

5DATA002W machine Learning & Data Mining (2021)

Week No:	Week Starting	Lecture	Staff	Tutorials
1	18/01/2021	Machine Learning & Data Mining overview (Introduction)	VK	Provide information on the software needed for this module as well as on the on-line tutorial style to be adopted for this module.
2	25/01/2021	Ethics and Bias in Machine Learning & Data Pre-processing	VK	Introduction to R programming, simple exercises using R
3	01/02/2021	Clustering Methods (Partitioning & Hierarchical)		Practical exercises on data pre-processing (manually and via R)
4	08/02/2021	Neural Networks & Applications	VK	Practical exercises on clustering (manually and via R) 10/02/2021: CWK to be issued
5	15/02/2021	Association Rules (Market Basket Analysis)	VK	Practical exercises on Neural Networks using R
6	22/02/2021	Engagement Week – No Lecture & Tutorial		
7	01/03/2021	Naïve Bayes (NB) Classification	VK	Practical exercises on Association Rules (manually and via R)
8	08/03/2021	K-Nearest Neighbour algorithm	VK	Practical exercises on NB classifier (manually and via R)
9	15/03/2021	Decision Trees	VK	Practical exercises on K-NN classifier using R
10	22/03/2021	Support Vector Machines (SVM) & a short introduction to Ensemble Systems	VK	Practical exercises on Decision Trees (manually and via R)
11	29/03/2021	Review	VK	Practical exercises on SVM using R Final Clarifications for CW assessment
12	05/04/2021	On-line Test <i>(60% of the module total mark)</i>	VK	Support material for ensemble systems 27/04/2021: CWK to be submitted via BB

Module Staff: Dr Vassilis S. Kontogiannis (VK) - kodogiv@westminster.ac.uk

Essential reading list

- *Introduction to Data Mining*, Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, Pearson, 2nd edition, 2018.
- *Data Mining: Practical Machine Learning Tools and Techniques*, Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal, Morgan Kaufmann, 2016
- *Pattern Recognition and Machine Learning*, C. Bishop, Springer-Verlag, 2007
- *Machine Learning for Absolute Beginners: A Plain English Introduction*, Oliver Theobald, Scatterplot Press, 2017
- *Machine Learning with R: Expert techniques for predictive modeling*, Brett Lantz, Packt Publishing, 3rd edition, 2019

Further Reading

- *The R Book*, Michael J. Crawley, Wiley-Blackwell, 2012

- ***Practical Data Science with R***, Nina Zumel and John Mount, Manning Publications, 2nd Edition, 2019
- ***An Introduction to Statistical Learning: with Applications in R***, Gareth James, Daniela Witten, Trevor Hastie, Springer; 2013.
- ***R in Action, Data Analysis and Graphics with R***, Robert I. Kabacoff , Manning Publications, 2011

On-line References

- [***https://www.r-bloggers.com/***](https://www.r-bloggers.com/)
- [***https://www.kaggle.com/***](https://www.kaggle.com/)
- [***https://www.r-exercises.com/***](https://www.r-exercises.com/)
- [***http://www.rdatamining.com/***](http://www.rdatamining.com/)
- [***www.kdnuggets.com***](http://www.kdnuggets.com)

Assessment

Assessment Method	Weighting	Qualifying Mark	Qualifying Set	Assessment Type
CWK	40%	30%		Individual coursework
Lab controlled assessment (on-line test)	60%	30%		Lab Exam