STAT 406 – Winter 2019 Term 1 - Course Syllabus

Course Title and Description

- Elements of Statistical Learning
- Flexible, data-adaptive methods for regression and classification models; regression smoothers; penalty methods; assessing accuracy of prediction; model selection; robustness; classification and regression trees; nearest-neighbour methods; ensembles; unsupervised learning.
- Pre-reqs: One of STAT 306, CPSC 340.

Contact Information

- Matias Salibian-Barrera
- Online communication through Piazza course website.
- A weekly office hour will be announced.

Course Structure

- There are 3 weekly instructional meetings: two lecture meetings and a computer lab session. Times and locations are available on-line from the Student Service Centre.

Course Content

- Review of Linear Regression; Goodness of Fit vs Prediction error, Cross Validation; Correlated predictors, Feature selection, AIC; Ridge regression, LASSO, Elastic Net; Smoothers (Local regression, Splines); Curse of dimensionality, Regression Trees; Bagging; Classification, LDA, LQA, Logistic Regression; Trees, Ensembles, Bagging; Random Forests; Boosting; Unsupervised learning, K-means, model-based clustering; Hierarchical clustering, Principal Components, Multidimensional Scaling.
- Schedule:
    - Week 1: Review of Linear Regression
    - Week 2: Goodness of Fit vs Prediction error, Cross Validation
    - Week 3: Correlated predictors, Feature selection, AIC
    - Week 4: Ridge regression, LASSO, Elastic Net
    - Week 5: Elastic Net, Smoothers (Local regression, Splines)
    - Week 6: Curse of dimensionality, Regression Trees
    - Week 7: Bagging
    - Week 8: Classification, LDA, LQA, Logistic Regression
    - Week 9: Trees, Ensembles, Bagging
    - Week 10: Random Forests
    - Week 11: Boosting, Neural Networks?
    - Week 12: Unsupervised learning, K-means, model-based clustering
    - Week 13: Hierarchical clustering, Principal Components, Multidimensional Scaling

Course Learning Outcomes

- At the end of the course, the student will be able to: assess the prediction properties of the supervised learning methods covered in class; correctly use regularization to improve predictions from linear models, and also to identify important explanatory variables; explain the practical

difference between predictions obtained with parametric and non-parametric methods, and decide in specific applications which approach should be used; select and construct appropriate ensembles to obtain improved predictions in different contexts; select sensible clustering methods and correctly interpret their output; correctly utilize and interpret principal components and other dimension reduction techniques.

Course Activities and Assessment

- Students are expected to come to class and to the lab sessions. They are also expected to have completed the assigned pre-class reading from the reference books.
- There will be four in-class exams (worth 12.5% each, scheduled for weeks 4, 7, 10 and 13), plus a final exam (50% of the course grade).
- All course updates & announcements will be made via Canvas.

Required Materials

- Recommended reference books:
  - James, G., Witten, D., Hastie, T. and Tibshirani, R. An Introduction to Statistical Learning. 2013. Springer-Verlag New York
  - Hastie, T., Tibshirani, R. and Friedman, J. The Elements of Statistical Learning. 2009. Second Edition. Springer-Verlag New York
  - Venables, W.N. and Ripley, B.D. Modern Applied Statistics with S. 2002. Fourth edition, Springer, New York.
- All the material is available on-line and on the university library. See links on the course's GitHub page.

Course Policies

- There will be no make-up opportunities. Any evaluation opportunity that is missed during the Term (for any reason) will have its weight transferred to the final exam.
- We follow UBC Policies regarding granting Academic Concessions. Please see
  - https://students.ubc.ca/enrolment/academic-learning-resources/academic-concessions
  - http://www.calendar.ubc.ca/vancouver/?tree=3,48,0,0
  - https://science.ubc.ca/students/advising/concession

University Policies

UBC provides resources to support student learning and to maintain healthy lifestyles but recognizes that sometimes crises arise and so there are additional resources to access including those for survivors of sexual violence. UBC values respect for the person and ideas of all members of the academic community. Harassment and discrimination are not tolerated nor is suppression of academic freedom. UBC provides appropriate accommodation for students with disabilities and for religious, spiritual and cultural observances. UBC values academic honesty and students ae expected to acknowledge the ideas generated by others and to uphold the highest academic standards in all of their actions. Details of the policies and how to access support are available here.