

UNIVERSIDADE DE SÃO PAULO
Pós-graduação em Sistemas de Informação
Disciplina: Aprendizado de Máquina
SYLABUS

1. DADOS INFORMATIVOS

- | | | | |
|------|----------------|---|---|
| 1.1. | Unidade | : | Escola de Artes, Ciências e Humanidades - EACH |
| 1.2. | Semestre | : | II – 2021 |
| 1.3 | Horas semanais | : | 4 horas teóricas – 9 horas de estudos: 12 créditos |
| 1.4. | Docente | : | Clodoaldo Aparecido de Moraes Lima – c.lima@usp.br |
| 1.5 | Código | : | SIN 5016 |
| 1.6 | Sala | : | Sala 335 (Prédio I1) |

2. MOTIVAÇÃO

Aprendizado de máquina envolve a automação da manipulação de informação e da representação do conhecimento. As fronteiras de aplicação de tais metodologias computacionais vêm se expandindo rapidamente ao longo dos últimos anos, e esta expansão é patrocinada por técnicas e conceitos derivados da estatística, da matemática aplicada e da ciência da computação. É neste cenário que se insere a síntese de ferramentas computacionais para treinamento supervisionado e não-supervisionado, a partir de dados amostrados. Atualmente, muitas aplicações têm utilizado técnicas de aprendizado de máquina, incluindo sistemas de previsão de séries temporais, reconhecimento de fala, face ou íris, ou análise extração de conhecimento de sinais biológicos.

3. OBJETIVOS

O objetivo do aprendizado de máquina concentra-se no estudo dos aspectos computacionais inerentes ao fenômeno do aprendizado, quer seja este observado em máquinas naturais ou artificiais. Um dos seus objetivos principais é o de propiciar a construção de sistemas capazes de adquirir e integrar, por conta própria, conhecimento acerca do domínio (classe de problemas) em que atuam, de modo a melhorarem progressivamente seu desempenho em relação ao cumprimento de seus objetivos. Nesse sentido, o aprendizado de máquina apresenta-se como uma metodologia alternativa a abordagens tradicionais voltadas para a especificação manual (de “tentativa-e-erro”) de tais sistemas. Este curso tem como objetivo apresentar uma introdução aos principais métodos de aprendizado de máquina; buscando compreender os fundamentos de aprendizagem de máquina, entender metodologias e algoritmos, apresentar e discutir exemplos e aplicações.

4. OBJETIVOS

Os tópicos abordados serão: teoria de aprendizado e generalização, a dimensão VC, trade-off bias-variância, modelos lineares, redes neurais, árvores de decisão, comitê de máquinas: estruturas estáticas e dinâmicas, regularização, métodos de kernel, máquinas de vetor de suporte, aprendizado extremo, algoritmos de agrupamento: k-means, SOM, fuzzy- c-means. Aplicações em: reconhecimento de padrões, aproximação de funções, identificação e controle de sistemas dinâmicos, predição de séries temporais.

5. METODOLOGIA

- Aulas expositivas pelos docentes
- Realização de trabalhos individuais e em grupo

6. CRONOGRAMA

	Data	Programação
1	19/08/2021	Alguns conceitos. Modelos paramétrico vs não paramétrico. Maldição da dimensionalidade. Teoria de aprendizado e generalização. Trade-off bias-variância. Sobre ajuste. A dimensão VC.
2	26/08/2021	Modelos lineares (classificação, regressão). Regressão logística. Seleção de modelos. No Free Lunch Theorem
3	02/09/2021	Revisão sobre otimização não linear irrestrita e restrita. Teorema de Karush-Kun-Tucker. Multiplicadores de Lagrange
4	09/09/2021	Técnicas de seleção de modelos. Modelos de Mistura e o Algoritmo EM. Comitê de máquinas: estruturas estáticas e dinâmicas. Ensemble
5	16/09/2021	Mistura de Especialista
6	23/09/2021	Máquinas de vetores Suporte. Problema linearmente separável, não linearmente separável. Treinamento
7	30/09/2021	Métodos de Decomposição. Algumas variações. Alguns exemplos
8	07/10/2021	Métodos de kernel. Condições. Teorema de Mercer. Kernel Definido Positivo. O truque do kernel
9	14/10/2021	Redes Neurais. Validação vs regularização. Teoria de Regularização. Aprendizado Extremo.
10	21/10/2021	Testes estatísticos para comparação de algoritmos. Aplicações em classificação regressão e previsão de series temporais
11	28/10/2021	Revisão sobre algumas distribuições. Aprendizado Bayesiano. Naive Bayes, Redes Bayesianas discretas e contínuas
12	04/11/2021	Árvores de decisão
13	11/11/2021	Agrupamento e Coagrupamento – conceitos e avaliação, K-means e variações, Self-Organizing Maps e variações
14	18/11/2021	Aprendizado Profundo
15	25/11/2021	Avaliação

7. BIBLIOGRAFIA GERAL

- [1] T. M. Mitchell. Machine Learning. McGraw–Hill Science/Engineering/Math, 432 páginas, 1997.
- [2] Duda, R. O., Hart, P. E. & Stork, D. G. Pattern Classification, 2nd Ed., John Wiley & Sons, 2001.
- [3] Vapnik, V. N. Statistical Learning Theory, Springer-Verlag, 1995 [4] Cristianini, N.; Shawe-Taylor, J Kernel Methods for Pattern Analysis. Cambridge University Press, 2004
- [5] Sharkey, A. Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems. Springer-Verlag London Ltd, 1999
- [6] Simon Haykin. Neural Networks and Learning Machines (3rd Edition). Prentice Hall, 3 edition, November 2008.
- [7] Christopher M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [8]. Periódicos na área de Aprendizado de Máquina (Machine Learning), Inteligência Computacional (Computational Intelligence), Inteligência Artificial (Artificial Intelligence), e afins.

8. TRABALHOS

Trabalhos de modelagem e implementação:

- a) **E1** (obrigatório) - implementação e teste de 02 técnicas de aprendizado de máquina tradicional e 01 de aprendizado profundo para reconhecimento biométrico baseado em face ou baseado em íris. A comparação entre os algoritmos deve ser realizada usando testes estatísticos.
- b) **E2** (Obrigatório) - implementação didática dos algoritmos de aprendizado supervisionado (em duplas) apresentados em aula.
- c) **E3** (opcional) - realização das atividades semanais disponibilizadas no Tidia

9. CRITÉRIOS DE AVALIAÇÃO

9.1. Avaliações

- Prova: P1 (síncrona) e P2 (assíncrona)

- Trabalhos: E1, E2 e E3

9.2. Médias

Nota da Prova	$NP = (0,6*P1+0,4*P2)$
Nota dos Trabalhos	$NT1 = [0,7* E1 + 0,3*E2]$ $NT2 = E3$
Média Final	$MF = (0,5*NP + 0,5*NT1)$
Bônus	$MF_{\text{bonus}} = (MF + 0,1*NT2)$

9.2. Mapeamento para conceitos

Média Final (MF)	Conceito
$8,5 \leq MF \leq 10,0$	A
$7,0 \leq MF < 8,5$	B
$5,0 \leq MF < 7,0$	C
$0,0 < MF < 5,0$	D
$MF = 0,0$	E

10. CONSIDERAÇÕES FINAIS

- A programação de aulas é preliminar e está sujeita a mudanças.
- Não existe abono de faltas e o aluno precisa ter, no mínimo, 75% de presença em aula, caso contrário, reprova por falta.
- A disciplina de Aprendizado de Máquina exige 3h teóricas (que correspondem às aulas em sala de aula) e mais **9h de estudo**, por semana, **no mínimo**