

Trabalho – Aprendizagem por reforço

Componente Curricular: GEX618 – Inteligência Artificial

Créditos: 4

Ano/semestre: 2025/2

Professor: Felipe Grando

1. Descrição

Trabalho individual ou em duplas cujo objetivo é o estudo e a aplicação de métodos de aprendizagem por reforço através de experimentos com a técnica Q-Learning e uma técnica avançada (A2C, DQN ou PPO).

O trabalho deverá usar o código disponibilizado pelo professor com o ambiente de simulação e o algoritmo Q-Learning já implementados (o código a ser usado no trabalho encontra-se dentro do IA_aula.zip, na subpasta Reinforcement). Use o material disponível no SIGAA (slides e vídeos) para a fundamentação teórica e para guiar suas escolhas no trabalho. Podem usar ou pesquisar outras fontes se julgarem necessário. Também está liberado o uso de qualquer ferramenta de auxílio para o desenvolvimento dos códigos.

O ambiente de simulação disponibilizado não deve ser modificado. Qualquer modificação no treino do algoritmo Q-Learning disponibilizado deve ser mencionada no relatório a ser entregue.

O trabalho deve seguir as seguintes etapas:

1. Configure adequadamente os parâmetros de recompensa do ambiente de simulação e do algoritmo Q-Learning.
 - a. Os valores das recompensas e parâmetros devem ser configurados no arquivo config.py.
 - b. Treine e teste ao menos 20 combinações de configurações (entre valores de recompensas e dos parâmetros do Q-Learning, podendo fixar alguns e mudar outros a cada configuração).
 - c. O valor que controla o número de simulações (SIMULATION_NUMBER) deve ser fixo para todas as configurações testadas. Realize alguns testes preliminares para verificar um valor adequado que economize tempo de processamento sem prejudicar o desempenho dos modelos.
2. Teste cada um dos 20 modelos gerados e summarize os resultados em uma tabela.
 - a. Cada uma das combinações geradas deve ser testada com 1000 simulações.
 - b. Computar a recompensa média e a taxa de sucesso no teste para cada modelo. Para isso será necessário modificar o arquivo RL-game.py.
 - c. Ao menos uma das configurações testadas deve atingir uma taxa de sucesso superior a 30%. Caso isso não ocorra, revise o material disponibilizado e teste mais configurações.
 - d. Crie um relatório com duas tabelas, uma contendo um identificador e os parâmetros para cada modelo e outra tabela com o identificador de cada modelo e seus resultados (recompensa média e taxa de sucesso).
3. Escolha e estude a documentação de um dos algoritmos da biblioteca Stable Baselines3.
 - a. <https://stable-baselines3.readthedocs.io/en/master/index.html>
 - b. Apenas os algoritmos A2C, DQN ou PPO podem ser escolhidos.

4. Aplique o algoritmo escolhido no ambiente de simulação disponibilizado pelo professor.
 - a. Treine 10 modelos, cada um com configurações diferentes de parâmetros (excluindo-se a quantidade de simulações que deve ser padronizada).
 - b. Deixe fixo as configurações de recompensa do ambiente. Use os valores que, no primeiro experimento com o Q-Learning, foram mais bem sucedidos.
5. Teste cada um dos 10 modelos gerados e summarize os resultados em uma tabela similarmente com o que fez anteriormente.
 - a. Cada uma das combinações geradas deve ser testada com 1000 simulações.
 - b. Computar a recompensa média e a taxa de sucesso no teste para cada modelo.
 - c. Adicione ao relatório anterior duas tabelas, uma contendo um identificador e os parâmetros para cada modelo e outra tabela com o identificador de cada modelo e seus resultados (recompensa média e taxa de sucesso).
6. Entregue no SIGAA um documento PDF contendo a identificação do grupo, e as 4 tabelas com as configurações e resultados conforme descrito anteriormente.
7. Apresente e discuta o trabalho com o professor.
 - a. Comentar brevemente os resultados obtidos. Não é necessário slides, podem usar o próprio relatório PDF.
 - b. Executar algumas simulações (com a parte gráfica ativada) usando o modelo que obteve os melhores resultados nos testes.

2. Entrega e prazos

A entrega do trabalho (formato PDF) deverá ser realizada no SIGAA até o dia indicado na tarefa **Trabalho – Aprendizagem por Reforço** que foi cadastrada no SIGAA. Serão desconsideradas as entregas do trabalho por outros meios de comunicação ou entregas realizadas fora do prazo.

3. Método de Avaliação

A avaliação do trabalho considerará a adequação aos requisitos solicitados na descrição do trabalho e a desenvoltura da apresentação.

Será atribuída nota 0 (zero) para todos os estudantes que não entregarem o trabalho dentro do prazo estipulado, que plagiarem o trabalho em sua totalidade ou em considerável parte de um colega ou de fonte externa, ou que não realizarem a apresentação.