

UNIVERSIDADE FEDERAL DO PAMPA

Anne Lize Lopes Escarrone

**Extração de perfis de usuários com e sem  
deficiência utilizando agrupamentos**

Alegrete  
2017



Anne Lize Lopes Escarrone

## Extração de perfis de usuários com e sem deficiência utilizando agrupamentos

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Software da Universidade Federal do Pampa como requisito parcial para a obtenção do título de Bacharel em Engenharia de Software.

Orientador: Prof. Me. Jean Felipe Patikowski Cheiran

Coorientador: Prof. Dr. Marcelo Resende Thielo

Alegrete  
2017

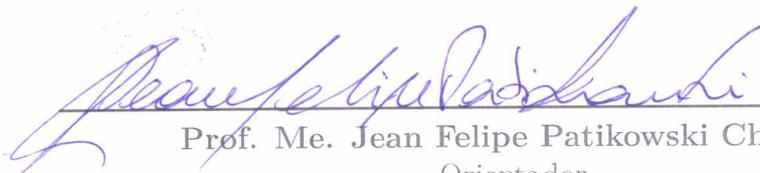


Anne Lize Lopes Escarrone

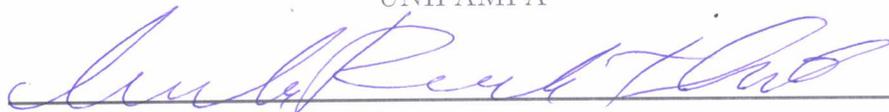
## Extração de perfis de usuários com e sem deficiência utilizando agrupamentos

Trabalho de Conclusão de Curso apresentado  
ao Curso de Graduação em Engenharia de  
Software da Universidade Federal do Pampa  
como requisito parcial para a obtenção do tí-  
tulo de Bacharel em Engenharia de Software.

Trabalho de Conclusão de Curso defendido e aprovado em 28. de novembro de 2017  
Banca examinadora:



Prof. Me. Jean Felipe Patikowski Cheiran  
Orientador  
UNIPAMPA



Prof. Dr. Marcelo Resende Thielo  
Coorientador  
UNIPAMPA



Prof<sup>ª</sup>. Me. Alice Fonseca Finger  
UNIPAMPA



Prof<sup>ª</sup>. Dr<sup>a</sup>. Amanda Meincke Melo  
UNIPAMPA



*A minha família que sempre me apoiou nas minhas escolhas independente de quais fossem. E dedico este trabalho especialmente para Runna que me ensinou que o amor existe independente do modo e do contexto em que ele é feito.*



## **AGRADECIMENTOS**

Primeiramente gostaria de agradecer a Deus por ter me dado força e teimosia para conseguir alcançar meus objetivos.

Gostaria de agradecer aos meus familiares pela paciência e compreensão pois sempre estiveram ao meu lado nos momentos difíceis.

Aos meus orientadores, Professores Jean Felipe Patikowski Cheiran e Marcelo Resende Thielo pela dedicação e apoio durante o trabalho e por acreditarem em mim.

Aos meus colegas e amigos Bruno, Cristiane, Diego, Joziele, Paula, Pedro, Karine, Wesley e Wolleson pela compreensão, amizade e companheirismo. E ao Luiz por ter me ensinado que única pessoa que impede o meu sucesso sou eu mesma.



*“Eu acredito que às vezes são as pessoas que  
ninguém espera nada que fazem as coisas  
que ninguém consegue imaginar.”  
(Alan Turing)*



## RESUMO

Nos dias atuais, pode-se afirmar que a web tem um papel de grande importância, permitindo o amplo acesso ao conhecimento e simplificando a realização de tarefas. Contudo, melhorias ainda se apresentam necessárias para que todas as pessoas, com ou sem deficiência, venham a ter pleno acesso à informação. Com base em tal necessidade são utilizados recursos de tecnologias assistivas, os quais possibilitam que pessoas com deficiência disponham de autonomia na realização de tarefas. Devido às diversas características de interação de pessoas com diferentes tipos de deficiência, é necessário que se realizem adaptações no conteúdo apresentado por páginas da web. Realizamos, então, uma revisão na literatura onde foram encontrados trabalhos que sugerem características de interação de usuários com e sem deficiência, separando-os em perfis distintos com base no seu modo de interação com *websites*. O presente trabalho tem como finalidade realizar o agrupamento de perfis de pessoas cegas, pessoas com deficiência motora e pessoas sem deficiências a partir do seu modo de interação com páginas da web. Para atingir tal objetivo, foi desenvolvido um sistema responsável por coletar as informações do usuário durante sua interação com uma página web. Foram então identificadas as características de interação de cada usuário com deficiência e criados descritores contendo as características de cada perfil de usuário. Em seguida, foram aplicados dois experimentos através de uma técnica de triagem (*screening*). Logo depois, todas as informações coletadas foram agrupadas com uma técnica de aprendizagem de máquina não-supervisionada (K-Means), implementada através da ferramenta WEKA. No primeiro experimento, alguns elementos não foram agrupados de modo correto pelo algoritmo, levando à correções no processamento dos dados coletados. Embora o segundo experimento tenha rotulado alguns elementos incorretamente, sua acurácia foi maior conforme medido pelo método de matriz de confusão. As contribuições deste trabalho se mostram promissoras, embora estudos mais amplos envolvendo pessoas com real deficiência e com uma quantidade maior de participantes sejam necessários para viabilizar futuramente sua integração com sistemas de interfaces adaptativas. Uma vez realizados os testes adicionais propostos e ajustados os parâmetros do sistema, pode-se considerar o seu uso para que, a partir da classificação do perfil do usuário, sejam aplicadas, em tempo real, alterações em páginas da web para facilitar a navegação.

**Palavras-chave:** Aprendizagem de Máquina. Acessibilidade web. Perfis.



## ABSTRACT

Nowadays, one can safely affirm that the web has a very important role in our society, allowing broad access to knowledge and simplifying the accomplishment of many different tasks. However, improvements are still needed to ensure that all people, having disabilities or not, enjoy full access to information. Based on this need, assistive technology resources are frequently used, enabling people with disabilities to have autonomy in performing tasks. Due to different interaction characteristics in people with different disabilities, it's necessary to adapt presented web content. We performed a literature review that found works which present interaction characteristics for users with disabilities and without disabilities, splitting them in distinct profiles based on interaction manners in websites. This work has the purpose of grouping profiles of blind people, people with motor disabilities and people without any disabilities from their way of interacting with web pages. In order to achieve this goal, a system was developed for collecting user information during its interaction with a web page. We then identified the interaction characteristics of each user with disability and created descriptors containing the characteristics of each user profile. Then, two experiments were applied through a screening technique. Soon after, all the information collected was grouped with a non-supervised machine learning technique (K-Means), implemented through the WEKA tool. In the first experiment, some elements were not correctly grouped by the algorithm, leading to corrections in the processing of the collected data. Although the second experiment had incorrectly labeled some elements, its accuracy was higher as measured by the matrix-of-confusion method. The contributions of this work are promising, although more extensive studies involving people with real disabilities and a larger number of participants are necessary to prove their viability in integrating into real adaptive system interfaces. Once we run additional proposed tests and tune system parameters, we can consider its use so, from classification of user profile, real-time web page adaptations could be applied to facilitate navigation.

**Key-words:** Machine Learning. Web accessibility. Profiles.



## LISTA DE FIGURAS

Figura 1 – <i>Mouse trackball</i> . . . . .	27
Figura 2 – <i>Mouse joystick</i> . . . . .	28
Figura 3 – Teclado colméia . . . . .	28
Figura 4 – Bastão de boca . . . . .	28
Figura 5 – Varinha de cabeça . . . . .	29
Figura 6 – String de busca utilizada . . . . .	39
Figura 7 – Processo para realização do trabalho . . . . .	45
Figura 8 – Diagrama de pacotes com suas classes . . . . .	55
Figura 9 – Modelo conceitual do banco de dados . . . . .	56
Figura 10 – Representação gráfica dos <i>clusters</i> sem modificações . . . . .	57
Figura 11 – Representação gráfica dos <i>clusters</i> . . . . .	59
Figura 12 – Representação gráfica dos <i>clusters</i> . . . . .	63



## LISTA DE TABELAS

Tabela 1 – Avaliação dos trabalhos relacionados. . . . .	43
Tabela 2 – Matriz de confusão . . . . .	50
Tabela 3 – Lista de requisitos funcionais do sistema. . . . .	54
Tabela 4 – Matriz de confusão contendo todos os <i>clusters</i> . . . . .	58
Tabela 5 – <i>Cluster</i> com deficiência visual . . . . .	60
Tabela 6 – <i>Cluster</i> com tremores . . . . .	60
Tabela 7 – <i>Cluster</i> sem deficiência . . . . .	61
Tabela 8 – Matriz de confusão contendo todos os <i>clusters</i> . . . . .	62
Tabela 9 – <i>Cluster</i> com tremores . . . . .	63
Tabela 10 – <i>Cluster</i> de pessoas sem deficiência . . . . .	64



## SUMÁRIO

1	INTRODUÇÃO . . . . .	21
1.1	Objetivo geral . . . . .	22
1.2	Objetivos específicos . . . . .	22
1.3	Organização do documento . . . . .	23
2	FUNDAMENTAÇÃO TEÓRICA . . . . .	25
2.1	Pessoas com deficiência . . . . .	25
2.1.1	Deficiência visual . . . . .	25
2.1.2	Deficiência física . . . . .	26
2.2	Acessibilidade para <i>web</i> . . . . .	30
2.2.1	<i>Design</i> universal . . . . .	31
2.2.2	Sistemas adaptativos . . . . .	32
2.3	Aprendizagem de máquina . . . . .	32
2.3.1	Aprendizagem supervisionada . . . . .	33
2.3.1.1	Regressão . . . . .	33
2.3.1.1.1	Regressão linear . . . . .	33
2.3.1.1.2	Regressão polinomial . . . . .	33
2.3.1.1.3	Regressão logística . . . . .	34
2.3.1.2	Redes neurais . . . . .	34
2.3.1.3	Árvore de decisão . . . . .	34
2.3.1.4	Naïve bayes . . . . .	35
2.3.2	Aprendizagem não-supervisionada . . . . .	35
2.3.2.1	K-Means . . . . .	36
2.3.2.2	Redes neurais . . . . .	36
2.3.2.3	Pairwise clustering . . . . .	36
2.4	Avaliação das técnicas de agrupamento . . . . .	36
3	TRABALHOS RELACIONADOS . . . . .	39
3.1	Resultados da pesquisa . . . . .	40
3.2	Análise dos trabalhos relacionados . . . . .	43
4	FERRAMENTAS E MÉTODOS . . . . .	45
4.1	Investigação das informações a serem coletadas . . . . .	45
4.2	Coleta de dados de interações . . . . .	46
4.3	Modelagem de descritores . . . . .	47
4.4	Escolha e implementação da técnica de agrupamento . . . . .	48
4.5	Comparação e análise de resultados . . . . .	48
4.6	Experimentos . . . . .	49
4.6.1	Experimento 1 . . . . .	50

4.6.2	Experimento 2 . . . . .	51
5	RESULTADOS . . . . .	53
5.1	Sistema de coleta de informações . . . . .	53
5.2	Experimento 1 . . . . .	57
5.3	Experimento 2 . . . . .	61
6	CONSIDERAÇÕES FINAIS . . . . .	67
	REFERÊNCIAS . . . . .	69

## 1 INTRODUÇÃO

A *web* tem se mostrado importante para todos, permitindo o acesso ao conhecimento em nível global, promovendo a comunicação e eliminando barreiras entre os indivíduos (RAMIRO et al., 2015). Com a Internet, é quase sempre possível o acesso a informações a qualquer hora ou lugar, permitindo que tarefas venham a ser realizadas de forma prática, gerando mais qualidade de vida para seus usuários.

O termo acessibilidade trata da inclusão de pessoas com deficiência em atividades realizadas diariamente. Por exemplo, prédios com rampas de acesso para cadeira de rodas e banheiros adaptados para pessoas com deficiência. De acordo com Santarosa et al. (2010) “a acessibilidade na *web* permite flexibilizar o acesso à informação e à interação de usuários que possuam algum tipo de necessidade especial ou encontram-se em desvantagem tecnológica”.

Contudo, ainda é necessário que ocorram modificações para que todos venham a ter acesso à informação. Com base em tal necessidade, foram estabelecidas diretrizes de acessibilidade que, juntamente com o *design* universal, visam garantir a inclusão de pessoas com deficiência. O *design* universal atua na concepção de espaços, artefatos e produtos que tem como propósito atender todas as pessoas com diferentes características, de forma autônoma, segura e confortável (FRANCISCO; MENEZES, 2011). Já as Diretrizes de Acessibilidade para Conteúdo *Web* 2.0, 2.0 (CALDWELL et al., 2015), oferecem recomendações de acessibilidade relacionadas à percepção, compreensão, operação e robustez para o conteúdo na *web*, tornando-o mais acessível para pessoas com deficiências.

Conforme Santarosa et al. (2010), as tecnologias assistivas auxiliam pessoas com deficiência na realização de tarefas, favorecendo a autonomia pessoal, total ou assistida, atuando como recursos de software ou *hardware*, para diversos tipos de necessidades (motora, sensorial ou de comunicação).

“O termo tecnologia assistiva refere-se um conjunto de recursos, que contribuem para proporcionar independência, qualidade de vida, inclusão social para pessoas com deficiência, por meio de um suplemento, da manutenção ou devolução de suas capacidades funcionais” (SANTAROSA et al., 2010)

As tecnologias assistivas são utilizadas por diversos tipos de pessoas com deficiência, com dificuldades motoras sérias, limitações de movimento, falta ou ausência de membros, pouca ou nenhuma visão, entre outras. Para cada tipo de deficiência são estabelecidos métodos, tecnologias ou ferramentas que facilitam a interação do usuário.

De acordo com Santarosa et al. (2010), o uso de tecnologias assistivas vinculadas ao computador apresenta uma possibilidade de adaptação mais abrangente, pois permite que o usuário possa utilizá-las em diferentes espaços. Durante a interação de pessoas com deficiência motora são aplicadas adaptações, como o aumento do tamanho da fonte, configuração da sensibilidade e velocidade do *mouse*, e configuração de teclas de atalho de sistemas de navegação por reconhecimento de voz. Já pessoas com deficiência visual

ou pessoas com baixa visão usam leitores ou ampliadores de tela que possuem diversos recursos de navegação através do teclado.

Diante do modo de interação de usuários com deficiência visual ou motora com o computador, são identificados eventuais padrões de uso ocasionados por tecnologias assistivas como, por exemplo, a utilização de teclas de atalho e navegação sequencial pela página, o que ocorre com pessoas que possuem cegueira e utilizam leitores de tela. Outro padrão de uso está relacionado aos movimentos lineares do *mouse* pela páginas, que são provocados pelo uso de sistemas de navegação por voz, que auxiliam pessoas com deficiência motora a interagir com o computador. Levando em consideração tais aspectos, é possível atribuir um perfil específico de interação para os usuários com deficiência com base em suas características de interação com o computador.

A partir do momento em que é estabelecido um perfil para usuários com deficiência, existem diversos contextos em que esses dados podem ser aplicados, como a realização de avaliações de acessibilidade com base no histórico de interações do usuário ou desenvolvimento de páginas da *web* adaptativas que modificam suas próprias características tornando-se mais adequadas às necessidades específicas de um usuário. Com base em tal particularidade, este trabalho visa identificar aspectos de interação e características de uso de pessoas com deficiência visual ou motora em páginas de *web*, estabelecendo um perfil que distinga cada tipo de usuário de acordo com sua deficiência, permitindo que somente através das características de interação seja identificado se o usuário possui ou não algum tipo de deficiência visual ou motora.

## 1.1 Objetivo geral

Agrupar o perfil de pessoas com deficiência visual, motora ou sem deficiência por meio do seu aspecto de interação com uma página *web*.

## 1.2 Objetivos específicos

- Identificar aspectos de interação que possam ser registrados na página HTML.
- Projetar um modelo de armazenamento das interações;
- Criar um sistema de registro de interações;
- Identificar e escolher técnicas de aprendizagem de máquina para agrupar perfis de interação semelhantes;
- Coletar dados de usuários com deficiência visual, com deficiência motora e sem deficiências;
- Avaliar os resultados por meio de testes preliminares e de um experimentos aplicando uma técnica de triagem (*screening*).

### 1.3 Organização do documento

O trabalho está organizado de acordo com a sequência de capítulos descritos a seguir:

No [Capítulo 2](#) são descritos os conceitos teóricos utilizados para o trabalho, que estão relacionados a pessoas com deficiência, tecnologias assistivas, acessibilidade para *web* e aprendizagem de máquina.

Os trabalhos relacionados são apresentados no [Capítulo 3](#), onde é descrito o processo realizado para a busca dos trabalhos relacionados, juntamente com a descrição de cada trabalho e sua contribuição.

O [Capítulo 4](#) apresenta o método utilizado para condução do trabalho juntamente com a descrição de cada etapa do processo.

No [Capítulo 5](#) são apresentados os resultados do trabalho que contém o desenvolvimento de software para coleta de informações dos usuários durante sua interação com o site e os resultados de dois experimentos.

O [Capítulo 6](#) contém as considerações finais do trabalho, juntamente com suas limitações e propostas que serão implementadas em trabalhos futuros.



## 2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os fundamentos teóricos utilizados como base para o desenvolvimento do presente trabalho; conceitos relacionados a pessoas com deficiência, tecnologias assistivas, acessibilidade para a *web* e aprendizagem de máquina.

### 2.1 Pessoas com deficiência

De acordo com [IBGE \(2010\)](#), 18% da população possui algum tipo de deficiência visual e 7% contém algum tipo de deficiência motora. Para garantir os direitos dessa porcentagem da população foi estabelecido o decreto nº 5.296/04 ([BRASIL, 2009](#)), que busca promover e proteger os direitos de todas as pessoas com deficiência e sua inerente dignidade. O Artigo 1º estabelece que:

“Pessoas com deficiência são aquelas que têm impedimentos de longo prazo de natureza física, mental, intelectual ou sensorial, os quais, em interação com diversas barreiras, podem obstruir sua participação plena e efetiva na sociedade em igualdades de condições com as demais pessoas”. ([BRASIL, 2009](#)).

São previstos no Artigo 4º meios para promover pesquisas para criação de tecnologias, incluindo tecnologias da informação e comunicação, ajudas técnicas para locomoção, dispositivos e tecnologias assistivas (TA) para pessoas com deficiência.

#### 2.1.1 Deficiência visual

Segundo [IBGE \(2010\)](#), pessoas com deficiência visual são caracterizadas com dificuldade permanente de enxergar, mesmo incluindo a utilização de óculos ou lentes de contato. De acordo com o decreto nº 5.296/04 ([BRASIL, 2004](#)), o Artigo 5º inciso I define a deficiência visual como:

“c) Cegueira na qual a acuidade visual é igual ou menor que 0,05 no melhor olho, com a melhor correção óptica; a baixa visão, que significa acuidade visual entre 0,3 e 0,05 no melhor olho, com a melhor correção óptica; os casos nos quais a somatória da medida do campo visual em ambos os olhos for igual ou menor que 60%; ou a ocorrência simultânea de quaisquer das condições anteriores” ([BRASIL, 2004](#)).

Para [Melo \(1991\)](#), a deficiência visual é definida por perdas parciais de visão ou totais que limitam o desempenho normal da visão. Conforme [Ferreira e Rodrigues \(2008\)](#), a deficiência visual é caracterizada pela perda (cegueira) ou diminuição da resposta visual (visão subnormal ou baixa visão).

Existem diversas barreiras para uma pessoa cega que podem interferir na sua participação em sociedade. A maioria dessas barreiras podem ser eliminadas com o uso de uma educação adaptada a sua realidade e o uso de tecnologias ([BORGES, 1996](#)). De acordo com [Conforto e Santarosa \(2000\)](#), o uso de ferramentas computacionais favorece as pessoas cuja a forma de aprendizado não segue os padrões típicos de desenvolvimento. Isso indica que pessoas com deficiências não são menos desenvolvidas; elas somente se

desenvolvem de forma diferenciada. Durante a interação do usuário cego com o computador são utilizadas tecnologias de apoio que são chamados de "leitores de tela" (*screen readers*), que estão associados a sintetizadores de voz. Com base na utilização dos leitores de tela, pessoas cegas têm acesso ao conteúdo da página utilizando somente teclas de atalho que auxiliam na navegação, sem a necessidade de um teclado especial (FERREIRA; RODRIGUES, 2008).

“Usuários com deficiências, ao acessarem um sistema, utilizam um ambiente bem diferente das pessoas sem deficiências; esses usuários criam modelos baseados na maneira peculiar com que interagem com o sistema a fim de suprir suas necessidades” (FERREIRA, 2007)

Por exemplo, no momento em que um usuário com deficiência acessa uma página da *web* com um leitor de tela os elementos da página são apresentados para o usuário de acordo com a sua posição no código fonte, diferente da forma visual do navegador (FERREIRA, 2007). Outro aspecto de interação é o modo de navegação sequencial utilizado pelo leitor de tela, que restringe o acesso do usuário a uma porção limitada da tela, perdendo a noção do contexto geral da página (LEPORINI; ANDRONICO; BUZZI, 2004). Já o uso de tecnologias adicionais como *Flash*<sup>1</sup> são detectadas pelos leitores de tela ocasionando conflitos e dificultando a compreensão do usuário em relação a página (FERREIRA, 2007). Conforme Leporini, Andronico e Buzzi (2004), quando sites apresentam menus e *banners* repetidos em todas as páginas antes do conteúdo principal, a navegação se torna lenta e cansativa. Outras barreiras encontradas pelos usuários cegos estão relacionadas a imagens que não contêm o equivalente textual (conhecido como texto alternativo), gráficos ou mapas que não possuem texto para entendimento, vídeos que não possuem descrição textual ou sonora, tabelas que não apresentem significados quando lidas de forma linear e formulários que não podem ser navegados em sequência lógica ou não estão devidamente rotulados (SONZA; KADE; FAÇANHA, 2013). De acordo com Ferreira (2007), quanto mais simples e linear a interface do site, melhor será processado o conteúdo pelo leitor de tela.

### 2.1.2 Deficiência física

Conforme (BRASIL, 2004), no Artigo 5º, inciso I, define a deficiência física como:

“a) Alteração completa ou parcial de um ou mais segmentos do corpo humano, acarretando o comprometimento da função física, apresentando-se sob a forma de paraplegia, paraparesia, monoplegia, monoparesia, tetraplegia, tetraparesia, triplegia, triparesia, hemiplegia, hemiparesia, ostomia, amputação ou ausência de membro, paralisia cerebral, nanismo, membros com deformidade congênita ou adquirida, exceto as deformidades estéticas e as que não produzam dificuldades para o desempenho de funções”.

<sup>1</sup> Plataforma multimídia de desenvolvimento de aplicações que contenham animações, áudio e vídeo; utilizado para construção de páginas *web* interativas.

De acordo com [Ferreira e Rodrigues \(2008\)](#), a deficiência física consiste na alteração completa ou parcial de uma ou mais partes do corpo, comprometendo funções físicas. Para [Sonza \(2008\)](#) são todas aquelas pessoas que são impedidas de realizar plenos movimentos. Segundo [Godinho et al. \(2004\)](#), algumas características frequentes são a redução no controle de movimentos, fraquezas musculares e a dificuldade na realização de ações que impliquem precisão e rapidez.

A principal barreira encontrada por pessoas com deficiência motora encontra-se no uso *mouse* no teclado ([SANTAROSA et al., 2010](#)). Para auxiliar a interação das pessoas com deficiência física, são utilizados recursos tecnológicos, como *mouse* ou apontadores, por exemplo, o *mouse trackball* apresentado na [Figura 1](#) e *mouse joystick* apresentado na [Figura 2](#), softwares de reconhecimento de voz, teclados alternativos, como o teclado colmeia apresentado na [Figura 3](#) e dispositivos em forma de ponteiro o qual é fixado na cabeça apresentado na [Figura 4](#) e o ponteiro de boca apresentado na [Figura 5](#) ([FERREIRA; RODRIGUES, 2008](#)).

Figura 1 – *Mouse trackball*



Fonte: [www.assist-it.org.uk](http://www.assist-it.org.uk) (2016)

Durante o processo de interação com o *hardware*, algumas pessoas possuem dificuldade como pressionar duas teclas ao mesmo tempo, pressionar várias teclas ou repetir letras ([GODINHO et al., 2004](#)). Acrescenta-se também o acesso a páginas *web* que possuem tempo de utilização limitado, páginas que possuem diversas janelas sobrepostas, navegadores ou ferramentas que não possuem suporte ao teclado alternativo ou botões que executem ações referentes a comandos do *mouse* e formulários que não podem ser navegados com o uso da tecla *tab* em sequência lógica ([SONZA, 2008](#)). O mesmo ocorre com pessoas diagnosticadas com a doença de Parkinson que possuem sintomas, como lentidão, rigidez e tremores em movimentos rítmicos geralmente nos membros superiores, o que por sua vez afeta a sua habilidade de utilizar os dispositivos de entrada do computador ([HARTIKAINEN; OVASKA, 2015](#)).

Figura 2 – *Mouse joystick*

Fonte: [www.webaccessivel.wordpress.com](http://www.webaccessivel.wordpress.com) (2016)

Figura 3 – Teclado colméia



Fonte: [www.clik.com.br](http://www.clik.com.br) (2016)

Figura 4 – Bastão de boca



Fonte: [www.brasilmedia.com](http://www.brasilmedia.com) (2016)

Figura 5 – Varinha de cabeça



Fonte: [www.brasilmedia.com](http://www.brasilmedia.com) (2016)

“Recursos de Tecnologia Assistiva (TA) potencializam as capacidades e as habilidades de seus usuários do desempenho das atividades cotidianas com autonomia e independência” (MELO, 2014), promovem maior qualidade de vida, autonomia e inclusão social. As formas de aplicação de tecnologias assistivas abrangem desde tarefas básicas de autocuidado até a realização de atividades profissionais (AJUDAS TÉCNICAS, 2006). Tratando-se de uma área multidisciplinar de conhecimento, são desenvolvidos estudos, produtos e pesquisas visando promover a inclusão social de pessoas com deficiência (SANTAROSA et al., 2010). Exemplos desses estudos estão relacionados ao *hardware* e *software* que visam garantir a inclusão social. De acordo com Santarosa et al. (2010), vão de uma simples bengala, cadeiras de rodas ou um par de óculos, até sistemas complexos que, permitindo domínio maior sobre o ambiente, possibilitam a caracterização de expressões do indivíduo.

Existem diversos tipos de tecnologias assistivas aplicadas sobre o contexto computacional. Por exemplo, leitores de tela, que auxiliam na leitura de informações textuais (MELO; COSTA; SOARES, 2006), juntamente com os sintetizadores de voz responsáveis por disponibilizar a informação em forma de som. Esse tipo de tecnologia é geralmente utilizada por pessoas que possuem deficiência visual acentuada ou total (FERREIRA; RODRIGUES, 2008). Já os ampliadores de tela são softwares que ampliam o conteúdo apresentado na tela do computador, de modo que pessoas com baixa visão, tenham acesso a gráficos ou elementos textuais em tamanho ampliado (MELO; COSTA; SOARES, 2006). Sistemas de reconhecimento de voz, que são softwares que permitem o acesso através da fala, com o reconhecimento de comandos (SANTAROSA et al., 2010), podem ser utilizados por pessoas que estejam com mobilidade dos membros superiores reduzida ou comprometida (MELO; COSTA; SOARES, 2006). Teclados alternativos que simulam o funcionamento do teclado convencional através de dispositivos físicos ou *softwares* de computador, oferecendo uma alternativa para o acionamento de teclas, alguns tipos desse

dispositivo possuem um espaçamento menor ou maior entre as teclas, como também protetores de tecla, para uma tecla ser acionada por vez.

## 2.2 Acessibilidade para *web*

O termo acessibilidade, para [Melo \(2014\)](#), trata da criação de ambientes, produtos e serviços que levem em consideração as diferenças de cada pessoa, visando atender todas as normas e padrões necessários, para que todos venham a ter acesso. De acordo com [Nicholl e Filho \(2001\)](#), a acessibilidade permite que qualquer pessoa independente da sua capacidade físico-motoras, perceptivas, culturais e sociais possam usufruir de uma vida em sociedade. Conforme o decreto nº 5.296/04 ([BRASIL, 2004](#)), no Artigo 8º, inciso I, estabelece a acessibilidade, como:

“Acessibilidade: condição para utilização, com segurança e autonomia, total ou assistida, dos espaços, mobiliários e equipamentos urbanos, das edificações, dos serviços de transporte e dos dispositivos, sistemas e meios de comunicação e informação, por pessoa portadora de deficiência ou com mobilidade reduzida”.

De acordo com [Ferreira e Rodrigues \(2008\)](#), a acessibilidade na *web* possibilita que “qualquer pessoa que apresente alguma deficiência ou necessidade especial venha a ter acesso aos seus recursos”. Para [Cusin e Vidotti \(2009\)](#), significa que pessoas com deficiência possam compreender, entender, navegar e interagir e contribuir com a *web*. Conforme [HENRY \(2005\)](#), a acessibilidade na *web* é definida como:

“Acessibilidade na *Web* é a possibilidade e a condição de alcance, percepção, entendimento e interação para a utilização, a participação e a contribuição, em igualdade de oportunidades, com segurança e autonomia, em sítios e serviços disponíveis na *web*, por qualquer indivíduo, independentemente de sua capacidade motora, visual, auditiva, intelectual, cultural ou social, a qualquer momento, em qualquer local e em qualquer ambiente físico ou computacional e a partir de qualquer dispositivo de acesso ([HENRY, 2005](#))”

A acessibilidade na *web* pode ser compreendida como uma forma de garantir a construção de páginas *web* que todos tenham acesso ao conteúdo independentemente do modo de interação do usuários com suas necessidades e preferências ([THATCHER; WADDELL; BURKS, 2002](#)). Com base nesse princípio, atua o *design* universal, na construção de espaços, artefatos e produtos que visam atender simultaneamente todas as pessoas, juntamente com suas características físicas, sensoriais e cognitivas ([FRANCISCO; ME-NEZES, 2011](#)).

Cabe destacar que garantir a acessibilidade na *web* não é algo trivial. Devem ser consideradas adaptações de forma a compensar as necessidades sensoriais e motoras das pessoas com deficiência ([SONZA, 2008](#)). Tendo em vista que a maior parte dos aplicativos na *web* são utilizados por diferentes tipos de usuários, dificilmente um site construído de forma estática poderia satisfazer seu usuário, o que sugere o desenvolvimento de interfaces

adaptativas ou sistemas adaptativos, que são estruturados de acordo com a forma de navegação do usuário (OLIVEIRA et al., 2003).

### 2.2.1 *Design* universal

O Desenho Universal ou *Design* Universal prevê a criação de produtos e ambientes que venham a ser utilizados pelo maior número de pessoas possível (BRAGA, 2014). Do mesmo modo, o conceito de acessibilidade *web* busca:

- Proporcionar o acesso à todos, independente da ferramenta utilizada ou modo de interação do usuário;
- Desenvolver sistemas acessíveis ao maior número de usuários (SONZA, 2008);
- Permitir o acesso de diferentes tipos de pessoas ao conteúdo do *website*, independente das suas eventuais características de interação;
- Promover a inclusão e beneficiando a todos e não somente uma minoria;
- Estabelecer produtos e ambientes que sejam adequados para um amplo número de pessoas, considerando as suas diferenças quanto a percepção auditiva, visual, forma de locomoção, controle de movimentos, altura, peso e maneira de se comunicar (PUPO; MELO; FERRÉS, 2006).

Segundo Correia, Correia e Frassinetti (2005) o conceito de *design* universal pressupõe o aumento da produtividade, prevenção de riscos relacionados à má utilização e desconforto, beneficiando não apenas uma minoria, mas a toda a população. Deve conter um conjunto de preocupações, conhecimentos, metodologias e práticas que visam à concepção de espaços, produtos e serviços, utilizáveis com eficácia, segurança e conforto para as pessoas independentemente das suas capacidades.

No entanto, de acordo com Dias (2003) devemos compreender que nem todos os recursos ou produtos podem ser utilizados por todas as pessoas. Sempre haverá uma pessoa com deficiência, com necessidades específicas de uso, que não será capaz de utilizar o recurso ou produto. “Apesar de ser possível modelar tipos de usuários, sob seu ponto de vista do funcionamento cognitivo e de seu comportamento, as pessoas são únicas em suas experiências, seus sentimentos e suas expectativas ” (CYBIS; BETIOL; FAUST, 2007). De acordo com Oliveira et al. (2003), devemos considerar que parte dos aplicativos da *web* são utilizados, por usuários com diferentes preferências e níveis de conhecimento. Por isso dificilmente um site construído de forma estática poderia satisfazer a todos seus usuários, existindo a necessidade de criar interfaces que acomodem as diferenças interpessoais de cada pessoa ou interfaces que adaptem de acordo com o usuário (CYBIS; BETIOL; FAUST, 2007).

### 2.2.2 Sistemas adaptativos

“Sistemas adaptativos são aqueles que podem alterar aspectos de sua estrutura, funcionalidades ou interface para acomodar as necessidades e diferenças individuais ou de um grupo de usuários” (SILVA; SILVA, 2007). De acordo com Oppermann (1994), diversas pesquisas vem sendo desenvolvidas para garantir que os usuários realizem adaptações em seus sistemas, denominando-os adaptáveis, customizáveis ou extensíveis.

Conforme Oliveira et al. (2003), “uma das principais razões que levam ao desenvolvimento de aplicativos com interfaces adaptativas refere-se à tentativa de estruturar a interface de forma a facilitar a navegação do usuário”. Para Cybis, Betiol e Faust (2007), interfaces adaptativas inteligentes, identificam estilos cognitivos, objetivos, necessidades e preferências, deficiências e outros aspectos dos usuários, baseando-se na análise de padrões e comportamento, mantendo diversas características dos seus usuários, tais como objetivos, interesses, preferências, nível de conhecimento, etc (PALALLO, 2000). Quando uma interface é bem definida, pode atuar como uma forma de motivar o usuário a utilizar o produto, caso contrário pode atuar como um fator negativo (VIEIRA; PONTES; PALAZZO, 2002). Percebe-se que a adaptação da interface é fundamental quando se trata de apoiar diversos grupos de usuários, incluindo pessoas com deficiência, proporcionando interação mais simples e satisfatória (ENCARNAÇÃO, 1997), possibilitando a criação de *web* sites de acordo com os padrões de acesso do usuário, com páginas mais acessíveis, ou destacando conteúdos interessantes (PERKOWITZ; ETZIONI, 2000).

No entanto, é necessário informar ao usuário sobre as mudanças que serão realizadas, permitindo que ele as desfaça ou até mesmo impeça novas adaptações. Sistemas adaptáveis e adaptativos ainda são um grande desafio para as atividades de relacionadas a Interação Humano-Computador (IHC), desafios esses que envolvem identificar as necessidades e oportunidades de adaptação, sendo adaptações manuais ou automáticas, em diferentes contextos. (SILVA; SILVA, 2007).

### 2.3 Aprendizagem de máquina

“Aprendizagem de máquina é uma área da inteligência artificial que tem por objetivo desenvolver algoritmos e técnicas computacionais que permitam que o computador seja capaz de aprender” (GOLDSCHMIDT, 2010). Para Monard e Baranauskas (2003) “trata-se de sistemas de computador que tomam decisões baseadas em experiências, que são acumuladas a partir de soluções bem sucedidas de problema anteriores”. Conforme Goldschmidt (2010), os algoritmos de aprendizagem de máquina são usualmente treinados através da apresentação de um conjunto de exemplos, os quais consistem em vetores de características ou atributos que descrevem estados das entradas do problema em questão. Um dos principais métodos utilizados para se obter conclusões genéricas sobre um determinado exemplo é a indução, sendo esta uma forma de inferência lógica (MONARD;

BARANAUSKAS, 2003). Durante o aprendizado indutivo, o conhecimento é abstraído a partir de sucessivas iterações sobre os exemplos disponíveis, que podem ocorrer através de uma aprendizagem supervisionada ou não-supervisionada.

### 2.3.1 Aprendizagem supervisionada

Na aprendizagem supervisionada, os exemplos disponíveis na etapa de treinamento do algoritmo contêm as amostras dos resultados que se espera produzir com o modelo de conhecimento que está sendo desenvolvido (GOLDSCHMIDT, 2010). O processo de treinamento requer supervisão, através da qual, após informados os exemplos de cada classe e seus resultados esperados, é realizado um ajuste interno de parâmetros do modelo. Ao final do treinamento, espera-se que o modelo seja capaz de classificar novos objetos nos grupos previamente apresentado (DOMINGUES, 2003). A seguir serão apresentados os conceitos sobre algumas técnicas de aprendizagem de máquina supervisionada.

#### 2.3.1.1 Regressão

De acordo com Stevenso (1981), a “ regressão compreende a análise de dados amostrais para saber se e como duas ou mais variáveis estão relacionadas umas com as outras numa população”. Já para Peternelli (2012), é usada para realizar estimativas entre variáveis dependentes, com uma ou mais variáveis independentes através de uma análise estatística. Ou seja, uma equação que tenta estabelecer a variação da variável dependente (X) pela variação do nível da variável independente (Y). Para isso é utilizado um diagrama de dispersão que mostra graficamente o comportamento das variáveis X em função das variáveis de Y que podem ser representadas de diversas maneiras, seja na forma linear, logística, polinomial, entre outras.

##### 2.3.1.1.1 Regressão linear

Segundo Stevenso (1981), “A regressão linear constitui uma tentativa de estabelecer uma equação matemática linear (linha reta) que descreva o relacionamento entre duas variáveis”. Conforme Hoffmann et al. (2015), a regressão linear determina características da função que relaciona X e Y que, no caso do modelo linear, é representada através de uma reta. De acordo com Stringhini (2005), o modelo linear ou equação é usado para encontrar o modelo matemático que se ajuste aos dados obtidos no passado e através deles simular e projetar possíveis cenários do futuro.

##### 2.3.1.1.2 Regressão polinomial

A regressão polinomial uma variação da regressão linear, podemos dizer que uma regressão linear simples é como uma regressão polinomial de um polinômio de grau um.

É uma forma de regressão em que a relação entre a variável independente  $X$  e a variável dependente  $Y$  é modelada como um grau  $N$ -polinomial.

### 2.3.1.1.3 Regressão logística

O modelo de regressão logística permite o ajuste de variáveis independentes a uma variável de resposta categórica (MENDES; VEGA, 2011). As variáveis categóricas podem assumir somente alguns valores como resposta, sendo binários (dicotômicos) cujo a resposta é falha ou sucesso ou politômicos que podem assumir múltiplos níveis de saída (JR; LEMESHOW, 2000). Para Mesquita e Nogueira (2015), “ a regressão logística descreve a relação entre uma variável dependente qualitativa binária, associada à um conjunto de variáveis independentes qualitativas ou métricas”. Ainda, segundo o autor, a regressão logística assume a forma parecida com um “S” em sua representação gráfica, existindo áreas onde a mudança é mais elevada e outras não.

### 2.3.1.2 Redes neurais

No aprendizado supervisionado, a rede neural do tipo *feed-forward* recebe um conjunto de entradas e um conjunto de saídas desejadas para cada entrada. Toda vez que uma nova entrada for apresentada a rede, é verificado se a saída obtida confere com a saída desejada para aquela entrada. Caso contrário, são ajustados os pesos da rede de forma que o aprendizado seja concluído. Para isso, são efetuados cálculos utilizando os valores dos pesos sinápticos da rede. Esse processo deverá ser repetido até que a taxa de acerto seja considerada satisfatória (ALVAREZ; LUQUE, 2003). Conforme Goldschmidt (2010), “ a cada par (entrada, saída desejada), um “professor” compara a saída desejada com a saída produzida pelo sistema, calculando um erro que é usado no ajuste dos pesos da rede a fim de minimizar tal erro”. De acordo com Silva (1998), as redes neurais podem ser aplicadas com sucesso a uma grande variedade de problemas de processamento de informação, identificação de padrões, aproximação de funções e previsão de séries temporais. No entanto, por se tratar do aprendizado supervisionado, é necessário completo domínio sobre os dados que serão utilizados durante o treinamento da rede.

### 2.3.1.3 Árvore de decisão

Tratam-se de modelos estatísticos aplicados sobre problemas de aprendizado de máquina e predição supervisionada. Para isso, é utilizado um conjunto de atributos capazes de predizer o valor do atributo de saída (ONODA; EBECKEN, 2001), adquirindo o valor de uma variável de destino com base em um conjunto de variáveis obtidas como entrada (BELL, 2014). O processo que ocorre entre as saída e as entradas é chamado de modelo preditivo (LIMA; ASSIS; SOUZA, 2010), onde os dados usados são um conjunto de casos, observações e exemplos (ONODA; EBECKEN, 2001). A estrutura do modelo

preditivo é semelhante a uma árvore ou expressa através de um conjunto de regras (LIMA; ASSIS; SOUZA, 2010).

A leitura de uma árvore de decisão sempre é iniciada pelo nó raiz, onde cada nó (não folha) indica um teste feito sobre um valor. As ligações entre os nós representam os valores possíveis do teste do nó superior, e as folhas indicam a classe à qual o registro pertence. Após montada a árvore de decisão, para classificarmos um novo registro, basta seguir o fluxo na árvore (mediante os testes nos nós não-folhas) começando no nó raiz até chegar a uma folha. O sucesso da árvore de decisão deve-se ao fato de ser uma técnica extremamente simples, não necessitando de parâmetros de configuração e geralmente tem um bom grau de assertividade. Apesar de ser uma técnica extremamente poderosa, é necessária uma análise detalhada dos dados que serão usados para garantir bons resultados (CAMILO; SILVA, 2009).

#### 2.3.1.4 Naïve bayes

O algoritmo *Naïve Bayes* faz uso de fórmulas estatísticas e cálculos probabilísticos para realizar as suas classificações. De acordo com o teorema de Bayes, é possível encontrar a probabilidade de um certo evento ocorrer, conforme a probabilidade de um outro evento que já ocorreu (CAMILO; SILVA, 2009). As classificações são baseadas em evidências fornecidas, que podem aumentar ou diminuir a probabilidade de as classes serem observadas sobre uma nova instância de classificação. São calculadas as probabilidades de todas as possíveis classes para determinar a classe mais provável para uma nova instância. No entanto, é atribuído o nome “ingênuo” porque todos os atributos classificados são considerados independentes, pois, sabe-se que na maioria dos casos essa suposição de independência dos atributos é falsa, e mesmo assim são frequentemente produzidos resultados satisfatórios ainda quando os atributos são consideravelmente independentes (PARDO; NUNES, 2002).

#### 2.3.2 Aprendizagem não-supervisionada

Durante a aprendizagem não-supervisionada, os exemplos são agrupados de acordo com as suas inter-similaridades, de forma que se auto organizem em um certo número de grupos de elementos mais semelhantes entre si (GOLDSCHMIDT, 2010). Neste caso, os exemplos não contêm resultados esperados (rótulos), e o número de classes pode ser desconhecido, sendo necessário que o algoritmo identifique como os objetos podem ser agrupados, ficando a cargo do sistema deduzir quais os objetos pertencentes as mesmas, além de que, em algumas abordagens, também o algoritmo deve inferir o número de classes (DOMINGUES, 2003). A seguir serão apresentados conceitos sobre técnicas de aprendizagem de máquina não-supervisionadas:

### 2.3.2.1 K-Means

É um dos algoritmos mais simples e difundidos, onde primeiramente é estabelecido o número de grupos que serão utilizados durante o processo. Posteriormente, são escolhidos de forma aleatória pontos responsáveis pela representação dos centróides de cada grupo. Na etapa subsequente os elementos são separados em grupos de acordo com o centróide mais próximo, que é estabelecido normalmente com base no cálculo da distância euclidiana comum. A cada iteração realizada pelo algoritmo é estabelecido um novo centróide conforme os elementos do grupo, que em seguida são reassociados ao novo centróide mais próximo. Esse processo é realizado até que os elementos sejam atribuídos aos mesmos grupos, de forma que os centróides sejam os mesmos (COSTA et al., 2013).

### 2.3.2.2 Redes neurais

Neste tipo de aprendizagem, o conhecimento é adquirido através de um conjunto de dados, de modo que não há para cada entrada uma saída desejada. No decorrer do processo é utilizada a similaridade dos dados para realizar o aprendizado, dados esses que são constantemente apresentados a rede, tornando a regularidade e a redundância das entradas características essenciais para o aprendizado. Durante o treinamento são modificados os pesos da rede de forma a produzir saídas que sejam consistentes, extraindo propriedades estatísticas do conjunto de treinamento e agrupando vetores similares em classes. Outro aspecto relacionado a esse modo de aprendizagem é não requerer um vetor alvo para saída e não realizar comparações para determinar a melhor resposta (GOLDSCHMIDT, 2010).

### 2.3.2.3 Pairwise clustering

Trata-se de método de otimização combinatória para agrupamento de dados que extrai estrutura a partir de dados de proximidade (HOFMANN; BUHMANN, 1997), apresentando a soma das dissimilaridade médias internas relacionadas a cada *cluster* (THIELO, 2000). Requer apenas um par de vizinhos ativos a cada iteração e estes atualizam as suas estimativas para a média de seus valores (AMARAL, 2013).

## 2.4 Avaliação das técnicas de agrupamento

Inicialmente foram identificadas quais as técnicas de agrupamento seriam relevantes para auxiliar na classificação do perfil dos usuários, que contém informações relacionadas ao modo de interação dos usuários com páginas da *web*. Foram então comparadas quais seriam as vantagens e desvantagem de aplicação de cada técnica de agrupamento e estabelecida qual seria a técnica utilizada para solucionar o problema.

Entre as técnicas analisadas estão a regressão linear, logística e polinomial que apresentam diferentes vantagens e desvantagens de implementação. Por exemplo, a re-

gressão linear é uma técnica de fácil implementação, mas que só deve ser aplicada sobre dados de comportamento linear. A regressão logística se aplica a classificações de informações simples, determinando se a informação analisada está ou não relacionada a um domínio específico. Já a regressão polinomial é usada para informações de comportamento mais complexo que a técnica de regressão linear, pois pode ser adaptada de acordo com a distribuição das informações analisadas. Por outro lado, as redes neurais supervisionadas apresentam uma grande precisão na sua classificação, tolerância a ruídos em amostras e não requerem modelos matemáticos do problema. No entanto, é difícil de incorporar os conhecimentos do domínio e realizar a interpretação dos pesos, além da frequente necessidade de um tempo de treinamento. A árvore de decisão auxilia na compreensão do problema proposto, sendo considerada simples de implementar e deve ser utilizada para um conjunto de tamanho médio de informações, pois quando aplicada a um grande número de dados perde a sua capacidade de classificação tornando o algoritmo pesado devido as *features*.

O *Naïve Bayes* possui modo de treinamento e classificação rápidos, lida com dados reais, discretos e contínuos e não é sensível a características irrelevantes, assumindo independência entre as características, obtendo resultados satisfatórios quando aplicado em problemas complexos. No *K-Means* todas as informações utilizadas são atribuídas a um grupo, sendo que a localização do centroíde do grupo pode variar, o que permite estabelecer condições iniciais de dependência. No entanto é necessário informar o número de grupos antes do início do algoritmo que atribuí cada informação a um grupo específico.

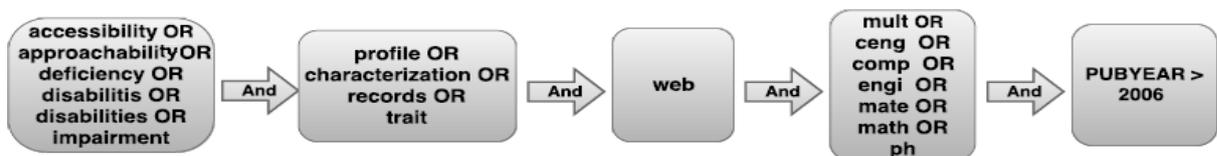


### 3 TRABALHOS RELACIONADOS

Nesta seção é descrito o processo de revisão bibliográfica usado para obter os trabalhos relacionados. O processo de revisão bibliográfica utilizado é uma adaptação de Petersen et al. (2008) aplicado na seleção dos trabalhos. Inicialmente, foi estabelecido o objetivo da pesquisa e logo depois foram formuladas as questões de pesquisa, responsáveis por direcionar o processo de busca.

A primeira questão **“Quais características dos usuários são identificadas a partir dos eventos de interação?”** tem como objetivo identificar a existência de trabalhos que extraíam as características do usuário somente através do seu modo de interação. Já a segunda questão **“É possível identificar uma pessoa com deficiência através dos eventos de interação com o site?”** consiste na identificação de trabalhos que apresentem padrões de uso de pessoas com deficiência, suficientes para identificá-las. A partir das questões foram identificados termos de sinônimos correspondentes e criada uma expressão de busca:

Figura 6 – String de busca utilizada



Fonte: a autora

Para seleção dos trabalhos foram aplicados critérios de inclusão e exclusão. De acordo com Petersen et al. (2008), “os critérios de exclusão são usados para excluir estudos que não são relevantes para responder às questões de pesquisa”, sendo critérios de inclusão requisitos que os trabalhos devem apresentar para serem incluídos na pesquisa. Foram adotados os seguintes:

- Os trabalhos devem conter no título ou no resumo pelo menos um dos termos relacionados ao domínio do objetivo da busca;
- Os trabalhos devem estabelecer um perfil para os usuários incluindo suas respectivas características;
- Trabalhos publicados a partir de 2006, disponíveis para download, de forma gratuita ou através de bases de dados acessíveis pela UNIPAMPA.

Já os critérios de exclusão correspondem a características que o estudo analisado não deve possuir. Foram estabelecidos os seguintes critérios de exclusão:

- Os trabalhos que não estejam escritos em português ou inglês;
- Os trabalhos que não apresentarem características de usuários com deficiência.

Em seguida foi definida a base de dados onde seriam realizadas as buscas dos trabalhos. A base de dados utilizada foi a Scopus, pois indexa outras bases de pesquisa como: IEEE (Institute of Electrical and Electronics Engineers) Xplore Digital Library <sup>1</sup>, ACM (Association Computing Machinery) Digital Library <sup>2</sup>, Science Direct <sup>3</sup> e Springer <sup>4</sup>.

Durante a realização da busca foram estabelecidas quatro etapas para guiar o processo de seleção dos trabalhos. A primeira etapa era a realização da busca com a expressão que obteve 201 resultados. Na segunda etapa foram analisados o título e resumo, na qual foram selecionados 40 resultados considerados relevantes. A terceira etapa tratou de avaliar os trabalhos com base na leitura da introdução e conclusão, dos quais somente 26 estavam disponíveis para download gratuito. Em seguida na quarta etapa foram analisados 5 trabalhos, onde foram verificados seus objetivos e tecnologias utilizadas e, a partir de tal análise, foi identificada sua relevância.

### 3.1 Resultados da pesquisa

Nesta seção são descritos os resultados obtidos durante a revisão bibliográfica. Na sequência são apresentadas informações referentes a cada trabalho, juntamente com seus objetivos, técnicas ou ferramentas e resultados, como também sua contribuição para o presente trabalho.

A ferramenta proposta por [Vigo et al. \(2008\)](#) tem como objetivo a coleta de dados sobre as tecnologias de apoio utilizadas pelos usuários durante a sua interação. Para isso são capturadas as informações referentes ao *hardware* ou *software*. Ainda, segundo o autor, a finalidade da ferramenta é a realização de avaliações de acessibilidade com base no usuário, como também a adaptação do conteúdo da *web* de acordo com o perfil do usuário. Durante o processo de coleta de dados é realizada uma consulta em um banco de dados remoto para verificar eventuais atualizações, que posteriormente são transferidas para o banco de dados local. Na sequência, é verificada a existência de alguma tecnologia assistiva pelo *Assistive Technologies Detector*, juntamente com informações relacionadas ao computador utilizado pelo usuário com o *Access Device features Detector*, como: plataforma de acesso, sistema operacional, tipo de teclado, resolução de tela dentre outras informações. As informações recuperadas são colocadas em um arquivo CC/PP <sup>5</sup> (*Composite Capability/Preference Profiles*), que descreve a capacidade de dispositivos

<sup>1</sup> Disponível em <<http://ieeexplore.ieee.org/Xplore/home.jsp>>

<sup>2</sup> Disponível em <<http://dl.acm.org/>>

<sup>3</sup> Disponível em <<http://www.sciencedirect.com/>>

<sup>4</sup> Disponível em <<http://www.springer.com/br/>>

<sup>5</sup> Disponível em <<https://www.w3.org/TR/CCPP-struct-vocab/>>

e preferências do usuário, que são enviadas para o perfil no servidor. No entanto, recolher informações relacionadas ao tipo de tecnologia assistiva utilizada pelo usuário não é considerado o suficiente para atribuí-lo à um grupo específico de pessoas com deficiência, devido a necessidades individuais de cada usuário. Complementando, este trabalho é relevante porque apresenta uma maneira de capturar informações de uso relacionadas aos usuários com deficiência, salientando que somente determinar a tecnologia assistiva não é suficiente para indicar a deficiência do usuário, evidenciando a necessidade de obter outras informações sobre o usuário para então poder classificá-lo.

O trabalho proposto por [Ferretti et al. \(2014\)](#) descreve um sistema capaz de adaptar automaticamente elementos textuais de páginas da *web* de acordo com o perfil do usuário, levando em consideração suas preferências ou tipo de deficiência, melhorando a legibilidade do conteúdo, como: tamanho da fonte, tipo de fonte contraste entre outros elementos da página. O *EXperiential TRAnscoding System* permite que o conteúdo da *web* seja alterado, monitorando o comportamento dos usuários. Com base nisso são modeladas as preferências de cada usuário. Para compreender a forma de interação do usuário, é explorado o mecanismo de aprendizagem de máquina. Para isso foi utilizado o algoritmo *Q-learning*, que se concentra no desenvolvimento de uma ação específica em um estado específico. Para testar a viabilidade do sistema foi usado um simulador PIQUE (*Platform for Implementing Q-Learning Experiments*) que reproduz a navegação do usuário com o conteúdo da *web*. A simulação foi realizada com base em usuários com dislexia, baixa visão e usuários idosos. De acordo com os resultados obtidos durante a simulação é constatada a viabilidade do sistema. No entanto, é necessária maior dedicação em relação ao perfil do usuário, ou seja, em descrever suas preferências sobre as características textuais. Como trabalho futuro é sugerida a combinação de *Q-learning* com outras técnicas, como árvore de decisão e a comparação de técnicas empregadas para geração de perfis. Complementando, esse trabalho é relevante pois utiliza a técnica de aprendizado de máquina para gerar perfil do usuário, reforçando a viabilidade da técnica que posteriormente será aplicada no presente trabalho.

Segundo [Bigham et al. \(2007\)](#), é proposto um estudo para analisar a acessibilidade na *web* a partir da perspectiva do usuário. São então analisados o modo de interação de usuário cegos e usuários com visão, medindo a acessibilidade das páginas visitadas. Para gravar a navegação foi usado o *UsaProxy* que utiliza a tecnologia *Javascript* para gravar as páginas visualizadas e as ações tomadas pelo usuário, como cliques em elementos, uso de botões entre outras ações. O estudo foi conduzido durante o período de uma semana, onde os participantes fizeram uso de tecnologias de apoio sem alterar a configuração já existente. Todos os dados foram armazenados em um banco de dados remoto e poderiam ser acessados pelos participantes, permitindo comentários sobre as páginas visitadas ou sobre a experiência e até mesmo excluir partes do seu histórico de navegação. No total foram recrutados 10 participantes cegos com a idade variando entre 18 e 63 anos e 10

participantes com visão na faixa etária de 19 a 61 anos. A abordagem centrada no usuário proporcionou uma visão de acessibilidade sobre problemas que realmente precisam ser solucionados, mas que somente poderiam ser identificados com a interação do usuário. Conclui-se que esse trabalho foi relevante, por identificar os usuários a partir dos eventos realizados sobre os elementos da página, sem a instalação de nenhum recurso adicional utilizando somente *Javascript* para gravar os eventos do usuário, comprovando a eficácia da tecnologia para recuperar informações do usuário.

A ferramenta WELFIT, proposta por [Santana e Baranauskas \(2015\)](#) implementada em *JavaScript* trata da avaliação de *websites*, sendo um projeto de longo prazo distribuído em diversos artigos. Apresenta diferenças entre o fluxo de eventos ocorridos durante a interação de pessoas com ou sem deficiência para com o *website*, permitindo que sejam capturados todos eventos do lado do cliente, como a identificação de comportamentos que sugerem o uso de tecnologias assistivas. Os dados recolhidos no lado cliente são enviados para o lado servidor que realiza o armazenamento. Durante a avaliação é necessário que o avaliador seja registrado e informe a página que deseja avaliar, juntamente com o consentimento do avaliado em participar da avaliação, informando se fará uso de alguma tecnologia assistiva durante o processo. Ao todo foram avaliados 180 participantes durante o período de 15 meses. A ferramenta apresentada neste trabalho abre novas possibilidades para a aplicação de técnicas de logs de eventos do lado do cliente. Com a avaliação contínua suportada pela ferramenta, foi possível a identificação de padrões de uso da *web* por pessoas com ou sem deficiência. Os trabalhos futuros consistem em explorar o uso de eventos para realizar adaptações na interface. Complementando, a ferramenta descrita anteriormente possui relevância para o presente trabalho, evidenciando a diferença entre os eventos realizados por pessoas com deficiência ou sem deficiência, permitindo que seja feita uma distinção entre os dois grupos de usuários, durante a interação com o conteúdo da *web*.

O trabalho proposto por [Arbelaitz et al. \(2014\)](#) tem como objetivo projetar um sistema para contribuir com a navegação dos usuários que utilizam o site *Discapnet*. O sistema proposto é baseado na forma de comportamento dos usuários juntamente com processo de mineração *web* com os dados obtidos, no log do servidor *web*. Foram implementadas duas abordagens para criação de perfis um sistema global construído com base em todo o site e uma abordagem modular para descobrir perfis dentro de cada zona de navegação do site. Para geração de perfil foi utilizado aprendizado de máquina e algoritmos de *clustering* para detectar padrões de navegação semelhantes entre os usuários. Para isso foi usado o algoritmo PAM (*Partitioning Around Medoids*), que se assemelha com o *K-means* e SPADE (*Sequential Pattern Discovery using Equivalence classes*) usado para extrair a sequência de cliques mais comuns no cluster de dados. Acredita-se que com as características de log das sessões é possível analisar o uso do site a partir de outro ponto de vista principalmente para detecção de problemas. Entende-se que esse trabalho possui

uma grande relevância, pois apresenta a utilização de técnicas de *clustering* e aprendizado de máquina e sua eficácia em relação à construção de perfis.

### 3.2 Análise dos trabalhos relacionados

Nesta etapa foi realizada uma análise sobre os trabalhos relacionados onde foi avaliado o grau de relevância presente em cada trabalho. O principal objetivo da análise foi destacar a relevância e o impacto gerado pelos trabalhos relacionados sobre o trabalho que está sendo proposto, destacando suas eventuais contribuições. Na [Tabela 1](#) são apresentados os trabalhos relacionados juntamente com as técnicas utilizadas e o seu grau de relevância.

Tabela 1 – Avaliação dos trabalhos relacionados.

Trabalho	Técnicas Utilizadas	Relevância
<a href="#">Vigo et al. (2008)</a>	<i>Assistive Technologies Detector</i> <i>Acess Device features Detector</i>	Baixa
<a href="#">Ferretti et al. (2014)</a>	Algoritmo <i>Q-learning</i> PIQLE <i>Platform for Implementing</i> <i>Q-Learning Experiments</i>	Alta
<a href="#">Arbelaitz et al. (2014)</a>	PAM( <i>Partitioning Around Me-</i> <i>doids</i> )	Média
<a href="#">Bigham et al. (2007)</a>	<i>Javascript</i>	Média
<a href="#">Santana e Baranauskas (2015)</a>	<i>JavaScript</i>	Média

Fonte: o autor

O primeiro trabalho proposto por [Vigo et al. \(2008\)](#), tem como objetivo coletar dados sobre as tecnologias de apoio utilizadas pelos usuários com o *Assistive Technologies Detector* e informações sobre o *hardware* e *software* com o *Acess Device features Detector*. Foi considerado com baixa relevância por coletar somente informações a respeito da tecnologia assistiva, *hardware* e *software*, diferente do trabalho que está sendo proposto que tem como objetivo coletar informações de eventos realizados pelos usuários na página da *web*, no entanto a sua contribuição está relacionada aos seus resultados que sugerem que somente a coleta de informações sobre o *hardware*, *software* e tecnologias de apoio não são o suficiente para atribuí-lo a um grupo específico de pessoas com deficiência.

O segundo trabalho proposto por [Ferretti et al. \(2014\)](#), tem como objetivo adaptar automaticamente elementos textuais da página de acordo com o perfil do usuário. Para isso é utilizado o algoritmo *Q-learning* para compreender a interação do usuário, juntamente com a *Platform for Implementing Q-Learning Experiments* que reproduzia a navegação do usuário com o conteúdo da *web*. O grau de relevância atribuído foi como

alta relevância, pois utiliza uma técnica de aprendizado de máquina, assim como o trabalho que está sendo proposto, indicando a viabilidade aprendizado de máquina para o monitoramento de interações do usuário com a *web*.

O terceiro trabalho proposto por [Arbelaitz et al. \(2014\)](#), apresenta um sistema para auxiliar na navegação do site *Discapnet* foi usado um algoritmo de clustering PAM (*Partitioning Around Medoids*). Esse trabalho foi classificado com relevância média para o trabalho proposto, por utilizar técnicas de aprendizagem de máquina para identificar padrões de uso por parte dos usuários para com um site, contendo de forma básica o mesmo objetivo do trabalho que está sendo proposto.

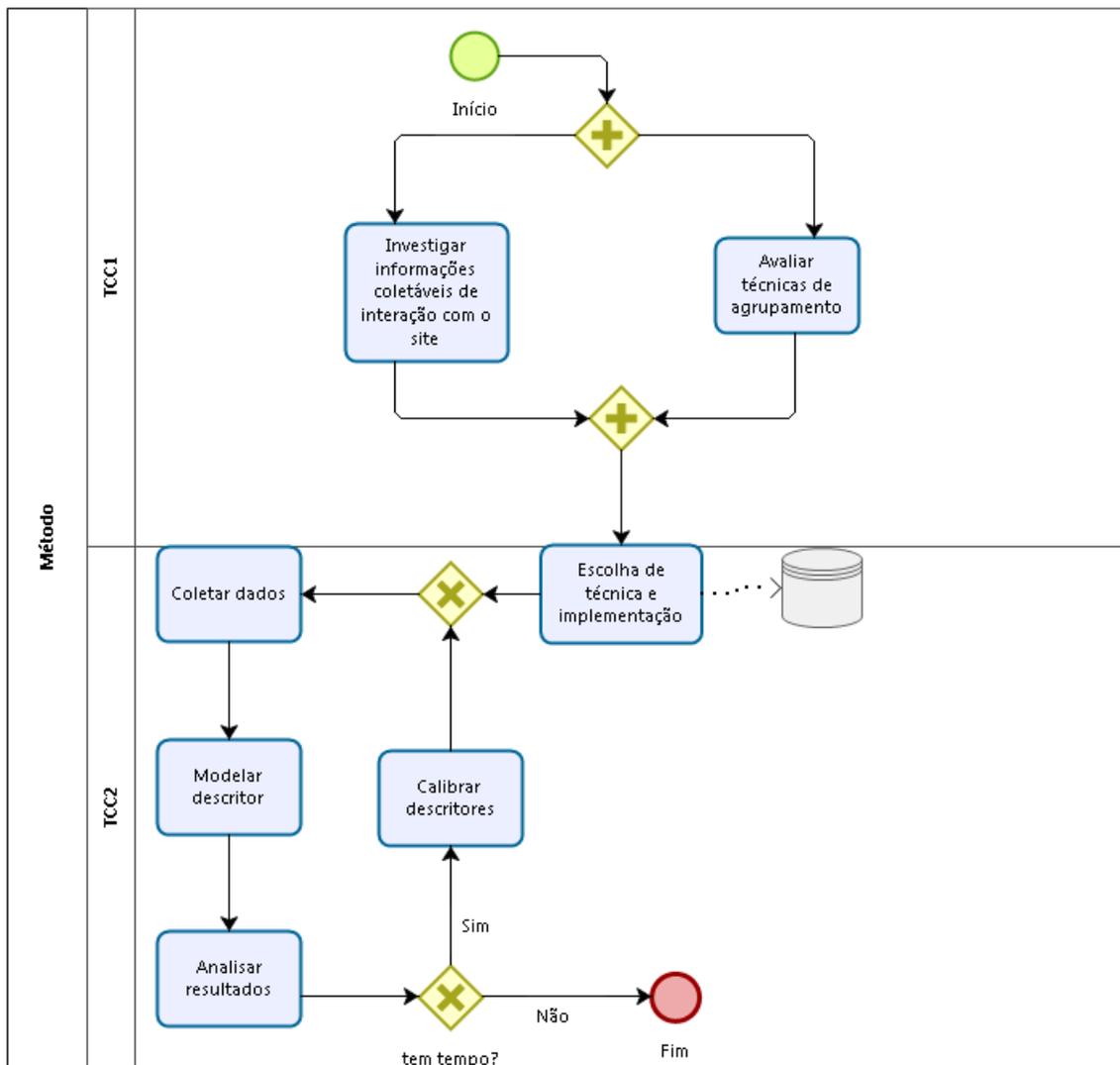
O quarto trabalho proposto por [Bigham et al. \(2007\)](#), apresenta um estudo para avaliar a acessibilidade a partir da perspectiva do usuário analisando seu modo de interação, para gravar as atividades realizadas pelos usuário foi usado a tecnologia *JavaScript* que coletou informações sobre os cliques, botões e elementos entre outras informações sobre o usuário. De acordo com a análise realizada esse trabalho foi considerado com uma relevância média, porque comprova a eficácia da tecnologia *JavaScript* para a coleta de informações dos usuários em páginas da *web*, sendo exatamente esta tecnologia que será utilizada no trabalho que está sendo proposto, para coletar informações do usuário.

O quinto trabalho proposto por [Santana e Baranauskas \(2015\)](#), trata da ferramenta WELFIT implementada em *JavaScript*, usada para a avaliação do fluxo de eventos realizados por usuários com deficiência e usuários sem deficiência, e seus eventuais padrões de uso possibilitando uma distinção entre os dois grupos de usuários. Como o trabalho descrito anteriormente, é usada a tecnologia *JavaScript* para capturar eventos comprovando novamente a eficácia da tecnologia para esse contexto, sendo considerado com relevância média para o presente trabalho.

## 4 FERRAMENTAS E MÉTODOS

Para o desenvolvimento trabalho foi utilizado um método com base em [Wazlawick \(2014\)](#). Na [Figura 7](#) é apresentado o processo contendo todas as etapas que foram realizadas durante o desenvolvimento do trabalho.

Figura 7 – Processo para realização do trabalho



Fonte: a autora

### 4.1 Investigação das informações a serem coletadas

Foi realizada uma investigação para identificar quais as informações seriam relevantes para caracterizar o perfil dos usuários durante a sua interação com páginas da *web*. Para isso foram utilizadas informações encontradas na literatura sobre ao modo de interação de pessoas com deficiência visual, apresentadas na seção [2.1.1](#) e informações sobre o

modo de interação de pessoas com deficiência motora nos membros superiores apresentadas na seção 2.1.2. Também foi realizada uma análise dos trabalhos relacionados na seção 3.2 a fim de identificar quais técnicas estavam sendo usadas para auxiliar na extração de perfis de pessoas com deficiência e sua relevância para o trabalho que está sendo proposto. A partir das informações apresentadas pela literatura e análise dos trabalhos relacionados foi realizada comparação sobre o modo de interação de pessoas com deficiência e pessoas sem deficiência com páginas da *web*.

Foram encontradas diferenças relacionadas de interação, como: o uso da tecla *tab* e *tab + shift*; setas direcionais para navegar pelo conteúdo da página; ausência de cliques com o *mouse* sobre os elementos da página; movimentos rítmicos com o *mouse* variando em uma frequência específica; período de tempo mais acentuado para realização de tarefas que normalmente seriam realizadas em menos tempo por pessoas sem deficiência; acionar atalhos da página (ir para conteúdo principal); movimentos lineares do *mouse* (movimentos perfeitamente retos horizontais, verticais ou diagonais).

## 4.2 Coleta de dados de interações

No decorrer do processo de coleta de dados, foi realizado um experimento aplicando uma técnica de triagem descrita na seção 4.6. Essa técnica tem como característica atribuir limitações para pessoas sem deficiência, com o objetivo de reproduzir um comportamento aproximado ao de uma pessoa com deficiência, por exemplo, permitindo somente a interação do participante via leitor de tela, que por sua vez teria como objetivo aproximar o comportamento de uma pessoa cega.

Durante a aplicação dos testes e dos experimentos foi estabelecido para cada participante uma sessão de uso, onde foram coletadas e armazenadas as informações sobre o modo de interação de cada participante no site. De acordo com [Cybis, Betiol e Faust \(2007\)](#) os dados de *log* coletados são recursos interessantes, pois são vestígios do uso real do site, proporcionando aos usuários em um ambiente natural, buscando realizar suas necessidades de forma autêntica. No início de cada sessão foi apresentada uma mensagem, informando que todos os dados fornecidos e eventos realizados no site seriam armazenados. Posteriormente, os participantes deveriam realizar uma ou mais tarefas no site de forma a interagir com uma variada quantidade de elementos de uma página *web*, com objetivos relacionados ao modo de interação de cada perfil, permitindo o registro de peculiaridades ocorridas durante cada sessão, incluindo informações como o tempo transcorrido para realizar a tarefa, eventos de *mouse* e teclado e percurso nos elementos nas subpáginas.

Para auxiliar o processo de coleta de informações, foi desenvolvido um site com diversos *widgets* <sup>1</sup> para identificar possivelmente o comportamento de diferentes perfis de usuário que interagem com os elementos da página. Foram estabelecidos ao todo 3 per-

---

<sup>1</sup> Componente de uma interface gráfica do usuário (GUI), o que inclui janelas, botões, menus, ícones, barras de rolagem

fis de usuários para cada participante durante a aplicação da técnica de triagem, sendo os perfis dividido entre pessoas cegas, pessoas com deficiência motora nos membros superiores e pessoas sem deficiência. O objetivo dessa divisão é identificar as diferenças de comportamento de cada participante com o site e posteriormente realizar a classificação a partir do seu modo de interação. Estabelecendo uma base de dados sobre cada participante para a aplicação do algoritmo de aprendizagem de máquina *K-Means*.

### 4.3 Modelagem de descritores

Na modelagem dos descritores foram identificados seis atributos considerados relevantes para caracterização de cada perfil de usuário. Esses atributos são um conjunto de informações coletadas ao longo da sessão, que foram analisadas e consideradas relevantes para indicar qual é o perfil de interação de cada usuário.

O primeiro atributo é o percurso realizado pelo *mouse* ao longo da sessão. Essa característica foi considerada relevante, pois dependendo do tipo de deficiência do usuário o *mouse* nem ao menos é utilizado, como ocorre com usuários cegos que fazem uso de leitores de tela para interagir com computador, sendo que a navegação do leitor de tela ocorre somente via teclado. No segundo atributo analisado é o tempo de sessão do usuário, devido ao modo de interação diferenciado entre pessoas sem deficiência, e pessoas com deficiência motora e pessoas cegas, que fazem uso de tecnologias assistivas para interagir com o computador e em virtude disso acabam levando um tempo maior para realização das tarefas. O terceiro atributo tem como finalidade identificar se a navegação do usuário pela página ocorreu de modo linear. Essa característica foi levada em consideração por causa do modo de navegação de pessoas com deficiência motora, que navegam pela página através de um dispositivo apontador (*mouse*) com movimentos perfeitamente lineares.

Já o quarto atributo tem como objetivo identificar se o usuário navegou de modo sequencial pela página, pois pessoas cegas utilizam leitores de tela para navegar pelas páginas, que são lidas de modo sequencial pelo leitor. Logo se o usuário realizou esse tipo de navegação existe a possibilidade que esse usuário venha a ser uma pessoa cega. O quinto atributo consiste em verificar o percentual de uso das teclas *delete* e *backspace*, que foi considerado relevante devido à dificuldade de digitação de algumas pessoas com deficiência, como ocorre com pessoas com deficiência motora com tremor nos membros superiores que possuem dificuldade em realizar movimentos precisos ocasionando o uso constante dessas teclas por causa de erros de digitação. O sexto atributo tem como finalidade identificar se os movimentos realizados pelo usuário com o *mouse* caracterizam algum tipo de tremor. Para identificar tal característica, foi realizada uma análise com o intuito de compreender quais frequências de movimentos eram considerados como tremor. Ao final desta análise foram constatados que tremores cuja a banda de 1 a 6 hz possuísem energia superior a 3000 u.m. seriam considerados tremores (HESS; PULLMAN, 2012). Para este fim foi utilizada a transformada rápida de *Fourier* que é uma implementação

otimizada da transformada de *Fourier* representada pela equação 4.1, usada para extrair as frequências dos movimentos realizados pelo usuário com o *mouse* durante a sessão.

$$F(w) = \int_{-\infty}^{\infty} f(x)e^{-iwt} dt \quad (4.1)$$

#### 4.4 Escolha e implementação da técnica de agrupamento

Na etapa de escolha e implementação da técnica de agrupamento foi determinada a técnica a ser utilizada para o processo de agrupamento das informações coletadas. Foi então realizada uma análise apresentada na seção 2.4 sobre as principais vantagens e desvantagens de aplicação de cada técnica.

Durante a escolha da técnica foi levada em consideração a complexidade das informações coletadas e quais seriam os seus critérios de agrupamento, como também, qual tipo de técnica seria mais adequada para o conjunto de informações. Com base na complexidade das informações e os seus critérios de agrupamento, foi escolhido o algoritmo de *clustering K-Means*. Esse algoritmo trata de uma técnica de aprendizagem de máquina não supervisionada, onde os critérios de agrupamento são estabelecidos pelo usuário, determinando a quantidade de grupos em que as informações coletadas devem ser classificadas. Também, trata-se de uma técnicas de simples implementação. Na seção 2.3.2 é descrito todo processo realizado pelo algoritmo.

Para o processo de implementação do algoritmo é utilizada a ferramenta *Waikato Environment for Knowledge Analysis*<sup>2</sup> (WEKA) que abrange algoritmos para preparação de dados de aprendizagem de máquina. O Weka foi desenvolvido na Universidade de *Waikato* na Nova Zelândia, escrito em *Java*, de código aberto disponível na *web*, possui uma interface gráfica amigável e seus algoritmos fornecem relatórios com dados analíticos e estatísticos do domínio minerado (SILVA, 2004).

#### 4.5 Comparação e análise de resultados

Durante esta etapa são verificados os resultados obtidos a partir da técnica de agrupamento de dados. Para isso foram analisados os resultados obtidos após a aplicação das duas técnicas de triagem, certificando-se de que o agrupamento das informações estaria de acordo com o gabarito com o perfil estabelecido a cada participante.

Para garantir que as informações estavam sendo agrupadas de maneira correta foram aplicados dois experimentos. O primeiro experimento está descrito na seção 4.6.1, e o segundo na seção 4.6.2. Para cada uma das informações obtidas, foi realizada uma comparação entre o perfil atribuído ao participante durante o experimento, e o *cluster* ao qual essa informação foi agrupada. Caso a informação não estivesse agrupada corretamente, então era realizada uma investigação para identificar o que estava causando essa inconsistência no agrupamento das informações.

<sup>2</sup> Disponível em <<https://www.cs.waikato.ac.nz/ml/weka/>>

## 4.6 Experimentos

Nesta seção são descritos os protocolos e técnicas usadas na aplicação do primeiro e segundo experimento. Durante a aplicação dos experimentos foi usada uma técnica de triagem (*Screening*) para geração de dados artificiais. Essa técnica tem como finalidade permitir que uma pessoa sem deficiência reproduza o comportamento de uma pessoa com deficiência, incluindo o uso de tecnologias assistivas, e até mesmo impor limitações de uso aos participantes durante a realização de tarefas.

Primeiramente foi feita uma apresentação aos participantes explicando do que se tratava a técnica e qual sua finalidade, juntamente com quais os tipos de deficiência os participantes deveriam reproduzir. Foi então solicitado aos participantes que realizassem quatro tarefas em um site fictício de uma cafeteria, sendo elas: comprar chocolate quente, descobrir o ano de criação da empresa, enviar uma mensagem elogiando um café e ver o café que estava em promoção quinta-feira. Para cada uma das tarefas realizadas no site os participantes deveriam reproduzir um dos quatro perfis de usuários, sendo eles: pessoas sem deficiência, pessoas cegas e pessoas com deficiência motora, que foi dividido entre pessoas com deficiência motora que navegam usando dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página e pessoas com deficiência motora com tremores nos membros superiores.

Posteriormente, foram indicadas quais as limitações para cada perfil de usuário, e quais seriam as tarefas por cada participante. Entre as limitações impostas ao participante, estavam desligar o monitor do computador e navegar pelo site somente utilizando teclado e o leitor de tela com a finalidade de representar usuários cegos. Já para os usuário que possuem deficiência motora nos membros superiores, foi usado um recurso de acessibilidade do sistema operacional *Windows* chamado “teclas do *mouse*”, que permite o movimento e ações do *mouse* através do teclado numérico, realizando movimentos lineares com o *mouse*, também foi realizada uma investigação sobre o modo de interação de pessoas com tremores e exibido um vídeo com uma pessoa tremendo ao fazer tarefas diárias e solicitado aos participantes que reproduzisse o tremor o mais semelhante possível com o que foi apresentado no vídeo durante sua interação com o site.

Para verificar o desempenho do algoritmo em classificar os elementos foi utilizado o método da matriz de confusão conceituado por Prati, Batista e Monard (2008). A matriz de confusão permite a avaliação da capacidade do algoritmo em relacionar classes reais e classes previstas. A classe real é a classe ao qual o elemento realmente pertence, representada por  $(\mathbf{p}, \mathbf{n})$  e a classe prevista é a classe à qual o algoritmo “acredita” que o elemento pertença, representada por  $(\mathbf{y}, \mathbf{n})$ . Na Tabela 2 é mostrado um exemplo de como é composta a matriz de confusão.

**True Positives** - Verdadeiro Positivo; **False Positives** - Falso Positivo;

**False Negatives** - Falso Negativo; **True Negatives** - Verdadeiro Negativo;

Tabela 2 – Matriz de confusão

		Classe Real	
		<b>p</b>	<b>n</b>
Classe Prevista	<b>y</b>	<i>True Positives</i>	<i>False Positives</i>
	<b>n</b>	<i>False Negatives</i>	<i>True Negatives</i>

A acurácia determina a taxa de predições corretas sem levar em consideração o que é positivo ou negativo.

$$\text{Acurácia} = \frac{TP + TN}{P + N}$$

A sensibilidade, por sua vez, mede a proporção de positivos que foram identificados como positivos. Analogamente a especificidade verifica a proporção dos negativos que estão corretamente identificados como negativos.

$$\text{Sensibilidade} = \frac{TP}{TP + FN}$$

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

A precisão positiva fornece a proporção dos verdadeiros positivos em relação a todas as predições positivas, e a precisão negativa fornece a proporção dos verdadeiros negativos em relação a todas as predições negativas.

$$\text{Precisão positiva} = \frac{TP}{TP + FP}$$

$$\text{Precisão negativa} = \frac{TN}{TN + FN}$$

Na seção subsequente são apresentados os objetivos, número de participantes e o modo de distribuição dos dados do primeiro e segundo experimento. Ficando a seção 4.6.1 relacionada ao primeiro experimento a seção 4.6.2 ao segundo experimento.

#### 4.6.1 Experimento 1

Nesta etapa foram gerados um conjunto de dados artificiais com o intuito de identificar de modo preliminar a eficácia do método de agrupamento para a resolução do problema que está sendo proposto, através da análise dos resultados obtidos com o agrupamento das informações. Durante o processo de criação dos dados artificiais houve somente um participante, a autora do trabalho, que ficou responsável por reproduzir de modo aproximado o comportamento de pessoas com deficiências e sem deficiência. Esse único participante já tinha conhecimentos prévios sobre o uso de tecnologias assistivas e o modo de interação de pessoas com deficiência. Todos os dados coletados ao longo do teste preliminar foram rotulados de acordo com o perfil de usuário que estava sendo reproduzido.

Cada uma das tarefas foram feitas com todos os perfis de usuários, incluindo o perfil de usuários sem deficiência. Ao todo foram obtidos 31 elementos que continham dados de *log*, que mais tarde foram divididos em 4 grupos, sendo 5 elementos tratavam-se de pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página, 6 elementos de pessoas com tremores nos membros superiores, e o restante dos elementos foi dividido entre pessoas cegas e pessoas sem deficiência ficando 10 elementos para cada grupo.

#### 4.6.2 Experimento 2

Após a realização do primeiro experimento foi identificada a necessidade de criação de mais dados artificiais com maior número de participantes possibilitando uma variabilidade maior entre os dados gerados. Ao todo o experimento contou com a participação de 10 pessoas, que ficaram encarregadas de reproduzir o comportamento de pessoas com deficiência e sem deficiência entre os participantes haviam pessoas com conhecimentos sobre tecnologias assistivas e pessoas sem conhecimento algum sobre o assunto.

Por fim, foram obtidos um total de 40 elementos contendo dados de *log* todos rotulados, divididos em 4 grupos: pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página, pessoas com tremores nos membros superiores, pessoas com cegueira e pessoas sem deficiência. O grupo de pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página, obteve 10 elementos, pessoas com tremores 10 elementos, pessoas cegas 10 elementos e pessoas sem deficiência 10 elementos.



## 5 RESULTADOS

Nesta são descritos os resultados obtidos com o trabalho. Na na seção 5.1 são apresentadas quais as tecnologias foram utilizadas para o desenvolvimento do sistema de coleta de informações, juntamente com seus requisitos e arquitetura. A seção 5.2 por sua vez descreve os resultados obtidos com o primeiro experimento e a seção 5.3 os resultados do segundo experimento.

### 5.1 Sistema de coleta de informações

Para armazenamento das informações obtidas em cada sessão do usuário foi desenvolvido um sistema <sup>1</sup> responsável por coletar as informações e armazená-las em um banco de dados. Durante o processo de desenvolvimento foram estabelecidos os requisitos funcionais ao longo de diversas reuniões com dois especialistas. Na sequência são apresentados os requisitos funcionais na Tabela 3, onde são descritas quais as funcionalidades que o sistema deve realizar (SOMMERVILLE et al., 2003), que foram documentadas em forma de lista.

Foram utilizadas para implementação do sistema as linguagens de programação *JavaScript* e a biblioteca *jQuery* que realizam a coleta dos dados do lado cliente através do navegador, juntamente com linguagem de programação PHP que utilizou o Framework *Laravel* <sup>2</sup> para armazenar os dados coletados no lado servidor em um banco de dados MySQL. Ambas as linguagens de programação são baseadas na arquitetura cliente e servidor. Na Figura 8 é apresentado o diagrama de pacote as relações entre os módulos do sistema e suas classes. Também foi criado um diagrama de classe do sistema <sup>3</sup>. Na Figura 9 é apresentado o modelo conceitual do banco de dados utilizado pelo sistema, informando os dados e os relacionamentos entre os dados que serão armazenados (HEUSER, 2009).

---

<sup>1</sup> Disponível em <<https://github.com/AnneEscarrone/SistemaDeColetaDeDados>>

<sup>2</sup> Disponível em <<https://laravel.com/>>

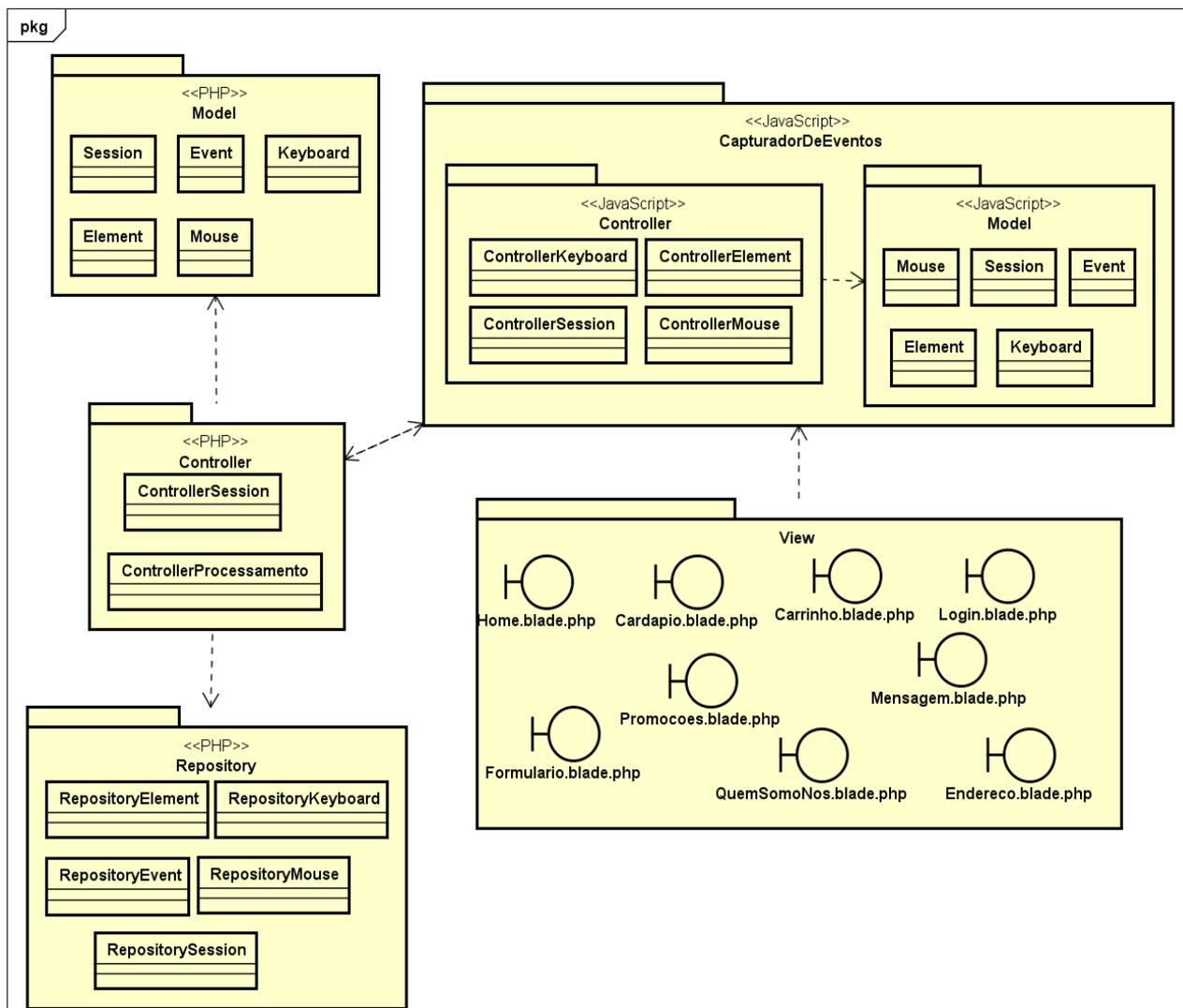
<sup>3</sup> Disponível em <<https://drive.google.com/open?id=1Gxdl-ARh52fxFKUcgcFlix4m1QV-mQDy>>

Tabela 3 – Lista de requisitos funcionais do sistema.

<b>Identificador</b>	<b>Requisitos Funcionais</b>
RF01	O sistema deve apresentar uma mensagem ao usuário informando que as atividades realizadas na página serão armazenadas.
RF02	O sistema deve permitir ao usuário uma opção para encerrar sua sessão.
RF03	O sistema deve armazenar o modo de navegação do usuário, sendo pelo cursor ou pelo teclado informando quais teclas ou botões foram utilizadas.
RF04	O sistema deve armazenar o id, url que o elemento pertence, tipo de elemento e texto caso seja possível e posição dos elementos na tela, sendo eles selecionados através do <i>mouse</i> ou teclado.
RF05	O sistema deve armazenar o número e código da tecla pressionada no teclado
RF06	O sistema deve armazenar o tempo duração de cada sessão e o ip do usuário
RF07	O sistema deve armazenar a posição dos cliques ocorridos na tela.
RF08	O sistema deve armazenar a posição dos cliques e do botão do <i>mouse</i> durante a sessão.
RF09	O sistema deve armazenar os movimentos do cursor pela pagina.

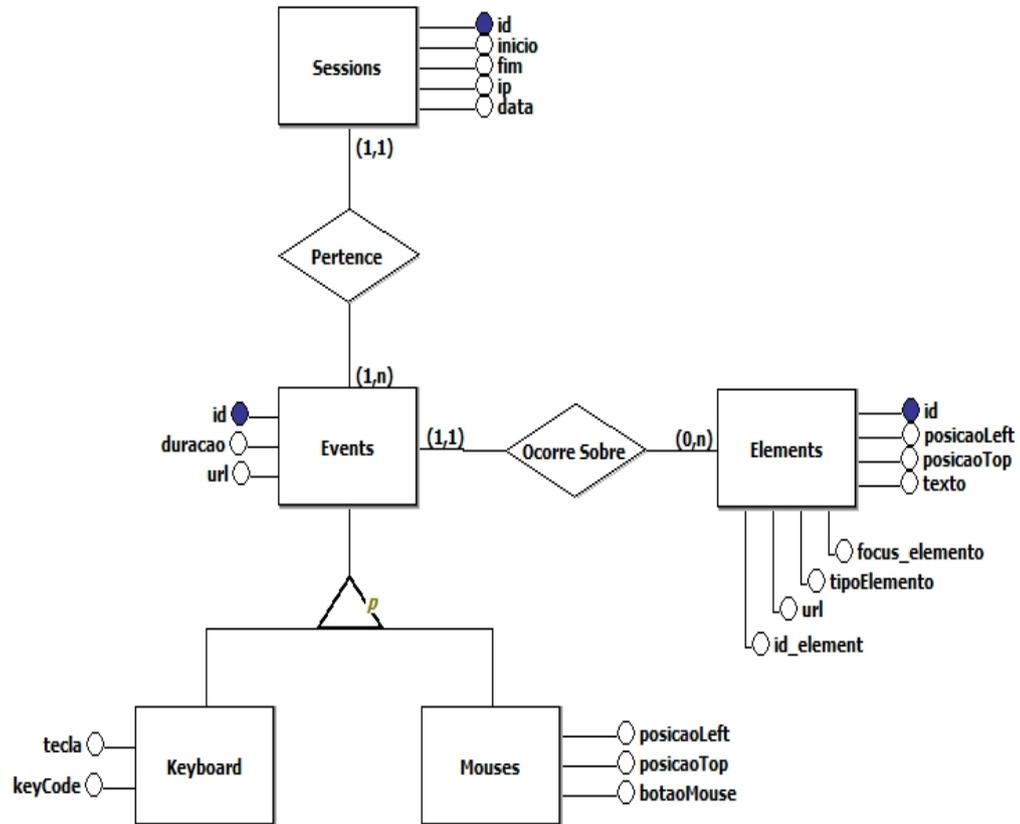
Fonte: a autora

Figura 8 – Diagrama de pacotes com suas classes



Fonte: a autora

Figura 9 – Modelo conceitual do banco de dados



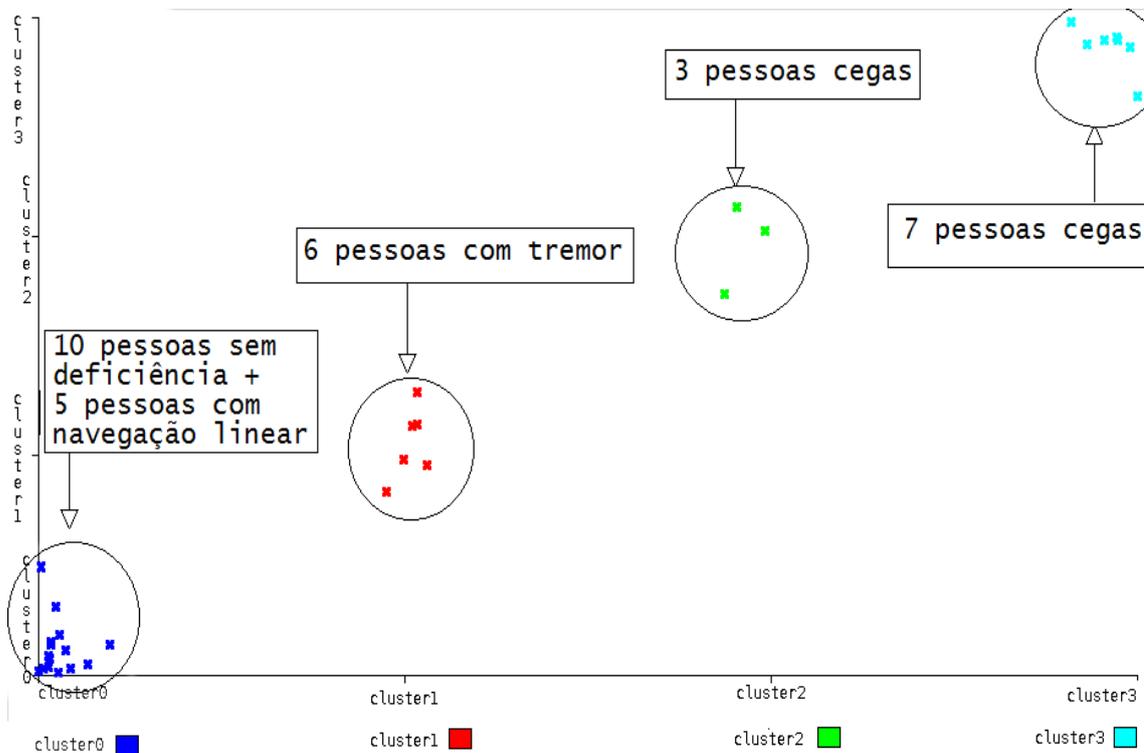
Fonte: a autora

## 5.2 Experimento 1

Durante o primeiro experimento foram coletados 31 elementos que são compostos por dados de *log* com interação dos usuários. Para cada elemento foi atribuído um perfil de usuário, somando um total de quatro perfis, sendo eles: pessoas com deficiência motora que navegam de modo linear, pessoas com deficiência motora que possuem tremor nos membros superiores, pessoas cegas e pessoas sem deficiência. Devido a diferenças encontradas entre os perfis de usuários, foi estabelecido para cada perfil um *cluster* e utilizado o algoritmo *K-Means* para o agrupamento dos elementos.

Depois de realizar o agrupamento dos elementos foi feita uma análise sobre a distribuição dos elementos em cada *cluster*. O primeiro *cluster*, ficou com um total de 15 elementos, sendo que 10 dos elementos contidos estavam rotulados como pessoas sem deficiência e os outros 5 estavam rotulados como pessoas com deficiência motora que realizam a navegação de modo linear. No segundo *cluster*, foram obtidos 6 elementos rotulados como pessoas com deficiência motora que possuem tremores nos membros superiores. O terceiro *cluster* possuía 3 elementos rotulados como pessoas cegas. O quarto um total de 7 elementos rotulados como pessoas cegas. Na Figura 10 são apresentados de modo gráfico os *clusters* e seus respectivos dados.

Figura 10 – Representação gráfica dos *clusters* sem modificações



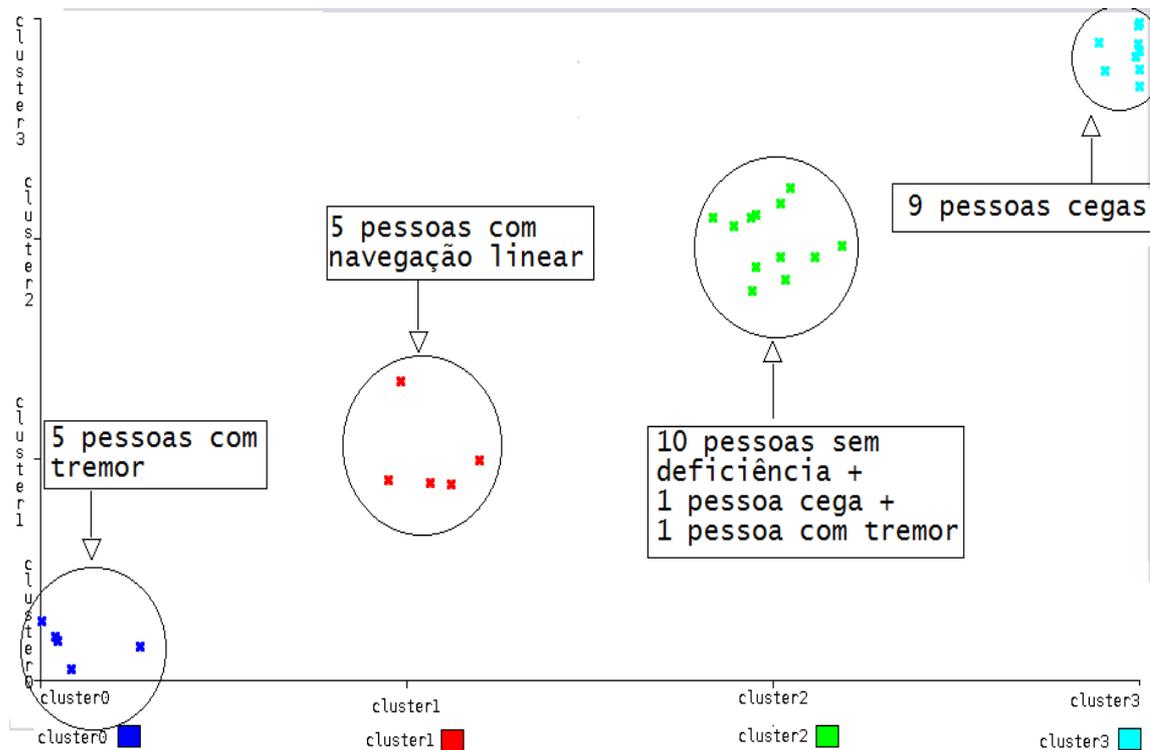
Com base nos resultados obtidos, com a análise dos agrupamentos dos dados foi possível identificar algumas falhas, como: o perfil de pessoas sem deficiência estar sendo agrupado juntamente com o perfil de pessoas que navegam de modo linear, e o perfil de pessoas cegas estar dividido em dois *clusters* diferentes. Devido a tal resultado, foi feita uma investigação sobre o que poderia estar causando essa falha no agrupamento dos elementos. Então, foram analisados cada um dos elementos que não foram agrupados de modo correto e averiguada a existência de alguma similaridade entre eles. Por fim, foram identificadas similaridades em relação aos atributos de navegação linear e sequencial. A similaridade encontrada estava relacionada ao percentual utilizado para classificar o modo de navegação do usuário, que estava muito baixo gerando uma classificação incorreta em alguns elementos, causando uma desorganização no agrupamento das informações. Foram então aplicadas algumas alterações no percentual de cada um desses atributos. O percentual de incerteza do atributo de navegação linear era maior que 40% e menor 80% e foi modificado para maior que 70% e menor que 80%, e o grau que determinava que não era navegação linear, era menor ou igual a 40% e mudou para menor ou igual a 70%. Já o percentual do atributo de navegação sequencial relacionado ao grau de incerteza, era maior que 20% e menor de 50 % foi modificado para maior que 40% e menor que 50% e o grau que determinava se era ou não navegação sequencial, era de 20 % e passou para 40%.

Assim que as alterações foram aplicadas, foi realizado novamente o agrupamento dos elementos e criadas quatro matrizes de confusão, para analisar o desempenho do algoritmo. Na [Tabela 4](#) é apresentada a primeira matriz de confusão contendo todos os elementos distribuídos ao longo de cada *cluster*, obtendo os seguintes resultados: o primeiro obteve 6 elementos rotulados de acordo com perfil de pessoas com deficiência motora com tremores nos membros superiores; o segundo *cluster* ficou com 5 elementos, todos com o rótulo do perfil pessoas com deficiência motora que realizam a navegação de modo linear; o terceiro *cluster* obteve 12 elementos, onde 10 elementos estavam rotulados como o perfil de pessoas sem deficiência e 1 elemento com o perfil de pessoas cegas e 1 elemento com o perfil de pessoas com deficiência motora com tremor leve; o quarto perfil ficou com 9 elementos, todos rotulados com o perfil de pessoas cegas. A [Figura 11](#) mostra de modo gráfico a distribuição dos elementos em seus respectivos *clusters*.

Tabela 4 – Matriz de confusão contendo todos os *clusters*

		Classe Real				
		T	DM	SD	DV	Total do <i>cluster</i>
Classe Prevista	T	<b>6</b>	0	0	0	6
	DM	0	<b>5</b>	0	0	5
	SD	0	1	<b>10</b>	1	12
	DV	0	0	0	<b>9</b>	9

**DM-** Deficiência Motora **DV-** Deficiência Visual **T-**Tremor **SD-** Sem Deficiência

Figura 11 – Representação gráfica dos *clusters*

Fonte: a autora

A segunda matriz de confusão apresentada na [Tabela 5](#), tem como finalidade mostrar os elementos relacionados ao *cluster* do perfil de pessoas cegas, assim como, também o desempenho do algoritmo em classificar esses elementos. Foram então analisados os percentuais do algoritmo em relação a sua acurácia, sensibilidade, especificidade e precisão. O algoritmo demonstrou eficiência em classificar de modo correto os elementos, atingindo um percentual de 96,77% de acurácia na classificação dos elementos. Ao passo que a sua sensibilidade foi de 90 %, em determinar quais elementos eram verdadeiros positivos. Foram obtidos somente 90% de sensibilidade e 96,77% de acurácia demonstrando uma classificação incorreta por parte do algoritmo, porque durante a coleta de dados a pesquisadora ao reproduzir o perfil de pessoas cegas, acabou movimentando o *mouse* e o movendo de modo linear, atingindo um percentual de navegação linear considerado incerto. No entanto, isso acabou causando uma diferença em relação aos outros perfis de pessoas cegas, fazendo com que o algoritmo classificasse esse elemento, como uma pessoa sem deficiência, devido a similaridade dos dados de *log* desse elemento.

Por outro lado, o algoritmo obteve 100% de especificidade em identificar quais elementos não estavam rotulados como pessoas cegas, e não foram classificados como pessoas cegas, demonstrando a qualidade do algoritmo em classificar quais elementos possuíam o rótulo do perfil de pessoas cegas. A precisão do algoritmo obteve 100% de acerto em

Tabela 5 – *Cluster* com deficiência visual

		Classe Real	
		DV	N-DV
Classe Prevista	DV	9	0
	N-DV	1	21

**DV-** Deficiência Visual **N-DV-** Não Deficiência Visual

identificar os verdadeiros positivos, evidenciando a capacidade do algoritmo em identificar quais elementos eram pessoas cegas, e realmente foram classificadas como pessoas cegas. Porém, o algoritmo não apresentou 100% de precisão em classificar os negativos verdadeiros, obtendo um total de 95,46% de acerto, em determinar quais elementos não eram pessoas cegas e realmente não foram classificados como pessoas cegas.

Tabela 6 – *Cluster* com tremores

		Classe Real	
		T	N-T
Classe Prevista	T	5	0
	N-T	1	25

**T-** Tremor **N-T-** Não Tremor

A terceira matriz de confusão exibida na [Tabela 6](#), contém os elementos relacionados ao *cluster* de pessoas com tremores. Assim como na segunda matriz de confusão foram verificados os percentuais do algoritmo em relação a classificação dos elementos. Foi obtido 96,77% de acurácia em classificar de forma correta os elementos, isto ocorreu porque, alguns elementos rotulados como pessoas com tremores nos membros superiores, possuíam dois tipos de intensidade de tremor. Essa dualidade entre os dados foi criada durante o processo de coleta de dados, em que a pesquisadora acabou reproduzindo dois tipos de tremores, de modo accidental, gerando tremores mais intensos e tremores mais leves, levando o algoritmo a classificar tremores considerados mais leves, como o perfil de pessoas sem deficiência.

A sensibilidade atingida pelo algoritmo foi de 83,34% na classificação de verdadeiros positivos, devido a existência de dois tipos de tremores. Apesar disso, foi obtido 100% de especificidade do algoritmo em determinar que elementos não eram rotulados como tremor, não foram classificados como tremor. Quanto ao grau de precisão do algoritmo em reconhecer quais elementos eram rotulados como tremor, e foram realmente classificados como tremores foi de 100% de acerto. Já em relação a precisão dos negativos verdadeiros foram obtidos 95,46% de acerto, esse valor percentual foi atingido devido a existência de dois tipos de tremores entre o conjunto de elementos.

A quarta matriz de confusão possui os elementos do *cluster* do perfil de pessoas sem deficiência, e é apresentada na [Tabela 7](#), como ocorreu, com as demais matrizes de confusão utilizadas ao longo desse experimento, também será verificado o percentual de

acurácia, sensibilidade, especificidade e precisão do algoritmo em relação aos elementos do *cluster* de pessoas sem deficiência. Foi então verificado o percentual de acurácia do algoritmo, que foi somente de 93,54% de acerto em classificar de modo correto os elementos, devido a classificação incorreta de dois elementos, um deles rotulado como o perfil de pessoas com tremores e outro rotulado com o perfil de pessoas cegas. Enquanto a sensibilidade do algoritmo em identificar quais elementos rotulados como pessoas sem deficiência, realmente foram classificadas como pessoas sem deficiência foi de 100%.

Em seguida foi verificada a especificidade do algoritmo em identificar quais elementos realmente não eram pessoas sem deficiência, que alcançou 90,47% de acerto. Logo depois foi averiguado o a precisão do algoritmo em identificar quais os elementos classificados como pessoas sem deficiência e se realmente tratava-se de elementos rotulados como pessoas sem deficiência, foram adquiridos 83,34% de precisão. Já a precisão do algoritmo em distinguir quais elementos pertenciam ao rótulo de pessoas com deficiência, e se de fato esses elementos não foram classificados como pessoas com deficiência. O *cluster* de pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página apresentou um percentual de 100% em relação à acurácia, sensibilidade, especificidade e precisão de verdadeiros positivos e falsos negativos.

Tabela 7 – *Cluster* sem deficiência

		Classe Real	
		SD	N-SD
Classe Prevista	SD	10	2
	N-SD	0	19

**SD-** Sem Deficiência **N-SD-** Não Sem Deficiência

Mais adiante foi realizada a soma da média aritmética relacionada a acurácia e sensibilidade de todos os *clusters* e obtido 96,77 % de acurácia em classificar de forma correta os elementos, e 93,33% de sensibilidade em identificar de modo correto os elementos de acordo com os seu respectivo perfil, evidenciando então a eficácia do algoritmo em classificar de modo correto a maioria dos elementos.

### 5.3 Experimento 2

No segundo experimento foram coletados um conjunto de 40 elementos que contêm dados de *log* de interações dos usuários durante as sessões. Esses elementos estavam rotulados de acordo com 4 perfis de usuários: pessoas sem deficiência, pessoas cegas, pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página e pessoas com deficiência motora que possuem tremores nos membros superiores. Assim como no primeiro experimento foi utilizado um algoritmo de agrupamento *K-Means* e determinado que os elementos seriam agrupados

em quatro *clusters*. Contudo, foram realizados testes com o algoritmo EM e o algoritmo CANOPY, a fim de verificar se o *K-Means* de fato era o ideal para o agrupamento desse tipo de informação. Ao final dos testes realizados com os dois algoritmos o *K-Means* obteve o resultado mais satisfatório em relação ao agrupamento dos elementos.

Em seguida os elementos foram rotulados de acordo com o seu perfil e realizado o agrupamento. Durante o processo de agrupamento dos elementos foi encontrado um erro no código, que estava relacionado ao método que determinava se o usuário estava realizando a navegação de modo linear. Logo após o erro ser corrigido, todos os elementos rotulados com o perfil de pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página, tiveram seus dados de *log* recalculados e reagrupados novamente. Devido ao erro encontrado foi necessário realizar ajustes em relacionados a percentual do método que determinava se a navegação estava sendo feita de modo linear ou não. Para identificar o percentual correto, foi calculada a média do percentual de navegação de todos os elementos rotulados com o perfil de pessoas com deficiência motora. Por fim o percentual foi alterado de 80% para 85%, permitindo uma precisão maior sobre o modo de navegação linear.

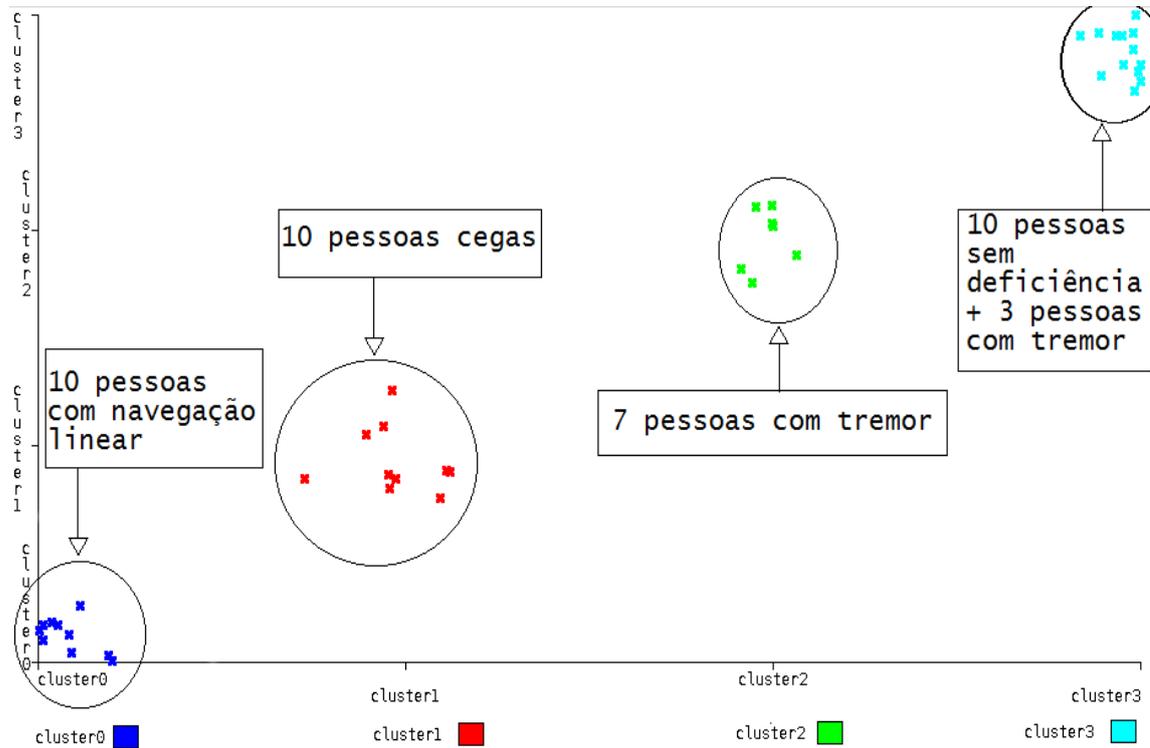
Mais adiante foi iniciada a etapa de análise dos resultados, onde foram utilizadas três matrizes de confusão para verificar o desempenho do algoritmo. A primeira matriz de confusão apresentada na [Tabela 8](#), foi criada com a finalidade mostrar uma visão geral dos elementos assim, como a distribuição dos mesmos em seus *clusters*. O primeiro *cluster* ficou com um total de 10 elementos, sendo que todos os elementos pertenciam ao rótulo do perfil de pessoas com deficiência motora que realizam a navegação de modo linear. No segundo *cluster* foram obtidos um total de 10 elementos, todos rotulados com o perfil de pessoas cegas. Já o terceiro *cluster* ficou com 7 elementos, todos os rótulos do perfil de pessoas com deficiência motora com tremor nos membros superiores. O quarto *cluster* obteve 13 elementos, sendo que 10 elementos correspondiam ao rótulo do perfil de pessoas sem deficiência e 3 elementos eram do rótulo do perfil de pessoas com deficiência motora com tremor nos membros superiores. Na [Figura 12](#) são apresentados de modo gráfico a distribuição dos dados em cada *cluster*.

Tabela 8 – Matriz de confusão contendo todos os *clusters*

		Classe Real				
		DM	DV	T	SD	TOTAL
Classe Prevista	DM	<b>10</b>	0	0	0	10
	DV	0	<b>10</b>	0	0	10
	T	0	0	<b>7</b>	0	7
	SD	0	0	3	<b>10</b>	13

**DM-** Deficiência Motora **DV-** Deficiência Visual **T-**Tremor **SD-** Sem Deficiência

A segunda matriz de confusão é exibida [Tabela 9](#) onde são apresentados somente os elementos rotulados com perfil de pessoas com deficiência motora com tremor nos

Figura 12 – Representação gráfica dos *clusters*

Fonte: a autora

Tabela 9 – *Cluster* com tremores

		Classe Real	
		T	N-T
Classe Prevista	T	7	0
	N-T	3	30

**T-** Tremor **N-T-** Não Tremor

membros superiores. Para determinar o desempenho do algoritmo em relação aos elementos, foi analisado seu percentual de acurácia, sensibilidade, especificidade e precisão. O algoritmo por sua vez demonstrou eficiência em relação à classificação correta dos elementos, obtendo 92,5% de acurácia. Porém, obteve 70% de sensibilidade de classificação de verdadeiros positivos. Esse percentual deve-se ao fato de que durante a realização das tarefas alguns participantes acabaram reproduzindo dois tipos de tremores: tremores leves e tremores mais intensos, causando diferenças entre os tipos de tremores gerando uma classificação distinta para cada tremor.

Por outro lado, o algoritmo demonstrou 100% de especificidade, isto é, todos os elementos que não eram rotulados como tremor não foram classificados como tremor, comprovando a qualidade do algoritmo em classificar os elementos que continham tremor. Do mesmo modo ocorreu com a precisão do algoritmo em relação aos verdadeiros positivos,

onde foram obtidos um percentual de 100%, demonstrando a precisão do algoritmo em identificar quais os dados eram tremores e realmente foram classificados como tremores. Entretanto, a precisão do algoritmo em identificar quais dados não eram tremores e realmente não foram classificados como tremores, foi de 90,91%. Esse valor percentual foi atingido devido à existência de dois tipos de tremores entre o conjunto de elementos, isso acabou afetando na precisão do algoritmo.

Tabela 10 – *Cluster* de pessoas sem deficiência

		Classe Real	
		SD	N-SD
Classe Prevista	SD	10	3
	N-SD	0	27

**SD-** Sem Deficiência **N-SD-** Não Sem Deficiência

Na [Tabela 10](#) é apresentada a terceira matriz de confusão que tem como propósito mostrar a distribuição dos dados rotulados como perfil de pessoas sem deficiência e o modo como foram classificados pelo algoritmo. Ao todo o algoritmo obteve um percentual de 92,5% de acurácia na classificação correta dos elementos rotulados como pessoas sem deficiência. Já referente a sensibilidade do algoritmo em identificar quais elementos eram rotulados como pessoas sem deficiência e realmente foram classificadas como pessoas sem deficiência foi de 100%.

Por outro lado o algoritmo obteve 90% de especificidade em reconhecer quais elementos realmente não pertenciam ao rótulo de pessoas com deficiência, como foi mencionado no parágrafo anterior durante a coleta de dados alguns participantes geraram dois tipos de tremores, levando o algoritmo a classificar incorretamente elementos com tremores leves como pessoas sem deficiência, devido a similaridade dos dados. Isto por sua vez acabou afetando a precisão do algoritmo em perceber se os elementos classificados como pessoas sem deficiência, realmente eram elementos rotulados com o perfil de pessoas sem deficiência, atingindo somente 76,92% de precisão.

No entanto, o algoritmo por sua vez apresentou 100% de precisão em identificar quais elementos se tratavam de pessoas com deficiência e foram classificados como pessoas com deficiência. Quanto aos demais *clusters*, todos apresentaram um percentual de 100% em relação à acurácia, sensibilidade, especificidade e precisão sobre os positivos e negativos. Então foi aplicada a soma da média aritmética relacionadas a acurácia e sensibilidade do algoritmo em relação a todos os *clusters*, e obtido um percentual de 96,25% de acurácia em classificar de modo correto os elementos e 92,5% de sensibilidade em classificar de modo correto os elementos de acordo com seu perfil.

Portanto, a partir dos resultados obtidos através da matriz de confusão e sobre o desempenho do algoritmo, podemos dizer que o algoritmo utilizado agrupou de maneira correta a maioria dos elementos. De fato, alguns elementos não foram agrupados de forma

correta, mas isso deve justamente ao modo como foi realizada a coleta de dados de cada perfil de usuário, que acabou gerando um ruído entre os elementos.



## 6 CONSIDERAÇÕES FINAIS

A Internet tem se mostrado cada dia mais presente no cotidiano das pessoas inclusive na realização de tarefas, e no modo de aprendizado. Isso por sua vez também inclui pessoas com deficiência que fazem uso de tecnologias assistivas para interagir com conteúdos de páginas da *web* que acaba sugerindo um modo de interação diferenciado das pessoas sem deficiência.

Devido a tal diferença o presente trabalho tem como finalidade realizar o agrupamento de perfis de usuários com deficiência visual, motora ou sem deficiência a partir da sua interação com páginas da *web*. Para capturar aspectos de interação dos usuários foi desenvolvido um sistema de coleta de informações, na linguagens *JavaScript* e PHP utilizando o framework *Laravel*, que ficou responsável por coletar e processar informações de interação do usuário com a página da *web* e posteriormente armazenar tais em informações em um banco de dados *MySQL*. Essas informações foram posteriormente agrupadas utilizando o algoritmo de agrupamento *K-Means*, que foi implementado através da ferramenta *Waikato Environment for Knowledge Analysis (WEKA)* que contém um conjunto de algoritmos para preparação de dados de aprendizagem de máquina, fornecendo relatórios com dados analíticos e estatísticos.

Durante a etapa de processamento dos dados foram encontrados alguns desafios, como identificar quais as informações coletadas ao longo da sessão seriam relevantes para caracterizar cada perfil de usuário. Outro desafio encontrado estava relacionado a identificar usuários com deficiência visual e motora para realização de testes e coleta de dados. Devido essa dificuldade em encontrar usuários com deficiências, foram gerados dados artificiais, através de uma técnica de triagem (*Screening*). Nessa técnica pessoas sem deficiência reproduzem o modo de interação de pessoas com deficiência, sendo permitido até mesmo o uso de tecnologias assistivas e tecnologias para impor limitações ao participante.

Assim, foram aplicados dois experimentos utilizando a técnica de triagem. No decorrer de cada um dos experimentos foi solicitado aos participantes que reproduzissem quatro perfis de usuários, sendo eles: pessoas sem deficiência, pessoas cegas, pessoas com deficiência motora que navegam usando o dispositivo apontador (*mouse*) com movimentos perfeitamente lineares pela página, e pessoas com deficiência motora com tremor nos membros superiores. Foi então utilizada uma matriz de confusão para investigar o grau de desempenho do algoritmo em relação à classificação de cada perfil de usuário, que estavam divididos em quatro *clusters*, contendo um perfil de usuário para cada *cluster*.

Ao final do primeiro experimento, alguns elementos não foram agrupados de modo correto pelo algoritmo. Entre os elementos agrupados de modo incorreto estavam um elemento rotulado com o perfil de pessoa com deficiência motora nos membros superiores e o outro rotulado no perfil de pessoas cegas. Ambos os elementos foram incorretamente agrupados junto com elementos do rótulo de pessoas sem deficiência. Do mesmo modo ocorreu com o segundo experimento que elementos rotulados com o perfil de pessoas

com o tremor foram agrupados com o perfil de pessoas sem deficiência. Após analisar os resultados obtidos com o agrupamentos de cada um dos experimentos, acreditamos que esse erro ocorreu porque durante o processo de coleta de dados foram gerados dois tipos de tremores: tremores leves e tremores mais intensos, levando o algoritmo a agrupar tremores mais leves como pessoas sem deficiência, devido a similaridade dos dados. O mesmo ocorreu com o elemento rotulado com o perfil de pessoas cegas, que por utilizar o *mouse* e realizar um movimento considerado linear, causando, uma diferença entre os elementos rotulados como pessoas cegas e esse elemento, pois os demais elementos nem ao menos faziam uso do *mouse*. Dessa forma, podemos dizer que algoritmo realizou o agrupamento correto na maioria dos casos, comprovando a eficácia do algoritmo *K-Means*.

No entanto, o presente trabalho apresenta limitações quanto à confiabilidade dos seus resultados, pois foram utilizados somente dados artificiais gerados a partir de uma técnica de triagem, e com um número limitado de participantes, que possuíam pouco ou nenhum conhecimento sobre técnicas de acessibilidade e uso de tecnologias assistivas. Outro aspecto que devemos levar em consideração está relacionado a quantidade atributos contido em cada descritor, e se de fato todos os atributos são uteis ou então suficientes para caracterizar o perfil dos usuários.

Acreditamos que em trabalhos futuros, será necessário realizar mais experimentos com usuários reais, que de fato possuam deficiência visual ou motora, para verificar a eficácia do algoritmo de agrupamento e dos descritores em agrupar de modo correto elementos de interação. Caso sejam obtidos resultados considerados satisfatórios o sistema de coleta de dados de interações, será integrado a uma página com a finalidade de classificar o perfil do usuário e sugerir modificações na páginas de acordo com o perfil desse usuário, como acionamento de recursos de acessibilidade destinados exclusivamente a pessoas com algum tipo de deficiência.

## REFERÊNCIAS

- AJUDAS TÉCNICAS, C. de. Tecnologia assistiva. 2006. Citado na página 29.
- ALVAREZ, A.; LUQUE, B. Rede neural de kohonen e outras técnicas para treinamento não-supervisionado. **Universidade Estadual de Campinas**, 2003. Citado na página 34.
- AMARAL, M. K. D. Avaliação de desempenho de algoritmos de consenso em canais sem fio usando modelo de outage. 2013. Citado na página 36.
- ARBELAITZ, O. et al. Global versus modular link prediction approach for discapnet: website focused to visually impaired people. In: **IEEE. Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on**. [S.l.], 2014. p. 51–58. Citado 3 vezes nas páginas 42, 43 e 44.
- BELL, J. **Machine Learning: Hands-on for developers and technical professionals**. [S.l.]: John Wiley & Sons, 2014. Citado na página 34.
- BIGHAM, J. P. et al. Webinsitu: a comparative analysis of blind and sighted browsing behavior. In: **ACM. Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility**. [S.l.], 2007. p. 51–58. Citado 3 vezes nas páginas 41, 43 e 44.
- BORGES, J. A. Dosvox—um novo acesso dos cegos à cultura e ao trabalho. **Revista Benjamin Constant**, n, 1996. Citado na página 25.
- BRAGA, J. Objetos de aprendizagem. 2014. Citado na página 31.
- BRASIL. Decreto nº 5.296/04, de 2 de dezembro de 2004. regulamenta as leis nos 10.048, de 8 de novembro de 2000, que da prioridade de atendimento às pessoas que especifica, e 10.098, de 19 de dezembro de 2000, que estabelece normas gerais e criterios basicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e da outras providências. **Diário Oficial da União**, 2004. Citado 3 vezes nas páginas 25, 26 e 30.
- BRASIL. Decreto nº 6.949, de 25 de agosto de 2009. promulga a convenção internacional sobre os direitos das pessoas com deficiência e seu protocolo facultativo, assinados em nova york, em 30 de março de 2007. **Diário Oficial da União**, 2009. Citado na página 25.
- CALDWELL, B. et al. **Diretrizes De Acessibilidade Para Conteúdo Web (WCAG) 2.0**. [S.l.]: Obtido em, 2015. Citado na página 21.
- CAMILO, C. O.; SILVA, J. C. d. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. **Universidade Federal de Goiás (UFG)**, p. 1–29, 2009. Citado na página 35.
- CONFORTO, D.; SANTAROSA, L. M. Acessibilidade à web: Internet para todos. **Revista de Informática na Educação: Teoria, Prática—PGIE/UFRGS**, v. 5, n. 2, p. 87–102, 2000. Citado na página 25.
- CORREIA, S.; CORREIA, P.; FRASSINETTI, E. d. P. Acessibilidade e desenho universal. **EP Frassinetti, Actas do Encontro Internacional Educação Especial**, p. 29–50, 2005. Citado na página 31.

- COSTA, E. et al. Mineração de dados educacionais: Conceitos, técnicas, ferramentas e aplicações. **Jornada de Atualização em Informática na Educação**, v. 1, n. 1, p. 1–29, 2013. Citado na página [36](#).
- CUSIN, C. A.; VIDOTTI, S. A. B. G. Inclusão digital via acessibilidade web| digital inclusion via web accessibility. **Liinc em Revista**, v. 5, n. 1, 2009. Citado na página [30](#).
- CYBIS, W. de A.; BETIOL, A. H.; FAUST, R. **Ergonomia e usabilidade: conhecimentos, métodos e aplicações**. [S.l.]: Novatec Editora, 2007. Citado 3 vezes nas páginas [31](#), [32](#) e [46](#).
- DIAS, C. **Usabilidade na web: criando portais mais acessíveis**. [S.l.]: Alta Books, 2003. Citado na página [31](#).
- DOMINGUES, M. L. C. S. **Mineração de Dados Utilizando Aprendizado Não-Supervisionado: um estudo de caso para bancos de dados da saúde**. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, 2003. Citado 2 vezes nas páginas [33](#) e [35](#).
- ENCARNAÇÃO, L. M. E. da. **Concept and realization of intelligent user support in interactive graphics applications**. Tese (Doutorado), 1997. Citado na página [32](#).
- FERREIRA, S. B. L. E-acessibilidade: tornando visível o invisível. **Revista Morpheus- Estudos Interdisciplinares em Memória Social**, v. 6, n. 10, 2007. Citado na página [26](#).
- FERREIRA, S. B. L.; RODRIGUES, R. N. **e-Usabilidade**. [S.l.]: Grupo Gen-LTC, 2008. Citado 5 vezes nas páginas [25](#), [26](#), [27](#), [29](#) e [30](#).
- FERRETTI, S. et al. Exploiting reinforcement learning to profile users and personalize web pages. In: IEEE. **Computer Software and Applications Conference Workshops (COMPSACW), 2014 IEEE 38th International**. [S.l.], 2014. p. 252–257. Citado 2 vezes nas páginas [41](#) e [43](#).
- FRANCISCO, P. C. M.; MENEZES, A. M. de. Design universal, acessibilidade e espaço construído. **CONSTRUINDO**, v. 3, n. 01, 2011. Citado 2 vezes nas páginas [21](#) e [30](#).
- GODINHO, F. et al. Tecnologias de informação sem barreiras no local de trabalho. **Vila Real: UTAD**, 2004. Citado na página [27](#).
- GOLDSCHMIDT, R. R. Uma introdução à inteligência computacional: fundamentos, ferramentas e aplicações. **Rio de Janeiro Brasil: IST-Rio**, 2010. Citado 5 vezes nas páginas [32](#), [33](#), [34](#), [35](#) e [36](#).
- HARTIKAINEN, M.; OVASKA, S. People with parkinson’s disease using computers. In: ACM. **Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility**. [S.l.], 2015. p. 407–408. Citado na página [27](#).
- HENRY, S. L. Education and outreach working group (eowg). **Introduction to Web Accessibility. W3C/WAI–World Wide Web Consortium/Web Accessibility Initiative**, 2005. Citado na página [30](#).

HESS, C. W.; PULLMAN, S. L. Tremor: clinical phenomenology and assessment techniques. **Tremor and other hyperkinetic movements**, Center for Digital Research and Scholarship, v. 2, 2012. Citado na página 47.

HEUSER, C. A. **Projeto de banco de dados: Volume 4 da Série Livros didáticos informática UFRGS**. [S.l.]: Bookman Editora, 2009. Citado na página 53.

HOFFMANN, R. et al. Análise de regressão: uma introdução à econometria. O autor, 2015. Citado na página 33.

HOFMANN, T.; BUHMANN, J. M. Pairwise data clustering by deterministic annealing. **Ieee transactions on pattern analysis and machine intelligence**, IEEE, v. 19, n. 1, p. 1–14, 1997. Citado na página 36.

IBGE, C. características gerais da população, religião e pessoas com deficiência. **Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística**, 2010. Citado na página 25.

JR, D. W. H.; LEMESHOW, S. **Applied logistic regression**. [S.l.]: John Wiley & Sons, 2000. Citado na página 34.

LEPORINI, B.; ANDRONICO, P.; BUZZI, M. Designing search engine user interfaces for the visually impaired. In: ACM. **Proceedings of the 2004 international cross-disciplinary workshop on Web accessibility (W4A)**. [S.l.], 2004. p. 57–66. Citado na página 26.

LIMA, C. F. L.; ASSIS, F.; SOUZA, C. Árvores de decisão baseadas nas entropias de shannon, rényi e tsallis para sistemas tolerantes a intrusão. In: **La Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informática: CISCI**. [S.l.: s.n.], 2010. Citado 2 vezes nas páginas 34 e 35.

MELO, A. M. Acessibilidade e inclusão digital. **Livro dos Tutoriais do XIII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais**, p. 29, 2014. Citado 2 vezes nas páginas 29 e 30.

MELO, A. M.; COSTA, J. B. da; SOARES, S. C. de M. Tecnologias assistivas. **ACESSIBILIDADE**, p. 62, 2006. Citado na página 29.

MELO, H. F. R. de. **Deficiência visual: lições práticas de orientação e mobilidade**. [S.l.]: UNICAMP, 1991. Citado na página 25.

MENDES, C. A. B.; VEGA, F. A. C. Técnicas de regressão logística aplicada a análise ambiental. **Geografia (Londrina)**, v. 20, n. 1, p. 5–30, 2011. Citado na página 34.

MESQUITA, P. S. B.; NOGUEIRA, R. T. Um modelo estatístico para gestão de programas de pós-graduação. **Gepros: Gestão da Produção, Operações e Sistemas**, Universidade Estadual Paulista-UNESP Bauru, Depto de Engenharia de Produção, v. 10, n. 2, p. 173, 2015. Citado na página 34.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas Inteligentes-Fundamentos e Aplicações**, v. 1, n. 1, 2003. Citado 2 vezes nas páginas 32 e 33.

- NICHOLL, A.; FILHO, J. J. B. O ambiente que promove a inclusão: conceitos de acessibilidade e usabilidade. **Assentamentos Humanos Magazine**, v. 3, n. 2, 2001. Citado na página 30.
- OLIVEIRA, F. L. D. et al. Modelo de interfaces adaptativas utilizando redes bayesianas. **ECOINFO: Encontro de Estudantes de Informática do Tocantins. Anais do ECOINFO/EIN. UFT. Palmas, TO**, p. 139–148, 2003. Citado 2 vezes nas páginas 31 e 32.
- ONODA, M.; EBECKEN, N. F. Implementação em java de um algoritmo de árvore de decisão acoplado a um sgbd relacional. In: **SBBD**. [S.l.: s.n.], 2001. p. 55–64. Citado na página 34.
- OPPERMANN, R. **Adaptive user support: ergonomic design of manually and automatically adaptable software**. [S.l.]: CRC Press, 1994. Citado na página 32.
- PALALLO, L. A. M. **Modelos Proativos para Hipermedia Adaptativa**. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, 2000. Citado na página 32.
- PARDO, T. A. S.; NUNES, M. d. G. V. Aprendizado bayesiano aplicado ao processamento de línguas naturais. **Série de Relatórios Técnicos do Instituto de Ciências Matemáticas e de Computação-ICMC, Universidade de São Paulo**, 2002. Citado na página 35.
- PERKOWITZ, M.; ETZIONI, O. Adaptive web sites. **Communications of the ACM**, ACM, v. 43, n. 8, p. 152–158, 2000. Citado na página 32.
- PETERNELLI, L. A. Capítulo 2: estatística descritiva. **Material didático. s/d. Disponível em:** < <http://www.dpi.ufv.br/~peternelli/inf162.www>, v. 16032004, 2012. Citado na página 33.
- PETERSEN, K. et al. Systematic mapping studies in software engineering. In: SN. **12th international conference on evaluation and assessment in software engineering**. [S.l.], 2008. v. 17, n. 1. Citado na página 39.
- PRATI, R.; BATISTA, G.; MONARD, M. Curvas roc para avaliação de classificadores. **Revista IEEE América Latina**, v. 6, n. 2, p. 215–222, 2008. Citado na página 49.
- PUPO, D. T.; MELO, A. M.; FERRÉS, S. P. Acessibilidade: discurso e prática no cotidiano das bibliotecas. **Campinas: Unicamp/Biblioteca Central Cesar Lattes**, 2006. Citado na página 31.
- RAMIRO, G. et al. ACESSWEB—uma perspectiva sobre a acessibilidade web em português. Tânia Rocha, 2015. Citado na página 21.
- SANTANA, V. F. de; BARANAUSKAS, M. C. C. Welfit: A remote evaluation tool for identifying web usage patterns through client-side logging. **International Journal of Human-Computer Studies**, Elsevier, v. 76, p. 40–49, 2015. Citado 3 vezes nas páginas 42, 43 e 44.
- SANTAROSA, L. M. C. et al. Tecnologias digitais acessíveis. **Porto Alegre: JSM Comunicação Ltda**, 2010. Citado 3 vezes nas páginas 21, 27 e 29.

- SILVA, L. N. d. C. **Análise e síntese de estratégias de aprendizado para redes neurais artificiais**. 1998. 210 p. Tese (Doutorado) — Dissertação (Mestrado em Engenharia Elétrica)—Faculdade de Engenharia Elétrica e Computação. Universidade Estadual de Campinas, Campinas, 1998. Citado na página 34.
- SILVA, M. P. d. S. Mineração de dados: Conceitos, aplicações e experimentos com weka. **Livro da Escola Regional de Informática Rio de Janeiro-Espírito Santo. Porto Alegre: Sociedade Brasileira de Computação**, v. 1, p. 1–20, 2004. Citado na página 48.
- SILVA, P. E. C.; SILVA, P. F. P. da. Interfaces adaptativas aplicadas a sistemas de informação—características desejáveis. **Revista Abstração Volume 2**, p. 5, 2007. Citado na página 32.
- SOMMERVILLE, I. et al. **Engenharia de software**. [S.l.]: Addison Wesley São Paulo, 2003. v. 6. Citado na página 53.
- SONZA, A.; KADE, A.; FAÇANHA, A. Acessibilidade e tecnologia assistiva: Pensando a inclusão sociodigital de pnes. **Bento Gonçalves: Instituto Federal do Rio Grande do Sul Campus Bento Gonçalves**, 2013. Citado na página 26.
- SONZA, A. P. Ambientes virtuais acessíveis sob a perspectiva de usuários com limitação visual. 2008. Citado 3 vezes nas páginas 27, 30 e 31.
- STEVENSON, J. Estatística. 1981. Citado na página 33.
- STRINGHINI, L. F. Regressão linear—uma aplicação em custos. 2005. Citado na página 33.
- THATCHER, J.; WADDELL, C.; BURKS, M. **Constructing accessible web sites**. [S.l.]: Glasshaus Birmingham, 2002. v. 34. Citado na página 30.
- THIELO, M. R. **Análise e classificação de séries temporais não estacionárias utilizando métodos não-lineares**. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, 2000. Citado na página 36.
- VIEIRA, A. C. H.; PONTES, A. M.; PALAZZO, L. A. M. Projetando interfaces adaptativas para comunidades virtuais de aprendizado. In: **XXII Congresso da Sociedade Brasileira de Computação**. vl. [S.l.: s.n.], 2002. v. 5, p. 383–387. Citado na página 32.
- VIGO, M. et al. Automatic creation of user profiles for achieving personal web accessibility. In: SPRINGER. **International Conference on Computers for Handicapped Persons**. [S.l.], 2008. p. 354–361. Citado 2 vezes nas páginas 40 e 43.
- WAZLAWICK, R. **Metodologia de Pesquisa para Ciência da Computação, 2ª Edição**. [S.l.]: Elsevier Brasil, 2014. v. 2. Citado na página 45.