Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2019 doi:10.20944/preprints201908.0203.v1

Advances in Machine Learning Modeling Reviewing Hybrid and Ensemble Methods

Sina Ardabili [1], Amir Mosavi [2,3]* and Annamaria R. Varkonyi-Koczy [2,4]

[1] Institute of Advanced Studies Koszeg, Koszeg, Hungary
[2] Kalman Kando Faculty of Electrical Engineering, Obuda University, Budapest, Hungary
[3] School of the Built Environment, Oxford Brookes University, Oxford OX3 0BP, UK
[4] Department of Mathematics and Informatics, J. Selye University, Komarno, Slovakia
* a.mosavi@brookes.ac.uk

Abstract. The conventional machine learning (ML) algorithms are continuously advancing and evolving at a fast-paced by introducing the novel learning algorithms. ML models are continually improving using hybridization and ensemble techniques to empower computation, functionality, robustness, and accuracy aspects of modeling. Currently, numerous hybrid and ensemble ML models have been introduced. However, they have not been surveyed in a comprehensive manner. This paper presents the state of the art of novel ML models and their performance and application domains through a novel taxonomy.

Keywords: machine learning; deep learning; ensemble models

Nomenclatures			
Artificial neural network	ANN	Bagging-based naïve bayes trees	BAGNBT
Extreme learning machine	ELM	Ensemble empirical mode decomposition	EEMD
Machine learning	ML	Grasshopper optimization algorithm	GOA
Support vector machine	SVM	Hybrid of linear regression-deep neural network	HybPAS
Wavelet neural networks	WNN	Trauma Severity model	TSM
Deep learning	DL	Gradient boosting decision tree	GBDT
Autoregressive integrated moving average	ARIMA	Evidential belief function and tree-based models	EBFTM
Ensemble empirical with adaptive noise technology	EE-ANT	Decision tree overfitting and neural network	DTFNN
Data assimilation Kalman filter-based	DA-KF	Improved complete ensemble empirical mode decomposition method with adaptive noise	ICEEMDMAN
Online sequential extreme learning machine	OSELM	Random forest	RF

1. Introduction

Machine learning (ML) methods are reported to outperform most of the physical and statistical methods in predictive modeling in terms of accuracy, robustness, uncertainty analysis, data efficiency, simplicity, and computation cost. Thus, ML methods have gained massive popularity during the past few years in a diverse range of applications, energy, hydrology, hazard prediction, finance, economics, computational mechanics, etc [1-9]. ML methods are numerous, and different classifications of methods have been recently given by researchers [10-13]. One of the popular classification methods is to divide the methods in three groups, i.e., single methods, hybrid methods, and ensembles [14-16].

The popular single ML methods which have been widely used include artificial neural networks (ANNs)-based methods [17-19], decision trees (DTs)-based methods [18, 20-23], support vector machines (SVM)-based methods [24-27], Bayesians-based [28-30], neuro-fuzzy-based [11, 12, 31, 32], classification and regression-based methods [33], and wavelet neural networks (WNNs)-based [12, 34]. Neuro-fuzzy methods and WNNs, although they are built upon two intelligent algorithms, have already been established as a single method. The ML methods are constantly progressing to hands-on higher performance algorithms [35-49]. The hybrid and ensemble methods are often identified to outperform single ML methods [50-60].

Ensemble and hybrid ML methods are the two major approaches toward more accurate, and reliable ML methods [61-63]. Hybrid ML models are made through integration of ML methods, with other ML methods, and/or with other soft computing, optimization techniques to improve the method in various aspects. While the ensemble methods are made using various grouping techniques such as bagging or boosting to use more than one ML classifier. It is suggested that the future success of ML highly depends on the advancement of novel ensemble and hybrids methods [32, 64-66]. Literature includes novel ML methods and various comparative analysis to identify the methods with the higher performance [67-69]. However, there is a gap in research in identifying the novel hybrid and ensemble ML methods and the applications they have been used in. Consequently, the contribution of this paper is to introduce these methods and highlight their applications.

2. Reviewing ensemble and hybrid ML methods

The use of ML methods, including singles, ensembles, and hybrids, have been dramatically increasing. Figure.1 shows the ever fast-growing trend of ML methods used from 2009 up until now. The widespread application areas include engineering, mathematics, physics, astronomy, earth and planetary sciences, medicine, materials science, biochemistry, genetics and molecular biology, environmental science, social sciences, energy, chemistry, decision sciences, agricultural and biological sciences.

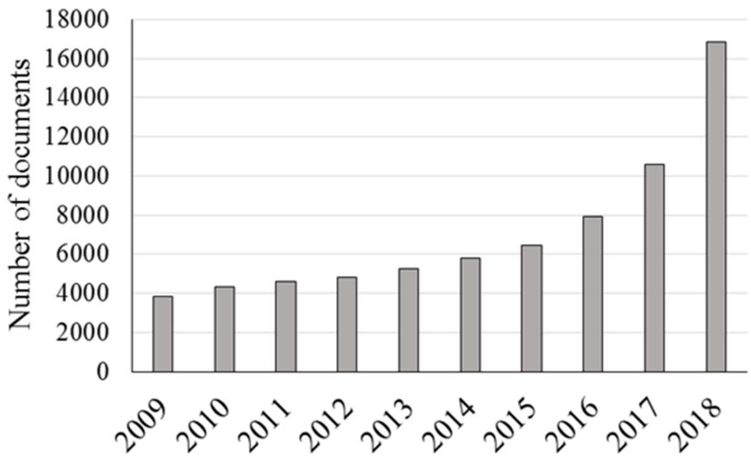

Figure 1. The growing trend of ML methods the past decade (source: web of science).

In the following, Figure.2 shows the popularity of ensemble and hybrid ML models in advancing the novel method with higher performance.

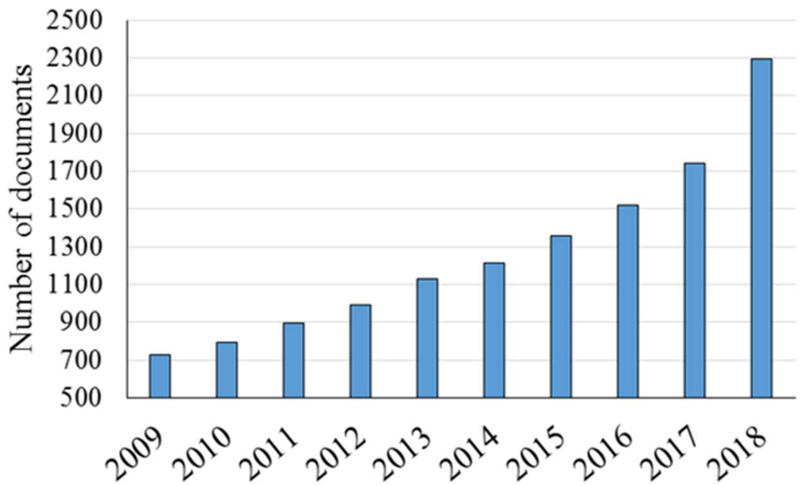

Figure 2. The growing trend of hybrid and ensemble ML methods (source: web of science).

2.1. Hybrid methods

Hybrid methods combine two or more ML and/or soft computing methods for higher performance and optimum results. In fact, hybrid methods benefit from the advantage of two or more methods reach better performance. Sometimes, hybrid methods contain one unit for prediction and one unit for the optimization of the prediction unit for reaching an accurate output. Therefore, it can be claimed that hybrid methods contain different single methods and form a method with higher flexibility with a high capability compared with single methods. Hybrid methods have become more popular due to their high potential and capability. Hybrid methods are the same as a company with different employees with different expertise to achieve a single goal.

Table 1 presents the top six studies developed by hybrid methods. Table 1 contains four columns including reference, contribution of each study, the developed method and application domain for presenting the key-point of each study for a quick look.

Table 1. Studies developed by hybrid methods.

Refeences	Contributions	Methods	Application domains
[70]	To develop an adaptive hybrid methodology for the estimation of urban traffic flow	ARIMA-WNN	-Urban traffic flow -Advanced hybrid machine learning
[71]	To develop an innovative hybrid method for the estimation and optimization of wind energy	EE-ANT-WNN	-Wind power -Hybrid machine learning
[72]	To develop a novel hybrid multi-stage method to be applied in credit scoring	Hybrid multi-stage method	-Classification -Multi-stage hybrid model
[73]	To develop a novel hybrid bagging based method for the assessment of the Landslide susceptibility	Hybrid BAGNBT	-Landslide susceptibility -Hybrid machine learning
[74]	To develop a hybrid method for the estimation of electricity load	EEMD-ELM-GOA	-Electrical load -Hybrid machine learning
[75]	To develop a hybrid linear regression-based deep learning method for the estimation of poly (A) signals in DNA	HybPAS	-Hybrid machine learning -Signal processing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2019 doi:10.20944/preprints201908.0203.v1

Hou et al. [70] developed a study in order to accurately estimate the urban traffic flow. The proposed method was an advanced hybrid wavelet neural network-integrated by autoregressive integrated moving average using a fuzzy method. The developed hybrid methods have been compared with the single form of each contributed methods in terms of mean absolute percentage error, and root mean square error. Results indicated about 60-70 % improvement in the estimation accuracy of the hybrid method over the single methods. In another study, Du et al. [71] developed a novel hybrid method for the estimation and optimization of wind power. The method was including an integrated ensemble empirical with adaptive noise technology for eliminating noise and extracting the main features of original data followed by an optimized wavelet neural network to take a high estimation accuracy. Results have been compared using mean absolute percentage error. Based on the results, the hybrid method could increase the accuracy of the estimation as well as increasing the sustainability of the prediction and optimization process.

Zhang et al. [72] proposed a novel hybrid method based on feature and classifier selections in order to take an optimal classifier and feature subset in credit scoring task. Improving the accuracy of the estimation phase was performed by the use of classifier ensemble as well as using an enhanced multi-population niche genetic algorithm. Evaluations have been performed using accuracy and area under the curve factors. Based on the results, the proposed hybrid method could successfully cope with the estimation and optimization tasks over the single methods.

Pham and Prakash [73] developed a novel bagging-based naïve Bayes trees for the assessment of landslide susceptibility. The proposed hybrid method was compared with single methods including Rotation forest-based Naïve Bayes Trees, Naïve Bayes Trees, and SVM in terms of area under the curve and statistical indexes. Based on results, the proposed hybrid BAGNBT method could successfully increase the accuracy and could be introduced as the best alternative model for the assessment of landslide susceptibility over the single methods.

Wu et al. [74] developed a novel hybrid method for improving the accuracy of the electricity load forecasting. The proposed method was including an advanced integration of ELM, ensemble empirical mode decomposition, and grasshopper optimization algorithm. The hybrid method has been compared with the necessary methods by employing the test data sets in terms of root mean square error, mean absolute error and mean absolute percentage error. Based on the results, the proposed hybrid method has a higher performance and accuracy compared with the necessary methods. Albalawi et al. [75] developed a hybrid HybPAS including the integration of linear regression-deep neural network models for the estimation of ply (a) signals in DNA in the presence of sequence-based features and signal processing-based statistical as input values. Based on the results, the hybrid method could successfully increase the accuracy and performance by 30.29 %.

As is clear from the above mentioned, brief literature, the hybrid methods are expanding and becoming popular due to their high potential and capability for increasing the estimation and optimizing performances. Table 2 represents a brief at the same time complete comparison for single and hybrid methods in terms of accuracy, reliability, and sustainability.

Table 2. the comparison results of Hybrid machine learning-based methods.

Method	*Application*	*Accuracy*	*Reliability*	*Sustainability*	*Reference*
Hybrid WNN-ARIMA	Estimation	+++	+++	+++	[70]
WNN	Estimation	++	++	++	[70]
ARIMA	Estimation	++	+	+	[70]
Hybrid EE-ANT-WNN	Estimation	+++	+++	+++	[71]
Hybrid the optimized multi-stage method	Estimation	+++	+++	+++	[72]
BAGNBT	Estimation	+++	+++	+++	[73]
SVM	Estimation	++	+	+	[73]
NBT	Estimation	+	+	+	[73]
RFNBT	Estimation	++	++	++	[73]
EEMD-ELM-GOA	Estimation	+++	+++	+++	[74]

HybPAS	Estimation	+++	+++	+++	[75]

2.2. Ensemble methods

Ensemble methods may use a series of ML classification trees instead of a single one. Through this technique, the accuracy of the model is substantially improved. Ensemble methods are considered as supervised learning algorithms. Ensemble methods benefit different training algorithms for increasing the training accuracy for reaching a higher testing accuracy. Ensemble method enables different training algorithms for making flexible training. Table 3 presents the top six studies developed by different Ensemble methods with different tasks.

Table 3. Studies developed by Ensemble methods.

References	Contributions	Method	Application domains
[76]	To develop an ensemble machine learning methodology for the estimation of risk	Ensemble TSM	-Risk prediction -Ensemble machine learning
[77]	To develop an ensemble model to estimate the churn in the relation of customers and search Ads.	Ensemble GBDT	-Customer churn -Ensemble machine learning
[78]	To employ rotation forest with DT as an ensemble methodology based on EBF and tree-based models for developing GPM	Ensemble EBFTM	-Hydrogeology -Ensemble machine learning
[79]	To develop a novel ensemble machine learning method integrated by ELM for the estimation of significant wave height	Ensemble ICEEMDAN-ELM	-Wave height forecasting -Ensemble machine learning
[80]	To develop a novel Ensemble data assimilation Kalman filter-based for the estimation of parameters of the system's state	Ensemble DA-KF	-System's state -Ensemble machine learning
[81]	To develop an ensemble estimation model for forecasting the thyroid	Ensemble Bagging-Boosting	-Thyroid forecasting -Ensemble machine learning

Gorczyca et al. [76] developed a Trauma Severity model as an ensemble machine learning for risk estimation. This method has been compared with the Harborview Assessment for Risk of Mortality, Bayesian Logistic Injury Severity Score, and the Trauma Mortality Prediction Model in terms of accuracy and F-score values. Based on the results, the proposed ensemble method could successfully increase the accuracy compared with that of the base method. Results also indicated that trauma is an essential predictor for this task.

Wang et al. [77] developed an ensemble gradient boosting decision tree model for the estimation of customer churn and its relation with search Ads in the presence of two types of features including dynamic and static features. Bing Ads platform dataset was employed in order to evaluate the developed method. The results were highly promising and could successfully cope with the related task with high sustainability. Naghibi et al. [78] developed a rotation forest with decision trees as an ensemble methodology based on evidential belief function and tree-based models (EBFTM) for developing groundwater potential maps. The developed ensemble method has been compared with boosted regression tree, random forest, and classification and regression tree in terms of the performance factors receiver operating characteristics and area under the curve. The highest performance was owned by the ensemble EBFTM method followed by random forest and boosted regression tree.

Ali and Prasad [79] developed a novel ensemble empirical mode decomposition method with adaptive noise integrated with extreme learning machine for accurately estimate the significant wave height. Evaluations have been performed geographically in the presence of the proposed method and ICEEMDAN-OSELM and ICEEMDAN-RF as the ensemble empirical mode decomposition method with adaptive noise integrated with online sequential extreme learning machine and random forest, respectively. Based on the results, the proposed ICEEMDAN-ELM method indicated the best performance over ICEEMDAN-OSELM and ICEEMDAN-RF with high accuracy and sustainability.

Yamanaka et al. [80] developed a novel ensemble method based on data assimilation-Kalman filter for the estimation of microstructure prediction using three-dimensional multi-phase-field as parameters of the system's state. Based on the results, the proposed method could successfully increase the accuracy with the lowest error, which indicates the capability of the model for directly applied to estimate the system parameters. Yadav and Pal [81] developed a novel ensemble method based on Bagging-Boosting for the estimation of women thyroid which is compared with decision tree overfitting and neural network (DTFNN) in the presence of root mean square error and mean absolute error. Based on results, the ensemble bagging-boosting method had about 65% higher accuracy over DTFNN method. As is apparent, the ensemble methods could successfully own a higher accuracy and sustainability, followed by higher attentions and trends. Therefore ensemble methods can be used by different policymakers. Accordingly, Table 4 presents a brief at the same time complete comparison for ensemble methods in terms of accuracy, reliability, and sustainability.

Table 4. the comparison results of Ensemble machine learning based methods.

Method	*Application*	*Accuracy*	*Reliability*	*Sustainability*	*Reference*
Ensemble TSM	Estimation	+++	+++	+++	[76]
Ensemble GBDT	Estimation	+++	+++	+++	[77]
Ensemble EBFTM	Estimation	+++	+++	+++	[78]
RF	Estimation	+++	++	++	[78]
BRT	Estimation	++	++	++	[78]
ICEEMDAN-ELM	Estimation	+++	+++	+++	[79]
ICEEMDAN-OSELM	Estimation	++	+	+	[79]
ICEEMDAN-RF	Estimation	++	++	++	[79]
Ensemble KF-DA	Estimation	++	++	++	[80]
Ensemble bagging-boosting	Estimation	+++	+++	+++	[81]
DTFNN	Estimation	++	+	+	[81]

3. Conclusions

The ensemble and hybrid models are the new generations of machine learning. They provide higher accuracy and outperform most of the conventional machine learning models. This paper presented the state of the art of hybrids and ensembles and listed their most famous algorithms and application domains. Ensembles are currently limited to decision trees; however, it is expected to expand to other machine learning methods. Bagging and boosting methods are reported as the most popular technique to build ensembles. The hybrid models are not built through the integration of optimization and/or soft computing methods to optimize the method. Health, energy, climate change, urban informatics, and hydrology are the primary application domains of ensemble and hybrid models. Consequently, future research trends are devoted to the novel hybrid and ensemble methods [82-91].

Acknowledgments

This publication has been supported by the Project: "Support of research and development activities of the J. Selye University in the field of Digital Slovakia and creative industry" of the Research & Innovation Operational Programme (ITMS code: NFP313010T504) co-funded by the European Regional Development Fund.

 doi:10.20944/preprints201908.0203.v1

References

1. Choubin, B., et al., *Snow avalanche hazard prediction using machine learning methods.* Journal of Hydrology, 2019. **577**.
2. Dehghani, M., et al., *Prediction of hydropower generation using Grey wolf optimization adaptive neuro-fuzzy inference system.* Energies, 2019. **12**(2).
3. Mosavi, A., Y. Bathla, and A. Varkonyi-Koczy, *Predicting the future using web knowledge: State of the art survey*, D. Luca, L. Sirghi, and C. Costin, Editors. 2018, Springer Verlag. p. 341-349.
4. Mosavi, A. and T. Rabczuk, *Learning and intelligent optimization for material design innovation*, D.E. Kvasov, et al., Editors. 2017, Springer Verlag. p. 358-363.
5. Mosavi, A., T. Rabczuk, and A.R. Varkonyi-Koczy, *Reviewing the novel machine learning tools for materials design*, D. Luca, L. Sirghi, and C. Costin, Editors. 2018, Springer Verlag. p. 50-58.
6. Qasem, S.N., et al., *Estimating daily dew point temperature using machine learning algorithms.* Water (Switzerland), 2019. **11**(3).
7. Taherei Ghazvinei, P., et al., *Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network.* Engineering Applications of Computational Fluid Mechanics, 2018. **12**(1): p. 738-749.
8. Torabi, M., et al., *A Hybrid clustering and classification technique for forecasting short-term energy consumption.* Environmental Progress and Sustainable Energy, 2019. **38**(1): p. 66-76.
9. Torabi, M., et al., *A Hybrid Machine Learning Approach for Daily Prediction of Solar Radiation*, in *Lecture Notes in Networks and Systems*. 2019, Springer. p. 266-274.
10. Dineva, A., et al., *Review of soft computing models in design and control of rotating electrical machines.* Energies, 2019. **12**(6).
11. Mosavi, A., P. Ozturk, and K.W. Chau, *Flood prediction using machine learning models: Literature review.* Water (Switzerland), 2018. **10**(11).
12. Mosavi, A., et al., *State of the art of machine learning models in energy systems, a systematic review.* Energies, 2019. **12**(7).
13. Mosavi, A. and A.R. Varkonyi-Koczy, *Integration of machine learning and optimization for robot learning*, R. Jablonski and R. Szewczyk, Editors. 2017, Springer Verlag. p. 349-355.
14. Cheng, L. and T. Yu, *A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems.* International Journal of Energy Research, 2019. **43**(6): p. 1928-1973.
15. Cheng, L., et al., *Machine Learning for Energy and Electric Power Systems: State of the Art and Prospects.* Dianli Xitong Zidonghua/Automation of Electric Power Systems, 2019. **43**(1): p. 15-31.
16. Chou, J.S. and D.S. Tran, *Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders.* Energy, 2018: p. 709-726.
17. Chatterjee, B., et al., *RF-PUF: Enhancing IoT Security Through Authentication of Wireless Nodes Using In-Situ Machine Learning.* IEEE Internet of Things Journal, 2019. **6**(1): p. 388-398.
18. Panesar, S.S., et al., *Machine Learning Versus Logistic Regression Methods for 2-Year Mortality Prognostication in a Small, Heterogeneous Glioma Database.* World Neurosurgery: X, 2019. **2**.
19. Thomas, P.B.M., et al., *Feasibility of simple machine learning approaches to support detection of non-glaucomatous visual fields in future automated glaucoma clinics.* Eye (Basingstoke), 2019. **33**(7): p. 1133-1139.
20. Alhajri, M.I., N.T. Ali, and R.M. Shubair, *Classification of Indoor Environments for IoT Applications: A Machine Learning Approach.* IEEE Antennas and Wireless Propagation Letters, 2018. **17**(12): p. 2164-2168.
21. Jamil, A. and B. Bayram, *The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods.* Geocarto International, 2019.
22. Maxwell, A.E., T.A. Warner, and F. Fang, *Implementation of machine-learning classification in remote sensing: An applied review.* International Journal of Remote Sensing, 2018. **39**(9): p. 2784-2817.
23. Sehgal, V., et al., *Machine learning creates a simple endoscopic classification system that improves dysplasia detection in barrett's oesophagus amongst non-expert endoscopists.* Gastroenterology Research and Practice, 2018. **2018**.
24. Azeem, M.I., et al., *Machine learning techniques for code smell detection: A systematic literature review and meta-analysis.* Information and Software Technology, 2019. **108**: p. 115-138.
25. Jabeen, A. and S. Ranganathan, *Applications of machine learning in GPCR bioactive ligand discovery.* Current Opinion in Structural Biology, 2019. **55**: p. 66-76.
26. Xu, C. and S.A. Jackson, *Machine learning and complex biological data.* Genome Biology, 2019. **20**(1).
27. Zhang, Z. and E. Sejdić, *Radiological images and machine learning: Trends, perspectives, and prospects.* Computers in Biology and Medicine, 2019. **108**: p. 354-370.
28. Bock, F.E., et al., *A review of the application of machine learning and data mining approaches in continuum materials mechanics.* Frontiers in Materials, 2019. **6**.

29. Ekins, S., et al., *Exploiting machine learning for end-to-end drug discovery and development.* Nature Materials, 2019. **18**(5): p. 435-441.
30. Woldaregay, A.Z., et al., *Data-driven blood glucose pattern classification and anomalies detection: Machine-learning applications in type 1 diabetes.* Journal of Medical Internet Research, 2019. **21**(5).
31. Najafzadeh, M. and A. Ghaemi, *Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods.* Environmental Monitoring and Assessment, 2019. **191**(6).
32. Singh, H., P.S. Rana, and U. Singh, *Prediction of drug synergy in cancer using ensemble-based machine learning techniques.* Modern Physics Letters B, 2018. **32**(11).
33. Choubin, B., et al., *River suspended sediment modelling using the CART model: A comparative study of machine learning techniques.* Science of the Total Environment, 2018. **615**: p. 272-281.
34. Zarkogianni, K., M. Athanasiou, and A.C. Thanopoulou, *Comparison of Machine Learning Approaches Toward Assessing the Risk of Developing Cardiovascular Disease as a Long-Term Diabetes Complication.* IEEE Journal of Biomedical and Health Informatics, 2018. **22**(5): p. 1637-1647.
35. Aram, F., et al., *Design and validation of a computational program for analysing mental maps: Aram mental map analyzer.* Sustainability (Switzerland), 2019. **11**(14).
36. Asadi, E., et al., *Groundwater Quality Assessment for Drinking and Agricultural Purposes in Tabriz Aquifer, Iran.* 2019.
37. Asghar, M.Z.; Subhan, F.; Imran, M.; Kundi, F.M.; Shamshirband, S.; Mosavi, A.; Csiba, P.; R. Várkonyi-Kóczy, A. Performance Evaluation of Supervised Machine Learning Techniques for Efficient Detection of Emotions from Online Content. Pre-prints 2019, 2019080019 (doi: 10.20944/preprints201908.0019.v1).
38. Bemani, A.; Baghban, A.; Shamshirband, S.; Mosavi, A.; Csiba, P.; Várkonyi-Kóczy, A.R. Applying ANN, ANFIS, and LSSVM Models for Estimation of Acid Sol-vent Solubility in Supercritical CO2. Preprints 2019, 2019060055 (doi: 10.20944/preprints201906.0055.v2).
39. Mosavi, A., T. Rabczuk, and A.R. Varkonyi-Koczy, *Reviewing the novel machine learning tools for materials design*, D. Luca, L. Sirghi, and C. Costin, Editors. 2018, Springer Verlag. p. 50-58.
40. Mosavi, A., et al., *Prediction of multi-inputs bubble column reactor using a novel hybrid model of computational fluid dynamics and machine learning.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 482-492.
41. Mosavi, A. and A.R. Varkonyi-Koczy, *Integration of machine learning and optimization for robot learning*, R. Jablonski and R. Szewczyk, Editors. 2017, Springer Verlag. p. 349-355.
42. Nosratabadi, S., et al., *Sustainable business models: A review.* Sustainability (Switzerland), 2019. **11**(6).
43. Rezakazemi, M., A. Mosavi, and S. Shirazian, *ANFIS pattern for molecular membranes separation optimization.* Journal of Molecular Liquids, 2019. **274**: p. 470-476.
44. Riahi-Madvar, H., et al., *Comparative analysis of soft computing techniques RBF, MLP, and ANFIS with MLR and MNLR for predicting grade-control scour hole geometry.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 529-550.
45. Shabani, S.; Samadianfard, S.; Taghi Sattari, M.; Shamshirband, S.; Mosavi, A.; Kmet, T.; R. Várkonyi-Kóczy, A. Modeling Daily Pan Evaporation in Humid Cli-mates Using Gaussian Process Regression. Preprints 2019, 2019070351 (doi: 10.20944/preprints201907.0351.v1).
46. Shamshirband, S.; Hadipoor, M.; Baghban, A.; Mosavi, A.; Bukor J.; Annamaria R. Varkonyi-Koczy, Developing an ANFIS-PSO Model to predict mercury emissions in Combustion Flue Gases. Preprints 2019, 2019070165 (doi: 10.20944/preprints201907.0165.v1).
47. Shamshirband, S., et al., *Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 91-101.
48. Shamshirband, S., A. Mosavi, and T. Rabczuk, *Particle swarm optimization model to predict scour depth around bridge pier.* arXiv preprint arXiv:1906.08863, 2019.
49. Torabi, M., et al., *A Hybrid Machine Learning Approach for Daily Prediction of Solar Radiation*, in *Lecture Notes in Networks and Systems*. 2019, Springer. p. 266-274.
50. Choubin, B., et al., *An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines.* Science of the Total Environment, 2019. **651**: p. 2087-2096.
51. Dineva, A., et al., *Multi-Label Classification for Fault Diagnosis of Rotating Electrical Machines.* 2019.
52. Farzaneh-Gord, M., et al., *Numerical simulation of pressure pulsation effects of a snubber in a CNG station for increasing measurement accuracy.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 642-663.

53. Ghalandari, M., et al., *Investigation of submerged structures' flexibility on sloshing frequency using a boundary element method and finite element analysis.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 519-528.
54. Ghalandari, M., et al., *Flutter speed estimation using presented differential quadrature method formulation.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 804-810.
55. Karballaeezadeh, N., et al., *Prediction of remaining service life of pavement using an optimized support vector machine (case study of Semnan–Firuzkuh road).* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 188-198.
56. Menad, N.A., et al., *Modeling temperature dependency of oil - water relative permeability in thermal enhanced oil recovery processes using group method of data handling and gene expression programming.* Engineering Applications of Computational Fluid Mechanics, 2019. **13**(1): p. 724-743.
57. Mohammadzadeh, S., et al., *Prediction of Compression Index of Fine-Grained Soils Using a Gene Expression Programming Model.* Infrastructures, 2019. **4**(2): p. 26.
58. Mosavi, A. and M. Edalatifar, *A Hybrid Neuro-Fuzzy Algorithm for Prediction of Reference Evapotranspiration*, in *Lecture Notes in Networks and Systems*. 2019, Springer. p. 235-243.
59. Mosavi, A., A. Lopez, and A.R. Varkonyi-Koczy, *Industrial applications of big data: State of the art survey*, D. Luca, L. Sirghi, and C. Costin, Editors. 2018, Springer Verlag. p. 225-232.
60. Mosavi, A. and T. Rabczuk, *Learning and intelligent optimization for material design innovation*, D.E. Kvasov, et al., Editors. 2017, Springer Verlag. p. 358-363.
61. Bui, D.T., et al., *Shallow landslide prediction using a novel hybrid functional machine learning algorithm.* Remote Sensing, 2019. **11**(8).
62. Pham, B.T., et al., *Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches.* Catena, 2019. **175**: p. 203-218.
63. Zhang, X. and S. Mahadevan, *Ensemble machine learning models for aviation incident risk prediction.* Decision Support Systems, 2019. **116**: p. 48-63.
64. Jaiswal, A. and R. Malhotra, *Software reliability prediction using machine learning techniques.* International Journal of Systems Assurance Engineering and Management, 2018. **9**(1): p. 230-244.
65. Khagi, B., G.R. Kwon, and R. Lama, *Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques.* International Journal of Imaging Systems and Technology, 2019.
66. Kumari, M., et al., *Comparative analysis of machine learning based QSAR models and molecular docking studies to screen potential anti-tubercular inhibitors against InhA of mycobacterium tuberculosis.* International Journal of Computational Biology and Drug Design, 2018. **11**(3): p. 209-235.
67. Bataineh, A.A., *A comparative analysis of nonlinear machine learning algorithms for breast cancer detection.* International Journal of Machine Learning and Computing, 2019. **9**(3): p. 248-254.
68. Manzoor, S.I. and J. Singla, *A comparative analysis of machine learning techniques for spam detection.* International Journal of Advanced Trends in Computer Science and Engineering, 2019. **8**(3): p. 810-814.
69. Odugu, K. and B. Rajasekar, *Comparative analysis on supervised machine learning models for future wireless communication networks.* International Journal of Innovative Technology and Exploring Engineering, 2019. **8**(6): p. 721-723.
70. Hou, Q., et al., *An adaptive hybrid model for short-term urban traffic flow prediction.* Physica A: Statistical Mechanics and its Applications, 2019. **527**.
71. Du, P., et al., *A novel hybrid model for short-term wind power forecasting.* Applied Soft Computing Journal, 2019. **80**: p. 93-106.
72. Zhang, W., H. He, and S. Zhang, *A novel multi-stage hybrid model with enhanced multi-population niche genetic algorithm: An application in credit scoring.* Expert Systems with Applications, 2019. **121**: p. 221-232.
73. Pham, B.T. and I. Prakash, *A novel hybrid model of Bagging-based Naïve Bayes Trees for landslide susceptibility assessment.* Bulletin of Engineering Geology and the Environment, 2019. **78**(3): p. 1911-1925.
74. Wu, J., et al., *A new hybrid model to predict the electrical load in five states of Australia.* Energy, 2019: p. 598-609.
75. Albalawi, F., et al., *Hybrid model for efficient prediction of poly(A) signals in human genomic DNA.* Methods, 2019.
76. Gorczyca, M.T., N.C. Toscano, and J.D. Cheng, *The trauma severity model: An ensemble machine learning approach to risk prediction.* Computers in Biology and Medicine, 2019. **108**: p. 9-19.
77. Wang, Q.F., M. Xu, and A. Hussain, *Large-scale Ensemble Model for Customer Churn Prediction in Search Ads.* Cognitive Computation, 2019. **11**(2): p. 262-270.
78. Naghibi, S.A., et al., *Application of rotation forest with decision trees as base classifier and a novel ensemble model in spatial modeling of groundwater potential.* Environmental Monitoring and Assessment, 2019. **191**(4).

 doi:10.20944/preprints201908.0203.v1

79. Ali, M. and R. Prasad, *Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition.* Renewable and Sustainable Energy Reviews, 2019: p. 281-295.
80. Yamanaka, A., Y. Maeda, and K. Sasaki, *Ensemble Kalman filter-based data assimilation for three-dimensional multi-phase-field model: Estimation of anisotropic grain boundary properties.* Materials and Design, 2019. **165**.
81. Yadav, D.C. and S. Pal, *To generate an ensemble model for women thyroid prediction using data mining techniques.* Asian Pacific Journal of Cancer Prevention, 2019. **20**(4): p. 1275-1281.
82. Ardabili, S., Mosavi, A., Mahmoudi, Mesri Gundoshmian, T, Nosratabadi, S., Var-konyi-Koczy, A., Modelling temperature variation of mushroom growing hall us-ing artificial neural networks, Preprints 2019.
83. Mesri Gundoshmian, T., Ardabili, S., Mosavi, A., Varkonyi-Koczy, A., Prediction of combine harvester performance using hybrid machine learning modeling and response surface methodology, Preprints 2019.
84. Ardabili, S., Mosavi, A., Varkonyi-Koczy, A., Systematic review of deep learning and machine learning models in biofuels research, Preprints 2019.
85. Ardabili, S., Mosavi, A., Varkonyi-Koczy, A., Advances in machine learning model-ing reviewing hybrid and ensemble methods, Preprints 2019.
86. Ardabili, S., Mosavi, A., Varkonyi-Koczy, A., Building Energy information: demand and consumption prediction with Machine Learning models for sustainable and smart cities, Preprints 2019.
87. Ardabili, S., Mosavi, A., Dehghani, M., Varkonyi-Koczy, A., Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review, Preprints 2019.
88. Mohammadzadeh D., Karballaeezadeh, N., Mohemmi, M., Mosavi, A., Varkonyi-Koczy A., Urban Train Soil-Structure Interaction Modeling and Analysis, Preprints 2019.
89. Mosavi, A., Ardabili, S., Varkonyi-Koczy, A., List of deep learning models, Preprints 2019.
90. Nosratabadi, S., Mosavi, A., Keivani, R., Ardabili, S., Aram, F., State of the art sur-vey of deep learning and machine learning models for smart cities and urban sustainability, Preprints 2019.
91. Perez, H.; Tah, J.H.M.; Mosavi, A. Deep Learning for Detecting Building Defects Using Convolutional Neural Networks. Sensors 2019, 19, 3556.