

INSTITUTO FEDERAL DE SANTA CATARINA

Bernardo Pires Mesko, Grégori Sabel, João Vitor Waldrich, Rafael Guilherme Onesko

**SOFTWARE PARA PREVISÃO DE PARTIDAS DE BASQUETE USANDO
ALGORITMOS GENÉTICOS EM PYTHON**

Gaspar
2020

Bernardo Pires Mesko, Grégori Sabel, João Vitor Waldrich, Rafael Guilherme Onesko

**SOFTWARE PARA PREVISÃO DE PARTIDAS DE BASQUETE USANDO ALGORITMOS
GENÉTICOS EM PYTHON**

Projeto Integrador apresentado ao Curso Técnico Integrado em Informática do Campus Gaspar do Instituto Federal de Santa Catarina como requisito parcial para aprovação na unidade curricular Projeto Integrador I.

Orientador: Thiago Lipinski Paes, Dr.

Co-orientador: Rafael Silvano Ferreira Silva

,
Software para Previsão de Partidas de Basquete usando
Algoritmos Genéticos em Python : / Bernardo Pires Mesko,
Grégori Sabel, João Vitor Waldrich, Rafael Guilherme
Onesko; orientador, Thiago Lipinski Paes, Dr. ;
co-orientador, Rafael Silvano Ferreira Silva. - Gaspar,
SC, Brasil 2020.
23 p.

- Instituto Federal de Santa Catarina, . . .

Inclui Referências

I. , . II. , . III. Instituto Federal de Santa
Catarina. . III. Software para Previsão de Partidas de
Basquete usando Algoritmos Genéticos em Python.

RESUMO

O basquete é provavelmente um dos esportes mais populares da atualidade. Toda essa fama é acompanhada de um mercado bastante lucrativo, que movimenta bilhões de dólares todos os anos. Os jogos de basquete não afetam somente seus times e outras organizações diretamente envolvidas com as partidas; afeta todo um contexto social, desde torcedores, empresas, dentre vários outros. Dentre todo esse nicho temos também as pessoas pertencentes ao crescente mercado de apostas esportivas. Uma grande demanda desse ramo é a existência de algum método que consiga prever, com segurança e da melhor forma, resultados de partidas: um problema com uma complexidade grande o suficiente para até mesmo recentemente não poder ser resolvido de maneira simples. Com a grande quantidade de ferramentas disponíveis, foi preciso avaliar quais linguagens e metodologias seriam as mais adequadas para o projeto e, com as informações disponíveis. Depois de uma vasta pesquisa foi definido a linguagem de programação Python como a base do projeto devido a sua simplicidade e agilidade, integrando-a com algoritmos de web scraping para coleta de dados. Este trabalho tem como objetivo o desenvolvimento de uma estratégia própria para a previsão de resultados de partidas, e espera que possam ser usados algoritmos genéticos para testar o impacto de diversos aspectos de um time para a inferência de seu desempenho no primeiro quarto tempo de uma partida na NBA no *software* desenvolvido.

Palavras-chave: Basquete. Previsão. Algoritmo Genético. Python.

ABSTRACT

Basketball is probably one of present day's most popular sports. All this fame is surrounded by a very lucrative market, which trades billions of dollars every year. The Basketball matches don't affect only their teams and other directly involved organizations with them; it concerns a whole social context, from fans of the sport, to businesses, between several others. Inside this niche we also have people belonging to the growing market of sports. One big demand of this branch is the existence of some method able to predict, with security and by the best possible manner, match results: a problem with a big enough complexity for it not to be solved in a simple way until recently. With the sheer amount of available tools, it was necessary to determine which languages and methodologies would be the most adequate for the project. After a wide search it was defined that the Python programming language would be the basis for the system, due to its simplicity and agility, integrating it with web scraping algorithms for data collection. This research project has as its objective the development of a proprietary strategy for match results prediction, and it is hoped that genetic algorithms will be used in the developed software for testing the impact of several aspects of a team on the inference of its performance in the first quarter time of a NBA match.

Keywords: Basketball. Prediction. Genetic algorithm. Python,

LIST OF FIGURES

Figure 1 – Medição de paridade (coeficiente de Gini) entre diversas competições de basquete. Fonte: (HARVARD SPORTS ANALYSIS, 2016)	11
Figure 2 – Imagem detalhando o fluxo do algoritmo genético. Fonte: Desenvolvido pelos autores.	15
Figure 3 – Diagrama de Casos de Uso	17
Figure 4 – Diagrama de Classes	18
Figure 5 – Diagrama de Sequência 1: Definir Partida para Prever	18
Figure 6 – Diagrama de Sequência 2: Gerar Vetor de Pesos	19
Figure 7 – Diagrama de Sequência 3: Manter Dados dos Times	19
Figure 8 – Diagrama de Sequência 4: Prever resultado da partida	20

LIST OF TABLES

Table 1 – Tabela de Requisitos	15
Table 2 – Cronograma	21

LISTA DE ABREVIATURAS E SIGLAS

IFSC	Instituto Federal de Santa Catarina
NBA	National Basketball Association
IA	Inteligência Artificial
HTML	Hypertext Markup Language
GTK	GIMP Toolkit
GIMP	GNU Image Manipulation Program
GNOME	GNU Network Object Model Environment
NBB	Novo Basquete Brasil
TCC	Trabalho de Conclusão de Curso
CSV	Comma Separated Values
AG	Algoritmo Genético
RF	Requisito Funcional
RNF	Requisito Não Funcional

CONTENTS

1	INTRODUÇÃO	9
1.1	OBJETIVOS	9
1.1.1	Objetivo geral	9
1.1.2	Objetivos específicos	9
1.2	JUSTIFICATIVA	10
2	FUNDAMENTAÇÃO TEÓRICA	11
2.1	HISTÓRIA E CONCEITOS DO BASQUETE	11
2.1.1	Paridade do basquete	11
2.1.2	Web scraping	11
2.1.3	Algoritmo Genético	12
2.1.4	Python	12
2.1.5	Interface Gráfica com GTK em Pytoon	12
2.1.6	Trabalhos Correlatos	12
3	MATERIAIS E MÉTODOS	14
3.1	DESCRIÇÃO DA SOLUÇÃO PROPOSTA	14
3.2	MATERIAIS	14
3.2.1	Python	14
3.2.2	Web Scraping	14
3.2.3	Algoritmos Genéticos	14
3.3	MÉTODOS	15
3.3.1	Requisitos Funcionais e Não-Funcionais	15
3.3.2	Diagrama de Casos de Uso	16
3.3.3	Diagrama de Classes	18
3.3.4	Diagramas de Sequência	18
3.4	VALIDAÇÃO	20
4	CRONOGRAMA	21
4.1	Estudo Teórico	21
4.2	Testes	21
4.3	Conceituação	21
4.4	Análise de Sistemas	21
4.5	Produção	21
4.6	Documentação	21
4.7	Apresentação	21
5	RESULTADOS ESPERADOS	22
	BIBLIOGRAPHY	23

1 INTRODUÇÃO

A *National Basketball Association* (NBA) é uma das principais ligas de basquetebol profissional da América do Norte, e de acordo com a *Harvard Sports Analysis* é um dos campeonatos desse esporte com maior paridade, ou seja, os times dessa liga possuem um nível de habilidade muito semelhante, o que por sua vez acarreta competições mais justas ([HARVARD SPORTS ANALYSIS, 2016](#)). Junto aos diversos fãs do esporte e acompanhadores da liga surge também uma grande quantidade de receita gerada: em torno de US\$ 8 bilhões na temporada de 2019 ([GOUGH, 2020](#)).

O mercado de apostas esportivas, outro setor importante, movimenta anualmente centenas de bilhões de dólares em todo o mundo. Esse fato atrai diversas pessoas interessadas em lucrar com este mercado. O site Terra estima que 2 bilhões de reais são movimentados por ano somente no Brasil, por exemplo ([DINO - DIVULGADOR DE NOTÍCIAS, 2018](#)).

Embora esse mercado tenha um grande potencial para a movimentação da economia, o Brasil se encontra relativamente atrasado neste setor. A regulamentação brasileira de apostas esteve em hiato por muito tempo ([SOARES, 2019](#)), até a criação da lei nº 13.756/18, que cria a definição de apostas de quota fixa (uma categoria que engloba as apostas esportivas) num contexto legislativo, e prevê um prazo de 2 anos para sua regulação pelo Ministério da Fazenda ([BRASIL, 2018](#)).

Há uma demanda muito grande de métodos confiáveis de previsão de resultados de partidas. Com esse intuito, diversas metodologias vêm sendo criadas, como a previsão com base em estatísticas prévias ([SUZUKI; TAVARES, 2018](#)) ou com mineração de dados ([FUKANO et al., 2018](#)), por exemplo. De forma geral, partidas de basquete envolvem muitas variáveis, e isso as torna difíceis de prever com segurança, ainda mais em uma liga tão competitiva quanto a NBA.

Unificando o mundo da NBA com o das apostas esportivas percebe-se que, além de uma boa previsão de resultados, é preciso levar em consideração as chamadas *odds*, as probabilidades de cada resultado ocorrer, que estão diretamente relacionadas ao valor pago pelas casas de aposta em caso de acerto. Configura-se, então, grande desafio: desenvolver uma solução que seja capaz de prever resultados e que, além disso, seja rentável a curto, médio ou longo prazo.

Deste modo, o presente trabalho tem como objetivo geral a criação de uma solução relativamente confiável, que emprega estatísticas pregressas em conjunto com o método computacional de algoritmos genéticos para fornecer indicações de vitória em partidas de basquete. O escopo inicial do trabalho ficará limitado apenas a times participantes da NBA, e a validação do trabalho, por sua vez, será feita a partir da comparação das previsões de resultados de partidas passadas feitas pelo programa com seus reais resultados. O programa também irá prever resultados de jogos futuros, os quais, à medida que acontecerem, servirão de indicador de sua precisão.

1.1 OBJETIVOS

Nessa seção serão mostrados o objetivo geral e os objetivos específicos do projeto proposto.

1.1.1 Objetivo geral

Desenvolver uma aplicação que analisa dados de jogos da NBA e que com eles tente prever o time vencedor do primeiro quarto tempo em uma partida.

1.1.2 Objetivos específicos

1. Coletar dados necessários para o programa sendo através de Web scraping;
2. Utilizar IA para formular uma relação entre esses dados e;
3. Apresentar os resultados da previsão de forma clara e precisa.

1.2 JUSTIFICATIVA

Este projeto foi escolhido por ter gerado muito interesse no grupo e por demandar tecnologias que oferecem muitas possibilidades de criação e várias formas de automação, por este motivo temos o intuito de utilizar deste projeto para aprendizado, visando no futuro suprimir esta e outras demandas do mercado. Outro ponto importante a se ressaltar é que como o mercado de apostas em esportes gira muito capital, tal projeto pode auxiliar novos públicos que pretendam começar a investir seu dinheiro nessa modalidade sem muito risco.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo tem como objetivo apresentar os principais conceitos relacionados ao basquete e ao seu funcionamento, assim como, noções básicas das tecnologias presentes no desenvolvimento do projeto.

2.1 HISTÓRIA E CONCEITOS DO BASQUETE

O basquetebol foi criado em 1891 pelo professor canadense de Educação Física James Naismith (1861-1940), surgindo como alternativa a esportes ao ar livre como futebol ou beisebol, visto que era incapaz de jogar esses anteriormente citados por conta dos rigorosos invernos na região. Além disso, a ideia era criar um jogo menos violento que o futebol americano, podendo ser usado durante aulas de educação física do próprio professor e criador como forma de integrar os alunos e estimular a coletividade e trabalho em grupo. Cada partida tem a duração de 40 minutos (na NBA são 48 minutos), sendo dividida em quatro tempos de 10 minutos (na NBA são 4 tempos de 12 minutos), chamados de "quartos". Em caso de empates ao final da partida, são permitidas prorrogações de 5 minutos ([TODAMATERIA, 2020](#)).

2.1.1 Paridade do basquete

A paridade é um conceito muito importante no meio esportivo, ainda mais quando o assunto envolve apostas. Ela nada mais é que o quão equivalentes em habilidade os times são. Por exemplo, quanto maior a paridade, mais difícil será prever qual time sairá vencedor de certa competição, pois a habilidade dos times são muito parecidas. De acordo com a *Harvard Sports Analysis*, a NBA é uma das ligas que conta com maior paridade, ou seja, alta competitividade e competições mais justas, levando em conta o nível similar dos times participantes ([HARVARD SPORTS ANALYSIS, 2016](#)). Foi feito um levantamento entre diversas ligas de basquete e suas respectivas paridades (representadas pelo coeficiente de Gini), cujos resultados se encontram na Figura 1.



Figure 1 – Medição de paridade (coeficiente de Gini) entre diversas competições de basquete. Fonte: ([HARVARD SPORTS ANALYSIS, 2016](#))

2.1.2 Web scraping

Web scraping é uma das formas de fazer coleta de dados da web de forma automática. A forma mais comum de utilizar a tecnologia é através da escrita de um programa automatizado que consulta um servidor web, solicita dados que são adquiridos na forma de HTML e outros arquivos que compõem páginas da web, e então analisa os dados para extrair a informação necessária. Em prática, *web scraping* abrange uma larga

variedade de técnicas e tecnologias de programação, como análise de dados, análise de linguagem natural e segurança de informação (MITCHELL, 2018).

2.1.3 Algoritmo Genético

O algoritmo genético é baseado nos princípios de seleção natural e evolução propostos por Charles Darwin. Na seleção natural, vários indivíduos de uma espécie, ao serem colocados em um ambiente hostil, terão que tentar sobreviver e se destacar para poderem se reproduzir e passar seus genes adiante. No algoritmo genético, um fenômeno similar acontece. São criados vários indivíduos, geralmente vetores, com valores predefinidos e que irão passar por um teste, que dará uma pontuação para cada indivíduo de acordo com sua performance. Com essa pontuação, será criado um ranking do indivíduo com melhor pontuação para o de menor pontuação. Os indivíduos mais abaixo do ranking serão deletados, e aqueles remanescentes passarão por um processo de cruzamento. O processo de cruzamento consiste em gerar novos indivíduos a partir daqueles selecionados. Nesta fase há troca de genes entre dois indivíduos, denominados pais, dando origem a outros indivíduos, os filhos. Essa mistura entre dois indivíduos pais se chama *crossing over*. Com os indivíduos sobreviventes e seus filhos, o processo de seleção natural reiniciará. Tal ciclo continuará até os indivíduos alcançarem uma performance desejada (LINDEN, 2008).

2.1.4 Python

Python é uma linguagem com foco na simplicidade que comunica-se de forma transparente com outros sistemas operacionais (Windows, Linux, MacOS, etc) devido ao fato de ser uma linguagem que opera através de um interpretador. Está muito presente em *back-ends*, simulações de engenharia, *machine learning* dentre outras utilidades (PYTHON FOUNDATION, 2012). Devido a estes fatores mencionados, utilizaremos desta linguagem como base do nosso projeto.

2.1.5 Interface Gráfica com GTK em Python

A Interface Gráfica é uma das partes mais importantes de qualquer *software*. Ela é responsável por criar todo elemento gráfico que aparecerá ao usuário: imagens, animações, botões, etc; simplificando a interação usuário-máquina de linhas de código para alguns cliques, um bom exemplo disso seriam os ícones no Windows.

O GTK é um *toolkit* gratuito e *open-source* utilizado para a criação de interfaces gráficas que funciona em múltiplas plataformas e é compatível com backends de diversas languages, incluindo Python. O GTK utiliza além de suas bibliotecas um aplicativo chamado *Glade*, o qual permite a visualização e edição da interface de forma simples e intuitiva.

GTK é a base da plataforma de desenvolvimento do GNOME, mas também pode ser usado para programar aplicações em ambientes Linux como também no Microsoft Windows e no MacOS da Apple.

2.1.6 Trabalhos Correlatos

No trabalho de conclusão de curso intitulado “Utilização de Redes Neurais e Regressão Linear na Predição de Resultados de Jogos do Novo Basquete Brasil (NBB)” (MICHELON et al., 2017) é feita uma proposta de estudo para descobrir qual técnica de redes neurais artificiais e regressão linear apresenta maior acurácia para previsão de resultados na liga NBB. O documento apresenta que a técnica de redes neurais provê em média, 74% de precisão, enquanto a técnica de regressão linear possui uma média de 75%.

Além disso, no artigo “Previsão de Resultados no Futebol por meio de Técnicas de Aprendizado de Máquina” (FERNANDES, 2019), foram reunidos dados sobre os times e os jogadores através de placares de partidas antigas. Estes foram aplicados em algoritmos de regressão e classificação como florestas de decisão aleatória. O método de *gradient boosting* teve a maior *F1-Score* (métrica usada para classificar a qualidade de um algoritmo de aprendizado de máquina) entre as estratégias testadas. Esta estratégia proveu uma precisão de 52,78%, o que lhe rendeu uma pontuação de 0,509.

Já os autores do TCC “Predição de Resultados de Jogos da NBA. Uma Abordagem de Mineração de Dados com Aprendizado de Máquina para Playoffs” ([FUKANO et al., 2018](#)) usaram aprendizado de máquina para tentar prever resultados de partidas de basquetebol chamadas *playoffs*, que são jogos ocorridos após o término da temporada oficial de uma liga, nos quais os times com a melhor classificação se enfrentam. Com este método de mineração de dados foi alcançada uma taxa de precisão das previsões para qual equipe venceria as partidas correspondente a 73,19%.

3 MATERIAIS E MÉTODOS

Essa seção apresentará os materiais e métodos do projeto proposto.

3.1 DESCRIÇÃO DA SOLUÇÃO PROPOSTA

Este projeto tem como objetivo principal a previsão de resultados do primeiro quarto tempo em partidas de basquete da NBA. O sistema terá uma interface intuitiva que possibilitará ao usuário escolher qual partida pretende prever ao informar os times que disputarão a partida. Após receber os times, o sistema irá buscar nos arquivos armazenados as informações necessárias das partidas que já ocorreram e, usando valores pré-determinados obtidos por meio de um algoritmo genético, irá calcular qual das duas equipes possui a maior chance de vencer a partida, e mostrará os resultados ao usuário, juntamente com o percentual de certeza daquela previsão.

O programa obterá os dados dos times usando uma técnica de *web scraping* no site da NBA e salvará essas informações em uma pasta local. Tais dados serão utilizados como entrada no momento de prever os resultados das partidas, como comentado anteriormente, mas também serão necessários no momento de execução do algoritmo genético para descobrir os pesos que multiplicarão estes dados a fim de atingir a previsão.

3.2 MATERIAIS

Nessa seção serão apresentadas as tecnologias que foram utilizadas ao longo do projeto proposto.

3.2.1 Python

Python é uma linguagem fácil de manipular, e é uma das linguagens que promove agilidade no desenvolvimento de *software*. Além disso, o fato de ter uma grande comunidade e ser muito utilizada facilita a resolução de problemas que possam vir a ocorrer, pois os fóruns sobre Python sempre provêm muito apoio aos desenvolvedores ([PYTHON FOUNDATION, 2012](#)).

O GTK será usado para a criação da *interface* gráfica, relacionando-se ao programa escrito em Python. Foi escolhido pelo fato de ser uma ferramenta poderosa (usada em vários programas famosos como o GIMP), gratuita e *open-source*, além de ser similar às ferramentas atualmente ensinadas no curso técnico em informática do IFSC, como o JavaFX.

Este toolkit permite uma fácil implementação do sistema e, sendo multiplataforma, pode ser usado com Python (por meio da *framework* PyGObject) permitindo homogeneidade ao resto do projeto ([THE GTK TEAM, 2010](#)).

3.2.2 Web Scraping

Em função de coletar os dados será utilizado o *web scraping*. Escolhemos esse método pois ele permite coletar uma grande quantidade de dados, que serão utilizados a fim de construir o vetor de pesos e a prever os resultados das partidas. Existem outras formas de coletar grandes quantidades de dados, mas com o *web scraping*, é possível coletar dados de basicamente quaisquer sites utilizando a estrutura HTML. Desta forma haverá uma área muito grande de atuação, e será possível, inclusive, coletar dados diretamente do site oficial da NBA. Isso permitirá o acesso às informações de forma muito atualizada. Após a coleta dos dados com o *web scraping*, os mesmos serão armazenados em uma pasta local, no formato .CSV ([MITCHELL, 2018](#)).

3.2.3 Algoritmos Genéticos

Para a previsão dos resultados das partidas, onde será previsto o vencedor no primeiro quarto tempo da partida, será utilizado um algoritmo genético. Ele foi escolhido pelo fato de não depender fortemente de bibliotecas, que muitas vezes são complexas e difíceis de entender, possibilitando um controle muito maior sobre

o algoritmo, e não necessitando de tanta calibragem quanto outros métodos estudados. Pode se ver o fluxo do algoritmo genético com mais detalhes na Figura 2.

A fim de conseguir os melhores resultados sem necessitar de muitas horas de processamento. Será necessário filtrar bem as informações que serão utilizadas para treinar o AG. Dessa forma utilizaremos os dados que aparentam mais decisivos, como por exemplo: porcentagem de vitórias no primeiro quarto tempo, média de pontos feitas e média de cestas de três pontos. Todas essas informações serão colunas de informação que serão utilizadas no período de treino do AG (LINDEN, 2008).

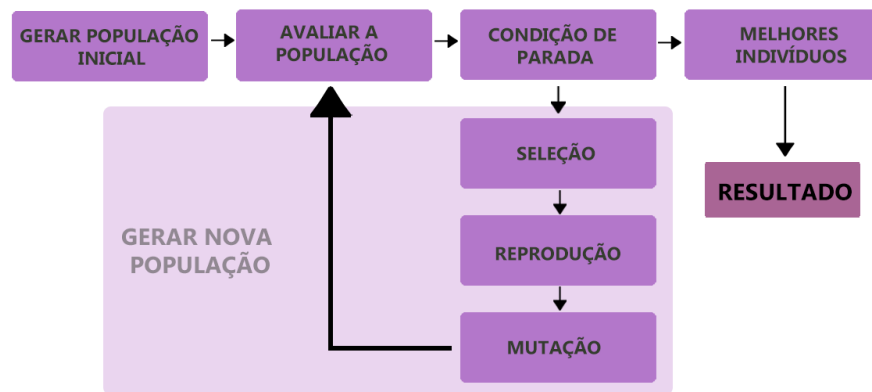


Figure 2 – Imagem detalhando o fluxo do algoritmo genético. Fonte: Desenvolvido pelos autores.

3.3 MÉTODOS

3.3.1 Requisitos Funcionais e Não-Funcionais

Os requisitos de um *software* são os aspectos essenciais que devem ser levados em consideração na hora do desenvolvimento e também servem como guias para manter o escopo do projeto. A Tabela 1 apresenta os requisitos elencados para o sistema.

Table 1 – Tabela de Requisitos

Identificação	Descrição	Prioridade	Levantamento
RF1	Prever resultados do primeiro quarto tempo de uma partida	Essencial	25/08/2020
RF2	Manter dados dos times	Essencial	25/08/2020
RF3	Gerar vetor de pesos	Importante	25/08/2020
RF4	Poder selecionar partidas para previsão	Importante	25/08/2020
RF5	Mostrar resultados em uma interface gráfica	Importante	25/08/2020
RNF6	Ter taxa de 70% de acerto nas previsões	Desejável	25/08/2020
RNF7	Ser feito na linguagem Python	Importante	25/08/2020

Fonte: Desenvolvido pelos Autores (2020)

Identificação: **RF1**

Descrição: **Prever resultados do primeiro quarto tempo de uma partida**

Detalhamento: O sistema realizará previsões dos resultados de partidas de basquete com base nos dados aos quais ele possui acesso. Esta é a função principal do programa, e grande parte dos outros requisitos servem para tornar esta especificação possível.

Prioridade: Essencial

Data de Levantamento: 25/08/2020

Identificação: **RF2**

Descrição: **Manter dados dos times**

Detalhamento: O sistema deve poder coletar, armazenar e editar arquivos com dados essenciais para realizar as previsões das partidas, tal como a taxa de vitórias de determinado time de basquete.

Prioridade: Essencial

Data de Levantamento: 25/08/2020

Identificação: **RF3**

Descrição: **Gerar vetor de pesos**

Detalhamento: Usando algoritmos genéticos, o sistema irá gerar uma série de números (pesos) para multiplicar os valores disponíveis nos arquivos com os dados de cada time. É esperado que cada peso corresponda à importância dos dados de um time para a previsão de seu desempenho na partida.

Prioridade: Importante

Data de Levantamento: 25/08/2020

Identificação: **RF4**

Descrição: **Poder selecionar partidas para previsão**

Detalhamento: O sistema deverá permitir que os usuários selecionem os times que jogariam na partida hipotética cujo resultado o programa ficará encarregado de prever.

Prioridade: Importante

Data de Levantamento: 25/08/2020

Identificação: **RF5**

Descrição: **Mostrar resultados em uma interface gráfica**

Detalhamento: Para promover a usabilidade do programa, ele deverá possuir uma interface gráfica, na qual exibirá maneiras do usuário interagir com o *software* de maneira intuitiva.

Prioridade: Importante

Data de Levantamento: 25/08/2020

Identificação: **RNF6**

Descrição: **Ter taxa de 70% de acerto nas previsões**

Detalhamento: Foi estabelecido uma taxa mínima de certezas para as previsões do programa para haver um critério concreto para diferenciar previsões "boas" de "ruins", além de tornar os resultados esperados mais realistas, levando em consideração que raramente haverá uma certeza absoluta nas previsões realizadas.

Prioridade: Desejável

Data de Levantamento: 25/08/2020

Identificação: **RNF7**

Descrição: **Ser feito na linguagem Python**

Detalhamento: O sistema deverá ser desenvolvido na linguagem de programação Python, por ela oferecer uma facilidade e agilidade de desenvolvimento.

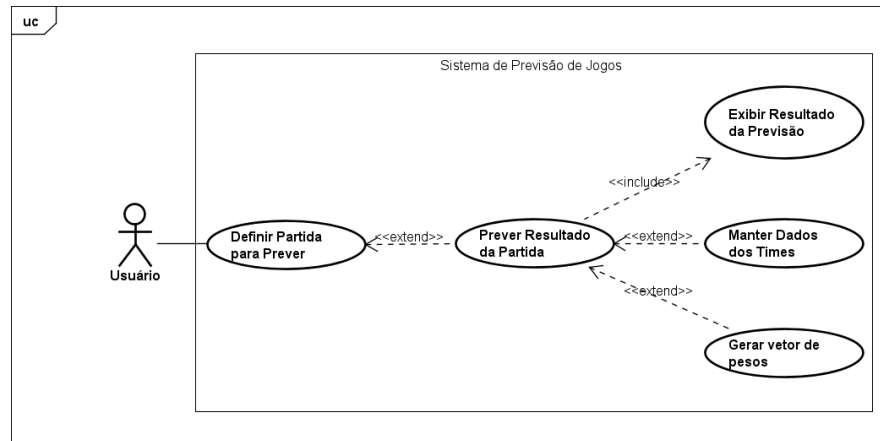
Prioridade: Importante

Data de Levantamento: 25/08/2020

3.3.2 Diagrama de Casos de Uso

Nessa seção é mostrado o diagrama de casos de uso. O diagrama de casos de uso mostra de maneira geral as funcionalidades de um sistema, além da maneira com a qual estas funcionalidades interagem entre si e com os atores (RIBEIRO, 2012). A Figura 3 apresenta o diagrama de casos de uso elencado no projeto.

Figure 3 – Diagrama de Casos de Uso

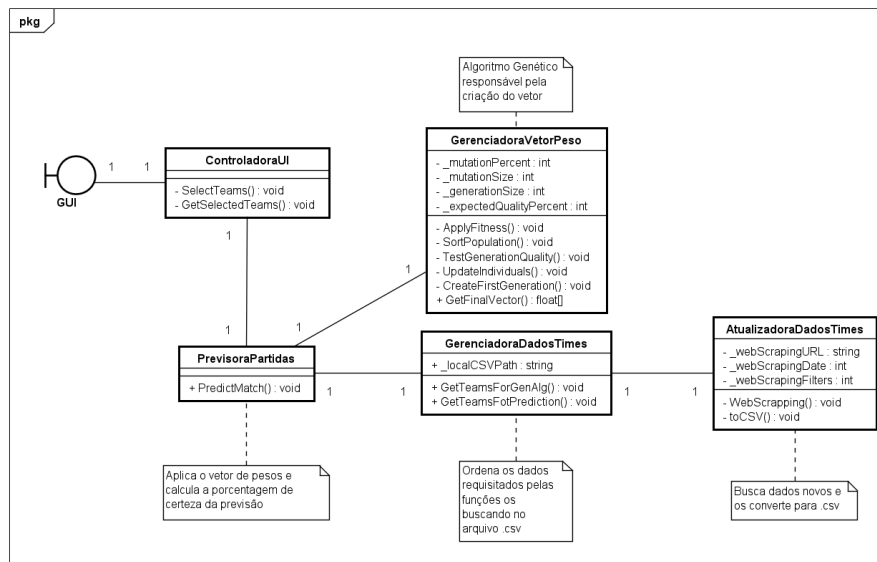


Fonte: Desenvolvido pelos Autores (2020)

3.3.3 Diagrama de Classes

Nessa seção é mostrado o diagrama de classes. Os diagramas de classes representam a estrutura e relações entre as classes do sistema que servem de modelo para a instanciação dos objetos (TYBELL, 2016). A Figura 4 apresenta o Diagrama de Classes elencado no projeto.

Figure 4 – Diagrama de Classes

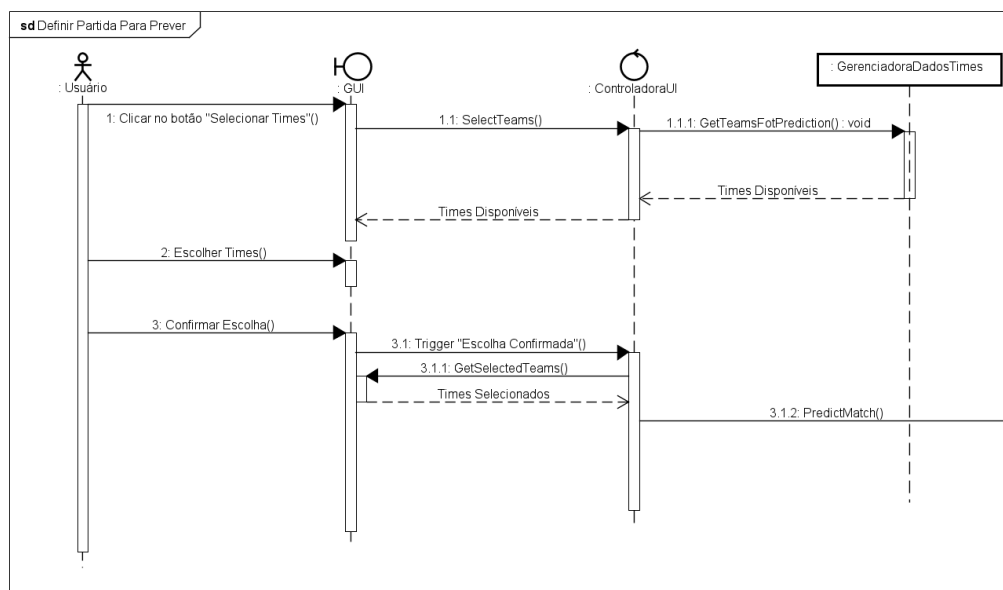


Fonte: Desenvolvido pelos Autores (2020)

3.3.4 Diagramas de Sequência

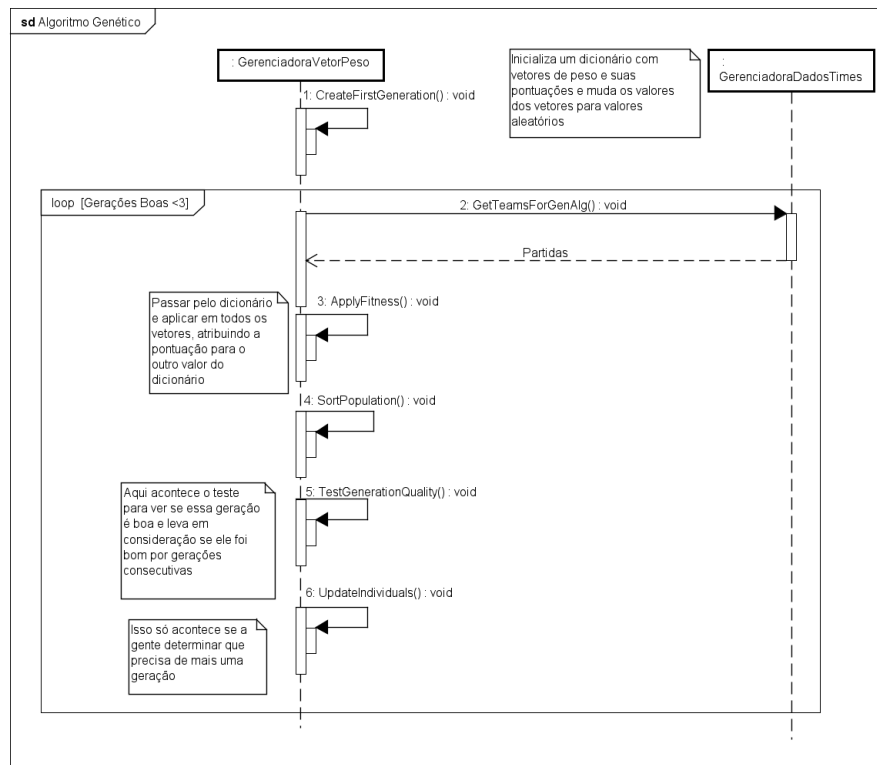
Nesta seção são mostrados os diagramas de sequência. Seu principal objetivo é demonstrar em linhas de tempo quais são as interações entre os objetos de um determinado cenário representado pelo diagrama (OLIVEIRA, 2013). As Figuras 5, 6, 7 e 8 apresentam os diagramas de sequência elencados no projeto.

Figure 5 – Diagrama de Sequência 1: Definir Partida para Prever



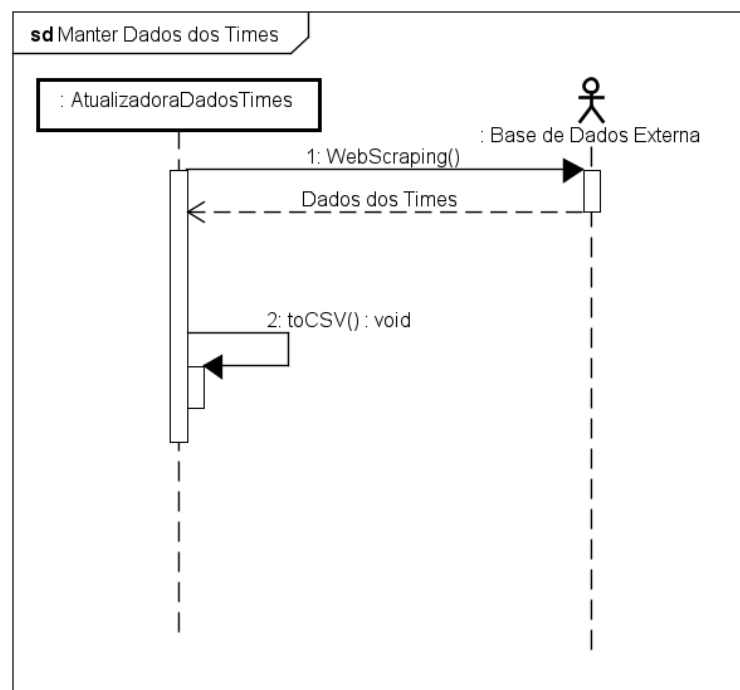
Fonte: Desenvolvido pelos Autores (2020)

Figure 6 – Diagrama de Sequência 2: Gerar Vetor de Pesos



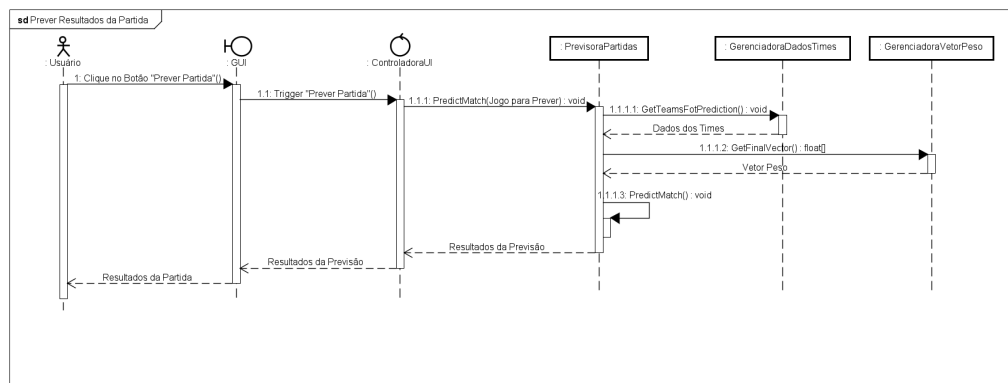
Fonte: Desenvolvido pelos Autores (2020)

Figure 7 – Diagrama de Sequência 3: Manter Dados dos Times



Fonte: Desenvolvido pelos Autores (2020)

Figure 8 – Diagrama de Sequência 4: Prever resultado da partida



Fonte: Desenvolvido pelos Autores (2020)

3.4 VALIDAÇÃO

A validação do projeto irá consistir em obter os times que jogarão em partidas futuras, simular essas partidas no *software* e anotar as previsões que ele realiza. Será colocado um cartaz detalhando as futuras partidas da NBA em mural no campus Gaspar do IFSC e toda vez que uma partida ocorresse, ele seria atualizado com a comparação entre o resultado real e o resultado previsto. Foi estabelecido como meta uma taxa de acerto de 70%.

Outro método para realizar a validação deste trabalho é convidar estudantes do campus para usar o programa e avaliar suas experiências com a intuitividade da interface e as impressões que eles tiveram acerca das previsões realizadas pelo sistema. Isto seria mensurado através de um questionário, realizado posteriormente.

4 CRONOGRAMA

Nesta seção será apresentado o cronograma para o desenvolvimento deste projeto, e em seguida, uma explicação do que consiste cada etapa desse processo.

Cronograma	2020											2021			
	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	JAN	FEV	MAR	ABR
Estudo Teórico		X	X		X	X	X	X	X						
Testes					X	X	X	X							
Conceituação	X	X				X									
Análise de Sistemas							X	X							
Produção										X	X	X	X	X	
Documentação		X	X	X	X	X	X	X	X	X	X	X	X	X	
Apresentação									X						X

Table 2 – Cronograma

4.1 Estudo Teórico

Nos períodos de estudo teórico, é feita uma análise das tecnologias disponíveis para selecionar quais serão propriamente utilizadas para o desenvolvimento do *software*, e o estudo das ferramentas escolhidas, para já acostumar os membros da equipe a trabalharem com elas quando forem implementar o projeto em si.

4.2 Testes

Na fase de testes são criados protótipos usando as tecnologias selecionadas para o desenvolvimento do sistema descrito nesse projeto, com a finalidade de verificar a factibilidade do que está sendo proposto.

4.3 Conceituação

Na parte de conceituação foram planejados aspectos básicos do sistema final desejado, como o fluxo de código, e levantados os requisitos do sistema para a análise de sistemas poder ser realizada.

4.4 Análise de Sistemas

Nesta etapa do projeto foram feitas tanto a avaliação e correção dos requisitos levantados na fase de conceituação, quanto a criação dos diagramas UML (Figuras 3 a 8).

4.5 Produção

Nesse período será feita a própria implementação do trabalho descrito neste documento, ou seja, esta seção do cronograma abrange os momentos em que será escrito código dentro do repositório do *software*.

4.6 Documentação

Essa fase abrange a duração do projeto na qual foi ou será feita a descrição e documentação do sistema, e se estenderá até o final do desenvolvimento do *software* e a escrita do artigo para a disciplina de Projeto Integrador II do IFSC.

4.7 Apresentação

No estágio de apresentação será feita a exposição do projeto e o que foi realizado até o momento como parte das disciplinas de Projeto Integrador I e II do IFSC.

5 RESULTADOS ESPERADOS

A princípio, espera-se que o *software* desenvolvido ao longo deste trabalho, sirva como uma maneira relativamente confiável de obter previsões para resultados do primeiro quarto de partidas da NBA, além de atuar como possível referencial para futuros desenvolvedores com ideias similares a este projeto, como a precisão dos algoritmos genéticos para inferência de dados

Fora isso , é desejado que os membros da equipe desenvolvedora aprofundem seu conhecimento nas tecnologias aplicadas, como Python, interface gráficas, *web scraping* e desenvolvimento de *software* de uma maneira geral.

BIBLIOGRAPHY

- BRASIL. *Lei 13.756, de 12 de Dezembro de 2018*. [S.l.], 2018. Disponível em: <https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/54976993/do1-2018-12-13-lei-n-13-756-de-12-de-dezembro-de-2018-54976737>. 9
- DINO - DIVULGADOR DE NOTÍCIAS. *Mercado de apostas esportivas movimenta R\$ 2 bilhões no Brasil, segundo pesquisa*. [S.l.], 2018. Disponível em: <<https://www.terra.com.br/noticias/dino/mercado-de-apostas-esportivas-movimenta-r-2-bilhoes-no-brasil-segundo-pesquisa,5e91353bb264cfb927b0b93d8a94e1a397u8wbih.html>>. Acesso em: 26 mar. 2020. 9
- FERNANDES, F. A. P. *PREVISÃO DE RESULTADOS NO FUTEBOL POR MEIO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA*. Belo Horizonte: [s.n.], 2019. Disponível em: <<https://bit.ly/2ER0cQKJ>>. Acesso em: 16 abr. 2020. 12
- FUKANO, A. M. K.; BRANCHER, J. D.; BONIDIA, R. P. *PREDIÇÃO DE RESULTADOS DE JOGOS DA NBA. UMA ABORDAGEM DE MINERAÇÃO DE DADOS COM APRENDIZADO DE MÁQUINA PARA PLAYOFFS*. Londrina: [s.n.], 2018. Disponível em: <http://www.uel.br/cce/dc/wp-content/uploads-/TCC_ALEXANDRE_MIKIO_KIMURA_FUKANO.pdf>. 9, 13
- GOUGH, C. *National Basketball Association total league revenue* from 2001/02 to 2018/19(in billion U.S. dollars)*. [S.l.]: [s.n.], 2020. Disponível em: <<https://www.statista.com/statistics/193467/total-league-revenue-of-the-nba-since-2005/>>. Acesso em: 16 set. 2020. 9
- HARVARD SPORTS ANALYSIS. *WHICH SPORTS LEAGUE HAS THE MOST PARITY?* Cambridge, MA, Estados Unidos, 2016. Disponível em: <<http://harvardsportsanalysis.org/2016/12/which-sports-league-has-the-most-parity/>>. Acesso em: 26 mar. 2020. 5, 9, 11
- LINDEN, R. *Algoritmos Genéticos*. 2. ed. Rio de Janeiro, RJ: O'Reilly Media, Inc., 2008. 12, 15
- MICHELON, A.; JUNIOR, O. O. B.; SCHOEFFEL, P. *Utilização de Redes Neurais e Regressão Linear na Predição de Resultados de Jogos do Novo Basquete Brasil (NBB)*. Ibirama, SC: [s.n.], 2017. Disponível em: <<http://periodicos.unesc.net/sulcomp/article/view/3149/2879>>. Acesso em: 20 abr. 2020. 12
- MITCHELL, R. *Web Scraping with Python*. 2. ed. Sebastopol, CA, Estados Unidos: O'Reilly Media, Inc., 2018. 12, 14
- OLIVEIRA, L. V. de. *UML: Diagramas de Sequência*. [S.l.]: [s.n.], 2013. Disponível em: <<http://www.theclub.com.br/restrito/revistas/201308/umld1308.aspx>>. Acesso em: 27 ago. 2020. 18
- PYTHON FOUNDATION. *Python: Site Oficial*. 2012. Disponível em: <<https://www.python.org/>>. Acesso em: 27 ago. 2020. 12, 14
- RIBEIRO, L. *DevMedia: O que é UML e Diagramas de Caso de Uso: Introdução Prática à UML*. Praia Grande, SP: [s.n.], 2012. Disponível em: <<https://www.devmedia.com.br/o-que-e-uml-e-diagramas-de-caso-de-uso-introducao-pratica-a-uml/23408>>. Acesso em: 27 ago. 2020. 16
- SOARES, I. de C. *Regulação e Tributação de apostas esportivas no Brasil: lei 13.756/18 e a compatibilidade com o ordenamento jurídico brasileiro*. 2019. Disponível em: <<https://repositorio.ufpb.br/jspui/handle/123456789/16211>>. Acesso em: 16 set. 2020. 9
- SUZUKI, A. K.; TAVARES, L. *Modelagem estatística para previsão esportiva: uma aplicação no futebol*. São Paulo, SP: [s.n.], 2018. Disponível em: <<https://repositorio.usp.br/bitstream/handle/BDPI/51336/2713238.pdf?sequence=1>>. 9
- THE GTK TEAM. *The GTK Project - A free and open source widget toolkit*. [S.l.], 2010. Disponível em: <<https://www.gtk.org/>>. Acesso em: 27 ago. 2020. 14
- TODAMATERIA. *Regras do Basquete*. [S. l.], 2020. Disponível em: <<https://www.todamateria.com.br/regras-do-basquete/>>. Acesso em: 02 mai. 2020. 11
- TYBELL, D. *DevMedia: Orientações básicas na elaboração de um diagrama de classes*. São Mateus, ES: [s.n.], 2016. Disponível em: <<https://www.devmedia.com.br/orientacoes-basicas-na-elaboracao-de-um-diagrama-de-classes-/37224>>. Acesso em: 27 ago. 2020. 18