrics that can be used to evaluate the model that is built. A standardized metric should be used so that no bias is involved in evaluating the models whether single or a hybrid model.

3.3. Summary of related works

Table 1 is the summary of the algorithms used in the related works section; the tick sign indicates the algorithm used in their paper. The paper that uses hybrid algorithm have more than one tick indicating the combination of the algorithms. There are some papers not included in the table of summary because the papers which were excluded are not prediction systems in event but are used to justify the argument made. In the others column, the algorithm can be Support Vector Machine (SVM), K-means, Regression Analysis or even algorithms specifically made to cater their specific problem.

Table 1. The summary of algorithms used in related works

Related Works	Algorithm(s) Used					
	Time Series Analysis	Neural Network	Genetic Algorithm	Decision Tree	Bayes	Others
[16]					✓	
[19]	✓					
[20]				✓		
[21]				✓		
[29]		✓	✓			
[38]	✓					✓
[8]				✓		
[22]		✓				
[16]				✓		✓
[30]			✓			✓
[31]		✓	✓			
[32]		✓	✓			
[39]		✓		✓		
[40]				✓		✓
[17]				✓		✓
[42]				✓		✓
[24]						✓
[43]			✓			✓

4. COMPARISON BETWEEN SINGLE AND HYBRID ALGORITHMS

In this section, we will look into the summarization of the previous part where in the event prediction systems, researchers have started to apply hybrid algorithms and the benefits of doing so. Apart from that, we would also look into the benefit of using a single algorithm. Table 2 is the generalization of comparison between the two types of algorithms used in the prediction model. The table is also created with in mind that they are in the scenario where a large dataset is used. Given a large dataset, a single algorithm can have a high time complexity, accuracy can be high but depends on the situation, efficiency is low as the time complexity is high and accuracy is not that good but it is easy to integrate into the prediction model. Less steps are needed to apply the model into the prediction model. As for hybrid algorithms, for a large dataset, it is very suitable as they can have high accuracy with a low time complexity but unlike a single algorithm, it is more complicated and more steps are needed to integrate the algorithm into the model.

A smaller dataset or a smaller scale project it is better to use a single algorithm for the prediction model. However, it is not really the case when the project expands to become bigger because a single algorithm has their own limitations. For example, the model may be overfitted into the training dataset and thus performance overall is affected. There are measures to prevent this such as more thorough data pre-processing, pruning and etc. A hybrid algorithm tackles the problem by enabling the benefits of the algorithms to be used properly. Given a decision tree and neural network [30], they complement each other and as a result, the model has the strength of both neural network and decision tree. Proper steps are needed for the deployment of model with hybrid algorithm as not all combination can work perfectly.

Table 2. The Comparison between single and hybrid algorithm

	Single	Hybrid
Time Complexity	High given a large dataset	Low event with large dataset
Accuracy	Average	High
Efficiency	Low	High
Implementation	Simple	Complex

5. PROPOSED MODEL FOR EFFECTIVE EVENT PREDICTION

The model proposed for event prediction system is one that is widely used in the various domains of event prediction which is the Genetic Algorithm – Neural Network (GA-NN). Proven to be efficient in other domains, this model shows potential in producing a good result in the event prediction. The goal of the event prediction system that uses this model is to predict what type of event will most likely be successful if the event is organized. Given the dataset contains information such as event participation, the review of the event, type of event, duration of the event and the date of the events held in the past. An event would be considered successful if the event has a participation rate of above 70% and has good reviews from the participant. A good review is when the average rating given by the participant is above 60% where the participants are required to rate the event from a scale of 1 to 5.

Genetic Algorithm will be used as the neural network parameter optimization as it has been shown effectiveness in boosting the performance of prediction model. The parameter is to undergo the genetic algorithm process of evaluation with fitness function, selection, crossover and mutation. The neural network to be used in the proposed model is the back propagation neural network which allows the model to back propagate and readjust the weights accordingly until the error produced is in the acceptable threshold. For the activation function of neural network, ReLU (Rectified Linear Unit) will be used. It takes less time to compute due to the simple math and it converges faster making it a viable choice.

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$$

Formula 1. ReLU

6. CONCLUSION

Based on the papers discussed and the results they achieved, an event prediction system is feasible both by a single algorithm or even a hybrid algorithm. It all depends on the application of the model, if it involves a huge amount of data then it is better to deploy a model with a hybrid algorithm due to its ability to handle large dataset and robustness. For a smaller scale project, a single algorithm is more suitable as it can give out an adequate accuracy with easier implementation. Although hybrid algorithm offers a more robust and efficient model, the factor that needs to be considered when using it is the suitable combination for the purpose of forecasting as certain combination are more suitable for it than the other.

As such, a new possible research field is to identify the combination between algorithms for event predictions. The suggested research might be included but not limited to the type of events we are trying to predict with various combinations of algorithms that returns the highest accuracy. There might be a correlation between the type of algorithms used and what kind of events are predicted.

ACKNOWLEDGEMENTS

This research is supported by TM R&D (RDTC/180969), the data was provided by Girl Guides Association Malaysia which was used validate the model.

REFERENCES

- [1] Dietterich, T. G., & Michalski, R. S. "Discovering patterns in sequences of events". *Artificial Intelligence*, vol. 25, no. 2, pp. 187-232, 1985.
- [2] Letham, B., Rudin, C. & Madigan, D. "Sequential event prediction". *Mach Learn* 93, pp. 357-380, 2013.
- [3] Russom, P. "Big data analytics". *TDWI best practices report, fourth quarter*, vol. 19, no. 4, pp. 1-34, 2011.
- [4] Raghupathi, W., & Raghupathi, V. "Big data analytics in healthcare: promise and potential". *Health information science and systems*, vol. 2, no. 1, p. 3, 2014.
- [5] Kambatla, K., Kollias, G., Kumar, V., & Grama, A. "Trends in big data analytics". *Journal of Parallel and Distributed Computing*, vol. 74, no. 7, pp. 2561-2573, 2014.
- [6] Pandya, R., & Pandya, J. C5. "0 algorithm to improved decision tree with feature selection and reduced error pruning". *International Journal of Computer Applications*, vol. 117, no. 16, pp. 18-21, 2015.
- [7] Nwulu, N. I. "A decision trees approach to oil price prediction". In *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, IEEE, pp. 1-5, September 2017.
- [8] Trafalis, T. B., & Ince, H. "Support vector machine for regression and applications to financial forecasting". In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, IEEE, vol. 6, pp. 348-353, July 2000.
- [9] Meyer, D., Leisch, F., & Hornik, K. "The support vector machine under test". *Neurocomputing*, vol. 55, no. 1-2, pp. 169-186, 2003.
- [10] Beal, M. J., Ghahramani, Z., & Rasmussen, C. E. "The infinite hidden Markov model". In *Advances in neural information processing systems*, pp. 577-584, 2002.
- [11] Müller, M. "What makes an event a mega-event? Definitions and sizes". *Leisure studies*, vol. 34, no. 6, pp. 627-642, 2015.
- [12] Chaudet, H. "Extending the event calculus for tracking epidemic spread". *Artificial Intelligence in Medicine*, vol. 38, no. 2, pp. 137-156, 2006.
- [13] Katzouris, N., Artikis, A., & Paliouras, G. "Online learning of event definitions". *Theory and Practice of Logic Programming*, vol. 16, no. 5-6, pp. 817-833, 2016.
- [14] Katzouris, N., Artikis, A., & Paliouras, G. "Incremental learning of event definitions with inductive logic programming". *Machine Learning*, vol. 100, no. 2-3, pp. 555-585, 2015.
- [15] Ma, X., Gao, X., & Chen, G. "Beep: A bayesian perspective early stage event prediction model for online social networks". In *2017 IEEE International Conference on Data Mining (ICDM)*, IEEE, pp. 973-978, November 2017.
- [16] Manna, S., Biswas, S., Kundu, R., Rakshit, S., Gupta, P., & Barman, S. "A statistical approach to predict flight delay using gradient boosted decision tree". In *2017 International Conference on Computational Intelligence in Data Science (ICCIDIS)*, IEEE, pp. 1-5, June 2017.
- [17] S. M. Molaei and M. R. Keyvanpour, "An analytical review for event prediction system on time series," *2015 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA)*, Rasht, pp. 1-6, 2015.
- [18] Panda, S., Sahu, S., Jena, P., & Chattopadhyay, S. "Comparing fuzzy-C means and K-means clustering techniques: a comprehensive study". In *Advances in computer science, engineering & applications*, Springer, Berlin, Heidelberg, pp. 451-460, 2012.
- [19] Karypis, M. S. G., Kumar, V., & Steinbach, M. "A comparison of document clustering techniques". In *TextMining Workshop at KDD2000*, May 2000.
- [20] Morissette, L., & Chartier, S. "The k-means clustering technique: General considerations and implementation in Mathematica". *Tutorials in Quantitative Methods for Psychology*, vol. 9, no. 1, pp. 15-24, 2013.
- [21] Animas, M. I., Byun, Y. C., Gerardo, B. D., & Concepcion, M. B. "Decision support system for agricultural management using prediction algorithm". In *2013 IEEE/ACIS 12th International Conference on Computer and Information Science (ICIS)*, IEEE, pp. 191-194, June 2013.
- [22] Tang, X., Liu, Z., Li, T., Wu, W., & Wei, Z. "The Application of Decision Tree in the Prediction of Winning Team". In *2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*, IEEE, pp. 239-242, August 2018.
- [23] Guo, J., Liu, H., Luan, Y., & Wu, Y. "Application of Birth Defect Prediction Model Based on C5. 0 Decision Tree Algorithm". In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, IEEE, pp. 1867-1871, July 2018.
- [24] Hu, W., Tao, Z., Guo, D., & Pan, Z. "Natural gas prediction model based on wavelet transform and BP neural network". In *2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC) IEEE*, pp. 952-955, May 2018.

- [25] Jadhav, S. D., & Channe, H. P. "Comparative study of K-NN, naive Bayes and decision tree classification techniques". *International Journal of Science and Research (IJSR)*, vol. 5, no. 1, pp. 1842-1845, 2016.
- [26] Abe, S., Shirakawa, M., Nakamura, T., Hara, T., Ikeda, K., & Hoashi, K. "Predicting the Occurrence of Life Events from User's Tweet History". In *2018 IEEE 12th International Conference on Semantic Computing (ICSC), IEEE*, pp. 219-226, 2018.
- [27] Lajevardi, S. B., & Minaei-Bidgoli, B. "Comparison between ANN and Decision Tree in Aerology Event Prediction". In *2008 International Conference on Advanced Computer Theory and Engineering, IEEE*, pp. 533-537, December 2008.
- [28] Xing, H., Sun, X., Wang, M., & Zheng, H. "Application of BP neural network and genetic algorithm in stress prediction of anchor bolt". In *2015 7th International Conference on Modelling, Identification and Control (ICMIC), IEEE*, pp. 1-4, December 2015.
- [29] Cheng, X., Sun, L., & Liu, P. "Application of regression analysis based on genetic particle swarm algorithm in financial analysis". In *2010 International Conference On Computer Design and Applications, IEEE*, vol. 4, pp. V4-335, June 2010.
- [30] Liu, Y., Tang, T., & Xun, J. "Prediction algorithms for train arrival time in urban rail transit". In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), IEEE*, pp. 1-6, October 2017.
- [31] Xufeng, J., Changying, G., Yuan, Z., & Jianbo, W. "Application of Genetic Algorithm-Based Artificial Neural Network in Prediction of Aircraft Engine Wear". In *2010 International Conference on Digital Manufacturing & Automation, IEEE*, vol. 1, pp. 260-262, December 2010.
- [32] Weiss, G. M., & Hirsh, H. "Learning to Predict Rare Events in Event Sequences". In *KDD*, vol. 98, pp. 359-363, August 1998.
- [33] Hou, T. H., Su, C. H., & Liu, W. L. "Parameters optimization of a nano-particle wet milling process using the Taguchi method, response surface method and genetic algorithm". *Powder technology*, vol. 173, no. 3, pp. 153-162, 2007.
- [34] Aly, A. A. "PID parameters optimization using genetic algorithm technique for electrohydraulic servo control system". *Intelligent Control and Automation*, vol. 2, no. 2, p. 69, 2011.
- [35] Ou, C., & Lin, W. "Comparison between PSO and GA for parameters optimization of PID controller". In *2006 International conference on mechatronics and automation, IEEE*, pp. 2471-2475, June 2006.
- [36] Zhao, M., Fu, C., Ji, L., Tang, K., & Zhou, M. "Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes". *Expert Systems with Applications*, vol. 38, no. 5, pp. 5197-5204, 2011.
- [37] Sun, W. X., Song, T., & Zhong, H. "Study on bus passenger capacity forecast based on regression analysis including time series". In *2009 International Conference on Measuring Technology and Mechatronics Automation, IEEE*, vol. 2, pp. 381-384, April 2009.
- [38] Wang, J., & Yu, L. "Use the combination of the decision tree and the artificial neural networks to predict the outcome of table tennis matches". In *2010 Sixth International Conference on Natural Computation, IEEE*, vol. 4, pp. 1929-1933, August 2010.
- [39] Nguyen, D. L., & Le, T. M. "Recommendation system for Facebook public events based on probabilistic classification and re-ranking". In *2016 Eighth International Conference on Knowledge and Systems Engineering (KSE), IEEE*, pp. 133-138, October 2016.
- [40] Caruana, R., Niculescu-Mizil, A., Crew, G., & Ksikes, A. "Ensemble selection from libraries of models". In *Proceedings of the twenty-first international conference on Machine learning, ACM*, p. 18, July 2004.
- [41] Ye, J., Chow, J. H., Chen, J., & Zheng, Z. "Stochastic gradient boosted distributed decision trees". In *Proceedings of the 18th ACM conference on Information and knowledge management, ACM*, pp. 2061-2064, November 2009.
- [42] Min, S. H., Lee, J., & Han, I. "Hybrid genetic algorithms and support vector machines for bankruptcy prediction". *Expert systems with applications*, vol. 31, no. 3, pp. 652-660, 2006.
- [43] Azeez, J., & Aravindhar, D. J. "Hybrid approach to crime prediction using deep learning". In *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE*, pp. 1701-1710, August 2015.
- [44] Furnkranz, J., & Flach, P. "An analysis of rule learning heuristics (Technical Report CSTR-03-002)". Department of Computer Science, University of Bristol, 2003.
- [45] M. Hall, G. Holmes, "Benchmarking Attribute Selection Techniques for Discrete Class Data Mining", *IEEE Trans. Knowledge and Data Eng.* Vol. 15, no. 6, pp. 1437-1447, 2003.
- [46] C. X. Ling, J. Huang, H. Zhang, "AUC: A Better Measure than Accuracy in Comparing Learning Algorithms", *Canadian Conference on Artificial Intelligence*, Halifax, Canada, pp. 329-341, 2003.
- [47] T. Mende, R. Koschke, "Revisiting the Evaluation of Defect Prediction Models", *Proc. 5th international Conference on Predictor Models in Software Engineering*, Vancouver, Canada, pp. 1-10, 2009.
- [48] T. Menzies, J. Greenwald, A. Frank, "Data Mining Static Code Attributes to Learn Defect Predictors", *IEEE Transactions on Software Engineering*, vol. 32, no. 1, pp. 2-13, 2007.
- [49] Catal, C. "Performance evaluation metrics for software fault prediction studies". *Acta Polytechnica Hungarica*, vol. 9, no. 4, pp. 193-206, 2012.
- [50] Gunawardana, A., & Shani, G. "A survey of accuracy evaluation metrics of recommendation tasks". *Journal of Machine Learning Research*, vol. 10, pp. 2935-2962, Dec 2009.