Self-Organizing Incremental Neural Networks
for Continual Learning

Chayut Wiwatcharakoses and **Daniel Berrar**

Data Science Laboratory, School of Engineering, Tokyo Institute of Technology, Japan
wiwatcharakoses.c.aa@m.titech.ac.jp, daniel.berrar@ict.e.titech.ac.jp

Abstract

Continual learning systems can adapt to new tasks, changes in data distributions, and new information that becomes incrementally available over time. The key challenge for such systems is how to mitigate catastrophic forgetting, i.e., how to prevent the loss of previously learned knowledge when new tasks need to be solved. In our research, we investigate self-organizing incremental neural networks (SOINN) for continual learning from both stationary and non-stationary data. We have developed a new algorithm, SOINN+, that learns to forget irrelevant nodes and edges and is robust to noise.

1 Introduction

Continual learning (or *lifelong learning*) denotes the field of machine learning that focuses on computational systems capable of adapting to new tasks and learning from data that become incrementally available over time [Parisi *et al.*, 2019]. A key characteristic of such systems is the ability to acquire new information without compromising previously learned knowledge, i.e., without *catastrophic forgetting* [McCloskey and Cohen, 1989], which is a major challenge even for state-of-the-art deep neural networks. Connectionist systems have to be sufficiently plastic to adapt to new input from non-stationary distributions, but on the other hand, too much adaptability can entail a loss of previously learned knowledge. This trade-off is known as the *stability-plasticity dilemma* [Ditzler *et al.*, 2015].

Humans are lifelong learners because they learn incrementally by fine-tuning their knowledge without completely forgetting consolidated knowledge. One of the hallmarks of human intelligence is the ability to use knowledge from one domain to solve problems in another domain in a process called *transfer learning*. By contrast, models based on neural networks are usually geared towards one specific task, for example, supervised learning from images with an *a priori* known number of classes. Such models cannot easily be extended to accommodate new classes, or to adapt to concept drifts due to changes in the underlying data distributions, let alone apply their knowledge to solve entirely new tasks.

2 Research Background

Continual learning represents a grand challenge for machine learning. One approach to mitigating catastrophic forgetting is a complete retraining from scratch; however, this approach is inefficient or even impossible for big data streams, or when the old training data are no longer available for rehearsal. In our research, we focus on *self-organizing incremental neural networks* (SOINN) [Shen and Hasegawa, 2006], which encompass a family of neural networks that learn a topological mapping of the input data by competitive Hebbian learning. SOINN have been used for a variety of tasks, including clustering [Shen and Hasegawa, 2006], classification [Shen and Hasegawa, 2008], and various applications in robotics [Okada *et al.*, 2010]. These studies have shown that SOINN can adapt well to changes in the data distributions without detrimentally affecting existing network structures.

3 Research Problem

Our research focuses on designing and implementing self-organizing incremental neural networks for learning from both stationary and non-stationary data. We consider applications where the data volume is too large for explicit storage, so each training case can be processed only once. In SOINN, a new training case is either merged with an existing node or added as a new node, depending on a similarity measure. Effective capacity control becomes an issue when SOINN is presented with training cases from new classes because these cases tend to be dissimilar to existing nodes and are therefore added as new nodes. Current SOINN address this problem by deleting nodes and edges at fixed, user-defined intervals; however, this approach entails an unnatural network evolution, where batches of nodes are deleted periodically. "Forgetting" is not an intrinsic part of the learning process in current SOINN. Another problem is how similarity between nodes and new training cases is measured. Current SOINN use the Euclidean distance, which is not suitable for high-dimensional feature spaces. Finally, it is unclear if (and how) SOINN can be used for transfer learning, and how we can effectively measure performance across different application domains. Our research will address these open problems.
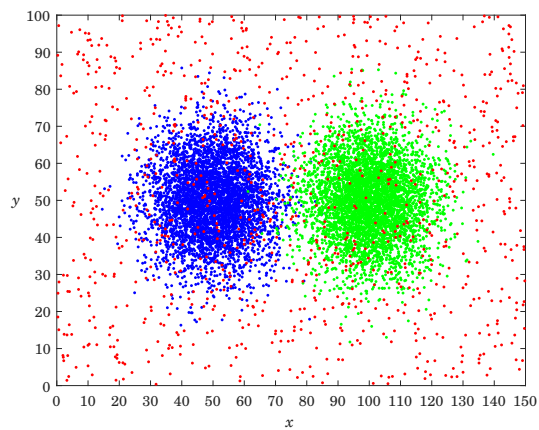
Figure 1: Training set of 9000 cases from two different classes (blue and green dots) and 1000 instances of noise (red dots).
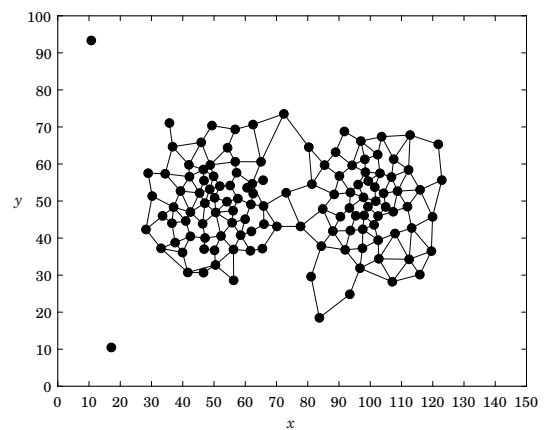

Figure 2: SOINN+ network with 125 nodes after processing the training data in Figure 1.

4 Research Contributions

We are currently working on a novel self-organizing incremental neural network, called SOINN+, that addresses several shortcomings of previous SOINN. Our method inherits some of the properties of the one-layer adjusted SOINN [Shen and Hasegawa, 2008]. However, SOINN+ is fundamentally different with respect to how nodes and edges are deleted. We consider the deletion of nodes as an intrinsic part of the learning process. To achieve this goal, we developed two new concepts: (1) the *trustworthiness* of a node reflects our degree of belief that a node represents a true signal and not noise; and (2) the *(un)utility* of a node, which is a measure of its (ir)relevance for the learning task at hand. Nodes that become untrustworthy or no longer useful are candidates for deletion. SOINN+ constantly monitors the state of its nodes and edges and then deletes them as necessary. As a result, SOINN+ becomes more resilient to noise and "learns to forget" more gracefully.

To illustrate the topological mapping by SOINN+, Figure 1 shows a training set of 4500 bivariate positive cases (blue dots), with $X_+, Y_+ \sim \mathcal{N}(50, 10)$ and 4500 bivariate negative cases (green dots), with $X_-, Y_- \sim \mathcal{N}(100, 10)$, plus 1000 cases of uniformly distributed noise (red dots). Figure 2 shows the resulting SOINN+ network after processing all 10 000 input cases. The final network consists of 128 nodes. The two unconnected nodes are noise and candidates for deletion when the training continues with new data.

The network in Figure 2 is a compact representation of the training data. It can be used for clustering and classification of new training cases. We are currently working on a nearest-neighbor classifier that uses the nodes as candidate prototypes. Our ongoing research has shown that SOINN+ can accommodate new classes without catastrophically interfering with the existing network structure.

5 Future Work

In our future work, we will investigate how our method can be improved. An open problem is the measure of similarity between new input cases and nodes. The network in Fig-

ure 2 is based on the Euclidean distance, which is appropriate for this example data set, but it is not suitable for high-dimensional data. Our next step is to investigate different distance metrics, notably the fractional distance, and analyze streams of high-dimensional data. Another goal is to extend the network by adding additional layers for the topological mapping, where each layer learns to represent a different hierarchical level of abstraction, similar to deep convolutional neural networks. As another direction for future work, we will investigate whether these extensions to SOINN+ can faciliate transfer learning.

References

[Ditzler *et al.*, 2015] Gregory Ditzler, Manuel Roveri, Cesare Alippi, and Robi Polikar. Learning in nonstationary environments: A survey. *IEEE Computational Intelligence Magazine*, 10(4):12–25, 2015.

[McCloskey and Cohen, 1989] Michael McCloskey and Neal J. Cohen. Catastrophic interference in connectionist networks: the sequential learning problem. *The Psychology of Learning and Motivation*, 24:104–169, 1989.

[Okada *et al.*, 2010] Shogo Okada, Yoichi Kobayashi, Satoshi Ishibashi, and Toyoaki Nishida. Incremental learning of gestures for human-robot interaction. *AI & Society*, 25(2):155–168, 2010.

[Parisi *et al.*, 2019] German I. Parisi, Ronald Kemker, Jose L. Part, Christopher Kanan, and Stefan Wermter. Continual lifelong learning with neural networks: A review. *Neural Networks*, 13:54–71, 2019.

[Shen and Hasegawa, 2006] Furao Shen and Osamu Hasegawa. An incremental network for on-line unsupervised classification and topology learning. *Neural Networks*, 19(1):90–106, 2006.

[Shen and Hasegawa, 2008] Furao Shen and Osamu Hasegawa. A fast nearest neighbor classifier based on self-organizing incremental neural network. *Neural Networks*, 21(10):1537–1547, 2008.