

Aprendizagem por Reforço

Como esta disciplina irá funcionar?

Objetivos de aprendizado

Ao final deste curso, o aluno deverá ser capaz de:

1. **criar um sistema de aprendizado por reforço** para tomada de decisões sequenciais;
2. entender como **formalizar seu problema como um problema de Aprendizado por Reforço**, como implementar uma solução e como avaliá-la;
3. entender o espaço de **algoritmos de RL** (Sarsa, Q-learning, Policy Gradients, e mais), e;
4. entender como RL se encaixa sob o guarda-chuva mais amplo de **aprendizado de máquina**, e como ele complementa o aprendizado supervisionado e não supervisionado.

Esta é uma disciplina mão na massa!

1. Em cada aula, vamos implementar algo.
2. Nós vamos ler alguns artigos e implementar algumas técnicas descritas em vários deles.
3. Vamos trabalhar com diferentes ferramentas. No entanto, todas as ferramentas serão pacotes **python**.

Requisitos

1. Você deve *saber* programar em **python**!
2. Você deve *gostar* de programar em **python**.
3. Você deve saber trabalhar com **Github**.

Conteúdo

1. Introdução ao Aprendizado por Reforço
2. Definição e conceitos-chave
3. Ferramentas e ambientes para RL
4. Como avaliar soluções de RL
5. Algoritmos Q-Learning e Sarsa
6. Como usar RL em diferentes ambientes (ambientes não-determinísticos, determinísticos e competitivos)
7. Algoritmos de Deep Q-Learning, Double Deep Q-Learning e Policy Optimization

Avaliação

A nota desta disciplina será calculada da seguinte forma:

- ▶ **Pequenas implementações.** A média dessas implementações valerá 30% da nota final da disciplina.
- ▶ **Dois projetos.** A média desses projetos valerá 50% da nota final da disciplina.
- ▶ **Uma prova final.** A prova final valerá 20% da nota final da disciplina. E tem um critério de corte. Se o aluno não tirar no mínimo **C** na prova final, ele não passa na disciplina.