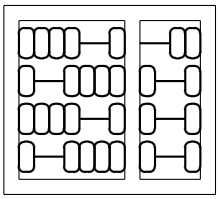

MO444/MC886
Aprendizado de Máquina e Reconhecimento de Padrões

Instituto de Computação — UNICAMP
2º Semestre de 2013
Prof.: Anderson Rocha
anderson.rocha@ic.unicamp.br

Descrição da Disciplina

Horário das aulas

Turmas	Dia	Horário	Sala
Todas	4ª	21 – 22:40	PB18
	6ª	19 – 20:40	CB02

Atendimento

Alunos que precisarem de atendimento extra-classe, favor, enviar um e-mail com 24 horas de antecedência.

Avaliação

A avaliação dessa disciplina se dará a partir de:

1. **Uma avaliação escrita**. **Peso**: 40% da nota total.
2. **Trabalhos práticos individuais** em que o professor passa os problemas e especificação aos alunos. Relatórios técnicos curtos devem ser submetidos. Os problemas resolvidos nos trabalhos práticos individuais podem fazer parte da avaliação final. Portanto, não deixem de fazê-los. **Peso**: 20% da nota total.
3. **Um trabalho prático estendido** de implementação. **Peso**: 40% da nota total.
   - Os grupos deverão ser de **três ou quatro pessoas**. Grupos devem ser formados respeitando as categorias: Graduação, Pós-Graduação e Especiais (e.g., alunos de graduação devem formar grupos com alunos de graduação).
   - A linguagem escolhida é livre (e.g., C, C++, Java, R, Python etc.).
   - O trabalho deverá conter a resolução (implementação e testes) de um problema na área de *Machine Learning* documentado por um relatório técnico.
   - O problema a ser resolvido deve ser previamente discutido com o professor. Relatório técnico e código devem ser submetidos.
   - Ao fim da terceira semana de aula, um documento de uma página deverá ser entregue (por e-mail) formalizando o problema a ser resolvido e a equipe. Esse documento servirá como um "contrato" do que será feito e avaliado.
   - Uma apresentação de 30 minutos deverá ser feita no fim do semestre descrevendo o problema e a solução desenvolvida.

Exame

- **Essa disciplina não possui exame.**

Programa do Curso

1. Introdução ao aprendizado de máquina, problemas, dados, ferramentas
2. Regressão linear, SSE, características
3. Overfitting, complexidade, treinamento, validação, dados de teste
4. Problemas de classificação, fronteiras de decisão, métodos de vizinhos mais próximos
5. Probabilidade e classificação
6. Naïve Bayes, distribuições
7. Classificadores lineares
8. Redução de Dimensionalidade: PCA e LDA
9. Algoritmos evolucionários
10. Redes neurais
11. Métodos de ensemble: bagging e boosting
12. Validação e comparação de algoritmos, validação cruzada, testes de Wilcoxon e Friedman, correção de Bonferroni-Dunn, etc.
13. Aprendizado não supervisionado: clustering, k-médias
14. Mineração de dados
15. Representações textuais e modelos multinomiais; clustering e espaços latentes
16. Dimensão VC, minimização do risco, métodos de classificação baseados em margens
17. Support Vector Machines para uma, duas ou mais classes
18. Deep learning

Avaliações

- Prova Escrita: 22 de Novembro (Sexta-feira)
- Projetos: Dezembro de 27 a 29 (Quarta a Sexta-feira) e Dezembro de 2 a 4 (Segunda a Quarta-feira).

Linguagem de Programação

Recomenda-se que os alunos utilizem a linguagem de programação Matlab para facilitar o desenvolvimento ou C/C++ com OpenCV. No entanto, os alunos são livres para utilizar outras linguagens desde que bem familiarizados com as mesmas e em comum acordo com o professor.

Página do Curso

http://www.ic.unicamp.br/~rocha/teaching/2013s2/mo444

Bibliografia

A seguir, encontram-se algumas referências consideradas importantes para o cumprimento do conteúdo proposto. As referências estão listadas na ordem de importância para o curso.

Livros

1. *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina.* Katti Faceli, Ana Carolina Lorena, João Gama, André C. P. L. F. de Carvalho. Genio. (2011) **Sigla**: IAAM

2. *Pattern Recognition and Machine Learning.* Christopher M. Bishop. Springer. (2006) **Sigla**: PRML

3. *Pattern Classification.* Richard O. Duda, Peter E. Hart e David G. Stork. Willey-Interscience. (2000) **Sigla**: PCDUDA

4. *Machine Learning.* Tom Mitchell. McGraw-Hill. (1997) **Sigla**: MLTM

Criado em 7 de agosto de 2013