

# Segundo Trabalho

-

## Inteligência Artificial e Sistemas Inteligentes

Marlon Moratti do Amaral

*Vitória, Brasil*

Universidade Federal do Espírito Santo

---

### Abstract

Este trabalho explora o uso de algoritmos genéticos para otimizar uma rede neural aplicada ao jogo Dino Game da Google, utilizando aprendizado por reforço. Os resultados mostram que o agente desenvolvido superou significativamente um agente base, demonstrando a eficácia dessa combinação na criação de agentes inteligentes para jogos.

*Keywords:* Aprendizado por Reforço, Algoritmo Genético, Redes Neurais.

---

### 1. Introdução

O aprendizado por reforço (AR) [1] é uma abordagem de inteligência artificial (IA) que permite a um agente aprender comportamentos ótimos através da interação com o ambiente e da obtenção de recompensas. A aplicação de AR em jogos tem se tornado comum para explorar novas estratégias de aprendizagem.

Neste trabalho, utilizamos o algoritmo genético (GA) [2] para otimizar os pesos de uma rede neural que joga o Dino Game da Google, um jogo ativado quando não há conexão com a internet. O GA, inspirado na seleção natural, utiliza reprodução, mutação e seleção para tentar encontrar soluções ótimas em espaços de busca complexos.

Nosso objetivo é avaliar a eficácia do GA na otimização de redes neurais no contexto de AR, demonstrando como a combinação de AR e GA pode melhorar

o desenvolvimento de agentes inteligentes em jogos. Esperamos que os resultados contribuam para a pesquisa em AR, oferecendo novas perspectivas sobre  
15 métodos de otimização e suas aplicações práticas em IA.

## 2. Descrição do Classificador

O classificador utilizado neste trabalho é uma rede neural projetada para maximizar a sobrevivência do dinossauro no Dino Game da Google. A rede neural é composta por três camadas principais, conforme ilustrado na Figura 1:

- 20 1. *Camada de Entrada*: Contém 4 neurônios, cada um representando uma das seguintes características do jogo:
  - *Speed*: Velocidade atual do jogo.
  - *Distance*: Distância até o próximo objeto.
  - *obHeight*: Altura do próximo objeto.
  - 25 - *obType*: Tipo do próximo objeto (e.g., cacto pequeno, cacto grande, pássaro).
2. *Camada Escondida*: Consiste em 4 neurônios com função de ativação ReLU, escolhida por sua capacidade de introduzir não-linearidades, permitindo a aprendizagem de relações complexas.
- 30 3. *Camada de Saída*: Possui um único neurônio com função de ativação sigmoide, produzindo um valor entre 0 e 1 que representa a probabilidade de o dinossauro precisar pular.

A saída da rede é comparada com um threshold de 0.55. Se a saída for menor que 0.55, a ação resultante é K\_DOWN, instruindo o dinossauro a se  
35 abaixar. Se a saída for maior ou igual a 0.55, a ação resultante é K\_UP, fazendo o dinossauro pular.

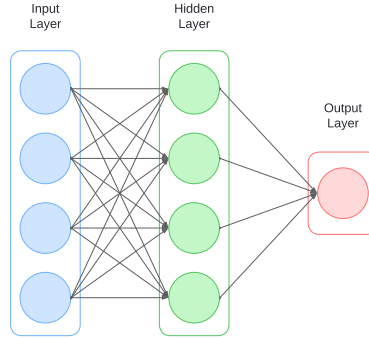


Figure 1: Topologia da rede neural

### 3. Descrição da Meta Heurística

Neste trabalho, utilizamos um GA para otimizar os pesos da rede neural que joga o Dino Game. A seguir, detalhamos os principais componentes e  
 40 parâmetros da metaheurística.

#### 3.1. Representação dos Estados

Os pesos da rede neural foram representados por um vetor de 25 posições. Cada 5 posições compreendem 1 bias e 4 pesos referente a cada neurônio da rede, garantindo uma representação compacta e eficiente.

#### 45 3.2. Inicialização dos Indivíduos

Para a camada escondida, utilizamos a *He initialization*, e para a camada de saída, aplicamos a *Xavier initialization*.

#### 3.3. Seleção dos Indivíduos

Adotamos a seleção por torneio com 5 indivíduos, até que 2 pais distintos sejam escolhidos para gerar 1 descendente. Para uma população de 100  
 50 indivíduos, a seleção é aplicada 100 vezes por geração.

### 3.4. Crossover

A taxa de crossover foi de 75%. O crossover realizado possui  $N - 1$  cortes equidistantes, onde  $N$  é o número de neurônios da rede. Dessa forma, ocorre  
55 uma troca de neurônios entre os indivíduos. A escolha dos neurônios a serem trocados é feita de maneira aleatória.

### 3.5. Mutação

A taxa de mutação foi de 25%. Similar ao crossover, a mutação é aplicada nos neurônios, que são escolhidos de forma aleatória. Aos neurônios selecionados, é  
60 somada uma distribuição normal com média 0 e desvio padrão 1.

### 3.6. Geração da Nova População

Garantimos a geração de indivíduos distintos na nova população e utilizamos o elitismo, mantendo o melhor indivíduo da população anterior para assegurar que as melhores soluções não sejam perdidas.

## 65 4. Resultados

Os resultados gerados são apresentados de quatro maneiras diferentes. Na Figura 2, é possível observar a evolução do agente proposto de acordo com as gerações. Na Tabela 1, apresentamos os 30 resultados agregados, seguidos de média e desvio padrão, tanto para professor quanto para aluno. Além disso, na  
70 Tabela 2, são apresentados os resultados dos testes de hipótese entre os agentes, utilizando o teste t pareado com amostras independentes na matriz triangular superior e o teste não paramétrico de Wilcoxon na matriz triangular inferior. Os testes que rejeitaram a hipótese nula foram destacados em negrito. Por fim, um boxplot é apresentado na Figura 3, facilitando a comparação entre os agentes.

## 75 5. Conclusões

Nesta seção, vamos apresentar uma síntese completa dos resultados deste estudo, organizada em três partes principais: uma análise geral dos resultados

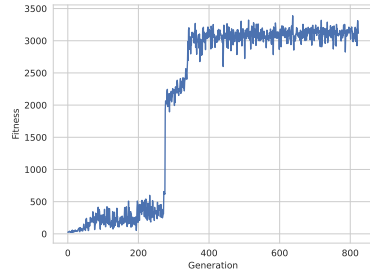


Figure 2: Evolução do agente por geração

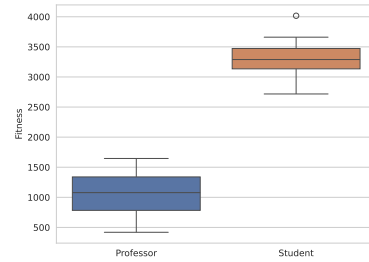


Figure 3: Comparação dos 30 resultados entre os agentes

Agente	Resultados						Média	Desvio Padrão
Professor	1214.0	759.5	1164.25	977.25	1201.0	930.0	1068.18	309.23
	1427.75	799.5	1006.25	783.5	728.5	419.25		
	1389.5	730.0	1306.25	675.5	1359.5	1000.25		
	1284.5	1350.0	751.0	1418.75	1276.5	1645.75		
	860.0	745.5	1426.25	783.5	1149.75	1482.25		
Aluno	3064.25	3390.25	3480.5	3451.25	3249.5	3301.0	3286.04	268.33
	3093.5	3609.75	3187.5	3277.75	2834.0	3381.5		
	3482.5	3511.5	4017.0	3360.25	3217.75	3314.0		
	3204.5	3127.0	3158.25	2808.5	3020.0	2717.5		
	3245.0	3065.0	3502.75	3337.0	3511.0	3661.0		

Table 1: Resultados agregados seguidos de média e desvio padrão

Professor	<b>0.0</b>
<b>0.0</b>	Aluno

Table 2: Testes de hipótese entre os agentes. Teste t pareado com amostras independentes na matriz triangular superior e teste não paramétrico de Wilcoxon na matriz triangular inferior. Em negrito, testes que rejeitaram a hipótese nula.

obtidos, as contribuições significativas deste trabalho e sugestões para melhorias e futuras direções de pesquisa.

### 80 5.1. Análise geral dos resultados

Ao analisar os resultados na Tabela 1 e na Figura 3, fica evidente que o agente proposto se destaca como o melhor método. Ele apresenta uma pontuação

média cerca de três vezes superior ao agente disponibilizado pelo professor, além de possuir um desvio padrão menor. Além disso, os testes de hipótese  
85 na tabela 2 mostram que nosso agente, rejeitou a hipótese nula em relação ao agente disponibilizado para ambos os testes, o que fortalece a evidência de sua superioridade.

### 5.2. Contribuições do Trabalho

Este estudo contribui significativamente ao demonstrar a eficácia dos algoritmos genéticos (GA) na otimização de redes neurais para aprendizado por  
90 reforço, aplicando-os com sucesso ao Dino Game da Google. Desenvolvemos um agente que superou o agente base com uma média de pontuação três vezes maior, validada por rigorosos testes estatísticos. Nossa metodologia detalha práticas de inicialização, seleção, crossover e mutação específicas, oferecendo um guia  
95 prático para problemas similares.

### 5.3. Melhorias e trabalhos futuros

Para aprimorar o desempenho do agente, futuras pesquisas podem explorar técnicas avançadas de seleção e mutação no algoritmo genético, além de arquiteturas de redes neurais mais complexas, como redes convolucionais ou  
100 recorrentes. Adicionalmente, aplicar a metodologia a outros jogos e ambientes simulados, bem como realizar estudos comparativos com outros métodos de otimização, como otimização por enxame de partículas (PSO) ou recozimento simulado (SA), pode validar e generalizar os resultados obtidos.

## References

- 105 [1] R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, 2nd Edition, MIT Press, 2018.
- [2] J. H. Holland, Adaptation in Natural and Artificial Systems, MIT Press, 1992.