

DISCIPLINA: PGENE 613 - Tópicos Avançados em Aprendizado de Máquina e Otimização

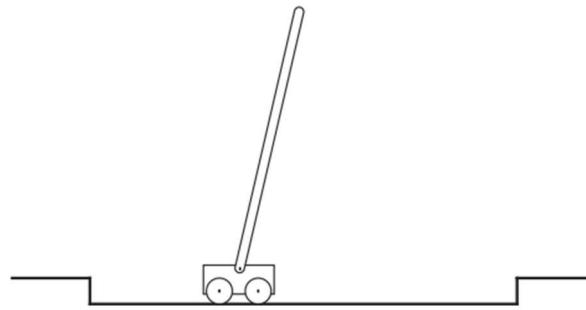
Tópicos em Aprendizado de Máquina e Otimização

2025-2

Trabalho 5: Aprendizado por Reforço

Parte 1:

Considere o problema do equilíbrio de uma haste em um carrinho, conforme mostrado na figura abaixo.



Pede-se:

Comparar o resultado de duas políticas diferentes para manter a haste na vertical:

Política 1: acelera para a esquerda quando o poste está inclinado para a esquerda e acelera para a direita quando o poste está inclinado para a direita.

Política 2: crie uma política de rede neural. Essa rede neural receberá uma observação como entrada e produzirá a ação a ser executada como saída. Utilize o algoritmo de treinamento proposto por Ronald William e descrito em [1].

Para ambas as políticas, considere episódios como no máximo de 500 tentativas. Para cada política, obtenha o valor médio e o desvio padrão do número de tentativas bem sucedidas, para uma sequência de 10 episódios.

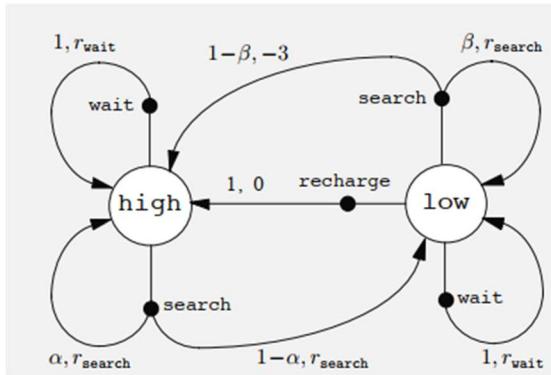
Parte 2:

Seja o problema do robô, visto em sala de aula, cujo diagrama de transição de probabilidades e recompensas é mostrado na figura abaixo. Para esse problema estimar a ação que deve ser executada em cada estado para se obter uma recompensa ótima. Para tanto, simule três situações.

- a) $\alpha = 0.2; \beta = 0.2$
- b) $\alpha = 0.4; \beta = 0.1$
- c) $\alpha = 0.1; \beta = 0.4$

Sugestão: Obtenha a matriz $q(s,a)$ conforme mostrado em problema similar em [1].

DISCIPLINA: PGENE 613 - Tópicos Avançados em Aprendizado de Máquina e Otimização



Relatório:

- 1) Introdução: Explicar no que consiste o trabalho
- 2) Metodologia: descrição dos métodos utilizados
- 3) Resultados: os que foram solicitados
- 4) Conclusões e Discussão: Comentar os resultados obtidos
- 5) Referências Bibliográficas

[1] Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow, Aurélien Géron, O'Reilly, 3rd Edition.

Forma de entrega: Upload na ferramenta de EAD