The nnlib2 library and nnlib2Rcpp R package for implementing neural networks

Vasilis N Nikolaidis

1 University of Peloponnesse

Summary

Artificial Neural Networks (ANN or NN) are computing models used in various data-driven applications. Such systems typically consist of a large number of processing elements (or nodes), usually organized in layers, which exchange data via weighted connections. An ever-increasing number of different neural network models have been proposed and used. Among the several factors differentiating each model are the network topology, the processing and training methods in nodes and connections, and the sequences utilized for transferring data to, within and from the model etc. The software presented here is a C++ library of classes and templates for implementing neural network components and models and an R package that allows users to instantiate and use such components from the R programming language.

Statement of need

A significant number of capable, flexible, high performance tools for NN are available today, including frameworks such as Tensorflow (Abadi et al., 2016) and Torch (Collobert et al., 2011), and related high level APIs including Keras (Chollet & others, 2015) and PyTorch (Paszke et al., 2019). Ready-to-use NN models are also provided by various machine learning platforms such as H2O (H2O.ai, 2020) or libraries, SNNS (Zell et al., 1994) and FANN (Nissen, 2003). Ready to use NNs are available in a large number of software packages (WEKA, SPSS, Matlab, NeuroSolutions, Noesis and many others). The R language (R Core Team, 2020), widely used for data analysis, includes package nnet (Venables & Ripley, 2002) in its standard packages, while a significant number of NN-related extension packages can be found in CRAN and other repositories. These include interfaces and bindings to aforementioned tools (keras (Chollet et al., 2017), torch (Falbel et al., 2020), rTorch (Reyes, 2020), h2o (LeDell & others, 2020), RSNNS (Bergmeir & Benítez, 2012)), as well as several other NN-specific packages that provide ready-to-use NN models (deepnet (Rong, 2014), deepNN (Taylor, 2020), RcppDL (Kou & Sugomori, 2014) etc.).

The nnlib2 library adds another tool for NN experimentation and implementation. It is a collection of C++ base classes and class-templates for defining NN parts, components and entire models. It contains base classes (and related functionality) for each part of a NN (processing nodes, connections, layers, groups of connections etc.) and for NN models that contain such parts. The library does not focus (in its current state) on high-performance computing but on being a versatile basis on which new NN parts and models can be defined with consistent, readable, maintainable and reusable code. Although the software does contain some ready-to-use NN models and their related reusable NN parts (currently for versions of Backpropagation, autoencoder, Learning Vector Quantization and Matrix Associative Memory), its main purpose is to aid the creation of new custom NN parts and the experimentation with them in arbitrary NN models and configurations. Overall the library (and related R package):
1. Allows the implementation of NN parts and models from reusable components that follow a common interface thus can easily be modified, combined with each other etc., making them suitable for educational or experimentation purposes.
2. Provides a basis for implementation of different NN parts and models, including classic NNs.
3. Is versatile enough for experimentation with new, custom components, configurations and models.
4. Has no dependencies with other software; to build or employ the nnlib2 parts and models only a standard C++ compiler is required, and the produced models are standalone, lightweight, highly portable and can be embedded in any C++ application.
5. Allows for the exact same NN parts (processing nodes, connections, layers, groups of connections etc.) defined using nnlib2 to be employed, combined, manipulated and monitored in R, via package nnlib2Rcpp which includes the entire nnlib2 library and also provides supporting R functions and modules for this purpose. Thus, new NN components can be developed entirely using R-related tools (such as Rtools and RStudio), and then be used in R or transferred to a pure C++ application if needed.

Discussion

As implied earlier, the nnlib2 library may appeal to NN students and experimenters who seek a basis for implementing various NN models in C++ and take advantage of the versatility and direct control of the model that this approach allows. Furthermore, the library lets new NN definitions be written in code that is comparatively simple and self-explanatory; most of the provided base classes and templates correspond to typical NN components (processing nodes, layers of nodes, connections, sets of connections between layers, entire neural networks etc.) and closely follow the conceptual model often found in related bibliography such as (McCord-Nelson & Illingworth, 1991; Pao, 1989; Simpson, 1991; Theodoridis & Koutroumbas, 2008) and elsewhere. Typical component functionality and features are provided in these base definitions, while diverse, model-specific components are to be implemented in custom sub-classes. By following a common interface, components (or even entire NN models) can be reused in other models. Finally, the library presents a high degree of target and compiler independence, with early versions used in various solutions targeting different operating systems, mainly for research purposes, for example in (Vasiliou Nikolaidis, 1999; V. N. Nikolaidis et al., 2013; Philippidis et al., 1999).

To take advantage of the R ecosystem, improve its usability and widen its audience, the nnlib2 library has also been integrated into an R package (a process supported by Rcpp (Eddelbuettel et al., 2020)). The package, named nnlib2Rcpp (available on CRAN (Vasiliou Nikolaidis, 2020b) and GitHub (Vasiliou Nikolaidis, 2020a)) includes the entire nnlib2 code as well as an additional R class (class NN) with methods that allow users to instantiate the nnlib2 NN components they define, combine them in custom NN configurations and generally control, monitor and employ them from R. Finally, the package provides a small collection of predefined NN components and complete ready-to-use NN models (for versions of Back Propagation, autoencoder, supervised/unsupervised Learning Vector Quantization and Matrix Associative Memory neural networks).

References


