



INSTITUTO FEDERAL GOIANO  
CAMPUS MORRINHOS

YAGO SILVA PEIXOTO

# **Mineração de Dados em Prol de Atividades Físicas**

Morrinhos  
2018

YAGO SILVA PEIXOTO

# **Mineração de Dados em Prol de Atividades Físicas**

Trabalho de Conclusão de Curso apresentado à  
Coordenação do Curso de Tecnologia em Sistemas  
para Internet do Instituto Federal Goiano – Campus  
Morrinhos, como requisito parcial para obtenção do  
grau em Tecnólogo em Sistemas para Internet.

**Área de concentração:** Mineração de Dados  
**Orientadora:** Prof.<sup>a</sup> Dra. Leila Roling Scariot da  
Silva

**Coorientador:** Prof.<sup>o</sup> Esp. Freddy Enrique Ramos  
Guimarães

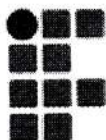
Morrinhos  
2018

Sistema desenvolvido pelo ICMC/USP  
Dados Internacionais de Catalogação na Publicação (CIP)  
**Sistema Integrado de Bibliotecas - Instituto Federal Goiano**

P379m Peixoto, Yago Silva  
Mineração de Dados em Prol de Atividades Físicas /  
Yago Silva Peixoto; orientadora Leila Roling Scariot  
da Silva; co-orientador Freddy Enrique Ramos  
Guimarães. -- Morrinhos, 2018.  
32 p.

Monografia (Graduação em Tecnologia em Sistemas  
Para Internet) -- Instituto Federal Goiano, Campus  
Morrinhos, 2018.

1. Mineração de dados. 2. Weka. 3. Pesquisa de  
Campo. I. Silva, Leila Roling Scariot da, orient.  
II. Guimarães, Freddy Enrique Ramos, co-orient. III.  
Título.



**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO**

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano, a disponibilizar gratuitamente o documento no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

**Identificação da Produção Técnico-Científica**

- Tese
- Dissertação
- Monografia – Especialização
- TCC - Graduação
- Produto Técnico e Educacional - Tipo: \_\_\_\_\_
- Artigo Científico
- Capítulo de Livro
- Livro
- Trabalho Apresentado em Evento

Nome Completo do Autor:

Matrícula:

Título do Trabalho:

**Restrições de Acesso ao Documento**

Documento confidencial:  Não  Sim, justifique: \_\_\_\_\_

Informe a data que poderá ser disponibilizado no RIIF Goiano: 21/03/2019

O documento está sujeito a registro de patente?  Sim  Não

O documento pode vir a ser publicado como livro?  Sim  Não

**DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA**

O/A referido/a autor/a declara que:

- o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Marcinhos 21/03/2019  
Local Data

Yago Silva Piscoto  
Assinatura do Autor e/ou Detentor dos Direitos Autorais

Ciente e de acordo:

Biba Neco  
Assinatura do(a) orientador(a)

Trabalho dedicado à minha orientadora, Prof.<sup>a</sup> Dr<sup>a</sup> Leila Roling Scariot da Silva, dedico também ao meu Coorientador, Prof.<sup>o</sup> Esp. Freddy Enrique Ramos Guimarães e à instituição de ensino Instituto Federal Goiano – Campus Morrinhos.

---

## Agradecimentos

---

Agradeço à Prof.<sup>a</sup> Dr<sup>a</sup> Leila Roling Scariot da Silva, por sua ajuda e prestação necessária para realização deste trabalho.

Agradeço ao Prof.<sup>o</sup> Esp. Freddy Enrique Ramos Guimarães, por sua ajuda e prestação necessária para realização deste trabalho.

Agradeço ao meu amigo José Osmar Martins Silva, que foi voluntário e ajudou para que o trabalho fosse realizado.

Agradeço a minha amiga Gabriela Ferreira Borba, que ajudou com ideias para que o trabalho fosse realizado.

Agradeço a todos que puderam contribuir para a pesquisa de forma direta ou indireta.

---

## Resumo

---

Peixoto, Yago. **Mineração de Dados em Prol de Atividades Físicas**. Morrinhos, 2018. 32p. Relatório de Graduação. Campus Morrinhos, Instituto Federal Goiano.

Fizemos este trabalho com base no artigo [Peixoto \(2017\)](#) “Mineração de Dados em Prol de Atividades Físicas”. Apresentamos e publicamos o mesmo no evento VI CEICT (VI Congresso Estadual de Iniciação Científica e Tecnológica do IF Goiano), realizado no campus de Urutaí em 2017. O trabalho consistiu em usar mineração de dados e tecnologias móveis para uma proposta de *app*, que possa ajudar pessoas a ter hábitos/vida mais saudáveis, e enfrentar a principal dificuldade, que é substituir um velho hábito por outro melhor e não desistir. A proposta é levantar dados para que possa ser usados em um *app* que busca ajudar o indivíduo a construir novos hábitos, fazendo exercícios que ele goste ou que sejam melhor aceitos por ele, auxiliando também na alimentação. Com isso, a transição para uma vida saudável se torna mais fácil, e após algum tempo, bons hábitos são criados. Para a coleta de dados, realizamos uma pesquisa de campo, a anamnese, com praticantes de diversas atividades físicas. Escolhemos programa *Weka* para fazer o tratamento e minerar os dados tendo em vista a experiência de uso da equipe, bem como a usabilidade e confiabilidade dos resultados desta ferramenta – já demonstrada em muitos artigos estudados. Logo após, fizemos uma análise acompanhado de um profissional da área, o Prof.º Esp. Freddy Enrique Ramos Guimarães. Com isso, obtivemos as informações para desenvolver um sistema que consiga realizar as tomadas de decisão de modo mais adequado para o praticante de atividades físicas.

### Palavras-chave

Mineração de Dados; *Weka*; Pesquisa de campo.

---

## Abstract

---

Peixoto, Yago. **Data Mining in Favor of Physical Activities**. Morrinhos, 2018.32p. Graduation report. Campus Morrinhos, Goiano Federal Institute.

We did this work based on the article [Peixoto \(2017\)](#) "Data Mining in Favor of Physical Activities". We present and publish the same in event VI CEICT (VI Congresso Estadual de Iniciação Científica e Tecnológica do IF Goiano), held at Urutaí campus in 2017. The work consisted in the use of data mining and mobile technologies for an application proposal, which can help people to have healthier habits / lives, and face the main difficulty, which is to replace one old habit with another better and not give up. The proposal is to raise data so that it can be used in an application that seeks to help the individual to build new habits, doing exercises that he likes or that are better accepted by him, also helping in the feeding. With this, the transition to a healthy life becomes easier and, after some time, good habits are created. For data collection, a field survey, anamnesis, was carried out with practitioners of various physical activities. We chose the Weka program to process and extract the data in view of the experience of using the team, as well as the usability and reliability of the results of this tool - already demonstrated in many articles studied. Soon after, we did an analysis accompanied by a professional of the area, Professor Esp. Freddy Enrique Ramos Guimarães. With this, we obtained the information to develop a system that is able to make decisions in a more appropriate way for the practitioner of physical activities.

### Keywords

Data Mining; Weka; Field research.



---

# Sumário

---

<b>CAPÍTULO 1 – Mineração de Dados em Prol de Atividades Físicas.....</b>	<b>11</b>
Introdução.....	12
Fundamentação teórica.....	13
Material e Métodos.....	15
Resultados e Discussão.....	17
Conclusão.....	27
<b>Referências Bibliográficas.....</b>	<b>28</b>
APÊNDICE A – Perguntas da Anamnese.....	30
APÊNDICE B – Algoritmo de Classificação da Planilha.....	32

---

## Índice de figuras

---

Figura 1: Classificação dos indivíduos no <i>Weka</i> , azul corresponde a avançado, vermelho a intermediário, azul-claro a não classificados e cinza a iniciantes.....	17
Figura 2: Idade em relação ao nível do praticante de atividades físicas.....	18
Figura 3: Motivos para praticar em relação ao nível do praticante de atividade física.....	19
Figura 4: Duração do tempo de treino em relação ao nível do praticante de atividade física.....	20
Figura 5: Frequência semanal em relação ao nível do praticante de atividade física.....	21
Figura 6: Alimentação balanceada em relação ao nível do praticante de atividades físicas.....	22
Figura 7: Árvore de decisão do programa J48.....	24
Figura 10: Resultado do ranqueamento do algoritmo de seleção de atributos <i>ReliefAttributeEval</i> .....	25
Figura 10: Resultado do ranqueamento do algoritmo de seleção de atributos <i>InfoGainAttributeEval</i> .....	25
Figura 10: Resultado do ranqueamento do algoritmo de seleção de atributos <i>OneRAttributeEval</i> .....	26

---

## Mineração de Dados em Prol de Atividades Físicas

---

A procura por formas tecnológicas que facilitam o dia a dia vem aumentando drasticamente, bem como o uso dos aplicativos móveis, contudo, a tecnologia induz ao sedentarismo (BARROS 2014). Por isso, surgiu a ideia de um aplicativo que pode auxiliar na realização de atividades físicas direcionadas de acordo com o perfil de cada usuário. Para isso é necessário coletar e minerar dados para assim, utilizá-los dentro do aplicativo, transformando estas informações em respostas mais adequadas aos usuários do aplicativo.

Foram coletadas respostas em uma pesquisa de campo usando uma anamnese que, de acordo com Rosa (2019) define o perfil psicossocial, físico e nutricional. Através da análise do seu histórico é feito o levantamento dos fatores de risco para o treino, do seu histórico físico, do seu estilo de vida e das recomendações médicas para a prática de atividade física. Tais respostas foram mineradas usando a ferramenta *Weka*. Tendo estes dados em mãos, realizamos uma análise para verificar a veracidade dos mesmos e obter mais dados que podem ser trabalhados em prol das atividades físicas. Este trabalho foi escrito com base no artigo de Peixoto (2017) publicado pelo autor desde.

---

## Introdução

---

A evolução tecnológica trouxe além das comodidades, empregos, redes sociais, entretenimento. É possível se entreter na internet facilmente navegando por redes sociais, assistindo vídeos, conversando com amigos, enfim, uma infinidade de opções. Grande parte das pessoas procuram por meio da tecnologia, formas para auxiliar e facilitar determinadas atividades, sendo que grande parte dessas buscas são voltadas para *softwares*, tais como aplicativos mobile encontrados em lojas virtuais. (ROSA 2018).

O objetivo deste trabalho foi coletar dados de praticantes de atividades físicas e usá-los como respostas para melhorar a qualidade de vida de pessoas ativas fisicamente ou sedentárias. Para isso realizamos uma pesquisa de campo usando uma anamnese e com isso, conseguimos coletar os dados de 110 pessoas praticantes de atividades físicas. Logo após, organizamos estas informações em uma planilha e geramos arquivos “.csv” e “.arff”. Deste modo, os arquivos puderam ser utilizados dentro da ferramenta *Weka*, que é utilizada para trabalhos que necessitem da mineração de dados.

Com o cruzamento destes dados, tínhamos outras informações que poderiam ser usadas para diversas finalidades. No nosso caso, utilizaremos esses dados em um aplicativo que terá como meta trazer e motivar pessoas que não são fisicamente ativas a desenvolverem uma melhor qualidade de vida, aderindo um programa de exercícios físicos que melhor se adapta a sua condição. Nos próximos subcapítulos estão descritas a fundamentação teórica, materiais e métodos, resultados e discussão e conclusão.

---

## Fundamentação teórica

---

*Data mining*, ou mineração de dados, é a prática de examinar dados que já foram coletados – utilizando diversos tipos de algoritmos, normalmente de forma automática, a fim de gerar novas informações e encontrar padrões. (PRATES 2018, HOPPEN 2018, SANTOS 2018).

A principal ferramenta utilizada para a mineração de dados é a ferramenta *Weka*. Ela se consolidou como sendo a ferramenta mais utilizada em ambiente acadêmico. É aplicada não apenas em pesquisas científicas, mas principalmente para fins didáticos. O recurso mais comum e consequentemente o mais utilizado é a tarefa de classificação, que consiste em realizar a associação automática de um objeto a uma determinada classe, pertencente a um conjunto predefinido de classes. (PERBONI 2013).

Tomar decisões baseadas em dados é um requisito nos dias atuais, algumas áreas e setores de empresa sabem o quanto isso é essencial, por isso, utilizam *data mining* (Mineração de dados) para obterem informações relevantes. Mais importante que ter acesso a um “oceano” de informação é saber como navegar nele. O uso de ferramentas adequadas para *data mining* permite que sua empresa cruze dados e os classifique de acordo com sua prioridade, visualizando somente o que importa. (HEKIMA 2014).

Para Dino (2018), não é de hoje que as áreas responsáveis pela coleta e análise de dados são vistas como setores-chave para o sucesso das companhias. E a tendência é que esses segmentos fiquem mais em voga. Segundo estimativa realizada pela consultoria Gartner, o grande volume de informações geradas e administradas pelas empresas ainda não atingiu todo o seu potencial, mas que se tornará mais importante com o passar do tempo, já que dentro de cinco anos esses dados passarão a serem contabilizados como parte do patrimônio corporativo. Se antecipando ao cenário, organizações nacionais já encaram os dados com maior seriedade e conseguem aproveitar o seu potencial com precisão, ganhando destaque em suas áreas de atuação e conquistando bons resultados.

A carreira de cientista de dados deve ganhar mais destaque no futuro. Uma pesquisa realizada pela Michael Page, líder mundial no recrutamento de executivos, indica que os gestores de negócios em *Business Intelligence* e *Big Data*, junto aos cientistas de dados, serão os mais procurados no segmento de tecnologia da informação nos próximos anos. (DINO 2018).

De acordo com Rocha (2015), por mais que o progresso traga uma série de vantagens para a nossa vida, tudo tem seu lado negativo. Um exemplo disso são os aplicativos feitos para criar treinamentos para quem deseja entrar em forma. Hoje, são tantas as opções no mercado, que a escolha daquele que será nosso guia na hora de fazer nossas atividades chega a ser um verdadeiro dilema.

Um grupo de pesquisadores analisou os 30 *apps* de exercícios gratuitos mais populares para *iPhone* e os comparou com as orientações do *American College of Sports Medicine* (ACSM) a respeito de boas práticas de atividades físicas. De acordo com Francois Modave, um dos autores do estudo, o problema com a maioria desses aplicativos vem do fato de acabarem colocando os usuários em risco. Segundo ele, os *apps* não preparam as pessoas para fazer as atividades da melhor forma, além de não usarem técnicas apropriadas e não tratarem de questões de segurança. (ROCHA 2015).

Como a busca constante por aplicativos para facilitar tarefas no geral do cotidiano de cada um é grande, os aplicativos já existentes para o nicho em questão poderiam ser melhorados ou substituídos (SILVA 2017). Em lojas virtuais de *apps* como a *play store* e *apple store* é fácil encontrar aplicativos para te auxiliar na prática de atividades físicas. *Apps* como o *Virtuagym* (VIRTUAGYN 2019), *MapMyFitness* (MAPMYFITNESS 2019), *Nike Training* (NIKE TRAINING 2019), entre outros, todos esses *apps* entregam muitas funcionalidades, porém entregam respostas de treino padrões para seus usuários, mesmo que eles tenha perfis e limitações diferentes. Com este fato, surgiu a ideia de usar mineração de dados para obter e analisar preferências de diferentes grupos de pessoas ativas fisicamente e usar essas informações dentro de um futuro *app* para entregar um treino que melhor se adapte a cada perfil, consequentemente ajudando a criar hábitos melhores.

---

## Material e Métodos

---

Para utilizar os dados na ferramenta, tivemos primeiramente que elaborar a anamnese. Para isso, usamos a ferramenta de formulários da *Google* chamada *Google Forms*, e coletamos as respostas através da mesma. O questionário ([Apêndice A](#)) continha 20 perguntas selecionadas em conjunto com um educador físico. Em seguida, realizamos uma pesquisa de campo em 3 cidades do estado de Goiás, sendo elas Buriti Alegre, Itumbiara e Morrinhos, na qual coletamos respostas de 110 pessoas fisicamente ativas, número suficiente para realizar a mineração e análise de dados para criar a proposta do aplicativo.

As atividades propostas partiram dos estudos feitos sobre a necessidade de um *app* inovador, que entrega a praticantes de atividades físicas, recomendações de treino que melhor combina com o seu perfil. Para isso, primeiramente tivemos que elaborar uma anamnese, criamos um formulário com as perguntas da anamnese utilizando a ferramenta de formulários da *Google*, e posteriormente fizemos uma pesquisa de campo para coletar dados, foi feito o tratamento desses dados dentro de uma planilha, gerando assim um arquivo “.csv”, para ser usado na mineração de dados, para encontrar padrões que nos gere informações para utilizar no futuro *app*.

Ao participar da pesquisa de campo, alguns dos praticantes de atividades físicas questionaram a usabilidade dos dados no *app*. De início os entrevistados estavam com a ideia que o *app* seria “mais do mesmo”, após compreenderem que o *app* leva em consideração as atividades físicas que deixa os usuários mais confortáveis em fazer, isso sem deixar de indicar um treinamento completo, um entrevistado comentou: “é interessante ter um *app* que leva em consideração tipos de exercícios mais alinhado com os nossos gostos pessoais”, assim fortalecendo ainda mais a ideia de utilizar mineração de dados em prol do aplicativo de atividades físicas.

Montamos uma planilha contendo os resultados do questionário(anamnese), sendo que nas colunas estavam as perguntas e nas linhas as respostas. Convertimos alguns valores da planilha em números, para que fosse possível a criação de um algoritmo classificador ([Apêndice B](#)) que usa as respostas colocadas na planilha para classificar as 110 pessoas fisicamente ativas que responderam a anamnese entre iniciantes, intermediários ou avançados. Esta classificação foi crucial para usar os dados no futuro aplicativo com os fins planejados.

Os três indicadores, de acordo com o autor [Junior \(2013\)](#), da anamnese para a classificação foram: o tempo de prática (sendo a quantidades de meses que o indivíduo pratica atividades físicas); a frequência semanal (sendo quantos dias na semana é praticado as atividades físicas); a duração do treino (sendo medidos em minutos). Os três indicadores precisam estar dentro de certos padrões, como praticar exercícios mais de uma vez por semana e mais de 15 minutos por dia, para que o indivíduo seja classificado como fisicamente ativo e que pertença a um certo grupo de praticantes, caso contrário será classificado como sedentário.

O cruzamento dos níveis de classificação com a frequência cardíaca não foi utilizado na análise, pois foi constatado que foram utilizados diferentes métodos para medir a frequência cardíaca do praticante de atividade física, e para validar o uso desses dados o método de medir tal frequência deveria ser padronizado para todos os indivíduos.

Foi necessária a conversão da planilha em um arquivo com a extensão “.csv” para ser usando dentro da ferramenta *Weka*.

Para realizar a mineração dos dados utilizamos a ferramenta *Weka* e abrimos o arquivo “.csv”. Utilizamos a classificação entre os 3 níveis de pessoas fisicamente ativas gerada pelo algoritmo criado na planilha e definimos como indicador de decisão dentro da ferramenta *Weka*, para assim poder cruzar essas classificações com as respostas da anamnese, encontrando padrões para os praticantes de atividades físicas: avançados, intermediários e iniciantes.

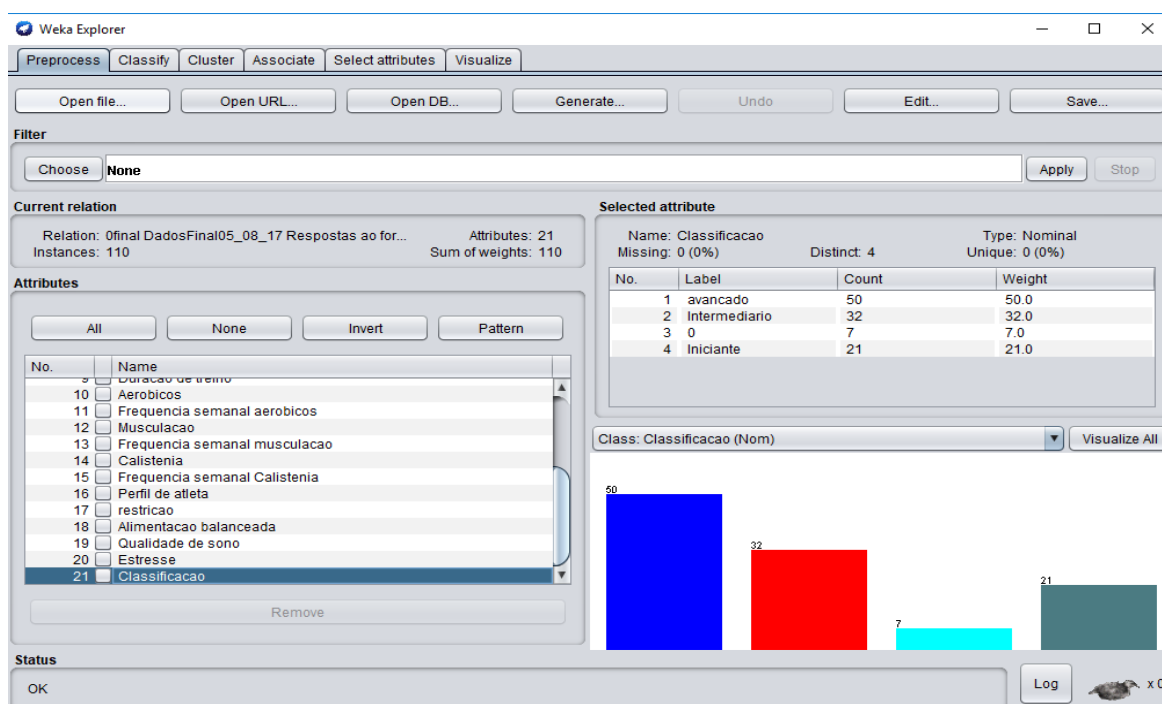
Os algoritmos de classificação testados foram os de árvore de decisão e o algoritmo de máquina de aprendizado representados na ferramenta *Weka* pelo programa *J48* e *SMO* respectivamente.

Também foram utilizados os algoritmos de seleção de atributos por ranqueamento: *ReliefAttributeEval*, *InfoGainAttributeEval*, *OneRAttributeEval*, com o objetivo de identificar os atributos que mais impactam na classificação de um usuário praticante de atividade física.



## Resultados e Discussão

Podemos notar na *Figura 1*, que 50 pessoas foram classificadas como avançadas, 32 pessoas como intermediárias e 21 como iniciantes. O algoritmo criado na planilha classificava apenas indivíduos fisicamente ativos, sendo que das 110 pessoas participantes, 103 foram classificadas como ativas fisicamente, pois as outras 7 pessoas foram classificadas como não ativas fisicamente e não pertenciam a nenhum grupo de praticantes.

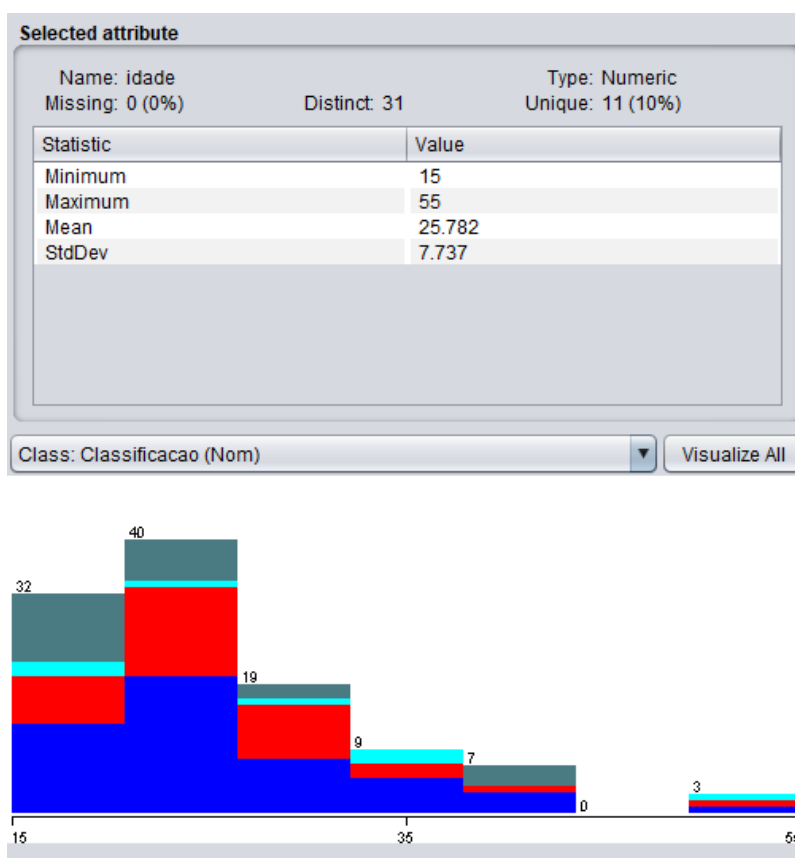


*Figura 1: Classificação dos indivíduos no Weka, azul corresponde a avançado, vermelho a intermediário, azul-claro a não classificados e cinza a iniciantes.*

Optamos por usar os níveis de classificação (iniciante, intermediário e avançado) como classificador, com a finalidade de ter as melhores respostas para utilizar em um *app*. O papel principal da mineração de dados é entregar dados de preferências de atividades físicas dos iniciantes, intermediários e avançados, o aplicativo utilizará as respostas e fará uma comparação com o perfil cadastrado do usuário dentro do futuro *app*, para indicar a atividade física que melhor se adapta aos usuários finais.

Nas Figuras mostradas abaixo, podemos notar o cruzamento da classificação dos níveis das pessoas que responderam a anamnese com as respostas de algumas das 20 perguntas. Em cada cruzamento, podemos ter diversas respostas, que analisadas em conjunto com um educador físico, nos permite criar as atividades físicas personalizadas para os usuários finais.

Como visto abaixo na *Figura 2*, podemos notar que a grande maioria das pessoas que responderam a anamnese têm menos que 35 anos e uma média 26 anos, boa parte destes indivíduos foram classificados como avançados ou intermediários, já a maioria dos iniciantes tem em média menos que 20 anos de idade. Assim encontramos um padrão de idade para maioria dos indivíduos avançados, intermediários e iniciantes.



*Figura 2: Idade em relação ao nível do praticante de atividades físicas*

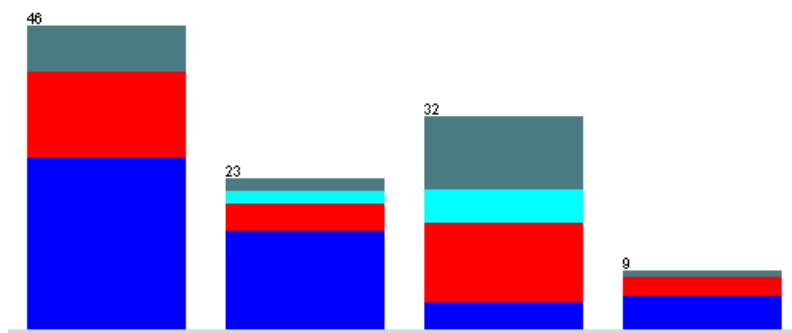
Abaixo na *Figura 3*, notamos que a maior parte das pessoas praticam atividades físicas por estética, seguido de saúde, por gostar de praticar e por último, outros. A maioria dos iniciantes praticam atividades físicas por saúde, os intermediários ficam divididos entre saúde e estética, já a maioria dos avançados praticam por estética, tendo uma parte considerável de praticantes que gostam de praticar.

**Selected attribute**

Name: Motivo para praticar      Type: Nominal  
Missing: 0 (0%)      Distinct: 4      Unique: 0 (0%)

No.	Label	Count	Weight
1	Estetica	46	46.0
2	Por gostar de praticar	23	23.0
3	Saude	32	32.0
4	Outros	9	9.0

Class: Classificacao (Nom)      Visualize All



*Figura 3: Motivos para praticar em relação ao nível do praticante de atividade física*

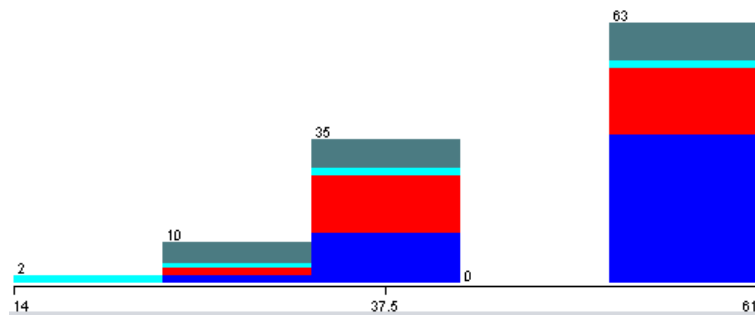
Já na *Figura 4*, notamos que o treino da maior parte de indivíduos avançados dura em média 60 minutos ou mais, dos indivíduos intermediários as respostas ficaram divididas entre 35 minutos a mais de 60 minutos, os indivíduos iniciantes também tem respostas de tempo de treino divididas, no caso em uma janela de 15 minutos a mais de 60 minutos.

**Selected attribute**

Name: Duracao de treino      Type: Numeric  
Missing: 0 (0%)      Distinct: 6      Unique: 2 (2%)

Statistic	Value
Minimum	14
Maximum	61
Mean	50.264
StdDev	12.39

Class: Classificacao (Nom)      Visualize All



*Figura 4: Duração do tempo de treino em relação ao nível do praticante de atividade física*

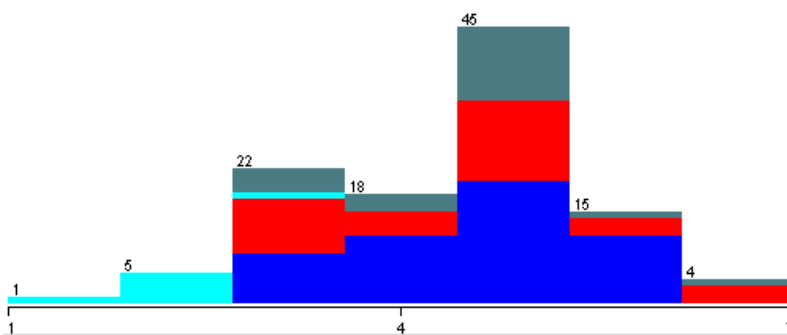
Logo abaixo, na *Figura 5*, podemos observar que a maior parte de cada grupo pratica atividades físicas 5 dias na semana. Podemos notar que a segunda maior parte dos indivíduos intermediários e iniciantes praticam 3 vezes na semana, já a segunda maior parte dos avançados praticam 6 vezes na semana.

**Selected attribute**

Name: Frequencia semanal      Type: Numeric  
Missing: 0 (0%)      Distinct: 7      Unique: 1 (1%)

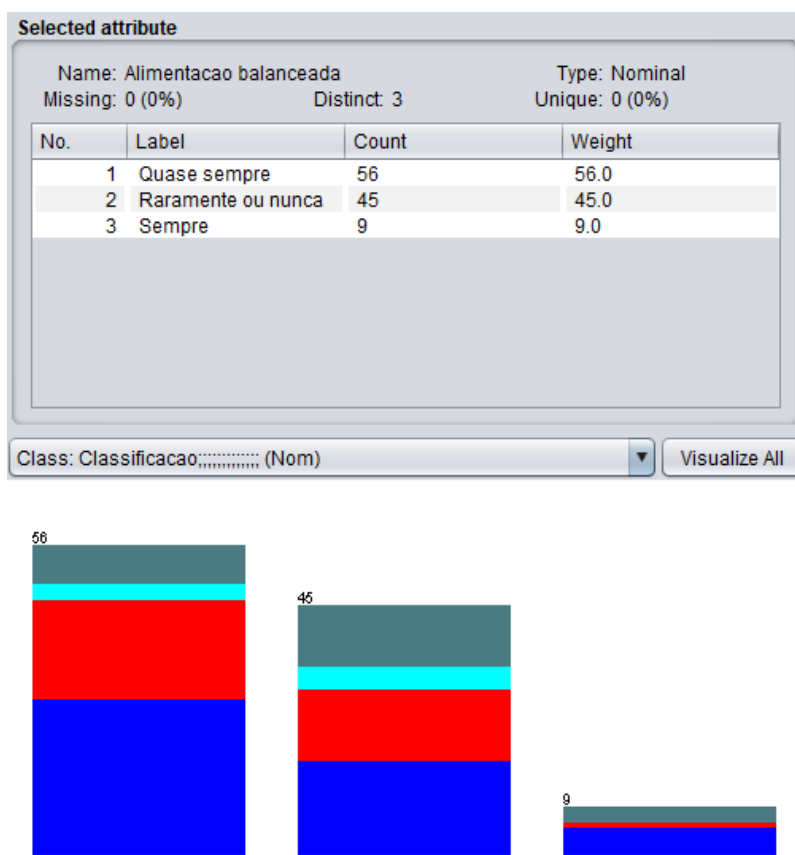
Statistic	Value
Minimum	1
Maximum	7
Mean	4.473
StdDev	1.239

Class: Classificacao (Nom)      Visualize All



*Figura 5: Frequência semanal em relação ao nível do praticante de atividade física*

Como visto abaixo na *Figura 6*, podemos notar que entre as pessoas responderam a anamnese aproximadamente a metade delas não segue uma alimentação equilibrada. Apenas 9 pessoas, sempre segue uma boa alimentação. Contudo nota-se que a maioria dos iniciantes estão no grupo que não segue uma boa alimentação, já a maioria dos indivíduos intermediários e avançados quase sempre seguem uma boa alimentação. Isso sugere que os indivíduos com o passar do tempo vão melhorando seus hábitos alimentares, acompanhado do seu nível de pessoa fisicamente ativa.



*Figura 6: Alimentação balanceada em relação ao nível do praticante de atividades físicas*

Apesar que alguns resultados nos entregou informações mais interessantes, podemos dizer que é possível encontrar padrões em praticamente todos os resultados, tendo como exceção somente a frequência cardíaca, que foi desconsiderada. Por exemplo: a partir dos resultados obtidos, percebemos que a maior parte dos iniciantes praticam atividades físicas por saúde, quando passam a ser intermediários e posteriormente avançados esse motivo para praticar, passa a ser na maioria das vezes por estética. Essa mudança na classificação do indivíduo acaba trazendo novos hábitos, como uma melhor alimentação.

Como dito por [Rocha \(2015\)](#), boa parte dos aplicativos de atividades físicas acabam colocando os usuários em risco. Com o uso de mineração de dados é possível melhorar a experiência dos que já são praticantes de atividades físicas, além de ajudar pessoas sedentárias a começar e a manter novos hábitos, com a prática de atividades físicas que melhor se adapta ao perfil desses usuários.

Os algoritmos classificadores conseguiram identificar as classes de acordo com os dados obtidos, deste modo é possível observar que os atributos possuem uma relação com as classificações.

Os resultados dos programas *J48* e *SMO* obtidos pelo *Weka* são:

Tabela 1: Resultados dos algoritmos classificadores

Programas	Classificação individual			Porcentagem total
	Avançada	Intermediária	Iniciante	
<i>J48</i>	0,911	0,787	0,773	84.5455
<i>SMO</i>	0,868	0,667	0,649	78.1818

A *tabela 1* apresenta o resultado dos algoritmos classificadores de árvore de decisão e máquina de aprendizado, ambos conseguiram realizar a classificação para todos os grupos (Avançado, intermediário e Iniciante) mostrando que é possível obter a classificação dos dados sem os indicadores de grupo. Sendo assim, é possível verificar identificar se o usuário é avançado, intermediário ou iniciante somente obtendo outras informações do atleta além do tempo de prática (sendo a quantidades de meses que o indivíduo pratica atividades físicas); a frequência semanal (sendo quantos dias na semana é praticado as atividades físicas); e a duração do treino (sendo medidos em minutos).

Estas informações são importantes pois muitas é possível direcionar as atividades praticadas pelo usuário não somente pelo tempo de prática (sendo a quantidades de meses que o indivíduo pratica atividades físicas), a frequência semanal (sendo quantos dias na semana é praticado as atividades físicas), e a duração do treino (sendo medidos em minutos) que são os indicadores, mas também pelo perfil do atleta, identificando melhor os objetivos de acordo com o nível que se encontra, bem como conseguindo melhores resultados com atividades mais assertivas.

A *Figura 7* apresenta a árvore de decisão como resultado do programa *J48*, com essa árvore de decisão é possível observar os principais fatores (atributos) classificadores são: Opinião pessoal do usuário, estresse, a duração do treino, o peso, os tipos de exercício (musculação, calistenia, exercícios aeróbicos), a qualidade do sono.

Os algoritmos de seleção de atributos também corroboraram com a árvore de decisão apresentando uma classificação de importância dos mesmos de modo semelhante. Como é apresentado nas figuras 8, 9 e 10. É importante observar que os algoritmos de seleção trabalham de modo diferente dos algoritmos classificadores. Apenas é um fator para complementar os resultados e nas discussões.

É possível observar que tanto a árvore de decisão como os algoritmos de seleção não consideraram como relevantes os atributos como idade, altura, gênero.

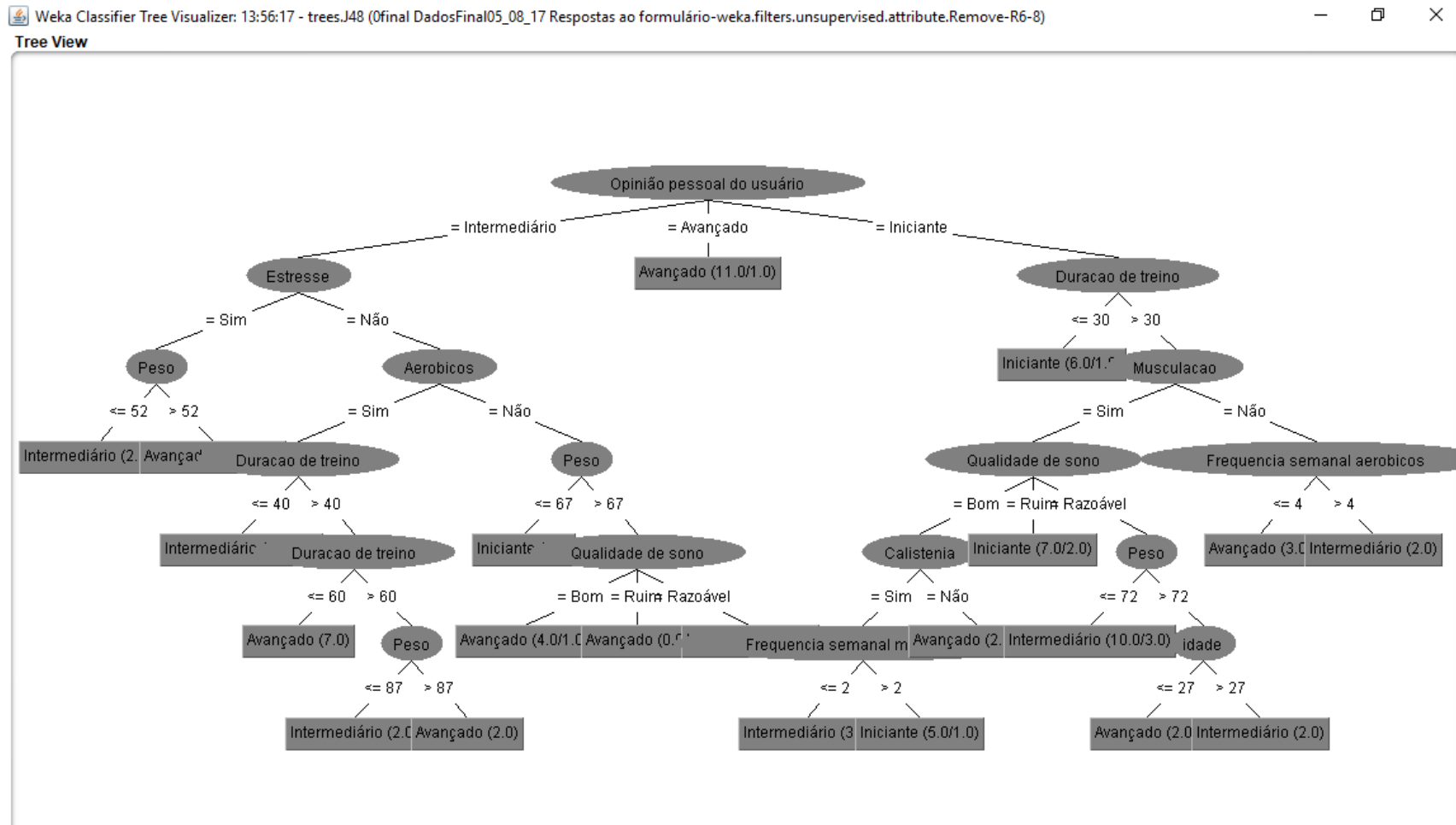


Figura 7: Árvore de decisão do programa J48



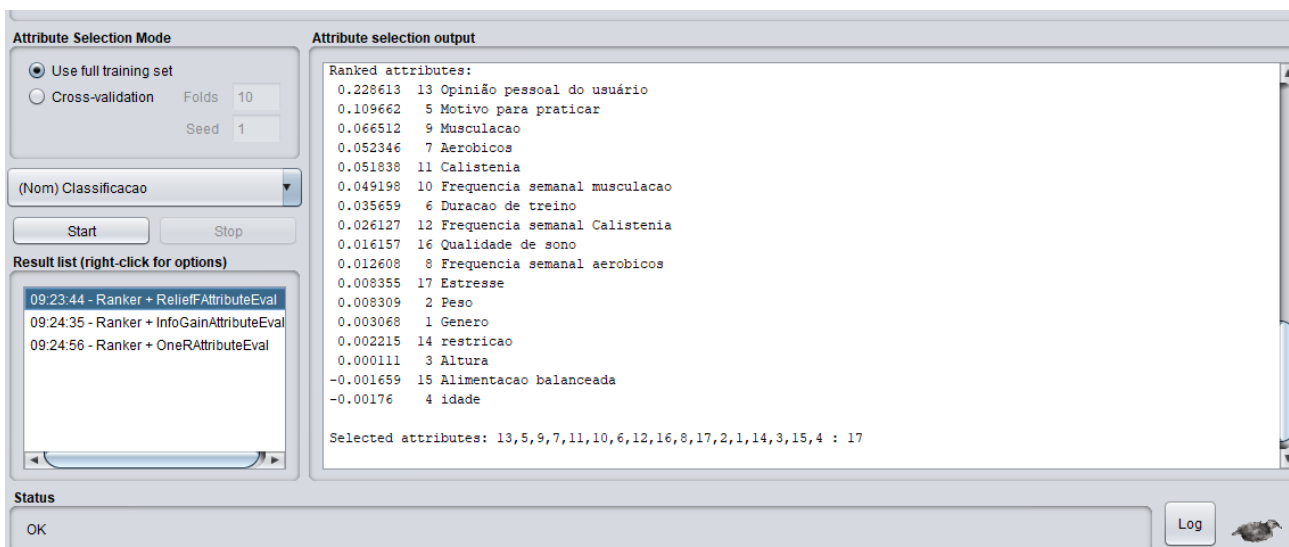


Figura 8: Resultado do ranqueamento do algoritmo de seleção de atributos *ReliefFAttributeEval*.

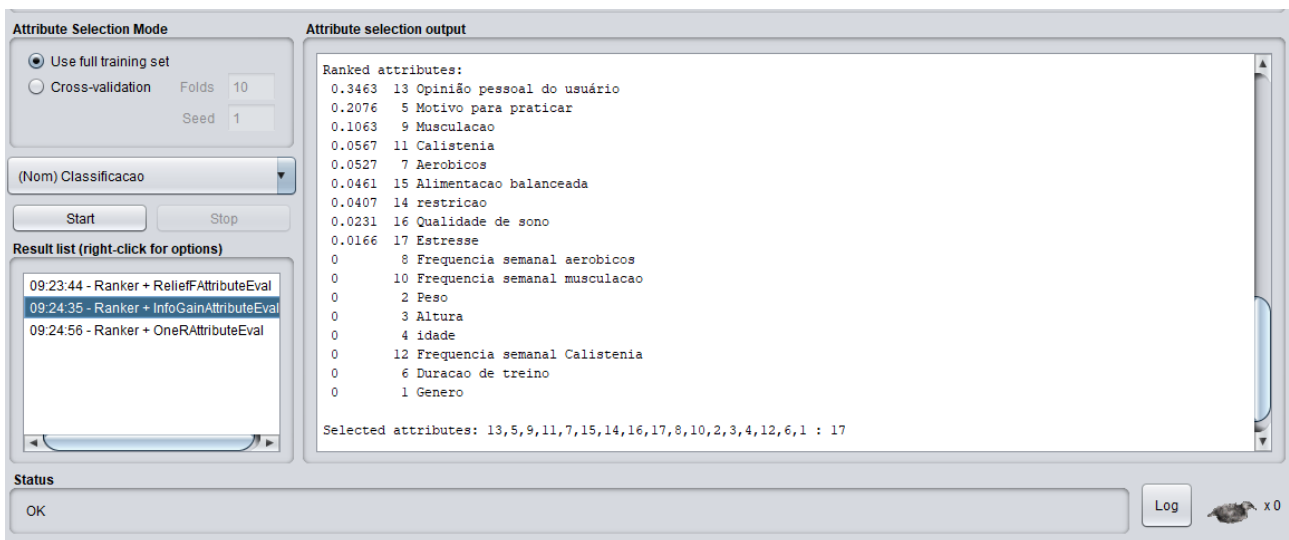


Figura 9: Resultado do ranqueamento do algoritmo de seleção de atributos *InfoGainAttributeEval*.

The screenshot displays a software window for attribute selection. On the left, the 'Attribute Selection Mode' section has 'Use full training set' selected, with 'Cross-validation' as an alternative. Parameters include 'Folds: 10' and 'Seed: 1'. Below this is a dropdown menu for '(Nom) Classificacao' and 'Start'/'Stop' buttons. A 'Result list (right-click for options)' shows three entries, with the third one, '09:24:56 - Ranker + OneRAttributeEval', highlighted. The main 'Attribute selection output' window contains a list of ranked attributes and a list of selected attributes.

**Attribute Selection Mode**

Use full training set  
 Cross-validation Folds: 10  
Seed: 1

(Nom) Classificacao

Start Stop

**Result list (right-click for options)**

- 09:23:44 - Ranker + ReliefAttributeEval
- 09:24:35 - Ranker + InfoGainAttributeEval
- 09:24:56 - Ranker + OneRAttributeEval

**Attribute selection output**

Ranked attributes:

52.7273	5	Motivo para praticar
51.8182	13	Opinião pessoal do usuário
50.9091	12	Frequencia semanal Calistenia
49.0909	11	Calistenia
49.0909	8	Frequencia semanal aerobicos
47.2727	7	Aerobicos
46.3636	2	Peso
45.4545	17	Estresse
45.4545	9	Musculacao
45.4545	1	Genero
45.4545	10	Frequencia semanal musculacao
45.4545	15	Alimentacao balanceada
45.4545	14	restricao
42.7273	16	Qualidade de sono
41.8182	3	Altura
41.8182	6	Duracao de treino
34.5455	4	idade

Selected attributes: 5,13,12,11,8,7,2,17,9,1,10,15,14,16,3,6,4 : 17

Status: OK Log x0

Figura 10: Resultado do ranqueamento do algoritmo de seleção de atributos OneRAttributeEval

---

## Conclusão

---

Com a elaboração da anamnese fizemos uma pesquisa de campo, onde fomos capazes de coletar respostas para as 20 perguntas do questionário (anamnese), e utilizando o algoritmo criado na planilha, classificamos as pessoas ativas fisicamente e que participaram da pesquisa entre iniciantes, intermediários e avançados. Por meio da ferramenta *Weka*, conseguimos cruzar essa classificação gerada pelo algoritmo com as respostas da anamnese. Esses resultados gerados têm como propósito serem utilizados para gerar um treino personalizado para diferentes grupos de praticantes de atividades físicas. Assim, a ideia é que quando a pessoa cadastrar o perfil no futuro *app*, o mesmo comparará com as preferências de perfis analisados resultando no treino que melhor combine com cada perfil.

Com o uso de mineração de dados, conseguimos informações para utilizar em nosso trabalho futuro (o desenvolvimento do *app*), buscando assim, aumentar o interesse das pessoas em relação a prática de atividades físicas, usando essas ferramentas para indicar um treino mais alinhado com o perfil dos usuários. Logo, o principal diferencial do *app* é facilitar a criação de treinos individuais, porém como o aplicativo não substitui um educador físico, buscaremos entregar o resultado mais próximo ao de um profissional.

O maior problema encontrado foi conseguir coletar um número de respostas (por meio da anamnese) suficiente para fazer a análise na ferramenta *Weka*, contudo, após 110 pessoas responderem a pesquisa, fizemos a mineração de dados e encontramos resultados para usarmos no fim planejado.

De acordo com os resultados e as discussões apresentados é possível concluir que com algumas informações aos usuários não é somente possível identificar a atividade física mais assertiva para o mesmo que modo que ele não somente consiga fazer as suas atividades físicas adequadamente com o seu perfil (iniciante, intermediário, avançado) como também atingir seus objetivos pessoais, como melhorar a saúde, estética, qualidade do sono, etc.

É possível realizar trabalhos futuros se utilizando desses resultados de mineração de dados, tal como a criação do *app* ou até mesmo utilizar os resultados de outras respostas por meio da ferramenta *Weka*, para conseguirmos confirmar ou não informações já conhecidas na ciência.

---

## Referências Bibliográficas

---

BARROS, C. (2014) “**As Facilidades da Tecnologia Avançada para Vida**”. Disponível em: [≤https://www.analiseagora.com/2014/05/as-facilidades-da-tecnologia-avancada.html/](https://www.analiseagora.com/2014/05/as-facilidades-da-tecnologia-avancada.html/)>. Acesso em: 18 de Março de 2019.

DINO (2018) “**Data Intelligence: 4 empresas que usam e provêm dados a seu favor**”. Disponível em: [≤https://www.terra.com.br/noticias/dino/data-intelligence-4-empresas-que-usam-e-provem-dados-a-seu-favor,2f8573f0f7bbc797892b8e6023e3faf6a1z270a2.html/](https://www.terra.com.br/noticias/dino/data-intelligence-4-empresas-que-usam-e-provem-dados-a-seu-favor,2f8573f0f7bbc797892b8e6023e3faf6a1z270a2.html/)>. Acesso em: 21 de Março de 2019.

HEKIMA (2014) “**Data mining é essencial pra sua empresa se destacar**”. Disponível em: [≤http://www.bigdatabusiness.com.br/por-que-a-mineracao-de-dados-e-essencial-para-as-empresas-que-querem-se-destacar/](http://www.bigdatabusiness.com.br/por-que-a-mineracao-de-dados-e-essencial-para-as-empresas-que-querem-se-destacar/)>. Acesso em: 18 de Março de 2019.

HOPPEN J., SANTOS M., PRATES W. R. (2018) “**O que é Data Mining (Mineração de Dados)?**”. Disponível em: [≤https://www.aquare.la/o-que-e-data-mining-mineracao-de-dados/](https://www.aquare.la/o-que-e-data-mining-mineracao-de-dados/)>. Acesso em: 18 de Março de 2019.

JUNIOR, A. G. M. (2017) “**Iniciante, Intermediário ou Avançado – Qual o seu nível de treino?**”. Disponível em: [≤http://alemdoanabolismo.com.br/iniciante-intermediario-ou-avancado-qual-o-seu-nivel-de-treino/](http://alemdoanabolismo.com.br/iniciante-intermediario-ou-avancado-qual-o-seu-nivel-de-treino/)>. Acesso em: 18 de Março de 2019.

MAPMYFITNESS (2019) “**MapMyFitness Personal Trainer**”. Disponível em: [≤https://play.google.com/store/apps/details?id=com.mapmyfitness.android2/](https://play.google.com/store/apps/details?id=com.mapmyfitness.android2/)>. Acesso em: 21 de Março de 2019.

MULLER, N. (2018) “**O Impacto da Tecnologia em Nossas Vidas**”. Disponível em: [≤https://www.oficinadanet.com.br/post/16174-o-impacto-da-tecnologia-em-nossas-vidas/](https://www.oficinadanet.com.br/post/16174-o-impacto-da-tecnologia-em-nossas-vidas/)>. Acesso em: 18 de Março de 2019.

NIKE TRAINING (2019) “**Nike Training Club – Treinos & Exercícios Fitness**”. Disponível em: [≤https://play.google.com/store/apps/details?id=com.nike.ntc/](https://play.google.com/store/apps/details?id=com.nike.ntc/)>. Acesso em: 21 de Março de 2019.

PEIXOTO, Y. S. (2017) “**MINERAÇÃO DE DADOS EM PROL DE ATIVIDADES FÍSICAS**”. Disponível em: [≤https://www.even3.com.br/Anais/ceictifgoiano/62251-MINERACAO-DE-DADOS-EM-PROL-DE-ATIVIDADES-FISICAS./>](https://www.even3.com.br/Anais/ceictifgoiano/62251-MINERACAO-DE-DADOS-EM-PROL-DE-ATIVIDADES-FISICAS./>). Acesso em: 18 de Março de 2019.

PERBONI, M. (2013) “**Mineração de Dados na Prática com Weka API**”. Disponível em: [<https://marcosvperboni.wordpress.com/2013/02/15/mineracao-de-dados-na-pratica-com-weka-api/>](https://marcosvperboni.wordpress.com/2013/02/15/mineracao-de-dados-na-pratica-com-weka-api/>). Acesso em: 18 de Março de 2019.

ROCHA, L. (2015) “**Maioria dos apps de exercícios do iPhone estão longe do ideal, diz estudo**”. Disponível em: [≤https://www.tecmundo.com.br/saude/86068-maioria-apps-exercicios-iphone-do-ideal-diz-estudo.htm/>](https://www.tecmundo.com.br/saude/86068-maioria-apps-exercicios-iphone-do-ideal-diz-estudo.htm/>). Acesso em: 21 de Março de 2019.

ROSA, R. (2019) “**Avaliação física e de desempenho: por que fazer?**”. Disponível em: [≤http://www.alvarorosa.com.br/atletas\\_2.asp/>](http://www.alvarorosa.com.br/atletas_2.asp/>). Acesso em: 20 de Março de 2019.

SILVA, J. O. M. (2017) “**COMPARAÇÃO DE SOFTWARES POPULARES PARA ATIVIDADES FÍSICAS COM O SYSAPPFITNESS**”. Disponível em: [≤https://www.even3.com.br/anais/ceictifgoiano/62431-comparacao-de-sofware-populares-para-atividades-fisicas-com-o-sysappfitness/>](https://www.even3.com.br/anais/ceictifgoiano/62431-comparacao-de-sofware-populares-para-atividades-fisicas-com-o-sysappfitness/>). Acesso em: 18 de Março de 2019.

VIRTUAGYN (2019) “**Virtuagyn**”. Disponível em: [≤https://play.google.com/store/apps/details?id=digifit.virtuagym.client.android&hl=pt\\_BR/>](https://play.google.com/store/apps/details?id=digifit.virtuagym.client.android&hl=pt_BR/>). Acesso em: 21 de Março de 2019.

---

## Perguntas da Anamnese

---

1- Qual o seu gênero?

Masculino  Feminino

2- Qual é seu peso? (em kg. ex: 80)

3- Qual é sua altura? (ex: 1,80)

4- Qual a sua idade?

5-O que mais motiva você a praticar atividades físicas?

Estética  Saúde  Por gostar de praticar  Outros

6- A quanto tempo pratica atividades físicas regularmente?

A menos de 6 meses  Entre 6 meses e 1 ano  A mais de 1 ano

7- Quantos dias na semana você pratica atividades físicas?

1 dia  2 dias  3 dias  4 dias  5 dias  6 dias  Todos os dias

8- Quanto tempo você treina no (s) dia (s) que pratica atividades físicas?

Menos de 15 min.  15 min.  30 min.  40 min.  1 h.  Mais de 1 h.

9- Você pratica exercícios aeróbicos?

Sim  Não

10- Em média qual sua frequência semanal de exercícios aeróbicos?

1 dia  2 dias  3 dias  4 dias  5 dias  6 dias  Todos os dias  Nenhuma opção

11- Você pratica musculação?

Sim  Não

12- Qual sua frequência semanal de musculação?

1 dia  2 dias  3 dias  4 dias  5 dias  6 dias  Todos os dias  Nenhuma opção

13- Você pratica exercícios com base na calistenia? (Ex: flexões, barras e outros exercícios que usa o peso do corpo).

Sim  Não

15- Perfil atleta: (Como você se define como atleta?).

Iniciante  Intermediário  Avançado

16- Você possui alguma restrição?

Sim  Não

17- Se na questão anterior sua resposta foi sim, qual é a sua restrição?

18- Em relação a sua alimentação, você segue uma dieta balanceada?

Sempre  Quase sempre  Raramente ou nunca

19- Qual sua qualidade de sono?

Bom  Razoável  Ruim

20- Você passa por muito estresse diariamente?

Sim  Não

---

## Algoritmo de Classificação da Planilha

---

```
=SE(G2=2;SE(H2>=3;SE(I2>=15;"avançado"));SE(G2=1;SE(H2>=3;SE(I2>=15;"Intermediario"));SE(G2=0;SE(H2>=3;SE(I2>=15;"Iniciante"));SE(G2=2;SE(H2<=2;SE(I2>=15;"Iniciante"));SE(G2=2;SE(H2>=3;SE(I2<=15;"Iniciante"))))))))
```

Observação:

G corresponde ao tempo de pratica, H corresponde a frequência semanal e I corresponde a duração do treino.