

DSCI351-351m-451: Class 01a, (CWRU, Pitt, UCF, UTRGV)

Profs: R. H. French, L. S. Bruckman, P. Leu, K. Davis, S. Cirlos

TAs: W. Oltjen, K. Hernandez, M. Li, M. Li, D. Colvin

21 March, 2023

Contents

7.2.1.0.1	Reading, Lab Exercises, SemProjects	1
7.2.1.0.2	Textbooks	1
7.2.1.0.3	Tidyverse Cheatsheets, Functions and Reading Your Code	2
7.2.1.0.4	Syllabus	2
7.2.1.1	Links	4

7.2.1.0.1 Reading, Lab Exercises, SemProjects

- Readings:
 - For today: ISLR9, (R4DS26-30)
 - For next class: ISLR10
- Laboratory Exercises:
 - LE4: Is due Thursday March 9th
 - LE5: Given out next Thursday
- Office Hours: (Class Canvas Calendar for Zoom Link)
 - Wednesdays @ 4:00 PM to 5:00 PM
 - Saturdays @ 3:00 PM to 4:00 PM
 - **Office Hours are on Zoom, and recorded**
- Semester Projects
 - Office Hours for SemProjs: Mondays at 4pm on Zoom
 - DSCI 453 Students Biweekly Updates Due
 - * Update # is Due ** **
 - DSCI 453 Students
 - * Next Report Out # is Due ** **
 - All DSCI 353/353M/453, E1453/2453 Students:
 - * **Peer Grading of Report Out #1 is Due Thursday March 2nd**
 - Exams
 - * MidTerm: **Thursday March 9th**, in class or remote, 11:30 - 12:45 PM
 - **CWRU Spring Break is March 13th to March 17, so NO CLASS**
 - * Final: **Thursday May 4th**, 2023, 12:00PM - 3:00PM, Nord 356 or remote

7.2.1.0.2 Textbooks

- Text Books for DSCI353/353M/453
 - [R4DS: Wickham: R for Data Science](#)
 - [ISLR: Intro to Statistical Learning with R, 2nd Ed.](#)
 - DLwR: Deep Learning with R, Chollet, Allaire,
 - [DLGB: Deep Learning, Goodfellow, Bengio, Courville](#)

- Magazine Articles about Deep Learning
 - DL1 to DL13 are “Deep Learning” articles in 3-readings/2-articles/
- Books from DSCI351/351M/451
 - Peng: R Programming for Data Science
 - Peng: Exploratory Data Analysis with R
 - Open Intro Stats, v4
 - R4DS: Wickham: R for Data Science

7.2.1.0.3 Tidyverse Cheatsheets, Functions and Reading Your Code

- Look at the Tidyverse Cheatsheet
 - **Tidyverse For Beginners Cheatsheet**
 - * In the Git/20s-dsci353-353m-453-prof/3-readings/3-CheatSheets/ folder
 - **Data Wrangling with dplyr and tidyr Cheatsheet**

Tidyverse Functions & Conventions

- The pipe operator `%>%`
- Use `dplyr::filter()` to subset data row-wise.
- Use `dplyr::arrange()` to sort the observations in a data frame
- Use `dplyr::mutate()` to update or create new columns of a data frame
- Use `dplyr::summarize()` to turn many observations into a single data point
- Use `dplyr::arrange()` to change the ordering of the rows of a data frame
- Use `dplyr::select()` to choose variables from a tibble,
 - * keeps only variables you mention
- Use `dplyr::rename()` keeps all the variables and renames variables
 - * `rename(iris, petal_length = Petal.Length)`
- These can be combined using `dplyr::group_by()`
 - * which lets you perform operations “by group”.
- The `%in%` matches conditions provided by a vector using the `c()` function
- The **forcats** package has tidyverse functions
 - * for factors (categorical variables)
- The **readr** package has tidyverse functions
 - * to read_..., melt_..., col_..., parse_... data and objects

Reading Your Code: Whenever you see

- The assignment operator `<-`, think “**gets**”
- The pipe operator, `%>%`, think “**then**”

7.2.1.0.4 Syllabus

Day:Date	Foundation	Practicum	Readings(optional)	Due(optional)
w01a:Tu:1/17/23 w01b:Th:1/19/23	Markov Cluster Stat. Learning, Approach	R, Rstudio IDE, Git Bash, Git, Class Repo	ISLR1,2 (R4DS-1-3)	(LE0)
w02a:Tu:1/24/23 w02b:Th:1/26/23	Lin. Regr. Bias-Var. Train/Test, Bias vs. Vari.	SemProjs; Regr. Ovrw Tidyverse Review	ISLR3,(R4DS-4-6) DL01 DL02 (R4DS-7,8)	(LE0:Due) LE1
w02Pr:Fr:1/27/23	ADD DROP	DEADLINE		453 Update 1
w03a:Tu:1/31/23	Logistic Regr. Classif	Pred. Analytics, Regr.	DL03,ISLR4	
w03b:Th:2/2/23 w03Sa:2/4/23	LDA/QDA	ggPlot2, Code Expect.	DL04, DL05	LE1:Due, LE2 LE1:Due
w04a:Tu:2/7/23 w04b:Th:2/9/23	Resample Cross-Valid. DL, ML Overview	ggplot Multilevel Mod.	ISLR5 ISLR6 (R4DS9-16)	
w04Pr:Fr:2/10/23				453 Update 2
w05a:Tu:2/14/23	Resampling: Bootstrap	Bootstrap Mixed Effects	DL2R1, DL06,07	LE2:Due, LE3
w05b:Th:2/16/23 w05Pr:Fr:2/17/23	Subset Selec., Shrink.	Dim. Red. PCA	DLwR2	453 Rep. Out 1
w06a:Tu:2/21/23 w06b:Th:2/23/23	ML with NNs Beyond Linear Modls	ggplot, clustering Feature Select., Caret	DLwR3 ISLR7 (R4DS22-25)	LE3:Due, LE4
w06Pr:Fr:2/24/23				453 Update 3
w07a:Tu:2/28/23	Dec. Trees, Rand. Forest	Tidy Modeling	ISLR8, DL08,09	
w07b:Th:3/2/23	MidTerm Review, SVM	SVM, SVR, ROC	ISLR9 (R4DS26-30)	Peer Review 1
w08a:Tu:3/7/23	ML Overview	, Keras/TF2, Torch	ISLR10.1,10.2	
w08b:Th:3/9/23 w08Pr:Fr:3/10/23	MIDTERM EXAM		DL10,11	LE4:Due LE5 453 Update 4
Tu:3/14/23 Th:3/16/23	SPRING SPRING	BREAK BREAK	ISLR10.3,10.4 ISLR10.5,10.6,	
w09a:Tu:3/21/23	Deep Learning	TF2 Keras Intro		ISLR10.7,10.8, DLwR3
w09b:Th:3/23/23 w09Pr:Fr:3/24/23	Computer Vision, CNN	CNN w/TF2, Overfit	DLwR4, DL12,13	453 Rep. Out 2
w10a:Tu:3/28/23	Deep Learn Intro	NN Types	DLwR5 Hinton ImageNet	
w10b:Th:3/30/23 w10Pr:Fr:3/31/23	DL CNN,RNN ImageNet	NN Types, CNN w/TF2		453 Upd.5 & PrRev 2
Sa:4/1/23				LE5:Due LE6
w11a:Tu:4/4/23	Fitting NNs	AUC,Prec,Recall Fruit		
w11b:Th:4/6/23	NLP, Graphs & ML		LeCun DL Rev. 2015	
w12a:Tu:4/11/23	Graphs & ML	NLP with sequences	DLwR6	
w12b:Th:4/13/23	NLP w attention	Graph Repr Proc Wrk-flw		LE6:Due LE7
w13a:Tu:4/18/23	DL Frameworks	Explaining DL w Lime		
w13b:Th:4/20/23 w13Pr:Fr:4/21/23	Linux Distros XGBoost	Explain Preds	Deep Dream	453 Rep. Out 3 Due
w14a:Tu:4/25/23 w14b:Th:4/27/23	Transformers Final Exam Review	Torch NN & DeepLearn		LE7:Due
w14Pr:Fr:4/28/23				Peer Rev 3 Due
	FINAL EXAM	Th. 5/4/23, 12-3pm	Nord 356 & Zoom	
	453 Final PDF Report	Fr. 4/29, 11:59pm		

Table 1: DSCI353-353M-453 Weekly Syllabus. R4DS-x.y, OISx.y, ISLRx.y, DLwRx.y, DLGBx.y refers to chapters and sections assigned as reading in our textbooks. DLx are deep learning articles. March 2, 2023

Figure 1: DSCI351-351M-451 Syllabus

7.2.1.1 Links