An introduction to statistical computing in Scala

Roxana Tesileanu

INCDS 2017

Table of Contents

[Getting started 1](#__RefHeading___Toc214_241067100)

# Getting started

This work gathers together my learning steps towards using Scala as an environment for statistical computing. My background in statistics and linear algebra helped a lot, so it does assume the interested people passed their statistics courses. The big advantage of Scala is it allows you to express your thoughts in a natural manner, so having a clear idea of what you want to implement really helps a lot.

In order to begin with statistical computing in Scala, you need to install the appropriate tools on your computer. Because different operating systems require different installing procedures, I will let you check the websites of the tools used and extract the information needed to install them on your computer. So, visit the website of the Scala language and of the sbt tool and follow the instructions found there. If you encounter big problems you can grasp to the introduction into the Scala language of Jason Swartz1 and come back later. This is a normal iterative approach used by anyone who wants to get into a new field. The main aim of this work is to serve as documentation for the scalaML package developed by me at INCDS. The GitHub repository of the project can be found at the following link:

<https://github.com/RoxanaTesileanu/multivariate_analyses/tree/master/DeepLearning/src/main/scala/com/incds/scalaML>.

As mentioned previously, in order to use the scalaML package, you need to install Scala and sbt. Supposed you passed through these initial steps successfully, you can create a new Scala project using sbt.