Cost-effective Artificial Neural Networks

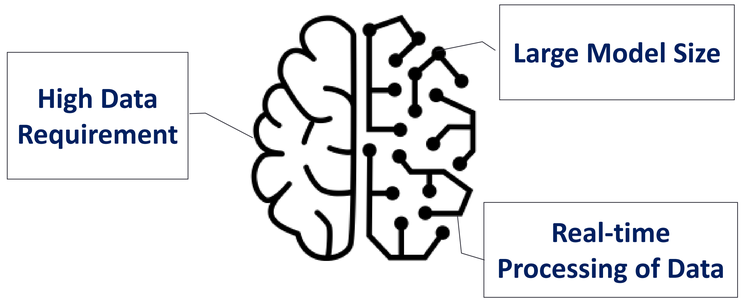
Zahra Atashgahi*[1] (z.atashgahi@utwente.nl), Decebal Constantin Mocanu [1, 2], Raymond Veldhuis [1], Mykola Pechenizkiy [2]
1 Department of Electrical Engineering, Mathematics and Computer Science (EEMCS), University of Twente, The Netherlands,
2 Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands,

Abstract

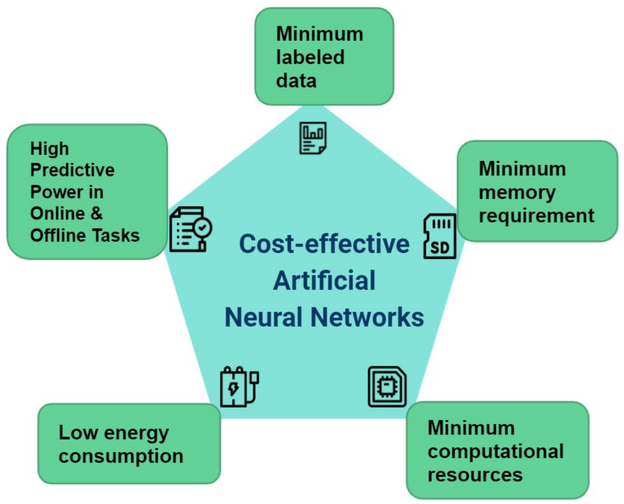
Artificial neural networks (ANNs) have gained huge attention over the last few years due to their promising results in a large variety of tasks. However, deep neural networks (DNNs) require plenty of annotated data and are recognized as being computationally demanding. Therefore, deep learning models are not well-suited to applications with limited computational resources, battery life, and labeled instances. Current solutions to reduce computation and annotation costs mostly focus on inference efficiency, while being resource-intensive during training.

My Ph.D. research aims to address these challenges by developing cost-effective neural networks that can achieve decent performance on various complex tasks using minimum computational resources and labeled samples, both during training and inference of the network.

Major challenges of Deep Learning



Research Goals

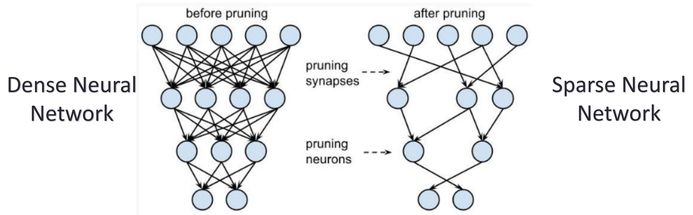


Research Question 1

How to maximize the learning capacity of a neural network with a *limited number of parameters*?

Focus on sparse neural networks and conceiving new algorithms to generate these sparse structures

Investigate whether sparse neural networks can be exploited to solve different problems more efficiently

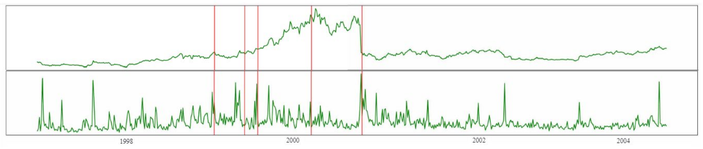


Research Question 2

How can we use neural networks to learn from data in a *memory-free manner continuously*?

Develop methods to prevent catastrophic forgetting

Detect event from data streams



Change-point Detection on Apple Stock Prices Dataset [2]

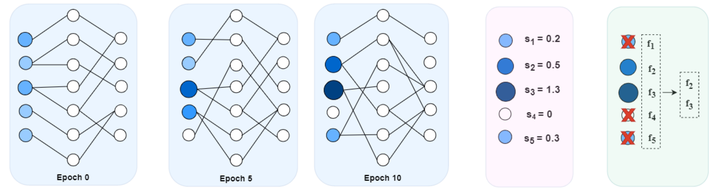
Research Question 3

How can we improve the performance of neural networks in *the tasks where labeled data is scarce, using minimum computational resources*?

Bridging the gap between human and machine intelligence

Benefit from unlabeled dataset either from the same domain or a related domain

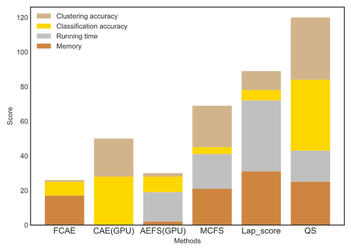
Quick and Robust Feature Selection

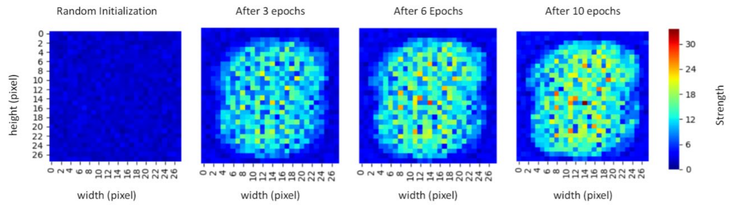


Overview of the proposed feature selection method, named QuickSelection [1]

Results

Evaluation of Feature selection
- Accuracy
- Running time
- Memory requirement





Neuron Strength Visualization on MNIST

Conclusion

Cost-effective neural networks can pave the way for reducing the ever-increasing computational costs of deep learning models. This will not only minimize the computation costs of processing data but also will ease the challenges of high energy consumption imposed on the environment by deep learning models.

References

[1] Atashgahi, Zahra, et al. "Quick and Robust Feature Selection: the Strength of Energy-efficient Sparse Training for Autoencoders." Accepted at the Machine Learning Journal (ECMLPKDD Journal Track), (2021).

[2] Atashgahi, Zahra, et al. "Unsupervised Online Memory-free Change-point Detection using an Ensemble of LSTM-Autoencoder-based Neural Networks." Accepted as an extended abstract to ACM womENcourage 2021.