Quick and Robust Feature Selection: the Strength of Energy-efficient Sparse Training for Autoencoders

Zahra Atashgahi*[1] (z.atashgahi@utwente.nl), Ghada Sokar[2], Time Van Der Lee [3], Decebal Constantin Mocanu[1,2], Elena Mocanu[1], Raymond Veldhuis[1], Mykola Pechenizkiy[2]

[1] Department of Electrical Engineering, Mathematics and Computer Science (EEMCS), University of Twente, The Netherlands, [2] Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands,
[3] Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands

1. Introduction

Major complications arise from the recent increase in the amount of high-dimensional data, including high computational costs and memory requirements. Feature selection, which identifies the most relevant and informative attributes of a dataset, has been introduced as a solution to this problem [1]. Most of the 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. We present a novel feature selection method, named **QuickSelection**, which introduces the strength of the neuron in sparse neural networks as a criterion to measure the feature importance. This criterion, blended with sparsely connected denoising autoencoders trained with the sparse evolutionary training procedure, derives the importance of all input features simultaneously. The corresponding paper is available online on arxiv [2].

2. Methodology

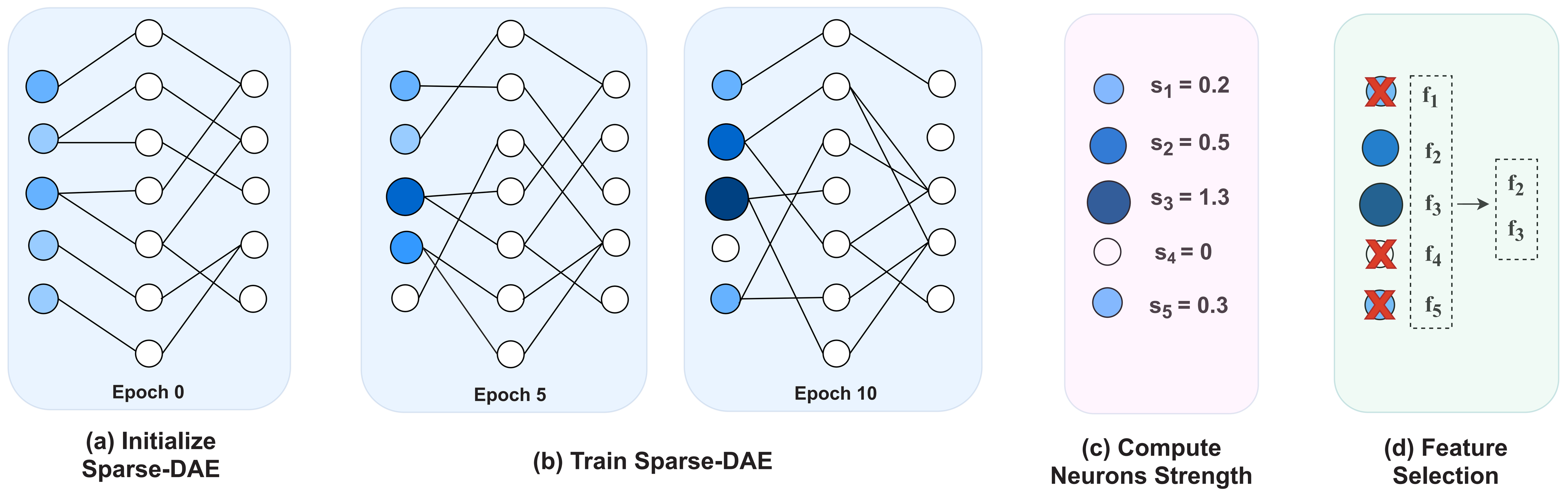

Figure 1. A high-level overview of the proposed method, **QuickSelection**.

- Training. We train a sparse denoising autoencoder (DAE) with sparse evolutionary training (SET) [3].
- Neuron Strength. We determine the importance of the neurons based on their **strength**.
- We estimate the strength as $s_i = \sum_{j=1}^{n^1} |W_{ij}^1|$, where n^1 is the number of neurons of the first hidden layer, and W_{ij}^1 denotes the weight of connection linking input neuron i to hidden neuron j.
- Feature Selection. We select the features corresponding to the neurons with K largest strength values.

3. Evaluation

To evaluate the methods, we compute a ranking-based score:

- On several dataset and for several value of K, we rank the methods based on the running time, memory requirement, clustering accuracy, and classification accuracy
- Give a score of 1 to the best and second-best performer
- The summation of scores are shown in Figure 2.
- **Our proposed method is able to achieve the best trade-off between accuracy, running time, and memory usage, among all these methods.**

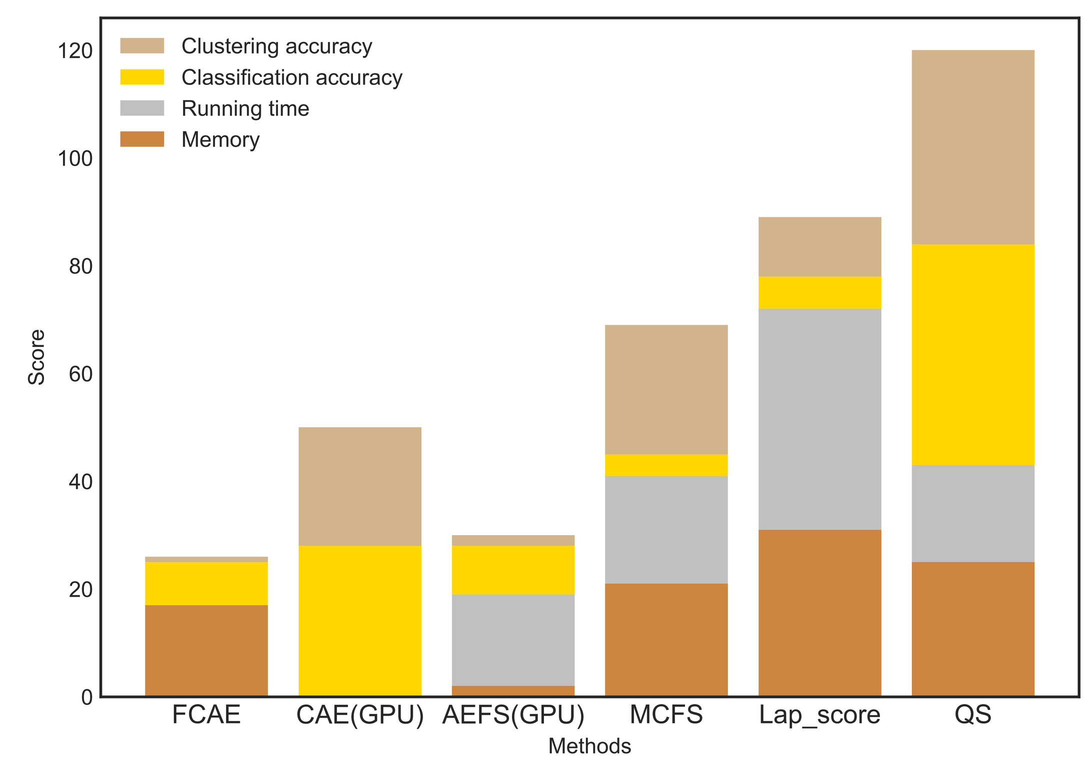

Figure 2. Feature selection comparison using a ranking-based score.

4. Visualization of Neurons's Strength

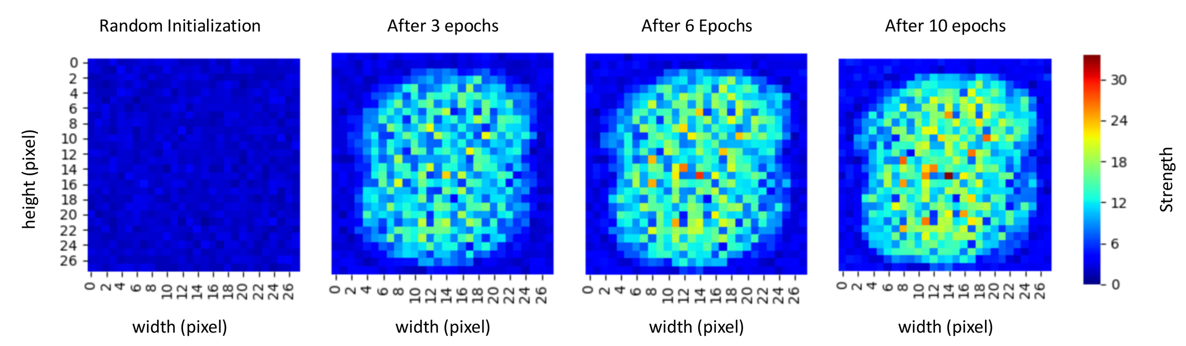

Figure 3. Neuron's strength on the MNIST dataset

- At the beginning of training: All the neurons have small strength due to the random initialization of each weight to a small value.
- During the network evolution: Stronger connections are linked to important features gradually, and therefore, the strength of important neurons/features increases.
- After 10 epochs: As can be seen in Figure 3, the neurons in the center of the map become stronger. This pattern is similar to the pattern of MNIST data in which most of the digits appear in the middle of the picture.

5. Conclusions

- Efficient Feature Selection. We show that feature selection can be performed using neural networks efficiently in terms of computational cost and memory requirement.
- Reducing Energy Cost. Efficient feature selection can pave the way for reducing the ever-increasing computational costs of deep learning models. 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.

6. References

[1] Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28, 2014.

[2] Zahra Atashgahi, Ghada Sokar, Tim van der Lee, Elena Mocanu, Decebal Constantin Mocanu, Raymond Veldhuis, and Mykola Pechenizkiy. Quick and robust feature selection: the strength of energy-efficient sparse training for autoencoders. *arXiv preprint arXiv:2012.00560*, 2020.

[3] Decebal Constantin Mocanu, Elena Mocanu, Peter Stone, Phuong H Nguyen, Madeleine Gibescu, and Antonio Liotta. Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. *Nature communications*, 9(1):2383, 2018.

UNIVERSITY OF TWENTE. TU/e SNN 2021 zgt PHILIPS NWO