

Análise de métodos de aprendizagem por reforço no jogo do dinossauro do *Google*

Daniel de Almeida Duque¹

Jardim Camburi, Vitória-ES. Brasil

Abstract

O artigo visa analisar o método de aprendizagem por reforço utilizando uma meta-heurística e um classificador para jogar o jogo do dinossauro do *Google* e atingir o máximo de pontos possível. Foi escolhido o Algoritmo Genético como meta-heurística e Rede Neural como classificador para o método de aprendizado por reforço. Após o aprendizado, realizou-se a comparação dos resultados, utilizando os testes estatísticos: t pareado e teste não paramétrico de *wilcoxon*, entre o método desenvolvido nesse artigo e o método desenvolvido pelo professor da disciplina de Inteligência Artificial, Flávio Miguel Varejão.

Keywords: aprendizado por reforço, Algoritmo Genético, Rede Neural, jogo do dinossauro, *Google*

1. Introdução

A aprendizagem por reforço consiste no treinamento, de maneira dinâmica, do classificador ou seja, utilizando tentativa e erro. O classificador recebe recompensas quando se aproxima do objetivo e penalidades caso contrário. Utiliza-se
5 tanto uma busca heurística quanto um classificador para que o aprendizado ocorra. Nesse artigo, foi utilizado o Algoritmo Genético (meta-heurística) que tenta achar os melhores parâmetros para maximizar o desempenho do classificador, no jogo, e a Rede Neural (classificador) que analisa o estado corrente do jogo para tomar as decisões (qual tecla pressionar) durante o jogo.

¹Aluno de Engenharia de Computação da Universidade Federal do Espírito Santo (UFES)

10 O desempenho do classificador (agente) foi avaliado executando, no mínimo,
3 vezes o jogo e foi determinado um limite de, no máximo, 24 horas de ex-
ecução do aprendizado. Após essa execução, o melhor agente é selecionado para
executar o jogo 30 vezes e computar os *scores* finais. Lançando mão desses
resultados, é realizada uma comparação com o resultado disponibilizado pelo
15 professor, utilizando teste t pareado e teste não paramétrico de *wilcoxon*.

2. Descrição do Classificador

Os classificadores de Rede Neural são inspirados na biologia, mais especifi-
camente nos neurônios do sistema nervoso. A Rede Neural é dividida em várias
camadas (conjunto de vários nós) de processamento que podem ter topologias
20 bem diferentes. As partes principais das Rede Neural são: camada de entrada,
camadas intermediárias (ou escondidas) e a camada de saída. A modelagem das
camadas para esse problema foi definida da seguinte maneira:

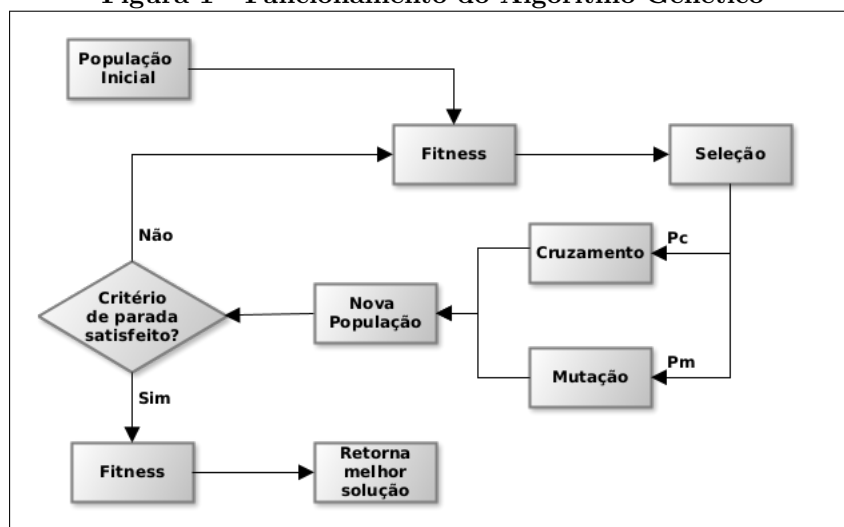
- **camada de entrada:** 1 camada com os sinais de entrada, ou seja, o estado atual(3 nós: *speed*, *distance*, *objHeight*) do jogo.
- 25 • **camadas intermediárias:** 4 camadas intermediárias com 3 nós cada que realizam grande parte do processamento.
- **camada de saída:** 1 camada com o resultado, (1 nó com um valor *float*), do classificador usado para escolher a tecla (menor que 0.5 escolhe *K_UP*, caso contrário *K_DOWN*).

30 3. Descrição da Meta Heurística

O Algoritmo Genético (AG) é uma meta-heurística inspirada na biologia, na seleção natural de Charles Darwin, ou seja, na sobrevivência do mais forte (nesse artigo, pontuação no jogo). Ela consiste em definir uma população inicial e calcular sua *fitness function* que quantifica o quanto a população é boa. A
35 partir disso, aplica-se operadores de seleção (cruzamento, elitismo) e modificação (mutação) dos indivíduos: o cruzamento consiste em combinar dois ou mais

indivíduos para gerar um melhor e a mutação consiste em modificar os genes de um indivíduo para que ele possa melhorar e, com isso, gerar uma nova população. Por fim, caso o critério de parada definido (número de gerações, tempo limite, etc..) for satisfeito, calcula-se o *fitness* e retorna a melhor solução, caso contrário executa tudo novamente utilizando a nova população definida. Abaixo há uma imagem do funcionamento padrão de um Algoritmo Genético (figura 1).

Figura 1 - Funcionamento do Algoritmo Genético



Fonte: Kato, Paiva e Izidoro, 2021, apud Izidoro et al. (2014).

Abaixo, outros conceitos do AG e os valores (em parenteses) usados:

- **gene (peso)**: caracter stica (peso da Rede Neural) que define o indiv duo
- **cromossomo (39 pesos)**: indiv duo com um conjunto de genes
- **popula  o (100 indiv duos)**: conjunto de indiv duos
- **% de muta  o (50%)**: probabilidade de um gene sofrer muta  o.
- **% de crossover (50%)**: probabilidade de dois indiv duos cruzarem e gerar um novo com a combina  o dos genes dos pais
- **% de elite (20%)**: porcentagem fixa da popula  o mantida nas pr ximas gera  es. Os melhores indiv duos s o escolhidos conforme o *fitness* que foi determinado como o *score* obtido durante o jogo.

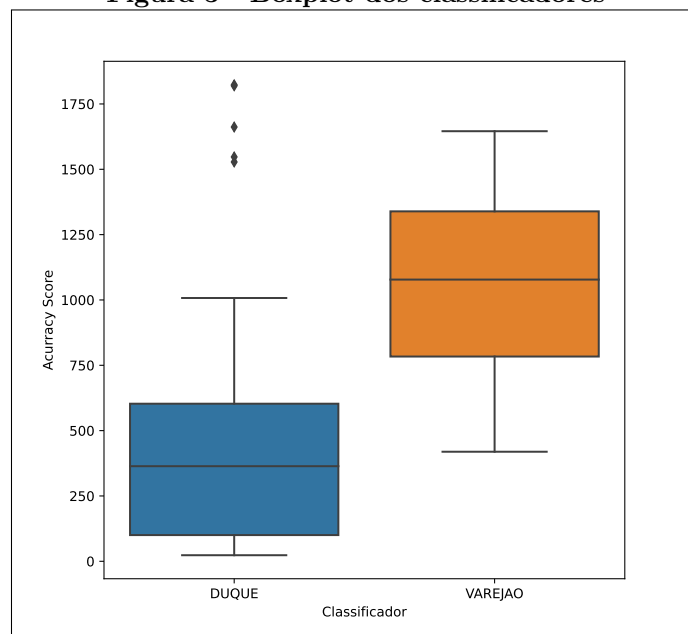
4. Resultados

Após a obtenção do melhor classificador (agente) da Rede Neural, os experimentos com ele foram realizados executando o jogo 30 vezes e, com o resultado, foram calculados a média e o desvio padrão. Além disso, os resultados obtidos pelo agente desenvolvido foram comparados com os resultados fornecidos pelo professor utilizando o teste t pareado e o teste não paramétrico de *wilcoxon* com uma significância de 95% (figura 2).

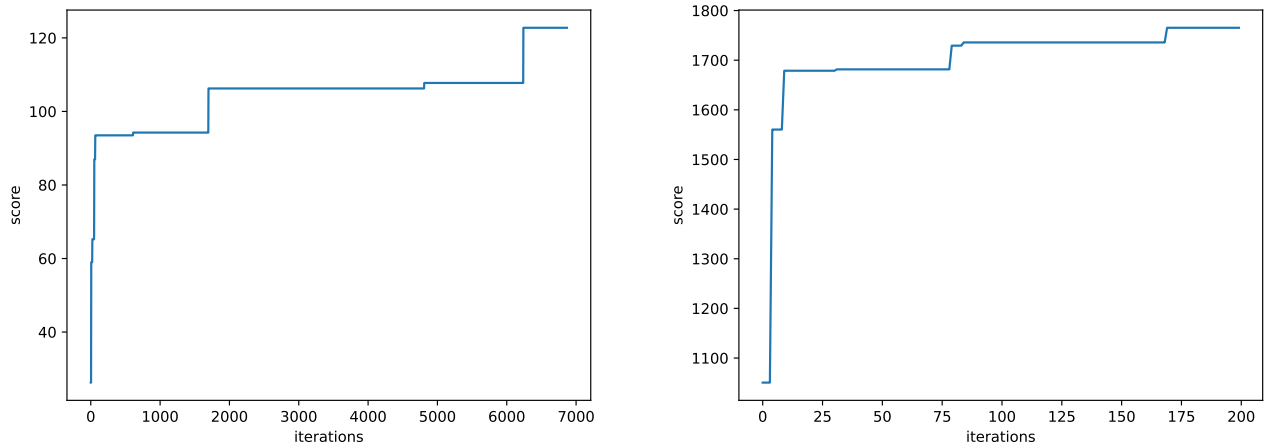
Figura 2 - Tabela pareada dos testes de hipótese

	DUQUE	VAREJAO
DUQUE	DUQUE	0.000
VAREJAO	0.001	VAREJAO

Figura 3 - Boxplot dos classificadores



Tendo em vista a característica evolutiva do Algoritmo Genético, decidi-se gerar gráficos que mostrassem a evolução da população em cada geração (iteração) para visualizar a melhora do *fitness*, ou seja, do *score* do dinossauro
70 ao longo da execução do aprendizado por reforço.



A imagem da esquerda foi a primeira vez que o dinossauro jogou, a partir de então, os melhores indivíduos foram pegos para criar uma nova população inicial de "clones", para que o jogo já iniciasse com os melhores indivíduos
75 achados anteriormente, e, fazendo isso mais algumas vezes, o melhor resultado foi o mostrado na imagem da direita.

5. Conclusões

5.1. Análise geral dos resultados

Percebe-se que o classificador criado pelo professor (VAREJAO) teve de-
80 sempenho melhor e mais estável comparado ao classificador criado no artigo (DUQUE). É perceptível isso quando analisamos a figura 3, que demonstra o resultado do classificador DUQUE com *outliers* e com resultado médio menor do que o classificador VAREJAO. Além disso, podemos confirmar que houve diferença estatística significativa entre os classificadores ao verificar a tabela

85 pareada da figura 2, com *p-values* menores que 0.05. Cabe salientar que o classificador DUQUE não foi executado durante 24h e os valores probabilidades de mutação e *crossover* do AG eram altos, o que pode ser um dos motivos do classificador DUQUE ficar com resultado médio inferior e mais instável.

5.2. Contribuições do Trabalho

90 A utilização da meta-heurística, Algoritmo Genético, em conjunto da Rede Neural foi interessante para explorar o trabalho combinado dessas duas ferramentas para desenvolver um agente através do aprendizado por reforço. Analisando os resultados, será possível comparar com outras meta-heurísticas e outros classificadores, e, também, verificar quais são os melhores métodos para essa
95 tarefa, jogar o jogo do dinossauro do *Google* e obter a maior quantidade de pontos possível.

5.3. Melhorias e trabalhos futuros

Devido à limitação de tempo e do poder de processamento do computador utilizado, percebe-se que, em trabalhos futuro, é importante que sejam utilizadas
100 bibliotecas do *Python* que permitam realizar os cálculos usando a placa gráfica (*GPU*) para melhorar o desempenho e, também, usar bibliotecas que permitam o processamento paralelo, ou seja, executar várias gerações dos dinossauros de forma simultânea. Consequentemente, o tempo de execução seria otimizado, pois percorreria várias soluções simultaneamente.

105 Referências Bibliográficas

- <https://bioinfo.com.br/algoritmos-geneticos/>
- <https://sites.icmc.usp.br/andre/research/neural/>
- Slides das aulas de inteligência artificial em 2022