Universidade Estadual de Campinas
Instituto de Computação

Plano de Desenvolvimento da Disciplina

MC934 - Tópicos em Engenharia de Computação
MO436 - Tópicos em Aprendizado de Máquina - Turma A
Oferecimento: Impactos sociais de inteligência artificial

Docentes: Prof. Jaques Wainer & Marcos M. Raimundo

Descrição

Este documento descreve, de forma sucinta, o plano de desenvolvimento da disciplina (PDD) de MC934 - Tópicos em Engenharia de Computação – abordando o impacto de inteligência artificial na sociedade, mais especificamente da turma A do primeiro semestre de 2023. Em particular, são destacados, de acordo com os requisitos do Regimento Geral de Graduação, o cronograma de atividades, os critérios de avaliação, punição para fraudes e plágios e a bibliografia a ser utilizada ao longo do semestre.

1. Programa da disciplina

O programa desta turma da disciplina cobrirá, em diferentes nıveis de profundidade, tópicos relacionados à impactos sociais de inteligência artificial. O objetivo terminal da disciplina é que o aluno seja ser capaz de: a) diagnosticar os potenciais de uso de ferramentas de aprendizado de máquina e suas limitações legais, b) identificar os riscos reais das ferramentas de aprendizado de máquina que gerem danos a indivíduos e grupos vulneráveis e, c) julgar quais ferramentas devem ser criadas dentro de contexto de ética e quando criada diagnosticar caminhos de redução de impactos danosos. Para isso, serão abordados os seguintes assuntos:

- O que é IA e seus principais usos.
- Leis regulatórias de Dados/IA.
- O papel de IA na reprodução e amplificação de vieses discriminatórios.
- Abordagens técnicas controle de vieses discriminatórios.
- O direito a explicação e IA explicável.
- Procurement (compra de sistemas de IA por governos).
- Capitalismo de vigilância.
- Controle/indução de comportamento.
- Vigilância governamental e policiamento.
- IA seu impacto em perda de empregos, desigualdade, e produtividade.

- Killer robots.
- Ameaça existencial.

2. Cronograma de atividades

O cronograma de atividades está descrito acima, consistindo numa série de palestras sobre diversos tópicos. A avaliação ocorrerá por resumos (de até 200 palavras) das palestras, entregues semanalmente antes da aula posterior.

3. Critérios de avaliação

Cada um dos resumos será avaliado através de conceitos, de A a F. A nota final NF consiste na mediana das notas. O sistema de conceitos irá se basear na qualidade do resumo. Seguindo os seguintes nortes:

A - Execução exemplar (raros equívocos), com pesquisa adicional com referências.

B - Execução exemplar (raros equívocos), sem pesquisa adicional com referências. Execução com falhas, mas com pesquisa adicional com referências.

C - Execução com falhas e sem pesquisa adicional referências.

D - Execução com falhas excessivas.

E - Execução completamente incoerente.

F - Não execução.

O resumo só poderá ser entregue se o aluno estiver presente na aula. O mapeamento entre conceito final e as notas numéricas será feito da seguinte forma: A - 10, B - 8, C - 6, D - 4, E - 2, F - 0.
Cláusulas:

- Cláusula 1: Um aluno com o conceito A será rebaixado para B se tiver mais que 1/4 dos trabalhos com conceito abaixo de D. E um aluno com conceito B será rebaixado para C se tiver mais que 1/3 dos trabalhos com conceito abaixo de C.

4. Punição para fraudes e plágios

Detecção de fraude ou plágio em um resumo implica em $N_i = 0$ para todos os envolvidos (quem recebeu ajuda e também quem ajudou). Reincidência implica em $NF = 0$ para todos os envolvidos.

5. Bibliografia

A bibliografia será apresentada durante o decorrer do semestre.