

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE EDUCAÇÃO FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM EDUCAÇÃO
FÍSICA**

SAMUEL ESTEVAM VIDAL

**O USO DE APRENDIZADO DE MÁQUINA NO AUXÍLIO DA
CLASSIFICAÇÃO DA PROGRESSÃO DA DOENÇA DE
PARKINSON POR MEIO DE TESTES FUNCIONAIS**

BRASÍLIA
2021

SAMUEL ESTEVAM VIDAL

**O USO DE APRENDIZADO DE MÁQUINA NO AUXÍLIO DA
CLASSIFICAÇÃO DA PROGRESSÃO DA DOENÇA DE PARKINSON POR
MEIO DE TESTES FUNCIONAIS**

Orientador:

Prof. Dr Ricardo Jacó de Oliveira

Tese apresentada ao Programa de Pós-graduação *Stricto Sensu* em Educação Física – FEF/UnB como requisito para obtenção do título de Doutor em Educação Física.

Para o Vini, meu esposo.

AGRADECIMENTOS

Diferente de como aconteceu no mestrado, eu não tenho certeza de como estruturar os agradecimentos. A sensação presente é que finalmente chega ao final uma Jornada de formação, que se deu em aspectos muito maiores e bem mais profundos do que a própria ciência.

Este trabalho traz a memória e o registro de um Programa desenvolvido na Universidade de Brasília; uma construção a muitas mãos, às quais eu reconheço e agradeço em nome do professor Ricardo Jacó de Oliveira, nosso orientador e coordenador do programa.

Agradeço ao professor Cristiano Miosso por sua generosidade durante as etapas de construção e teste dos classificadores. Graças a tecnologia, nos momentos de telas compartilhadas eu pude assistir à sua maestria conduzindo a programação utilizada nesta tese. Estas serão, sem dúvidas, algumas das minhas memórias mais queridas desta época.

Agradeço ao UNICEPLAC, instituição que me acolheu profissionalmente e que pude contribuir durante os anos do doutorado. Em especial, agradeço ao curso de educação física, que alicerçou em mim o sentido do esforço para levar Educação de qualidade para todas as pessoas. Agradeço em especial ao professor Igor Fernandes Cunha, apoio fundamental em todas as etapas, desde que chegamos juntos à IES.

Por muitas vezes eu acreditei que a conclusão deste trabalho não seria possível... E foi me apoiando em amigos e amigas que consegui chegar até aqui. Assim, agradeço ao professor Jake do Carmo, um amigo que a Academia me trouxe, por ser peça chave na articulação deste trabalho.

Às amigadas da caminhada, todo o meu amor e gratidão! Pandinha e Carol, obrigado por tanto, por sempre. Cris, Lucília e Gabi, obrigado por tanto esforço coletivo dedicado a mim! Agradeço à irmandade do CELF, pelo suporte espiritual e por me ensinar sobre firmeza, amor, dedicação e Vontade de fazer o bem.

Agradeço a Deus, esta Luz Divina que a todos vem clarear. A maior conquista na formação acadêmica foi minha rendição à esta espiritualidade organizada, assertiva e sutilmente científica. O potencial da ciência de trazer luz e conhecimento pode ser uma forma coerente de salvar toda a humanidade.

À minha família, o maior agradecimento que posso fazer é retornar de onde saí para trilhar este caminho. É chegado o tempo de desfrutar desse chão, dessa casa, dessa

construção agora feita! Amo profundamente cada um de vocês. E sou grato por ter chegado até aqui.

Por fim, o agradecimento mais especial é para meu esposo Vini. Obrigado por sua doçura e por sua firmeza. Pelo seu amor que se provou incondicional. Eu não imaginava os rumos que a minha vida tomaria, e hoje sou feliz e grato ao próprio Caminho, por ter te encontrado. Vida longa ao nosso amor, meu amor.

Queremos saber
O que vão fazer
Com as novas invenções
Queremos notícia mais séria
Sobre a descoberta da antimatéria
E suas implicações
Na emancipação do homem (e das mulheres!)
Das grandes populações
Homens pobres das cidades
Das estepes dos sertões
Queremos saber
Quando vamos ter
Raio laser mais barato
Queremos, de fato, um relato
Retrato mais sério do Mistério da Luz
Luz do disco voador
Pra iluminação do homem
Tão carente e sofredor
Tão perdido na distância
Da morada do Senhor
Queremos saber
Queremos viver
Confiantes no futuro
Por isso se faz necessário prever
Qual o itinerário da Ilusão
A ilusão do poder
Pois se foi permitido ao homem
Tantas coisas conhecer
É melhor que todos saibam
O que pode acontecer...

Queremos saber, todos queremos saber
(EnCantada na voz de Cássia Eller)

RESUMO

A doença de Parkinson (DP) é um distúrbio motor com alta prevalência entre pessoas idosas; afeta aspectos cognitivos, sociais e psicológicos, além dos prejuízos físico-motores inerentes da doença (POEWE, 2017). Tanto o diagnóstico quanto o monitoramento da progressão da doença se dão através da combinação de exames motores que quantificam o acometimento da DP (BRASIL, 2010). Neste sentido, os testes funcionais vêm sendo amplamente aplicados para auxiliar no tratamento administrado (OPARA, 2017), bem como o uso de tecnologias avançadas, como é o caso do aprendizado de máquina (ZHAN, 2018). Assim, o objetivo desta tese foi verificar a possibilidade do auxílio na predição de classificação da doença de Parkinson a partir de aspectos funcionais de pessoas acometidas, utilizando técnicas do aprendizado de máquina. Foram considerados os dados provenientes do Programa de Atividades Físicas para Pessoas com Doenças Neurodegenerativas (PAFPDN) da Faculdade de Educação Física da Universidade de Brasília. A amostra foi composta por 82 homens e 31 mulheres, com idade média de 64,80 (\pm 10,18) anos. Os aspectos funcionais foram quantificados através dos escores do Timed Up and Go Test (TUG), do 30s Chair-Stand Test (T30), do Ten Meters Walk Test (TMW), e da seção II da UPDRS. As características individuais analisadas foram idade, massa e estatura; e como aspectos clínicos, a seção III da UPDRS, além da escala de Hoehn e Yahr e do tempo de diagnóstico. Quatro classificadores de aprendizado supervisionado de máquina foram testados: SVM, SVM com data augmentation, RUSBoost e RUSBoost com data augmentation. Os aspectos funcionais, clínicos e as características individuais foram combinados e testados com diferentes valores de referência da UPDRS III para verificar a acurácia da máquina em classificar a progressão da doença. O TESTE 2 (TUG, TMW, T30 e Hohen e Yahr) e o TESTE 4 (TUG, TMW, T30, idade e tempo de diagnóstico) apresentaram a melhor avaliação de classificação considerando as métricas de Fawcett (2006), ambas utilizando a pontuação 15 na UPDRS III como valor de referência. Conclui-se, então, que é possível que aspectos funcionais possam auxiliar na predição da classificação da doença de Parkinson, utilizando técnicas do aprendizado de máquina.

ABSTRACT

Parkinson's disease (PD) is a motor disorder with a high prevalence among elderly people; it affects cognitive, social and psychological aspects, in addition to the physical-motor damage inherent to the disease (POEWE, 2017). Both diagnosis and monitoring of disease progression occur through a combination of motor tests that quantify the involvement of PD (BRASIL, 2010). In this sense, functional tests have been widely applied to assist in administered treatment (OPARA, 2017), as well as the use of advanced technologies, such as machine learning (ZHAN, 2018). Thus, the objective of this thesis was to verify the possibility of helping to predict the classification of Parkinson's disease from functional aspects of affected people, using machine learning techniques. Data from the Physical Activities Program for People with Neurodegenerative Diseases (PAFPDN) of the Faculty of Physical Education of the University of Brasília were considered. The sample consisted of 82 men and 31 women, with a mean age of 64.80 (\pm 10.18) years. Functional aspects were quantified through the scores of the Timed Up and Go Test (TUG), the 30s Chair-Stand Test (T30), the Ten Meters Walk Test (TMW), and section II of the UPDRS. The individual characteristics analyzed were age, mass, and height; and as clinical aspects, section III of the UPDRS, in addition to the Hoehn and Yahr scale and time of diagnosis. Four supervised machine learning classifiers were tested: SVM, SVM with data augmentation, RUSBoost and RUSBoost with data augmentation. Functional, clinical and individual characteristics were combined and tested with different UPDRS III reference values to verify the machine's accuracy in classifying disease progression. TEST 2 (TUG, TMW, T30 and Hohen and Yahr) and TEST 4 (TUG, TMW, T30, age and time since diagnosis) presented the best classification assessment considering the metrics of Fawcett (2006), both using the score 15 in UPDRS III as a reference value. We conclude, then, that it is possible that functional aspects can help predict the classification of Parkinson's disease, using machine learning techniques.

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 - Testes funcionais utilizados no estuúdo e suas respectivas variáveis para análise | 47 |
| Tabela 2 - Caracterização da amostra: gênero e classificação de Hoehn e Yahr. | 52 |
| Tabela 3 - Caracterização da amostra: valores antropométricos e tempo de diagnóstico. | 52 |
| Tabela 4 – TESTE 1:Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW, T30)..... | 53 |
| Tabela 5 – TESTE 2: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30) e aspectos clínicos (Hoehn e Yahr)..... | 54 |
| Tabela 6 – TESTE 3: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW, T30 e UPDRS II), aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade)..... | 55 |
| Tabela 7 – TESTE 4: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade) | 56 |
| Tabela 8 – TESTE 5: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade, massa e estatura) | 57 |
| Tabela 9 – TESTE 6: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico e Escala de Hoehn e Yahr) e características individuais (Idade, massa e estatura)..... | 58 |
| Tabela 10 – TESTE 7: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico e Escala de Hoehn e Yahr) e características individuais (Idade)..... | 59 |
| Tabela 11 – TESTE 8: Classificação da progressão da doença a partir de aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade, massa e estatura)..... | 60 |
| Tabela 12 – TESTE 9: Classificação da progressão da doença a partir de aspectos clínicos (Tempo de diagnóstico e Escala de Hoehn e Yahr) e características individuais (Idade, massa e estatura) | 61 |

LISTA DE SIGLAS E ABREVIACOES

| | |
|----------------|--|
| AM | Aprendizado de Mquina |
| BDI | Inventrio de Depresso de Beck |
| DP | Doena de Parkinson |
| FEF | Faculdade e Educao Fsica |
| FN | False Negatives |
| FP | False Positives |
| H&Y | Escala de Hoehn e Yahr |
| IPAQ | Questionrio Internacional de Atividade Fsica |
| MDS | Movement Disorders Society |
| MEEM | Mini Exame do Estado Mental |
| MMII | Membros inferiores |
| MMSS | Membros superiores |
| PAFPDN | Programa de Atividades Fsicas para Pessoas com Doenas Neurodegenerativas |
| PDQL | Parkinson`s Disease Quality of Life Questionnaire |
| PPGEF | Programa de Ps Graduao Stricto Sensu em Educao Fsica |
| SVM | Support Vector Machine |
| T30 | 30 seconds Chair-Stand Test |
| TMW | Ten Meters Walk Test |
| TN | True Negatives |
| TP | True Positives |
| TUG | Timed Up and Go Test |
| UnB | Universidade de Braslia |
| UPDRS | Unified Parkinson`s Disease Rating Scale |

SUMÁRIO

| | |
|---|-----------|
| RESUMO | 1 |
| ABSTRACT | 1 |
| 1 INTRODUÇÃO | 10 |
| 2 OBJETIVOS | 14 |
| 2.1 OBJETIVO GERAL | 14 |
| 2.3 OBJETIVOS ESPECÍFICOS | 14 |
| 3 HIPÓTESE | 14 |
| 4 REFERENCIAL TEÓRICO | 15 |
| 4.1 DOENÇA DE PARKINSON | 15 |
| 4.1.1 <i>Epidemiologia</i> | 17 |
| 4.1.2 <i>Descrição clínica da doença</i> | 18 |
| 4.1.3 <i>Classificação da doença de Parkinson</i> | 23 |
| 4.1.4 <i>Avaliações funcionais em pessoas com DP</i> | 27 |
| 4.2 APRENDIZADO DE MÁQUINA | 34 |
| 4.2.1 <i>Aprendizado de máquina no contexto da doença de Parkinson</i> | 42 |
| 5 MÉTODOS | 46 |
| 5.1 DELINEAMENTO DA PESQUISA | 46 |
| 5.2 LEVANTAMENTO DE DADOS | 47 |
| 5.3 CURADORIA DOS DADOS | 50 |
| 5.4 ANÁLISE DOS DADOS | 50 |
| 6 RESULTADOS | 52 |
| 6.1 CLASSIFICAÇÃO A PARTIR DE ASPECTOS FUNCIONAIS | 53 |
| 6.2 CLASSIFICAÇÃO A PARTIR DE ASPECTOS FUNCIONAIS E CLÍNICOS | 54 |
| 6.3 CLASSIFICAÇÃO A PARTIR DE ASPECTOS FUNCIONAIS, CLÍNICOS E CARACTERÍSTICAS INDIVIDUAIS . | 54 |
| 6.4 CLASSIFICAÇÃO A PARTIR DE CARACTERÍSTICAS INDIVIDUAIS E ASPECTOS CLÍNICOS | 59 |
| 7 DISCUSSÃO | 62 |
| 8 CONCLUSÃO | 69 |
| REFERÊNCIAS | 70 |

1 INTRODUÇÃO

A doença de Parkinson (DP) é uma doença neurodegenerativa que apresenta a perda da funcionalidade como um sintoma frequente entre as pessoas acometidas. Pela doença são afetados aspectos físico-motores, cognitivos, sociais e psicológicos (BRASIL, 2010; CANO-DE-LA-CUERDA ET AL, 2011; DIBBLE ET AL, 2009;).

É comum que a pessoa com DP apresente tremor a partir de algum estágio da doença, mas este não é o único e nem o principal sintoma desta condição (BRASIL, 2010). Por sua origem idiopática, a DP pode se apresentar de forma diferente entre as pessoas acometidas, e a determinação destas manifestações não é clara na literatura. Um outro aspecto a se observar é que não existe um exame laboratorial para o diagnóstico da DP. O processo se dá através do exame clínico e da combinação de sintomas apresentados; a lentidão de movimentos é a principal característica observada durante a investigação para o fechamento do diagnóstico (GREENBERG, 2014).

A DP foi descrita pela primeira vez em 1817 por James Parkinson (1817) e em 1992 Hughes et al (1992) publicaram a primeira sistematização para o diagnóstico da doença. Hoje, o padrão de referência para este diagnóstico consiste em uma combinação de critérios positivos verificados em avaliação clínica. Esta metodologia foi desenvolvida no Hospital Nacional de Neurologia e Neurocirurgia, de Londres. Por falta de respaldo laboratorial, é possível ocorrer imprecisões, justificadas por outras doenças com sintomas parecidos, porém com causa distinta (BRASIL, 2017).

Recentemente, durante o período de existência do Programa de Atividades Físicas para Pessoas com Doenças Neurodegenerativas da Faculdade de Educação Física da Universidade de Brasília (UnB), foram atendidas, principalmente, pessoas com doença de Parkinson. Era comum presenciar questionamentos entre os participantes, que desencadeavam na solicitação de revisão do diagnóstico de DP por parte dos alunos e das respectivas famílias e acompanhantes. Eram solicitadas, então, recomendações de médicos mais especializados em DP, e houve casos de diagnósticos diferentes. Este processo, no entanto, não foi protocolado, e este parágrafo serve como registro memorial do ocorrido, e que reforça a insegurança nos diagnósticos de DP.

Existem evidências de que os sintomas não motores aparecem até 20 anos antes do diagnóstico. Desordens do sono, depressão, ansiedade, episódios relevantes de constipação, disfunções no olfato e no sistema nervoso autonômico são alguns destes sintomas não motores comumente relatados pelas pessoas acometidas por DP e que se

apresentam antes das primeiras características físicas (POSTUMA, et al 2012; POEWE, 2017; ISHIHARA-PAUL, 2008). Todavia, todas estas características, apesar de frequentes, não são específicas para DP e podem ser consequência de outros quadros patológicos.

Desta forma, o aparecimento dos sintomas motores é fundamental para o diagnóstico da doença. São observados o tremor, a rigidez, um início unilateral, a preservação da assimetria ao longo da progressão da doença, a velocidade de progressão da doença, o surgimento de efeitos colaterais consequentes do tratamento medicamentoso, entre outros (GREENBERG, 2014; POEWE, 2017).

Considerando estes aspectos, foram criadas escalas para classificação da progressão da DP. A Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (GOETZ et al, 2008) é a escala com maior frequência de utilização para monitorar a debilitação motora dos indivíduos. Antes da MDS-UPDRS, a escala utilizada para classificação de pessoas com DP era a escala de Hoehn e Yahr (1967), que categoriza a progressão em cinco estágios da doença; começando pelo nível I, o mais brando e com sintomas presentes pouco incapacitantes, e indo até o nível V, que se refere ao estágio caquético (enfraquecido e irreversível) do indivíduo, em que o uso de cadeira de rodas é indispensável. Não obstante, a explanação do aparecimento de sintomas por Poewe (2017) é feita classificando a doença em três grandes estágios a partir do diagnóstico: estágio inicial, estágio moderado e estágio avançado.

Independente da escala utilizada, as classificações atribuídas aos indivíduos se utilizam das mensurações clínicas e funcionais que trazem a subjetividade da pessoa que avalia como limitação comum. Apoiado nesta justificativa, Zhan et al. (2018) apresentam uma possibilidade de monitoramento da doença utilizando a tecnologia. Os autores criaram um score de severidade da doença através do aprendizado de máquina (*Machine Learning*), em uma abordagem inovadora que trouxe componentes objetivos que podem dar complementação à avaliação clínica periódica.

Para o desenvolvimento de um novo biomarcador, os autores coletaram dados gerados por smartphones de 129 pessoas, numa tentativa de mensurar objetivamente sintomas que indicassem a severidade da doença. Foram realizados testes clínicos que aferiram as mesmas variáveis monitoradas pelos smartphones para comparação entre os dados, e verificou-se que os smartphones são capazes de contribuir para a rotina clínica através da alimentação remota de informações. Zhan et al (2018), então, desenvolveram uma nova forma de avaliação através de um aprendizado supervisionado que foi capaz de

interpretar os números provenientes das avaliações físicas e funcionais, no caso do estudo: voz, movimento de pinça, marcha, equilíbrio e tempo de reação.

É relevante aproximar a tecnologia das avaliações funcionais tradicionais para que, a partir dos dados coletados – e que irão alimentar os algoritmos dos equipamentos, seja possível auxiliar na prática clínica, no monitoramento da evolução dos sintomas, na prescrição de exercícios, e em diversas outras terapias adjuntas aos medicamentos (ZHAN ET AL 2018; BRATIC ET AL, 2018).

A aprendizagem supervisionada, citada acima, é um dos três tipos de aprendizagem de máquina existentes, além da aprendizagem não supervisionada e da aprendizagem mista. De forma generalizada, trata-se de dar ao computador a habilidade de aprender a partir de dados, através do desenvolvimento de algoritmos que dão sentido a estes mesmo dados, ou seja, é uma inteligência artificial capaz de transformar dados em conhecimento (RASHCKA, 2015).

Traduzindo em termos não técnicos, o computador tem mais capacidade de cálculo que um ser humano e, por isso, é mais sensível para identificar informações que não eram possíveis de se perceber por métodos tradicionais. Por sua capacidade de aprendizado e simulação de um cenário real, esse tipo de técnica possibilita um aprofundamento no contexto em que é utilizada e, no caso da área saúde, pode ajudar a aprimorar tratamentos e respaldar o processo de diagnóstico daquelas condições que ainda não possuem biomarcadores aferíveis em laboratórios (RASHCKA, 2015).

Este tipo de tecnologia exige um número grande de informações para que a máquina possa aprender a classificar o comportamento daquela condição, e então, fazer previsões que podem auxiliar no tratamento, acompanhamento e conhecimento da severidade da doença nos indivíduos. A partir deste raciocínio surgiu a oportunidade de resgatar os dados coletados durante o período de existência do Programa de Atividades Físicas para Pessoas com Doenças Neurodegenerativas (PAFPDN) para submeter a esta nova forma de análise.

O PAFPDN foi criado em 2012 na Faculdade de Educação Física (FEF) da Universidade de Brasília (UnB). Trata-se de uma ação acadêmica desenvolvida pelos membros do Grupo de Estudos em Atividade Física, Genética e Doenças Crônicas Não Transmissíveis e de outros estudantes e colegas de áreas afins. Por alguns anos, o PAFPDN ofereceu à comunidade intervenções de diversas modalidades de atividades físicas, como equoterapia, dança e outros exercícios para pessoas com doença de Parkinson.

Para as pesquisas envolvidas no Programa, foram coletados dados clínicos e funcionais de dezenas de pessoas. Até hoje, estes dados foram utilizados de forma separada, atendendo ao foco estrito das pesquisas e estudos provenientes da época. Desta maneira, viu-se a oportunidade de, assessorado por métodos e ferramentas mais modernas e precisas, revisitar o banco de dados geral do programa, submetendo os dados à técnica do aprendizado de máquina.

Então, o que se pretende com esta pesquisa é organizar o banco de dados do PAFPDN, no que tange aos aspectos clínicos e funcionais, considerando os dados de caracterização da amostra, submetendo as informações ao treinar de máquina a ponto de se obter um resultado para auxiliar na classificação clínica de pessoas com DP.

2 OBJETIVOS

2.1 Objetivo geral

Quantificar a possibilidade do auxílio na predição de classificação da doença de Parkinson a partir de aspectos funcionais de pessoas acometidas, utilizando técnicas do aprendizado de máquina.

2.3 Objetivos específicos

- Identificar o nível de progressão da doença de Parkinson de acordo com a escala de UPDRS 3 a partir da classificação dos aspectos funcionais de pessoas acometidas;
- Identificar o nível de progressão da doença de Parkinson de acordo com a escala de UPDRS 3 a partir da classificação dos aspectos funcionais e aspectos clínicos;
- Identificar o nível de progressão da doença de Parkinson de acordo com a escala de UPDRS 3 a partir da classificação dos aspectos funcionais, características individuais e aspectos clínicos de pessoas acometidas;

3 HIPÓTESE

O aprendizado de máquina sendo treinado com biomarcadores serve de auxílio na classificação da severidade da doença de Parkinson.

4 REFERENCIAL TEÓRICO

4.1 Doença de Parkinson

A doença de Parkinson (DP) é considerada um distúrbio motor, por sua origem no corpo humano e por seus sintomas principais e mais aparentes. É uma doença incapacitante, degenerativa e com cura desconhecida. Para além de aspectos físicos, a existência de sintomas não motores como depressão, declínio cognitivo, disfunções autonômicas, distúrbios do sono, entre outras, evidenciam que a DP é uma doença mais grave e mais abrangente do que se imaginava no início de sua investigação (POEWE, W. et al, 2017; GREENBERG, 2014).

A DP ainda não tem cura (POEWE, 2017) e, de acordo com a portaria 07 de 2017 do Ministério da Saúde (BRASIL, 2017) a redução da progressão da doença é o foco inicial do tratamento medicamentoso. A partir daí, o tratamento sintomático deve ser administrado de acordo com a necessidade do indivíduo, a fim de incrementar seu desempenho funcional, mas levando em consideração o risco de efeitos colaterais adversos. Neste sentido, leva-se em consideração a progressão da doença e o tempo de acometimento. Isso significa que é preferível que a pessoa conviva com alguns sintomas da doença de Parkinson para que seja possível uma administração da medicação por mais tempo. Existe risco do desenvolvimento de efeitos adversos que podem ultrapassar os efeitos deletérios da própria doença. A doença de Parkinson apresenta mortalidade baixa, principalmente na primeira década de diagnóstico (PINTER et al, 2015), justificando o pensamento a longo prazo da administração medicamentosa.

Além dos medicamentos, e contextualizando as pesquisas envolvendo a doença entre os objetos estudados pela Educação Física, existe consenso sobre a prescrição do exercício como tratamento ajunto (POEWE, 2017). A prática de exercícios é reconhecida como uma estratégia terapêutica para retardar a progressão da doença (BOUÇA-MACHADO, 2020). Já existem consensos sobre a prescrição de exercícios para pessoas com Parkinson (MIRELMAN, 2019)

As prescrições de exercício focam em cinco principais áreas: condicionamento físico, mobilidade, atividades manuais, equilíbrio e marcha (BOUÇA-MACHADO, 2020), e há de se levar em consideração que, quanto mais avançada é a doença, menor é a eficácia do exercício, em função da própria degeneração da doença. Estabelecer uma prescrição correta de exercícios nos estágios iniciais da doença é determinante para a preservação da autonomia e qualidade de vida do indivíduo acometido pela doença.

O exercício como intervenção terapêutica, também chamado de exercício clínico, se utiliza de ciências do esporte que envolvem, além de questões propriamente físicas, aspectos psicológicos e outros associados, focando em respostas em curso e longo prazo (BOUÇA MACHADO 2020). Os autores argumentam que o exercício vem sendo considerado uma ferramenta útil e importante em diversos tratamentos médicos.

Existem críticas da comunidade científica envolvida na saúde quanto à prescrição de exercícios para pessoas com Parkinson. Os argumentos trazidos pela literatura dizem que não é claro de que forma o exercício interfere nos aspectos clínicos da doença de Parkinson, além de serem frequentemente publicados com baixo rigor metodológico ou com técnicas de mensuração distintas (RAMAZZINA, 2017).

As avaliações clínicas, assim com avaliações físicas e de performance de uma forma geral, estão sujeitas à interferência do avaliador (DA FONTOURA, 2011). A partir disto, instrumentos e equipamentos com um maior e mais apurado método de mensuração poderiam ser uma forma de objetivar as avaliações, significando um acompanhamento mais criterioso dos efeitos do exercício no retardamento da doença, por exemplo. É possível, também, vislumbrar que este tipo de estratégia possa vir a auxiliar na construção, ou definição, do diagnóstico precoce de DP, configurando o estabelecimento de biomarcadores um evento relevante para a rotina clínica.

De acordo com o protocolo do Ministério da Saúde (BRASIL, 2010):

“A natureza progressiva da DP e suas manifestações clínicas (motoras e não motoras), associadas a efeitos colaterais precoces e tardios da intervenção terapêutica, tornam o tratamento da doença bastante complexo. Estima-se que a taxa de morte dos neurônios dopaminérgicos da substância nigra se situe ao redor de 10% ao ano. Consequentemente, com o tempo, a sintomatologia parkinsoniana piora e a necessidade de medicamentos sintomáticos aumenta. O grau de resposta aos medicamentos vai decrescendo com a progressão da doença e novos sintomas vão surgindo. Um objetivo desejado é reduzir ou interromper essa progressão. A prevenção primária não é possível devido à ausência de marcadores biológicos ou fatores de risco identificáveis, excetuando-se o envelhecimento e a transmissão genética em raras famílias. A prevenção secundária, uma vez que a DP tenha sido diagnosticada, busca reduzir a taxa de progressão, parar ou mesmo reverter a morte neuronal.”

Não se pretende, com este trabalho, estabelecer um novo biomarcador definitivo para o diagnóstico da doença, mas fornecer à comunidade científica análises preliminares, com utilização de tecnologia avançada. Se as disfunções motoras se apresentam mais visivelmente no estágio moderado e avançado da doença, em consequência de lesões no

sistema nervoso já afetados desde os primeiros estágios da doença, pode ser possível a observação de sinais motores discretos, iniciais, mas presentes.

Assim, ao levar em consideração às características gerais da doença apresentadas anteriormente, respalda-se a necessidade de investigação e desenvolvimento de novas ferramentas para o diagnóstico, bem como para a prescrição de tratamentos – tradicionais e adjuntos, visando a manutenção da qualidade de vida e funcionalidade de indivíduos acometidos pela doença.

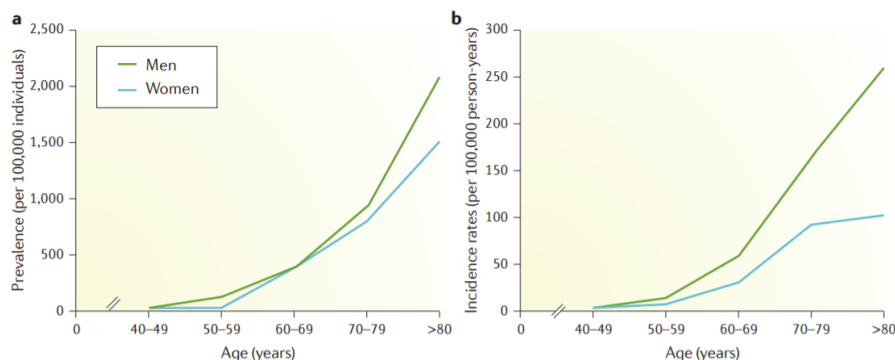
4.1.1 *Epidemiologia*

A DP é uma doença neurodegenerativa comum, mais prevalente em homens do que em mulheres e acomete em geral pessoas com idade avançada (WIRDEFELDT, 2011). São poucos e antigos os estudos que tratam dos números de incidência e prevalência de forma objetiva no Brasil, uma vez que a notificação desta doença não é obrigatória (BOVOLENTA et al, 2017).

A meta análise de Pringsheim (2014) apresenta que a prevalência é maior de acordo com a idade: são estimados 41 casos por 100 mil habitantes na faixa etária entre 40 e 49 anos, em contraste com 428 casos por 100.000 hab. em pessoas entre 60 e 69 anos e 1087 casos por 100.000 hab. em pessoas acima de 70 anos. Estes números são maiores do que os estimados para o Brasil no estudo de Barbosa (2006). Os dados governamentais não são atualizados, visto que o Ministério da Saúde referencia um estudo de 1996 (TANNER, 1996) no protocolo de diagnóstico e tratamento, publicado em 2017 (BRASIL, 2017), no que tange à epidemiologia da doença.

A cartilha publicada pela revista Nature em 2017 (POEWE, 2017) reporta que a prevalência global da doença é de 0,3% da população mundial (mais de 23 milhões de pessoas, aproximadamente). Considerando todas as faixas etárias, acomete cerca de 3% das pessoas acima de 80 anos e se apresenta mais prevalente em homens que em mulheres (Figura 1). A cartilha também faz referência à estudos que sugerem que a DP é menos prevalente em pessoas de etnias afro americanas e asiáticas (VAN DEN EEDEN, 2003).

Figura 1 - Prevalência e incidência da doença de Parkinson, por sexo (POEWE, 2017)



4.1.2 Descrição clínica da doença

A DP foi descrita há mais de 200 anos e é classificada como uma desordem do movimento, uma vez que os principais sintomas da doença são verificados por alterações no padrão motor de indivíduos acometidos (GREENBERG, 2014; HAINES, 2019; BRASIL, 2010; PARKINSON, 1817). Por muito tempo os estudos que tratam da sintomatologia da DP estiveram focados apenas nos aspectos motores, mesmo com a descrição dos sintomas não motores ainda nos primeiros relatos de Parkinson e Charcot (DELLA COLETTA, 2016). Atualmente, é listada no capítulo VI do CID-10, capítulo referente a doenças do sistema nervoso, juntamente com o grupo de doenças extrapiramidais e transtornos do movimento (CLASSIFICAÇÃO INTERNACIONAL DE DOENÇAS – CID). Caracteriza-se como distúrbio motor qualquer falha na transmissão de sinais entre as partes que compõem o sistema motor, desde as áreas mais superiores do sistema nervoso central até a placa terminal motora, localizada junto aos músculos (GREENBERG, 2014).

No Brasil, o protocolo clínico para diagnóstico e tratamento foi atualizado em 2017 pelo Ministério da Saúde, revogando, assim, o protocolo até então utilizado de 2010 (BRASIL, 2017). Os critérios para a definição do diagnóstico da DP são a presença de lentidão dos movimentos (bradicinesia) associada à rigidez muscular e/ou tremor de repouso (critérios necessários). E, para além da verificação desses sintomas, é necessária a presença de pelo menos três critérios de suporte positivo, a saber:

- Início unilateral;
- Quadro progressivo de sintomas;
- Persistência da assimetria dos sintomas;
- Boa resposta a levodopa;

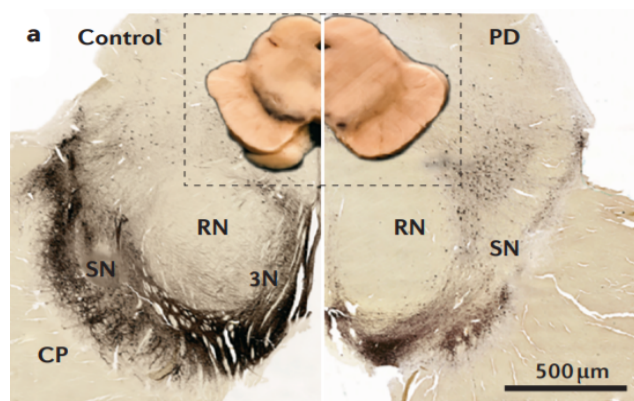
- Presença de discinesias induzidas por levodopa;
- Resposta a levodopa por 5 anos ou mais;
- Evolução clínica de 10 anos ou mais.

Além dos critérios de suporte positivo, a presença de qualquer critério negativo para a DP sugere diagnóstico de alguma outra forma de parkinsonismo; são os critérios negativos:

- História de acidente vascular cerebral (AVC) de repetição;
- História de trauma craniano grave;
- História definida de encefalite;
- Crises oculogíricas;
- Tratamento prévio com neurolépticos;
- Remissão espontânea dos sintomas;
- Quadro clínico estritamente unilateral após 3 anos;
- Paralisia supranuclear do olhar;
- Sinais cerebelares;
- Sinais autonômicos precoces;
- Demência precoce;
- Liberação piramidal com sinal de Babinski;
- Presença de tumor cerebral ou hidrocefalia comunicante;
- Resposta negativa a altas doses de levodopa;
- Exposição ao metilfeniltetrapiridínio (MPTP)

Apesar dos sintomas muito parecidos, a doença de Parkinson se difere de outros tipos de parkinsonismo por sua neuropatologia, que consiste na perda neuronal de dopamina acumulada na substância negra e na acumulação intracelular de α -sinucleína (GREENBERG, 2014; DICKSON et al, 2009). Este processo causa a despigmentação da substância negra de pessoas acometidas pela doença, conforme mostrado na Figura 2 . Entretanto, só é possível verificar esta despigmentação em análises *post-mortem*.

Figura 2 - Despigmentação da substância negra em pessoas com doença de Parkinson (à direita) em comparação com controle (à esquerda) (POEWE, 2017)

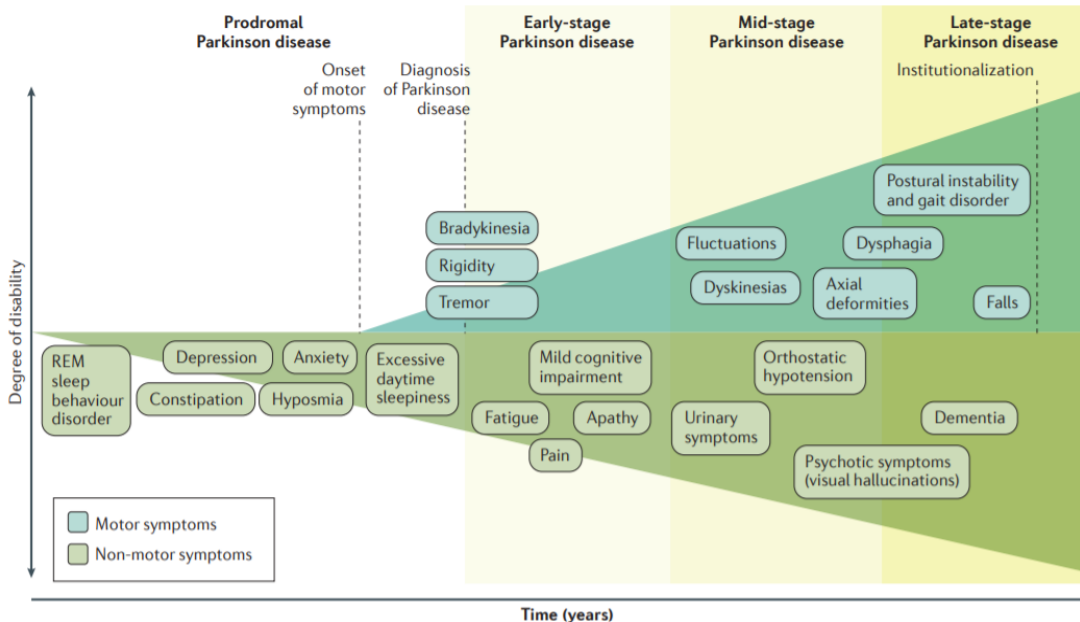


Na doença de Parkinson são verificados distúrbios hipercinéticos – que são movimentos involuntários anormais, e distúrbios hipocinéticos, caracterizados pela pobreza de movimentos (GREENBERG, 2014). Esses distúrbios se originam em estruturas profundas do SNC, chamadas núcleos da base e, portanto, não afetam diretamente a força, a sensibilidade ou a função cerebelar.

Há divergência entre autores sobre a participação da substância negra como parte integrante dos núcleos da base. Ao levar em consideração apenas critérios funcionais, é possível associar a substância negra aos núcleos da base (GREENBERG, 2014), mas, do ponto de vista anatômico, a substância negra faz parte do tronco encefálico, logo, região do mesencéfalo, enquanto os núcleos da base estão posicionados na base do cérebro (MACHADO, 2014).

A doença de Parkinson é diagnosticada através da combinação de avaliações clínicas, e o aparecimento dos sintomas motores é fundamental para o diagnóstico da doença, mesmo com evidências de uma fase prodrômica (prévia) da doença, apresentada por Hawkes (2010) e Poewe (2017), que sinalizam o aparecimento de sintomas não motores característicos da DP até 20 anos antes do diagnóstico (**Figura 3**).

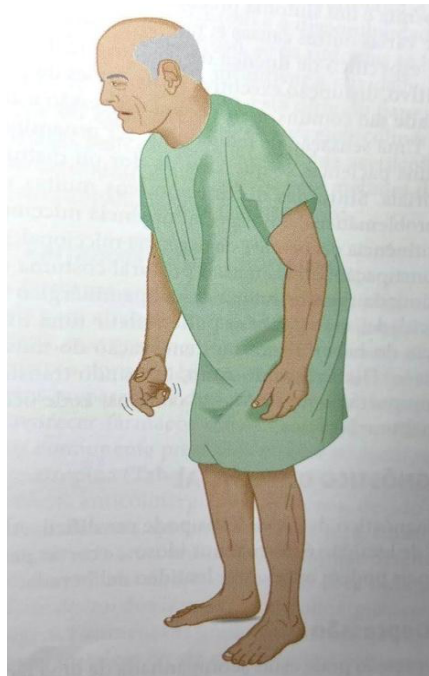
Figura 3 - Sintomas clínicos associados com a progressão da doença de Parkinson (POEWE, 2017).



Os sintomas não motores mais comuns são distúrbios do sono, ansiedade, depressão, ansiedade, constipação forte, perda ou diminuição do olfato e prisão de ventre. A publicação de Poewe (2017) ratifica os achados de Hawkes (2010), confirmando a relação destes sintomas com a doença. Entretanto, não é possível estabelecer uma relação determinante entre estes sintomas e a DP, visto que podem ser manifestações consequentes de outros diagnósticos e/ou quadros clínicos. (ISHIHARA-PAUL, 2008)

Os sintomas motores afetam, principalmente, a marcha, a estabilidade dinâmica, o controle de movimentos e a autonomia funcional. Greenberg (2014) caracteriza uma pessoa com Parkinson ao apresentar a postura fletida, consequência da baixa capacidade de iniciar a marcha, inclinando a coluna cada vez para a frente enquanto caminha no mesmo lugar. Essa dificuldade de iniciar a movimentação é chamada de *freezing*.

Figura 4 - Postura fletida típica de indivíduo com doença de Parkinson (GREENBERG, 2014)



Uma deficiência ou desordem do movimento, como é o caso da doença de Parkinson, é caracterizada pela substituição de um mecanismo neural rítmico, ou a troca de uma automaticidade saudável para um controle compensatório voluntário para execução de tarefas (HERCULANO-HAZEL; KANDEL, 2014). Em outras palavras, à medida que a pessoa aprende e aprimora determinados padrões de movimentos, estes padrões podem ser executados de forma automática enquanto a atenção da pessoa pode estar voltada para uma outra atividade. Por exemplo: é possível prestar atenção em uma conversa enquanto se caminha, sem que o movimento da caminhada requeira atenção. Entretanto, no caso da pessoa com doença de Parkinson, este movimento automático é inviabilizado pela falha de conexão das redes neurais dos núcleos da base devido à diminuição da quantidade de dopamina concentrada na substância negra.

Neste sentido, as alterações motoras da DP são manifestadas a partir da combinação dos sintomas clássicos, a saber: bradicinesia, rigidez, diminuição da amplitude de movimento e tremor. Essas alterações aumentam com a progressão da doença, afetando diretamente a independência e a qualidade de vida da pessoa acometida (MIRELMAN, 2019). Cabe destacar que em estágios iniciais nem todos os sintomas que

podem ser percebidos são específicos da doença. A redução da velocidade da marcha e a diminuição do tamanho da passada são características da DP, mas podem ser consequências de diferentes fatores, enquanto a redução do balanço dos braços e assimetria entre os membros já indicam a confirmação do diagnóstico.

Cabe citar que alterações motoras também podem surgir através de respostas não adequadas ao tratamento farmacológico em estágios avançados da doença, bem como disfunções autonômicas e demência (BRASIL, 2010). Greenberg et al. (2014) defendem a não utilização de fármacos em parkinsonismo precoce, sob a justificativa de que os efeitos colaterais dos medicamentos podem se apresentar mais severos que os próprios sintomas da doença. Borovac (2016) identifica que a administração medicamentosa da DP pode afetar doenças cardiovasculares, psicoses e depressão, além do risco do comprometimento do sistema renal e hepático.

A partir desta vastidão de sintomas motores e não motores, e considerando os efeitos colaterais do próprio medicamento, Opara (2017) diz que monitorar a progressão da doença passa a ser uma tarefa difícil. Ademais, a DP se manifesta de forma diferente entre as pessoas acometidas; os sintomas principais são chamados de cardinais, indicando a existência de um espectro de características, que são influenciadas pela individualidade biológica e também pelo contexto social e estilo de vida da pessoa acometida (JANKOVIC, 2013)

Após a publicação dos princípios para diagnóstico de doença de Parkinson (Hughes, 1992) foram desenvolvidos diferentes instrumentos para subsidiar a prática clínica, principalmente escalas de avaliação. Jankovic (2008) explica que uma vez que não existe um teste único e definitivo para determinar a presença da doença de Parkinson, fazem-se necessários critérios clínicos para o respaldo deste diagnóstico.

4.1.3 Classificação da doença de Parkinson

A Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (GOETZ, 2008) é a escala com maior frequência de utilização para monitorar a debilitação motora da DP. A International Parkinson's Disease and Movement Disorders Society (MDS) é uma associação mundial de médicos, cientistas e outros agentes da saúde, responsável pela elaboração, validação e capacitação de avaliadores da MDS-UPDRS.

De acordo com a explicação da própria sociedade:

“A MDS UPDRS tem quatro partes: Parte I (aspectos não motores da vida diária), Parte II (aspectos motores da vida diária), Parte III (avaliação motora) e Parte IV (complicações motoras). A Parte I tem dois componentes: IA refere-se a um determinado número de comportamentos avaliados pelo investigador através de toda a informação pertinente obtida a partir dos pacientes e cuidadores e IB, que é preenchida pelo paciente, com ou sem a ajuda do cuidador, mas de forma independente do investigador. Este componente pode, no entanto, ser revisto pelo avaliador, para garantir que todas as perguntas foram respondidas de modo claro, podendo o avaliador ajudar a explicar qualquer ambiguidade encontrada. A Parte II é desenhada para ser um questionário de autopreenchimento, como a parte IB, porém pode ser revista pelo investigador para garantir o seu claro e completo preenchimento. Deve ser salientado que as versões oficiais das Partes 1A, 1B e 2 da MDS UPDRS não têm avaliações *on* e *off* separadas. Contudo, para programas ou protocolos específicos, pode-se utilizar as mesmas questões para os estados *on* e *off* separadamente. A Parte III tem instruções para o avaliador fornecer ou demonstrar ao paciente e é preenchida pelo avaliador. A Parte IV tem instruções para o avaliador e também instruções para serem lidas ao paciente. Esta parte integra a informação obtida do paciente com as observações e julgamentos clínicos do avaliador, sendo preenchida pelo avaliador.”

O Quadro 1 apresenta um resumo dos aspectos investigados pela MDS-UPDRS.

Quadro 1 - Estrutura da avaliação da MDS-UPDRS

| | | | |
|---|--|--|--|
| UPDRS - ESCALA UNIFICADA DE CLASSIFICAÇÃO DA DOENÇA DE PARKINSON | UPDRS I Aspectos não motores da vida diária | UPDRS IA Comportamentos complexos avaliados pelo investigador | Disfunção cognitiva |
| | | | Alucinações e psicose |
| | | | Humor depressivo |
| | | | Ansiedade |
| | | | Apatia |
| | | | Aspectos da Síndrome de desregulação dopaminérgica |
| | | UPDRS IB Respostas autorreferidas | Problemas do sono |
| | | | Sonolência diurna |
| | | | Dor e outras sensações |
| | | | Problemas urinários |
| | | | Prisão de ventre |
| | | | Tontura a se levantar |
| | | | Fadiga |
| | UPDRS II Aspectos motores da vida diária (respostas autorreferidas) | Fala | |
| | | Saliva e baba | |
| | | Mastigação e deglutição | |
| | | Tarefas para comer | |
| | | Vestir | |
| | | Higiene | |
| | | Escrita | |
| | | Passatempos e outras atividades | |
| | | Virar-se na cama | |
| | | Tremor | |
| | | Sair da cama, do carro ou de uma cadeira baixa | |
| | | Marcha e equilíbrio | |
| | | Bloqueios na marcha | |
| | UPDRS III Avaliação Motora (considera o estadiamento de escala de Hoehn e Yahr e o impacto de discinesias durante a avaliação da parte III) | Fala | |
| | | Expressão facial | |
| | | Rigidez | |
| | | Movimento de pinça (finger tapping) | |
| | | Movimento das mãos | |
| | | Movimentos de pronação-supinação das mãos | |
| | | Bater dos dedos dos pés | |
| | | Agilidade das pernas | |
| | | Levantar-se da cadeira | |
| | | Marcha | |
| | | Freezing (bloqueio na marcha) | |
| Estabilidade postural | | | |
| Postura | | | |
| Bradicinesia (espontaneidade global de movimento) | | | |
| Tremor postural das mãos | | | |
| Tremor cinético das mãos | | | |
| Amplitude do tremor de repouso | | | |
| Persistência do tremor de repouso | | | |
| UPDRS IV Complicações motoras | Discinesias (excluindo distonia do período off) | Tempo com discinesias | |
| | | Impacto funcional das discinesias | |
| | Flutuações motoras | Tempo em off | |
| | | Impacto funcional das flutuações | |
| | Distonia em off | Complexidade das flutuações motoras | |
| | Distonia dolorosa em período off | | |

Antes da MDS UPDRS, a escala de Hoehn e Yahr (1967) já classificava a progressão da doença em pessoas com DP. Margaret Hoehn e Melvin Yahr conduziram uma análise dos sintomas de 856 pessoas diagnosticadas com doença de Parkinson, paralisia agitante e parkinsonismos entre 1949 e 1964 e desenvolveram uma classificação prática da progressão da doença (Figura 5).

Figura 5 - Recorte da descrição inicial da classificação da doença de Parkinson (Hoehn e Yahr, 1967)

Stage I. Unilateral involvement only, usually with minimal or no functional impairment.
Stage II. Bilateral or midline involvement, without impairment of balance.
Stage III. First sign of impaired righting reflexes. This is evident by unsteadiness as the patient turns or is demonstrated when he is pushed from standing equilibrium with the feet together and eyes closed. Functionally the patient is somewhat restricted in his activities but may have some work potential depending upon the type of employment. Patients are physically capable of leading independent lives, and their disability is mild to moderate.
Stage IV. Fully developed, severely disabling disease; the patient is still able to walk and stand unassisted but is markedly incapacitated.
Stage V. Confinement to bed or wheelchair unless aided.

A escala de Hoehn e Yahr divide a progressão da doença em cinco estágios, a saber: (1) início unilateral da doença; (2) acometimento bilateral sem déficit de equilíbrio; (3) acometimento bilateral com sintomas importantes e instabilidade postural, mas independência funcional; (4) Incapacidade grave e possibilidade de locomoção, e; (5) Acamado com locomoção restrita à cadeira de rodas ou auxílio de terceiros.

Em 2004 a MDS publicou uma crítica, incluindo análise da escala de Hoehn e Yahr e também recomendações para na avaliação da doença. Os autores destacam que mesmo não contemplando todas as especificidades da DP a escala de Hoehn e Yahr continuava sendo utilizada comumente. Não se sabe ao certo a origem da adaptação da escala que incluía estágios escalonados em pontuações de 0,5. A Figura 6 mostra a comparação das duas escalas, incluindo 1,5 como classificação pelo acometimento

unilateral axial e 2,5 para classificar como doença bilateral leve, com recuperação no “teste do empurrão¹” (GOETZ, et al 2004).

Figura 6 - Recorte da tabela de comparação entre a Escala de Hoehn e Yahr tradicional e modificada (GOETZ, et al. 2004)

| Hoehn and Yahr scale | Modified Hoehn and Yahr scale |
|--|--|
| 1: Unilateral involvement only usually with minimal or no functional disability | 1.0: Unilateral involvement only |
| 2: Bilateral or midline involvement without impairment of balance | 1.5: Unilateral and axial involvement |
| 3: Bilateral disease: mild to moderate disability with impaired postural reflexes; physically independent ^a | 2.0: Bilateral involvement without impairment of balance |
| 4: Severely disabling disease; still able to walk or stand unassisted | 2.5: Mild bilateral disease with recovery on pull test |
| 5: Confinement to bed or wheelchair unless aided | 3.0: Mild to moderate bilateral disease; some postural instability; physically independent |
| | 4.0: Severe disability; still able to walk or stand unassisted |
| | 5.0: Wheelchair bound or bedridden unless aided |

^aStage 3 is a summary of the authors' original, more narrative description.

Estas e outras escalas buscam traduzir o grau de manifestação dos sintomas para uma régua de progressão da doença de Parkinson. Logo, é possível justificar que a investigação considere os aspectos motores da doença para determinação do nível de progressão, ao considerar que o estágio avançado da doença é caracterizado pelo aparecimento de disfunções do equilíbrio e da marcha, de quedas e pelo aumento da disfagia (dificuldade de deglutição, ou de engolir e de deformidades axiais (GREENBERG, 2014).

4.1.4 Avaliações funcionais em pessoas com DP

Da mesma forma que uma amostra de sangue permite identificar e quantificar índices na corrente sanguínea que representam a saúde de uma pessoa, para avaliar características como mobilidade, equilíbrio, força, performance etc., são utilizadas amostras dos comportamentos motores dos indivíduos, coletadas durante uma avaliação que utiliza testes funcionais para traduzir numericamente o padrão de movimento que está sendo avaliado.

¹ Empurra-se bruscamente a pessoa para trás, tocando nos ombros, a pessoa dá mais que três passos, mas recupera o equilíbrio sem ajuda.

O movimento do corpo humano, de uma maneira geral, é o resultado de complexos processos de programação e monitoramento que envolvem diversas regiões cerebrais e terminam na contração das fibras musculares (LENT, 2001). Em condições normais, os movimentos acontecem dentro dos padrões motores desenvolvidos na pessoa. Funcionalmente é possível dividir o sistema motor em quatro elementos: (a) os efetadores – músculos, que executam os movimentos; (b) os ordenadores – motoneurônios, que recebem as informações dos controladores e comandam os efetadores; (c) os controladores – principalmente os núcleos da base, que cuidam da correta execução dos comandos motores e; (d) os planejadores – córtex, responsáveis pela programação e pelo planejamento coordenado de sequências de instruções às demais estruturas (LENT, 2001; DUARTE, 2010) .

Um exemplo de tarefa executada por essas estruturas é o controle da locomoção. Herculano-Houzel (2001) caracteriza a natureza da locomoção como cíclica e semiautomática. Este conceito refere-se à interação entre os movimentos voluntários e involuntários. O movimento de marcha, por exemplo, é rítmico e acontece quase que involuntariamente: a transferência de peso entre as duas pernas acontece de maneira automática. A revisão de Clark (2015) categoriza este processo descrito como automaticidade do movimento. O conceito gira em torno da capacidade que o Sistema Nervoso Central tem de coordenar alguns movimentos, como a marcha, utilizando o mínimo de ativação de células, ou seja, com pouca atenção específica demandada.

Outros autores (SCHNEIDER E CHEIN, 2003) utilizam uma nomenclatura diferente nos estudos sobre os processos neurais que resultam em tarefas. Entendendo que toda tarefa parte de uma iniciativa voluntária, os processos não automáticos foram chamados de processos controlados. Schneider e Chein (2003) são enfáticos em explicar que o desenvolvimento da automaticidade de uma tarefa demanda tempo e treino. Takakuzaki (2013) traz, do ponto de vista desta pesquisa, a definição mais adequada ao dizer que esse tipo de movimento é resultado de um controle automático do processo.

Neste contexto, Schneider e Chein (2003) entenderam que a velocidade está relacionada à automaticidade, e demonstraram que o processamento automático é mais rápido que o processamento voluntário. Neurologicamente, isso se dá pela frequente ativação de um mesmo conjunto de células neurais em uma situação habitual, necessitando de menor atenção despendida para a execução da tarefa (CLARK, 2015).

Ainda no exemplo da marcha, o controle da velocidade e da amplitude da passada iniciam como movimento de processamento voluntário, o qual depende de controle consciente do indivíduo; a partir daí, a velocidade e a amplitude da passada se mantêm,

então, de forma controlada automaticamente. Mas, se no percurso desta marcha surgir um obstáculo, é necessário que haja um processamento da situação para que a pessoa possa desviar. O evento demanda uma ação voluntária, e esta interação entre os dois tipos de comando do movimento influencia na execução do próprio movimento.

Clark (2015) identifica que as lesões do Sistema Nervoso Central também geram alterações de automaticidade, e isto pode ser verificado por alterações nos padrões dos movimentos. Uma vez que a doença de Parkinson afeta a comunicação dentro dos núcleos da base e suas conexões com as vias diretas e indiretas, é esperado que uma das consequências do acometimento seja a alteração dos padrões normais de movimento. Nas palavras dos autores: “a categorização de um movimento anormal é, em geral, o primeiro passo em direção ao diagnóstico neurológico” (GREENBERG, 2014).

Desta forma respaldados, os testes funcionais são capazes de estabelecer esta categorização através da determinação dos valores de base da mobilidade funcional, na predição do risco de quedas, no monitoramento do tratamento e no resultado de intervenções. Estes testes vêm sendo aplicados ao longo da prática clínica como ferramenta de auxílio e suporte para definição de diagnósticos e prognósticos (OPARA, 2017). A partir da observação da manifestação motora é possível indicar alterações do estado de saúde não detectáveis em exames laboratoriais.

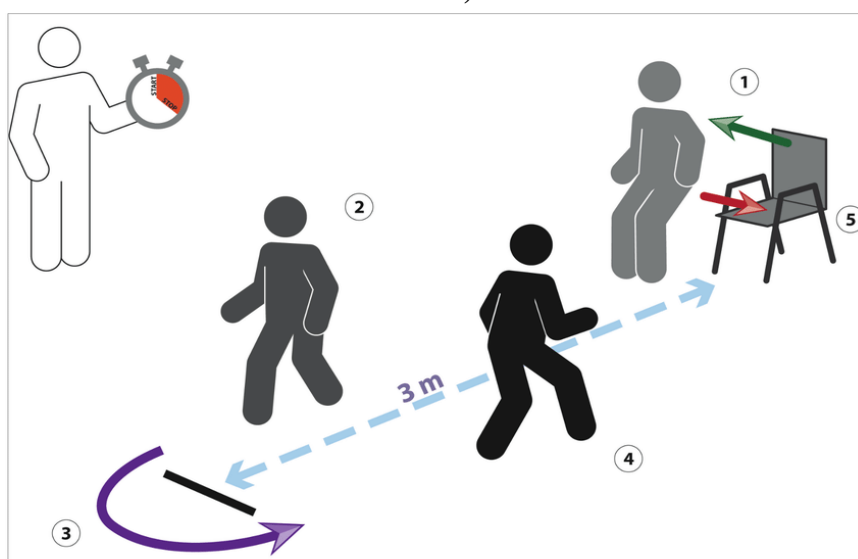
É evidente que nos dias atuais existam recursos avançados e tecnológicos para uma avaliação mais aprofundada da mobilidade humana, como as plataformas de equilíbrio, por exemplo (GIARDINI, 2018; DE JSEUS SANTOS, 2017). Entretanto, este tipo de equipamento continua sendo uma alternativa cara, e por isso também é pequena sua utilização no cotidiano clínico da maioria das pessoas com distúrbios motores.

Como alternativa viável, então, os testes funcionais foram sendo propostos e levam em consideração aspectos relacionados a autonomia funcional e atividades da vida diária para criar uma espécie de um score ou uma escala para a avaliação das manifestações motoras. As tarefas motoras básicas são objeto de investigação dos testes funcionais de forma geral para tentar identificar possíveis debilidades físicas em pessoas idosas ou que apresentem outros fatores que possam comprometer o comportamento motor do ser humano.

Os testes funcionais são populares. Em 2017 foi publicada uma compilação de avaliações motoras (OPARA, 2017) que citou mais de 25 testes e avaliações funcionais apenas no contexto da doença de Parkinson. É possível justificar essa diversidade e quantidade de testes funcionais pela facilidade de aplicação, o baixo custo e geração de resultados objetivos.

O Timed Up and Go Test (TUG) (PODSIADLO, 1991), é um teste funcional elaborado sob a justificativa de quantificar as atividades da vida diária. O TUG é um teste de simples condução e rápida execução; consiste na seguinte tarefa: a pessoa, sentada em uma cadeira, deve se levantar, andar por três metros, se virar e retornar para a cadeira, se sentando novamente. É registrado o tempo em segundos que a pessoa leva para executar toda a tarefa.

Figura 7 - Representação gráfica da execução do TUG (NIÉRAT et. al, 2016)



Em 1989 o TUG foi publicado como um instrumento de avaliação funcional de pessoas idosas frágeis, e contou com a participação de sujeitos diagnosticados com DP. Entretanto, foram excluídos do estudo as pessoas que apresentavam nível \geq IV na escala de Hoehn e Yahr (ver item 4.1.3.), e não foi observado um comportamento específico entre os participantes com Parkinson que participaram da pesquisa. Yoo (2020) traz a reflexão de que se os sintomas da DP afetam frequentemente a postura e a marcha, é esperado que o TUG seja um teste plausível no monitoramento da progressão da doença.

Na época de sua publicação, o teste foi relacionado com a escala de equilíbrio de Berg (1989), com velocidade de marcha e com a Escala de Barthel para Atividades da Vida Diária (MAHONEY, 1965), apresentando relação com todas as variáveis. Desta forma, o teste foi apresentado à comunidade científica como um quantificador da mobilidade funcional do indivíduo, sendo então, útil no acompanhamento da progressão do envelhecimento e de diversas doenças.

Em 2001, Morris et al (2001) publicaram os achados positivos sobre a confiabilidade da aplicação do TUG para pessoas com DP. O TUG se mostrou sensível em quantificar os prejuízos motores na pessoa, inclusive sem a medicação; os sujeitos da pesquisa foram avaliados na fase *on* e *off* (ver item 4.1.2).

Em 2020 foi publicado um estudo de coorte que apresenta o TUG como teste de monitoramento na fase prodrômica da doença. Os autores apontam que indivíduos que levam entre 10 e 20 segundos para completar o TUG tem mais chances de desenvolver DP do que indivíduos que levam menos de 10 segundos. Para aqueles que executaram o TUG com 20 segundos ou mais, a chance do desenvolvimento da doença é dobrada.

O TUG vem sendo usado frequentemente até os dias de hoje em avaliações com indivíduos de Parkinson. Composto a avaliação motora de indivíduos acometidos, o TUG é um indicativo de manutenção da postura e do equilíbrio (OPARA, 2017). Na mesma perspectiva, para analisar a mobilidade através da velocidade de marcha, pesquisas da área vêm usando o 10-meters Walk Test (TMW).

Os autores do TMW defendem que para a maioria dos indivíduos, independência significa a habilidade de locomoção (BOHANNON, 1996). No artigo original, a avaliação da velocidade de marcha se dava pelo registro, em segundos, do tempo que a pessoa levava para percorrer 25 pés (7,62m). Outras distâncias, no entanto, vêm sendo utilizadas na aplicação deste mesmo teste (PETERS, 2013).

Os resultados do TMW são expressos em metros por segundo e podem ser indicativos de qualidade de marcha (LINA, 2020), equilíbrio (CANO PORRAS, 2019) e mobilidade no geral (SOUSA, 2019). Steffen e Seney (2014) investigaram a velocidade de marcha como uma variável relevante para pessoas com DP. Os autores atestaram que o TMW é uma avaliação sensível para o monitoramento de pessoas com Parkinson.

Outro teste funcional abordado nesta pesquisa é o 30-s Chair-Stand Test (T30) (JONES, RIKLI E BEAM, 1999). Este teste tem como objetivo avaliar a força de membros inferiores e foi proposto em 1999 como uma alternativa inclusiva para pessoas idosas e frágeis. O teste consiste em que a pessoa sentada levante-se de uma cadeira e sente-se novamente quantas vezes conseguir durante 30 segundos. Os braços são posicionados cruzados à frente do tronco.

Jones, Rikli e Beam (1999) identificaram que nos testes da época que avaliavam a força de membros inferiores eram registrados o tempo em que a pessoa levava para executar 5 ou 10 repetições. Os autores argumentaram que um número considerável de sujeitos idosos não conseguia executar o número de repetições exigida pelo teste. No

estudo de confiabilidade do T30, uma das observações foi a capacidade de execução do teste por 100% da amostra recrutada.

No contexto da doença de Parkinson, Greenberg et al (2014) afirmam que os distúrbios motores não afetam diretamente a força dos indivíduos, e trazem um ponto que deve ser considerado, que diz assim:

“Pacientes com déficits motores geralmente se queixam de fraqueza, sensação de peso, rigidez, desajeitamento, distúrbio do controle muscular ou dificuldade na execução dos movimentos. O termo fraqueza algumas vezes é usado de um modo não específico para denotar fadiga ou perda de energia, motivação ou entusiasmo, e deve-se tomar o cuidado de esclarecer o que o paciente quer dizer.”

Em 2015 Tillman et al (2015) publicaram uma meta análise que considerou estudos de força para membros inferiores (MMII) e concluiu que, apesar dos resultados positivos no que tange ao fortalecimento muscular, não foi possível verificar diferenças significativas na velocidade de marcha, ou equilíbrio da pessoa com DP. Ramazzina (2017) endossa o argumento em uma revisão sistemática, e adiciona que não é clara a relação entre os exercícios de força e as alterações na manifestação clínica da doença de Parkinson.

Em contraponto, uma meta-análise publicada em 2020 (LI et al 2020) justifica que talvez a diversidade de testes e de formas de exposição dos resultados seja um fator confundidor dos estudos que envolvem exercício e DP. A cartilha mais atual da Nature sobre a doença de Parkinson (POEWE, 2017) reconhece o número constante e crescente de estudos e publicações envolvendo exercícios e Parkinson.

Moraes Filho et al (2020) demonstraram que a força dos membros inferiores foi capaz de alterar sintomas motores e funcionais, como a bradicinesia. Mollinedo (2007) utilizou o T30 para investigar alterações no equilíbrio dinâmico. Skinner et al (2019) destaca que a força de MMII é fundamental para o controle motor, mobilidade e qualidade de vida dos indivíduos.

Ao levar em consideração as atividades da vida diária, é notável identificar que para a execução de algumas tarefas como andar, vestir-se, dirigir automóveis, subir escadas, etc. sendo a força de MMII diretamente envolvida (JONES, RIKLI E BOEM, 1999).

De uma forma geral, os testes funcionais são indicativos que tentam, em conjunto, expressar em escores o nível de autonomia ou incremento motor para o indivíduo. Alguns autores trabalham com a perspectiva do *Minimal Detectable Change* (mínimo de

mudança detectável), que é a relevância da alteração avaliada pelos testes funcionais. Isso significa que resultados de testes funcionais podem ter valores alterados por questões cotidianas, sem relação direta com a progressão da doença. Nesse sentido, é importante uma alta confiabilidade de cada um dos testes funcionais utilizados.

4.2 Aprendizado de máquina

Aprendizado de máquina (AM) pode ser entendido como uma área e ferramenta da Inteligência Artificial capaz de aprender padrões a partir de conjuntos de dados. A partir da identificação destes padrões, a técnica possibilita que os dados analisados sejam classificados em etiquetas pré-definidas ou não. As técnicas mais modernas são capazes de generalizar estes modelos, permitindo sua aplicação em novos dados, ou novas situações (RASHCKA, 2015; ZHOU, 2012; HANDELLMAN, 2018).

Handelman et al (2019) chamam atenção para a especificação dos conceitos. Os autores discutem que IA tem uma abrangência maior, e diz respeito à automação das tomadas de decisão feitas por máquinas; no aprendizado de máquina, no entanto, é dado ao computador a função de aprender, sem que o processo seja detalhadamente programado (SAMUEL, 1959).

O AM se dá pelo desenvolvimento de um modelo (algoritmo ou função) que aprende de forma automática através do registro e da avaliação das próprias experiências (MITCHELL, 1977). Souza (2020) explica que no contexto de máquinas, aprender significa otimizar um modelo para que este seja capaz de classificar, por exemplo, um conjunto de dados. Quando o que se busca com o aprendizado de máquinas consiste em uma classificação, Zhou (2012) chama o algoritmo desenvolvido como classificador.

Apesar de envolver tecnologia avançada, o assunto não é recente. Em 1950, na famosa publicação do Jogo da Imitação², Alan Turing já falava sobre o aprendizado de máquinas, no mesmo trabalho em que ainda se discutia sobre a possibilidade da criação de computadores digitais (TURING, 1950). O primeiro programa de autoaprendizagem, no entanto, é atribuído a Arthur L. Samuel, que elaborou uma modelo capaz de aprender movimentações em um jogo de damas (SAMUEL, 1959).

O processamento do aprendizado de máquina é complexo, mas o raciocínio central do mecanismo é acessível, e inicia por uma função básica da computação: *if/then* (se/então) (KUBAT, 2015). Para qualquer processamento que se execute, o primeiro passo é sempre a constatação das características do dado de entrada. A partir daí, se (*if*) esse dado de entrada obedece à uma regra estabelecida no código do programa, então (*then*) a saída desse dado vai ser representada por uma classificação positiva; entretanto, se o dado não obedece a esta mesma regra, a saída será uma classificação negativa.

² O Jogo da Imitação é um filme vencedor do Óscar® de Melhor Roteiro Adaptado em 2015, que retrata parte da vida e da obra de Alan Turing.

Ávila (CIÊNCIA DA COMPUTAÇÃO 9, 2020) comenta que a extração de informações mediada pela computação, no início deste tipo de literatura, era relativamente simples uma vez que eram fornecidos exatamente quais valores se desejava extrair, mas o avanço tecnológico trouxe novos desafios e expectativas às possibilidades da computação. A autora explica que um ser humano identifica rapidamente se uma imagem retrata um momento do dia ou da noite. Para que uma máquina faça esta mesma identificação é possível programar para que seja considerado apenas o espectro de cores da imagem para se chegar na classificação. Entretanto, identificar se uma pessoa na imagem é uma criança ou um adulto continua sendo uma tarefa simples para um ser humano, mas bem mais complexa para uma máquina.

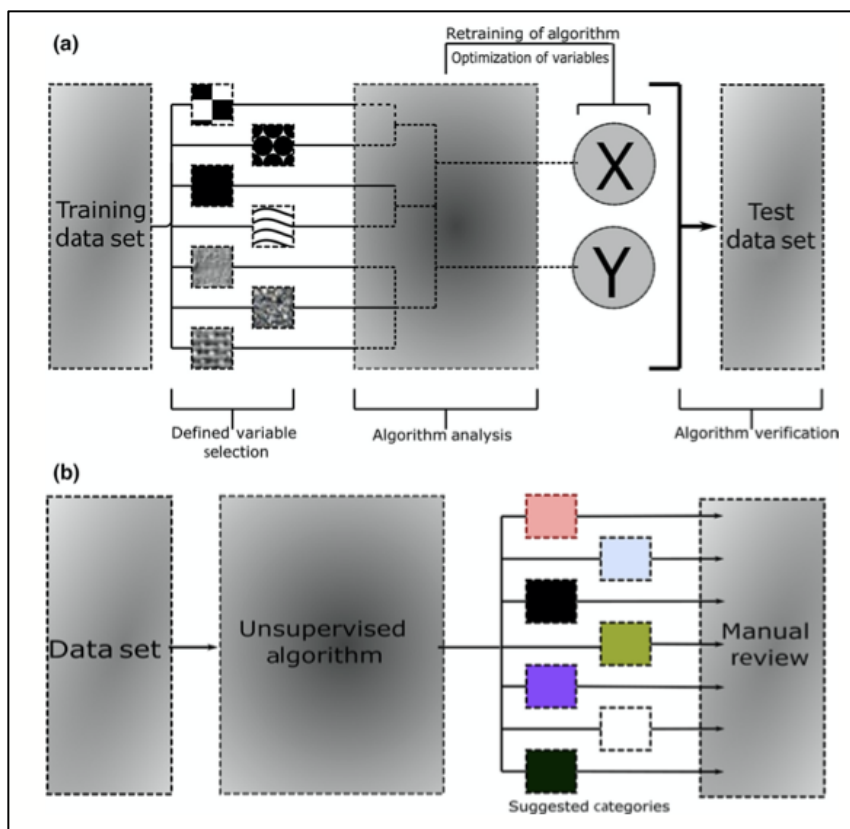
Escrever manualmente todas as regras de uma programação se torna impraticável à medida em que se consideram mais características a serem avaliadas e interpretadas. O processo da aprendizagem de máquina surge, então, como uma alternativa para diminuir o esforço da construção da programação (SAMUEL, 1959). E segundo Raschka (2015), as máquinas aprendem a partir de três formulações básicas: o aprendizado supervisionado, o aprendizado não supervisionado e o aprendizado de reforço.

No aprendizado supervisionado os dados que serão analisados (entradas ou *inputs*) já possuem uma classificação conhecida (saída ou *output*) e o foco é programar condições da máquina aprender a ponto de classificar novas entradas que podem acontecer.

Um bom exemplo para entender o aprendizado supervisionado é a classificação de e-mails como spam. Uma vez que o usuário marca um e-mail ou um emissor como spam, o algoritmo aplicado consegue identificar padrões que auxiliam o sistema a classificar futuros e-mails semelhantes ou do mesmo emissor como spam. Neste exemplo, cada e-mail é uma entrada, ou um dado novo a ser analisado e a marcação *spam* ou *não spam* são etiquetas de classificação (RASCHKA, 2015; HANDELMAN et al, 2018).

No aprendizado não supervisionado é possível considerar o mesmo exemplo da classificação de e-mails, mas sem a determinação prévia de quais e-mails ou emissores seriam spam. Neste processo acontece um agrupamento dos e-mails a partir de características deles mesmos, por exemplo: mensagens escritas, mensagens com fotos, mensagens com links, etc. O método de separação em grupos sem uma prévia classificação é conhecido como clusterização e a classificação desses *clusters* é possível através do aprendizado não supervisionado. Este tipo de aprendizado é capaz de explorar um conjunto de dados sem uma prévia definição e ainda assim extrair classificações e informações relevantes (RASCHKA, 2015; HANDELMAN et al, 2018).

Figura 8 - Diagrama de fluxo do aprendizado de máquina supervisionado e não supervisionado (HANDELMAN et al., 2018).



Na imagem acima, Handelman e colaboradores (2018) trazem em formato de esquema o fluxo dos processos do aprendizado supervisionado (a) e do aprendizado não supervisionado.

No primeiro exemplo (a), são fornecidas ao computador informações das *features* (características) das entradas a serem analisadas e também informações das classificações que se deseja, chamadas de *targets*, representadas na imagem pelas letras X e Y. Uma vez que a técnica computacional é mais eficiente que métodos estatísticos tradicionais por sua capacidade de aprofundamento dos cálculos (RASCHKA, 2015), o destaque principal do aprendizado supervisionado está na identificação das relações entre estes dois conjuntos de dados (HANDELMAN, 2018).

O exemplo (b) traz a sequência do processo do aprendizado não supervisionado. Neste caso, o conjunto de dados é submetido ao algoritmo sem que haja uma prévia classificação. Desta forma, o que o programa devolve (saída) são categorizações sugeridas, determinadas a partir das características dos dados. Assim, o foco do aprendizado não supervisionado é avaliar possíveis agrupamentos naturais, calculados de maneira abstrata e sem um objetivo programado.

Todas as saídas, no entanto, precisam ser validadas por um olhar especialista da área envolvida no caso do AM não supervisionado; o agrupamento dos dados, por si só, pode não significar algo relevante. Bengio (2012) argumenta que os algoritmos do aprendizado de máquina servem para estruturar dados e identificar relações preditivas entre as variáveis. A capacidade de cálculo de um computador associada à capacidade de interpretação dos dados dos especialistas de uma determinada área funcionam como um acelerador no entendimento dos fenômenos, trazendo contribuições relevantes para a prática acadêmica, científica, clínica, etc.

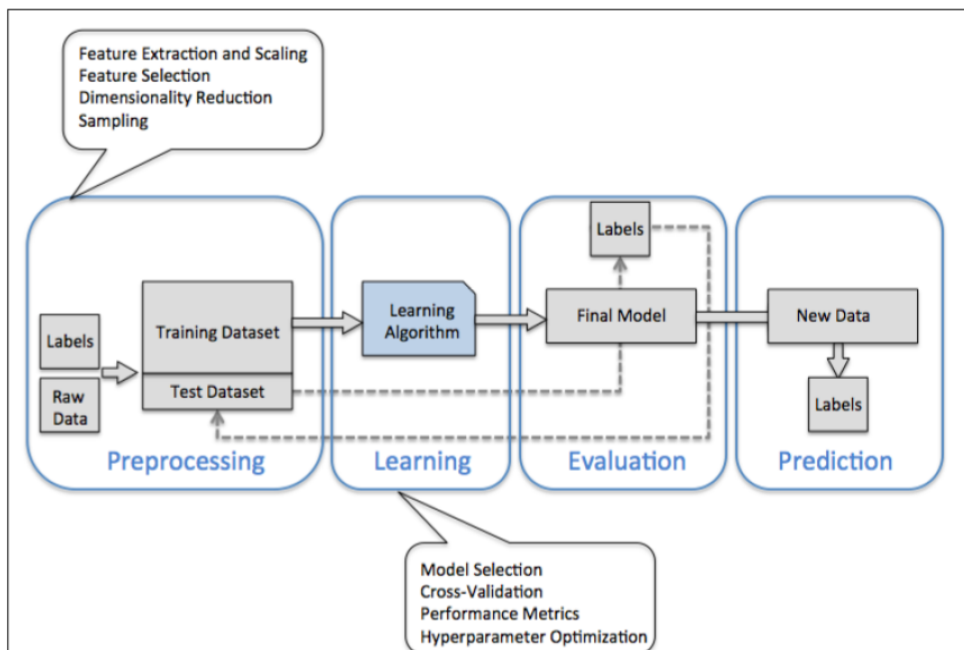
Numa terceira perspectiva, o aprendizado de reforço traz uma certa autonomia para a máquina, através de um sistema capaz de melhorar o seu próprio desempenho, levando em consideração as interações com o ambiente. Isso significa que é possível maximizar o alcance e a capacidade de cálculo de uma máquina através da retroalimentação de informações (RASCHKA, 2015).

Por exemplo: é comum hoje em dia que aplicativos de compras, de música, etc., venham com avaliações sobre a experiência do usuário. Curtir uma música ou bloquear um tipo de conteúdo são feedbacks positivos ou negativos sobre aquela interação. A partir daí, o algoritmo tem condições de agrupar uma resposta individual à outras respostas semelhantes e, por generalização, sugerir novas músicas ou produtos para aquele usuário. É desta forma que acontece o aprimoramento da capacidade preditiva e da autonomia em eventos futuros de um sistema de aprendizado de reforço.

As etapas da construção do algoritmo envolvem critérios éticos para que os resultados do SVM e RUSBoost sejam confiáveis. O fluxo da construção de um algoritmo de aprendizagem de máquina está ilustrado na Figura 9. A primeira etapa consiste no pré-processamento dos dados, ou seja, na organização dos dados brutos, levando em consideração a identificação das *features* que serão processadas. Esta etapa deve ser acompanhado por um especialista da área que abrange o objeto de estudo, e levar em conta as informações disponíveis na respectiva literatura.

O passo seguinte é a separação de dados para treino e dados para teste. Os dados para treino são dados separados que serão utilizados durante o aprendizado da máquina. Nesta etapa é possível modular e corrigir aspectos do algoritmo, chegando assim em um modelo final, que será validado utilizando os dados de teste. Esses dados de teste asseguram que o algoritmo desenvolvido é eficiente em generalizações, incluindo valores não utilizados durante o processo do aprendizado (RASCHKA, 2015).

Figura 9 - Fluxograma para utilização do aprendizado de máquina (RASCHKA, 2015)



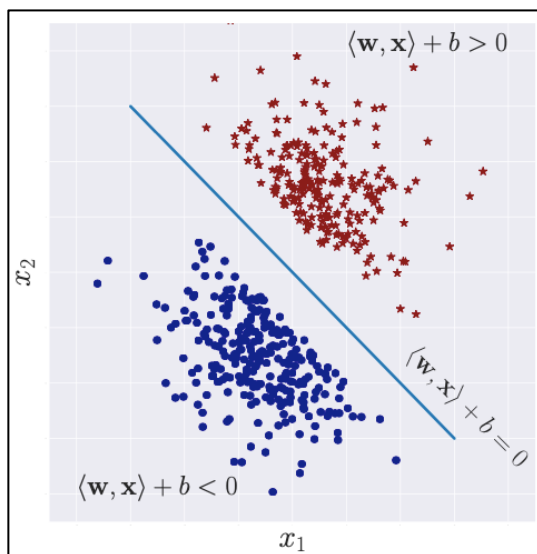
Com o modelo pronto é possível fazer previsões em eventos futuros, ou identificar novos aspectos, mais profundos ou assertivos, a depender do objetivo da construção do algoritmo.

Para cada um dos tipos de aprendizado existem técnicas que foram desenvolvidas para devolver um resultado que expresse a realidade daquilo que se propôs investigar. No âmbito desta pesquisa foram utilizados dois modelos do AM supervisionado, a saber: Support Vector Machine (SVM) e RUSBOOST.

SVM é um classificador elaborado para separar conjuntos de dados em duas classes diferentes, mesmo considerando pequenas variações entre os sinais ou valores de entrada. É possível obter uma regressão através da utilização de SVM, resultando em uma avaliação linear dos dados obtidos (ZHOU, 2012; FIELD, 2009).

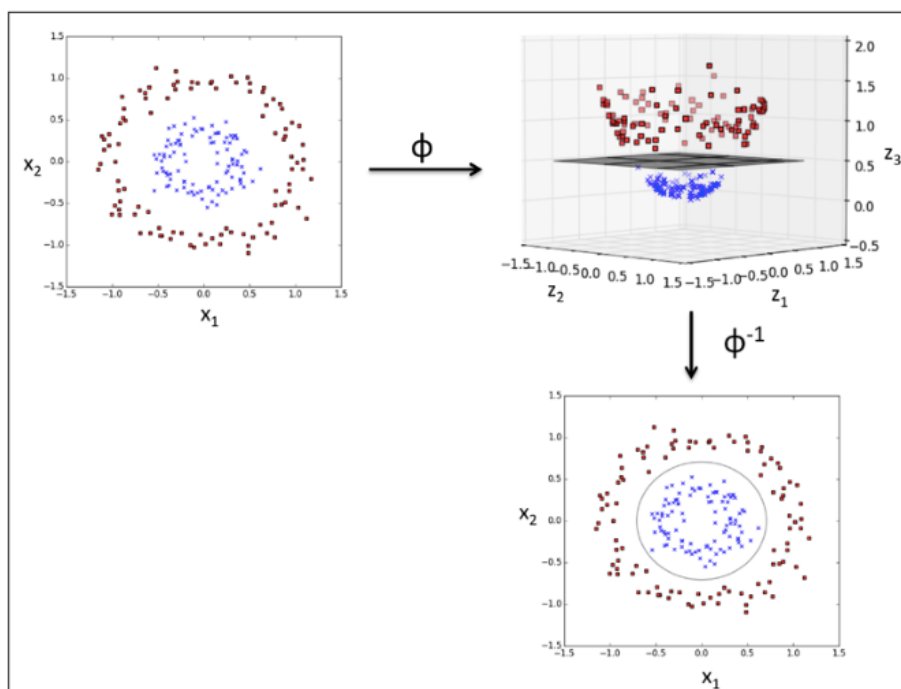
No exemplo a seguir, Souza (2020) calculou um banco de dados aleatório a fim de identificar a separação linear em uma classificação binária. O cálculo separou, sem erros, os dados pertencentes ao grupo das estrelas vermelhas (valor positivo) dos dados não pertencentes ao grupo das estrelas vermelhas (valor negativo), ou grupo de círculos azuis.

Figura 10 - Exemplo de um conjunto de dados linearmente separável (SOUZA, 2020)



Entretanto, dados reais apresentam uma infinidade de variáveis nem sempre controláveis, e muitas vezes os critérios de separação dos grupos não são lineares e necessitam de cálculos mais aprofundados. O objetivo é que o modelo perceba, ou considere, os dados a partir de uma perspectiva espacial com mais de duas dimensões. Para calcular esta função, o Truque de Kernel é um recurso (ZHOU, 2012), apresentado na figura a seguir.

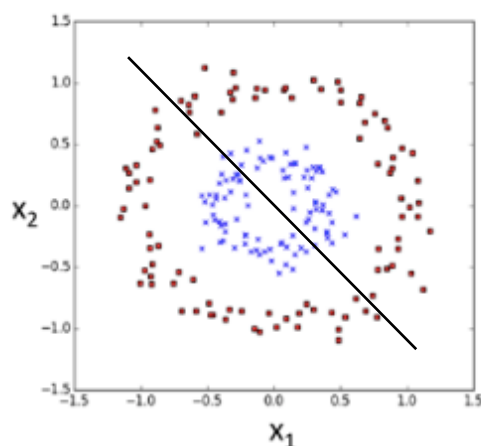
Figura 11 - Aplicação do truque de Kernel em uma situação não linear aleatória (Raschka, 2015)



Ao exemplo da figura 10, alguns conjuntos de dados são denominados não linearmente separáveis. Isto não significa que não existam informações a serem obtidas por percepções diferentes. O truque de Kernel (ou função de Kernel), então, é uma modificação da abordagem do algoritmo, para que seja possível identificar o traço de separação de classes entre os dados avaliados.

Se hipoteticamente fosse aplicado o cálculo utilizado por Souza (2020) – Figura 10, no banco de dados utilizado por Raschka (2015) – Figura 11, provavelmente o modelo não seria considerado um bom classificador, resultando em uma separação não eficiente dos dados analisados.

Figura 12 - Manipulação da imagem de Raschka (2015) para ilustração de uma separação linear em dados não lineares.



Ao passo que SVM é uma abordagem que busca construir um classificador único e assertivo, RUSBoost é um método de *ensembles* (conjuntos) que se utiliza de muitos classificadores mais simplificados para tomadas de decisões individuais, que são combinadas para a classificação final de saída (GALAR, 2012). A técnica do RUSBoost faz uma randomização dos dados de base durante a etapa de treinamento e por isso cada interação pode gerar uma classificação diferente, que serão combinadas resultando em uma única classificação de saída.

Essa reamostragem aleatória diferencia o RUSBoost pela capacidade de impulsionar (*to boost*) dados em menor quantidade. A função é trazer equilíbrio para a análise pela ponderação dos dados disponíveis. Este *ensemble* também pode ser utilizado para aumentar a acurácia alcançada por um único classificador (GALAR, 2012; ZHOU, 2012). A necessidade deste tipo de abordagem se justifica pelas dificuldades do recrutamento de alguns tipos de amostra e sua respectiva influência na classificação desejada.

A avaliação da qualidade de um classificador se dá através da análise dos gráficos de ROC (FAWCETT, 2006). A técnica permite visualizar e avaliar um classificador baseado na sua própria performance, sendo muito utilizada no contexto de avaliação por distribuição por classes. Para cada teste, o classificador tem quatro possibilidades de resposta, conforme disposto abaixo:

| | Total de valores positivos (P) | Total de valores negativos (N) |
|-------------------------------|---|--|
| Classificação positiva | True Positives (TP) (positivos verdadeiros) | False Positives (FP) (positivos falsos) |
| Classificação negativa | False Negatives (FN) (negativos verdadeiros) | True Negatives (TN) (negativos falsos) |

Se um valor positivo é classificado pela máquina como sendo positivo, esta classificação é contada como *true positive* (TP). Seguindo a mesma lógica, é considerada como *false positive* (FP) se um valor negativo é classificado como positivo; *false negative* (FN) consiste em um valor positivo ser considerado negativo; e por fim, *true negative* (TN) são os valores negativos que foram considerados negativos.

A avaliação do classificador, então, se dá a partir de cinco indicadores: acurácia (accuracy), precisão (precision), sensibilidade (sensitivity), especificidade (specificity) e F1-measure (FAWCETT, 2006). As fórmulas e aplicações destes indicadores estão descritos a seguir.

A Acurácia é a medida principal da classificação. Apresenta uma relação geral entre todos os valores e todas as classificações, e se dá pela seguinte fórmula:

$$Acurácia = \frac{TP + TN}{P + N}$$

A Precisão é a proporção de acertos dentre os valores considerados positivos, calculado por:

$$Precisão = \frac{TP}{TP + FP}$$

A Sensibilidade também é chamada de recall e apresenta a relação entre a quantidade de valores positivos preditos em relação ao total de valores positivos:

$$Sensibilidade = \frac{TP}{P}$$

A Especificidade apresenta a relação entre a quantidade de valores negativos preditos em relação ao total de valores negativos. Sua fórmula é:

$$\text{Especificidade} = \frac{N}{FP + TN}$$

Por fim, a *F1-measure* estabelece uma relação entre a Precisão e a Sensibilidade, calculada a partir de:

$$F1 - measure = \frac{2}{1/Precisão + 1/Sensibilidade}$$

A observação dos valores de todas estas métricas é importante para verificar se a alta acurácia de um classificador representa a capacidade de classificação ou se o resultado positivo se deve ao acaso. Medidas desbalanceadas acontecem quando o valor de sensibilidade, por exemplo, é muito alto quando comparado ao valor de especificidade. Pode ser que a máquina tenha classificado muitos valores positivos corretamente, mas não porque aprendeu com os dados de treino, senão por uma assertividade aleatória.

4.2.1 Aprendizado de máquina no contexto da doença de Parkinson

No tocante às áreas que se relacionam nesta pesquisa, Bi et al (2019) explicam que termos técnicos da tecnologia podem dificultar a compreensão por profissionais da saúde e vice-versa. Assim, os autores criaram um glossário de termos do aprendizado de máquina, com sua respectiva equivalência em termos epidemiológicos/estatísticos utilizados pela área da saúde, trazido a seguir de forma adaptada e com tradução livre:

| Termos do aprendizado de máquina | Termos da área da saúde | Considerações |
|--|-------------------------|--|
| Feature (características); <i>attribute, predictor, field</i> . | Variável independente | Chamada também de variável preditora ou de causa (FIELD, 2009). É o detalhamento das características dos inputs. |
| Input (entradas) | Variável independente | Em aprendizado de máquina, input se refere a todos os preditores (ou <i>features</i>) utilizados no modelo a ser calculado. |

| | | |
|-------------------------|--|--|
| Output (saídas) | Variáveis dependentes | São aquelas variáveis afetadas pela(s) variável(eis) independente(s). São consideradas como as classificações a partir dos valores preditos. |
| Label (etiqueta) | Valores das variáveis dependentes (ponto de corte) | A classificação das variáveis dependentes se dá através de <i>labels</i> . Ao considerar que as variáveis dependentes devolvem valores resultantes da influência da variável independente, é necessário classificar, ou dar sentido, aos valores encontrados. Este processo se dá através de uma rotulação dos resultados. |
| Loss Function | Erro de medida | Quando o cálculo de predição construído não apresenta conclusões ou valores que representam a realidade. |

O aprendizado de máquinas vem sendo amplamente utilizado em diferentes aplicações no cenário da saúde. Alguns exemplos são: a predição do diagnóstico de doenças a partir de imagens (KOMURA, 2019); a detecção de metástase de células cancerígenas (BEJNORDI, 2017); a interpretação de exames de eletrocardiograma (DEO, 2015); o monitoramento de saúde mental (GRAHAM, 2020), a predição de acometimentos de natureza cognitiva (BRATIC, 2018); a predição do desenvolvimento de depressão (BYEON, 2020), entre outros.

No contexto da doença de Parkinson, Falchetti et. al (2020) conduziram uma meta-análise para tentar identificar genes preditores de DP; Youngmann et. al (2019) desenvolveram um classificador virtual para identificação da doença a partir do registro da utilização de mouse e teclado dos avaliados. Existem também investigações da progressão da doença através dos valores de citocina no organismo (RASTEGAR et al, 2019), do efeito da doença nas características cognitivas (SALMANPOUR et al 2019), além dos estudos conduzidos de modo a revelar dependências escondidas durante a progressão da doença e o aparecimento dos sintomas (ZHAN et al, 2018).

Em uma outra abordagem, o aprendizado de máquina pode favorecer pesquisas que têm se concentrado na determinação inicial da doença de Parkinson, oportunizando assim uma administração da combinação mais correta e assertiva de medicamento com tratamentos adjuntos. Levando em consideração que o transcurso da doença pode levar muitos anos (POEWE, 2017), um melhor entendimento dos sintomas e sua respectiva relação pode aumentar a qualidade de vida das pessoas com DP.

O Parkinson Progression Markers Initiative (PPMI), por exemplo, é uma iniciativa financiada pela Michael J. Fox Foundation; trata-se de um biorrepositório com dados de

peças com DP. Este banco de dados pode ser solicitado para estudos que se propõem a testar modelos preditivos que contribuam de alguma forma para o entendimento da progressão da doença.

Em 2018, por exemplo, Zhan e colaboradores (2018) conduziram um experimento de avaliação da progressão da DP utilizando smartphones. Através do aplicativo baixado pelos participantes da pesquisa, foram coletadas informações diárias a respeito de cinco *features*: voz, *finger tapping*³, marcha, equilíbrio e tempo de reação. Para comparação, os autores avaliaram as pessoas através da MDS-UPDRS, do TUG e da escala de Hoehn e Yahr;

Os resultados mostraram coerência entre os resultados coletados pelas duas vias. Os autores não só defenderam que o score gerado pelo aplicativo pode ser considerado uma medida confiável para o monitoramento da doença, como identificaram que a variação dos valores, em função da coleta diária de dados, é um aspecto que poderia ser analisado, além dos tradicionais valores coletados pelos exames clínicos e funcionais periódicos.

Uma outra vantagem da utilização de equipamentos de tecnologias é a possibilidade da avaliação em ambientes reais. A diminuição do cenário de laboratório ou clínica em que são feitos os testes e exames tradicionais podem trazer medidas mais reais do dia-a-dia do indivíduo (ZAHN, 2018; DYAGILEV, 2016).

Outro experimento que utilizou sinais coletados de smartphones foi o de Bot et al (2016). Participaram do estudo 9520 sujeitos que completaram formulários online via iPhone. A proposta consistiu em disponibilizar alguns componentes da avaliação clínica para que fossem preenchidos de forma auto referida e também atividades via app para que fossem completadas três vezes por dia, abarcando tarefas de memória, *tapping*, voz e marcha. Todo o estudo foi conduzido virtualmente, sem que os sujeitos tivessem contato presencial com os pesquisadores.

Posteriormente, Abujrida et al (2019) publicaram um outro estudo utilizando estes mesmos dados, desta vez comparando com avaliações motoras realizadas em parte destes sujeitos; foi verificado que é possível criar uma classificação automática da severidade da doença e também utilizar as medidas de avaliação de alterações da marcha.

³ Em tradução livre: bater de dedos. Um teste neurológico em que se pede para o avaliado realizar o movimento de pinça com os dedos indicador e polegar repetidamente. É o quarto item da MDS-UPDRS III e utilizada no contexto da DP por ser uma ferramenta de avaliação da motricidade humana (LI et al, 2020).

Existem muitas características que podem ser observadas com a finalidade de avaliar a qualidade motora de pessoas com Parkinson (REHMAN, 2019). Extrapolar os dados com técnicas avançadas é uma tentativa de distinguir objetivamente características determinantes, identificadoras da doença de Parkinson.

Perto da conclusão deste manuscrito, Fraiwan e Hassanin (2021) publicaram análises de classificadores que buscavam identificar doenças degenerativas neuromusculares utilizando *ensembles*. O estudo buscou identificar padrões de distúrbios na marcha em pessoas com diferentes tipos de doenças neurais, incluindo a doença de Parkinson.

Os sinais foram coletados durante a marcha e classificadores como o RUSBoost demonstram ser excelentes máquinas de classificação deste tipo de aspecto. As métricas de Fawcett (2006) apresentadas no estudo indicaram que este tipo de Inteligência Artificial é promissor na investigação clínica deste tipo de doença.

5 MÉTODOS

5.1 Delineamento da pesquisa

De acordo com Hulley (2015), trata-se de pesquisa secundária, já que serão utilizados dados previamente existentes; o estudo tem corte transversal visto que as características de causa permanente (doença de Parkinson) e os respectivos efeitos (alterações motoras) estavam presentes ao mesmo tempo no momento da coleta (CAMPANA, 1999).

Serão considerados para esta pesquisa os dados coletados no Programa de Atividade Física para Pessoas com Doenças Neurodegenerativas (PAFPDN) entre os anos de 2012 e 2013. O PAFPDN foi fomentado pelo Programa de Pós-Graduação *Stricto Sensu* em Educação Física da Faculdade de Educação Física da Universidade de Brasília – PPGEF/FEF/UnB.

Para simplificação do entendimento e unificação da linguagem serão considerados nesta pesquisa:

- **Dados de entrada:** todas as características da amostra (idade, altura, massa), bem como os escores dos testes funcionais e aspectos clínicos, que sob o ponto de vista do objetivo desta pesquisa compõem as variáveis independentes; variáveis previsoras.
- **Dados de saída:** as classificações resultantes da análise dos dados de entrada.

O desenvolvimento do algoritmo de aprendizado de máquina para calcular a hipótese deste trabalho foi desenvolvido em parceria com o Programa de Pós-Graduação em Engenharia Biomédica do Campus Gama da Universidade de Brasília – FGA/UnB. Desta maneira, os aspectos técnicos do desenvolvimento e testagem do algoritmo serão tratados neste manuscrito de forma simplificada, sendo a interpretação das respostas dos cálculos o objetivo principal desta investigação.

Cabe registrar que o delineamento desta pesquisa foi diretamente afetado pela eclosão da pandemia *Sars-Cov-2*. A impossibilidade de novas coletas foi estabelecida como medida de segurança contra o coronavírus (COVID-19). Uma pesquisa secundária pode apresentar fragilidades decorrentes do não controle da investigação sobre a coleta de dados, entretanto, esta tese respalda-se na observação de Hulley et. al (2015) que defendem que estudos prévios conduzidos pela mesma instituição, e neste caso pelo

mesmo investigador, são considerados “uma das fontes mais ricas de de dados secundários”. Desta forma, a fim de minimizar – ou explicitar – os fatores confundidores, faz-se necessária a descrição da coleta de dados que compuseram os dados de entrada do presente estudo.

5.2 Levantamento de dados

Os testes funcionais escolhidos para integrar o banco de dados desta investigação foram aqueles com maior frequência de aplicação durante o período considerado. A relação dos testes funcionais, bem como sua descrição e aplicação estão na Tabela 1.

Tabela 1 - Testes funcionais utilizados no estuado e suas respectivas variáveis para análise

| Instrumento | Descrição | Autoria | Variáveis avaliadas no contexto da doença de Parkinson | Referências |
|--|---|-------------------------------|---|---------------------|
| TUG Timed Up and Go Test | É coletado quantos segundos a pessoa leva para levantar-se, caminhar por 3 metros, retornar e se sentar novamente. | Podsiadlo e Richarlison, 1991 | Mobilidade funcional | Morris et al., 2001 |
| | | | Bradicinesia da marcha | Moraes Filho, 2013 |
| T30 Teste de sentar e levantar | Conta-se quantas vezes a pessoa consegue se levantar e se sentar de um banco durante 30 segundos. | Jones, Rikli e Beam, 1999 | Força dos membros inferiores | Marcelino, 2017 |
| TMW Ten meters walk test | É coletado em quanto tempo a pessoa caminha dez metros. O percurso total do teste tem 16m, visto que os primeiros e últimos 3m são desconsiderados pela variação da aceleração e desaceleração. | Bohannon et al., 1996 | Deambulação (qualidade de marcha) | Opara, 2017 |
| | | | Velocidade de marcha | Homem, 2016 |
| UPDRS II | Questionário autorreferido das atividades da vida diária. | Fahn e Elton, 1987 | Mobilidade corporal e atividades da vida diária | Opara, 2017 |

Em relação às características clínicas, foram consideradas para esta pesquisa a escala de Hoehn e Yahr e a pontuação da MDS-UDPRS III. Sobre as demais características das pessoas admitidas no programa, além do diagnóstico da DP, todas as pessoas não deviam apresentar nenhuma outra doença neurológica, problemas ortopédicos ou cardíacos que poderiam afetar a mecânica ou sua segurança ao caminhar.

Todos os testes foram conduzidos na fase *on* de medicamentos antiparkinsonianos. Foi verificado também a pontuação no Mini Exame do Estado Mental (MEEM) como critério de exclusão dos estudos, caso o resultado fosse indicativo de demência ou outra condição cognitiva deletéria.

Como estratégia de recrutamento de amostra, à época foram veiculadas divulgações sobre o Programa e as pesquisas abrangidas. Foram afixados cartazes nos murais do Campus Darcy Ribeiro da Universidade de Brasília, do Hospital Universitário de Brasília e seus arredores. Também ocorreram chamadas em rádios e telejornais de grande audiência na cidade.

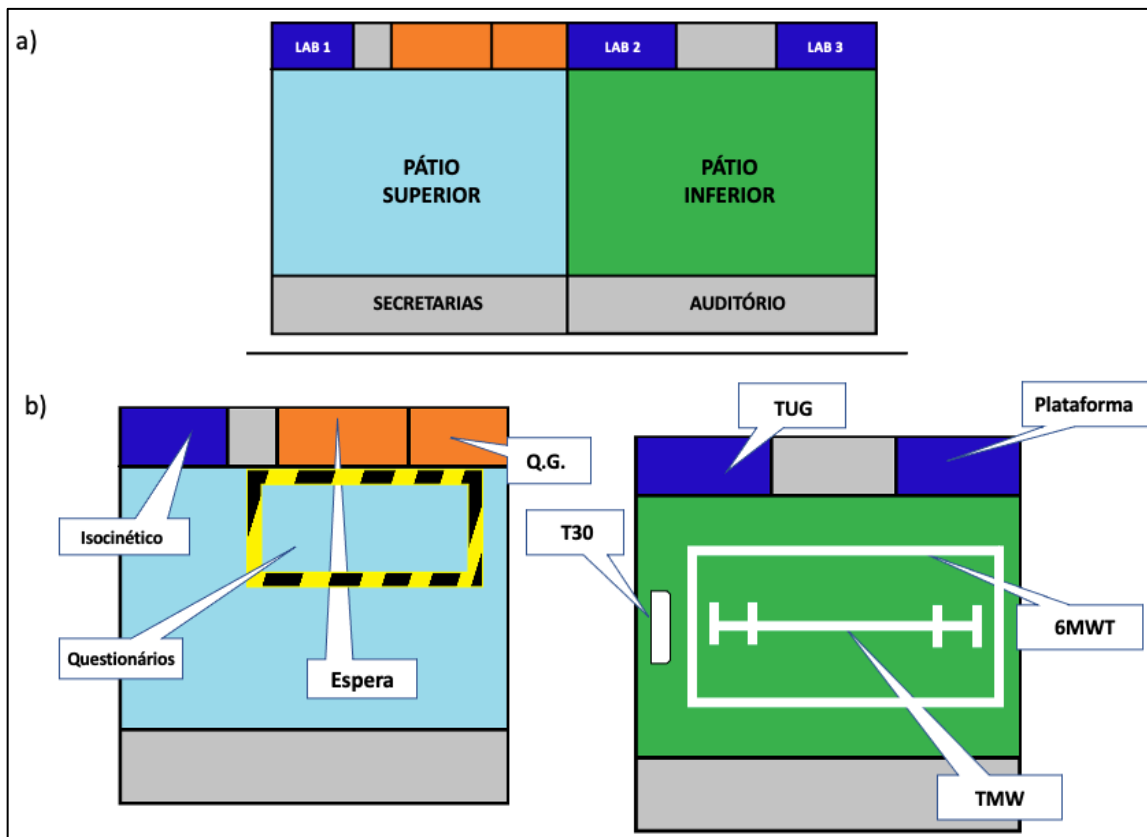
Foram admitidas pessoas sem restrição de gênero, residentes do Distrito Federal, em amostragem por conveniência. A baixa restrição no processo de recrutamento da amostra se deu em função da dificuldade do recrutamento e da adesão de pessoas com DP nas experiências precedentes do Grupo de Estudos em Atividade Física, Genética e Doenças Crônicas Não Transmissíveis, responsável pela gestão do PAFPDN.

As coletas de dados aconteceram no pátio da Faculdade de Educação Física (FEF) da UnB, e contavam com uma quantidade grande de voluntários e pesquisadores. Os testes foram divididos em três baterias, significando que cada voluntário precisava comparecer à Faculdade em três dias diferentes.

A bateria 1 era composta pelos testes funcionais TUG, T30 e TMW, além do questionário sobre qualidade de vida de pessoas com Parkinson (PDQL-BR) e do Inventário de Depressão de Beck (BDI). A bateria 2 era constituída pelos testes de força de membros inferiores, aferida por um dinamômetro isocinético, e de membros superiores, aferida por um dinamômetro de preensão manual. A bateria 3 era composta pela avaliação do equilíbrio, feita por uma plataforma de força e também pela avaliação motora dos sujeitos voluntários (MDS-UPDRS III).

Durante o período de espera para a realização dos testes em cada uma das baterias, eram coletados os dados relacionados ao MEEM, ao relatório de atividade física (IPAQ), ao relatório de atividades da vida diária (MDS-UPDRS II), aos dados antropométricos (massa, altura) e ao registro da administração medicamentosa. A estrutura física da FEF precisava ser adaptada durante o período de coletas, de acordo com a figura a seguir:

Figura 13 - Distribuição logística dos locais de testes.



a) Representação esquemática da planta baixa da FEF, considerando os locais a serem utilizados para a coleta de dados. b) Alocação dos diferentes testes nos espaços físicos da FEF.

A avaliação motora (UPDRS 3) foi realizada na sala de apoio do PAFPDN, localizada no Centro Olímpico da UnB. A equipe de pesquisadores envolvia estudantes de pós-graduação em níveis de mestrado e doutorado, estagiários da graduação e professores do quadro da FEF e do PPGEF/FEF. As avaliações aconteciam nos intervalos dos períodos letivos, para que fosse possível utilizar os espaços comuns da faculdade.

A classificação de Hoehn e Yahr era geralmente compilada dos laudos médicos entregues como parte dos documentos necessários para a inscrição. A classificação de Hoehn e Yahr era feita por um dos pesquisadores responsáveis pelo PAFPDN caso não constasse no relatório médico. A avaliação da MDS-UPDRS III foi conduzida pelos pesquisadores do programa. Cabe destacar que os testes e demais avaliações foram orientados por uma neurologista, membro do grupo por determinado período, além do material de apoio fornecido pela *Movement Disorders Society*.

5.3 Curadoria dos dados

De acordo com Raschka (2015), antes da aplicação dos cálculos ou do desenvolvimento do algoritmo de aprendizagem faz-se necessário qualificar os dados brutos que serão analisados. Quanto mais características forem informadas na entrada do processamento dos dados, mais profundas e assertivas serão as classificações de saída.

Os dados contidos nas planilhas arquivadas foram organizados em uma única planilha. Os valores relacionados aos aspectos funcionais e às características da amostra foram checados com arquivos de um outro pesquisador do PAFPDN. Os dados relacionados aos aspectos clínicos (UPDRS III) foram conferidos com os arquivos físicos. Valores divergentes eram corrigidos, quando possível, ou excluídos nos demais casos.

O PAFPDN, por se tratar de um programa estruturado a partir de pesquisas experimentais, avaliava de maneira constante os participantes que frequentavam as atividades oferecidas, como aulas de dança de salão (VIDAL, 2014), musculação (FISCHER, 2014; MORAES FILHO, 2014; SANTOS, 2014), equoterapia (HOMEM, 2014), etc. As avaliações eram feitas pré- e pós-intervenção, e também foram feitas avaliações durante o período de destreino (após as férias, por exemplo).

Levando em consideração que o foco desta pesquisa é a investigação da relação entre os aspectos funcionais e os aspectos clínicos, foram considerados dados coletados das mesmas pessoas em períodos diferentes. Isto respalda-se pela identificação de variação nos valores brutos das planilhas e pela manifestação diversificada da doença, apresentada no item 4.1.2 desta tese. Ao final da curadoria, foram considerados 203 conjuntos dados (coletas transversais das variáveis clínicas e funcionais) provenientes de 112 pessoas com Parkinson avaliadas pelo PAFPDN entre 2012 e 2013.

5.4 Análise dos dados

Foram utilizados dois classificadores de SVM e dois classificadores de RUSBoost para tentar detectar o nível, em pontos, da UPDRS III, a partir da combinação das variáveis descritas a seguir. Em todos os casos, foram separados aleatoriamente 60% dos dados para treinamento e 40% dos dados para testes.

Em todas os classificadores foi realizada a divisão dos dados para treino e testes de forma aleatória, obedecendo a proporção de 60% dos dados para treinamento e 40% dos dados para teste. No processo de divisão da amostra, foram consideradas as múltiplas entradas de uma mesma pessoa. Neste sentido, o algoritmo foi configurado para alocar

todas os dados referentes da mesma pessoa no mesmo grupo, seja ele o de treinamento ou de teste.

A função de data augmentation foi construída para criar dados capazes de equilibrar as classes, a partir de valores de referência das variáveis utilizadas. Num primeiro momento, houve a identificação da maior quantidade de dados – se positivos ou negativos, e valores os adicionados consideram a média de casos minoritários, gerando dados com ruído. A partir de uma quantidade menor de dados, a estratégia serve para aumentar a quantidade de dados e aumentar a acurácia da avaliação. É importante ressaltar que foi utilizado data augmentation apenas nos dados para treino. Não foram feitas alterações ou manipulações de qualquer tipo nos dados de teste.

Nos classificadores do tipo SVM, foram rodados 1000 trials na configuração normal e 10000 trials na configuração com data augmentation. Os dois classificadores do tipo RUSBoost utilizaram 10 learners para todas as tentativas, mantendo as 1000 trials para a configuração normal e 10000 com data augmentation.

A avaliação dos classificadores se deu pela análise das métricas que apresentam valores de acurácia, sensibilidade, especificidade, precisão e F1-measure (FAWCETT, 2006). A escolha da combinação de variáveis tentou atender aos objetivos dispostos neste trabalho, de acordo com os dados disponíveis no banco de dados utilizados e com indicações da literatura. As combinações também foram numeradas por ordem de teste e serão referidas desta maneira, para facilitar o entendimento ao longo do texto. As informações estão relacionadas abaixo:

| OBJETIVO | TESTE | COMBINAÇÃO DE VARIÁVEIS DE ENTRADA |
|---|---------|--|
| Aspectos funcionais | TESTE 1 | TUG, TMW e T30. |
| Aspectos funcionais e clínicos | TESTE 2 | TUG, TMW, T30 e H&Y. |
| Aspectos funcionais, clínicos e características individuais | TESTE 3 | TUG, TMW, T30, UPDRS II, tempo de diagnóstico e idade. |
| | TESTE 4 | TUG, TMW, T30, tempo de diagnóstico e idade. |
| | TESTE 5 | TUG, TMW, T30, tempo de diagnóstico, idade, massa e estatura. |
| | TESTE 6 | TUG, TMW, T30, tempo de diagnóstico, idade, massa, estatura e H&Y. |
| | TESTE 7 | TUG, TMW, T30, tempo de diagnóstico, idade e H&Y. |
| Aspectos clínicos e individuais | TESTE 8 | Tempo de diagnóstico, idade, massa e estatura. |
| | TESTE 9 | Tempo de diagnóstico, idade, massa, estatura e H&Y. |

6 RESULTADOS

A estatística descritiva de caracterização da amostra presente no banco de dados utilizado está apresentada a seguir. A Tabela 2 organiza os dados categóricos, indicando a porcentagem de cada valor e a tabela 3 apresenta os valores mínimos, máximos, média e desvio padrão das demais características pessoais.

Tabela 2 - Caracterização da amostra: gênero e classificação de Hoehn e Yahr.

| Características da amostra | N | % |
|-----------------------------------|----------|----------|
| Participantes | | |
| N Total | 113 | 100 |
| Homens | 82 | 72,56 |
| Mulheres | 31 | 27,44 |
| Hoehn & Yahr | | |
| Sem informação | 46 | 40,7 |
| 1 | 14 | 12,38 |
| 1,5 | 18 | 15,92 |
| 2 | 25 | 22,12 |
| 2,5 | 11 | 9,73 |
| 3 | 16 | 14,15 |
| 4 | 15 | 13,27 |
| 5 | 1 | 0,08 |

Tabela 3 - Caracterização da amostra: valores antropométricos e tempo de diagnóstico.

| | Mínimo | Máximo | Média \pm Desvio Padrão |
|-----------------------------|--------|--------|---------------------------|
| Idade (anos) | 36 | 90 | 64,80 \pm 10,18 |
| Estatura (m) | 1,44 | 1,83 | 1,65 \pm 0,08 |
| Massa (kg) | 44,4 | 110 | 73,65 \pm 13,21 |
| Tempo de diagnóstico (anos) | 0,8 | 27 | 6,27 \pm 4,67 |

Os resultados dos treinos e testes realizados pelos diferentes classificadores estão apresentados a seguir. Cada combinação de *features* foi testada com cinco valores de referência diferentes: 15, 17, 18, 19 e 21. Estes valores de referência dizem respeito à maior quantidade de dados disponíveis para os cálculos. Em testes iniciais durante o desenvolvimento do algoritmo foi percebido que em pontuações mais altas ou mais baixas que os valores de referência apresentados, os retornos dos indicadores retornavam valores absolutos (1 e 0), indicando *over fitting* e a perda da sensibilidade na classificação.

Os resultados estão agrupados de acordo com os objetivos, e para facilitar a identificação daqueles classificadores que apresentaram os melhores resultados, estes foram destacados da seguinte forma:

- Em cinza: valores $> 0,5$ em todos os indicadores
- Em azul: valores $> 0,7$ em todos os indicadores

6.1 Classificação a partir de aspectos funcionais

Considerando o TUG, TMW e T30 como aspectos funcionais para a classificação da doença, foram calculados os valores dos testes. O RUSBoost foi o classificador com maior quantidade de acertos entre os outros classificadores nesta combinação de features, e também o que apresentou a melhor classificação, utilizando 19 como valor de referência, conforme mostra a tabela a seguir:

Tabela 4 – TESTE 1: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW, T30)

| TESTE 1 | | TUG, TMW, T30 | | | | |
|---------------------|--------------------------------|-----------------------|----------|---------------|----------------|------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,64 | 0,62 | 0,92 | 0,26 | 0,74 |
| | SVM com data augmentation | 0,54 | 0,72 | 0,49 | 0,63 | 0,58 |
| | RUSBoost | 0,52 | 0,79 | 0,41 | 0,76 | 0,54 |
| | RUSBoost com data augmentation | 0,47 | 0,56 | 0,56 | 0,33 | 0,56 |
| UPDRS 17 | SVM | 0,62 | 0,67 | 0,59 | 0,65 | 0,63 |
| | SVM com data augmentation | 0,53 | 0,52 | 0,65 | 0,40 | 0,58 |
| | RUSBoost | 0,54 | 0,50 | 0,58 | 0,52 | 0,54 |
| | RUSBoost com data augmentation | 0,58 | 0,50 | 0,81 | 0,41 | 0,62 |
| UPDRS 18 | SVM | 0,45 | 0,71 | 0,24 | 0,83 | 0,36 |
| | SVM com data augmentation | 0,61 | 0,50 | 0,67 | 0,58 | 0,57 |
| | RUSBoost | 0,56 | 0,56 | 0,56 | 0,56 | 0,56 |
| | RUSBoost com data augmentation | 0,48 | 0,54 | 0,52 | 0,43 | 0,53 |
| UPDRS 19 | SVM | 0,59 | 0,25 | 0,13 | 0,81 | 0,17 |
| | SVM com data augmentation | 0,46 | 0,20 | 0,25 | 0,56 | 0,22 |
| | RUSBoost | 0,71 | 0,69 | 0,79 | 0,64 | 0,73 |
| | RUSBoost com data augmentation | 0,42 | 0,33 | 0,19 | 0,65 | 0,24 |
| UPDRS 21 | SVM | 0,69 | 0,00 | 0,00 | 0,94 | 0,00 |
| | SVM com data augmentation | 0,52 | 0,00 | 0,00 | 0,72 | 0,00 |
| | RUSBoost | 0,70 | 0,50 | 0,54 | 0,77 | 0,52 |
| | RUSBoost com data augmentation | 0,61 | 0,47 | 0,41 | 0,72 | 0,44 |

6.2 Classificação a partir de aspectos funcionais e clínicos

Em adição aos aspectos funcionais TUG, TWM e T30, foi calculado a capacidade de classificação da máquina considerando o estágio de Hoehn e Yahr. A técnica de SVM com data augmentation apresentou um dos melhores resultados. Todos os indicadores estão acima de 75%, com destaque para a precisão da máquina no valor de referência 15.

Tabela 5 – TESTE 2: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30) e aspectos clínicos (Hoehn e Yahr)

| TESTE 2 | | TUG, TMW, T30 e Hoehn e Yahr | | | | |
|---------------------|--------------------------------|------------------------------|----------|---------------|----------------|------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,71 | 0,81 | 0,76 | 0,57 | 0,79 |
| | SVM com data augmentation | 0,79 | 0,94 | 0,77 | 0,86 | 0,85 |
| | RUSBoost | 0,48 | 0,60 | 0,38 | 0,63 | 0,46 |
| | RUSBoost com data augmentation | 0,66 | 0,77 | 0,59 | 0,75 | 0,67 |
| UPDRS 17 | SVM | 0,57 | 0,57 | 0,47 | 0,67 | 0,52 |
| | SVM com data augmentation | 0,53 | 0,45 | 0,63 | 0,45 | 0,53 |
| | RUSBoost | 0,50 | 0,58 | 0,35 | 0,69 | 0,44 |
| | RUSBoost com data augmentation | 0,56 | 0,67 | 0,50 | 0,64 | 0,57 |
| UPDRS 18 | SVM | 0,60 | 0,64 | 0,36 | 0,81 | 0,46 |
| | SVM com data augmentation | 0,57 | 0,53 | 0,69 | 0,47 | 0,60 |
| | RUSBoost | 0,71 | 0,79 | 0,58 | 0,85 | 0,67 |
| | RUSBoost com data augmentation | 0,60 | 0,60 | 0,57 | 0,64 | 0,59 |
| UPDRS 19 | SVM | 0,72 | 0,50 | 0,22 | 0,91 | 0,31 |
| | SVM com data augmentation | 0,56 | 0,33 | 0,06 | 0,92 | 0,10 |
| | RUSBoost | 0,56 | 0,44 | 0,36 | 0,69 | 0,40 |
| | RUSBoost com data augmentation | 0,60 | 0,39 | 0,70 | 0,56 | 0,50 |
| UPDRS 21 | SVM | 0,72 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,74 | 0,00 | 0,00 | 0,95 | 0,00 |
| | RUSBoost | 0,65 | 0,32 | 0,70 | 0,63 | 0,44 |
| | RUSBoost com data augmentation | 0,43 | 0,55 | 0,46 | 0,38 | 0,50 |

6.3 Classificação a partir de aspectos funcionais, clínicos e características individuais

Neste primeiro teste, foram consideradas as features: idade, tempo de diagnóstico, TUG, TMW, T30 e UPDRS II; O SVM com data augmentation apresentou valores acima de 50% nos indicadores do valor de referência 15, mas não se trata de uma medida positiva. É possível que a máquina tenha realizado uma classificação abstrata demais, pouco sensível em testes em populações maiores.

O RUSBoost com saída de UPDRS valor 21 retornou valores entre 50% e 70%, não demonstrando assim força suficiente para ser considerado um bom classificador. Em todos os valores de referência, o RUSBoost com data augmentation apresentou os piores indicadores na avaliação de ROC.

Tabela 6 – TESTE 3: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW, T30 e UPDRS II), aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade)

| TESTE 3 | | TUG, TMW, T30, UPDRS II, tempo de diagnóstico e idade | | | | |
|---------------------|--------------------------------|---|----------|---------------|----------------|------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,50 | 0,53 | 0,80 | 0,13 | 0,64 |
| | SVM com data augmentation | 0,62 | 1,00 | 0,50 | 1,00 | 0,67 |
| | RUSBoost | 0,59 | 0,61 | 0,73 | 0,42 | 0,67 |
| | RUSBoost com data augmentation | 0,26 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 17 | SVM | 0,47 | 1,00 | 0,06 | 1,00 | 0,11 |
| | SVM com data augmentation | 0,58 | 0,54 | 0,44 | 0,70 | 0,48 |
| | RUSBoost | 0,53 | 0,25 | 0,14 | 0,75 | 0,18 |
| | RUSBoost com data augmentation | 0,50 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 18 | SVM | 0,56 | 0,67 | 0,31 | 0,83 | 0,42 |
| | SVM com data augmentation | 0,45 | 0,00 | 0,00 | 0,69 | 0,00 |
| | RUSBoost | 0,47 | 0,56 | 0,45 | 0,50 | 0,50 |
| | RUSBoost com data augmentation | 0,60 | 0,40 | 0,40 | 0,70 | 0,40 |
| UPDRS 19 | SVM | 0,74 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,69 | 1,00 | 0,17 | 1,00 | 0,29 |
| | RUSBoost | 0,75 | 0,33 | 0,25 | 0,88 | 0,29 |
| | RUSBoost com data augmentation | 0,80 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 21 | SVM | 0,96 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,50 | 0,00 | 0,00 | 0,80 | 0,00 |
| | RUSBoost | 0,70 | 0,50 | 0,70 | 0,50 | 0,58 |
| | RUSBoost com data augmentation | 0,75 | 0,00 | 0,00 | 1,00 | 0,00 |

Ainda no mesmo objetivo, foram calculadas a idade e o tempo diagnóstico junto com os aspectos funcionais (TUG, TMW e T30) mas sem a utilização dos dados da UPDRS II. Apesar de ser um questionário que traz informações relevantes no que tange as atividades da vida diária, a quantidade de dados disponíveis dessa feature é pequena. Isso pode ter restringido a capacidade de classificação das máquinas utilizadas nos cálculos apresentados anteriormente.

Mais uma vez o SVM com data augmentation foi o classificador que apresentou os maiores acertos quando considerada o valor de referência 15. Com 0,95 de precisão e 0,9 de especificidade, a F1-measure resultou 0,82, indicando uma boa relação entre todas as variáveis.

Tabela 7 – TESTE 4: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade)

| TESTE 4 | | TUG, TMW, T30, tempo de diagnóstico e idade | | | | |
|---------------------|--------------------------------|---|----------|---------------|----------------|------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,58 | 0,61 | 0,79 | 0,25 | 0,69 |
| | SVM com data augmentation | 0,78 | 0,95 | 0,73 | 0,91 | 0,83 |
| | RUSBoost | 0,44 | 0,38 | 0,45 | 0,43 | 0,42 |
| | RUSBoost com data augmentation | 0,29 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 17 | SVM | 0,52 | 0,37 | 0,70 | 0,43 | 0,48 |
| | SVM com data augmentation | 0,62 | 0,53 | 0,57 | 0,65 | 0,55 |
| | RUSBoost | 0,34 | 0,34 | 1,00 | 0,00 | 0,51 |
| | RUSBoost com data augmentation | 0,52 | 0,44 | 0,88 | 0,31 | 0,58 |
| UPDRS 18 | SVM | 0,63 | 0,60 | 0,25 | 0,89 | 0,35 |
| | SVM com data augmentation | 0,63 | 0,67 | 0,36 | 0,85 | 0,47 |
| | RUSBoost | 0,64 | 0,50 | 0,83 | 0,52 | 0,63 |
| | RUSBoost com data augmentation | 0,62 | 0,50 | 0,60 | 0,63 | 0,55 |
| UPDRS 19 | SVM | 0,75 | 0,57 | 0,33 | 0,91 | 0,42 |
| | SVM com data augmentation | 0,50 | 0,07 | 0,50 | 0,50 | 0,12 |
| | RUSBoost | 0,73 | 0,25 | 0,25 | 0,83 | 0,25 |
| | RUSBoost com data augmentation | 0,60 | 0,38 | 0,38 | 0,71 | 0,38 |
| UPDRS 21 | SVM | 0,82 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,74 | 0,00 | 0,00 | 0,90 | 0,00 |
| | RUSBoost | 0,32 | 0,10 | 0,50 | 0,30 | 0,16 |
| | RUSBoost com data augmentation | 0,65 | 0,00 | 0,00 | 1,00 | 0,00 |

Foi investigado também se características físicas, como massa e estatura, estariam relacionadas na determinação da progressão da doença. A tabela abaixo traz a configuração das features Idade, tempo de diagnóstico, estatura, massa, TUG, TMW e T30. Os resultados mais interessantes foram: RUSBoost com valor 15 e SVM com data augmentation e valor 18.

Tabela 8 – TESTE 5: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade, massa e estatura)

| TESTE 5 | | TUG, TMW, T30, tempo de diagnóstico, idade, massa e estatura | | | | |
|----------------------------|--------------------------------|---|-----------------|----------------------|-----------------------|-------------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,56 | 0,56 | 0,93 | 0,08 | 0,70 |
| | SVM com data augmentation | 0,46 | 0,39 | 0,78 | 0,27 | 0,52 |
| | RUSBoost | 0,61 | 0,63 | 0,85 | 0,23 | 0,72 |
| | RUSBoost com data augmentation | 0,73 | 0,83 | 0,76 | 0,67 | 0,79 |
| UPDRS 17 | SVM | 0,61 | 1,00 | 0,13 | 1,00 | 0,24 |
| | SVM com data augmentation | 0,52 | 0,45 | 0,71 | 0,37 | 0,56 |
| | RUSBoost | 0,56 | 0,00 | 0,00 | 1,00 | 0,00 |
| | RUSBoost com data augmentation | 0,63 | 0,53 | 0,80 | 0,50 | 0,64 |
| UPDRS 18 | SVM | 0,65 | 1,00 | 0,10 | 1,00 | 0,18 |
| | SVM com data augmentation | 0,82 | 0,60 | 0,75 | 0,85 | 0,67 |
| | RUSBoost | 0,35 | 0,25 | 0,29 | 0,40 | 0,27 |
| | RUSBoost com data augmentation | 0,52 | 0,20 | 0,13 | 0,73 | 0,15 |
| UPDRS 19 | SVM | 0,75 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,68 | 0,00 | 0,00 | 0,81 | 0,00 |
| | RUSBoost | 0,77 | 0,55 | 0,75 | 0,77 | 0,63 |
| | RUSBoost com data augmentation | 0,78 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 21 | SVM | 0,73 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,67 | 0,17 | 0,40 | 0,71 | 0,24 |
| | RUSBoost | 0,47 | 0,06 | 0,50 | 0,46 | 0,11 |
| | RUSBoost com data augmentation | 0,76 | 0,00 | 0,00 | 1,00 | 0,00 |

No próximo teste foram consideradas as features: idade, tempo de diagnóstico, estatura, massa, TUG, TMW, T30 e Hoehn e Yahr. Todos os classificadores, entretanto, apresentaram baixa capacidade de identificação, considerando todos os valores de referência e todas os tipos de classificadores.

Tabela 9 – TESTE 6: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico e Escala de Hoehn e Yahr) e características individuais (Idade, massa e estatura)

| TESTE 6 Valor de referência | Classificador | Métricas de avaliação | | | | |
|--------------------------------|--------------------------------|-----------------------|----------|---------------|----------------|------------|
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,48 | 0,57 | 0,75 | 0,00 | 0,65 |
| | SVM com data augmentation | 0,61 | 0,55 | 0,75 | 0,50 | 0,63 |
| | RUSBoost | 0,44 | 0,44 | 0,44 | 0,44 | 0,44 |
| | RUSBoost com data augmentation | 0,42 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 17 | SVM | 0,44 | 0,50 | 0,20 | 0,75 | 0,29 |
| | SVM com data augmentation | 0,56 | 0,50 | 0,42 | 0,67 | 0,45 |
| | RUSBoost | 0,53 | 0,50 | 0,38 | 0,67 | 0,43 |
| | RUSBoost com data augmentation | 0,67 | 0,60 | 0,75 | 0,60 | 0,67 |
| UPDRS 18 | SVM | 0,42 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,58 | 1,00 | 0,11 | 1,00 | 0,20 |
| | RUSBoost | 0,67 | 0,25 | 0,25 | 0,79 | 0,25 |
| | RUSBoost com data augmentation | 0,53 | 0,20 | 0,25 | 0,64 | 0,22 |
| UPDRS 19 | SVM | 0,68 | 1,00 | 0,13 | 1,00 | 0,22 |
| | SVM com data augmentation | 0,75 | 0,50 | 0,71 | 0,76 | 0,59 |
| | RUSBoost | 0,76 | 0,38 | 0,60 | 0,79 | 0,46 |
| | RUSBoost com data augmentation | 0,60 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 21 | SVM | 0,78 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,75 | 0,67 | 0,40 | 0,91 | 0,50 |
| | RUSBoost | 0,47 | 0,11 | 0,50 | 0,47 | 0,18 |
| | RUSBoost com data augmentation | 0,88 | 0,00 | 0,00 | 1,00 | 0,00 |

A próxima tabela apresenta os cálculos rodados considerando idade, tempo de diagnóstico, TUG, TMW, T30 e Hoehn e Yahr como features. Nesta tentativa, foram excluídas massa e estatura. Os resultados apresentaram uma maior quantidade de classificadores com mais de 50% de assertividade nos valores de referência 15 e 17 da UPDRS, mas ainda assim, sem força para ser considerado um bom classificador.

Tabela 10 – TESTE 7: Classificação da progressão da doença a partir de aspectos funcionais (TUG, TMW e T30), aspectos clínicos (Tempo de diagnóstico e Escala de Hoehn e Yahr) e características individuais (Idade)

| TESTE 7 | | TUG, TMW, T30, tempo de diagnóstico, Hoehn e Yahr e idade | | | | |
|---------------------|--------------------------------|---|----------|---------------|----------------|------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,64 | 0,76 | 0,70 | 0,50 | 0,73 |
| | SVM com data augmentation | 0,65 | 0,73 | 0,76 | 0,40 | 0,74 |
| | RUSBoost | 0,61 | 0,68 | 0,72 | 0,40 | 0,70 |
| | RUSBoost com data augmentation | 0,54 | 0,79 | 0,55 | 0,50 | 0,65 |
| UPDRS 17 | SVM | 0,48 | 0,67 | 0,13 | 0,92 | 0,22 |
| | SVM com data augmentation | 0,53 | 0,50 | 0,50 | 0,56 | 0,50 |
| | RUSBoost | 0,54 | 0,53 | 0,62 | 0,46 | 0,57 |
| | RUSBoost com data augmentation | 0,55 | 0,50 | 0,54 | 0,56 | 0,52 |
| UPDRS 18 | SVM | 0,56 | 0,31 | 0,67 | 0,53 | 0,42 |
| | SVM com data augmentation | 0,51 | 0,41 | 0,47 | 0,55 | 0,44 |
| | RUSBoost | 0,72 | 0,00 | 0,00 | 1,00 | 0,00 |
| | RUSBoost com data augmentation | 0,45 | 0,43 | 0,30 | 0,60 | 0,35 |
| UPDRS 19 | SVM | 0,84 | 1,00 | 0,17 | 1,00 | 0,29 |
| | SVM com data augmentation | 0,61 | 0,29 | 0,17 | 0,81 | 0,21 |
| | RUSBoost | 0,48 | 0,24 | 0,50 | 0,47 | 0,32 |
| | RUSBoost com data augmentation | 0,87 | 0,00 | 0,00 | 1,00 | 0,00 |
| UPDRS 21 | SVM | 0,91 | 0,00 | 0,00 | 0,97 | 0,00 |
| | SVM com data augmentation | 0,70 | 0,00 | 0,00 | 0,89 | 0,00 |
| | RUSBoost | 0,61 | 0,00 | 0,00 | 0,77 | 0,00 |
| | RUSBoost com data augmentation | 0,87 | 0,67 | 0,67 | 0,92 | 0,67 |

6.4 Classificação a partir de características individuais e aspectos clínicos

Mesmo não sendo o objetivo inicial desta investigação, foram calculadas combinações de características individuais e aspectos clínicos sem considerar os resultados dos testes funcionais realizados. Na primeira tentativa, foram calculadas as features: idade, tempo de diagnóstico, estatura e massa, conforme apresentado na tabela a seguir:

Tabela 11 – TESTE 8: Classificação da progressão da doença a partir de aspectos clínicos (Tempo de diagnóstico) e características individuais (Idade, massa e estatura)

| TESTE 8 | | Tempo de diagnóstico, idade, massa e estatura | | | | |
|---------------------|--------------------------------|---|----------|---------------|----------------|------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,47 | 0,59 | 0,62 | 0,18 | 0,60 |
| | SVM com data augmentation | 0,50 | 0,71 | 0,55 | 0,36 | 0,62 |
| | RUSBoost | 0,46 | 0,76 | 0,54 | 0,00 | 0,63 |
| | RUSBoost com data augmentation | 0,59 | 0,75 | 0,68 | 0,29 | 0,71 |
| UPDRS 17 | SVM | 0,57 | 0,55 | 0,65 | 0,50 | 0,59 |
| | SVM com data augmentation | 0,56 | 0,50 | 0,65 | 0,48 | 0,57 |
| | RUSBoost | 0,47 | 0,40 | 0,50 | 0,45 | 0,44 |
| | RUSBoost com data augmentation | 0,54 | 0,50 | 0,25 | 0,79 | 0,33 |
| UPDRS 18 | SVM | 0,43 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,48 | 0,53 | 0,47 | 0,50 | 0,50 |
| | RUSBoost | 0,26 | 0,11 | 0,14 | 0,33 | 0,13 |
| | RUSBoost com data augmentation | 0,40 | 0,21 | 0,75 | 0,31 | 0,33 |
| UPDRS 19 | SVM | 0,68 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,59 | 0,33 | 0,18 | 0,81 | 0,24 |
| | RUSBoost | 0,64 | 0,29 | 0,18 | 0,82 | 0,22 |
| | RUSBoost com data augmentation | 0,48 | 0,09 | 0,25 | 0,52 | 0,13 |
| UPDRS 21 | SVM | 0,60 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,60 | 0,36 | 0,33 | 0,72 | 0,34 |
| | RUSBoost | 0,64 | 0,20 | 0,29 | 0,72 | 0,24 |
| | RUSBoost com data augmentation | 0,75 | 0,36 | 0,67 | 0,77 | 0,47 |

Mesmo sendo a classificação de Hoehn também um valor indicativo da progressão da doença, essa *feature* foi testada separada do tempo de diagnóstico. A tabela abaixo apresenta os valores obtidos por cada um dos classificadores, não apresentando nenhuma combinação de máquina e valor de referência relevantes ou eficientes para investigações em outros bancos de dados:

Tabela 12 – TESTE 9: Classificação da progressão da doença a partir de aspectos clínicos (Tempo de diagnóstico e Escala de Hoehn e Yahr) e características individuais (Idade, massa e estatura)

| TESTE 9 | | Features: tempo de diagnóstico, Hoehn e Yahr, idade, massa e estatura. | | | | |
|----------------------------|--------------------------------|---|-----------------|----------------------|-----------------------|-------------------|
| Valor de referência | Classificador | Métricas de avaliação | | | | |
| | | Acurácia | Precisão | Sensibilidade | Especificidade | F1-measure |
| UPDRS 15 | SVM | 0,50 | 0,60 | 0,71 | 0,09 | 0,65 |
| | SVM com data augmentation | 0,49 | 0,53 | 0,53 | 0,44 | 0,53 |
| | RUSBoost | 0,38 | 1,00 | 0,29 | 1,00 | 0,44 |
| | RUSBoost com data augmentation | 0,43 | 0,60 | 0,39 | 0,50 | 0,47 |
| UPDRS 17 | SVM | 0,44 | 0,27 | 0,60 | 0,38 | 0,38 |
| | SVM com data augmentation | 0,40 | 0,38 | 0,60 | 0,25 | 0,46 |
| | RUSBoost | 0,47 | 0,40 | 0,24 | 0,68 | 0,30 |
| | RUSBoost com data augmentation | 0,38 | 0,42 | 0,63 | 0,13 | 0,50 |
| UPDRS 18 | SVM | 0,44 | 0,20 | 0,09 | 0,71 | 0,13 |
| | SVM com data augmentation | 0,48 | 0,67 | 0,53 | 0,33 | 0,59 |
| | RUSBoost | 0,35 | 0,08 | 0,20 | 0,39 | 0,12 |
| | RUSBoost com data augmentation | 0,41 | 0,21 | 0,33 | 0,45 | 0,26 |
| UPDRS 19 | SVM | 0,62 | 0,50 | 0,19 | 0,88 | 0,27 |
| | SVM com data augmentation | 0,58 | 0,00 | 0,00 | 0,78 | 0,00 |
| | RUSBoost | 0,58 | 0,14 | 0,20 | 0,68 | 0,17 |
| | RUSBoost com data augmentation | 0,68 | 0,17 | 0,33 | 0,74 | 0,22 |
| UPDRS 21 | SVM | 0,71 | 0,00 | 0,00 | 1,00 | 0,00 |
| | SVM com data augmentation | 0,65 | 0,29 | 0,80 | 0,62 | 0,42 |
| | RUSBoost | 0,63 | 0,38 | 0,33 | 0,76 | 0,35 |
| | RUSBoost com data augmentation | 0,69 | 0,33 | 0,22 | 0,85 | 0,27 |

7 DISCUSSÃO

Antes de discutir os resultados propriamente ditos, é necessário fazer considerações sobre a coleta dos dados utilizados nesta pesquisa. Desta forma, a diferença inter avaliadores deve ser considerada. Toda a equipe voluntária do PAFPDN que participava da coleta dos testes era treinada, e recebia instruções específicas para as tarefas que executariam nos dias de testes. Ainda assim, variações de registro por parte dos avaliadores, possíveis erros de registro, entre outras falhas passíveis do trabalho com uma equipe grande podem ter interferido no registro das pontuações.

Um outro aspecto a ser observado é que os dados são provenientes de pesquisas experimentais. Em diferentes modalidades, as pessoas foram testadas antes e após intervenção com algum tipo de exercício. As modalidades ofertadas pelo PAFPDN na época dos dados que foram utilizados nesta pesquisa eram: musculação, dança de salão e equoterapia. Em função disto, é possível que os resultados dos testes funcionais tenham sido alterados por conta da intervenção, mesmo que isto não signifique que os exercícios tenham alterado o estado clínico dos participantes (Ramazzina et al 2017).

Em relação à escolha das variáveis de entrada, as combinações foram resultado da análise dos objetivos da pesquisa, bem como de sugestões encontradas na literatura que envolve a doença de Parkinson.

A classificação a partir de aspectos funcionais analisou se apenas as combinações dos testes funcionais TUG, TMW e 30 seriam suficientes para classificar a progressão da pessoa com DP (TESTE 1). Em todos os níveis de referência e em todas as máquinas utilizadas, nenhum classificador apresentou indicadores altos o bastante para uma classificação confiável. O resultado mais promissor foi o RUSBoost com saída de UPDRS III valor 19. Este classificador apresenta indicadores entre 65% e 80% aproximadamente.

Fraiwan e Hassanin (2021) utilizaram as mesmas métricas ROC para avaliar os classificadores desenhados para identificar doenças neurodegenerativas a partir de dados de marcha. Os autores consideraram outras doenças além da DP, e outros métodos de aprendizado de máquina. Nos testes envolvendo o RUSBoost, os indicadores de acurácia, sensibilidade, especificidade e F1-measure tiveram porcentagens compatíveis com as encontradas neste estudo. Três dos quatro testes utilizando RUSBoost retornaram valores entre 50% e 80%, aproximadamente, e um dos classificadores foi avaliado acima de 90%, sendo um dos destaques da publicação. A avaliação destes classificadores foi considerada positiva, e reforça a qualidade dos classificadores desta pesquisa.

O segundo conjunto de combinações de variáveis diz respeito aos aspectos funcionais e clínicos. Neste caso, além dos testes funcionais TUG, TMW e T30 a classificação de Hoehn e Yahr também foi atribuída nas variáveis de entrada (TESTE 2). A tentativa retornou com um dos melhores resultados desta pesquisa, utilizando SVM com data augmentation e saída de UPDRS III valor 15, apresentando todos os indicadores acima de 77%.

Este resultado positivo talvez seja a maior contribuição deste trabalho, não apenas pelos bons indicadores alcançados, mas também pela praticidade de sua aplicação no transcurso da doença. Além disso, parece razoável indicar a progressão de uma doença motora a partir da associação de uma característica clínica com os valores dos testes funcionais.

Percebam que a classificação da escala de Hoehn e Yahr é relativamente simples de se realizar, uma vez que não considera aspectos mais refinados da doença. Os primeiros números da escala apontam apenas a localização da manifestação da doença no corpo da pessoa acometida, e os demais, uma progressão muito demarcada da perda da autonomia funcional. Ou seja, apesar da simples utilização, a escala não é sensível o suficiente para descrever clinicamente o paciente (GOETZ, 2004).

Em termos práticos, é possível que duas pessoas com estágios diferentes de sintomas estejam com a mesma classificação de Hoehn e Yahr. Por exemplo: uma pessoa muito tremulante ou com repetidos episódios de freezing pode ser classificada como 2,5 na escala Hoehn e Yahr; e a mesma classificação pode acontecer para uma pessoa com maior rigidez, mas pouco tremulante. Uma constatação adicional é que a escala de Hoehn e Yahr apresenta um espectro de pontuação reduzido, com 7 níveis de referência no total. Isso significa que toda mudança de nível representa um avanço significativo da DP.

Levando em consideração essas limitações, a UPDRS foi uma escala criada na perspectiva de se estabelecer um instrumento padrão de avaliação da DP, pensado de forma abrangente para ser suficiente para monitorar a doença; Goetz et al (2008) aprofundaram a especificação dos sintomas na publicação da MDS-UPDRS, versão vigente até os dias de hoje. Os autores esmiuçaram detalhes dos sintomas motores e redistribuíram algumas pontuações as versões antigas da UDPRS, para que fosse possível quantificar com mais precisão os sintomas da doença.

Apesar de ambas as escalas, Hoehn e Yahr e UPDRS, se proporem a monitorar a progressão da DP, existe uma diferença conceitual entre as duas, sendo necessário esclarecer os tipos diferentes de resultados fornecidos. A escala de Hoehn e Yahr traz

valores categóricos e progressivos, considerando o desenvolvimento da doença em uma relação direta: quanto maior a progressão da doença, maior a pontuação da escala.

Por um outro lado, a escala de UPDRS também apresenta valores categóricos, mas não foi desenhada para ser interpretada como um escore contínuo. A quantificação dos sintomas pela UPDRS torna prática a interpretação da situação clínica realizada para especialistas da área, mas são dados para serem observados de forma separada (GOETZ et al, 2004).

Em 2015, Martínez-Martín et al (2015) tentou calcular um possível escalonamento da UPDRS a partir de outras escalas, entre elas a Hoehn e Yahr. Os autores perceberam que mesmo não sendo criada para esta finalidade, a UPDRS pode ser considerada uma medida numérica de indicação contínua e progressiva da DP. A publicação sugere, inclusive, pontos de corte para determinação dos estágios leve, moderado e avançado da doença.

O estudo indica pontos de corte para todas as seções da UPDRS. Em relação ao exame motor (UPDRS III), os resultados sugerem 32/33 pontos como indicadores da transição do estágio leve para moderado e 58/59 pontos como transição do estágio moderado para avançado. Não foi possível calcular estes valores como valores de referência pela pequena quantidade de dados disponível no banco de dados desta pesquisa, conforme apontado na descrição dos resultados.

O terceiro conjunto de combinações de variáveis abrange os aspectos funcionais, clínicos e características individuais. No TESTE 3, as variáveis de entrada submetidas aos classificadores foram: TUG, TMW, T30, UPDRS II, idade e tempo de diagnóstico. Essa combinação surgiu a partir da publicação de Latourelle et al (2017), que identificou a utilização dos escores das partes II e III da UPDRS.

Este classificador, entretanto, não apresentou resultados satisfatórios em nenhuma combinação, inclusive nas destacadas nos resultados, acima de 50%. No teste de SVM com data augmentation e saída de UPDRS valor 15, a interpretação é de uma classificação desbalanceada, visto que é grande o intervalo entre os valores da sensibilidade e a especificidade do classificador.

Uma possível explicação para a baixa qualidade destes dados foi o pequeno conjunto de dados do UPDRS II. Esta foi a medida com mais valores ausentes entre todas as variáveis utilizadas na pesquisa. Quando o tamanho da amostra não é grande o suficiente para treinar bem um classificador, dois eventos podem acontecer: resultados muito extremos e desbalanceados, ou valores baixos e pouco relevantes (PRATI, 2006).

Os dois cenários são percebidos nesta combinação de aspectos funcionais, clínicos e individuais.

Desta forma, para tentar superar a limitação apresentada, foi testada a combinação anterior com a exclusão do UPDRS II. As variáveis de entrada, presentes no TESTE 4 foram TUG, TMW, T30, idade e tempo de diagnóstico. Os testes retornaram avaliações positivas, incluindo o SVM com data augmentation e saída de UPDRS valor 15. Além dos valores de todos os indicadores acima de 70%, o valor expressivo pode indicar que o classificador bem avaliado no TESTE 2 foi não aleatório.

Na perspectiva desta comparação, cabe a provocação a respeito do que se entende sobre condição clínica e condição funcional das pessoas com doença de Parkinson. Greenberg (2014) coloca que a condição clínica não é modificável pela natureza da própria doença, e as publicações da Movement Disorders Society a respeito da classificação da doença de Parkinson sempre tiveram um foco muito grande na análise dos sintomas (GOETZ, 2003; GOETZ, 2004; GOETZ, 2008).

Inspirado na teoria de SVM, que demonstra que uma mesma situação observada de um outro ângulo pode apresentar informações diferentes (REF), uma sugestão trazida é que talvez fosse mais claro observar a progressão da doença a partir da condição funcional dos indivíduos acometidos, e não apenas da condição clínica. Para sustentar o argumento, é necessário levar em consideração que se trata de um distúrbio motor (GREENBERG et al 2014), que se caracteriza como doença de Parkinson a partir das manifestações motoras (POEWE et al, 2017), em que vem sendo praticada uma grande quantidade de testes funcionais para a avaliação condição funcional de pessoas com DP (OPARA, 2017).

Neste sentido, os classificadores desta pesquisa com avaliação positiva contribuem não pela determinação de um valor de referência ou uma sistematização definitiva do monitoramento da DP a partir de aspectos funcionais, mas pelo cálculo da possibilidade deste tipo de relação. Da mesma forma que Martínez-Martín (2015) percebeu que a MDS-UPDRS pode ser observada a partir de um ponto de vista numérico, estes achados positivos podem representar uma forma de quantificar a progressão da doença, a partir da avaliação da condição funcional das pessoas com DP.

Reforçando o raciocínio, a linha do tempo de Poewe (2017) explicita uma fase prodromica ao início dos sintomas motores da DP. Pelo argumento de Hawkes (2010), a doença de Parkinson inicia aproximadamente 20 anos antes do diagnóstico, mas sem uma possibilidade de definição apenas pelos sintomas não motores. Em 2020 Yoo et al (2020) publicou um estudo de coorte com 1.196.614 participantes entre os anos de 2009 e 2014.

O TUG, um dos principais testes funcionais utilizados na investigação com DP (OPARA, 2017), pode ser um marcador prodrômico de risco para o desenvolvimento da doença de Parkinson.

Uma outra consideração necessária é perceber que não foi utilizada a classificação de Hoehn e Yahr como métrica clínica. Neste sentido, existe a indicação de que a idade e o tempo de diagnóstico podem ser considerados aspectos classificadores da progressão da doença. Por ser uma doença progressiva (GREENBERG et al 2014), é esperado que quanto maior for o tempo de diagnóstico, maior será o nível de acometimento da doença.

A idade, por sua vez, não é considerada um aspecto clínico, mas uma característica individual; mesmo assim, que pode estar relacionada com o desenvolvimento do quadro clínico da pessoa com DP. A progressão da DP pode acontecer de forma lenta ou rápida, e não são claros quais os fatores que determinam esta diferença (LAROURETTE, 2017). Mesmo assim, Lees et al (2009) defendem que a idade é um fator de risco para a doença. Desta forma, nos resultados encontrados nesta pesquisa, a idade e o tempo de diagnóstico conseguiram apresentar valores de qualidade semelhantes entre os classificadores com diferentes variáveis clínicas.

O TESTE 5 foi conduzido de forma semelhante ao teste anterior, sendo adicionado informações antropométricas de massa e estatura. Neste caso, dois classificadores chamaram atenção: o RUSBoost com data augmentation e saída UPDRS de valor 15 e o SVM com data augmentation e saída UPDRS de valor 18. Apesar dos indicadores positivos, não foi encontrada respaldo na literatura para considerar massa e estatura como medidas influenciadas pela DP.

Um estudo de 1997 (SAKAJIRI e TAKAMORI, 1997) identificou um padrão de perda de massa em pessoas diagnosticadas com DP. Na época da publicação os autores associaram a perda de peso como experiência decorrente do desenvolvimento da doença. Em 2016 Lindskov et al (2016) concluiu que não existem evidências que associem alterações da composição corporal como etapa comum durante o curso da DP. Desta maneira, os interesses da investigação sobre composição corporal em pessoas com Parkinson giram em torno da influência destas variáveis diretamente sobre os sintomas, como o equilíbrio (DUARTE, 2010), por exemplo.

No TESTE 6 foram consideradas como variáveis de entrada TUG, TMW, T30, tempo de diagnóstico, idade, massa, estatura e a classificação de Hoehn e Yahr. Entre os dois classificadores com avaliação positiva nesta combinação, mais uma vez o SVM com data augmentation e saída UPDRS de valor 15 foi destaque. O classificador RUSBoost com data augmentation e saída UPDRS de valor 17 também retornou com avaliação

favorável. A diferença deste teste para o TESTE 7 consiste na retirada das variáveis antropométricas.

O TESTE 7 considerou TUG, TMW, T30, tempo de diagnóstico, idade e a classificação de Hoehn e Yahr, e apresentou mais quantidades de classificadores com avaliação positiva do que o teste anterior. O SVM com saída UPDRS de valor 15 e o RUSBoost com data augmentation e saída UPDRS de valor 15 foram os classificadores mais precisos desta combinação de variáveis. Ainda assim, a especificidade de 50% indica que estes classificadores não são sensíveis o suficiente.

Era esperado que este classificador conseguisse avaliar de forma mais assertiva todos os dados a serem classificados. Reforçados pelos indicadores dos TESTE 2 e TESTE 4, ao considerar os aspectos funcionais e um conjunto de aspectos que podem trazer informações sobre os aspectos clínicos, os resultados do TESTE 7 ficaram aquém do que era previsto. É possível que a quantidade de dados, de uma forma geral, não tenha sido grande o suficiente para máquina alcançar os padrões entre as variáveis e a sobreposição de variáveis para se chegar em uma avaliação positiva da classificação.

Os TESTE 8 e TESTE 9 foram conduzidos para investigar se os aspectos funcionais realmente eram determinantes para a construção de um bom classificador. Mesmo o SVM com saída UPDRS de valor 17 apresentando indicadores acima de 50%, todos os resultados apresentados apontam para uma classificação fraca da progressão da doença. Uma vez que a seção III da UPDRS é dedicada ao exame motor da pessoa com DP (GOETZ, et al 2008), estes achados servem para reforçar protagonismo dos testes funcionais no monitoramento da doença de Parkinson.

A revisão de Ramazzina et al (2017) reconhece a qualidade dos artigos que recomendam exercícios como tratamento adjunto, mas diz não ser claro como os exercícios podem afetar o status clínico das pessoas acometidas. A dificuldade de achar esta resposta talvez seja porque os exercícios não afetem o aspecto clínico em si. A explicação se dá no contexto da automaticidade de movimento, que estabelece padrões neurais a fim de otimizar a movimentação e diminuir a concentração de atenção despendida para a tarefa (KANDEL, 2014).

Levando em consideração as etapas necessárias para a execução do movimento (ver item 4.1.4) e o aspecto orgânico da doença de Parkinson (ver item 4.1.1), percebe-se que a origem do problema são as falhas de comunicação neural protagonizadas pela perda de dopamina nos núcleos da base. Uma vez que estas estruturas fazem parte do controle ou refinamento dos movimentos, o envolvimento de níveis neurologicamente mais altos podem compensar o prejuízo trazido pela doença no que tange ao controle motor de uma

tarefa específica. Mas isso não significa nenhum tipo de alteração no aspecto clínico. Existem duas situações que podem exemplificar o argumento, descritas abaixo.

Alguns estudos que investigaram os efeitos da dança em pessoas com Parkinson (VIDAL, 2014, EAHART, 2009) justificaram que a dança poderia ser uma boa ferramenta para a locomoção de pessoas com DP; não por reverter o quadro da doença, mas por possibilitar uma mudança de caminho neural durante a marcha, por exemplo. Nos estudos conduzidos pela professora Gammon Eahrart (EAHART, 2009; DUNCAN e EAHART, 2014), os indivíduos apresentavam dificuldades de marcha, mas executavam os passos do tango argentino, semelhantes à passada da marcha, sem a mesma dificuldade. Neste exemplo, uma possível explicação é a troca da circuitaria neural para a realização das tarefas. O ritmo, a concentração e a musicalidade são elementos que se somam à marcha durante a dança do tango argentino.

É oportuno, também, discutir uma percepção a respeito de uma nova tendência ao se fazer ciência. A exemplo da própria doença de Parkinson, a primeira pessoa que descreve ou descobre um fenômeno vai ser por muito anos referenciada por outros investigadores a respeito do mesmo assunto. Também pelos cuidados relacionados à propriedade intelectual, os dados coletados ao longo das pesquisas são protegidos e divulgados apenas após muitas análises, através das conclusões obtidas.

Entretanto, iniciativas como a Michael J. Fox Foundation disponibilizam biorrepositórios para serem estudados, e isto significa um novo momento nas pesquisas sobre doença de Parkinson, sendo um reflexo deste novo formato de fazer ciência. Admitir que chegamos, como sociedade, a uma estrutura mais tecnológica e de conhecimento mais acessível pode acelerar descobertas. Neste sentido, é necessário, também, reconhecer dificuldades de construir ou obter tecnologias mais avançadas e/ou a bancos de dados maiores. Uma vez que o aprendizado de máquinas é uma boa ferramenta para analisar grandes bancos de dados, a quantidade de dados se mostra tão fundamental como a qualidade destes dados para uma análise representativa e isso sugere que o trabalho coletivo e colaborativo é um instrumento forte a ser utilizado pelo método científico.

8 CONCLUSÃO

Por tudo o que foi exposto até aqui, conclui-se que é possível que aspectos funcionais possam auxiliar na predição da classificação da doença de Parkinson, utilizando técnicas do aprendizado de máquina.

A classificação dos aspectos funcionais foi positiva principalmente nos classificadores do tipo RUSBoost. Foi possível a identificação de classificações nos valores de referência 17, 18, 19 e 20 da UPDRS III.

Quando combinados os aspectos funcionais e clínicos, o SVM com 1000 trial e saída de UPDRS III valor 15 parece ser um ótimo classificador. Outras classificações com diferentes combinações dos testes funcionais com a escala de Hoehn e Yahr, idade e tempo de diagnóstico apresentaram métricas positivas com o mesmo classificador e o mesmo valor de referência para a UPDRS III.

Adicionalmente, foram submetidos aos classificadores as variáveis clínicas e características individuais, sem a presença dos valores dos testes funcionais. O padrão de todas as métricas retornou com valor desbalanceado, indicando a necessidade de se considerar os testes funcionais neste tipo de investigação.

Sugere-se que estudos futuros possam submeter bancos de dados maiores a classificadores, para que as combinações propostas sejam testadas, a fim de auxiliar pessoas acometidas durante o transcurso da doença de Parkinson.

REFERÊNCIAS

- ABUJRIDA, Hamza; AGU, Emmanuel; PAHLAVAN, Kaveh. Machine learning-based motor assessment of Parkinson's disease using postural sway, gait and lifestyle features on crowdsourced smartphone data. **Biomedical Physics & Engineering Express**, v. 6, n. 3, p. 035005, 2020.
- BARBOSA, E. R.; LIMONGI, J. C. Tratamento das complicações neuropsiquiátricas na doença de Parkinson. **Andrade LAF, Barbosa RE, Cardoso F, Teive HAG. Doença de Parkinson: estratégias atuais de tratamento**, v. 2, p. 143-53, 2006.
- BEJNORDI, Babak Ehteshami et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. **Jama**, v. 318, n. 22, p. 2199-2210, 2017.
- BENGIO, Yoshua. Deep learning of representations for unsupervised and transfer learning. In: **Proceedings of ICML workshop on unsupervised and transfer learning**. 2012. p. 17-36.
- BERG, Katherine et al. Measuring balance in the elderly: preliminary development of an instrument. **Physiotherapy Canada**, v. 41, n. 6, p. 304-311, 1989.
- BI, Qifang et al. What is machine learning? A primer for the epidemiologist. **American journal of epidemiology**, v. 188, n. 12, p. 2222-2239, 2019.
- BOHANNON, Richard W.; ANDREWS, A. Williams; THOMAS, Michael W. Walking speed: reference values and correlates for older adults. **Journal of Orthopaedic & Sports Physical Therapy**, v. 24, n. 2, p. 86-90, 1996.
- BOROVAC, Josip Anđelo. Focus: The Aging Brain: Side effects of a dopamine agonist therapy for Parkinson's disease: A mini-review of clinical pharmacology. **The Yale journal of biology and medicine**, v. 89, n. 1, p. 37, 2016.
- BOT, Brian M. et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. **Scientific data**, v. 3, n. 1, p. 1-9, 2016.
- BOUÇA-MACHADO, Raquel et al. Physical activity, exercise, and physiotherapy in Parkinson's disease: defining the concepts. **Movement Disorders Clinical Practice**, v. 7, n. 1, p. 7-15, 2020.
- BOVOLENTA, Tânia M. et al. Average annual cost of Parkinson's disease in São Paulo, Brazil, with a focus on disease-related motor symptoms. **Clinical interventions in aging**, v. 12, p. 2095, 2017.
- BRASIL. Saúde. Protocolo Clínico e Diretrizes Terapêuticas. **PORTARIA N° 228, DE 10 DE MAIO DE 2010**.
- BRASIL. Saúde. Protocolo Clínico; Diretrizes Terapêuticas–Doença de Parkinson. **PORTARIA CONJUNTA N° 10, DE 31 DE OUTUBRO DE 2017**.
- BRATIĆ, Brankica et al. Machine learning for predicting cognitive diseases: methods, data sources and risk factors. **Journal of medical systems**, v. 42, n. 12, p. 243, 2018.
- BYEON, Haewon. Development of a depression in Parkinson's disease prediction model using machine learning. **World Journal of Psychiatry**, v. 10, n. 10, p. 234, 2020.

- CAMPANA, Álvaro Oscar. Metodologia da investigação científica aplicada à área biomédica: 2. Investigações na área médica. **Jornal de Pneumologia**, v. 25, p. 84-93, 1999.
- CANO PORRAS, Desiderio et al. Advanced virtual reality-based rehabilitation of balance and gait in clinical practice. **Therapeutic advances in chronic disease**, v. 10, p. 2040622319868379, 2019.
- CANO-DE-LA-CUERDA, Roberto et al. Axial rigidity and quality of life in patients with Parkinson's disease: a preliminary study. **Quality of Life Research**, v. 20, n. 6, p. 817-823, 2011.
- CID – Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde <http://cid.ninsaude.com/capitulo/vi/g20-g26/#.XrsT9WhKhPY> acessado em 12/05/2020
- Ciência da Computação 9: Aprendizado de Máquina e Visão Computacional com Sandra Avila, Entrevistadores: Vanderson M. Rosário e Diogo Gonçalves. Podcast. Disponível em: https://open.spotify.com/episode/3rdDct6V1KUNJFQbefwtD0?si=gvhpfO6wRHq-3KX53Zil3g&dl_branch=1 . Acessado em: abril de 2021.
- CLARK, David J. Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. **Frontiers in human neuroscience**, v. 9, p. 246, 2015.
- DA FONTOURA, Andréa Silveira; FORMENTIN, Charles Marques; ABECH, Everson Alves. Guia prático de avaliação física: uma abordagem didática, abrangente e atualizada. **Phorte Editora LTDA**, 2011.
- DE JESUS SANTOS, Taynan et al. AVALIAÇÃO DO EQUILÍBRIO NA DOENÇA DE PARKINSON: UMA REVISÃO SISTEMÁTICA. In: **Congresso Internacional de Atividade Física, Nutrição e Saúde**. 2017.
- DELLA COLETTA, M. V.; et al. Doença de Parkinson: Outros Sintomas Não-Motores. Cap. 8, 2016.
- DEO, Rahul C. Machine learning in medicine. **Circulation**, v. 132, n. 20, p. 1920-1930, 2015.
- DIBBLE, Leland E.; ADDISON, Odessa; PAPA, Evan. The effects of exercise on balance in persons with Parkinson's disease: a systematic review across the disability spectrum. **Journal of Neurologic Physical Therapy**, v. 33, n. 1, p. 14-26, 2009.
- DICKSON, Dennis W. et al. Neuropathology of non-motor features of Parkinson disease. **Parkinsonism & related disorders**, v. 15, p. S1-S5, 2009.
- DUARTE, Marcos; FREITAS, Sandra MSF. Revisão sobre posturografia baseada em plataforma de força para avaliação do equilíbrio. **Brazilian Journal of Physical Therapy**, v. 14, n. 3, p. 183-192, 2010.
- DUNCAN, Ryan P.; EARHART, Gammon M. Are the effects of community-based dance on Parkinson disease severity, balance, and functional mobility reduced with time? A 2-year prospective pilot study. **The Journal of Alternative and Complementary Medicine**, v. 20, n. 10, p. 757-763, 2014.
- DYAGILEV, Kirill; SARIA, Suchi. Learning (predictive) risk scores in the presence of censoring due to interventions. **Machine Learning**, v. 102, n. 3, p. 323-348, 2016.

- EARHART, Gammon M. Dance as therapy for individuals with Parkinson disease. **European journal of physical and rehabilitation medicine**, v. 45, n. 2, p. 231, 2009.
- FALCHETTI, Marcelo; PREDIGER, Rui Daniel; ZANOTTO-FILHO, Alfeu. Classification algorithms applied to blood-based transcriptome meta-analysis to predict idiopathic Parkinson's disease. **Computers in Biology and Medicine**, v. 124, p. 103925, 2020.
- FAWCETT, Tom. An introduction to ROC analysis. **Pattern recognition letters**, v. 27, n. 8, p. 861-874, 2006.
- FIELD, Andy. **Descobrimdo a estatística usando o SPSS-5**. Penso Editora, 2009.
- FISCHER, Bruno Leonardo. Efeitos do treinamento de potência na força muscular de indivíduos com doença de Parkinson. 2014.
- FRAIWAN, Luay; HASSANIN, Omnia. Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers. **Plos one**, v. 16, n. 6, p. e0252380, 2021.
- GALAR, Diego et al. Maintenance decision making based on different types of data fusion. **Eksploatacja i niezawodnosc-Maintenance and Reliability**, v. 14, n. 2, p. 135-144, 2012.
- GIARDINI, Marica et al. Instrumental or physical-exercise rehabilitation of balance improves both balance and gait in Parkinson's disease. **Neural plasticity**, v. 2018, 2018
- GOETZ, C. G. et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. **Movement disorders: official journal of the Movement Disorder Society**, v. 23, n. 15, p. 2129-2170, 2008.
- GOETZ, Christopher G. et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations the Movement Disorder Society Task Force on rating scales for Parkinson's disease. **Movement disorders**, v. 19, n. 9, p. 1020-1028, 2004.
- GRAHAM, Sarah et al. Artificial intelligence for mental health and mental illnesses: an overview. **Current psychiatry reports**, v. 21, n. 11, p. 1-18, 2019.
- GREENBERG, David A.; AMINOFF, Michael J.; SIMON, Roger P. **Neurologia clínica-8**. AMGH Editora, 2014.
- HAINES, Duane E.; FAAA, PhD Faaas; MIHAILOFF, Gregory A. (Ed.). **PRINCIPIOS DE NEUROCIENCIA: aplicaciones básicas y clínicas**. Elsevier, 2019.
- HANDELMAN, G. S. et al. eD octor: machine learning and the future of medicine. **Journal of internal medicine**, v. 284, n. 6, p. 603-619, 2018.
- HAWKES, Christopher H.; DEL TREDICI, Kelly; BRAAK, Heiko. A timeline for Parkinson's disease. **Parkinsonism & related disorders**, v. 16, n. 2, p. 79-84, 2010.
- HERCULANO-HOUZEL, S. A Frenologia eo nascimento da Neurociência Experimental. **LENT, R. Cem bilhões de neurônios**. São Paulo: Atheneu/Faperj, p. 20-21, 2001.
- HOEHN MM, YAHR MD. Parkinsonism: onset, progression and mortality. **Neurology** 1967;17:427– 442.

- HOMEM, Rita de Cassia Pereira Pinto. Efeitos da equoterapia no desempenho funcional e na qualidade de vida de pessoas com doença de Parkinson. 2016.
- HUGHES, Andrew J. et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. **Journal of Neurology, Neurosurgery & Psychiatry**, v. 55, n. 3, p. 181-184, 1992.
- HUGHES, Andrew J. et al. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. **Brain**, v. 125, n. 4, p. 861-870, 2002.
- HULLEY, Stephen B. et al. **Delineando a pesquisa clínica-4**. Artmed Editora, 2015.
- ISHIHARA-PAUL, L. et al. Prospective association between emotional health and clinical evidence of Parkinson's disease. **European Journal of Neurology**, v. 15, n. 11, p. 1148-1154, 2008.
- JANKOVIC, J. Pathophysiology and assessment of parkinsonian symptoms and signs. In: PAHWA, Rajesh; LYONS, Kelly E. (Ed.). **Handbook of Parkinson's disease**. Crc Press, 2013.
- JONES, C. Jessie; RIKLI, Roberta E.; BEAM, William C. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. **Research quarterly for exercise and sport**, v. 70, n. 2, p. 113-119, 1999.
- KANDEL, Eric et al. **Princípios de neurociências-5**. AMGH Editora, 2014.
- KOMURA, Daisuke; ISHIKAWA, Shumpei. Machine learning approaches for pathologic diagnosis. **Virchows Archiv**, v. 475, n. 2, p. 131-138, 2019.
- KUBAT, Miroslav. Artificial neural networks. In: **An Introduction to Machine Learning**. Springer, Cham, 2015. p. 91-111.
- LATOURELLE, Jeanne C. et al. Large-scale identification of clinical and genetic predictors of Parkinson's disease motor progression in newly-diagnosed patients: a longitudinal cohort study and validation. **The Lancet. Neurology**, v. 16, n. 11, p. 908, 2017.
- LEES AJ, Hardy J, Revesz T. Parkinson's disease. **Lancet**. 2009 Jun 13;373(9680):2055-66. doi: 10.1016/S0140-6736(09)60492-X. Erratum in: **Lancet**. 2009 Aug 29;374(9691):684. PMID: 19524782.
- LENT, Robert. Cem Bilhões de Neurônios. Conceitos fundamentais de neurociência. **Atheneu-FAPERJ**, 2001.
- LI, Junjie et al. Three-Dimensional Pattern Features in Finger Tapping Test for Patients with Parkinson's disease. In: **2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)**. IEEE, 2020. p. 3676-3679.
- LI, Xiaoyan et al. Lower Limb Resistance Training in Individuals With Parkinson's Disease: An Updated Systematic Review and Meta-Analysis of Randomized Controlled Trials. **Frontiers in Neurology**, v. 11, 2020.
- LINA, Chen et al. The Effect of Virtual Reality on the Ability to Perform Activities of Daily Living, Balance During Gait, and Motor Function in Parkinson Disease Patients: A Systematic Review and Meta-Analysis. **American journal of physical medicine & rehabilitation**, v. 99, n. 10, p. 917-924, 2020.
- LINDSKOV, Susanne et al. Weight stability in Parkinson's disease. **Nutritional neuroscience**, v. 19, n. 1, p. 11-20, 2016.

- LINDSKOV, Susanne et al. Weight stability in Parkinson's disease. **Nutritional neuroscience**, v. 19, n. 1, p. 11-20, 2016.
- MACHADO, A. B. M.; HAERTEL, L. M. Neuroanatomia funcional (Functional neuroanatomy). **Editora Atheneu, São Paulo**, 2014.
- MAHONEY, Florence I.; BARTHEL, Dorothea W. Functional evaluation: the Barthel Index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. **Maryland state medical journal**, 1965.
- MARCELINO, Camila Wells Damato. Efeitos do treinamento com Kettlebell no desempenho funcional, estabilidade postural e força isocinética de membros inferiores em indivíduos com doença de Parkinson. 2017.
- MARTÍNEZ-MARTÍN, Pablo et al. Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale. **Parkinsonism & related disorders**, v. 21, n. 1, p. 50-54, 2015.
- MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.
- MEUR, A.; STAES, L. Psicomotricidade: educação e reeducação. Tradução: Ana Maria Izique Galuban e Setsuko Ono. 1989.
- MIRELMAN, Anat et al. Gait impairments in Parkinson's disease. **The Lancet Neurology**, v. 18, n. 7, p. 697-708, 2019.
- MITCHELL, Tom M. Version spaces: A candidate elimination approach to rule learning. In: **Proceedings of the 5th international joint conference on Artificial intelligence-Volume 1**. 1977. p. 305-310.
- MOLLINEDO-CARDALDA, Irimia; CANCELA-CARRAL, José María; VILA-SUÁREZ, María Helena. Effect of a Mat pilates program with TheraBand on dynamic balance in patients with Parkinson's disease: feasibility study and randomized controlled trial. **Rejuvenation research**, v. 21, n. 5, p. 423-430, 2018.
- MORAES FILHO, Ariel Vieira de. Efeitos do treinamento de força sobre a bradicinesia, força muscular e desempenho funcional em indivíduos com Doença de Parkinson. 2013.
- MORRIS, Susan; MORRIS, Meg E.; IANSEK, Robert. Reliability of measurements obtained with the Timed "Up & Go" test in people with Parkinson disease. **Physical therapy**, v. 81, n. 2, p. 810-818, 2001.
- MOVEMENT DISORDER SOCIETY TASK FORCE ON RATING SCALES FOR PARKINSON'S DISEASE. The unified Parkinson's disease rating scale (UPDRS): status and recommendations. **Movement Disorders**, v. 18, n. 7, p. 738-750, 2003.
- NIERAT, Marie-Cécile et al. When breathing interferes with cognition: experimental inspiratory loading alters timed up-and-go test in normal *humans*. **PloS one**, v. 11, n. 3, p. e0151625, 2016.
- NOYCE, Alastair J. et al. Meta-analysis of early nonmotor features and risk factors for Parkinson disease. **Annals of neurology**, v. 72, n. 6, p. 893-901, 2012.
- OPARA, Józef et al. Motor assessment in Parkinson's disease. **Ann Agric Environ Med**, v. 24, n. 3, p. 411-415, 2017.
- PARKINSON J. An Essay on the Shaking Palsy. London: **Whittingham and Rowland for Sherwood, Neely, and Jones**; 1817.

- PETERS, Denise M.; FRITZ, Stacy L.; KROTISH, Debra E. Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults. **Journal of geriatric physical therapy**, v. 36, n. 1, p. 24-30, 2013.
- PINTER, Bernadette et al. Mortality in Parkinson's disease: a 38-year follow-up study. **Movement Disorders**, v. 30, n. 2, p. 266-269, 2015.
- PODSIADLO, Diane; RICHARDSON, Sandra. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. **Journal of the American Geriatrics Society**, v. 39, n. 2, p. 142-148, 1991.
- POEWE, Werner et al. Parkinson disease. **Nature reviews Disease primers**, v. 3, n. 1, p. 1-21, 2017.
- POSTUMA, Ronald B. et al. Identifying prodromal Parkinson's disease: pre-motor disorders in Parkinson's disease. **Movement Disorders**, v. 27, n. 5, p. 617-626, 2012.
- PRATI, Ronaldo Cristiano. **Novas abordagens em aprendizado de máquina para a geração de regras, classes desbalanceadas e ordenação de casos**. 2006. Tese de Doutorado. Universidade de São Paulo.
- PRINGSHEIM, T. et al. A prevalência da doença de Parkinson: Uma revisão sistemática e meta-análise. **Mov Disord**, v. 29, n. 13, p. 1583-90, 2014.
- RAMAZZINA, Ileana; BERNAZZOLI, Benedetta; COSTANTINO, Cosimo. Systematic review on strength training in Parkinson's disease: an unsolved question. **Clinical Interventions in Aging**, v. 12, p. 619, 2017.
- RASCHKA, Sebastian. **Python machine learning**. Packt publishing ltd, 2015.
- RASTEGAR, Diba Ahmadi et al. Parkinson's progression prediction using machine learning and serum cytokines. **NPJ Parkinson's disease**, v. 5, n. 1, p. 1-8, 2019.
- RASTEGAR, Farzaneh; FOOLADI, Mehdi Dehghan Takht. Online virtual machine assignment using multi-armed bandit in cloud computing. In: **2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)**. IEEE, 2019. p. 40-46.
- REHMAN, Rana Zia Ur et al. Selecting clinically relevant gait characteristics for classification of early parkinson's disease: A comprehensive machine learning approach. **Scientific reports**, v. 9, n. 1, p. 1-12, 2019.
- SAKAJIRI, K.; TAKAMORI, M. Body fat loss in patients with Parkinson's disease. **Rinsho shinkeigaku= Clinical neurology**, v. 37, n. 7, p. 611-614, 1997.
- SAKAJIRI, K.; TAKAMORI, M. Body fat loss in patients with Parkinson's disease. **Rinsho shinkeigaku= Clinical neurology**, v. 37, n. 7, p. 611-614, 1997.
- SALMANPOUR, Mohammad R. et al. Optimized machine learning methods for prediction of cognitive outcome in Parkinson's disease. **Computers in biology and medicine**, v. 111, p. 103347, 2019.
- SALMANPOUR, Mohammad R. et al. Optimized machine learning methods for prediction of cognitive outcome in Parkinson's disease. **Computers in biology and medicine**, v. 111, p. 103347, 2019.
- SAMUEL, Arthur L. Some studies in machine learning using the game of checkers. **IBM Journal of research and development**, v. 3, n. 3, p. 210-229, 1959.

- SANTOS, Tácio Rodrigues da Silva. Função autonômica cardíaca e desempenho no exercício resistido em indivíduos com doença de Parkinson. 2014.
- SCHNEIDER, Walter; CHEIN, Jason M. Controlled & automatic processing: behavior, theory, and biological mechanisms. **Cognitive science**, v. 27, n. 3, p. 525-559, 2003.
- SILVA JUNIOR, Jorge Luiz Ferreira da. Estadiamento da doença de Parkinson com a utilização de técnicas de aprendizado de máquinas. 2018.
- SKINNER, Jared W.; CHRISTOU, Evangelos A.; HASS, Chris J. Lower extremity muscle strength and force variability in persons with Parkinson disease. **Journal of Neurologic Physical Therapy**, v. 43, n. 1, p. 56-62, 2019.
- SOUSA, Nariana Mattos Figueiredo; MACEDO, Roberta Correa. Relationship between cognitive performance and mobility in patients with Parkinson's disease: A cross-sectional study. **Dementia & neuropsychologia**, v. 13, n. 4, p. 403-409, 2019.
- SOUZA, Fernando Vinicius Goncalves de. Desenvolvimento e avaliação de um sistema de espectroscopia funcional de infravermelho próximo para detecção de movimento intencional com base na atividade cerebral. 2020.
- STEFFEN, Teresa; SENEY, Megan. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. **Physical therapy**, v. 88, n. 6, p. 733-746, 2008.
- TAKAKUSAKI, Kaoru. Neurophysiology of gait: from the spinal cord to the frontal lobe. **Movement Disorders**, v. 28, n. 11, p. 1483-1491, 2013.
- TANNER, Caroline M.; GOLDMAN, Samuel M. Epidemiology of Parkinson's disease. **Neurologic clinics**, v. 14, n. 2, p. 317, 1996.
- TILLMAN, Alex et al. Lower limb progressive resistance training improves leg strength but not gait speed or balance in Parkinson's disease: a systematic review and meta-analysis. **Frontiers in aging neuroscience**, v. 7, p. 40, 2015.
- TURING, Alan M.; HAUGELAND, J. **Computing machinery and intelligence**. Cambridge, MA: MIT Press, 1950.
- VAN DEN EEDEN, Stephen K. et al. Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. **American journal of epidemiology**, v. 157, n. 11, p. 1015-1022, 2003.
- VIDAL, Samuel Estevam. Efeitos da dança sobre a manutenção da estabilidade dinâmica e qualidade de vida de indivíduos com doença de Parkinson. 2014.
- WIRDEFELDT, Karin et al. Heritability of Parkinson disease in Swedish twins: a longitudinal study. **Neurobiology of aging**, v. 32, n. 10, p. 1923. e1-1923. e8, 2011.
- YOO, Jung Eun et al. Timed Up and Go Test and the Risk of Parkinson's Disease: A Nation-wide Retrospective Cohort Study. **Movement Disorders**, v. 35, n. 7, p. 1263-1267, 2020.
- YOUNGMANN, Brit et al. A machine learning algorithm successfully screens for Parkinson's in web users. **Annals of clinical and translational neurology**, v. 6, n. 12, p. 2503-2509, 2019.
- ZHAN, Andong et al. Using smartphones and machine learning to quantify Parkinson disease severity: the mobile Parkinson disease score. **JAMA neurology**, v. 75, n. 7, p. 876-880, 2018.

ZHOU, Z. H. Ensemble Methods: Foundations and algorithms. Chapman & Hall/Crc Machine Learning. 2012.