Evolving Systems manuscript No.
(will be inserted by the editor)

On ensemble techniques of weight-constrained neural networks

Ioannis E. Livieris · Lazaros Iliadis · Panagiotis Pintelas

Abstract Ensemble learning constitutes one of the most fundamental and reliable strategies for building powerful and accurate predictive models, aiming to exploit the predictions of a number of multiple learners. In this paper, we propose two ensemble prediction models which exploit the classification performance of Weight-Constrained Neural Networks (WC-NNs). The proposed models are based on Bagging and Boosting, which constitute two of the most popular strategies, to efficiently combine the predictions of WCNN classifiers. We conducted a series of experiments using a variety of benchmarks from UCI repository in order to evaluate the performance of the two proposed models against other state-of-the-art ensemble classifiers. The reported experimental results illustrate the prediction accuracy of the proposed models providing empirical evidence that the hybridization of ensemble learning and WCNNs can build efficient and powerful classification models.

Keywords Weight-constrained neural networks · ensemble learning · bagging · boosting-AdaBoost.

I.E. Livieris
Department of Mathematics, University of Patras, GR 265-00, Greece. E-mail: livieris@upatras.gr

L. Iliadis
Department of Civil Engineering, Democritus University of Thrace, GR 67100, Greece. E-mail: liliadis@civil.duth.gr

P. Pintelas
Department of Mathematics, University of Patras, GR 265-00, Greece. E-mail: pintelas@upatras.gr

References

Baldi, P. and Sadowski, P. J. (2013), Understanding dropout, *in* ‘Advances in neural information processing systems’, pp. 2814–2822.

Bauer, E. and Kohavi, R. (1999), ‘An empirical comparison of voting classification algorithms: Bagging, boosting, and variants’, *Machine learning* **36**(1-2), 105–139.

Breiman, L. (1996), ‘Bagging predictors’, *Machine learning* **24**(2), 123–140.

Bühlmann, P. (2012), Bagging, boosting and ensemble methods, *in* ‘Handbook of Computational Statistics’, Springer, pp. 985–1022.

Demuth, H., Beale, M., De Jess, O. and Hagan, M. (2014), *Neural network design*, Martin Hagan.

Dolan, E. and Moré, J. (2002), ‘Benchmarking optimization software with performance profiles’, *Mathematical Programming* **91**, 201–213.

Dua, D. and Karra Taniskidou, E. (2017), ‘UCI machine learning repository’.
URL: *http://archive.ics.uci.edu/ml*

Farzaneh, V., Ghodsvali, A., Bakhshabadi, H., Dolatabadi, Z., Farzaneh, F., Carvalho, I. and Sarabandi, K. (2018), ‘Screening of the alterations in qualitative characteristics of grape under the impacts of storage and harvest times using artificial neural network’, *Evolving Systems* **9**(1), 81–89.

Freund, Y. and Schapire, R. (1996), Experiments with a new boosting algorithm, *in* ‘icml’, Vol. 96, Citeseer, pp. 148–156.

Goodfellow, I., Bengio, Y. and Courville, A. (2016), *Deep learning*, MIT press.

Hager, W. W. and Zhang, H. (2006), ‘Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent’, *ACM Transactions on Mathematical Software (TOMS)* **32**(1), 113–137.

Hara, K., Saitoh, D. and Shouno, H. (2016), Analysis of dropout learning regarded as ensemble learning, *in* ‘International Conference on Artificial Neural Networks’, Springer, pp. 72–79.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. R. (2012), ‘Improving neural networks by preventing co-adaptation of feature detectors’, *arXiv preprint arXiv:1207.0580* .

Hinton, G., Vinyals, O. and Dean, J. (2014), ‘Dark knowledge’, *Presented as the keynote in BayLearn* **2**.

Iliadis, L., Mansfield, S., Avramidis, S. and El-Kassaby, Y. (2013), ‘Predicting Douglas-fir wood density by artificial neural networks (ANN) based on progeny testing information’, *Holzforschung* **67**(7), 771–777.

Khwaja, A., Naeem, M., Anpalagan, A., Venetsanopoulos, A. and Venkatesh, B. (2015), ‘Improved short-term load forecasting using bagged neural networks’, *Electric Power Systems Research* **125**, 109–115.

Khwaja, A., Zhang, X., Anpalagan, A. and Venkatesh, B. (2017), ‘Boosted neural networks for improved short-term electric load forecasting’, *Electric Power Systems Research* **143**, 431–437.

Kim, M. and Kang, D. (2010), ‘Ensemble with neural networks for bankruptcy prediction’, *Expert systems with applications* **37**(4), 3373–3379.

Kotsiantis, S. and Pintelas, P. (2004), ‘Combining bagging and boosting’, *International Journal of Computational Intelligence* **1**(4), 324–333.

Lango, M. and Stefanowski, J. (2018), ‘Multi-class and feature selection extensions of roughly balanced bagging for imbalanced data’, *Journal of Intelligent Information Systems* **50**(1), 97–127.

Leung, K. and Parker, D. (2003), Empirical comparisons of various voting methods in bagging, *in* 'Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining', ACM, pp. 595–600.

Livieris, I. (2019*a*), 'Forecasting economy-related data utilizing constrained recurrent neural networks', *Algorithms* **12**(85).

Livieris, I. (2019*b*), 'Improving the classification efficiency of an ANN utilizing a new training methodology', *Informatics* **6**(1).

Livieris, I., Kotsilieris, T., Stavroyiannis, S. and Pintelas, P. (2019), 'Forecasting stock price index movement using a constrained deep neural network training algorithm', *Intelligent Decision Technologies, (accepted for publication)* .

Livieris, I. and Pintelas, P. (2019*a*), 'An adaptive nonmonotone active set -weight constrained- neural network training algorithm', *Neurocomputing* .

Livieris, I. and Pintelas, P. (2019*b*), 'An improved weight-constrained neural network training algorithm', *Neural Computing and Applications* .

Maren, A., Harston, C. and Pap, R. (2014), *Handbook of neural computing applications*, Academic Press.

Odior, A. (2013), 'Application of neural network and fuzzy model to grinding process control', *Evolving Systems* **4**(3), 195–201.

Opitz, D. W. and Maclin, R. F. (1997), An empirical evaluation of bagging and boosting for artificial neural networks, *in* 'Proceedings of International Conference on Neural Networks (ICNN'97)', Vol. 3, IEEE, pp. 1401–1405.

Patel, J., Fioranelli, F., Ritchie, M. and Griffiths, H. (2018), 'Multistatic radar classification of armed vs unarmed personnel using neural networks', *Evolving Systems* **9**(2), 135–144.

Phaisangittisagul, E. (2016), An analysis of the regularization between l2 and dropout in single hidden layer neural network, *in* '2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)', IEEE, pp. 174–179.

Reed, R. and Marks II, R. J. (1999), *Neural smithing: supervised learning in feedforward artificial neural networks*, Mit Press.

Rokach, L. (2010), 'Ensemble-based classifiers', *Artificial Intelligence Review* **33**(1-2), 1–39.

Sesmero, M., Ledezma, A. and Sanchis, A. (2015), 'Generating ensembles of heterogeneous classifiers using stacked generalization', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **5**(1), 21–34.

Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. R. (2014), 'Dropout: a simple way to prevent neural networks from overfitting', *The journal of machine learning research* **15**(1), 1929–1958.

Suzuki, K. (2013), *Artificial neural networks: Architectures and applications*, BoD–Books on Demand.

Wu, X. and Kumar, V. (2009), *The top ten algorithms in data mining*, CRC press.