



Plano de Desenvolvimento da Disciplina

MC826(MO431) - Fundamentos de Álgebra Linear e Otimização para Aprendizado de Máquina

Docente: Prof. Marcos M. Raimundo

Descrição

Este documento descreve o plano de desenvolvimento da disciplina (PDD) de MC826(MO431) - Fundamentos de Álgebra Linear e Otimização para Aprendizado de Máquina. Em conformidade com o Regimento Geral da Unicamp, este plano detalha o programa do curso, o cronograma dos blocos de conteúdo, as datas das avaliações, os critérios de nota, a política contra fraudes/plágio e a bibliografia recomendada.

1 Programa da disciplina

O programa da disciplina adota uma abordagem orientada por problemas de Aprendizado de Máquina (ML), na qual cada conceito de Álgebra Linear e Otimização é introduzido em simbiose com sua utilidade em modelos preditivos reais.

O objetivo terminal é capacitar o aluno a: a) Compreender a geometria subjacente à representação de dados e transformações no espaço de parâmetros; b) Derivar analiticamente e analisar rigorosamente algoritmos de otimização contínua; c) Implementar algoritmos do zero em frameworks modernos, compreendendo preconditionadores, estocasticidade e tratamento de não-suavidade.

O curso é estruturado em três blocos fundamentais:

- **Bloco 1: Álgebra Linear Guiada por Dados (Mapeamento e Geometria)**

Objetivo: Estabelecer a fundamentação algébrico-geométrica para a representação de dados, redução de dimensionalidade e regressões.

Conteúdo: Espaços vetoriais, normas (L_1, L_2, L_∞), métricas e produtos internos aplicados a modelos de similaridade (k -NN); sistemas de equações lineares e posto aplicados à Regressão Linear Múltipla; projeções ortogonais e a derivação de Mínimos Quadrados Clássicos (OLS); autovalores, autovetores e matrizes simétricas para análise de variância; Decomposição em Valores Singulares (SVD) aplicada a PCA e Sistemas de Recomendação.

- **Bloco 2: O Cálculo da Otimização Diferenciável (Superfícies de Erro e Sensibilidade)**

Objetivo: Transicionar da geometria linear para a otimização contínua, capacitando o aluno a navegar em superfícies de custo e construir diferenciação computacional.

Conteúdo: Derivadas parciais, vetor gradiente e matriz Jacobiana aplicados à sensibilidade de modelos; convexidade, matriz Hessiana e garantias de mínimo global; Gradiente Descendente (GD) aplicado à Regressão Logística (Entropia Cruzada); séries de Taylor e regra da cadeia multivariada para propagação de erro em redes neurais; Grafos de Computação e Diferenciação Automática (Backpropagation em modo reverso).

- **Bloco 3: Dinâmica Avançada e Escala (Otimização Estocástica, Precondicionamento e Esparsidade)**

Objetivo: Estudar o comportamento de otimizadores de escala em contextos com ruído estocástico, mal condicionamento do espaço e penalidades não-diferenciáveis.

Conteúdo: Gradiente Descendente Estocástico (SGD) e impacto de mini-batches; aceleradores de momento (Polyak e Nesterov); métodos de segunda ordem (Newton) e preconditionamento diagonal adaptativo (RMSProp, Adam); preconditionamento estruturado de tensores (Shampoo), ortogonalização via iterações de Newton-Schulz e o otimizador Muon; otimização não-suave com subgradientes, operadores proximais e o algoritmo ISTA/FISTA para indução de esparsidade (L_1).

2 Atividades e Avaliação

A avaliação da disciplina será composta por três testes teóricos/práticos individuais (T_1, T_2, T_3), aplicados ao longo do semestre após a conclusão de cada bloco, e um Trabalho Final prático (TF) focado na implementação e análise comparativa dos algoritmos estudados.

A nota final (NF) será calculada de forma ponderada pela seguinte fórmula:

$$NF = 0.20T_1 + 0.20T_2 + 0.20T_3 + 0.40TF$$

onde:

- T_1 é a nota do Teste 1 (com foco no Bloco 1), aplicado no dia **16 de setembro**.
- T_2 é a nota do Teste 2 (com foco no Bloco 2), aplicado no dia **21 de outubro**.
- T_3 é a nota do Teste 3 (com foco no Bloco 3), aplicado no dia **02 de dezembro**.
- TF é a nota do Trabalho Final (40% da nota), a ser entregue ao fim do semestre.

Para os alunos de pós-graduação (MO431) e alunos de graduação matriculados simultaneamente, a conversão para conceitos obedecerá à seguinte tabela:

A - Nota Final $\geq 8,5$.

B - Nota Final entre 7,0 e 8,4.

C - Nota Final entre 5,5 e 6,9.

D - Nota Final abaixo de 5,5.

3 Punição para fraudes e plágios

Detecção de fraude ou plágio em qualquer avaliação ou no trabalho final implicará em nota zero na atividade específica para todos os envolvidos. A reincidência ou a constatação de plágio grave acarretará em nota final zero ($NF = 0$) e atribuição de conceito D para todos os envolvidos, além da notificação às instâncias administrativas do IC/Unicamp.

4 Bibliografia Recomendada

- Deisenroth, Marc Peter, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for machine learning*. Cambridge University Press, 2020.
- Wright, Stephen J., and Benjamin Recht. *Optimization for data analysis*. Cambridge University Press, 2022.
- Boyd, Stephen, and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.
- Boyd, Stephen, and Lieven Vandenberghe. *Introduction to applied linear algebra: vectors, matrices, and least squares*. Cambridge University Press, 2018.