

UNIVERSIDADE NOVE DE JULHO – UNINOVE  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA E GESTÃO DO  
CONHECIMENTO

SAULO VINICIUS STOPA DE LIMA

FRAMEWORK DE INTELIGÊNCIA ARTIFICIAL PARA COMPREENSÃO E  
PREDIÇÃO DE FATORES QUE MOTIVAM A PRÁTICA DE ATIVIDADES  
FÍSICAS

São Paulo  
2023

UNIVERSIDADE NOVE DE JULHO – UNINOVE  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA E GESTÃO DO  
CONHECIMENTO

SAULO VINICIUS STOPA DE LIMA

FRAMEWORK DE INTELIGÊNCIA ARTIFICIAL PARA COMPREENSÃO E  
PREDIÇÃO DE FATORES QUE MOTIVAM A PRÁTICA DE ATIVIDADES  
FÍSICAS

Dissertação apresentada ao Programa  
de Pós-Graduação em Informática e  
Gestão do Conhecimento da UNINOVE  
como parte dos requisitos para  
obtenção do título de Mestre em  
Informática.

Orientador: Prof. Dr. Sidnei A. de Araújo.

Linha de pesquisa: LP2 – Sistemas  
Inteligentes

São Paulo  
2023

Lima, Saulo Vinicius Stopa de.

Framework de inteligência artificial para compreensão e  
predição de fatores que motivam a prática de atividades físicas.  
/ Saulo Vinicius Stopa de Lima. 2023.

92 f.

Dissertação (Mestrado) - Universidade Nove de Julho -  
UNINOVE, São Paulo, 2023.

Orientador (a): Prof. Dr. Sidnei Alves de Araújo.

## **PARECER – EXAME DE DEFESA**

Parecer da Comissão Examinadora designada para o exame de defesa do Programa de Pós-Graduação em Informática e Gestão do Conhecimento a qual se submeteu o aluno Saulo Vinicius Stopa de Lima

Tendo examinado o trabalho apresentado para obtenção do título de "Mestre em Informática e Gestão do Conhecimento", com Dissertação intitulada "FRAMEWORK DE INTELIGÊNCIA ARTIFICIAL PARA COMPREENSÃO E PREDIÇÃO DE FATORES QUE MOTIVAM A PRÁTICA DE ATIVIDADES FÍSICAS", a Comissão Examinadora considerou o trabalho:

Aprovado

Aprovado condicionalmente

Reprovado com direito a novo exame

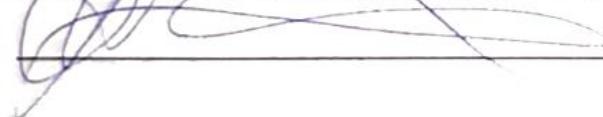
Reprovado

## **EXAMINADORES**

Prof. Dr. Sidnei Alves de Araújo - PPGI/Uninove (Orientador)



Prof. Dr. Paulo Roberto Garcia Lucareli - PPGCR/Uninove (Membro Externo)



Prof. Dr. André Felipe Henriques Librantz - PPGI/UNINOVA (Membro Interno)



São Paulo, 11 de agosto de 2023

## RESUMO

A ciência tem demonstrado que a prática de atividades físicas melhora a saúde, contribuindo na redução de riscos para doenças crônicas como diabetes, hipertensão, certos tipos de câncer, osteoporose e depressão. Contudo, um desafio comum para boa parte das pessoas é ter motivação para praticar atividades físicas regularmente, especialmente em virtude de cada vez menos elas terem necessidade de se deslocarem para realização de suas atividades cotidianas, por conta dos diversos recursos tecnológicos à sua disposição e o aumento das práticas de trabalho remoto, principalmente após a pandemia de COVID 19. Nesse cenário, torna-se importante o estudo e o desenvolvimento de ferramentas que facilitam a compreensão dos fatores que motivam a prática de atividades físicas visando auxiliar os profissionais de saúde na personalização dos atendimentos. Nesta pesquisa propõe-se um *framework* de Inteligência Artificial (IA) para auxiliar na compreensão e predição dos fatores que motivam as pessoas para a prática de atividades físicas, a partir de seus perfis socioeconômicos. Inicialmente, por meio de Mineração de Dados (MD), padrões descritos por regras do tipo SE...ENTÃO são gerados por Árvores de Decisão (AD) e pelo algoritmo Apriori. Na sequência, essas regras são empregadas na construção de um mecanismo de inferência fuzzy (MIF), o qual compõe um Sistema de Recomendações (SR) para indicar os fatores que motivam uma pessoa e as atividades mais adequadas, com base no seu perfil. Nos experimentos conduzidos foi utilizada uma base de dados de 140 pessoas, cedida por uma empresa do ramo de desenvolvimento de aplicativos para treinos online. Os resultados obtidos na MD (acurárias  $\geq 90\%$  e índices Kappa  $\geq 62\%$ ) evidenciando a presença de padrões consistentes e os desfechos alcançados pelo MIF, quando acoplado ao SR, mostram que o *framework* desenvolvido pode ser útil para os profissionais de educação física na orientação e elaboração de treinos personalizados para cada pessoa.

**Palavras-chave:** Atividades Físicas, Fatores Motivacionais, Inteligência Artificial, Árvores de Decisão, Mineração de Dados, Lógica Fuzzy.

## ABSTRACT

Science has demonstrated that practicing physical activities improves health, contributing to reducing the risk of chronic diseases such as diabetes, hypertension, certain types of cancer, osteoporosis and depression. However, a common challenge for most people is to have the motivation to practice physical activities regularly, especially as they have less and less need to travel to carry out their daily activities, due to the various technological resources at their disposal and the increase in remote work practices, especially after the COVID 19 pandemic. In this scenario, it is important to study and develop tools that facilitate the understanding of the factors that motivate the practice of physical activities in order to assist health professionals in customizing of services. In this research, an Artificial Intelligence (AI) framework is proposed to help understand and predict the factors that motivate people to practice physical activities, based on their socioeconomic profiles. Initially, through Data Mining (DM), patterns described by IF...THEN rules are generated by Decision Trees (DT) and the Apriori algorithm. Subsequently, these rules are used to build a *fuzzy* inference mechanism (FIM), which forms a Recommendation System (RS) to indicate the factors that motivate a person and the most appropriate activities, based on their profile. In the experiments conducted, a database of 140 people was used, provided by a company in the field of developing applications for online training. The results obtained in the DM (accuracies  $\geq 90\%$  and Kappa indices  $\geq 62\%$ ) showing the presence of consistent patterns and the outcomes achieved by the FIM, when coupled with the RS, show that the developed framework can be useful for education professionals physics in the guidance and development of personalized training for each person.

**Keywords:** Physical Activities, Motivational Factors, Artificial Intelligence, Decision Trees, Data Mining, *Fuzzy Logic*.

## LISTA DE SIGLAS

AD	Árvores de Decisão
AG	Algoritmos Genéticos
AF	Atividades Físicas
AM	Aprendizagem de Máquina
CSS	<i>Cascading Style Sheets</i> (folhas de estilo em cascata)
CSV	<i>Comma Separated Values</i> (valores separados por vírgulas )
GPS global)	<i>Global Positioning System</i> (sistema de posicionamento global)
HIIT	<i>High-intensity interval training</i> (treinamento intervalado de alta intensidade)
HTML	<i>HyperText Markup Language</i> (linguagem de marcação de hipertexto)
IA	Inteligência Artificial
IHC	Interação Humano Computador
IMC	Índice de Massa Corpórea
IMPRAF	Inventário de Motivação à Prática Regular de Atividade Física
JavaScript	Linguagem de programação interpretada estruturada
KDD	<i>Knowledge-discovery in Databases</i> (descoberta de conhecimento em bancos de dados)
Kernel	Núcleo do Sistema Operacional
MATLAB	<i>Matrix Laboratory</i> (software para cálculo numérico)
MD	Mineração de Dados
MIF	Mecanismo de Inferência <i>Fuzzy</i>
MongoDB	Software de banco de dados orientado a documentos

NextJS	Estrutura da <i>web</i> de desenvolvimento estático
OMS	Organização Mundial da Saúde
PLN	Processamento de Linguagem Natural
QMPAFS	Questionário de Motivação para a Prática de Atividade Física Sistematizada
QSDS	Questionário Sociodemográfico Simples
RNA	Redes Neurais Artificiais
RNC	Redes Neurais Convolucionais
SPSS	<i>Statistical Package for the Social Sciences</i> (pacote estatístico para as ciências sociais)
SR	Sistema de Recomendação
SVM	<i>Support-vector Machine</i> (máquinas de vetores de suporte)
TIC	Tecnologia da Informação e Comunicação
TypeScript	Linguagem de programação fortemente tipada

## LISTA DE TABELAS

Tabela 1 – Base de dados sobre as condições para jogar ou não tênis.....	29
Tabela 2 – Base de dados para representar o exemplo do funcionamento do algoritmo Apriori.....	31
Tabela 3 – Exemplo de matriz de confusão .....	33
Tabela 4 – Níveis de concordância do coeficiente Kappa .....	34
Tabela 5 – Distribuição dos respondentes de acordo com os atributos socioeconômicos .....	37
Tabela 6 – Conjuntos <i>fuzzy</i> para variáveis de entrada.....	42
Tabela 7 – Matriz de confusão obtida na classificação dos dados para o fator “Condicionamento Físico” .....	49
Tabela 8 – Matriz de confusão obtida na classificação dos dados para o fator “Estética” .....	50
Tabela 9 – Matriz de confusão obtida na classificação dos dados para o fator “Saúde” .....	51
Tabela 10 – Matriz de confusão obtida na classificação dos dados para o fator “Integração Social” .....	52
Tabela 11 – Matriz de confusão obtida na classificação dos dados para o fator “Ansiedade/Estresse” .....	53
Tabela 12 – Matriz de confusão obtida na classificação dos dados para o fator “Disponibilidade de Tempo” .....	54

## LISTA DE FIGURAS

Figura 1 – Uma taxonomia da motivação humana .....	20
Figura 2 – Etapas do Processo <i>KDD</i> .....	23
Figura 3 – Árvore de decisão sobre “jogar tênis ou não” .....	30
Figura 4 – Exemplo de um Mecanismo de Inferência <i>Fuzzy</i> (MIF) aplicado à saúde .....	32
Figura 5 – Diagrama esquemático do encadeamento das técnicas de IA no <i>framework</i> proposto.....	38
Figura 6 – Arquitetura do MIF .....	41
Figura 7 – Funções de pertinência associadas à variável de entrada Estado Civil .....	43
Figura 8 – Funções de pertinência associadas à variável de entrada Cor .....	43
Figura 9 – Funções de pertinência associadas à variável de entrada Escolaridade .....	43
Figura 10 – Funções de pertinência associadas à variável de entrada Faixa Salarial .....	44
Figura 11 – Funções de pertinência associadas à variável de entrada Faixa Etária .....	44
Figura 12 – Funções de pertinência associadas à variável de entrada Gênero	44
Figura 13 – Funções de pertinência associadas às variáveis de saída .....	45
Figura 14 – Protótipo esquemático do Sistema de Recomendação .....	46
Figura 15 – Exemplos de telas de entrada dos dados do SR .....	47
Figura 16 – Exemplo de predição do MIF .....	59
Figura 17 – Alguns resultados do MIF para o fator Condicionamento Físico ...	59
Figura 18 – Alguns resultados do MIF para o fator Saúde .....	60
Figura 19 – Alguns resultados do MIF para o fator Estética.....	60
Figura 20 – Alguns resultados do MIF para o fator Ansiedade/Estresse .....	60
Figura 21 – Alguns resultados do MIF para o fator Integração Social .....	61
Figura 22 – Alguns resultados do MIF para o fator Disponibilidade de Tempo	61
Figura 23 – Alguns resultados do MIF para o fator Custo .....	61
Figura 24 – Tela de resultado do SR.....	63

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>10</b>
1.1	Contextualização do tema .....	10
1.2	Trabalhos correlatos e lacunas de pesquisa .....	13
1.3	Objetivos .....	16
1.3.1	Objetivo geral .....	16
1.3.2	Objetivos específicos .....	16
1.4	Justificativa e motivação.....	16
1.5	Limitações da pesquisa .....	17
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>18</b>
2.1	Motivações para a prática de atividades físicas .....	18
2.2	Inteligência Artificial (IA) .....	21
2.2.1	Diferenças entre IA e Estatística Descritiva na análise de dados	21
2.3	<i>Knowledge Discovery in Databases – KDD</i> .....	22
2.3.1	Etapas do <i>KDD</i> .....	23
2.4	Algumas técnicas de IA empregadas em MD.....	25
2.4.1	Redes Neurais Artificiais (RNA).....	25
2.4.2	Algoritmos Genéticos (AGs) .....	26
2.4.3	<i>K-means</i> .....	26
2.4.4	<i>Support Vector Machines (SVM)</i> .....	27
2.4.5	Redes Bayesianas .....	27
2.4.6	Árvores de Decisão .....	27
2.4.7	Algoritmo Apriori.....	30
2.4.8	Lógica <i>Fuzzy</i> .....	31
2.4.9	Medidas de desempenho de algoritmos de classificação empregados em MD .....	33
2.5	Sistemas de recomendação.....	34
<b>3</b>	<b>MATERIAIS E MÉTODOS.....</b>	<b>36</b>
3.1	Base de dados .....	36
3.2	<i>Framework</i> proposto.....	37

3.3	Sistema de recomendação (SR) .....	46
<b>4</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>48</b>
4.1	Mineração de Dados (MD).....	48
4.1.1	Análise dos dados por meio de Árvores de Decisão (classificação) 48	
4.1.2	Geração de regras de associação usando o algoritmo Apriori....	55
4.2	Mecanismo de Inferência <i>Fuzzy</i> (MIF).....	59
4.3	Protótipo do Sistema de Recomendação .....	62
4.4	Discussão.....	64
<b>5</b>	<b>CONCLUSÕES E PROPOSTA PARA TRABALHOS FUTUROS .....</b>	<b>68</b>
	<b>REFERÊNCIAS.....</b>	<b>70</b>
	<b>APÊNDICE I – AD relativa ao fator motivacional CONDICIONAMENTO FÍSICO .....</b>	<b>78</b>
	<b>APÊNDICE II – AD relativa ao fator motivacional ESTÉTICA .....</b>	<b>79</b>
	<b>APÊNDICE III – AD relativa ao fator motivacional SAÚDE.....</b>	<b>80</b>
	<b>APÊNDICE IV – AD relativa ao fator motivacional INTEGRAÇÃO SOCIAL</b>	<b>81</b>
	<b>APÊNDICE V – AD relativa ao fator motivacional ANSIEDADE/ESTRESSE</b> <b>82</b>	
	<b>APÊNDICE VI – AD relativa ao fator motivacional DISPONIBILIDADE DE TEMPO .....</b>	<b>83</b>
	<b>APÊNDICE VII – AD relativa ao fator motivacional CUSTO .....</b>	<b>84</b>
	<b>ANEXO A – “Pesquisa Motivacional para Prática de Atividades Físicas”</b>	<b>85</b>

## 1 INTRODUÇÃO

Este capítulo contempla, primeiramente, a contextualização do tema explorado nesta pesquisa (seção 1.1). Na sequência são apresentados trabalhos da literatura explorando a aplicação de estatística descritiva e de IA na prática de atividades físicas, e as lacunas de pesquisa (seção 1.2). Por fim, são apresentados os objetivos (seção 1.3), a justificativa (seção 1.4) e as limitações deste trabalho (seção 1.5).

### 1.1 Contextualização do tema

O exercício ou as atividades físicas são benéficos na redução dos riscos cardiovasculares em pacientes com doenças crônicas, como diabetes e hipertensão. De acordo com um relatório da OMS de 2022, o sedentarismo pode levar 500 milhões de pessoas a desenvolverem doenças cardíacas, obesidade, diabetes e outras doenças não transmissíveis até 2030 (OMS, 2022). Somente 30% das nações têm diretrizes nacionais de atividades físicas para todas as faixas etárias. Normalmente 50% dos indivíduos que iniciam um programa de exercícios param nos primeiros seis meses. O desafio geralmente enfrentado é ter a motivação certa para incluir exercícios físicos na rotina diária, encontrar tempo para isso em uma agenda lotada e seguir o regime por um longo período (SHETTIGAR *et al.*, 2019).

Nos últimos anos, as pessoas têm cada vez menos a necessidade de se deslocar, por conta do trabalho remoto e em virtude da realização de várias atividades cotidianas à distância, através de sites ou de aplicativos para dispositivos móveis (PATRI *et al.*, 2017).

As recomendações globais da Organização Mundial da Saúde – OMS, sobre atividades físicas e seus benefícios, preconizam que crianças e adolescentes devem praticar atividades físicas pelo menos 60 minutos diárias, 3 vezes por semana. Para adultos a recomendação é de 150 a 300 minutos de atividades físicas aeróbica de moderada intensidade, ou pelo menos 75 a 150 minutos de atividades físicas aeróbica de vigorosa intensidade; ou ainda uma combinação equivalente de atividades físicas de moderada e vigorosa intensidade ao longo da semana. Já os idosos devem realizar pelo menos 150 a

300 minutos de atividades físicas aeróbica de moderada intensidade, ou pelo menos 75 a 150 minutos de atividades físicas aeróbica de vigorosa intensidade (OMS, 2022).

A falta de saúde pode acarretar diversos prejuízos financeiros para a sociedade, impactando tanto os indivíduos quanto o sistema de saúde como um todo. De acordo com um estudo da Organização Pan-Americana da Saúde – OPAS, publicado em setembro de 2018, mais de 1,4 bilhão de adultos correm risco de desenvolver doenças por inatividade física em todo o mundo (OPAS, 2018). Assim, muitas pesquisas científicas têm investigado essa questão e apontado algumas consequências econômicas significativas e prejuízos financeiros associados à falta de saúde como custos diretos de cuidados médicos.

De acordo com um estudo do Ministério da Saúde (MS) publicado em maio de 2022 sobre os impactos econômicos da inatividade física nos sistemas de saúde do Brasil, 15% das internações pelo Sistema Único de Saúde (SUS) pode ser atribuído à inatividade física, resultando em um custo aproximado de R\$275.646.877,64 (MS, 2022). Assim, o MS tem proposto ações na atenção primária à saúde entre as quais está a elaboração do Guia de Atividade Física para População Brasileira, que aborda diretrizes para prática de atividades físicas.

De acordo com Nunes (2022), cerca de 47% dos brasileiros são sedentários, sendo que na população jovem esse percentual pode atingir 84%, o que é alarmante. O autor destaca ainda que isso pode ser um grande problema para a sociedade tendo em vista que quanto mais doente uma população, maiores também são os custos com absenteísmo (falta ao trabalho).

Em situações em que pessoas enfrentam problemas de saúde que demandam cuidados ou assistência a longo prazo, é comum que familiares ou amigos intervenham como cuidadores informais. Essa circunstância pode acarretar em perda de renda e diminuição da produtividade dos cuidadores, pois muitas vezes eles precisam reduzir a carga horária de trabalho ou abrir mão do emprego para fornecer os cuidados necessários (DELAVANDE *et al.*, 2013).

Condicionamento físico, estética, saúde, integração social, ansiedade, disponibilidade de tempo e custo são alguns dos fatores que durante o treinamento físico tem um efeito maior nos componentes psicológicos, por exemplo, o de sentimentos positivos como o de alegria e de eficácia, que estão intimamente associados ao componente do bem-estar subjetivo na influência da prática de atividades físicas (GERBER *et al.*, 2021; ARCAS *et al.*, 2021; RYAN & DECI 2017; SCHÜLER, 2020).

Os fatores que motivam as atividades físicas regulares têm sido divididos em intrínsecos e extrínsecos, sendo os intrínsecos classificados em 3 tipos: para saber (experiência no aprendizado), para realizar (na realização da mesma pelo prazer da simples execução) e para experiência (situações que o estimulam). A motivação intrínseca está relacionada à prevenção de doenças, condição física, sentimentos de compaixão, diversão, alegria, eficácia e autonomia (BRIÈRE *et al.*, 1995).

A amotivação é também conhecida da literatura como uma construção motivacional identificada em certas pessoas incapazes de perceberem algum motivo para a prática de atividades físicas. Elas não veem nenhum benefício, ou não conseguem executar tais atividades de maneira satisfatória, no seu ponto de vista (BRIÈRE *et al.*, 1995).

Os fatores que motivam as pessoas a praticarem atividades físicas têm sido investigados nos últimos anos com auxílio de técnicas de estatística descritiva e de inteligência artificial (IA). Entre as diversas pesquisas estão as de SHETTIGAR *et al.* 2019; LIMA *et al.*, 2020; GERBER *et al.*, 2021; 2020; ARCAS *et al.*, 2021; RYAN & DECI 2017 e SCHÜLER, 2020.

Enquanto a estatística descritiva tem por objetivo descrever os dados analisados, a IA pode ser caracterizada como uma tecnologia de elaboração de sistemas computacionais capazes de reproduzir ou emular comportamentos inteligentes e pertinentes aos seres humanos. O aprendizado de máquina e o aprendizado profundo são tópicos da IA que podem ser aplicados em diversos campos do conhecimento (BHAVE *et al.*, 2020).

Em comparação com a estatística descritiva, amplamente empregada nos estudos para sumarizar conjuntos de dados por meio de medidas padronizadas

como média, variância, desvio padrão e histograma, a IA tem como principal vantagem o poder de descobrir padrões e de propiciar previsões com grande precisão a partir dos dados analisados. Contudo, atualmente ainda são poucos os estudos da literatura propondo a aplicação de IA na temática aqui investigada, sendo alguns deles apresentados na seção a seguir, na qual também estão identificadas as lacunas de pesquisa.

## 1.2 Trabalhos correlatos e lacunas de pesquisa

Para identificação das lacunas de pesquisa, o primeiro passo foi realizar uma revisão da literatura considerando o período de 2018 a 2022, levando em conta as seguintes palavras-chave (combinadas em Português e em Inglês): “Atividade Física”, “Fator Motivacional”, “Inteligência Artificial”, “Aprendizagem de Máquina”, e “Mineração de Dados”. A pesquisa foi feita nas bases Science Direct, Scopus e também no Google Scholar, resultando em 563 trabalhos, boa parte deles concentrada nos últimos 7 anos.

Após uma filtragem para selecionar os trabalhos mais relevantes (com base no número de citações), restaram 12 trabalhos, entre os quais estão aqueles que empregam Inteligência Artificial (IA), Aprendizagem de Máquina (AM) e Mineração de Dados (MD).

Dentre os trabalhos relacionados, os estudos utilizando estatística descritiva exploraram a prescrição de exercícios físicos e realização de diagnósticos, visando identificar falta de tempo, falta de motivação, motivação autônoma, e experiência de aprendizado com o uso de games na educação física. Entretanto, de acordo com os autores, a inclusão das atividades físicas na rotina diária para mudança de estilo de vida e a possibilidade da criação de novas perspectivas de aprendizado não trouxeram contribuições significativas de novos conhecimentos para a prática de atividades físicas (SHETTIGAR *et al.*, 2019; WELTER *et al.*, 2020; EMONSON *et al.*, 2020; ANDERSEN *et al.*, 2021; NUNES *et al.*, 2022).

No que tange aos trabalhos empregando IA, Baker *et al.* (2019) realizaram um estudo entre a relação no uso do Facebook (FB), relacionamento e motivação para a prática do exercício físico, porém o estudo se limitou a estudantes de

graduação e um número muito baixo de participantes de atividades físicas, que não deixou claro a identificação do potencial lado desadaptativo do FB resultante de comparações sociais, frustração de necessidades ativas (por exemplo, *bullying*) e negligência de necessidades (por exemplo, sentir-se excluído).

Um outro estudo utilizando IA foi realizado com o uso de processamento de linguagem natural (PLN) para entender perfis motivacionais que mantém a prática de atividades físicas utilizando aplicativo móvel (ASWANI *et al.*, 2018), no qual os resultados mostraram que a motivação e os objetivos seguidos pela crença nos resultados e na saúde são fortes preditores da manutenção das atividades físicas, porém apesar da pesquisa fazer uso de técnicas de IA através do algoritmo K-Means para agrupamento, o modelo usado neste estudo para tarefas de PLN não leva em consideração a ordem em que as palavras aparecem em uma frase, nem o discurso dos rótulos, gerando difícil interpretação para classificação e agrupamento das palavras.

Uma abordagem de IA utilizando tecnologia como *wearables* (relógios inteligentes) combinadas com recursos digitais de mudança de comportamento foi proposta por Ball *et al.* (2020) para analisar o aumento da prática de atividades físicas. Os resultados apontaram que os adolescentes usaram os *wearables* diariamente após a abordagem, com a percepção do aumento da motivação para a prática de atividades físicas, porém a pesquisa apresentou dados através de estatística descritiva e dados qualitativos analisados por tema que não descrevem os fatores da baixa adesão e do engajamento na prática de atividades físicas principalmente por não terem conseguido diferenciar os comportamentos de acordo com o sexo dos participantes.

Na educação física, a IA está presente também em avaliações e prescrições de exercícios físicos, através de aplicativos que realizam avaliação de composição corporal, controle de gasto calórico, dentre outros, os quais vêm se multiplicando nos últimos anos produzindo novos modos de vida e alternativas para a prática de atividades físicas (FRAGA & OLIVEIRA, 2020), que a partir desses instrumentos tecnológicos, se faz possível a orientação da prática das atividades físicas com planejamento e particularidades do usuário para os seus treinos.

Pereira (2019) desenvolveu um trabalho com o objetivo de descrever o uso de aplicativos possibilitados pelas Tecnologias da Informação e Comunicação (TICs) e das técnicas de Interação Humano-Computador (IHC), para facilitar o controle da alimentação e da prática de exercícios. Entretanto, os resultados encontrados identificaram a necessidade da utilização de elementos mobilizadores que gerem compartilhamento de dados que podem ser fatores contemporâneos com potencial mobilizador e que colabora para que as pessoas mudem seu estilo de vida e possam usufruir de diversos fatores benéficos, como tabelas comparativas entre os usuários do *software* do aplicativo mantendo um caráter colaborativo.

Um trabalho sobre a IA e seu espaço de inovação na educação física durante o COVID-19 foi proposto por Pilay *et al.* (2021) para mostrar que a educação física virtual se tornou um instrumento eficaz que pode contribuir para a qualidade de vida das pessoas isoladas. O estudo aponta que determinadas situações podem maximizar o potencial do aluno, apesar de não mencionarem quais são as medidas exigidas.

O uso de IA por meio de árvores de decisão (ADs) foi proposto no estudo de Sulino (2020), que explora a prática de atividades físicas e os benefícios trazidos pela disponibilidade de dispositivos e aplicativos móveis capacitados a registrar e armazenar as informações da atividade, tais como coordenadas de GPS, batimento cardíaco, altimetria e velocidade, como ocorre no popular aplicativo Strava, que oferece ao profissional de educação física uma ferramenta de análise para auxiliá-lo no planejamento, acompanhamento e ajustes relacionados ao treinamento esportivo. No entanto, a pesquisa se limitou a um grupo específico de corrida e ciclismo com análise de fatores de desempenho, clima e qualidade de vida, não analisando os fatores motivacionais para a prática de atividades físicas.

Analizando-se os estudos apresentados, pode-se observar que boa parte deles possui foco em informações analisadas por meio de estatística descritiva, incluindo melhores orientações sobre práticas de exercícios e orientações sobre potenciais que cada pessoa pode alcançar. Porém, são poucos os trabalhos que exploram os graus de motivação das pessoas (praticantes ou não) para a prática de atividades físicas. Observa-se uma carência de pesquisas que visam a

aplicação de IA para identificação e/ou compreensão dos fatores que levam a adesão e engajamento na prática de atividades físicas, principalmente levando em conta o perfil socioeconômico da pessoa.

Neste contexto, emerge a seguinte questão de pesquisa: como aplicar IA na compreensão e predição de fatores que motivam uma pessoa para a prática de atividades físicas com base no seu perfil socioeconômico?

### **1.3 Objetivos**

#### **1.3.1 Objetivo geral**

Desenvolver um *framework* de Inteligência Artificial para auxiliar na compreensão e predição dos fatores que motivam a prática de atividades físicas por praticantes e não praticantes.

#### **1.3.2 Objetivos específicos**

1.3.2.1 Investigar a aplicabilidade e limitações das abordagens encontradas na literatura para compreensão dos fatores que motivam a prática de atividades físicas;

1.3.2.2 Aplicar técnicas de Mineração de Dados (MD) para identificar, a partir de um conjunto dados de praticantes e não praticantes, os fatores que motivam a prática de atividades físicas em função dos perfis socioeconômicos das pessoas;

1.3.2.3 Desenvolver um Mecanismo de Inferência *Fuzzy* (MIF) usando os resultados obtidos na MD para fazer predições sobre fatores motivacionais a partir dos dados que descrevem o perfil socioeconômico de uma pessoa;

1.3.2.4 Desenvolver um Sistema de Recomendações (SR), cujo núcleo é composto pelo MIF, para auxiliar os profissionais de saúde no que tange as orientações sobre as práticas de atividades físicas;

### **1.4 Justificativa e motivação**

A identificação dos fatores que motivam a prática de atividades físicas exerce um papel na obtenção de conhecimento na área da saúde, contribuindo nos atendimentos realizados por profissionais de saúde que atuam no combate ao sedentarismo e excesso de peso (ANDRADE, 2021). Dessa forma, a MD

torna-se essencial para a descoberta de padrões nos dados sobre a prática de atividades físicas. Os dados sobre atividades físicas são diversos e difíceis de serem associados. Por este motivo, técnicas de MD como Árvores de Decisão (AD) e o algoritmo Apriori mostram-se adequadas na associação dessas informações.

Em adição, o emprego da lógica *fuzzy* tem o objetivo de predizer os fatores motivacionais para a prática de atividades físicas regular com base nas preferências e no perfil socioeconômico de cada pessoa.

Por fim, o SR proposto pode auxiliar profissionais de saúde a orientar atividades físicas, personalizando sugestões com base em dados médicos, dados pessoais e metas individuais. Isso pode propiciar a melhoria da prescrição de exercícios, monitoramento do progresso do paciente e promoção de um estilo de vida ativo de forma segura e eficaz.

## 1.5 Limitações da pesquisa

Embora esta pesquisa tenha como desfecho o uso de IA para auxiliar na compreensão e predição dos fatores que motivam a prática de atividades físicas, como qualquer outra pesquisa, esta não esgota o assunto e possui algumas limitações, destacadas a seguir, que abrem perspectivas para trabalhos futuros.

- A proposta do trabalho não foi avaliar as características das rotinas das pessoas, mas aplicar IA na compreensão e predição dos fatores que motivam a prática de atividades físicas com base nos seus perfis.
- Os atributos de faixa etária, cor, estado civil e escolaridade, foram categorizados com base em intervalos de valores visando reduzir o número de regras geradas pelas árvores de decisão. Assim, as inferências foram produzidas a partir de algumas generalizações, o que pode diminuir a precisão das previsões do MIF.
- Como as inferências produzidas pelo MIF são baseadas em regras geradas a partir de uma base de dados limitada (140 pessoas), os resultados produzidos pelo SR podem ser imprecisos e/ou potencialmente enviesados.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos teóricos necessários para o entendimento do *framework* proposto neste trabalho. Primeiro são apresentadas algumas definições sobre os fatores que motivam a prática de atividades físicas (seção 2.1). Em seguida são apresentados conceitos sobre IA (seção 2.2), *Knowledge Discovery in Databases – KDD* e Mineração de Dados (seção 2.3), Técnicas de IA aplicadas em MD (seção 2.4), e, por fim, sobre Sistemas de Recomendação (seção 2.5).

### 2.1 Motivações para a prática de atividades físicas

O "Inventário de Motivação à Prática Regular de Atividade Física" (BARBOSA e BALBINOTTI, 2006) é um instrumento baseado em pressupostos da Teoria da Auto-Determinação (*Self-Determination Theory*), descrita por Ryan e Deci (1985, 2000a). Trata-se de uma teoria amplamente utilizada e aceita (BRIÈRE *et al.*, 1995; DECI, RYAN e KOESTNER, 1999; DUDA *et al.*, 1995; FREDERICK e RYAN, 1995; REEVE e DECI, 1996; RYAN e DECI, 2000b; RYAN *et al.*, 1997; FORTIER *et al.*, 1997; ANSHEL *et al.*, 2000). De acordo com esta teoria um sujeito pode ser motivado em diferentes níveis intrínsecos ou extrínsecos, ou ainda, sem um fator motivacional definido durante a prática de uma atividade.

Um sujeito intrinsecamente motivado é aquele que ingressa na atividade pela sua própria vontade, pelo prazer e pela satisfação do processo de conhecer uma nova atividade. Comportamentos intrinsecamente motivados são associados com bem-estar psicológico, interesse, prazer, alegria e persistência (RYAN e DECI, 2000b).

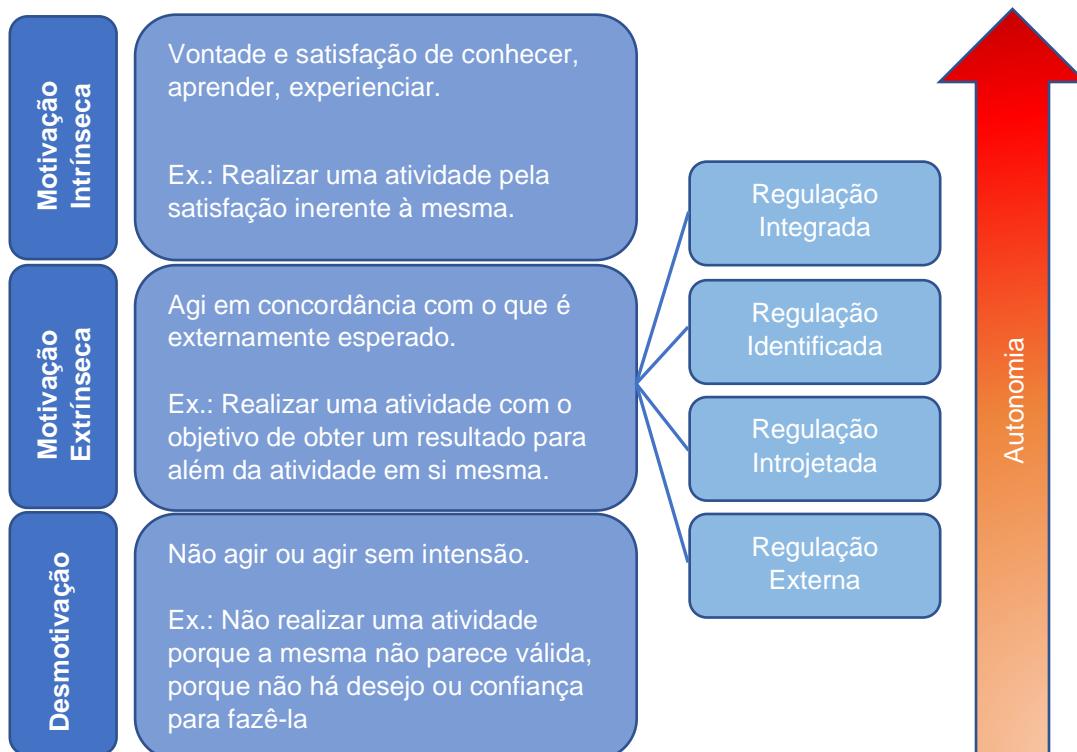
A motivação intrínseca tem sido subdividida em 3 tipos: "para saber", "para realizar" e "para experiência". A motivação intrínseca "para saber" pode ser explicada como executar uma atividade para satisfazer uma experiência enquanto se aprende a atividade; a motivação "para realizar" ocorre quando um indivíduo realiza uma atividade pelo prazer de executá-la, e a motivação "para experiência" ocorre quando um indivíduo realiza uma atividade para experienciar as situações estimulantes inerentes à tarefa (BRIÈRE *et al.*, 1995).

Já uma motivação extrínseca a uma atividade, segundo Ryan e Deci (2000a), ocorre quando esta atividade é efetuada com outro objetivo que não o inerente à própria pessoa. Entretanto, estes motivos podem variar grandemente em relação ao seu grau de autonomia, criando três categorias de motivação extrínseca: a) regulação externa: é o comportamento regulado por meios como premiação material ou medo de consequências negativas, como críticas do técnico ou pais. Este tipo de motivação pode ser observado no âmbito esportivo quando o treinador impõe penas aos atletas que não realizarem as tarefas propostas; b) regulação interiorizada, na qual uma fonte de motivação externa é internalizada, como comportamentos reforçados por pressões internas como a culpa, ou como a necessidade de ser bem-quisto. Este comportamento pode ser visto quando alguém realiza uma atividade por “desencargo de consciência”; c) regulação identificada, quando um sujeito realiza uma tarefa, a qual não lhe é disponível a escolha, mesmo que este comportamento não o seja interessante. Este comportamento é visualizado, por exemplo, no diálogo de um atleta que diz que aulas de alongamento são necessárias porque seu treinador disse, e mesmo não gostando de executar ele o realiza.

Ryan e Deci (2000), também citam um tipo de construção motivacional percebida em indivíduos que não são aptos a identificar nenhum bom motivo para realizar alguma atividade, chamada de amotivação. Segundo estes indivíduos, a atividade ou não lhes trará nenhum benefício, ou eles não conseguirão realizá-la de modo satisfatório, no seu ponto de vista (BRIÈRE *et al.*, 1995).

Entretanto, convém se fazer algumas ressalvas. Petherick e Weigand (2002) sugerem que a simples divisão entre motivação intrínseca e motivação extrínseca gera uma dicotomia simplista entre as duas. Também é necessário que se diga que ser motivado extrinsecamente não corresponde a um comportamento negativo. De acordo com Ryan e colaboradores (1997), os motivos extrínsecos possuem um grande grau de autonomia. Porém, motivos intrínsecos possuem caráter unicamente auto determinável. Esta relação de autonomia pode ser melhor verificada em uma taxonomia organizada por Ryan e Deci (2000), ilustrada na Figura 1.

**Figura 1 – Uma taxonomia da motivação humana**



Fonte: Própria autoria

Desta maneira, a identificação tanto de uma motivação intrínseca quanto extrínseca não se torna tão relevante quanto as correlações entre as variáveis motivacionais para a prática de atividades físicas deste estudo, já que todas elas podem ser classificadas por regulação integrada, como é o caso do condicionamento físico, ou uma regulação identificada, como é o caso da prática da atividades físicas motivada por busca de saúde ou alívio de ansiedade, ou uma regulação introjetada, como é o caso da prática da atividades físicas motivada por estética, ou ainda por uma regulação externa, como é o caso da atividades físicas motivada pela integração social.

## 2.2 Inteligência Artificial (IA)

A criação de sistemas computacionais com capacidade de simular ou imitar comportamentos inteligentes e relevantes aos seres humanos pode ser classificada como uma tecnologia de inteligência artificial (IA), mesmo que suas ações se deem de forma diferente dos seres humanos (BHAVE *et al.*, 2020).

Aprendizado de máquina (MITCHELL, 2010) e aprendizado profundo, estão entre as áreas de pesquisa da IA e podem ser empregadas em todas as áreas do conhecimento (BHAVE *et al.*, 2020). Na educação física por exemplo, a IA está presente em avaliações e prescrições de exercícios físicos, através de aplicativos que realizam avaliação de composição corporal, controle de gasto calórico, dentre outros, os quais vêm se multiplicando nos últimos anos produzindo novos modos de vida e alternativas para a prática de atividades físicas (OLIVEIRA; FRAGA, 2020). Esses instrumentos tecnológicos auxiliam na orientação da prática das atividades físicas com planejamento e particularidades do usuário para os seus treinos.

Entre as técnicas de IA mais utilizadas no desenvolvimento de sistemas inteligentes estão as redes neurais artificiais (RNA), as redes neurais convolucionais (RNC), *Support Vector Machine* (SVM), Lógica Fuzzy e Árvores de Decisão (AD).

### 2.2.1 Diferenças entre IA e Estatística Descritiva na análise de dados

A estatística descritiva é uma técnica de análise de dados que usa medidas de estatística descritiva para resumir e descrever conjuntos de dados. Ela se concentra em dados históricos e tende a ser mais voltada para o passado. Alguns exemplos de medidas estatísticas incluem a média, o desvio padrão e a mediana. A estatística descritiva é útil para fornecer informações sobre como um conjunto de dados se comportou no passado e pode ser usada para tomar decisões com base nessas informações. Ela diferencia-se da estatística inferencial (ou estatística indutiva), a qual se ocupa com a generalização dos dados, ou seja, é usada para fazer inferências sobre um conjunto de dados (Ferreira, 2005; Correia *et al.*, 2021).

Tanto a estatística descritiva quanto IA são abordagens bastante empregadas para obter informações a partir da análise de conjuntos de dados. No entanto, embora ambas possam ser úteis, a IA oferece algumas vantagens em relação à estatística descritiva.

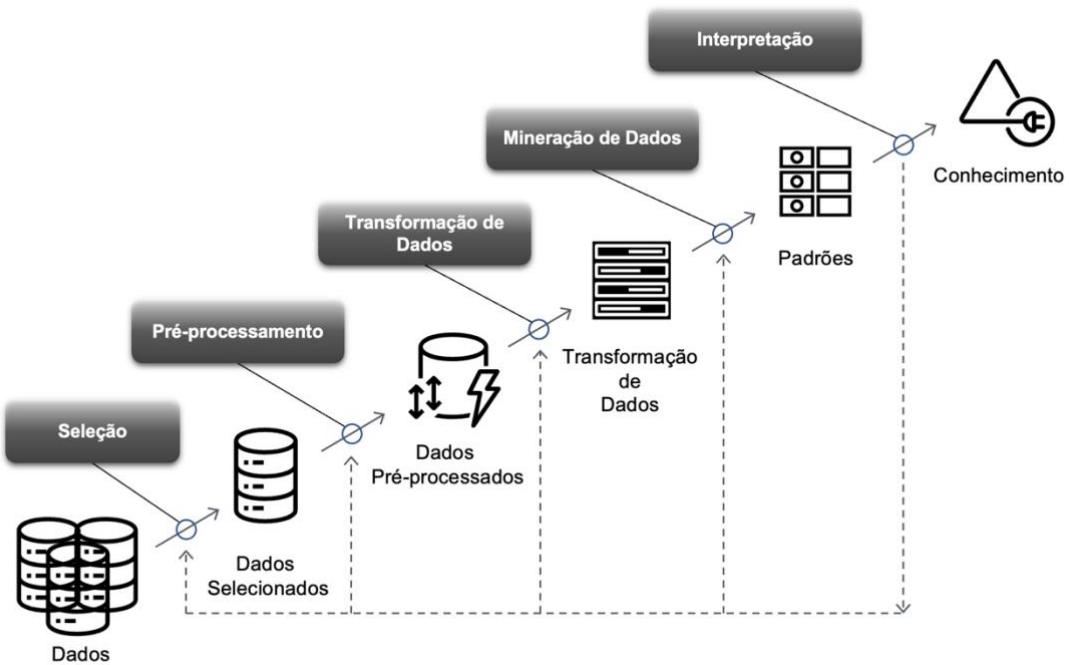
IA é uma abordagem de análise de dados mais sofisticada que usa algoritmos de aprendizado de máquina para identificar padrões e tendências em grandes conjuntos de dados. Ela é mais voltada para o futuro, pois pode ser usada para prever resultados e tomar decisões com base nessas previsões (Correia et al., 2021). Algumas das vantagens da IA sobre a estatística descritiva incluem:

- Previsões mais precisas: capacidade de identificar padrões e tendências que podem ser difíceis de detectar com a estatística descritiva.
- Escalabilidade: poder para lidar com conjuntos de dados muito maiores do que a estatística descritiva, o que é especialmente útil em setores como a indústria, saúde, educação e finanças.
- Velocidade: capacidade de processar grandes conjuntos de dados muito mais rapidamente do que a estatística descritiva, que em setores que requerem a tomada de decisões em tempo real empregam.

### **2.3 *Knowledge Discovery in Databases – KDD***

O *KDD* pode ser descrito como um processo que envolve a extração de informações úteis, padrões, conhecimentos e insights a partir de bases de dados (FAYYAD, 1996). Conforme ilustrado na Figura 2, o *KDD* contempla todas as etapas para a descoberta do conhecimento a partir da base de dados analisada, as quais estão descritas na subseção 2.3.1, sendo a Mineração de Dados uma das etapas mais importantes do processo.

**Figura 2 – Etapas do Processo KDD**



Fonte: Própria autoria

### 2.3.1 Etapas do KDD

#### a) Seleção

A etapa de seleção consiste em selecionar o(s) conjunto(s) de dados que será(ão) analisado(s). As fontes de dados podem ser bancos de dados, planilhas ou *data warehouses* contendo dados em diversos formatos (estruturados, semiestruturados e não-estruturados). Trata-se de uma etapa crucial para identificar e coletar dados relevantes para a análise posterior. Durante essa etapa, também é considerada a representatividade e qualidade dos dados selecionados para garantir que os resultados da análise sejam confiáveis. Além disso, é fundamental avaliar a capacidade de processamento dos dados selecionados para que possam ser manipulados adequadamente pelos algoritmos de mineração de dados utilizados na etapa seguinte. Por fim, documentar todas as decisões tomadas durante a seleção de dados para garantir a transparência e reproduzibilidade do processo (MURTHY, 1998).

### **b) Pré-processamento**

A etapa de pré-processamento de dados do *KDD* é responsável por preparar os dados selecionados para a análise posterior. Nessa etapa, são realizadas diversas atividades como limpeza de dados para remover ruídos e dados inconsistentes ou anomalias (*outliers*). O objetivo principal desta etapa é melhorar a qualidade dos dados para que possam ser analisados pelos algoritmos de MD com maior precisão e eficiência (WITTEN; FRANK, 2005).

### **c) Transformação dos dados**

A etapa de transformação de dados do *KDD* é responsável por aplicar técnicas de manipulação e modificação dos dados pré-processados, como discretização de dados contínuos, agregação de dados, normalização de dados (para garantir que todos os dados tenham a mesma escala e distribuição), combinação de dados e geração de novas características. O objetivo principal desta etapa é criar uma representação mais útil e significativa dos dados para a análise. Ela é crucial para obter resultados precisos e confiáveis na etapa de MD, pois a qualidade dos dados transformados influencia diretamente a eficácia dos algoritmos de análise (GOLDSCHMIDT e PASSOS, 2005).

### **d) Mineração de dados**

O conceito de Mineração de Dados (MD) surgiu na década de 1990 advindo da evolução da tecnologia de banco de dados que possibilitou às organizações realizarem coleta e armazenamento contínuo de grandes volumes de dados sobre clientes, fornecedores, produtos e serviços de dados que estavam sendo subutilizados (FAYYAD, 1996; ARAÚJO, *et al.*, 2022).

A MD pode ser entendida como o processo de exploração e análise de dados por meio do uso de algoritmos de IA e de estatística descritiva, com o objetivo de descoberta de padrões e regras que forneçam compreensão mais apurada das informações que estes conjuntos de dados “escondem” (FAYYAD, 1996).

As tarefas mais essenciais da MD são: classificação, descoberta de regras de associação e agrupamento. A classificação tem a finalidade de estudar o

conjunto de dados de entrada e compor a descrição das propriedades desses dados para cada classe (FAYYAD, 1996).

As regras de associação revelam padrões descobertos na forma de regras do tipo SE...ENTÃO (FAYYAD, 1996). De forma geral, os algoritmos empregados nesse processo geram um conjunto consideravelmente amplo de regras, deixando ao usuário a tarefa de selecionar aquelas mais pertinentes para uma aplicação específica. Finalmente, o processo de clusterização tem como objetivo agrupar os dados analisados em dois ou mais grupos, utilizando uma métrica para evidenciar a similaridade entre os padrões com base em suas características individuais.

As técnicas e algoritmos de IA mais comumente empregados em MD são: Redes Neurais Artificiais (RNA), Algoritmos Genéticos (AG), *K-means*, Máquinas de Vetores de Suporte (*Support Vector Machines* – SVM), Redes Bayesianas, Árvores de Decisão (AD), Apriori e Lógica *Fuzzy*. Tais técnicas são apresentadas na seção 2.4 a seguir, sendo as empregadas na presente pesquisa (AD, algoritmo Apriori e Lógica *Fuzzy*) descritas de forma mais detalhada.

#### **e) Interpretação/Avaliação**

O objetivo principal da interpretação de dados é obter insights e conhecimentos a partir dos dados analisados visando solucionar problemas complexos. Além disso, é fundamental elaborar avaliações sobre os padrões encontrados na etapa anterior, por meio de análises de testes e validações usando medidas estatísticas como taxa de acerto, acurácia, precisão, índice Kappa, entre outras (FAYYAD, 1996).

### **2.4 Algumas técnicas de IA empregadas em MD**

#### **2.4.1 Redes Neurais Artificiais (RNA)**

Uma RNA é um modelo computacional inspirado no funcionamento do cérebro humano que tem sido amplamente utilizado no campo da IA. Ela é composta por várias unidades de processamento interconectadas, conhecidas como neurônios artificiais, que colaboram entre si para realizar tarefas complexas de aprendizado de máquina supervisionado e processamento de informações.

Essas redes são capazes de aprender a partir de exemplos e experiências, ajustando automaticamente os pesos das conexões entre os neurônios para melhorar o desempenho em determinada tarefa (MCCULLOCH e PITTS 1943; ROSENBLATT, 1958; RUMELHART *et al.* 1986). As RNAs têm sido aplicadas em uma ampla gama de aplicações, como reconhecimento de padrões, processamento de linguagem natural, visão computacional e previsões de séries temporais, demonstrando sua eficácia em diversos domínios (CAVALCANTE *et al.* 2023).

#### 2.4.2 Algoritmos Genéticos (AGs)

Os AGs são uma abordagem computacional inspirada no processo de evolução biológica e tem sido amplamente utilizados para resolver problemas complexos em diversas áreas. Esses algoritmos são fundamentados na ideia de que uma população de soluções candidatas evolui ao longo do tempo, passando por processos de seleção, recombinação e mutação, de forma análoga ao que ocorre na natureza (GOLDBERG, 1989; DEB, 2009).

Os avanços demonstrados têm impulsionado o uso dos AGs em problemas complexos de otimização, como em roteamento de veículos, projetos de redes de comunicação, otimização de sistemas energéticos, entre outros.

#### 2.4.3 *K-means*

O algoritmo *K-means* é uma técnica de aprendizado de máquina não supervisionada amplamente utilizada para a clusterização de dados em grupos distintos. Ela busca partitionar um conjunto de pontos em *K clusters*, onde *K* é um número pré-definido. O objetivo do *K-means* é minimizar as distâncias entre os pontos e seus respectivos centroides, sendo que cada centróide representa o centro de um *cluster*.

A partir de 2018, várias pesquisas foram realizadas para melhorar o desempenho e a eficiência do algoritmo *K-means*, como por exemplo o estudo de Ganti *et al.* (2018) que propôs uma abordagem baseada em técnicas de amostragem para reduzir a complexidade computacional do algoritmo em grandes conjuntos de dados.

#### 2.4.4 Support Vector Machines (SVM)

O SVM é um algoritmo de aprendizado de máquina supervisionado que tem como objetivo principal a classificação e regressão de dados. Essa abordagem utiliza conceitos de otimização matemática e teoria do aprendizado estatístico para criar modelos capazes de separar ou mapear dados em espaços multidimensionais.

As SVMs procuram encontrar um hiperplano de separação que maximize a distância entre as diferentes classes de dados, permitindo uma melhor generalização para novos exemplos. Além disso, as SVMs possuem a capacidade de lidar com dados não lineares por meio do uso de funções de *kernel*, que transformam o espaço original dos dados em um espaço de maior dimensionalidade (KHOBRAGADE *et al.*, 2016).

#### 2.4.5 Redes Bayesianas

As Redes Bayesianas são modelos estatísticos que permitem representar e analisar relações de dependência entre variáveis. Seu objetivo principal é fornecer um mecanismo para realizar inferências probabilísticas em ambientes incertos, permitindo a tomada de decisões informadas.

As Redes Bayesianas têm sido amplamente aplicadas em diversas áreas, incluindo medicina, engenharia, finanças e ciência da computação. Recentemente, várias pesquisas destacam o potencial das Redes Bayesianas na análise de dados complexos, como genômica, dados de saúde e sistemas de recomendação. Por exemplo, um estudo publicado por Chen *et al.* (2006) aplicou Redes Bayesianas na análise de dados de expressão genética para prever risco de doença arterial coronariana.

#### 2.4.6 Árvores de Decisão

Dentre os algoritmos de aprendizado supervisionado, a AD pode ser considerado um dos mais simples. Uma AD é descrita por meio de um conjunto de regras do tipo SE...ENTÃO (QUINLAN, 1986; MITCHELL, 2010). Em virtude da simplicidade, as ADs têm sido largamente empregadas em tarefas de classificação, sendo um caminho eficaz na construção de classificadores que

predizem classes com base nos valores dos atributos que representam os padrões. Desta forma, elas podem ser empregadas em diversas aplicações como diagnósticos médicos e análise de risco em créditos, entre outras.

A chave para o sucesso de um algoritmo AD é como gerar a árvore, ou seja, como escolher os atributos mais significativos para gerar as regras, e quais regras podem ser descartadas da árvore (MITCHELL, 1997). O atributo mais significativo é posicionado na raiz da árvore. De acordo com Araújo *et al.* (2018), tendo em vista de que as ADs são construídas com base na importância de cada atributo, nem sempre elas se valem de todos os atributos que representam um padrão para geração do conjunto de regras, trazendo o benefício de reduzir o tempo computacional nas tarefas de classificação.

Entre os algoritmos mais conhecidos de AD estão ID3 (QUINLAN, 1986; MITCHELL, 2010), C4.5 (QUINLAN, 1993) e CART (BRAMER, 2007), sendo o primeiro um dos mais básicos. O algoritmo C4.5, empregado neste trabalho, constrói AD a partir de um conjunto de dados da mesma forma que o algoritmo ID3, usando os conceitos de entropia e ganho de informação (equações 1 e 2) para definir a importância dos atributos. Em cada nó da árvore, o algoritmo C4.5 escolhe o atributo que melhor particiona o conjunto de amostras em subconjuntos, tendendo a uma categoria ou outra. O atributo com maior ganho de informação normalizado é escolhido para tomar a decisão (QUINLAN, 1993).

Para realizar o cálculo do ganho de informação, primeiro obtém-se a entropia, que pode ser definida como sendo o grau de pureza do conjunto de dados. Dado um conjunto de entrada  $S$  que pode ter  $n$  classes distintas, a entropia de  $S$  é dada por:

$$Entropia(S) = \sum_{i=1}^n -P_i \log_2 P_i \quad (1)$$

onde:  $P_i$  é a proporção de dados em  $S$  que pertencem à classe  $i$ .

$$Gain(S, A) = Entropia(S) - \sum_{v \text{ valores}(A)} \frac{|S_v|}{|S|} Entropia(S_v) \quad (2)$$

$v$  é um elemento dos valores que o atributo  $A$  pode assumir e  $S_v$  é o subconjunto de  $S$  formado pelos dados em que  $A = v$ .

A técnica de AD pode ser facilmente compreendida por meio do exemplo apresentado a seguir, cujo objetivo é determinar se um jogador irá praticar tênis. Para isso, é necessário levar em consideração certos elementos do ambiente, como a aparência do céu, a temperatura, a umidade relativa do ar e o vento. Cada um desses atributos possui diversas opções (representadas por valores). Por exemplo, para a temperatura, os valores possíveis são “Ameno”, “Fresco” ou “Quente”. A decisão de praticar tênis (Sim) ou não praticar tênis (Não) é o resultado da classificação. Para criar a estrutura de decisão são considerados exemplos extraídos de dias anteriores, mostrados na Tabela 1.

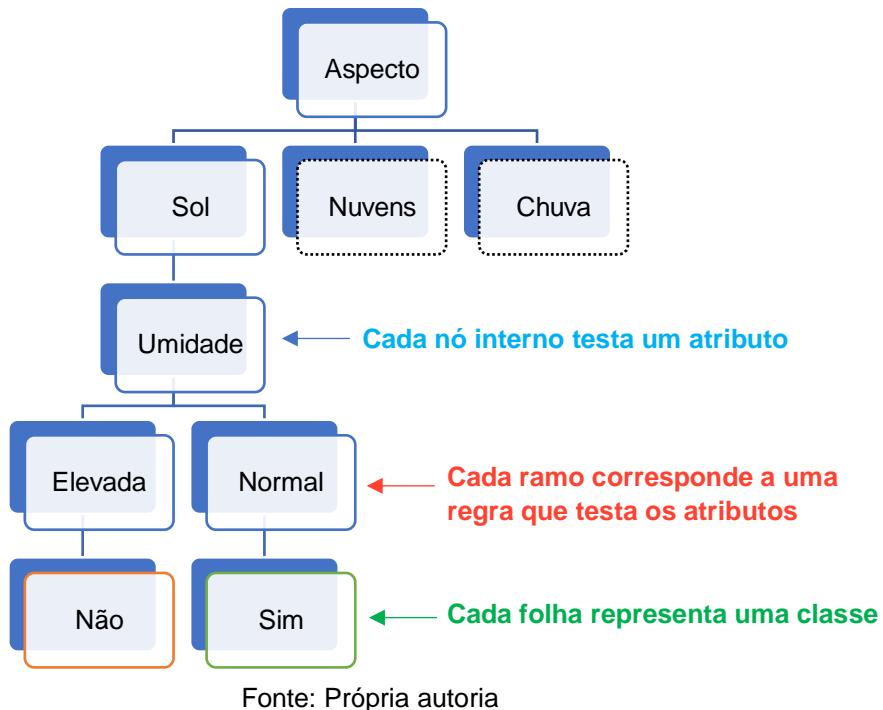
**Tabela 1 – Base de dados sobre as condições para jogar ou não tênis**

Dia	Aspecto	Temperatura	Umidade	Vento	Jogar Tênis
D1	Sol	Quente	Elevada	Fraco	Não
D2	Sol	Quente	Elevada	Forte	Não
D3	Nuvens	Quente	Elevada	Fraco	Sim
D4	Chuva	Ameno	Elevada	Fraco	Sim
D5	Chuva	Fresco	Normal	Fraco	Sim
D6	Chuva	Fresco	Normal	Forte	Não
D7	Nuvens	Fresco	Normal	Fraco	Sim
D8	Sol	Ameno	Elevada	Fraco	Não
D9	Sol	Fresco	Normal	Fraco	Sim
D10	Chuva	Ameno	Normal	Forte	Sim
D11	Chuva	Ameno	Normal	Forte	Sim
D12	Nuvens	Ameno	Elevada	Forte	Sim
D13	Nuvens	Quente	Normal	Fraco	Sim
D14	Chuva	Ameno	Elevada	Forte	Não

Fonte: Própria autoria

Aplicando os conceitos anteriormente descritos na base de dados da Tabela 1, é possível construir a AD mostrada na Figura 3, a qual ilustra a relação entre os elementos da árvore (nós e folhas) e os atributos (valores e classificações).

**Figura 3 – Árvore de decisão sobre “jogar tênis ou não”**



As ADs são largamente empregadas em MD em virtude do funcionamento e resultados gerados por elas serem de fácil interpretação por pessoas de qualquer área do conhecimento (JUNG *et al.*, 2017).

#### 2.4.7 Algoritmo Apriori

O algoritmo Apriori, sugerido por Agrawal e Srikant (1994), é um dos mais antigos algoritmos supervisionados amplamente conhecidos para mineração de regras de associação em grandes conjuntos de dados. O nome do algoritmo é derivado do fato de que ele utiliza o conhecimento prévio das propriedades frequentes do conjunto de itens.

Ele identifica todos os conjuntos de itens frequentes, conhecidos como conjuntos frequentes de itens. A primeira etapa do algoritmo Apriori envolve a contagem das ocorrências dos itens para determinar os conjuntos frequentes de itens de tamanho único. As etapas subsequentes compõem duas fases. Primeiramente, os conjuntos frequentes de itens encontrados na primeira etapa são empregados para gerar conjuntos candidatos de itens. Posteriormente, uma nova busca é realizada no conjunto de dados, contabilizando o suporte de cada candidato (AGRAWAL e SRIKANT, 1994).

Para exemplificar o funcionamento do algoritmo Apriori, considere a base de dados apresentada na Tabela 2, na qual cada linha é uma transação e cada célula é um item individual da transação:

Tabela 2 – Base de dados para representar o exemplo do funcionamento do algoritmo Apriori

alfa	beta	épsilon
alfa	beta	teta
alfa	beta	épsilon
alfa	beta	teta

Fonte: Própria autoria

As regras de associação que podem ser determinadas dessa base de dados pelo algoritmo Apriori são as seguintes:

- 100% dos conjuntos com alfa também contêm beta
- 50% dos conjuntos com alfa, beta também contêm épsilon
- 50% dos conjuntos com alfa, beta também contêm teta

#### 2.4.8 Lógica *Fuzzy*

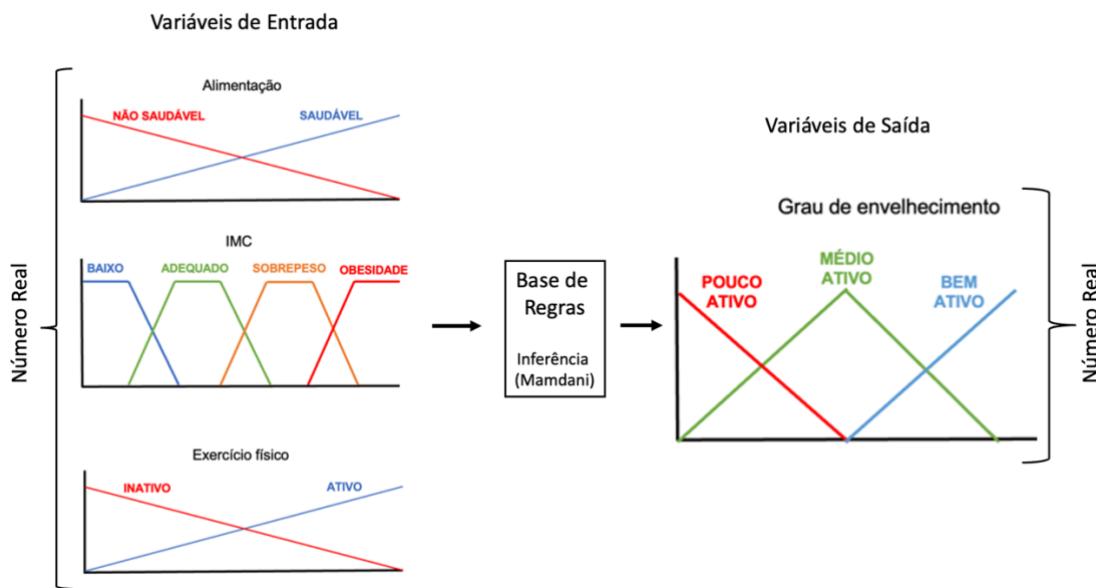
A palavra *Fuzzy* em inglês pode ter diversos significados, que mudam dependendo do contexto. "Nebuloso" e "difuso" são algumas das traduções mais comuns para *Fuzzy* (MARRO, 2010), mas o sentido fundamental desse adjetivo está sempre relacionado a algo vago, indistinto e incerto.

A lógica *Fuzzy*, também conhecida como lógica nebulosa, é considerada uma forma de lógica de lidar com modelos de raciocínios imprecisos ou aproximados (ZADEH, 1973). Essa lógica permite trabalhar com problemas de decisão que não são facilmente representados em modelos matemáticos convencionais. Um controle *Fuzzy* procura "imitar" um agente humano por meio de uma descrição descritiva e experimental de um processo específico (ZADEH, 1973).

No campo da IA, a lógica *Fuzzy* é considerada um mecanismo fundamental para representar de forma apropriada o conhecimento e o raciocínio devido à sua habilidade de trabalhar com incertezas, raciocínio aproximado,

termos vagos e ambíguos, que são maneiras pelas quais as pessoas expressam seus pensamentos. Por isso, os mecanismos de inferência *Fuzzy* (Figura 4) possuem a capacidade de mostrar etapas levando em conta perspectivas relacionadas à incerteza e aos processos práticos, tornando-os mais semelhantes a um comportamento "humano" (MARRO, 2010).

**Figura 4 – Exemplo de um Mecanismo de Inferência *Fuzzy* (MIF) aplicado à saúde**



Fonte: Própria autoria

A seguir é apresentada uma lista não-exaustiva de domínios de aplicação da Lógica *Fuzzy*.

- Sistemas especialistas;
- Sistemas multiagentes;
- Reconhecimento de padrões;
- Robótica;
- Sistemas de controle inteligentes;
- Sistemas de apoio à tomada de decisão;
- Algoritmos genéticos;
- *Data mining*.

McNeil e Thro (1994) relacionam algumas características de sistemas nos quais a aplicação da Lógica *Fuzzy* é benéfica:

- Sistemas complexos que são difíceis ou impossíveis de modelar;
- Sistemas controlados por especialistas (humanos);
- Sistemas com entradas e saídas complexas e contínuas;
- Sistemas que se utilizam da observação humana como entradas ou como base para regras;
- Sistemas que são naturalmente “vagos”, como os que envolvem ciências sociais e comportamentais, cuja descrição é extremamente complexa.

#### 2.4.9 Medidas de desempenho de algoritmos de classificação empregados em MD

Em tarefas de MD diversas métricas são comumente empregadas para avaliar e medir a qualidade e o desempenho dos modelos e algoritmos utilizados em tarefas de classificação.

Algumas das métricas mais utilizadas incluem a acurácia, que mede a proporção de instâncias corretamente classificadas; a precisão, que indica a proporção de instâncias classificadas como positivas que são realmente positivas; a revocação (*recall* ou taxa de acerto), que representa a proporção de instâncias positivas classificadas corretamente; e o coeficiente Kappa, comumente referido como pontuação Kappa de Cohen (PIO, 2023), que é uma estatística amplamente usada para avaliar a eficácia dos modelos de classificação em aprendizado de máquina.

Todas essas métricas podem ser obtidas a partir de uma matriz de confusão, como a apresentada na Tabela 3, a qual reflete os acertos (verdadeiros positivos (VP) e verdadeiros negativos (VN)) na sua diagonal principal. Os demais elementos da matriz contemplam os erros do classificador (falsos positivos (FP) e falsos negativos (FN)).

**Tabela 3 – Exemplo de matriz de confusão**

<b>Classe real</b>		<b>Classe predita</b>
<b>+</b>	<b>-</b>	
<b>VP</b>	<b>FP</b>	<b>+</b>
<b>FN</b>	<b>VN</b>	<b>-</b>

Fonte: Própria autoria

Neste trabalho a avaliação das classificações foram feitas com base na acurácia e no índice Kappa, cujos cômputos são apresentados nas equações 3 e 4.

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (3)$$

$$K = \frac{Po - Pe}{1 - Pe} = 1 - \frac{1 - Po}{1 - Pe} \quad (4)$$

Na definição de  $K$ ,  $Po$  é proporção de concordâncias observadas e  $Pe$  é proporção de concordâncias esperadas, e podem ser obtidos como segue:

$$Po = (VP + VN) / (VP + FP + VN + FN)$$

$$Pe = (VP + FP) \times (VP + FN) + (FN + VN) \times (FP + VN) / (VP + FP + FN + VN)^2$$

O coeficiente Kappa é uma maneira muito utilizada para avaliar a concordância dos resultados ou ainda expressar a confiabilidade de um classificador, conforme indicado na Tabela 4.

**Tabela 4 – Níveis de concordância do coeficiente Kappa**

K	Nível de concordância
0 – 0,20	Nenhuma
0,21 – 0,39	Mínima
0,40 – 0,59	Fraca/Sofrível
0,60 – 0,79	Substancial
0,80 – 0,90	Forte/Boa
0,91 – 0,99	Ótima
1,00	Perfeita

Fonte: adaptado de MCHUGH (2012)

## 2.5 Sistemas de recomendação

Na era atual enfrentamos um volume praticamente ilimitado de dados prontamente disponíveis, especialmente na web. Isso resulta em um cenário em que diversos indivíduos se deparam com uma ampla gama de possibilidades, mesmo que possuam escassa ou praticamente nenhuma vivência pessoal para fazer escolhas diante das diversas opções apresentadas a eles. No cotidiano, é

frequente que indivíduos façam seleções com base nas sugestões de terceiros. Essas sugestões podem surgir de interações sociais, análises de estabelecimentos ou produtos, propagandas em meios de comunicação, e assim por diante (ISINKAYE, FOLAJIMI & OJOKOH, 2015).

Sistemas de recomendação (SR) podem ser vistos como técnicas e algoritmos, normalmente oriundos do campo de IA, utilizados para sugerir itens, produtos ou informações relevantes aos usuários com base em seus interesses, histórico de interações ou características pessoais. Tais sistemas têm ampla aplicação em diversos setores, como comércio eletrônico, streaming de mídia, redes sociais e muito mais. Em geral, o objetivo de um SR é melhorar a experiência do usuário, aumentar o engajamento e facilitar a descoberta de novos conteúdos, produtos ou serviços relevantes (RESNICK; VARIAN, 1997).

Balabanovic' e Shoham (1997) categorizam os Sistemas de Recomendação com base no tipo de dados que analisam. Essas categorias incluem Sistemas de Recomendação Baseados em Conteúdo, Sistemas de Recomendação Colaborativos e Sistemas de Recomendação Híbridos.

Os Sistemas de Recomendação Baseados em Conteúdo sugerem ao usuário um item semelhante aos que ele aprovou anteriormente, conforme demonstrado por Pazzani e Billsus (1997). Por outro lado, os Sistemas de Recomendação Colaborativos propõem itens que já foram utilizados por outros usuários com perfis semelhantes, conforme apresentado por Goldberg *et al.* (1992). Já os Sistemas de Recomendação Híbridos utilizam uma combinação das duas categorias anteriores, aproveitando as vantagens e mitigando os pontos fracos de cada abordagem (FALK, 2019).

### 3 MATERIAIS E MÉTODOS

Neste capítulo são apresentados os materiais e métodos empregados na descoberta de padrões e predições dos fatores que motivam a prática de atividades físicas. Primeiro é apresentada a base de dados na seção 3.1 e, em seguida, são apresentados o *framework* de IA (seção 3.2) e o Sistema de Recomendação (seção 3.3) propostos.

#### 3.1 Base de dados

Para a realização dos experimentos foi empregada uma base de dados contendo 140 registros de indivíduos de todo Brasil, que foi cedida pela empresa S2Pro Data Health<sup>1</sup>, a qual construiu a base utilizando o formulário de pesquisa online apresentado no ANEXO A, implementado por meio do aplicativo online coleta dos dados FORMS.APP<sup>2</sup>.

A base contempla 14 atributos (colunas) sendo 7 socioeconômicos (faixa etária, sexo, cor/etnia, estado civil, escolaridade, faixa salarial e profissão) e outros 7 relativos aos fatores de motivação para a prática de atividades físicas (condicionamento físico, estética, saúde, integração social, ansiedade, disponibilidade de tempo e custo), para os quais havia 5 graus possíveis de resposta: “nulo”, “baixo”, “médio”, “alto” e “muito alto”.

Os respondentes tinham idade mínima de 18 e máxima de 70 anos, com maior frequência de mulheres (56%), sendo 78 indivíduos casados, 47 solteiros, 15 divorciados; escolaridade variando de primeiro grau à doutorado; faixa salarial de 1 até mais de 10 salários mínimos; e a seguinte distribuição de cor: branca (66,43%), parda (25.71%), outras (7.86%). A distribuição dos respondentes de acordo com as respostas para os atributos socioeconômicos é apresentada na Tabela 5, na qual os números de respondentes são indicados entre parênteses.

---

<sup>1</sup> <https://s2pro.com.br>

<sup>2</sup> <https://my.forms.app/s2proapp/atividades-fisicas-e-motivacao>

**Tabela 5 – Distribuição dos respondentes de acordo com os atributos socioeconômicos**

Atributo	Faixa de respostas				
Faixa Etária	18 a 25 anos (23)	26 a 35 anos (38)	36 a 45 anos (54)	46 a 55 anos (20)	56 a 99 anos (05)
Sexo	Masculino (61)	Feminino (79)			
Cor/Etnia	Branco (93)	Pardo (36)	Outros (11)		
Estado Civil	Casado 78	Separado 15	Solteiro 47		
Escolaridade	SG (27)	ES (46)	PG (67)		
Faixa Salarial	Até 1 SM (03)	De 1 a 5 SM (49)	De 6 a 10 SM (30)	Mais de 10 SM (22)	Prefiro não responder (36)

**SM = salário mínimo; SG = segundo grau; ES = ensino superior; PG = pós-graduação**

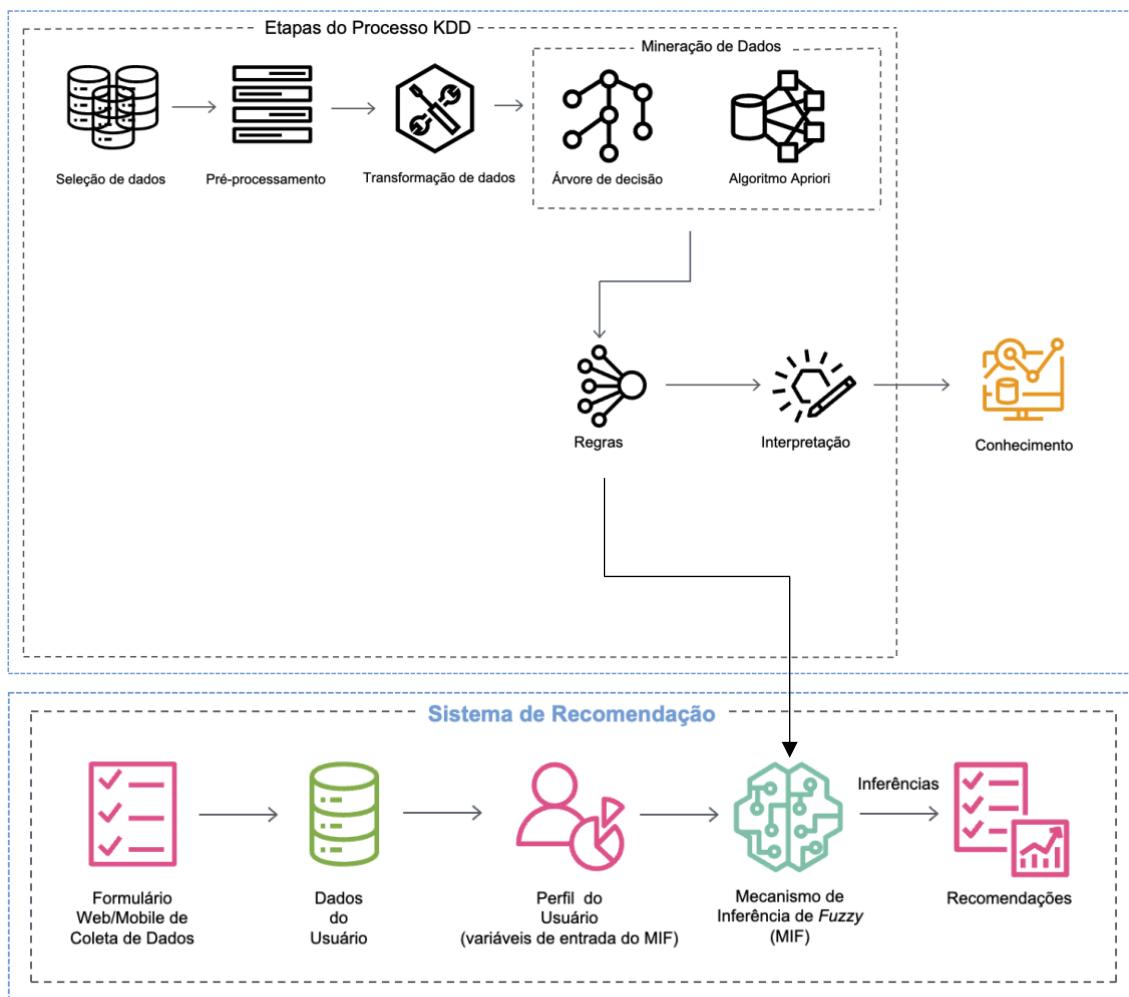
Fonte: Autoria Própria com base na Pesquisa Motivacional para Prática de Atividades Físicas

Os dados foram recebidos em uma planilha no formato CSV (dados separados por vírgula), para posterior análise usando o software WEKA – *Waikato Environment for Knowledge Analysis*, uma ferramenta que fornece diversos algoritmos de IA e aprendizagem de máquina para tarefas de mineração de dados (UNIVERSITY OF WAIKATO, 2010).

### 3.2 **Framework proposto**

O *framework* proposto neste trabalho, cujo diagrama de funcionamento é ilustrado na Figura 5, consiste em um encadeamento de técnicas de IA que permitem desde a descoberta de padrões para auxiliar compreensão dos fatores que motivam a prática de atividades físicas até a predição dos graus associados a cada fator motivacional e recomendação de atividades físicas.

**Figura 5 – Diagrama esquemático do encadeamento das técnicas de IA no framework proposto**



Fonte: Própria autoria

Na primeira etapa do *framework* são realizadas tarefas de classificação e geração de regras de associação aplicando-se as etapas do *KDD*.

A classificação teve como objetivo investigar a existência de padrões que podem determinar a importância dos fatores motivacionais para prática de atividades físicas em diferentes grupos de pessoas. Em adição, as regras de associação permitem descobrir relações entre os dados, a fim de derivar novos conhecimentos a partir do conjunto de dados analisado. Com o algoritmo C4.5 foram feitos experimentos utilizando todo conjunto de dados, com poda da AD.

Na segunda etapa foi desenvolvido um Mecanismo de Inferência *Fuzzy* (MIF) para auxiliar a identificação dos fatores motivadores para prática de atividades físicas a partir do perfil de um indivíduo e um sistema de

recomendação no qual o usuário entra com seus dados socioeconômicos e recebe as devidas recomendações de forma automática.

As regras do MIF são aquelas descritas pela AD e pelas regras de associação geradas na primeira etapa. O objetivo do MIF é predizer o grau de cada fator motivador com base nos valores das variáveis que descrevem o perfil socioeconômico de uma pessoa. Ele foi implementado em linguagem *Python*, por meio do Jupyter Notebook do Google (Google Colab).

Posteriormente, os perfis dos usuários testados foram modelados para serem utilizados no Sistema de Recomendação (SR), desenvolvido em linguagem *JavaScript* com *NextJS framework*, *HyperText Markup Language (HTML)*, *TypeScript*, e *Cascading Style Sheets (CSS)* para prover a coleta dos dados por meio de um formulário. O *Backend*, desenvolvido em *TypeScript*, cria o perfil do usuário através de um algoritmo que modela o perfil do usuário (cria representação adequada ao MIF) a fim de fazer a recomendação através do acionamento do MIF. Em outras palavras, o MIF constitui o núcleo do SR.

### **3.2.1 Aplicação das etapas do KDD**

Seleção de dados: Primeiramente tentou-se buscar o conjunto de dados na literatura e por meio de diversos contatos com responsáveis pelos dados apresentados nos trabalhos relacionados apresentados na seção 1.2, mas não se obteve êxito nesta tarefa. Diante disso, outras soluções foram investigadas até se conseguir o conjunto de dados da empresa S2Pro Data Health. Durante a seleção dos dados, foram aplicados critérios para filtrar informações que não representavam qualidade nos dados e outras informações não relevantes.

Pré-processamento: Nenhum pré-processamento de dados foi necessário em virtude da coleta de dados ter sido feita através de formulário *web* com validações, e pelo fato da S2Pro Data Health ter feito a checagem dos dados após a coleta.

Transformação: Para utilizar a base de dados foi necessário categorizar ou recategorizar os atributos faixa etária, cor, estado civil, escolaridade e faixa salarial visando a geração de ADs menores e mais eficientes. Além disso, foi excluído o atributo profissão em virtude da falta de padronização nas respostas.

Para o atributo faixa etária foram adotadas as seguintes categorias: “18 à 25 anos”, “26 à 35 anos”, “36 à 45 anos”, “46 à 55 anos” e “acima de 55 anos”. O atributo de escolaridade foi recategorizado como segue: “2º Grau”, “Ensino Superior”, e “Pós-graduação”. O atributo faixa salarial foi recategorizado como segue: “Até 1 salário”, “De 1 à 5 salários”, “De 6 à 10 salários”, “Mais de 10 salários”, e “Não respondido”.

Mineração de dados: Primeiro, a partir dos dados pré-processados e transformados foram criados sete arquivos ".ARFF" (extensão usada pelo WEKA), um para cada fator motivacional.

Na sequência foram geradas regras para relacionar as diversas variáveis sob este estudo com o emprego de Árvores de Decisão (AD) e o algoritmo Apriori. No caso das ADs, utilizou-se o algoritmo C4.5, que representa as relações entre as variáveis no formato de uma árvore, a partir da qual puderam ser lidas as regras.

O algoritmo C4.5 foi aplicado com os seguintes parâmetros: número mínimo de instâncias por nó igual a 2 e fator de confiança mínima utilizada em podas igual a 0,25. O fator de confiança de uma regra de associação corresponde ao nível em que a regra se comprova ao analisar as entradas de maneira própria. Esse valor é obtido ao dividir a quantidade de transações que corroboram a relação pelo número de transações que sustentam exclusivamente a primeira porção da regra.

O experimento foi configurado para testar o modelo obtido utilizando-se a mesma base de dados, com aplicação do método de poda. Tanto as regras geradas pelas ADs quanto pelo algoritmo Apriori representam padrões descobertos a partir da base de dados e que constituirão a base de regras do MIF. Nos experimentos com o algoritmo Apriori foram empregados os seguintes parâmetros para avaliar os resultados: grau mínimo de confiabilidade de 40% e todas as regras têm como consequente um dos fatores de motivação.

A classificação dos atributos teve como objetivo investigar a existência de padrões que podem levar a identificar perfis em relação aos graus de motivação para a prática de atividades físicas. Já as regras de associação permitiram

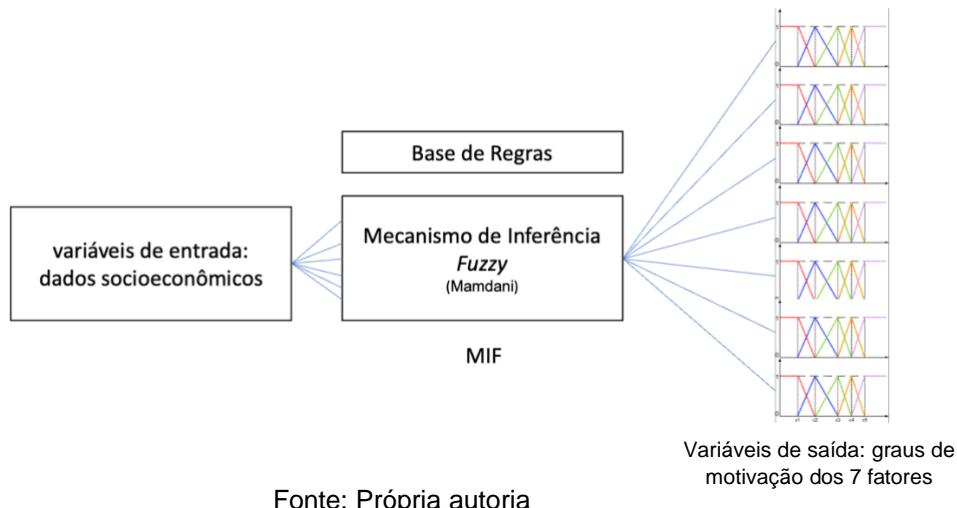
descobrir correlações entre os dados, a fim de derivar novo conhecimento do conjunto de dados analisado.

Interpretação/avaliação: A interpretação dos resultados produzidos na MD foi feita a partir da análise dos padrões descritos pelas regras geradas, os quais foram avaliados com base na acurácia, no índice Kappa, e no fator de confiança (no caso das regras de associação). Quanto maior a medida de avaliação, maior a consistência dos padrões encontrados.

### 3.2.2 Mecanismo de inferência *Fuzzy* (MIF)

O MIF, cuja arquitetura é ilustrada na Figura 6, foi implementado na linguagem de programação *Python* com o uso do Google Colab e a biblioteca SciKit-Fuzzy, e seu objetivo é predizer o grau de cada fator motivador com base nos valores das variáveis que representam um perfil socioeconômico.

Figura 6 – Arquitetura do MIF



A base de regras do MIF foi constituída a partir das regras produzidas na etapa de MD. A base contempla 403 regras do tipo **SE... ENTÃO**, como a mostrada no exemplo a seguir: **SE** Sexo = Masculino **E** Cor = Pardo **ENTÃO** grau de condicionamento físico = **Alto**. O MIF foi projetado com 6 variáveis de entrada representando dados pessoais do usuário (estado civil, cor, escolaridade, faixa salarial, faixa etária, e gênero) e 7 variáveis de saída representando os fatores motivacionais, quais sejam:

grau\_condicionamento\_fisico, grau\_estetica, grau\_saude, grau\_integracao\_social, grau\_ansiedade\_estresse, grau\_disponibilidade\_de\_tempo e grau\_custo.

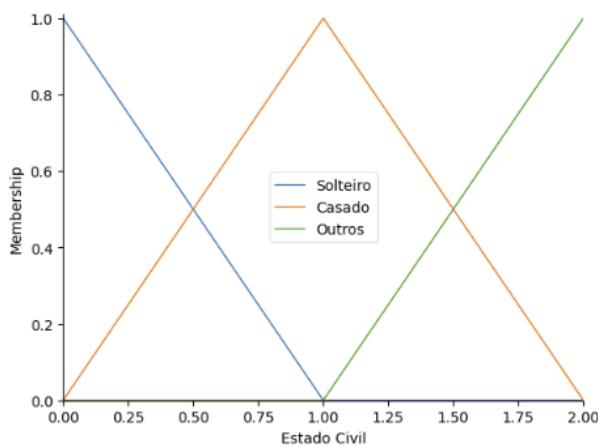
Para modelar as variáveis de entrada e de saída, ilustradas nas figuras 7 a 13, foram empregadas funções de pertinência dos tipo triangular e trapezoidal, pois possuem simplicidade na forma da função e permitem uma satisfatória representação dos termos linguísticos (PEDRYCZ, 1994). Em adição, os conjuntos *fuzzy* associados às variáveis de entrada foram modelados com base nos dados listados na Tabela 6.

**Tabela 6 – Conjuntos *fuzzy* para variáveis de entrada**

Estado Civil (C, 3 conjuntos)	Cor (COR, 4 conjuntos)	Escolaridade (E, 4 conjuntos)
C1, C2, C3 ['Solteiro', 'Casado', 'Separado']	COR1, COR2, COR3, COR4 ['Branco', 'Pardo', 'Preto', 'Amarelo']	E1, E2, E3, E4 ['1Grau', '2Grau', 'Sup', 'Pós-grad']
Faixa Salarial (S, 5 conjuntos)	Faixa Etária (I, 5 conjuntos)	Gênero (G, 3 conjuntos)
S1, S2, S3, S4, S5 ['<1SM', '>1-5SM', '6-10SM', '>10SM', 'Sem-info']	I1, I2, I3, I4, I5 ['18-25A', '26-35A', '36-45A', '46-55A', '56-99A']	G1, G2, G3 ['F', 'NB', 'M']

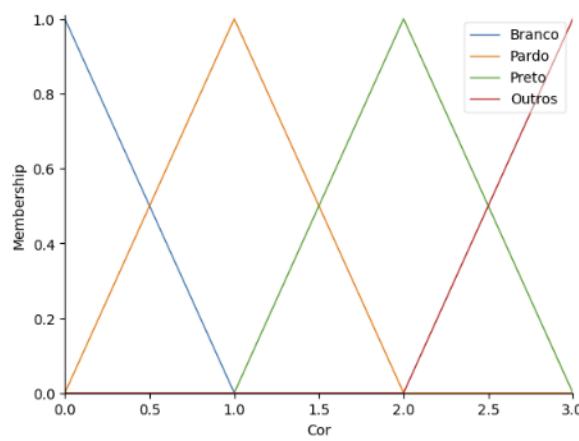
Fonte: Própria autoria

**Figura 7 – Funções de pertinência associadas à variável de entrada Estado Civil**



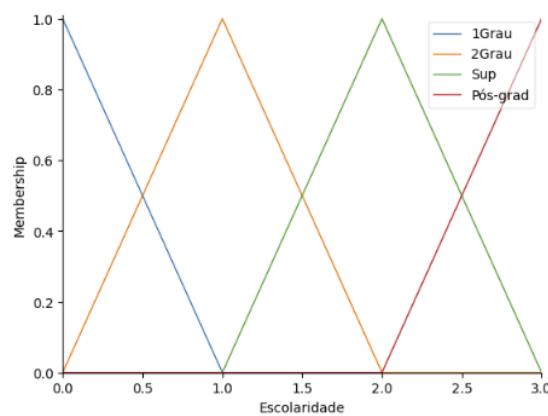
Fonte: Autoria própria

**Figura 8 – Funções de pertinência associadas à variável de entrada Cor**



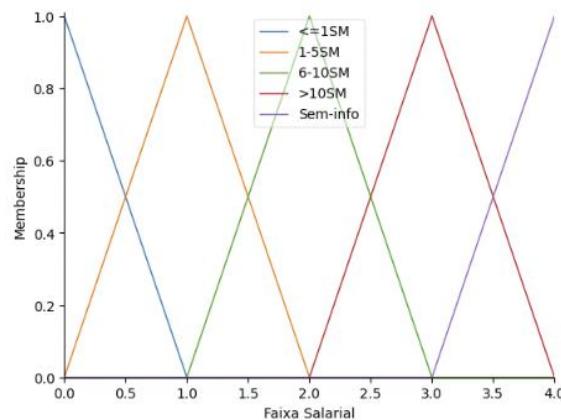
Fonte: Autoria própria

**Figura 9 – Funções de pertinência associadas à variável de entrada Escolaridade**



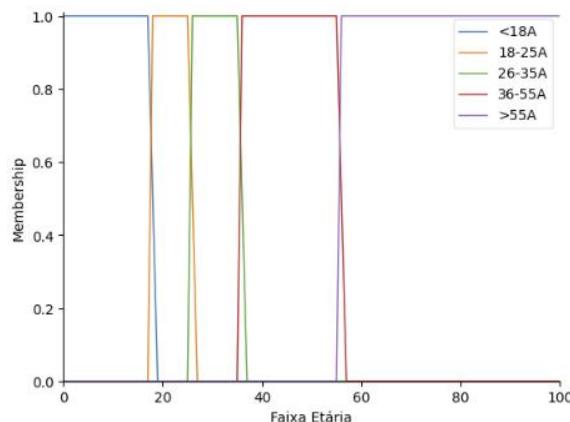
Fonte: Autoria própria

**Figura 10 – Funções de pertinência associadas à variável de entrada Faixa Salarial**



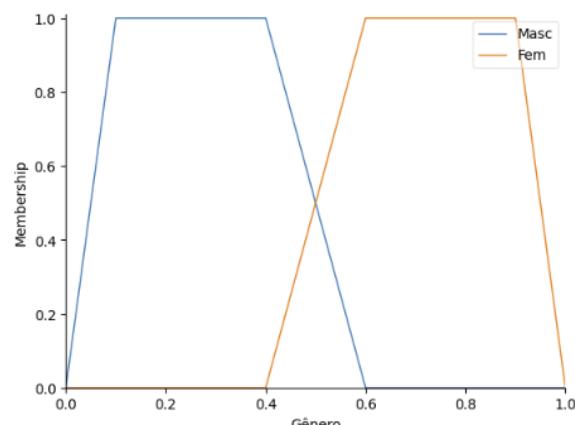
Fonte: Autoria própria

**Figura 11 – Funções de pertinência associadas à variável de entrada Faixa Etária**



Fonte: Autoria própria

**Figura 12 – Funções de pertinência associadas à variável de entrada Gênero**

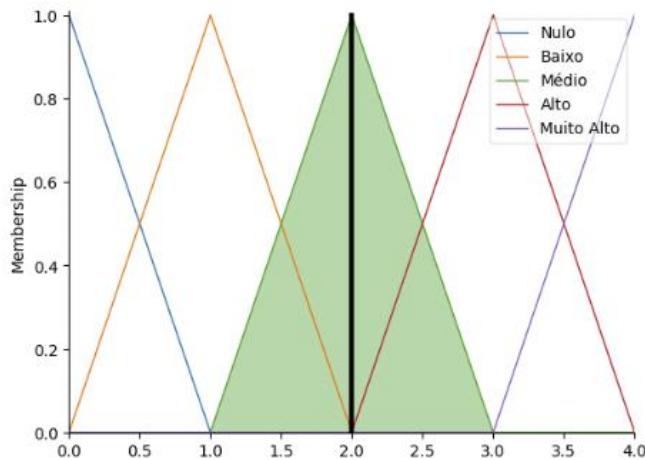


Fonte: Autoria própria

As variáveis idade e estado civil contemplam funções de pertinência correspondentes às categorias dos dados originais ou transformados, conforme descrito na seção 3.1. Para a variável gênero, uma adaptação foi feita na variável sexo, a fim de representar feminino, masculino e não-binário. Mesmo essas variáveis não sendo naturalmente nebulosas, a aplicabilidade do MIF permanece válida.

As 7 variáveis de saída foram modeladas com funções de pertinência idênticas, quais sejam: nulo (0); baixo (1); médio (2); alto (3) e muito alto (4), correspondentes aos graus associados aos 7 fatores motivacionais. Um exemplo da modelagem das funções de pertinência para as variáveis de saída é ilustrado na Figura 13.

**Figura 13 – Funções de pertinência associadas às variáveis de saída**



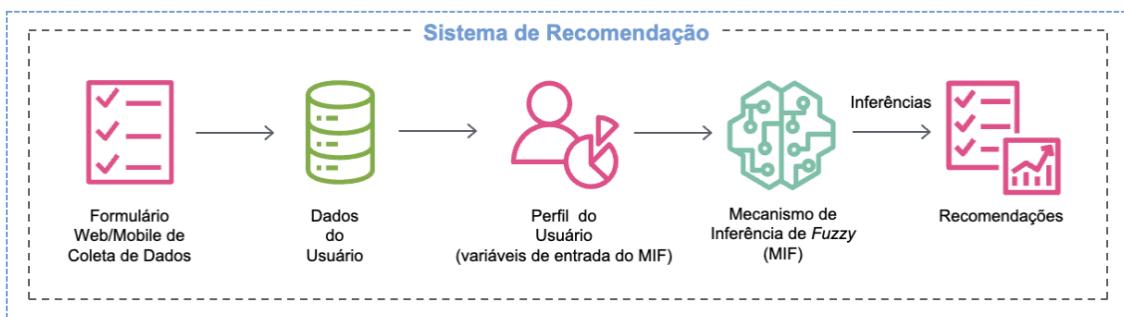
**Fonte:** Autoria própria

Por fim, para o processo de defuzzyficação foi empregado o método de Mamdani, o qual traz como resposta um conjunto *fuzzy* originado da combinação dos valores de entrada com os seus respectivos graus de pertinência, por meio do operador mínimo e, em seguida, pela superposição das regras pelo operador máximo. Em resumo, a defuzzificação realiza a “tradução” do valor linguístico em valor numérico *crisp*.

### 3.3 Sistema de recomendação (SR)

Foi desenvolvido um protótipo do Sistema de Recomendação (SR), que tem como foco o usuário final e cujo diagrama de funcionamento é ilustrado na Figura 14. Para tanto, foi utilizado o *JavaScript* com *framework NextJS*, *HTML*, *TypeScript* e *CSS*. O SR contempla um formulário que o torna uma aplicação *Web* responsiva, que poderá ser utilizada também em celulares e *tablets*. Para armazenar os perfis dos usuários utilizou-se o banco de dados *MongoDB*, que pode ser facilmente acessado por rotinas escritas em *JavaScript*.

**Figura 14 – Protótipo esquemático do Sistema de Recomendação**



Fonte: Autoria própria

Assim, a coleta dos dados é realizada por meio do formulário e um algoritmo modela o perfil do usuário adequando-o às variáveis de entrada do MIF. Após a inserção dos dados do usuário, o SR aciona o MIF para fazer as previsões e recomendar exercícios adequados de acordo com os graus preditos para os fatores motivacionais relacionados ao usuário. Algumas telas de formulário do SR estão ilustradas na Figura 15 a seguir.

O desenvolvimento do protótipo do SR foi conduzido com o auxílio de dois profissionais da área de educação física, com conhecimentos teóricos (fisiologia do exercício e avaliação física e prescrição de exercícios) e práticos (ministração de aulas de musculação, aeróbico, ciclismo e maratona).

Figura 15 – Exemplos de telas de entrada dos dados do SR

The figure consists of three vertically stacked screenshots of a mobile application interface, likely for a fitness or health tracking system. All three screenshots feature a top navigation bar with the 'S2 PRO' logo, 'Meus Treinos', 'Treinos do Sistema', 'Feed', 'Perfil', 'Dashboard' (highlighted in green), and 'Sair' buttons.

**Screenshot 1: Pratica Atividade Física?**

This screen asks the user if they practice physical activity. It contains two buttons: 'Sim' (selected) and 'Não'. A green 'Próximo >>' button is at the bottom.

**Screenshot 2: Data de Nascimento**

This screen displays the user's birth date as '04/10/1978' in a text input field, with a clear 'X' icon to its right. Below the input are two green buttons: '<< Anterior' and 'Próximo >>'.

**Screenshot 3: Sexo**

This screen asks for the user's gender. It shows 'Masculino' (selected) and 'Feminino' buttons. A green 'Próximo >>' button is at the bottom.

Fonte: Captura de telas do Sistema de Recomendação

## 4 RESULTADOS E DISCUSSÃO

Este capítulo contempla os resultados da descoberta de padrões relacionados aos fatores que motivam a prática de atividades físicas, tanto por meio das ADs (seção 4.1) quanto pelo algoritmo Apriori (seção 4.2). Também são apresentados os resultados obtidos pelo MIF (seção 4.3) e pelo Sistema de Recomendação (seção 4.4).

### 4.1 Mineração de Dados (MD)

#### 4.1.1 Análise dos dados por meio de Árvores de Decisão (classificação)

Nesta seção são apresentadas as análises feitas para cada fator motivacional por meio de AD gerada pelo algoritmo C4.5 (J48 no WEKA), mostrando o seu desempenho com base no número de instâncias classificadas corretamente (acurácia). Foram criadas sete ADs, apresentadas nos itens (a) a (g), sendo uma para cada fator motivacional, quais sejam: condicionamento físico, estética, saúde, integração social, ansiedade/estresse, disponibilidade de tempo e custo.

##### a) Condicionamento físico

A AD gerada para análise do fator motivacional condicionamento físico está apresentada no Apêndice I. Na classificação das 140 instâncias de treinamento, 103 foram corretamente classificadas, levando a uma acurácia de 73,6% e uma concordância substancial pelo índice Kappa (62,9%). A partir da matriz de confusão apresentada na Tabela 7 pode-se observar que as maiores taxas de erro de classificação foram obtidas nos graus de motivação extremos (“nulo” e “muito alto”), o que pode indicar algum viés nas respostas para esses graus induzido pelo baixo número de respostas dos participantes da pesquisa realizada pela empresa S2Pro para os graus extremos. Esse problema foi observado em todos os fatores.

**Tabela 7 – Matriz de confusão obtida na classificação dos dados para o fator “Condicionamento Físico”**

<b>Classe predita</b>					<b>Classe real</b>
<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>	
40	4	1	1	1	<b>a = Médio</b>
6	29	0	4	0	<b>b = Baixo</b>
2	2	4	0	0	<b>c = Nulo</b>
7	5	1	27	0	<b>d = Alto</b>
1	2	0	0	3	<b>e = Muito alto</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

Ainda com base na mesma AD, pode-se derivar os seguintes conhecimentos sobre o fator motivacional condicionamento físico:

- Pessoas brancas, solteiras, com ensino superior e faixa salarial de 1 a 5 salários mínimos, têm baixa motivação com relação a este fator.
- Pessoas com pós-graduação, faixa etária entre 36 e 45 anos e faixa salarial de 6 a 10 salários mínimos (ou que não informaram faixa salarial), têm média motivação com relação a este fator.
- Pessoas casadas, de cor branca, com faixa salarial de mais de 10 salários mínimos, têm alta motivação com relação a este fator.

#### b) Estética

A AD gerada para análise do fator motivacional estética está apresentada no Apêndice II. Na classificação das 140 instâncias de treinamento, 90 foram corretamente classificadas, levando a uma acurácia de 64,3% e uma concordância substancial pelo índice Kappa (49,5%). A partir da matriz de confusão apresentada na Tabela 8 pode-se observar que as maiores taxas de erro de classificação foram obtidas nos graus de motivação extremos (“nulo”, “baixo” e “muito alto”).

**Tabela 8 – Matriz de confusão obtida na classificação dos dados para o fator “Estética”**

Classe predita					Classe real
a	b	c	d	e	
42	1	3	3	0	<b>a = Médio</b>
4	7	2	1	0	<b>b = Muito Alto</b>
14	2	29	1	0	<b>c = Alto</b>
5	1	4	10	1	<b>d = Baixo</b>
1	0	4	3	2	<b>e = Nulo</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

A partir da AD mostrada no Apêndice II, pode-se derivar os seguintes conhecimentos sobre o fator motivacional estética:

- Pessoas do sexo feminino, solteiras, com faixa etária entre 26 e 35 anos, têm alta motivação com relação a este fator.
- Pessoas casadas, com faixa etária entre 26 e 35 anos, e faixa salarial entre 1 e 5 salários mínimos, têm média motivação com relação a este fator.
- Pessoas do sexo masculino, com faixa etária entre 36 e 45 anos, e faixa salarial com mais de 10 salários mínimos, não possuem nenhuma motivação com relação a este fator.
- Pessoas da cor branca, com faixa etária entre 46 e 55 anos, e pós-graduação, têm média motivação com relação a este fator.

### c) Saúde

A AD gerada para análise do fator motivacional saúde está apresentada no Apêndice III. Na classificação das 140 instâncias de treinamento, 92 foram corretamente classificadas, produzindo uma acurácia de 65,7% e uma concordância substancial pelo índice Kappa (50%). A partir da matriz de confusão apresentada na Tabela 9 pode-se observar que as maiores taxas de erro de classificação foram obtidas nos graus de motivação extremos (“nulo” e “baixo”).

**Tabela 9 – Matriz de confusão obtida na classificação dos dados para o fator “Saúde”**

Classe predita					Classe real
a	b	c	d	e	
38	3	11	0	0	<b>a = Alto</b>
7	21	2	0	1	<b>b = Médio</b>
11	5	30	0	1	<b>c = Muito alto</b>
0	2	0	0	0	<b>d = Nulo</b>
3	2	0	0	3	<b>e = Baixo</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

A partir da AD mostrada no Apêndice III, pode-se derivar os seguintes conhecimentos sobre o fator motivacional saúde:

- Pessoas com nível superior, de cor parda, têm alta motivação com relação a este fator.
- Pessoas com faixa salarial entre 6 e 10 salários mínimos, de cor branca, com pós-graduação, têm altíssima motivação com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, e faixa salarial com mais de 10 salários mínimos, casadas, com pós-graduação, têm motivação média com relação a este fator.
- Pessoas com nível superior, de cor branca, solteiras, com faixa salarial entre 1 e 5 salários mínimos, têm alta motivação com relação a este fator.

#### d) Integração social

A AD gerada para análise do fator motivacional integração social está apresentada no Apêndice IV. Na classificação das 140 instâncias de treinamento, 90 foram corretamente classificadas, levando a uma acurácia de 64,3% e uma concordância substancial pelo índice Kappa (53,5%). A partir da matriz de confusão apresentada na Tabela 10 pode-se observar que as maiores taxas de erro de classificação foram obtidas nos graus de motivação extremos (“alto” e “muito alto”).

**Tabela 10 – Matriz de confusão obtida na classificação dos dados para o fator “Integração Social”**

<b>Classe predita</b>					<b>Classe real</b>
<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>	
30	1	1	0	0	<b>a = Baixo</b>
6	13	5	0	1	<b>b = Nulo</b>
3	4	29	1	3	<b>c = Médio</b>
4	2	0	5	0	<b>d = Muito alto</b>
7	3	8	1	13	<b>e = Alto</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

A partir da AD mostrada no Apêndice IV, pode-se derivar os seguintes conhecimentos sobre o fator motivacional integração social:

- Pessoas com nível superior, de cor branca, faixa salarial de 1 a 5 salários mínimos, faixa etária entre 26 e 35 anos, têm baixa motivação com relação a este fator.
- Pessoas de cor parda, faixa salarial de 1 a 5 salários mínimos, faixa etária entre 26 e 35 anos, têm motivação média com relação a este fator.
- Pessoas com faixa etária entre 26 e 35 anos, cor branca, e faixa salarial entre 6 e 10 salários mínimos, não possuem nenhuma motivação com relação a este fator.
- Pessoas de cor branca, sexo masculino, com faixa salarial com mais de 10 salários mínimos, faixa etária entre 36 e 45 anos, não possuem nenhuma motivação com relação a este fator.

#### e) Ansiedade/Estresse

A AD gerada para análise do fator motivacional ansiedade/estresse está apresentada no Apêndice V. Na classificação das 140 instâncias de treinamento, 88 foram corretamente classificadas, levando a uma acurácia de 62,9% e uma concordância substancial pelo índice Kappa (50,6%). A partir da matriz de confusão apresentada na Tabela 11 pode-se observar que as maiores taxas de erro de classificação foram obtidas nos graus de motivação nulo.

**Tabela 11 – Matriz de confusão obtida na classificação dos dados para o fator “Ansiedade/Estresse”**

<b>Classe predita</b>					<b>Classe real</b>
<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>E</b>	
33	3	0	0	1	<b>a = Alto</b>
8	31	0	1	1	<b>b = Médio</b>
3	1	4	1	1	<b>c = Nulo</b>
12	5	2	8	1	<b>d = Baixo</b>
4	5	2	1	12	<b>e = Muito alto</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

A partir da AD mostrada no Apêndice V, pode-se derivar os seguintes conhecimentos sobre o fator motivacional ansiedade/estresse:

- Pessoas de cor parda, faixa etária entre 18 e 25 anos, têm média motivação com relação a este fator.
- Pessoas de cor parda, com pós-graduação, casadas, faixa etária entre 26 e 35 anos, têm motivação alta com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, cor branca, e faixa salarial entre 1 e 5 salários mínimos, e ensino superior, têm alta motivação com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, cor branca, e faixa salarial com mais de 10 salários mínimos, têm média motivação com relação a este fator.

#### f) Disponibilidade de tempo

A AD gerada para análise do fator motivacional disponibilidade de tempo está apresentada no Apêndice VI. Na classificação das 140 instâncias de treinamento, 99 foram corretamente classificadas, levando a uma acurácia de 70,7% e uma concordância substancial pelo índice Kappa (54,6%). A partir da matriz de confusão apresentada na Tabela 12 pode-se observar que as maiores taxas de erro de classificação foram obtidas nos graus de motivação extremos (“nulo”, “alto” e “muito alto”).

**Tabela 12 – Matriz de confusão obtida na classificação dos dados para o fator “Disponibilidade de Tempo”**

<b>Classe predita</b>					<b>Classe real</b>
<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>E</b>	
41	4	1	1	0	<b>a = Baixo</b>
12	47	1	2	0	<b>b = Médio</b>
4	4	5	1	0	<b>c = Nulo</b>
2	5	0	5	0	<b>d = Alto</b>
0	4	0	0	1	<b>e = Muito alto</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

A partir da AD mostrada no Apêndice VI, pode-se derivar os seguintes conhecimentos sobre o fator motivacional disponibilidade de tempo:

- Pessoas de cor branca, faixa etária entre 26 e 35 anos, solteiras, com faixa salarial entre 1 e 5 salários mínimos, têm alta motivação com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, com ensino superior, com faixa salarial entre 1 e 5 salários mínimos, têm baixa motivação com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, com ensino superior, com faixa salarial entre 6 e 10 salários mínimos, têm média motivação com relação a este fator.
- Pessoas de faixa salarial com mais de 10 salários mínimos, cor branca, e pós-graduação, têm baixa motivação com relação a este fator.

g) Custo

A AD gerada para análise do fator motivacional custo está apresentada no Apêndice VII. Na classificação das 140 instâncias de treinamento, 99 foram corretamente classificadas, levando a uma acurácia de 70,7% e uma concordância substancial pelo índice Kappa (52,5%). A partir da matriz de confusão apresentada na Tabela 13 pode-se observar que as maiores taxas de erro de classificação de foram obtidas nos graus de motivação alto e muito alto.

**Tabela 13 – Matriz de confusão obtida na classificação dos dados para o fator “Custo”**

Classe predita					Classe real
a	b	c	d	E	
12	4	1	0	1	<b>a = Nulo</b>
0	21	14	0	0	<b>b = Baixo</b>
5	8	61	0	0	<b>c = Médio</b>
1	1	2	1	1	<b>d = Muito alto</b>
0	1	2	0	4	<b>e = Alto</b>

Fonte: Autoria própria com base na matriz de confusão gerada pelo software WEKA 3.8.6

A partir da AD mostrada no Apêndice VII, pode-se derivar os seguintes conhecimentos sobre o fator motivacional custo:

- Pessoas com faixa etária entre 26 e 35 anos e ensino superior, de cor branca, têm média motivação com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, casados, ensino superior, têm média motivação com relação a este fator.
- Pessoas com faixa etária entre 26 e 35 anos, de cor parda e pós-graduação, têm média motivação com relação a este fator.
- Pessoas com faixa etária entre 36 e 45 anos, pós-graduação, faixa salarial com mais 10 salários mínimos, de cor branca, têm baixa motivação com relação a este fator.

#### 4.1.2 Geração de regras de associação usando o algoritmo Apriori

Os conhecimentos descritos por meio das regras apresentadas a seguir foram gerados pelo algoritmo Apriori para os 7 fatores considerados nesta pesquisa.

##### a) Condicionamento Físico

- 53% dos respondentes com faixa salarial entre 6 e 10 salários mínimos associam um médio grau de motivação a este fator.
- 47% dos respondentes com faixa salarial entre 1 e 5 salários mínimos e cor branca, associam um baixo grau de motivação a este fator.

- 45% dos respondentes solteiro(a)s associam um alto grau de motivação a este fator.
- 43% dos respondentes com faixa etária entre 36 e 45 anos, associam um médio grau de motivação a este fator.
- 42% dos respondentes de cor parda, associam um alto grau de motivação a este fator.
- 41% dos respondentes casado(a)s com faixa etária entre 36 e 45 anos, associam um médio grau de motivação a este fator.
- 41% dos respondentes casado(a)s com pós-graduação, associam um médio grau de motivação a este fator.
- 41% dos respondentes casado(a)s do sexo feminino, associam um médio grau de motivação a este fator.

b) Estética

- 50% dos respondentes do sexo feminino, solteiros, associam um alto grau de motivação para este fator.
- 48% dos respondentes com faixa etária entre 36 e 45 anos, do sexo feminino, associam um médio grau de motivação para este fator.
- 47% dos respondentes com faixa etária entre 26 e 35 anos associam um alto grau de motivação para este fator.
- 47% dos respondentes solteiros associam um alto grau de motivação para este fator.
- 47% dos respondentes do sexo feminino, com pós-graduação, associam um médio grau de motivação para este fator.
- 46% dos respondentes com faixa etária entre 36 e 45 anos, casados, associam um médio grau de motivação para este fator.
- 46% dos respondentes casados, com pós-graduação, associam um médio grau de motivação para este fator.
- 46% dos respondentes do sexo feminino, casados, associam um médio grau de motivação para este fator.

c) Saúde

- 53% dos respondentes com faixa salarial entre 6 e 10 salários mínimos, associam um alto grau de motivação para este fator.
- 52% dos respondentes com ensino superior, associam um alto grau de motivação para este fator.
- 51% dos respondentes solteiros, associam um alto grau de motivação para este fator.
- 50% dos respondentes do sexo feminino, solteiros, associam um alto grau de motivação para este fator.
- 49% dos respondentes do sexo feminino com ensino superior, associam um alto grau de motivação para este fator.
- 48% dos respondentes com faixa etária entre 36 e 45 anos, do sexo feminino, associam um alto grau de motivação para este fator.
- 47% dos respondentes de cor parda, associam um alto grau de motivação para este fator.
- 46% dos respondentes com pós-graduação, casados, associam um alto grau de motivação para este fator.
- 45% dos respondentes de cor branca, solteiros, associam um alto grau de motivação para este fator.

d) Integração Social

- 50% dos respondentes de cor parda, associam um médio grau de motivação para este fator.
- 50% dos respondentes de cor branca, com faixa salarial entre 1 e 5 salários mínimos, associam um baixo grau de motivação para este fator.

e) Ansiedade/Estresse

- 58% dos respondentes do sexo feminino, de cor branca, com ensino superior, associam um alto grau de motivação para este fator.
- 47% dos respondentes de cor branca, com ensino superior, associam um alto grau de motivação para este fator.

- 44% dos respondentes com faixa etária entre 36 e 45, de cor branca, associam um alto grau de motivação para este fator.
- 43% dos respondentes do sexo feminino, com ensino superior, associam um alto grau de motivação para este fator.

f) Disponibilidade de Tempo

- 75% dos respondentes de cor branca, com faixa salarial entre 6 e 10 salários mínimos, associam um médio grau de motivação para este fator.
- 70% dos respondentes com faixa salarial entre 6 e 10 salários mínimos, associam um médio grau de motivação para este fator.
- 64% dos respondentes com mais de 10 salários mínimos de faixa salarial, associam um baixo grau de motivação para este fator.
- 60% dos respondentes do sexo feminino, com pós-graduação, associam um médio grau de motivação para este fator.
- 56% dos respondentes do sexo feminino, casados, associam um médio grau de motivação para este fator.
- 55% dos respondentes com faixa etária entre 36 e 45, do sexo feminino, associam um médio grau de motivação para este fator.

g) Custo

- 76% dos respondentes com ensino superior, associam um médio grau de motivação para este fator.
- 74% dos respondentes do sexo feminino, com ensino superior, associam um médio grau de motivação para este fator.
- 72% dos respondentes de cor branca, com ensino superior, associam um médio grau de motivação para este fator.
- 58% dos respondentes de faixa etária entre 26 e 35, associam um médio grau de motivação para este fator.
- 55% dos respondentes solteiros associam um médio grau de motivação para este fator.
- 54% dos respondentes do sexo feminino associam um médio grau de motivação para este fator.

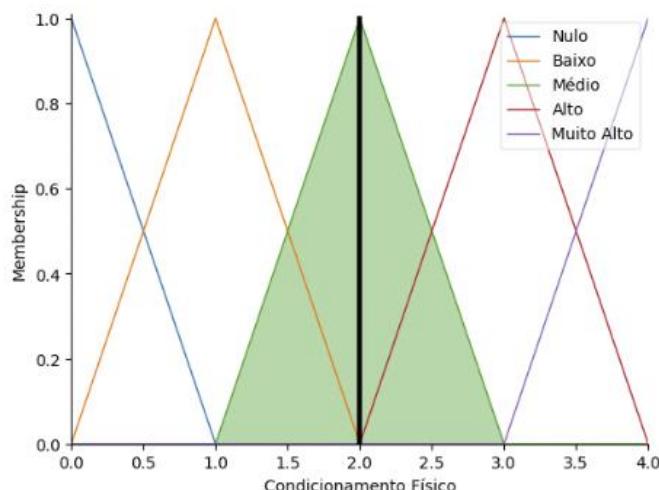
- 51% dos respondentes do sexo masculino associam um médio grau de motivação para este fator.

#### 4.2 Mecanismo de Inferência Fuzzy (MIF)

A avaliação do MIF foi realizada por meio de comparação entre os graus preditos pelo MIF para os fatores motivacionais (linha vermelha) e os graus reais contidos na base de dados (linha azul). Devido ao baixo número de dados disponíveis, foram selecionados aleatoriamente dez registros da base que foram submetidos aos testes de predição.

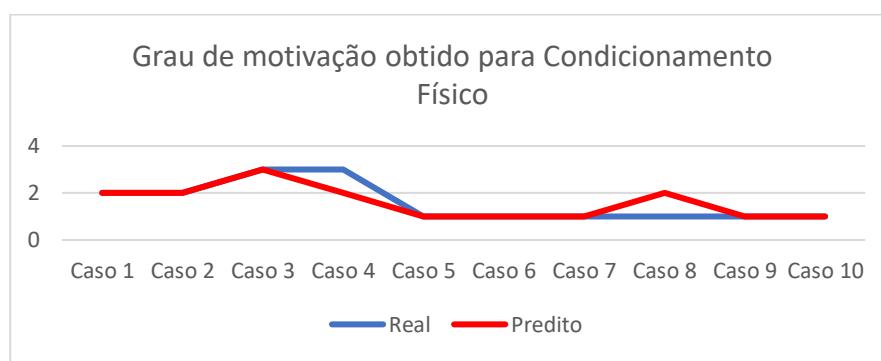
A Figura 16 ilustra a forma como as previsões são obtidas (exemplificando com a variável condicionamento físico), enquanto as figuras 17 a 23 ilustram os resultados de algumas previsões feitas pelo MIF.

**Figura 16 – Exemplo de previsão do MIF**

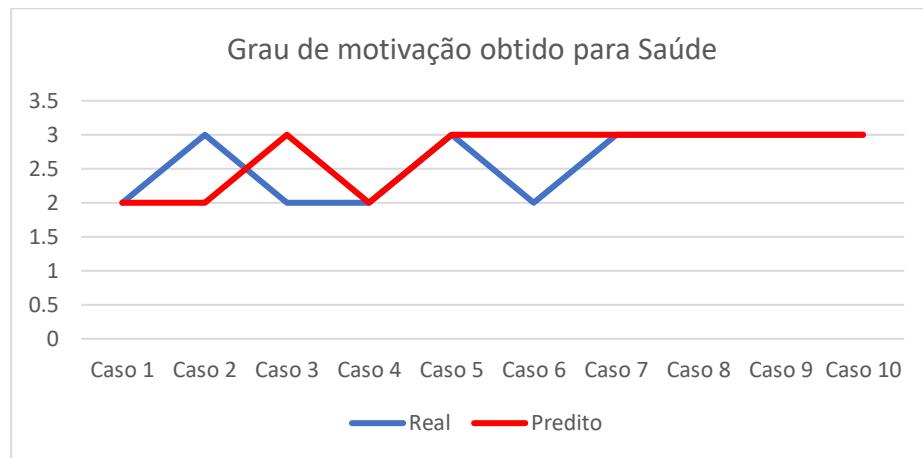


Fonte: Autoria própria

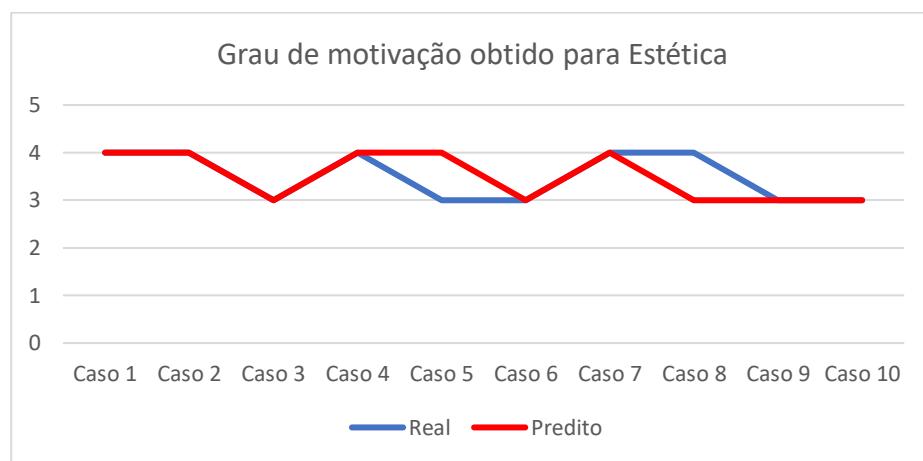
**Figura 17 – Alguns resultados do MIF para o fator Condicionamento Físico**



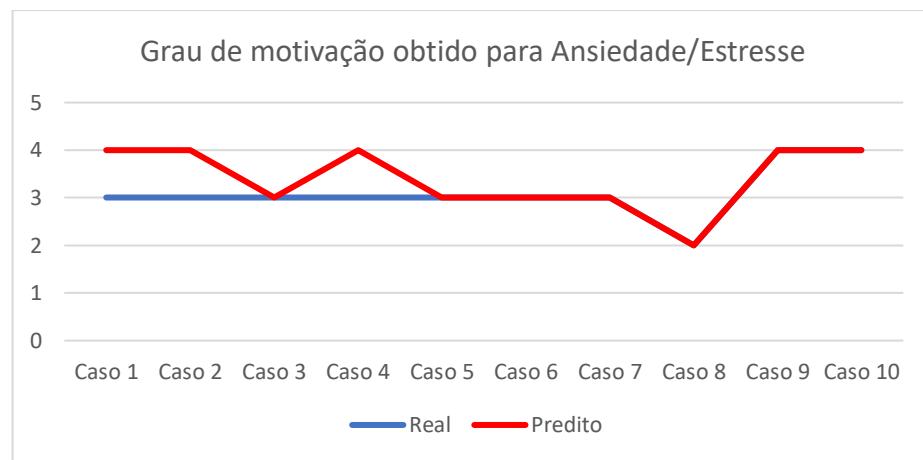
Fonte: Autoria própria

**Figura 18 – Alguns resultados do MIF para o fator Saúde**

Fonte: Autoria própria

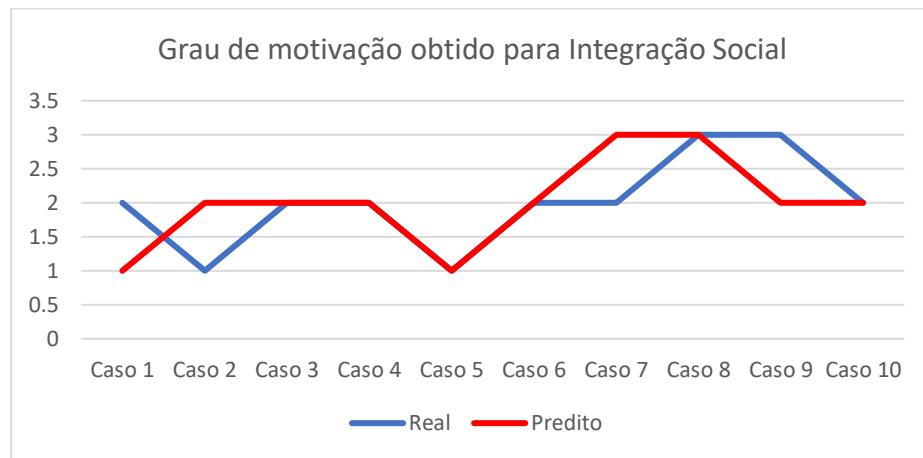
**Figura 19 – Alguns resultados do MIF para o fator Estética**

Fonte: Autoria própria

**Figura 20 – Alguns resultados do MIF para o fator Ansiedade/Estresse**

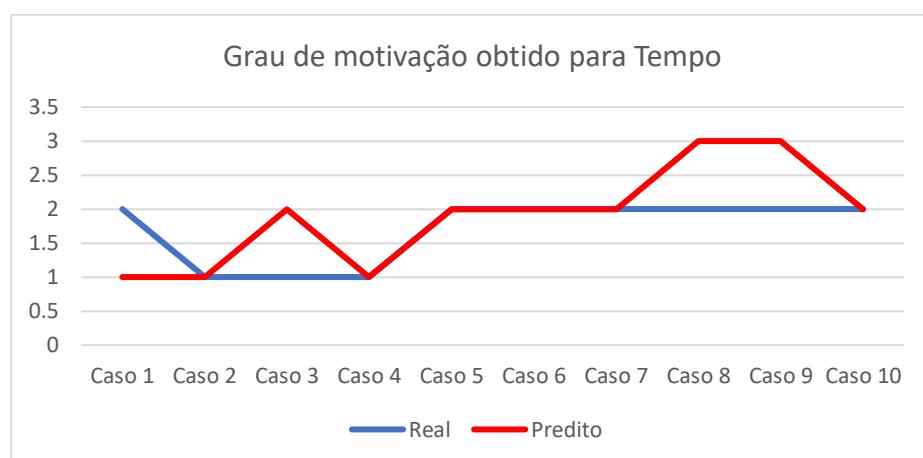
Fonte: Autoria própria

**Figura 21 – Alguns resultados do MIF para o fator Integração Social**



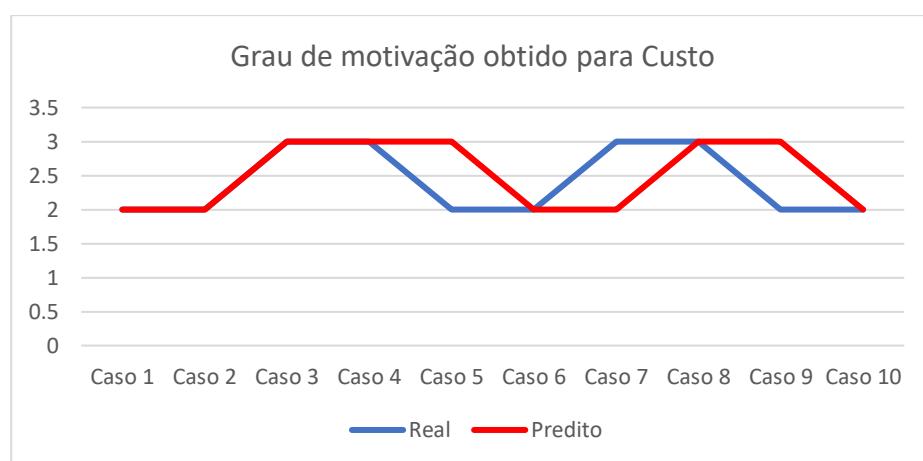
Fonte: Autoria própria

**Figura 22 – Alguns resultados do MIF para o fator Disponibilidade de Tempo**



Fonte: Autoria própria

**Figura 23 – Alguns resultados do MIF para o fator Custo**



Fonte: Autoria própria

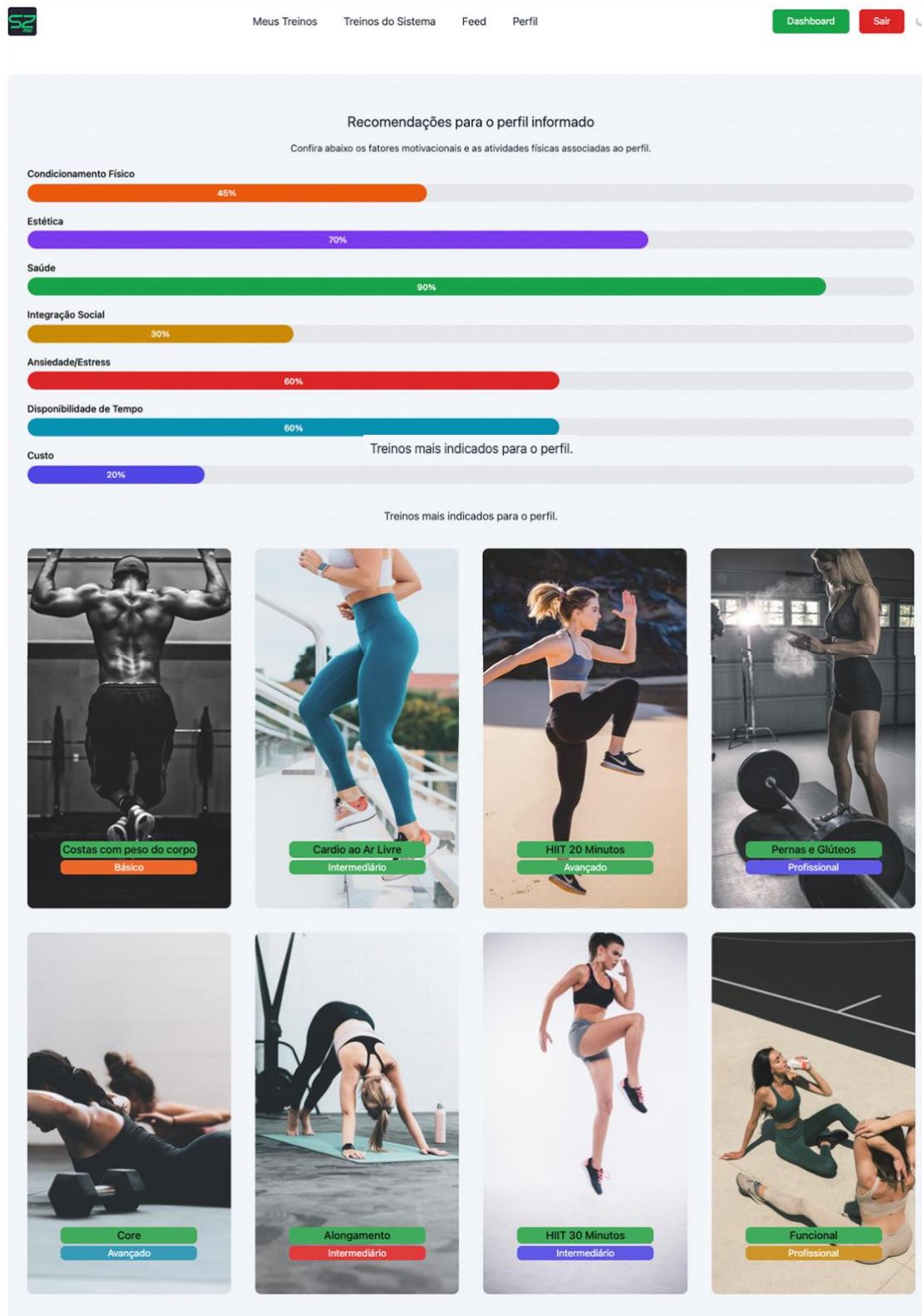
Conforme evidenciado nos exemplos apresentados nas figuras 17 a 23, para o fator motivacional condicionamento físico os resultados não satisfatórios puderam ser observados no caso 4 e no caso 8, para os quais as predições foram erradas. Já para o fator motivacional saúde, o MIF inferiu erroneamente 3 dos 10 casos testados. Para o fator estética, o MIF inferiu apenas 2 casos errados ao teste realizado, além disso, este fator indicou graus de motivação alto para a prática de atividades físicas, diferentemente de ansiedade/estresse que apesar de ter graus de interesses altos para a prática de atividades físicas relacionados, o MIF errou 3 dos 10 casos testados. Para integração social e disponibilidade de tempo, o MIF inferiu 4 casos errados dos 10 casos testados, e para o fator custo, o MIF inferiu 3 casos errados dos 10 casos testados.

Apesar dos erros observados, de maneira geral pode-se afirmar que o MIF alcançou um rendimento satisfatório. Observa-se que os resultados não foram melhores porque as regras não conseguem generalizar devido a quantidade limitada de dados empregada para produzir os padrões. Assim, uma análise mais profunda torna-se necessária visando aprimorá-lo ainda mais em termos de eficácia.

#### **4.3 Protótipo do Sistema de Recomendação**

O SR foi desenvolvido numa versão inicial onde um usuário informa seus dados socioeconômicos, para que ele acione o MIF e faça as recomendações de atividades físicas para o perfil informado, como ilustrado na Figura 24.

**Figura 24 – Tela de resultado do SR**



Fonte: Captura de tela do Sistema de Recomendação

Na Figura 24 pode ser visto um exemplo de recomendação de treinos voltados para o perfil indicado por barras coloridas mostrando em termos

percentuais os fatores motivacionais associados aos treinos. No exemplo é sugerido um treino de costas com peso do corpo indicado para condicionamento físico, um treino de pernas e glúteos indicado para estética, dois treinos de aeróbio (Cárdio ao Ar Livre e *HIT* 20 Minutos) recomendados para saúde, já que o MIF indicou para este fator um alto grau.

O treino de funcional foi associado ao fator motivacional de integração social, já que o mesmo pode ser feito ao ar livre e com uma companhia, sendo indicado para melhorar a flexibilidade e aumentar a resistência e o equilíbrio corporal para os movimentos do dia a dia.

Para o fator de ansiedade e estresse, um treino de alongamento foi indicado, o qual promove o estiramento das fibras musculares, fazendo com que elas aumentem o seu comprimento aliviando assim o estresse do dia a dia, um treino de core foi indicado ao fator tempo o qual está associado à potência, força e estabilização, justamente para quem precisa que seu tempo seja produtivo, e um treino de *HIT* de 30 minutos ao fator custo já que pode ser feito em qualquer lugar para quem precisa de economizar tempo.

Entretanto, apesar dos testes realizados demonstrarem a potencialidade do SR para orientar a pessoas na prática de atividades físicas, é importante ressaltar que apenas um protótipo foi desenvolvido e necessita de profissionais da área para validarem constantemente os resultados sugeridos pelo SR. Outra limitação do SR está relacionada ao MIF que compõe o seu núcleo. Tendo em vista que as regras usadas na modelagem do MIF foram geradas a partir de um conjunto de dados limitado, suas inferências podem apresentar vieses e outros problemas de generalização indicados na seção anterior.

#### **4.4 Discussão**

Os resultados apresentados na seção 4.1 demonstram como a MD pode ser uma ferramenta valiosa na área da educação física, permitindo analisar bases de dados e descobrir padrões e *insights* significativos, como mencionado no trabalho de Song (2022). Através da análise das ADs foi possível identificar, por exemplo, que os fatores de motivação com um maior grau para a para a prática de atividades físicas no grupo de pessoas de cor branca, com mais de 6 salários mínimos foram condicionamento físico e saúde, enquanto no grupo de

pessoas com faixa etária entre 26 e 45 anos, o fator motivacional de ansiedade/estresse normalmente tem alto grau de motivação.

Fatores como interação social, disponibilidade de tempo e custo mostraram menores estímulos para a prática de atividades físicas na população considerada. Em adição, o condicionamento físico é o fator que mais motiva a prática de atividades físicas por pessoas com mais de 5 salários mínimos ou pós-graduação.

Também foi possível identificar, a partir das matrizes de confusão, que a maioria dos fatores apresentavam uma baixa frequência de classificação nos extremos de motivação (“nulo” e “muito alto”), o que pode sugerir um possível enviesamento nas respostas relacionadas a esses graus. Isso pode estar relacionado ao fato das pessoas terem uma tendência natural de querer se encaixar nas normas sociais e serem aceitas em determinados grupos, o que as levam a evitar respostas extremas para evitar se destacar ou parecerem “diferentes”. Obviamente, a quantidade de dados considerada nesta pesquisa também é um ponto de limitação para o processo de descoberta de padrões.

Com relação ao MIF, cujas inferências são realizadas com base nas regras originadas pelas ADs e pelo algoritmo Apriori, pode-se verificar a sua capacidade de determinar o grau de motivação para cada um dos sete fatores considerados, permitindo um tratamento diferenciado para cada pessoa. Ressalta-se que, diferente de outros métodos de aprendizagem de máquina, um MIF possibilita previsões levando em conta a variação e a subjetividade que podem estar presentes tanto nos dados de entrada quanto de saída, e daí a sua escolha para compor o *framework* apresentado neste estudo. No entanto, é importante mencionar que as inferências feitas pelo MIF desenvolvido podem apresentar vieses e problemas de generalização em virtude das regras terem sido geradas a partir de um conjunto de dados limitado (apenas 140 respondentes de uma pesquisa).

O SR proposto neste trabalho pode desempenhar um papel significativo no incentivo e aprimoramento da prática de atividades físicas, principalmente no que tange aos seguintes aspectos:

- I. Personalização das orientações: O SR pode oferecer orientações personalizadas e adequadas às necessidades específicas de cada pessoa levando em consideração as características individuais, descritas pelos dados socioeconômicos e validados por profissionais de saúde;
- II. Motivação e engajamento: ao fornecer feedback contínuo, o SR pode manter os praticantes de atividades físicas engajados e motivados. Isso pode ser feito por meio de lembretes, reconhecimentos por conquistas e estabelecimento de metas alcançáveis;
- III. Variedade e Diversidade: o SR pode apresentar uma ampla variedade de atividades físicas. Isso ajuda o usuário descobrir novas formas de se exercitar, evitando a monotonia e aumentando o interesse na prática regular;
- IV. Integração com Dispositivos e Aplicativos: o SR pode ser integrado a aplicativos de rastreamento de atividades físicas e outras tecnologias de saúde, permitindo que os usuários tenham um registro mais detalhado de suas atividades.

Isso corrobora o estudo de Nogueira (2022), o qual relata que os SR podem ser muito úteis em inúmeras aplicações na área de saúde. Contudo, apesar do potencial do SR proposto para orientar e poder ser usado por indivíduos na prática de atividades físicas, contribuindo para um estilo de vida mais saudável e ativo, trata-se de um protótipo que ainda precisa ser concluído e avaliado por profissionais de saúde.

Cabe destacar que o SR poderia ter sua efetividade incrementada ao considerar outras informações do usuário como nível de aptidão física, objetivos e histórico de saúde. A conclusão e avaliação do SR por especialistas na área de educação física, bem como o incremento da sua efetividade constituem tarefas a serem conduzidas em trabalhos futuros.

Por fim, o *framework* de IA proposto representa uma contribuição significativa tendo em vista a forma inovadora de oferecer uma solução, não apenas para o problema investigado, mas também para uma variedade de problemas nos campos de saúde, educação, finanças entre outros. Não obstante, a ferramenta proposta pode trazer contribuições para o cumprimento

dos ODS estabelecidos na agenda 2030 da Organização das Nações Unidas (ONU), especialmente o ODS 3 (Saúde e Bem-Estar), ao promover hábitos de vida ativos visando reduzir doenças relacionadas ao sedentarismo.

## 5 CONCLUSÕES E PROPOSTA PARA TRABALHOS FUTUROS

O uso de técnicas de MD permitiu produzir regras que representam novos conhecimentos acerca da base de dados considerada nesta pesquisa. Esses conhecimentos representam descobertas de padrões que possibilitaram a implementação de um MIF capaz de determinar fatores motivacionais com maior grau de interesse para a prática de atividades físicas, de acordo com o perfil socioeconômico de uma pessoa. A vantagem do MIF em relação aos classificadores comumente empregados nos trabalhos da literatura é a possibilidade de obtenção de inferências a partir de informações imprecisas.

Sistemas de recomendação como o proposto neste trabalho podem auxiliar profissionais de saúde no oferecimento de orientações mais precisas, adaptadas às preferências, condições médicas e objetivos em atividades físicas. Essas recomendações personalizadas melhoram a adesão dos usuários, monitorando seu progresso ao longo do tempo e fazendo ajustes conforme necessário. Isso capacita os profissionais de saúde na promoção de um estilo de vida ativo de maneira segura e eficaz, contribuindo para a saúde geral dos usuários. Contudo, é importante mencionar que foi desenvolvido apenas um protótipo que apresenta limitações, e que ainda precisa ser concluído e avaliado por profissionais de saúde.

O uso da IA na compreensão e aprimoramento de atividades físicas desempenha um papel fundamental na concretização dos Objetivos de Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas (ONU). Ao motivar os indivíduos a adotarem estilos de vida mais saudáveis por meio de orientações personalizadas, a ferramenta apresentada neste trabalho pode contribuir para o ODS 3 (Saúde e Bem-Estar), no que tange a promoção de hábitos de vida ativos e redução de doenças relacionadas ao sedentarismo. Além disso, ao utilizar tecnologia de ponta para engajar a população em atividades físicas, essa ferramenta contribui para o ODS 9 (Indústria, Inovação e Infraestrutura) ao demonstrar a aplicação prática da tecnologia para melhorar a saúde. Ao incentivar práticas sustentáveis de exercício e fomentar a conscientização sobre a importância do bem-estar físico, essa ferramenta se alinha também ao ODS 11 (Cidades e Comunidades Sustentáveis) e ao ODS 17

(Parcerias e Meios de Implementação), já que pode ser usada para promover comunidades saudáveis e colaboração entre setores para alcançar objetivos de saúde e desenvolvimento duradouros.

No que tange as contribuições da pesquisa, os resultados obtidos a partir da investigação da literatura acerca da aplicabilidade e limitações das abordagens para a identificação de fatores que motivam a prática de atividades físicas podem ser citados como uma contribuição teórica. O *framework* proposto, que consiste em um encadeamento de técnicas de IA, pode ser considerado a principal contribuição científica desta pesquisa tendo em vista que apesar de algumas limitações nos resultados produzidos, em virtude da limitação da base de dados empregada, ele é aplicável tanto para solução de problemas correlatos ao aqui investigado quanto problemas de outras áreas do conhecimento. Não obstante, o SR constitui um artefato de software que representa uma contribuição tecnológica visto que, quando concluído, poderá ser registrado junto ao INPI (Instituto Nacional da Propriedade Industrial).

Em trabalhos futuros pretende-se empregar uma base mais ampla de dados a fim de melhorar a consistência dos padrões encontrados, melhorando o desempenho do MIF e consequentemente do SR. Pretende-se ainda concluir a implementação do SR, submetê-lo a avaliação de profissionais de saúde, especialmente os do campo de educação física, além de disponibilizá-lo para utilização no campo prático.

## REFERÊNCIAS

- AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules. In: **Proc. 20th int. conf. very large data bases, VLDB**, vol. 1215, p. 487- 499. 1994.
- ANDERSEN V.; PEDERSEN H.; FIMLAND M. S.; SHAW M.; SOLSTAD TEJ; STIEN N.; TOLDNES K.; SAETERBAKKEN A. H. Comparison of Muscle Activity in Three Single-Joint, Hip Extension Exercises in Resistance-Trained Women. **The Journal of Sports Science and Medicine, JSSM**, p. 181-187, 2021.
- ANDRADE, F. B.; FERREIRA, C. S. **Desigualdades socioeconômicas associadas ao excesso de peso e sedentarismo em adolescentes brasileiros**. SciELO Public Health, 2021.
- ANSHEL, M.; FOGARTY, G.; GROVE, R.; JACKSON, S.; MCKENZIE, A.; TENENBAUM, G.; WEINBERG, R. **Motivation for youth participation in sport and physical activity: relationships to culture, self-reported activity levels, and gender**. International Journal of Sport Psychology, 2000
- ARAÚJO, A. S.; BISSACO, A.S. M.; DELGADO, S.; DOMINGUES, M. A.; Computational Intelligence in Serious Games: A Case Study to Identify Patterns in a Game for Children with Learning Disabilities. **International Journal of Computer Applications**, v. 184, p. 40-44, 2022.
- ARCAS, M. M.; GRANIZO, I. R.; IBÁÑEZ, E. M.; JIMÉNEZ, J. L. U.; ORTEGA, F. Z.; PIPÓ, J. C.; VALERO, G. G. **Physical Self-Concept Changes in Adults and Older Adults: Influence of Emotional Intelligence, Intrinsic Motivation and Sports Habits**. International Journal of Environmental Research and Public Health, DOI: <https://doi.org/10.3390/ijerph18041711>, 2021
- ASWANI A.; FUKUOKA Y.; HOOPER J.; LINDGREN TG.; MINTZ TD. **Applying Natural Language Processing to Understand Motivational Profiles for Maintaining Physical Activity After a Mobile App and Accelerometer-Based Intervention: The mPED Randomized Controlled Trial**. JMIR Mhealth Uhealth 6(6): e10042, 2018.

BALABANOVIC', M; SHOHAM, Y. **Fab: content-based, collaborative recommendation.** *Communications of the ACM*, ACM New York, NY, USA, v. 40, n. 3, p. 66-72, 1997.

BALL K.; BROWN H.; CHAPPEL S.; KOORTS H.; LAI SK.; LEWIS M.; MACFARLANE S.; RIDGERS ND.; SALMON J.; TIMPERIO A. **Translatability of a Wearable Technology Intervention to Increase Adolescent Physical Activity: Mixed Methods Implementation Evaluation.** *J Med Internet Res* 22(8): e13573, 2020.

BAKER S.; DIVINE A.; HALL C. M.; WATSON M. **Facebook, relatedness and exercise motivation in university students: A mixed methods investigation.** Elsevier. 19, p. 138-150. <https://doi.org/10.1016/j.chb.2018.09.037>, 2019.

BALBINOTTI, A. A. M. **Inventário de Motivação à Prática Regular de Atividade Física (IMPRAF-126).** Laboratório de Psicologia do Esporte – Universidade Federal do Rio Grande do Sul. Porto Alegre, 2004.

BARBOSA, M. L. L. **Propriedades Métricas do Inventário de Motivação à Prática Regular de Atividade Física (Impraf-126).** p. 151. Dissertação (Mestrado em Ciência do Movimento Humano) - Escola de Educação Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

BHAVE, D. P.; TEO, L. H.; & DALAL, R. S. Privacy at work: A review and a research agenda for a contested terrain. **Journal of Management**, 46(1), 127-164. <https://doi.org/10.1177/0149206319878254>, 2020.

BRAMER, M. **Principles of data mining.** Springer, London, 2007.

BRIÈRE, N.; VALLERAND, R.; BLAIS, M.; PELLETIER, L. Developpement et Validation d'une Mesure de Motivation Intrinséque, Extrinséque et d'Amotivation en Contexte Sportif: l'échelle de motivation dam les sports. **International Journal of Sport Psychology**, v.26 n.4 p.465-489, 1995.

CAVALCANTE, R. C.; SOBRINHO, G. F. L. **Computational intelligence in the financial market: a review of techniques for automating operations.** Research, Society and Development. DOI: 10.33448/rsd-v12i5.41793, 2023

CHEN, Q.; LI, G.; HAN, B.; HENG, C.; LEONG, T. **Coronary Artery Disease Prediction with Bayesian Networks and Constraint Elicitation.** The National University of Singapore, 2006.

CORREIA, L.; LOPES, D.; PORTO, J.V.; LACERDA, Y.F.; CORREIA, V.C.A.; BAGANO, G.O.; PONTES, B.S.B.; MELO, M. H.V.; SILVA, T.E.A.; MEIRELES, A. C. **Validação de um Algoritmo de Inteligência Artificial para a Predição Diagnóstica de Doença Coronariana: Comparação com um Modelo Estatístico Tradicional.** Arq. Bras. Cardiol., v. 117, n. 6, p. 1061-1070, 2021.

DEB, K. **Multi-Objective Optimization Using Evolutionary Algorithms.** Wiley, 2009.

DECI, E. L.; RYAN, R. M. **Intrinsic Motivation and self-determination in human behavior.** New York: Plenum, 1985.

DECI, E. L.; RYAN, R. M. **Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being.** American Psychologist. 2000.

DECI, E. L.; OLAFSEN, A. H.; RYAN, R. M. **Self-Determination Theory in Work Organizations: The State of a Science.** Annual Reviews. DOI: <https://doi.org/10.1146/annurev-orgpsych-032516-113108>, 2017.

DECI, E. L.; DOESTNER, R.; RYAN, R. M. **A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation.** American Psychological Association. DOI: <https://psycnet.apa.org/doi/10.1037/0033-2909.125.6.627>, 1999.

DECI, E. L.; REEVE, J. **Elements of the competitive situation that affect intrinsic motivation.** Personality and Social Psychology Bulletin. DOI: <https://doi.org/10.1177/0146167296221003>, 1996.

DELAVANDE, A.; HURD, M. D.; LANGA, K. M.; MARTORELL, P.; MULLEN, K. J. Monetary costs of dementia in the United States. **New England Journal of Medicine**, 368(14), p. 1326-1334, 2013.

EMONSON C.; FARMER E.; FUELSCHER I.; HYDE C.; MCGILLIVRAY J.; OLIVE L.; PAPADOPOULOS N.; PESCE C.; RINEHART N. **A Preliminary**

**Investigation of the Relationship between Motivation for Physical Activity and Emotional and Behavioural Difficulties in Children Aged 8-12 Years: The Role of Autonomous Motivation.** Environmental Research and Public Health. 17(15), 5584; <https://doi.org/10.3390/ijerph17155584>, 2020.

FALK, K. **Practical recommender systems.** [S.I.]: Manning Publications, 2019.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; & UTHURUSAMY, R. **Advances in knowledge discovery and data mining. American Association for Artificial Intelligence.** Menlo Park: MIT Press, 1996.

FERREIRA, P. L. **Estatística descritiva e inferencial: breves notas**, 2005. Disponível em: <https://baes.uc.pt/bitstream/10316/9961/1/AP200501.pdf>. Acesso em 20 Abr 2023.

FOLAJIMI, Y. O.; ISINKAYE, F. O.; OJOKOH, B. A. **Recommendation systems: Principles, methods and evaluation.** Egyptian Informatics Journal. DOI: <https://doi.org/10.1016/j.eij.2015.06.005>, 2015

FORTIER, M. S.; GUAY, F.; VALLERAND, R. R.J.; **Self-determination and persistence in a real-life setting: toward a motivational model of high school dropout.** Journal of Personality and Social psychology. DOI: <https://psycnet.apa.org/doi/10.1037/0022-3514.72.5.1161>, 1997

GANTI, V.; GEHRKE, J.; Ramakrishnan, R. **RainForest - A Framework for Fast Decision Tree Construction of Large Datasets.** Data Mining and Knowledge Discovery, DOI: 10.1023/A:1009839829793, 2000.

GERBER, M.; GROB, A.; MEYER, S. **No fun, no gain: The stress-buffering effect of physical activity on life satisfaction depends on adolescents' intrinsic motivation.** Elsevier, DOI: <https://doi.org/10.1016/j.psychsport.2021.102004>, 2021.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization, and Machine Learning.** Addison-Wesley Professional, 1989.

GOLDBERG, D.; NICHOLS, D.; OKI, B. M.; TERRY, D. **Using collaborative filtering to weave an information tapestry.** Communications of the ACM, ACM New York, NY, USA, v. 35, n. 12, p. 61-70, 1992.

GOLDSCHMIDT, R.; PASSOS, E. **Data Mining um Guia Prático. Conceitos, Técnicas, Ferramentas, Orientações e Aplicações.** Ed. Campus, Rio de Janeiro, 2005.

JUNG, J.; SONG, H.; KIM, Y.; IM, H.; OH, S. Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. **Computers in Human Behavior**, 71, 291–298. 2017.

KHOBragade, A.; N., MALIK, L.; RAGHUWANSHI, M. M. Evaluating Kernel Effect on Performance of SVM Classification using Satellite Images. **International Journal of Scientific & Engineering Research**, v. 7, Issue 3, 2016.

LIMA, G. S.; PAPINI, C. B.; SILVA, D. B.; SOUZA, M. V. C. **Motivos e barreiras para a prática de exercício físico em idosos frequentadores de academias.** Arquivos de Ciências do Esporte. Universidade Federal do Triângulo Mineiro, UFTM, Uberaba, Minas Gerais, Brasil. <https://seer.ufmt.edu.br/revistaelectronica/index.php/aces/article/view/3264>, 2020.

MARRO, A. A.; SOUZA, A. D. C.; CAVALCANTE, E. D. S.; BEZERRA, G. S.; NUNES, R. O. Lógica fuzzy: conceitos e aplicações. **Natal: Universidade Federal do Rio Grande do Norte (UFRN)**, p. 2, 2010.

MCCULLOCH, W.; PITTS, W. **A Logical Calculus of the Ideas Immanent in Nervous Activity**, The Bulletin of Mathematical Biophysics, v. 5, n. 4, p. 115-133, 1943.

MCHUGH, M. L. Interrater reliability: the kappa statistic. **Biochemia medica**, v. 22, n. 3, p. 276-282, 2012.

MITCHELL, T. M. **Machine Learning**. New York: McGraw-Hill, 1997.

MS – Ministério da Saúde, Secretaria de Atenção Primária à Saúde. Com investimento de R\$ 100 milhões, ação do Ministério da Saúde estimula atividade física em 5 mil municípios. 16 maio 2022. Disponível em: <https://aps.saude.gov.br/noticia/17256>. Acesso em: 30/06/23.

MURTHY, S. K. **Automatic Construction of Decision Trees from Data: a Multi-Disciplinary Survey**, Data Mining and Knowledge Discovery, v. 2, n. 4, p. 345-389, 1998.

NOGUEIRA, A. O uso da inteligência artificial como ferramenta de apoio à gestão das ações em saúde na Secretaria de Estado da Saúde de Goiás. **REVISTA CIENTÍFICA DA ESCOLA ESTADUAL DE SAÚDE PÚBLICA DE GOIÁS "CÂNDIDO SANTIAGO"**, v. 8, p. 1-15, 2022.

NUNES, S. Impacto Econômico da Inatividade Física. 12 set 2022. Disponível em: <https://asemananews.com.br/2022/09/12/impacto-economico-da-inatividade-fisica/>. Acesso em: 30/06/23.

NUNES, C. C. G.; CHAVES, C. M. C. B.; DUARTE, J. C. Motivação para a prática de atividade física em estudantes de enfermagem. **Revista de Enfermagem Referência**, v. 6, n. 1, 2022.

FRAGA, A. B.; OLIVEIRA, B. N. **Uso das tecnologias digitais para a prática de exercícios físicos: uma revisão integrativa**. Conexões: Educ. Física Esporte e Saúde, Campinas, v. 18, p. 1-19, 2020.

OMS - Organização Mundial da Saúde **Global status report on physical activity 2022** - Disponível em: <https://www.who.int/publications/i/item/9789240059153>. Acesso em: 30/06/23.

OPAS – Organização Pan-Americana da Saúde. Mais de 1,4 bilhão de adultos correm risco de desenvolver doenças por inatividade física em todo o mundo. 5 set 2018. Disponível em: <https://www.paho.org/pt/noticias/5-9-2018-mais-14-bilhao-adultos-correm-risco-desenvolver-doencas-por-inatividade-fisica-em>. Acesso em: 30/06/23.

PATRI, R., SURESH, M. **Agility assessment using fuzzy logic approach: a case of healthcare dispensary**. BMC Health Services Research, DOI 10.1186/s12913-017-2332-y, 2017

PAZZANI, M.; BILLSUS, D. **Learning and revising user profiles: The identification of interesting web sites.** Machine learning, Springer, v. 27, n. 3, p. 313-331, 1997.

PEREIRA, R. C. **Aplicativos para controle alimentar e prática de exercícios físicos utilizando as tecnologias de informação e comunicação e interação humano-computador.** Monografia apresentada ao curso de Bacharelado em Sistemas e Mídias Digitais do Instituto Universidade Virtual da Universidade Federal do Ceará. UFC, 2019.

PEDRYCZ, W. **Why triangular membership functions?** Fuzzy Sets and Systems, v. 64, p. 21-30, 1994.

PETHERICK, C. M.; WEIGAND, D. A. **The relationship of dispositional goal orientations and perceived motivational climates on indices of motivation in male and female swimmers.** International Journal of Sport Psychology, 2002

PILAY, M. A. T.; CASTRO, M. I. R.; CASTRO, R. M. R.; CASTRO, V. F. R. **La inteligencia artificial en la Educación Física en tiempo de COVID 19.** Horizontes, 2021.

PIO, F. P. B. **Avaliação da técnica de modelo linear de mistura espectral como subsídio à classificação do uso e ocupação do solo.** UNIVERSIDADE FEDERAL DE MINAS GERAIS. DOI: <http://hdl.handle.net/1843/54383>, 2023.

QUINLAN, J. R. **C4.5: programs for machine learning.** Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.

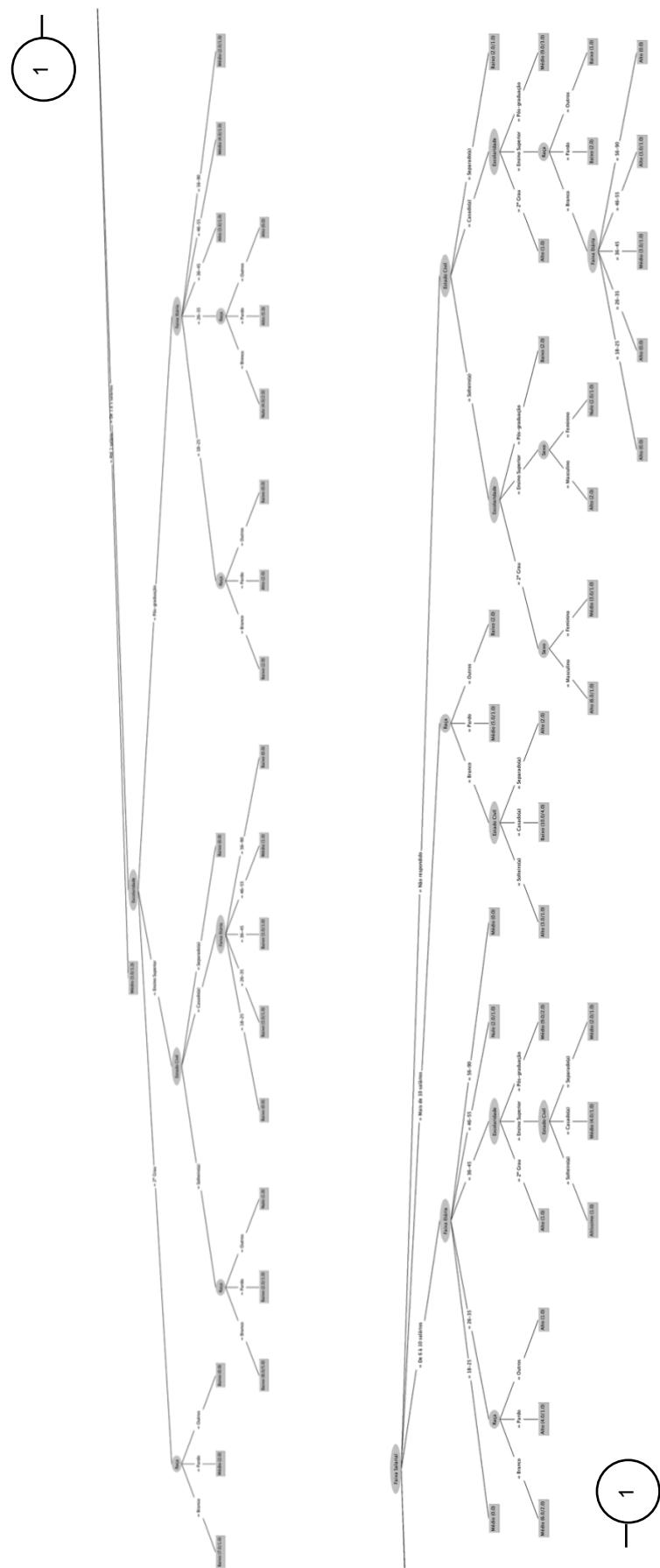
QUINLAN, J. R. **Induction of decision trees.** Machine Learning, v. 1, p. 81-106, 1986.

ROSENBLATT, F. **The perceptron: A probabilistic model for information storage and organization in the brain.** Psychological Review, v. 65, n. 6, p. 386-408, 1958.

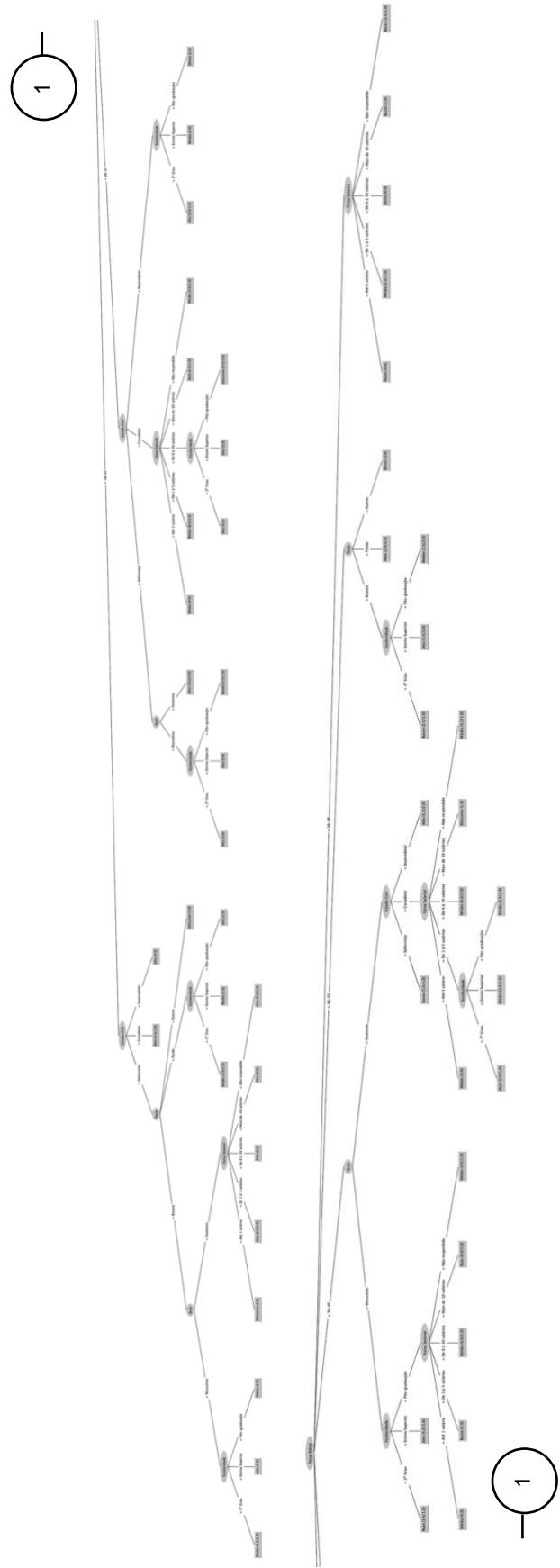
RESNICK, P.; VARIAN, H. R. **Recommender systems.** Communications of the ACM, ACM New York, NY, USA, v. 40, n. 3, p. 56-58, 1997.

- RUMELHART, D. E.; MCCLELLAND, J. **Parallel Distributed Processing**, v. 1, MIT Press, 1986.
- SHETTIGAR S.; SHIVARAJ K. A **Study to Assess the Factors Affecting Adherence to Exercise in the Indian Population**. Cureus 6062, 2019.
- SONG, X. Discussão sobre a aplicação da tecnologia de mineração de dados na gestão do desempenho esportivo. **Revista Brasileira de Medicina do Esporte**, v. 28, p. 460-464, 2022.
- SULINO, R. M. **Mineração de dados e descoberta de conhecimento a partir de dados de atividade física, saúde, clima e qualidade de vida**, 2020.
- UNIVERSITY OF WAIKATO. WEKA 3 - **Machine Learning Software in Java**, 2010.
- WELTER, F. H.; SZINVELSKI, C. A.; BORGES, R. M. **A utilização da inteligência artificial no campo da educação física**. Universidade Nacional de Ijuí, 2020.
- WITTEN, I.; FRANK, E. **Data Mining: practical machine learning toolsand techniques**. 2th. ed. USA: Elsevier, 2005.
- ZADEH, L. A. **Outline of a New Approach to the Analysis of Complex Systems and Decision Processes**, IEEE Transactions on System, Man and Cybernetics, v. 3, p. 2844, 1973.

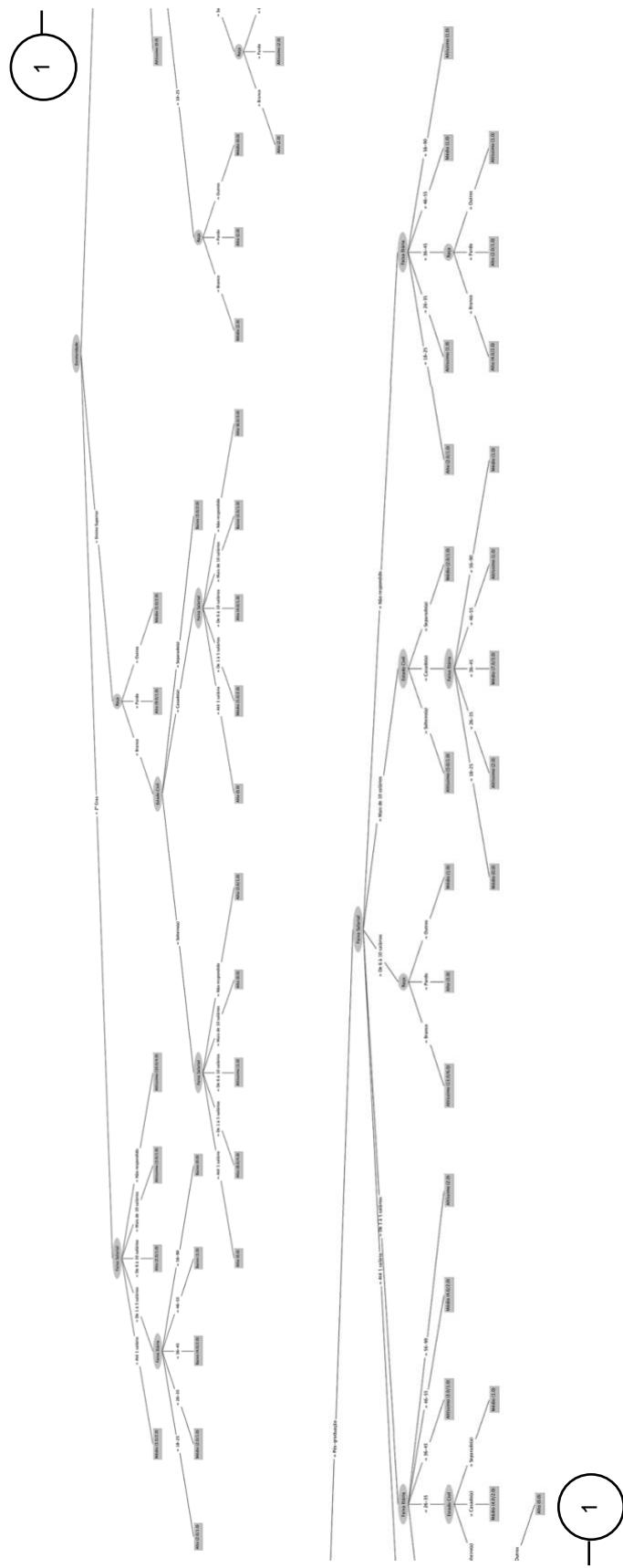
## **APÊNDICE I – AD relativa ao fator motivacional CONDICIONAMENTO FÍSICO**



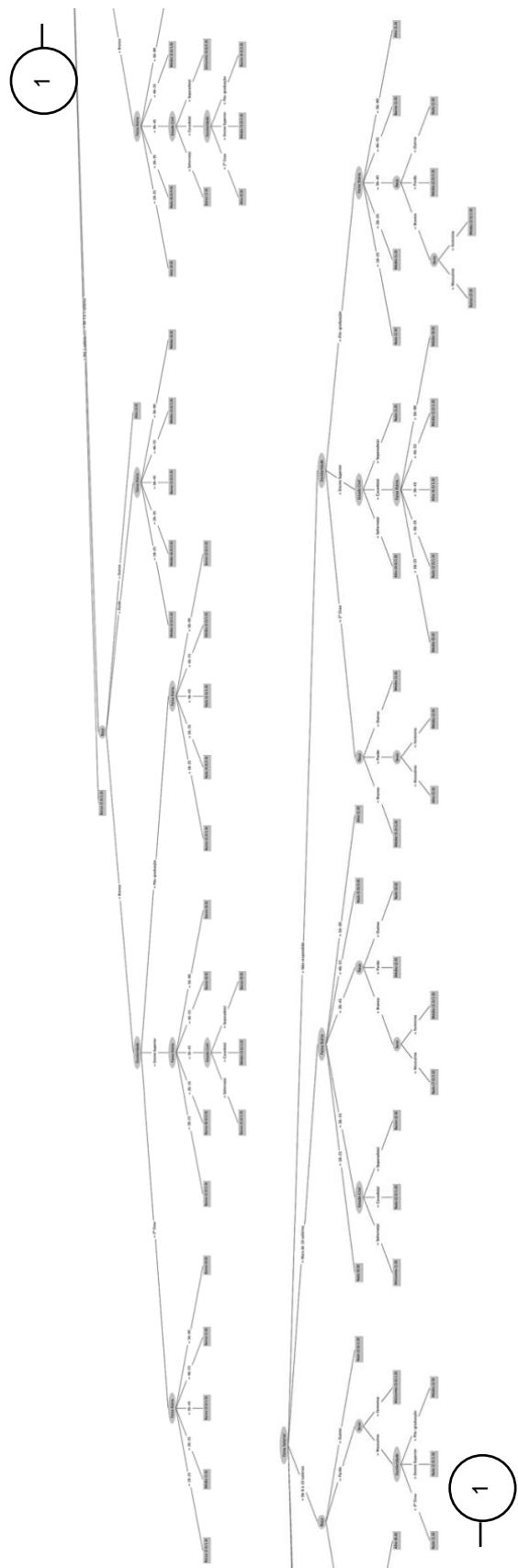
## **APÊNDICE II – AD relativa ao fator motivacional ESTÉTICA**

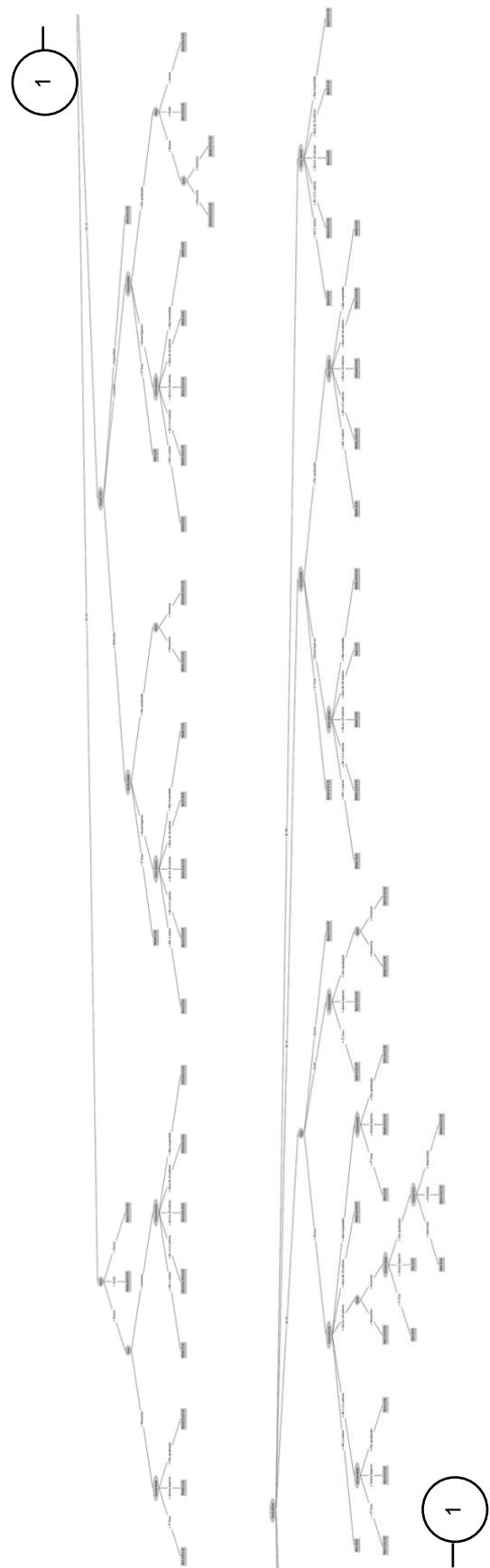


**APÊNDICE III – AD relativa ao fator motivacional SAÚDE**

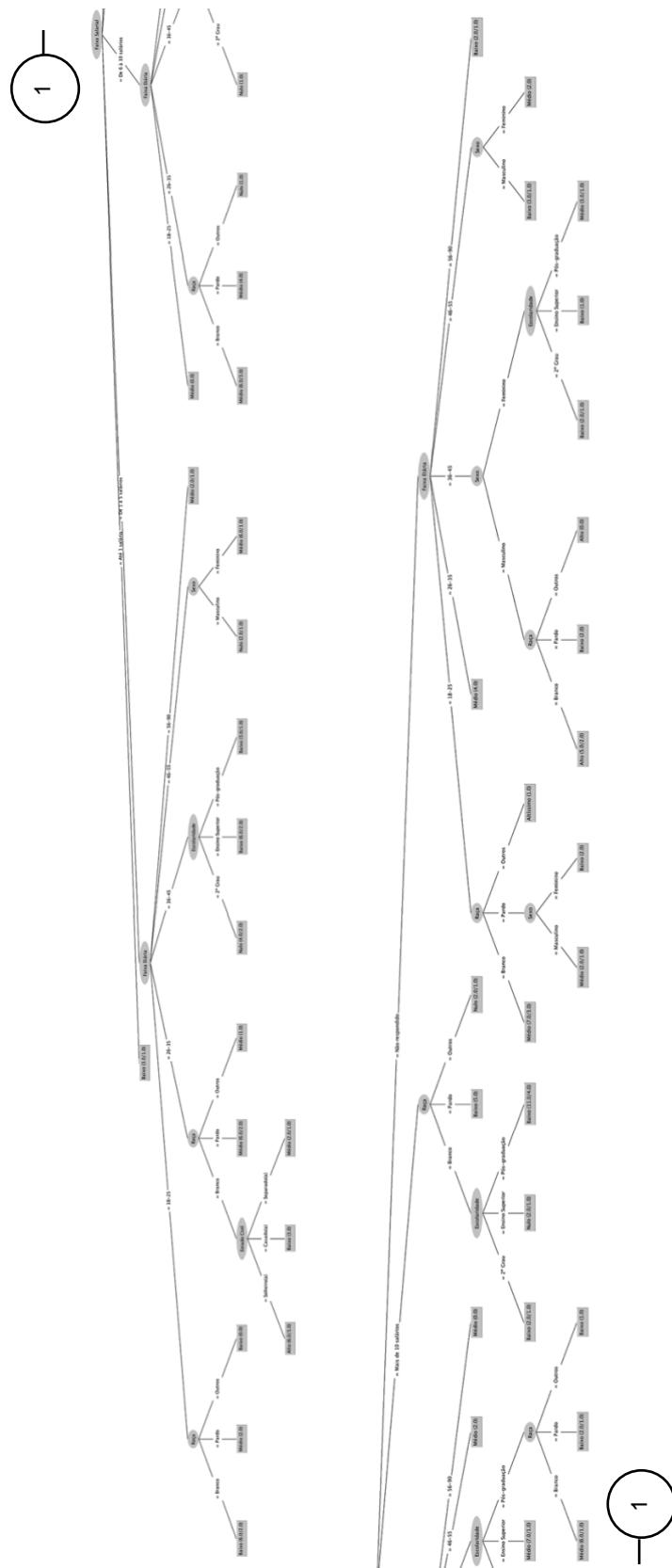


**APÊNDICE IV – AD relativa ao fator motivacional INTEGRAÇÃO SOCIAL**

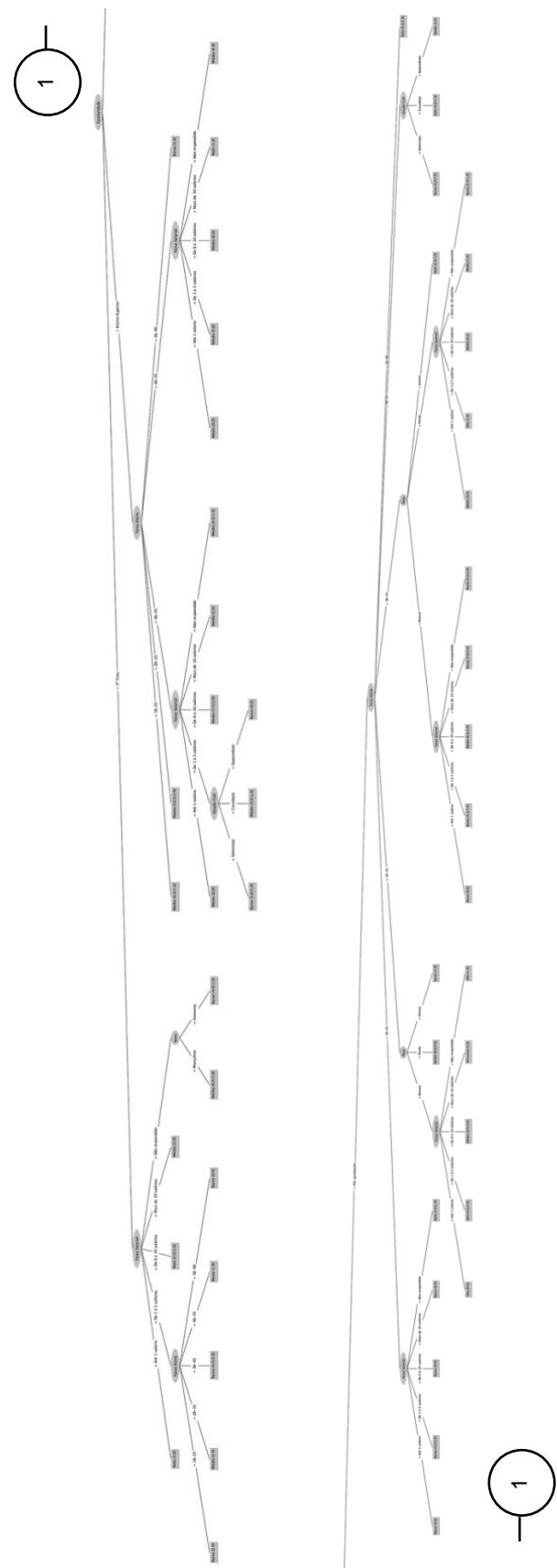


**APÊNDICE V – AD relativa ao fator motivacional ANSIEDADE/ESTRESSE**

## **APÊNDICE VI – AD relativa ao fator motivacional DISPONIBILIDADE DE TEMPO**



## **APÊNDICE VII – AD relativa ao fator motivacional CUSTO**



**ANEXO A – “Pesquisa Motivacional para Prática de Atividades Físicas”****1. Deseja se identificar?**

As primeiras 200 respostas identificadas com email e nome, receberão acesso grátis ao App de treinos da S2Pro por 1 ano! (O uso do aplicativo será disponibilizado após o término da coleta de dados)

 Sim Não**2. E-mail***Descrição***3. Nome***Descrição*

T

**4. Faixa Etária \****Descrição* 18 a 20 anos 21 a 25 anos 26 a 30 anos 31 a 35 anos 36 a 40 anos 41 a 45 anos 46 a 50 anos 51 a 55 anos 56 a 60 anos acima de 60 anos

**5. Sexo \****Descrição* Masculino Feminino**6. Como você se declara quanto à Raça ou Etnia? \****Descrição* Branco Afrodescendente Pardo Amarelo Indígena**7. Estado Civil \****Descrição*

Por favor selecione

**8. Escolaridade \****Descrição*

Por favor selecione

**9. Profissão \****Descrição*

T

**10. Faixa Salarial***Descrição*

- Não estou trabalhando
- Até 1 salário mínimo (R\$ 1.212,00)
- Mais que 1 à 2 salários mínimos (R\$ 1.212,01 a R\$ 2.424,00)
- Mais que 2 à 5 salários mínimos (R\$ 2.424,01 a R\$ 6.060,00)
- Mais que 5 à 8 salários mínimos (R\$ 6.060,01 a R\$ 9.696,00)
- Mais que 8 à 10 salários mínimos (R\$ 9.696,01 a R\$ 12.120,00)
- Mais de 10 salários mínimos (acima de R\$ 12.120,01)
- Prefiro não responder

**11. Estado***Descrição*

- SP x

**12. Cidade***Descrição*

- Por favor selecione ▼

**40.** Como você classifica o seu grau de MOTIVAÇÃO para praticar atividades físicas, com relação a cada um dos fatores abaixo?\*

*Condicionamento Físico (Ex.: competir com os outros; ser o melhor no esporte; vencer competições) | Estética (Ex.: ficar com o corpo mais bonito, sentir-se bonito(a), manter o corpo em forma) | Saúde (Ex.: adquirir saúde; ficar livre de doenças; viver mais) | Integração Social (Ex.: encontrar amigos; estar com outras pessoas; fazer novos amigos) | Ansiedade/Estresse (Ex.: diminuição de irritação; diminuir ansiedade; diminuir estresse) | Tempo (Ex.: tenho tempo livre; sou muito dinâmico com meu tempo) | Custo (Ex.: acho preço da academia baixo pelos benefícios que obtenho em saúde; a empresa paga parte ou valor total da minha academia; pratico atividade em uma academia em que eu ou um membro da minha família é sócio)*

	Nulo	Baixo	Médio	Alto	Muito Alto
Condicionamento Físico	<input type="radio"/>				
Estética	<input type="radio"/>				
Saúde	<input type="radio"/>				
Integração Social	<input type="radio"/>				
Ansiedade/Estresse	<input type="radio"/>				
Disponibilidade de Tempo	<input type="radio"/>				
Custo	<input type="radio"/>				

**41.** Existem outros fatores que te MOTIVAM a praticar atividades físicas?

*Escreva abaixo os fatores separados por vírgula*

T

**42.** Como você classifica o seu grau de DESMOTIVAÇÃO para praticar atividades físicas, com relação a cada um dos fatores abaixo?\*

*Condicionamento Físico (Ex.: sinto fraqueza muscular, fadiga, cansaço) | Estética (Ex.: não me sinto confortável com minha aparência física; sinto-me velho demais) | Saúde (Ex.: limitação por doença, dor, lesão, falta de ar) | Integração Social (Ex.: sou muito reservado/tímido; tenho dificuldade em fazer novas amizades) | Ansiedade/Estresse (Ex.: Sinto-me ansioso/estressado; tenho preguiça, falta de disposição/falta de vontade) | Tempo (Ex.: falta de tempo; trabalho muito; tenho agenda desorganizada) | Custo (Ex.: falta de dinheiro; acho o preço da academia alto, não cabe no meu orçamento)*

	Nulo	Baixo	Médio	Alto	Muito Alto
Condicionamento Físico	<input type="radio"/>				
Estética	<input type="radio"/>				
Saúde	<input type="radio"/>				
Integração Social	<input type="radio"/>				
Ansiedade/Estresse	<input type="radio"/>				
Disponibilidade de Tempo	<input type="radio"/>				
Custo	<input type="radio"/>				

**43.** Existem outros fatores que te DESMOTIVAM a praticar atividades físicas?

*Escreva abaixo os fatores separados por vírgula*

T