

Escola Universitària Politécnica de Mataró

Centre adscrit a:



**UNIVERSITAT POLITÈCNICA
DE CATALUNYA**

Grau en enginyeria informàtica

Machine Learning with H2o and R: A Big Data problem.

Appendices

**Ricard Gardella Garcia
TUTOR: Dr. Xavier Font**

Spring 2017



**TecnoCampus
Mataró-Maresme**

Index.

Appendix one. DVD Content.....	1
--------------------------------	---

Appendix one. DVD Content

- Documentation: Memory, Feasibility Study and Appendixes.
- Models: The Machine Learning models.
- Project:
 - GardellaRicardMarkdown.html: The report given by the RMarkdown.
 - GardellaRicardMarkdown.rmd: The rmd file used to create the report. Contains all the code with explanations of every step and can be executed by parts.
 - FunctionsTFG: Contain functions separated of the general code.
 - GeneralCode.R: Contains all the code. The same code that appear in the RMD file.
 - SessionMarkdown.RData: A session stored necessary to execute the rmd.
- Data:
 - DataHolidays.csv: File with all the flights in holidays for year 2006 to 2008
 - Airports.csv: File that contains all the data of all the airports in the USA.

It is important to have in mind that the SessionMarkdown.Rdata is necessary in order to execute the rmd without modifying the file. The GeneralCode.R contains all the code.

The data files than contain all the flights of a year, can be found here: <http://stat-computing.org/dataexpo/2009/the-data.html>

