



**Data
Schools**

Teaching Artificial Neural Networks

August, 2020

CODATA RDA Data Science School

Teaching ANN



1. Please register in google docs - the link will appear in chat
2. menti interactive session to establish
 1. Who are you teaching?
 2. What are you teaching? Syllabus/learning outcomes
 3. How are you going to do it?
 1. Which language?
 2. Which neural network package (if any)?
 3. What data to use?
3. Discussion in breakout rooms on Data, Language and Package. Appoint scribe (use same link) and rapporteur
4. Report back



Which language?

First priority: a language your students know.

Second priority: a language you know.

You can't go wrong with R or Python

C is fine. So is Fortran (though horrible in many ways)

C++ and C# enable you to bring in neat Object-Oriented code. So does Java.

Matlab is OK if you are forced to use it (but somebody has to pay for it).

I have not tried LabView. COBOL is dead, thank goodness. Excel is not a language.

Which package (or write your own)

Decide on language first

Weigh up reasons for and against getting them to write their own Neural Net code (driving lessons analogy)

- The coding is easier than people think
- They will learn more about what's really happening
- They'll get a big kick out of writing their own code
- It will not be so smooth or performant or complete
- It does take a little longer to get results

If you decide to use a package, consult google. Whatever the language, there are lots of good ones out there. Don't just pick the first one, try two or three.

Data for your students to work on

Try and make it relevant for your students

Artificial data is OK

How many input variables? 2-10

How many output variables? 1

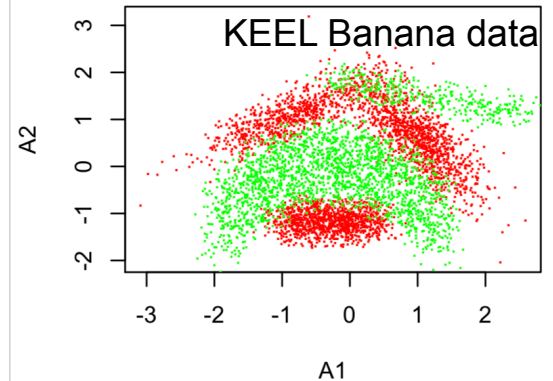
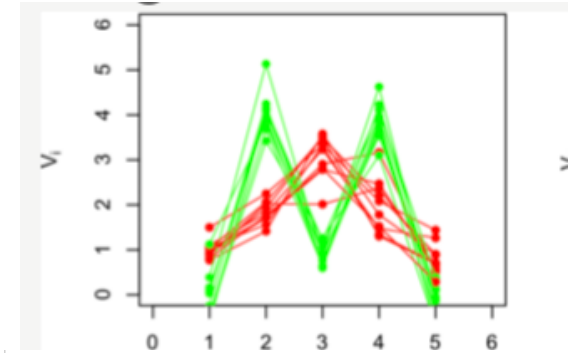
Data values:

- Real numbers are good.
- Integers are OK.
- Boolean is OK. (0 or 1)
- Ordinal is OK.
- Categorical is probably not. Re-cast 1...N categorical as N booleans

For examples and ideas see <https://sci2s.ugr.es/keel/datasets.php>

76 lovely datasets to play with

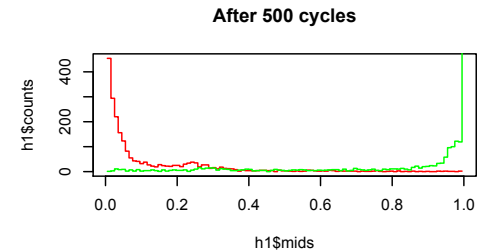
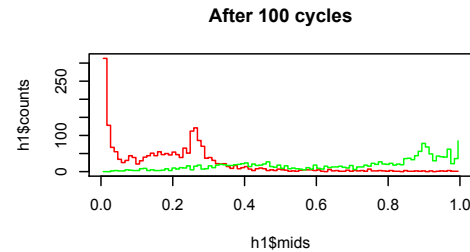
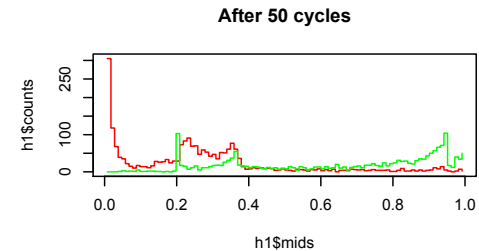
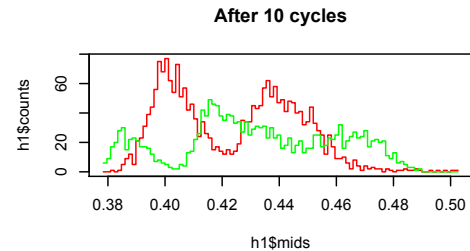
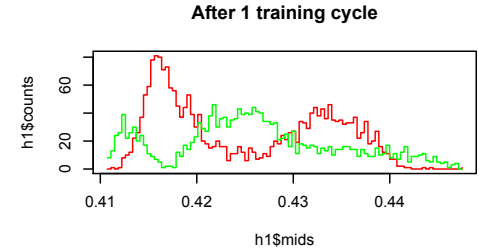
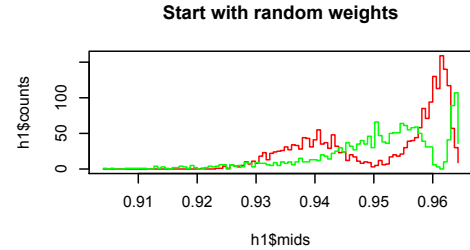
Trieste Camel-Dromedary data



Some results

Shows the output on the KEEL banana dataset (5300 entries) with the standard ANN program I wrote for the Trieste school = 2-9-9-1 topology, learning parameter $\alpha=0.05$

A few lines needed to read datafile.
Complete program available on google drive - same folder as talks





Data Schools

Let's get started!

Good luck to everyone.