|  |  |  |
| --- | --- | --- |
|  | INSTITUTO TECNOLÓGICO DE AERONÁUTICA  CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq |  |

**PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO CIENTÍFICA - PIBIC**

**Título do Projeto**

**André Abreu Moreno**

**RELATÓRIO PARCIAL DE ATIVIDADES**

Orientador: Paulo André L. de Castro

**abr / 2020**

|  |  |  |
| --- | --- | --- |
|  | INSTITUTO TECNOLÓGICO DE AERONÁUTICA  CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq |  |

**PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO CIENTÍFICA - PIBIC**

**Relatório Parcial**

**Construção de Agentes Autônomos de Trading utilizando Inteligência Artificial e Linguagem MQL5**

São José dos Campos, 10 / 04 / 2020

|  |  |
| --- | --- |
| Nome do aluno | André Abreu Moreno |
| Assinatura do aluno |  |

|  |  |
| --- | --- |
| Nome do orientador | Paulo André L. de Castro |
| Assinatura do orientador |  |

INSTITUTO TECNOLÓGICO DE AERONÁUTICA

**PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO CIENTÍFICA - PIBIC**

Formulário de Aprovação de Relatório pelo Orientador

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
| **Relatório:** | **x** | Rel. Parcial |  | Rel. Final |

|  |
| --- |
| **1- CONSIDERO O RELATÓRIO APROVADO COM BASE NOS SEGUINTES ASPECTOS** |
|  |
| **2-** **APRECIAÇÕES DO ORIENTADOR SOBRE O DESEMPENHO DO BOLSISTA NA EXECUÇÃO DO TRABALHO DE INICIAÇÃO CIENTÍFICA** |
|  |

**Local e data:**

**Assinatura do Orientador:**

Sumário

[1. Introdução 6](#_Toc38215412)

[2. Resumo do Plano Inicial 6](#_Toc38215413)

[3. Resumo das Atividades Realizadas 6](#_Toc38215414)

[3.1 Algoritmo SARSA 7](#_Toc38215415)

[3.2 Modelagem do problema e implementação do agente 7](#_Toc38215416)

[4. Plano de Trabalho e Cronograma das Etapas Seguintes 9](#_Toc38215417)

[Bibliografia 11](#_Toc38215418)

# Introdução

A busca de novas tecnologias que possam otimizar e melhorar o desempenho nas atividades de compra e venda de ações no mercado financeiro é algo constante para o setor financeiro. Nesse contexto, surge o conceito *Quantitative Trading*, método no qual a compra e venda de títulos é feita a partir de algoritmos focados na coleta de dados e capazes de aplicar informações obtidas em modelos matemáticos. Os agentes responsáveis pela tomada de decisão, comumente chamados de Traiding Robots, possuem autonomia para realizar diversas transações no mercado financeiro, desempenhando um papel muito importante hoje.

Além dos objetivos esperados por esses algoritmos, como identificar melhores oportunidades de compra e venda, eles são capazes de operar com bom desempenho em conjuntos de dados distintos, aprendendo padrões a partir de regras definidas pelo investidor,

Sendo assim, a pesquisa de tecnologias que resultem em um melhor desempenho desses robôs tem se intensificado, principalmente na área de inteligência artificial, por contemplar mecanismos de aprendizado.

Nesse sentido, a presente pesquisa visa realizar uma pesquisa sobre técnicas de Inteligência artificial que possa ser promissor nessa área e implementá-la e realizar teste utilizando a plataforma de negociação *Metatrader 5* a fim de verificar o seu desempenho como *Traiding Robots*.

# Resumo do Plano Inicial

O plano inicial para esta pesquisa consiste nas seguintes etapas:

* Familiarização com o software Metatrader e com a linguagem MQL5
* Estudo de técnicas de Machine Learning.
* Modelagem do problema para a operação simulada
* Desenvolvimento da aplicação e implementação de agente baseado em Aprendizado por reforço
* Confecção do relatório parcial
* Avaliação do modelo desenvolvido
* Realização de testes, depuração e otimização dos agentes criados
* Confecção do relatório final

# Resumo das Atividades Realizadas

Em um primeiro momento, realizou-se um estudo sobre a software *Metatrader5*. Este sistema de *trader* de ativos permite utilizar algoritmos que automatizam a sua posição na bolsa de ações em tempo real. Para tal, fez uso da documentação disponível no site www.mql5.com.

Em seguida, estudou-se sobre as técnicas de *Machine Learning* e implementou-se um agente baseado em Aprendizado por Reforço utilizando o algoritmo *SARSA*. Esse algoritmo apresentou-se promissor nessa área por ser *model-free*, visto que algoritmos baseados em modelo se tornam impraticáveis à medida que o espaço de estado e o espaço de ação aumentam.

Além disso, algoritmos *model-free* dependem de tentativa e erro para atualizar o conhecimento. Dessa forma, não exige espaço para armazenar toda a combinação de estados e ações.

### 3.1 Algoritmo SARSA

O algoritmo SARSA é uma modificação do algoritmo Q-*learning,* que utiliza um mecanismo de iteração política (Sutton, 1998). Esse algoritmo permite estabelecer autonomamente uma política de ações de maneira interativa.

Esse algoritmo observa a diferença entre a estimativa atual da função de estado-ação, o valor descontado da função de estado, a ação para o próximo estado e a recompensa obtida para então corrigir a estimativa anterior (Oliveira, et al., 2019).

O algoritmo pode ser descrito da seguinte forma:

Inicializa todos os valores de arbitrariamente;

**foreach** *episodio* do

Inicializa ;

Escolhe ação em Q usando -*greedy*;

**repeat**

Executa ação ;

Obtém recompensa e próximo estado ;

Escolhe próxima ação em Q-*greedy*;

//Atualiza

;

**until** *e é estado terminal*;

**end**

Onde:

* é a função estado-ação
* é o estado no instante *t*
* é a ação no instante *t*
* é a recompensa no instante *t*
* é o fator de desconto
* é a taxa de desconto

### 3.2 Modelagem do problema e implementação do agente

Inicialmente, como necessário em todo problema de aprendizado por reforço, definiu-se um espaço de estados, um conjunto de ações e a função recompensa descritos a seguir, conforme modelagem descrita em (Oliveira, et al., 2019).

A base de dados (histórico de valor das ações) utilizada para o aprendizado e testes contém, a cada instante de tempo, os valores referentes a Abertura, Fechamento, Preço de Máxima, Preço de Mínimo, Variação e Volume da ação. Na modelagem, será utilizada as seguintes siglas:

* *A(t)* – Preço de abertura da ação no instante t
* *F(t)* – Preço de fechamento da ação no instante t
* *PMax(t)* – Preço máximo da ação no instante t
* *PMin(t)* – Preço mínimo da ação no instante t
* *Vol(t)* – Volume da ação no instante t

1. **Espaço de Estados**

Para cada instante de tempo *t ∈* {0, 1, 2, 3, ..., T}, tem-se um estado , pertencente a um espaço de estados discreto, que será definido por uma tupla de 9 variáveis descritas a seguir.

1. Tipo de posição: (*LONG, SHORT, NPOS*). Essa variável descreve a posição atual do agente na ação, onde:

* *LONG* refere-se a uma posição comprada;
* *SHORT* refere-se a uma posição vendida;
* *NPOS* informa que o agente não está posicionado.

1. Variação de preço em *t – 2*: (*UP, DOWN*). Essa variável descreve a direção da variação do preço da ação no tempo *t – 2*.

* *UP* se *F(t – 2) > A(t – 2);*
* *DOWN* se *F(t – 2) < A(t – 2).*

1. Variação de preço em *t – 1*: (*UP, DOWN*). Analogamente a variável anterior, porém no instante *t – 1*.

* *UP* se *F(t – 1) > A(t – 1);*
* *DOWN* se *F(t – 1) < A(t – 1).*

1. Ação tomada em *t – 1*: (*BUY, SELL, NOP*). Essa variável descreve a decisão que foi tomada pelo agente no tempo anterior.

* *BUY* denota que o agente comprou a ação;
* *SELL* denota que o agente vendeu a ação;
* *NOP* significa que o agente não operou.

1. Variação do *On-Balance Volume* (OBV) em *t – 2*: (*UP, DOWN*). Essa variável descreve a variável do indicador técnico OBV no tempo *t – 2*.

* *UP* se *Vol(t – 2) > Vol(t – 3),* ou seja, indica que houve um aumento do volume financeiro negociado na ação;
* *DOWN* se *Vol(t – 2) < Vol(t – 3),* ou seja, indica uma diminuição.

1. Variação do OBV em *t – 1*: (*UP, DOWN*). Analogamente a variável anterior, porém no instante *t – 1*.

* *UP* se *Vol(t – 1) > Vol(t – 2);*
* *DOWN* se *Vol(t – 1) < Vol(t – 2).*

1. Extremo mais próximo do preço de fechamento em *t – 2*: (*MAX, MIN*). Essa variável indica se o preço de fechamento da ação no tempo *t – 2* estava mais próximo do preço máximo (*MAX*) da ação ou do preço mínimo (*MIN*).

Seja

* *MAX* se *F(t – 2) > média;*
* *MIN* se *F(t – 2) < media.*

1. Extremo mais próximo do preço de fechamento em *t – 1*: (*MAX, MIN*). Analogamente a variável anterior, porém no instante *t – 1*.

Seja

* *MAX* se *F(t – 1) > média;*
* *MIN* se *F(t – 1) < média.*

1. Retorno na posição em *t – 1*:(*UP, DOWN, NPOS*).Indica se a posição atual do agente em relação ao preço da ação em *t – 1*.

* *NPOS* se ele não estiver posicionado
* *UP* se:
  + Sua posição for *LONG* e *F(t) > F(t – 1);*
  + Sua posição for *SHORT* e *F(t) < F(t – 1).*
* *DOWN* se:
  + Sua posição for *LONG* e *F(t) < F(t – 1);*
  + Sua posição for *SHORT* e *F(t) > F(t – 1).*

1. **Conjunto de Ações do Agente**

O conjunto de ações que o agente pode tomar a cada estado irá depender da posição que ele estiver. A Fig. 1 ilustra as possíveis ações para cada posição do agente.

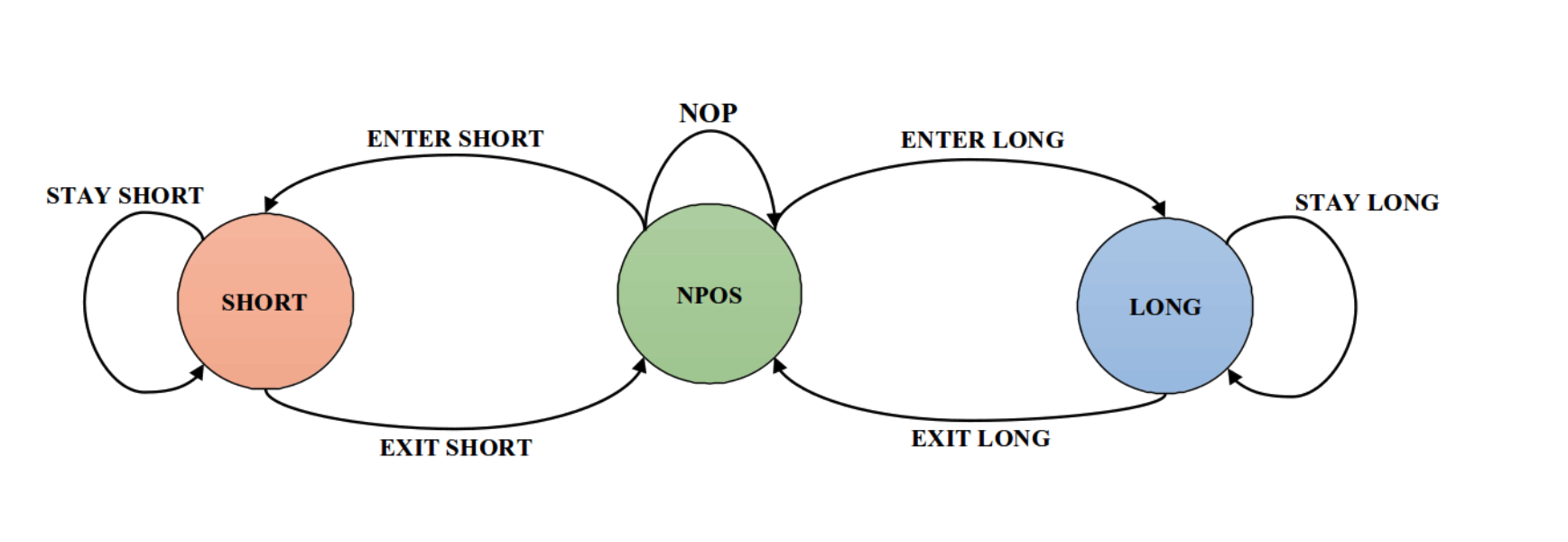


Figura 1: Máquina de estados do agente. Fonte: (Oliveira, et al., 2019)

1. *COMPRAR*:

* *ENTER\_LONG* – inicia uma posição *LONG*;
* *EXIT\_SHORT* – sai da posição *SHORT*.

1. VENDER:

* *ENTER\_SHORT* – inicia uma posição *SHORT*;
* *EXIT\_LONG* – sai de uma posição *LONG*.

1. *NOP*:

* *STAY\_LONG* – permanece na posição *LONG*;
* *STAY\_SHORT* – permanece na posição *SHORT*.
* *NPOS* – permanece não posicionado.

Por fim, cabe ressaltar que as operações de COMPRA e VENDA são executadas a mercado pelo preço de fechamento atual.

1. **Função Recompensa**

Como Fig.1, o conjunto de ações possíveis que o agente poderá tomar, assim como a sua recompensa irá depender do tipo de posição assumida no instante atual. Dessa forma, para ação *at* tomara pelo agente no estado *et* no instante *t,* o agente receberá uma recompensa rt+1 no instante *t+1,* conforme definido a seguir:

# Plano de Trabalho e Cronograma das Etapas Seguintes

A próxima etapa do projeto consiste na realização de teste com algoritmo implementado utilizando histórico de valores na bolsa disponibilizado pela plataforma *Metatrader5*. Em seguida a depuração e otimização dos agentes criados.

|  |  |
| --- | --- |
| **Mai** | * Realização de testes. * Análise parcial dos resultados |
| **Jun** | * Depuração e otimização dos agentes criados * Confecção do relatório final |

# Bibliografia

**Agente de Negociação de Ações Utilizando Aprendizado Por Reforço** [http://www.comp.ita.br/labsca/waiaf/papers/RenatoOliveira\_paper\_8.pdf] / A. Oliveira Renato e Pereira Adriano. - São José dos Campos : [s.n.], 2019.

**MQL5** [Online]. - https://www.mql5.com.

**Reinforcement learning: An introduction** [Livro] / A. Sutton R. S. e Barto, A. G. - Massachusetts : MIT Press, 1998.

**The Julia Programming Language** [Online] // Julia. - https://julialang.org.