Quick and Robust Feature Selection: the Strength of Energy-efficient Sparse Training for Autoencoders (Extended Abstract)⋆ ⋆⋆

Zahra Atashgahi[1], Ghada Sokar[2], Tim van der Lee[3], Elena Mocanu[1], Decebal Constantin Mocanu[1,2], Raymond Veldhuis[1], and Mykola Pechenizkiy[2,4]

[1] EEMCS Faculty, University of Twente, the Netherland
[2] M&CS Faculty, Eindhoven University of Technology, the Netherlands
[3] EE Faculty, Eindhoven University of Technology, the Netherlands
[4] Faculty of Information Technology, University of Jyväskylä, Finland

1 Introduction

Feature selection, which identifies the most relevant and informative attributes of a dataset, has been introduced to address the challenges raised by the emerge of high-dimensional data [3]. Most existing feature selection methods are computationally inefficient; inefficient algorithms lead to high energy consumption, which is not desirable for devices with limited computational and energy resources. In [1], a novel and flexible method for unsupervised feature selection is proposed. This method, named "QuickSelection"[5], introduces the strength of the neuron in sparse neural networks as a criterion to measure the feature importance. When tested on several benchmark datasets, the proposed method is able to achieve the best trade-off of classification and clustering accuracy, running time, and memory usage, among widely used approaches for feature selection.

2 Proposed Method

QuickSelection is capable of selecting the most informative attributes of the data efficiently. The overview of the method is presented in Figure 1. This algorithm consists of two main phases: **1. Training Sparse DAE.** We use the ability of Denoising autoencoders (DAEs) to learn a robust representation of the data and select the most important features. We introduce for the first time sparse training in the world of denoising autoencoders, and we name the newly introduced model sparse denoising autoencoder (sparse DAE). We train the sparse DAE with the Sparse Evolutionary Training (SET) [4] algorithm to keep the number of parameters low during the training. **2. Feature Selection.** In the second phase, we use the trained network to derive the hierarchical importance of the

⋆ This research has been partly funded by the NWO EDIC project.
⋆⋆ The full paper corresponding to this abstract has been accepted for publication in the Machine Learning Journal (ECML-PKDD 2022 Journal Track)
[5] The code is available at: `https://github.com/zahraatashgahi/QuickSelection`

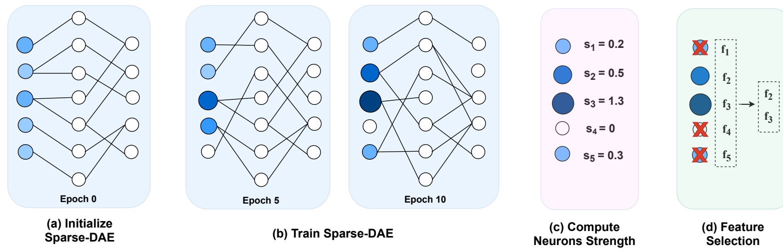

Fig. 1. A high-level overview of the proposed method, "QuickSelection".

features. We select the most important features of the data based on the weights of their corresponding input neurons of the trained sparse DAE. Inspired by node strength in graph theory [2], we determine the importance of each neuron based on its *strength*. We estimate the strength of each neuron by the summation of absolute weights of its outgoing connections. We select the features corresponding to the neurons with K largest strength values as the K important features.

Results. In order to verify the validity of our proposed method, we carry out several experiments to measure its performance in terms of the running time, memory requirement, clustering accuracy, and classification accuracy. To analyze the trade-off of the methods between accuracy and efficiency, we compute a ranking-based score (Figure 2): on several datasets and for several values of K, we rank the methods based on the aforementioned metrics. Then, we give a score of 1 to the best and second-best performers. As can be seen in Figure 2, our proposed method can achieve the best trade-off between accuracy, running time, and memory usage, among all the considered methods.

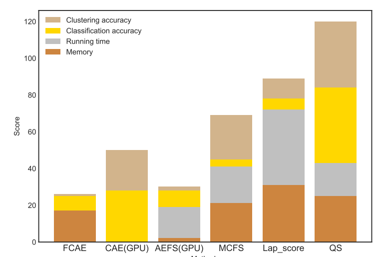

Fig. 2. Feature selection comparison using a ranking-based score.

3 Concluding Remarks

In this paper [1], a novel method (QuickSelection) for energy-efficient unsupervised feature selection has been proposed. We introduced *neuron strength* as a metric to measure the importance of the input neurons in a sparse neural network. By adopting this metric in a sparsely connected denoising autoencoder, we are able to derive the importance of all input features simultaneously. By using sparse layers instead of dense ones from the beginning, the number of parameters drops significantly. As a result, QuickSelection requires much less memory, computational resources, and energy consumption than its competitors. This will not only save the energy costs of processing high-dimensional data but also will ease the challenges of high energy consumption imposed on the environment.

References

1. Atashgahi, Z., Sokar, G., van der Lee, T., Mocanu, E., Mocanu, D.C., Veldhuis, R., Pechenizkiy, M.: Quick and robust feature selection: the strength of energy-efficient sparse training for autoencoders. Accepted at Machine Learning Journal (ECML-PKDD 2022 Journal Track) preprint arXiv:2012.00560 (2020)
2. Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A.: The architecture of complex weighted networks. Proceedings of the National Academy of Sciences **101**(11), 3747–3752 (2004)
3. Chandrashekar, G., Sahin, F.: A survey on feature selection methods. Computers & Electrical Engineering **40**(1), 16–28 (2014)
4. Mocanu, D.C., Mocanu, E., Stone, P., Nguyen, P.H., Gibescu, M., Liotta, A.: Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. Nature Communications **9**(1), 2383 (2018)