DATA ANALYTICS AND MACHINE LEARNING WITH R

SYLLABUS

LUIS GUSTAVO NARDIN

INTERNET TECHNOLOGY

BRANDENBURG UNIVERSITY OF TECHNOLOGY

SUMMARY

- Course Information
- Learning Objectives
- Overview of Tools

COURSE INFORMATION

- **Effort and ECTS Points**
  Seminar: 2 SWS
- **Timetable**
  Tuesday, 15:30 - 17:00, room VG 1C/1.21

ORGANIZATIONAL ISSUES

- Course language: English
- Register for the course at our LMS **https://hydrogen.informatik.tu-cottbus.de/moodle/** using your BTU login (user name and password) to the course "2018 Data Analysis and Machine Learning with R" using the enrollment key **DML18**.
- Please **upload a picture of yourself** – it helps to improve communication via the LMS!
- Ask all **questions** about the course in the LMS discussion forum!
- Get all **learning materials** (including the lecture slides) from the LMS

COURSE FORMAT

This course is composed of a set of lectures and short assignments, and one final project.

- **Lectures** - an overview of the tools and concepts in data analysis and machine learning
- **Short Assignments** - an introductory hands-on experience with the tools and concepts
- **Final Project** - to demonstrate the use of the tools and concepts learned in a more elaborated project

GRADING

- 75% for final project
  P + 2 R + 3 C
  - S - Project code
  - R - Technical report
  - P - Presentation
- 25% for short assignments

CODE OF CONDUCT

- We encourage you to discuss questions about the assignments with each other
- BUT: **NEVER** show or give your solution code to others!
- You may explain your solution approach to other students, but do not show or give them your code!
- If yo do not comply with this rule, you will have a problem!

LEARNING OBJECTIVES

- how to manipulate data and perform statistical analysis using R
- how to analyze and communicate data analysis results using R
- how to discover patterns and build predictive models using machine learning techniques in R
- how to build interactive web apps straight from R

TOPICS

1. Introduction to Data Science
2. Essential of R Programming
3. Exploratory Data Analysis
4. Inferential Analysis
5. Machine Learning
6. Interactive WebApps

OVERVIEW OF TOOLS

- R
- RStudio
- Tensorflow for R
- Shiny

R

https://www.r-project.org

- Integrated suite of software facilities for data manipulation, calculation and graphical display

RSTUDIO IDE

https://www.rstudio.com/

RStudio is an integrated development environment (IDE) for R

- Open source software
- Syntax highlighting, code completion, and smart indentation
- Execute R code directly from the source editor
- Interactive debugger to diagnose and fix errors quickly
- Integrate several tools for use with R into a single environment: R help and documentation, Git and Subversion, and authoring tools
- Supports interactive graphics with Shiny and ggvis

TENSORFLOW FOR R

https://tensorflow.rstudio.com/

- **TensorFlow™** is an open source framework for Machine Learning
  - perform numerical computation using data flow graphs
  - nodes represent mathematical operations
  - edges represent the multidimensional data arrays communicated between them
- Tensorflow for R provides several high-level APIs to access different functions of Tensorflow from R

SHINY

https://www.rstudio.com/products/shiny/

Shiny is an open source R package that provides an elegant and powerful web framework for building web applications using R.