



MO444/MC886
Machine Learning and Pattern Recognition

INSTITUTE OF COMPUTING — UNICAMP
2ND SEMESTER, 2015

Prof.: Anderson Rocha
anderson.rocha@ic.unicamp.br

Course Description

Days and Times

Class	Day	Time	Room
Every	Wednesdays	21 – 22:40	PB18
	Fridays	19 – 20:40	PB18

Extra-class aid

Students whom need extra-class help from the professor must send an e-mail with 24 hours in anticipation. However, every week an extra-class option will be available and broadly released to the class.

- Wednesdays @18:00 through 20:00 (Office #79, IC/Unicamp).
- Send an e-mail beforehand confirming you are coming.

Evaluation

This class will be evaluated according to the following criteria:

1. **One written test.** **Weight:** 40% of final grade.
2. **Individual practical assignments** in which the professor gives the problems and specs to the students. Short technical reports must be turned in. The solved problems in the practical assignments may be part of the final evaluation (written) test. Do not miss them.
Weight: $4 \times 10\%$ of the final grade. The implementation language is free of choice.
3. **One extended practical assignment.** **Weight:** 20% of final grade.
 - This assignment must be done in groups of two people. The groups must be formed respecting the categories: undergrads, graduates, and special students (e.g., undergrad students must form groups with other undergrad students).
 - The programming language of choice is free (e.g., C, C++, Java, R, Python etc.).
 - The work must contain the resolution (implementation and tests) of one problem in the area of Machine Learning documented by a technical report in the form of an article (model will be given later). The report and source code must be submitted.
 - At the end of the third week of class, a 1-page PDF document must be turned in (by e-mail) formalizing the problem to be solved and the team who will work on such problem. This document will serve as a “contract” about what will be done and what will be considered when grading.
 - Each team will present their final problem and solution to a committee in the end of the semester.

Exam

- **This discipline does not have an extra exam after the end of the semester.**

Syllabus

1. Introduction to Machine Learning, problems, data, tools
2. Linear regression, SSE, features
3. Overfitting, model complexity, training, validation, test data
4. Classification problems, decision boundaries, nearest neighbour methods
5. Classification and probabilities
6. Naïve Bayes, distributions
7. Linear classifiers
8. Dimensionality reduction: PCA and LDA
9. Evolutionary algorithms
10. Neural networks
11. Ensemble methods: bagging and boosting
12. Validation and comparison of algorithms, cross validation, Wilcoxon and Friedman tests, Bonferroni-Dunn corrections, etc.
13. Unsupervised learning: clustering, k-means
14. Data mining
15. Textual representation and multinomial models; clustering and latent spaces
16. VC dimension, risk minimisation, margin-based classification methods
17. Support Vector Machines for two and more classes
18. Deep learning

Evaluation

- Written test: November 27th (Friday)
- Projects: November 30th to December 2nd (Monday to Wednesday).

Programming Language

It is recommended that the students use a Machine Learning friendly language such as R, Python or Matlab to make the development easier and faster. However, the implementation language is free of choice.

Webpage

<http://www.ic.unicamp.br/~rocha/teaching/2015s2/mo444>

Bibliography

Here are some important references for the course. They are complementary.

Books

1. *Pattern Recognition and Machine Learning*. Christopher M. Bishop. Springer. (2006) **Acronym:** PRML
2. *Elements of Statistical Learning*. Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Springer. (2009) **Acronym:** Elements
3. *Pattern Classification*. Richard O. Duda, Peter E. Hart e David G. Stork. Willey-Interscience. (2000) **Acronym:** PCDUDA
4. *Machine Learning*. Tom Mitchell. McGraw-Hill. (1997) **Acronym:** MLTM
5. *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina*. Katti Faceli, Ana Carolina Lorena, João Gama, André C. P. L. F. de Carvalho. Genio. (2011) **Acronym:** IAAM

Created on August 10, 2015