

# Usando K-Means para Descoberta de Perfis de Aplicativos Educacionais no Repositório Google Play Store

Matheus Máximo R. Campos, Lesandro Ponciano

Pontifícia Universidade Católica de Minas Gerais  
Bacharelado em Sistemas de Informação

matheus.maximo@sga.pucminas.br, lesandrop@pucminas.br

**Resumo.** *Este trabalho tem o objetivo de realizar uma análise de aplicativos mobile que auxiliam o estudante no processo de aprendizagem. Busca-se agrupá-los, utilizando algoritmos de mineração de dados, por suas características funcionais e de interação a fim de identificar perfis através de um algoritmo de agrupamento, analisar os grupos encontrados de modo a entender sua semelhanças e diferenças. Através deste processo, são definidas seis métricas para os aplicativos, que tratam de três diferentes aspectos: Nível de avaliação do aplicativo; Nível de atenção dada ao aplicativo; Tempo de disponibilidade e atualização do aplicativo. É realizada a coleta de dados de aplicativos de Auxílio à Aprendizagem com foco nas matérias de Matemática e Física, esses aplicativos são agrupadas em cinco grupos distintos, nomeados como: Populares; Impopular; Bem Avaliado Desatualizado; Bem Avaliado Antigo; Bem Avaliado Novo. A análise de aplicativos representantes destes grupos, é feita por estudantes e professores de cursos pré-vestibulares e estudantes do primeiro e segundo período de cursos de graduação seguindo a escala System Usability Scale (SUS).*

**Palavras-Chaves:** Agrupamento, TIC, Dispositivos Móveis, Processo de Aprendizagem, SUS

## 1. Introdução

A quantidade de celulares no Brasil já é maior que a quantidade de habitantes, dados<sup>1</sup> mostram que em janeiro de 2018 o país já possuía mais de 236 milhões de telefones móveis, sendo estes um dos meios mais utilizados para acesso à internet, comunicação e busca rápida de informações. Os dispositivos móveis possuem um número considerável de usuários entre 10 e 30 anos [Oliveira 2014], ou seja, são utilizados por pessoas que se encontram desde o ciclo básico de ensino até o superior. O uso dos celulares para acesso à informação e aprendizagem, o *M-Learning (Mobile Learning)*, vem expandindo as formas de aprender, onde vemos a possibilidade de um processo de aprendizagem colaborativo. Estudantes podem interagir extraclasse e realizarem a promoção do conhecimento de forma conjunta, saindo assim do conceito implantado de aprendizagem apenas vinda do professor e apenas dentro da escola, aumentando os caminhos existentes no processo de aprendizagem [Alencar et al 2015].

---

<sup>1</sup> TELEFONIA MÓVEL - ACESSOS disponível em:

<<http://www.anatel.gov.br/dados/destaque-1/283-brasil-tem-236-2-milhoes-de-linhas-moveis-em-janeiro-de-2018>> . Acessado em: 16 abril 2018.

O **problema** investigado na pesquisa apresentada neste artigo, é que apesar do uso cada vez maior dos recursos de TIC (Tecnologias de Informação e Comunicação), especialmente o celular, pouco se sabe sobre as características dos aplicativos para dispositivos móveis (*mobile*) usados pelos estudantes no auxílio de seu processo de aprendizagem. Tem-se hoje uma enorme gama de aplicativos disponíveis com o intuito de auxiliar neste ponto, porém, não há informações suficientes sobre o sua eficácia e sua variedade. Pouco se sabe sobre as características de aplicativos que tenham uma melhor aceitação por parte dos estudantes.

O objetivo principal deste trabalho é descobrir quais os perfis típicos de aplicativos para dispositivos móveis voltados para o auxílio do processo de aprendizagem. Ênfase é dada a aplicativos associados às matérias de Física e Matemática disponíveis no repositório *Play Store* do Google, pois para ambas as matérias há um número consideravelmente maior comparado às demais matérias lecionadas no ensino médio, proporcionando uma maior massa de dados para análise. Para alcançar este objetivo, realiza-se um levantamento de todos os aplicativos disponíveis na *Play Store* do Google, que encaixam no escopo definido. Após realizar este levantamento, é aplicado um algoritmo de agrupamento (*clustering*) sobre os dados do aplicativos obtidos, para levantar todos os perfis existente dentre estes aplicativos e analisar cada grupo a fim de identificar perfis de aplicativos. Por fim, investiga-se a opinião dos alunos e professores acerca de cada aplicativo representante dos perfis levantados.

Analisou-se o perfil de aplicativos de acordo com os dados coletados do repositório *Play Store* do Google para aplicativos de estudo para as matérias de Matemática e Física. Em ambas as matérias, são encontrados cinco diferentes perfis baseados em três diferentes aspectos: nível de avaliação do aplicativo; nível de atenção dada ao aplicativo; tempo de disponibilidade/atualização do aplicativo. Os perfis são nomeados da seguinte maneira: Popular; Impopular; Bem Avaliado Desatualizado; Bem Avaliado Antigo; Bem Avaliado Novo. Em seguida, é realizada uma análise de distribuição dos aplicativos pelos perfis, onde a maior porcentagem de aplicativos (47% para a matéria de Matemática e 53% para a matéria de Física) aparece no perfil Bem Avaliado Novo, enquanto o menor número de aplicativos (0,5%) aparece no perfil Impopular para os aplicativos de Física e para o perfil Popular (1%) nos aplicativos de Matemática. Realizada também a escolha de um aplicativo representante para cada perfil, tendo como critério os aplicativos com a menor distância euclidiana do centróide de cada perfil. Utilizando os aplicativos representantes de cada perfil, pode-se realizar um teste com os usuários alvo desses aplicativos, para analisar a resposta dos usuários sobre os aplicativos representantes.

Por fim, neste estudo, pode-se citar duas principais contribuições. Primeira, a proposta de métricas e grupos para os perfis de aplicativos de auxílio no processo de aprendizagem para as matérias de Física e Matemática, realizando uma análise quantitativa desses aplicativos. Segunda, um levantamento qualitativo da opinião do público alvo desses aplicativos sobre cada perfil de aplicativo, podendo assim, propor qual o melhor modelo de aplicativo para este tipo de aplicativo.

O restante do artigo está organizado da seguinte forma: na Seção 2 é apresentado a contextualização sobre os ambientes computacionais de aprendizagem móvel e

aprendizagem colaborativa por meio de TIC e o tema de agrupamento de dados com definições relevantes para o entendimento do trabalho. Na Seção 3 são apresentados os trabalhos relacionados. Em seguida, a seção 4 apresenta detalhes da metodologia utilizada. Os resultados do trabalho são apresentados na Seção 5. Por fim, na Seção 6 são apresentados as conclusões e trabalhos futuros.

## **2. Referencial Teórico**

Esta seção está dividida em três subseções. A primeira realiza a contextualização sobre ambientes computacionais móveis de aprendizagem. A segunda aborda a aprendizagem colaborativa por meio de TIC. Por fim, terceira aborda o tema de agrupamento de dados com definições relevantes para o entendimento deste trabalho.

### **2.1. Ambientes Computacionais de Aprendizagem *Mobile***

Os ambientes computacionais móveis vêm proporcionando uma nova modalidade de ensino, a aprendizagem móvel. Este novo modelo, caracteriza-se pela capacidade de proporcionar uma forte interação entre os estudantes e professores, como também, uma grande interação entre estudantes e estudantes, caracterizando um ambiente de aprendizagem colaborativa. Além destes modelos de interação, possibilita aos usuários desses aplicativos, participar e acessar o ambiente de ensino por meio de dispositivos móveis (celulares, *tablets*, *laptops*, entre outros) a qualquer momento e em qualquer lugar [Laine et al. 2010]

A fim de desenvolver estes ambientes de aprendizagem para que sejam sensíveis ao contexto do estudante, é essencial que os conteúdos educacionais sejam criados de maneira padronizada. Desta forma, é possível que um ambiente de suporte à aprendizagem exiba o conteúdo de forma adequada e reutilize conteúdos em diferentes contextos e diversos repositórios. Além disso, é possível combinar conteúdos distintos, o que, por sua vez, melhora o processo de produção de conteúdo, como consequência, reduz os seus custos [Rodolpho 2009].

Para melhorar a absorção de conhecimento pelos estudantes na aprendizagem móvel, deve-se levar em consideração as características particulares de cada estudante. Isso é necessário não apenas para fornecer um conteúdo que atenda às características cognitivas dos estudantes, mas também para prover conteúdos de forma adequada às restrições dos seus dispositivos móveis, uma vez que os mesmos possuem recursos distintos e limitados. Surge então o conceito de ambientes sensíveis ao contexto. Esse tipo de ambiente se adequa ao perfil do usuário, levando em consideração informações fornecidas pelo próprio usuário, além daquelas captadas dinamicamente a partir de sua interação com os dispositivos computacionais [Moore, Hu e Wan 2011].

### **2.2. Aprendizagem Colaborativa Através de TICs**

Para Souza (2012), a aprendizagem é promovida pelo compartilhamento e pelo uso da informação, os quais, como resultado, possibilitam novos olhares e novos aprendizados, cujos mais significativos são os novos conhecimentos e habilidades adquiridos. As redes sociais, por se constituírem em espaços favoráveis ao compartilhamento da informação e do conhecimento, podem também se configurar como espaços de aprendizagem, tornando-se um fértil ambiente para o desenvolvimento e inovações pedagógicas. Isso porque as redes sociais permitem a partilha de conteúdos em múltiplos suportes. A

familiaridade dos alunos com as redes sociais também é um fator relevante quando se trata de sua utilização em contextos escolares de aprendizagem.

Souza (2012) ainda afirma que é importante proporcionar ao docente uma visão mais abrangente da inserção das tecnologias, a partir de suas próprias vivências e das de seu alunado. O desenvolvimento de estratégias para a formação em serviço, através da pesquisa e da busca por novas formas de ensinar e aprender, é uma maneira eficiente de se estabelecer a articulação entre teoria e prática, tão necessária para a superação dos desafios educacionais do século XXI.

Para Lupion Torres (2004), o sucesso de uma proposta de aprendizagem colaborativa depende fundamentalmente que todas as atividades sejam planejadas de modo a provocar rupturas e desafiar os educandos. Desta forma pode-se levá-los a formar uma comunidade de aprendizagem coesiva e reflexiva, cujos membros trabalham para alcançar objetivos comuns enquanto respeitam a diversidade de ideias, valores, crenças e estilos de vida. Com estas as atividades planejadas, pretende-se um movimento espiral, uma sucessão de equilíbrios e desequilíbrios, um estar aberto para o novo. Busca-se por meio de uma sucessão de atividades organizadas e com regras preestabelecidas promover rupturas que gerem situações de desequilíbrio com força suficiente para provocar mudanças, que levem o aluno a se assumir sujeito pesquisador e autônomo.

### **2.3. Agrupamento de Dados**

Agrupamento é a classificação não-supervisionada de dados, grupos de dados baseado em suas características em comum. Ela representa uma das principais etapas de processos de análise de dados, denominada análise de grupos. A análise de grupos envolve, portanto, a organização de um conjunto de padrões (usualmente representados na forma de vetores de atributos ou pontos em um espaço multidimensional – espaço de atributos) em grupos, de acordo com alguma medida de similaridade. Intuitivamente, padrões pertencentes a um dado grupo, devem ser mais similares entre si do que em relação a padrões pertencentes a outros grupos [Doni 2004].

É necessário distinguir agrupamento de análise discriminante. Em análise de discriminante ocorre uma classificação supervisionada, são fornecidos padrões rotulados e o problema é rotular novos padrões, ainda não-rotulados. Em agrupamento, é realizada uma classificação não-supervisionada, o problema aqui, é agrupar um conjunto de padrões não-rotulados em grupos que possuam algum significado, ou seja, de tal modo que os padrões apresentem alguma propriedade comum. Sendo assim, uma vez definidos os grupos, os padrões também estarão “rotulados”, mas o rótulo aqui é ditado pelos próprios padrões que compõem cada grupo e não por um padrão pré-estabelecido pelo pesquisador [Moscatto e Fernando 2002].

O algoritmo de agrupamento K-Means, baseia-se na distância euclidiana, tendo como critério de agregação das respostas, o critério centróide. Esse algoritmo de agrupamento, utiliza um método de agrupamento não-hierárquico por repartição, que consiste num procedimento em que, dado um número de clusters previamente determinado, calcula-se pontos que representam os “centros” desses grupos, espalhados homogeneamente no conjunto de dados e movidos, heurísticamente, até grupos preestabelecidos, e a melhor partição dos ‘n’ casos será aquela que otimize o critério

escolhido. Em outras palavras, o procedimento inicia-se usando os valores dos primeiros 'k' casos como estimativas temporárias das médias dos grupos. Os centróides iniciais são formados pela designação de cada caso ao cluster de centro mais próximo. Com a inclusão de cada caso, a média altera-se, alterando assim o centróide. Um processo iterativo é usado para achar os centróides finais de cada cluster. A cada passo, os casos são agrupados ao cluster de centro mais próximo, e novamente as médias são recalculadas. Este processo continua até que não haja mais alterações nas médias ou que um número predeterminado de iterações aconteça, encerrando-se o processo [HAIR et al., 2005].

Um dos métodos utilizados para determinar qual o melhor número de grupos para o algoritmo K-Means é o método Elbow [Kodinariya & Makwana, 2013]. A idéia do método é realizar o agrupamento para cada valor contido em um intervalo previamente definido pelo analista dos dados, com quantidades de grupos mínima e máxima. Para cada valor contido no intervalo é realizada a soma dos quadrados dos erros (SSE) das distâncias dos objetos de cada classe em relação ao centróide da respectiva classe. Após a plotagem do gráfico que define a SSE em relação ao número de classes, verifica-se o ponto de inflexão a partir do qual o decréscimo da SSE passa a ser menor quando do aumento de uma unidade no número de classes [Morais, 2016].

### **3. Trabalhos Relacionados**

Nesta seção são apresentados os trabalhos relacionados a análise de características de aplicativos para dispositivos móveis e avaliação de softwares educacionais voltados para a processo de aprendizagem. Embora os trabalhos não tratam diretamente do agrupamento de perfis de aplicativos já existentes, eles apresentam diversas informações que nortearão a abordagem a ser proposta neste trabalho.

Cruz (2015) propõe a criação de um *guideline* para a elaboração de aplicativos educacionais a serem utilizados em dispositivos móveis de comunicação, visando uma maior aproximação dos jovens e adolescentes que possuem esses aparelhos. O artigo parte do pressuposto que não existe um *guideline* ideal para o desenvolvimento de um aplicativo para dispositivos móveis e da necessidade de se criar uma. Apresenta-se recomendações de um *guideline* baseado em recomendações dos principais fabricantes.

Filho e Ellen (2012) apresentam um conjunto de características e requisitos específicos para ambientes de aprendizagem móvel. O artigo baseia-se no fato de que não há indícios de um conjunto de características e requisitos bem formulado no contexto de desenvolvimento de ambientes de aprendizagem móvel, ressaltando a necessidade de se definir criteriosamente características e requisitos de desenvolvimento. O artigo apresentou um conjunto de características e requisitos específicos para ambientes de aprendizagem móvel, estabelecido através da condução de uma Revisão Sistemática [Filho e Ellen 2012].

Gomes (2012) propõe uma metodologia alternativa para a avaliação de softwares para o ensino da Matemática. O artigo parte do pressuposto que as metodologias tradicionais não vêm atingindo o objetivo de gerar bons softwares para as propostas de ensino existentes. Para conclusão do estudo, foi realizado um minicurso com cinco professores de Matemática do ensino fundamental, para avaliar o uso de duas

ferramentas educacionais, Aritmética TICK-TACK-TOE e KidMath, no processo de aprendizagem dos alunos, ao utilizarem as ferramentas.

Nascimento (2012) apresenta a maior facilidade para resolver e aprender, por parte dos alunos, ao utilizarem o software GeoGebra em problemas de geometria. O artigo baseia-se na necessidade dos alunos de recursos tecnológicos para auxiliarem na resolução de problemas de geometria de certa complexidade. Para obtenção dos dados do artigo, foi realizada uma pesquisa experimental, onde foram apresentados aos alunos e professores de Matemática os recursos e a potencialidade do GeoGebra para resolução destes problemas.

O presente trabalho reúne as principais ideias dos trabalhos citados anteriormente. O que difere este trabalho dos demais, é o fato de se considerar os perfis de utilização dos aplicativos, no escopo escolhido, em uma linha de utilização para auxílio no processo de aprendizagem de maneira mais eficaz, segundo os próprios usuários. Não sendo comparados com modelos pré-estabelecidos de avaliação e de desenvolvimento de aplicativos, mas sim, analisando a opinião de estudantes e professores que realizam o uso desses aplicativos já em circulação. Avaliado-se, também, a quantidade de download e *rating* de avaliação que contabilizada na plataforma do Google Play pelos usuários que realizam o *download* desses aplicativos.

## **4. Metodologia**

Esta seção está dividida em quatro subseções. Na primeira serão exibidos os métodos e ferramentas utilizados para obtenção dos dados dos aplicativos de Matemática e Física. Serão apresentados também, as métricas propostas para o nível de avaliação do aplicativo, nível de atenção dada os aplicativos e tempo de disponibilidade e atualização dos aplicativos. Na segunda seção serão apresentadas as estratégias de agrupamento baseadas nos valores das métricas dos aplicativos. Estes grupos permitirão identificar os perfis a partir de suas similaridades. Na terceira seção serão apresentados os métodos de medição da usabilidade de aplicativos. Por fim, serão apresentados os métodos de coleta de opiniões dos usuários alvos dos aplicativos foco deste trabalho.

### **4.1. Obtenção de Dados e Métricas**

Os dados dos aplicativos voltadas para as matérias de Física e Matemática disponíveis para a plataforma Android são obtidos do repositório do Google Play utilizando a API disponibilizada pela Google, onde é possível ter acesso às informações dos aplicativos. O acesso a API é realizado utilizando a ferramenta de código aberto (*Open Source*) desenvolvida na tecnologia NodeJS, GooglePlay Scraper<sup>2</sup>, onde é possível realizar a pesquisa dos aplicativos através vários critérios pré-programados na ferramenta, além de consolidar as informações dos aplicativos e manifestações dos usuários que fizeram o download dos aplicativos. Os aplicativos de cada uma das matérias (Matemática e Física), são pré-filtrados para que fossem obtidos apenas as mais novos aplicativos não pagos (*Top Free Applications*), recomendadas para usuários acima de nove anos (última faixa etária disponível para filtragem no *storage*) e que estejam na língua portuguesa.

---

<sup>2</sup> GooglePlayScraper - Acesso disponível em: <<https://github.com/facundoolano/google-play-scraper>>

Para análise das informações dos aplicativos, são definidos seis métricas, divididas em três aspectos: nível de avaliação do aplicativo, que é dada pelo número de instalações e número de comentários; nível de atenção dada aos aplicativos, que é dado pelo número de avaliações e a média das avaliações; tempo de disponibilidade e atualização dos aplicativos, onde são analisados o tempo, em dias, desde que o aplicativo foi lançado e o tempo, em dias, desde a última atualização disponibilizada. As métricas descritas anteriormente têm como objetivo auxiliar no processo de agrupamento dos aplicativos correlacionados, que possuam o mesmo padrão dentro da Google Play Store.

## 4.2. Normalização e Agrupamento

É usado o algoritmo de agrupamento K-Means para encontrar grupos de aplicativos que possuem valores similares as métricas definidas. A entrada de dados é uma matriz  $|N| \times 6$  onde cada linha refere a um aplicativo e as colunas a uma das seis métricas. Como o resultado dos agrupamentos depende de que os valores relativos dos parâmetros sejam agrupados, uma normalização dos parâmetros antes do agrupamento é necessária. Como normalização foi utilizado um intervalo definido entre 0 e 1. Para obter esta escala, foi utilizado a fórmula  $x_i = (x_i - x_{min}) \div (x_{max} - x_{min})$ , onde  $x$  representa o valor da variável e  $i$  os aplicativos.

Para selecionar o número adequado de grupos, primeiro foi fixado o centroide inicial do algoritmo de K-Means a fim de diminuir a variabilidade no resultado do algoritmo, uma vez que o algoritmo de K-Means inicia o cálculo do centróide de maneira aleatória, quando não fixado. Logo após, foi executado o algoritmo K-Means variando o número de grupos ( $k$ ) em um intervalo pré-selecionado, avaliando a qualidade dos agrupamentos através do algoritmo de Elbow, para assim selecionar o número adequado de grupos.

## 4.3. Escala de Usabilidade de Sistemas

A escala SUS (*System Usability Scale*) foi desenvolvida em 1986, por John Brooke, no laboratório da *Digital Equipment Corporation*, no Reino Unido. É um questionário composto por 10 itens, com 5 opções de respostas. SUS é uma tecnologia independente e já foi testado em hardware, software de consumo, sites, telemóveis, unidade de resposta audível (URA) e até páginas amarelas. Tornou-se um padrão da indústria com referências em mais de 600 publicações [Jeff, S. 2009].

O respondente do questionário da SUS, assinala sua resposta para as perguntas, disponíveis na Tabela 1, em uma escala Likert que varia de: Discordo Totalmente a Concordo Totalmente. A versão em português das questões são extraídas de Tenório (2011).

**Tabela 1. Perguntas aplicadas no formulário da SUS.**

Nº.	Pergunta
1	Eu acho que gostaria de usar esse sistema com frequência.
2	Eu acho o sistema desnecessariamente complexo.
3	Eu achei o sistema fácil de usar.
4	Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema.
5	Eu acho que as várias funções do sistema estão muito bem integradas.
6	Eu acho que o sistema apresenta muita inconsistência.
7	Eu imagino que as pessoas aprenderão como usar esse sistema rapidamente.
8	Eu achei o sistema atrapalhado de usar.
9	Eu me senti confiante ao usar o sistema.
10	Eu precisei aprender várias coisas novas antes de conseguir usar o sistema.

O resultado da SUS é a soma da contribuição individual de cada item. Para os itens ímpares deve-se subtrair 1 à resposta do usuário, ao passo que para os itens pares o score é 5 menos a resposta do usuário. Depois de obter o score de cada item, somam-se os scores e multiplica-se o resultado por 2,5. Desta forma, o resultado obtido será um índice de satisfação do usuário, que varia de 0 a 100 [Brooke, J. 1986].

#### **4.4. Obtenção e Avaliação da Opinião dos Usuários**

Para a escolha da Técnica de Coleta de Opiniões a ser utilizada, foi considerado o fato dos estudantes e professores que utilizam os aplicativos que compõem o escopo desta pesquisa, estarem dispersos por vários níveis de escolaridade. Desta forma, a utilização de um questionário para esta pesquisa evitaria custos em deslocamentos para entrevistas e, ainda, possibilitaria atingir usuários de vários grupos.

O Questionário da SUS foi escolhido como instrumento para avaliação de usabilidade deste estudo, por ser um instrumento gratuito e devido ao pequeno número de questões, fato que poderia facilitar a aderência à pesquisa. Por fim, outro fator relevante para a escolha do SUS foi a necessidade de um instrumento objetivo para decidir um perfil de aplicativo que tem maior aceitação por parte dos estudantes e professores, a escala escolhida apresenta uma métrica para avaliação da Usabilidade e Aderência de Conteúdo. Para facilitar a coleta de dados, foi disponibilizado através do GoogleForms<sup>3</sup> o questionário SUS para estudantes e professores de Instituições de

---

<sup>3</sup> Formulário de Física <<https://goo.gl/forms/eB2kPD41R5FIv5C52>> - Formulário de Matemática <<https://goo.gl/forms/h5NGy2EGhpzBiPgy2>>



Pré-Vestibular e Preparação para concursos e estudantes universitários do primeiro e segundo período.

## 5. Resultados

Nesta seção, apresenta-se o resultado da aplicação do método proposto para analisar os perfis de aplicativos das matérias de Matemática e Física e opiniões dos usuários alvos desses aplicativos. Primeiro, apresentamos os resultados obtidos ao realizar o agrupamento e cálculo do aplicativo que representa o grupo. Após apresentamos os resultados da pesquisa qualitativa com os alunos e professores do Ensino Médio, sobre os aplicativos que representam cada grupo.

### 5.1. Normalização e Agrupamento

Ao realizar a obtenção dos dados, pôde-se observar que a unidade de medida afetaria a análise e agrupamento dos dados. Para que não houvesse problemas de agrupamento, devido a proporção entre as variáveis, estas sofreram normalização. Foi aplicado o método de normalização proposto na metodologia, para que fosse limitado a escala de todas as variáveis entre 0 e 1, tornando o agrupamento mais assertivo. Na Figura 1 e Figura 2, pode-se verificar uma análise de dispersão, utilizando Boxplot, das variáveis já normalizadas.

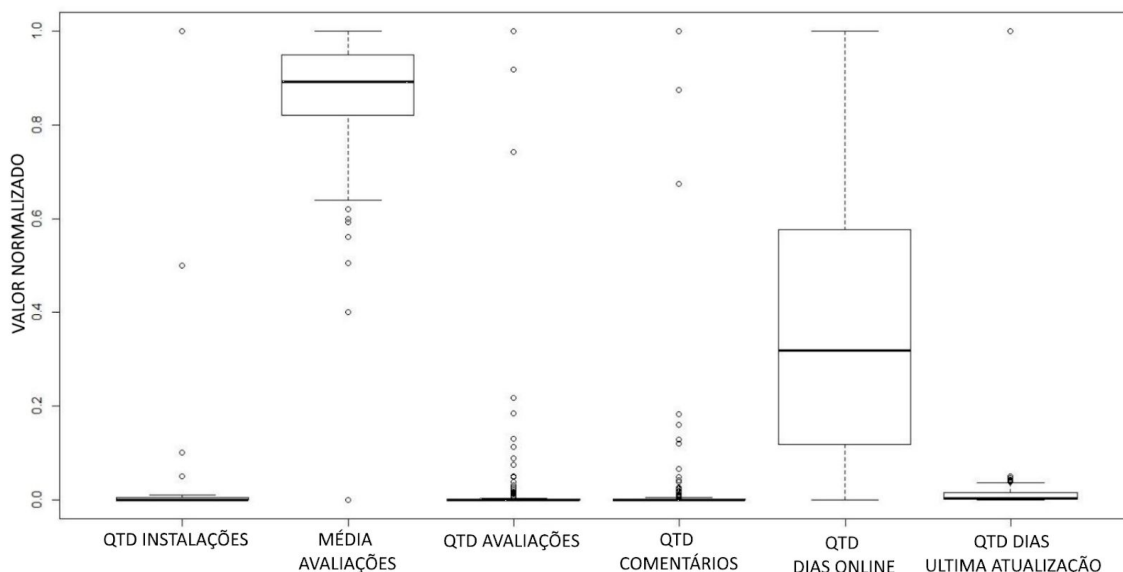
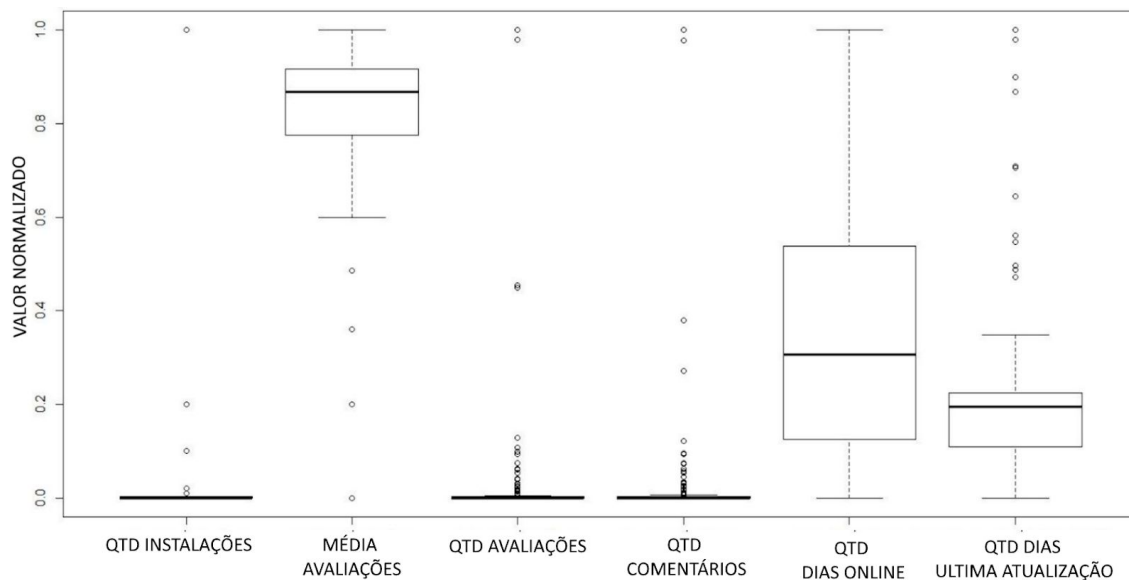


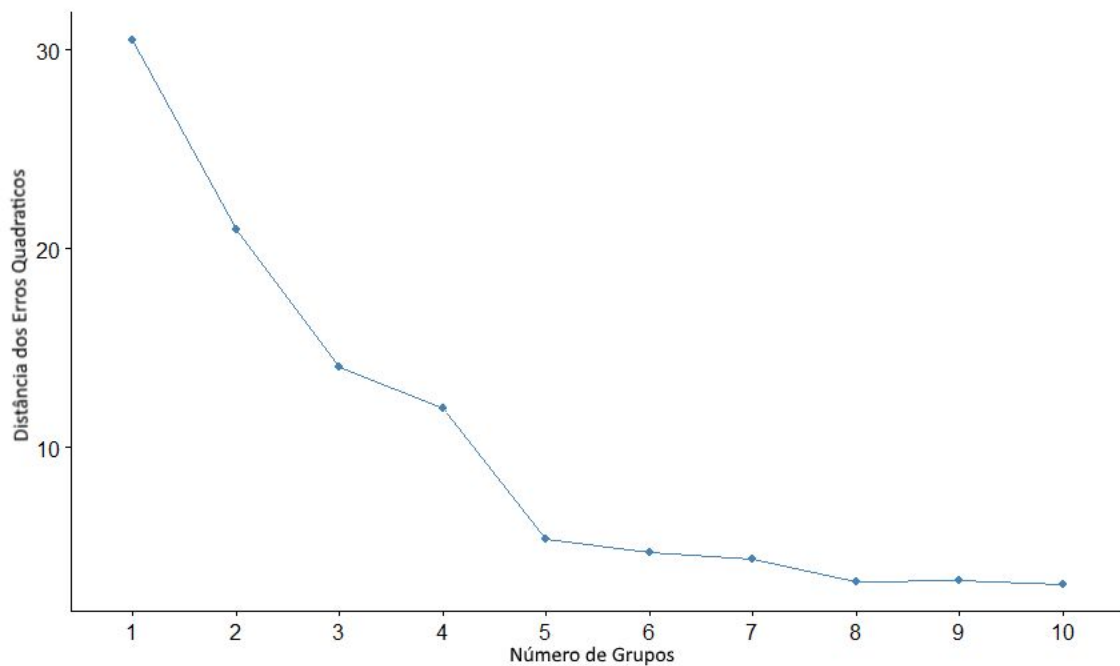
Figura 1. Boxplot mostrando a distribuição das métricas normalizadas dos aplicativos de Física.



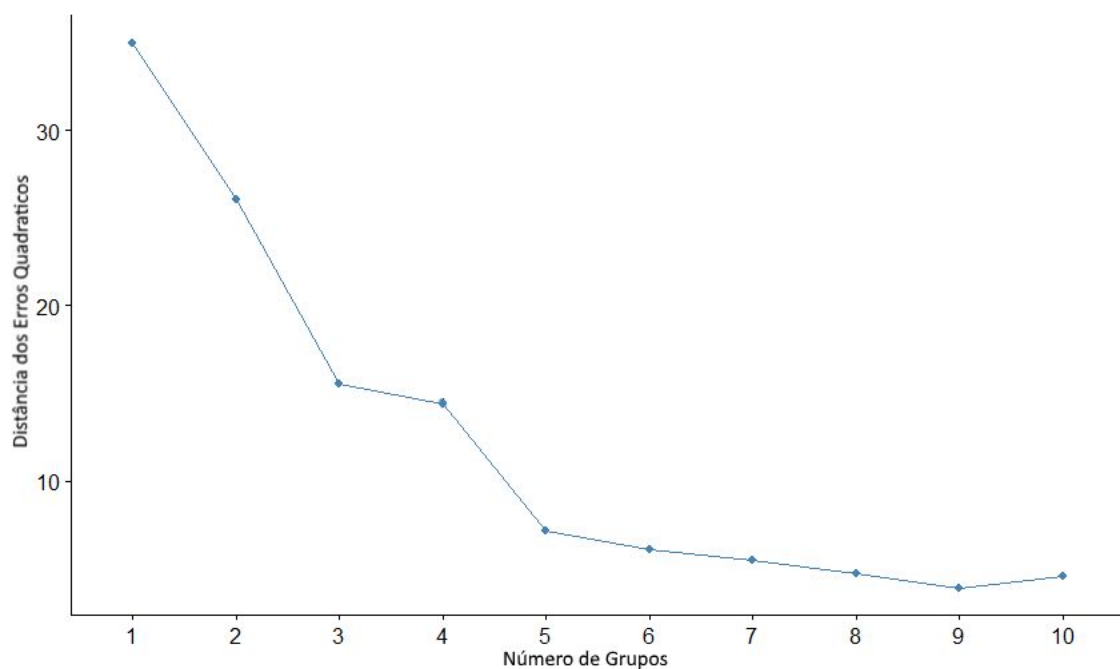
**Figura 2. Boxplot mostrando a distribuição das métricas normalizadas dos aplicativos de Matemática.**

Após a normalização é possível ver a dispersão das variáveis de forma clara. É possível observar que as variáveis dos aplicativos de ambas as matérias (Matemática e Física), se comportam de forma quase semelhante. Percebemos que para as métricas: quantidade de instalações, quantidade de avaliações e quantidade de comentários, o primeiro, segundo e terceiro quartil da distribuição das variáveis, encontram-se concentrados próximos de 0. Para as médias de avaliações, os aplicativos de ambas as matérias possuem seus quartis acima de 0,8 e com um limite inferior em ou acima 0,6. Para a quantidade de dias online, os aplicativos de ambas as matérias possuem uma maior variação, mas com todos os quartis da distribuição dos dados abaixo de 0,6.

Para realizarmos o agrupamento dos dados normalizados, utilizamos o algoritmo de Elbow para encontrar o melhor número de grupos para as variáveis. O resultado da qualidade dos agrupamentos, quando o número de grupos varia entre 1 e 10 pode ser visto nas Figuras 3 e Figura 4.



**Figura 3. Análise do agrupamento *K-Means* para os dados dos aplicativos de Física utilizando algoritmo de *Elbow*.**



**Figura 4. Análise do agrupamento *K-Means* para os dados dos aplicativos de Matemática utilizando algoritmo de *Elbow*.**

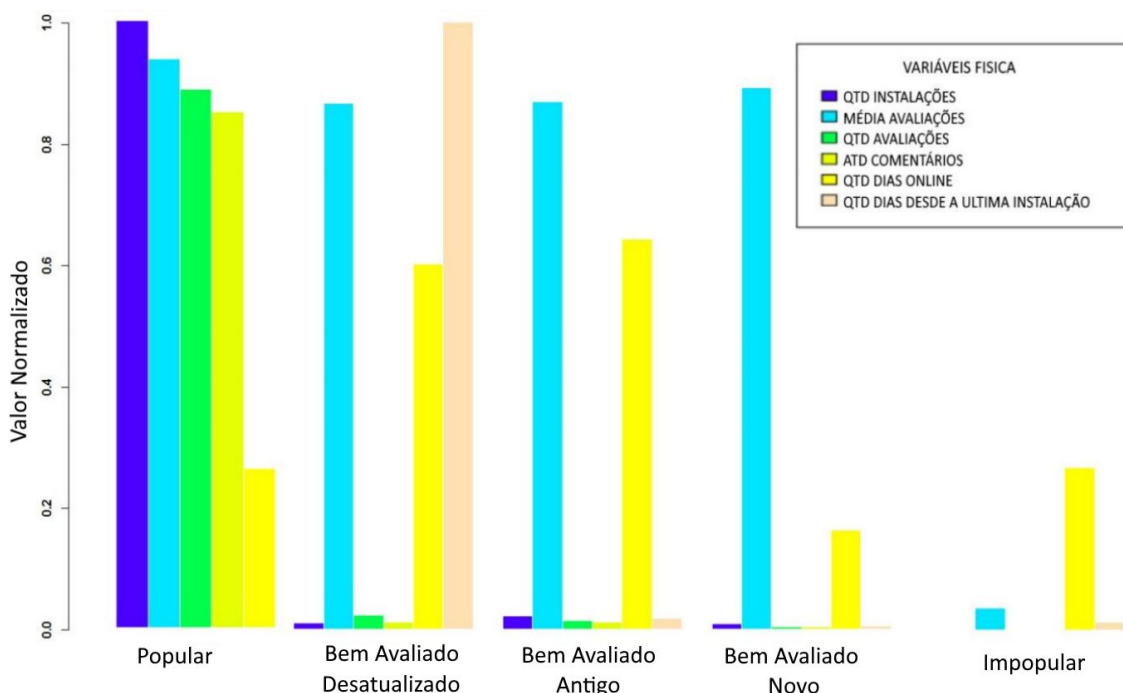
Nas Figura 3 e Figura 4 podemos verificar o somatório da variância dos dados em relação ao número de grupos, é possível verificar a partir de que ponto o aumento do número de grupos não há uma queda considerável. A partir de cinco grupos, tanto para os dados de aplicativos de Matemática, quanto para os aplicativos de Física, as

distâncias dos erros quadráticos praticamente se estabiliza. Após a análise do gráfico pode-se concluir claramente que o número ideal de grupos para os dados dos aplicativos de ambas as matérias é cinco.

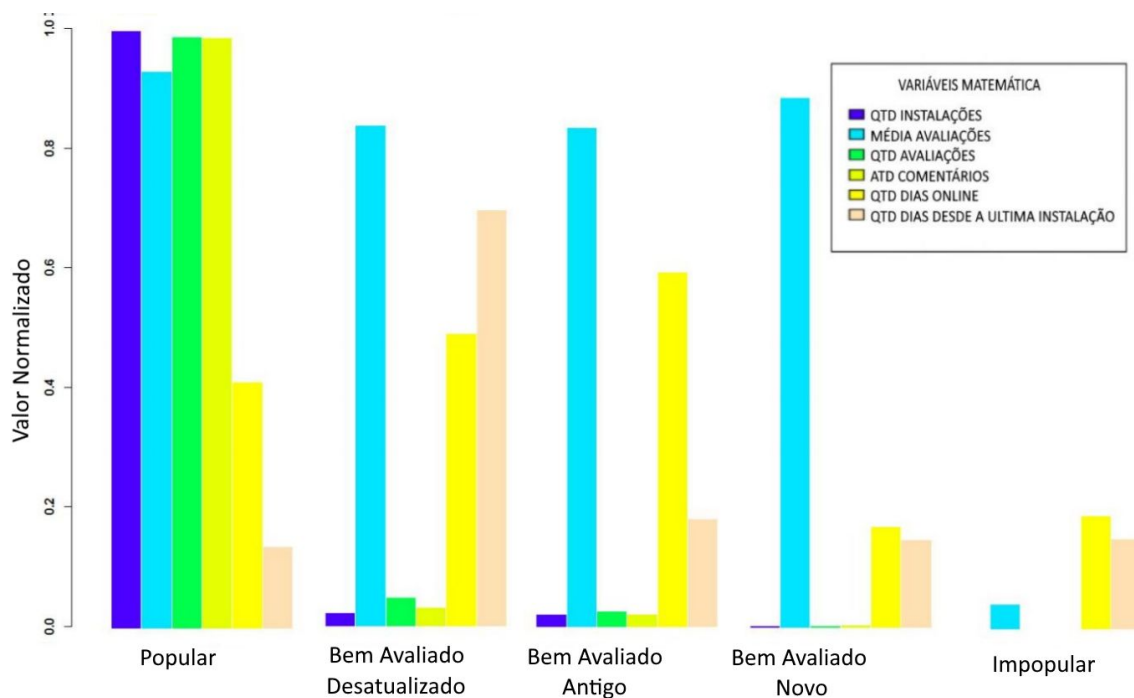
## 5.2. Perfis

Para entendermos os diferentes grupos descobertos pelo algoritmo, foram analisados, três itens importantes. Primeiro, os centróides que representam cada grupo. Segundo, como as diferentes métricas agrupadas explicam o grupo. Terceiro, a opinião de professores e alunos do ensino médio sobre aplicativos de cada grupo.

Nesta análise, estabelecemos nomes para os grupos que representa a principal característica de cada grupo. Estes grupos representam diferentes tipos de aplicativos, nomeados como: Populares; Impopular; Bem Avaliado Desatualizado; Bem Avaliado Antigo; Bem Avaliado Novo. Nas Figura 5 e Figura 6, são mostrados os centróides que representam cada um dos perfis e como eles se diferem. Em cada imagem, o eixo horizontal apresenta os perfis de aplicativos e cada barra representa uma métrica de avaliação, o eixo vertical indica como os perfis pontuam em cada uma das métricas.



**Figura 5. Medidas de cada métrica em cada perfil dos aplicativos de Física. Os perfis são representados pelos centróides de cada grupo gerado pelo K-Means.**



**Figura 6. Medidas de cada métrica em cada perfil dos aplicativos de Matemática. Os perfis são representados pelos centróides de cada grupo gerado pelo K-Means.**

**Perfil Popular.** Este Perfil se caracteriza pelo alto número de instalações (Mais de 10 milhões de instalações para aplicativos de Física e aplicativos de Matemática) , com alta média de avaliações (4,69 para aplicativos de Física e 4,72 para aplicativos de Matemática) e alto número de comentários (Mais de 406 mil comentários para aplicativos de Física e 472 mil comentários para aplicativos de Matemática), além de estar a um tempo mediano (839 dias para aplicativos de Física e 1010 dias para aplicativos de Matemática) no *storage* da Google e possuir um nível de atualização relativamente alto (48 dias para aplicativos de Física e 27 dias para aplicativos de Matemática). Estes aplicativos tem uma boa consistência em relação às opiniões dos usuários e uma taxa de atualização por parte de seus desenvolvedores. Para os aplicativos de ambas as matérias há uma baixa porcentagem de aplicações para este perfil: 1% para os aplicativos de Matemática e 6,4% para os aplicativos de Física (Tabela 2).

**Perfil Impopular.** Este Perfil se caracteriza pelo baixo número de instalações (16 instalações para aplicativos de Física e 45 instalações para aplicativos de Matemática), baixa média de avaliações (0,18 para aplicativos de Física e 0,25 para aplicativos de Matemática), nenhum ou quase nenhum comentario (Média de menos de 2 comentários para aplicativos de Física e aplicativos de Matemática). Estes aplicativos além de estarem há pouco tempo no mercado (854 dias para aplicativos de Física e 832 dias para aplicativos de Matemática), pode-se verificar que estes aplicativos não despertou interesse de usuários. Para os aplicativos de ambas as matérias há uma baixa porcentagem de aplicações para este perfil: 9,5% para os aplicativos de Matemática e 0,5% para os aplicativos de Física (Tabela 2).

Perfil **Bem Avaliado Desatualizado**. Este perfil se caracteriza pela alta média de avaliações entre os usuários (4,33 para aplicativos de Física e 4,52 para aplicativos de Matemática), porém possui uma baixa taxa de atualização por parte dos desenvolvedores (2200 dias para aplicativos de Física e 1912 dias para aplicativos de Matemática), além de estarem há um tempo considerado (2200 dias para aplicativos de Física e 1912 dias para aplicativos de Matemática) no *storage* da Google. Para os aplicativos deste perfil, há uma porcentagem relevante para os aplicativos de Física (38,4%) e uma baixa porcentagem de aplicações de Matemática (6,3%) (Tabela 2).

Perfil **Bem Avaliado Antigo**. Este perfil se caracteriza pela alta média de avaliações entre os usuários (4,34 para aplicativos de Física e 4,33 para aplicativos de Matemática), possui uma boa taxa de atualização por parte dos desenvolvedores (793 dias para aplicativos de Física e 685 dias para aplicativos de Matemática), além de estarem a mais tempo (2207 dias para aplicativos de Física e 2236 dias para aplicativos de Matemática) no *storage* da Google. São representadas por aplicativos que tiveram uma boa avaliação pelos usuários que a usaram, porém não se popularizaram de forma considerável. Para os aplicativos deste perfil, há uma porcentagem relevante para os aplicativos de Matemática (35,9%) e uma baixa porcentagem de aplicações de Física (1,7%) (Tabela 2).

Perfil **Bem Avaliado Novo**. Este perfil se caracteriza pela alta média de avaliações entre os usuários (4,46 para aplicativos de Física e 4,42 para aplicativos de Matemática), possui uma boa taxa de atualização por parte dos desenvolvedores (212 dias para aplicativos de Física e 195 dias aplicativos de Matemática) e estão a menos tempo no *storage* da Google (529 dias para aplicativos de Física e 489 dias aplicativos de Matemática). Este perfil é composto por aplicativos que foram postadas recentemente e já tiveram uma boa avaliação dos usuários. A maior parte dos aplicativos de ambas as matérias estão presentes neste perfil: 47% dos aplicativos de Matemática e 52,9% dos aplicativos de Física (Tabela 2).

A Tabela 2 apresenta como cada um dos perfis se diferem em termos de números de aplicativos e como cada grupo se difere em total de aplicativos por perfil, conforme descrito nos perfis acima.

**Tabela 2. Número de aplicativos por perfil.**

Perfis	Matemática	Física
	Nº. de aplicativos	Nº. de aplicativos
Popular	2 (1%)	11 (6,4%)
Impopular	18 (9,5%)	1 (0,5%)
Bem Avaliado Desatualizado	12 (6,3%)	66 (38,4%)
Bem Avaliado Antigo	68 (35,9%)	3 (1,7%)

Bem Avaliado Novo	89 (47%)	91 (52,9%)
-------------------	----------	------------

### 5.3. Opiniões dos Usuários

Os usuários dos aplicativos participantes do teste eram compostos por estudantes e professores de cursos de pré-vestibular e estudantes do primeiro e segundo período de cursos de graduação. Os participantes responderam a um questionário, onde foi coletado a idade, nível de escolaridade, se eles usam aplicativos para estudar ou lecionar, se já conheciam alguma dos aplicativos propostas para o teste e um questionário de perguntas da escala SUS para cada aplicativo. Foi disponibilizado um *tablet* e celulares com os aplicativos previamente instaladas, para que os participantes pudessem realizar o uso das mesmas e pudessem responder ao questionário.

Quanto a quantidade e perfil dos participantes dos formulários aplicados, para os aplicativos da matéria de Matemática houve: um total de 17 participantes, com uma idade média de 20 anos. O participante com maior idade tem 32 anos e o de menor idade tem 18 anos. Quanto ao nível de escolaridade, 41,2% possuem Superior Incompleto, 47,1% possuem Ensino Médio Completo e 11,8% possuem Superior Completo. Sobre o conhecimento de algum dos aplicativos, 9,4% dos participantes não conhecem nenhum dos aplicativos apresentadas, 41,1% dos participantes conhecem 1 dos aplicativos apresentadas, 23,5% dos participantes conhecem 2 dos aplicativos apresentadas, 5,8% dos participantes conhecem 3 dos aplicativos apresentadas. Quanto a utilizar aplicativos de celular para estudar e/ou lecionar, 88,2% dizem utilizar aplicativos de celular para estudar e/ou lecionar e 11,8% dizem não utilizar.

Quanto a quantidade e perfil dos participantes dos formulários aplicados, para os aplicativos da matéria de Física houve: um total de 20 participantes, com uma idade média de 21,5 anos. O participante com maior idade tem 51 anos e o de menor idade tem 18 anos. Quanto ao nível de escolaridade, 35% possuem Superior Incompleto, 25% possuem Ensino Médio Completo e 35% possuem Superior Completo e 5% possuem Ensino Médio Incompleto. Sobre conhecer algum dos aplicativos, 20% dos participantes não conhecem nenhum dos aplicativos apresentadas, 25% dos participantes conhecem 1 dos aplicativos apresentadas, 30% dos participantes conhecem 2 dos aplicativos apresentadas, 25% dos participantes conhecem 3 dos aplicativos apresentadas. Quanto a utilizar aplicativos de celular para estudar e/ou lecionar, 65% dizem utilizar aplicativos de celular para estudar e/ou lecionar e 35% dizem não utilizar.

Sobre os resultados obtidos sobre o *rating* de cada aplicativo, foi realizado a aplicação da SUS para realizar a eleição do aplicativo com melhor avaliação dos usuários. A Tabela 3 apresenta os aplicativos, seu *rating* e a que matéria pertence e seu perfil representante.

**Tabela 3. Rating dos aplicativos representantes por perfil e matéria.**

<b>Perfil</b>	<b>Aplicativo Matemática</b>	<b>Rating</b>	<b>Aplicativo Física</b>	<b>Rating</b>
Popular	Matemática	72,33	Circuit Builder	44,15
Impopular	Math Studio	72,72	Simulado ENEM 2018	73,57
Bem Avaliado Desatualizado	MyScript Calculator	70,72	Calculadora de Física	64,18
Bem Avaliado Antigo	Frações Matemática	61,71	Electric Circuit Simulation	42,72
Bem Avaliado Novo	Math Calculator-Solve Math Problems by Camera	73,21	Brainly - Estude com a Gente	77,1

Após aplicar o cálculo do rating da SUS sobre os aplicativos, pode-se observar que o aplicativo mais bem avaliado para a matéria de Matemática foi o aplicativo: Math Calculator-Solve Math Problems by Camera, com um rating de 73,21. Seguindo o perfil de agrupamento: Bem Avaliado Novo. Para a matéria de Física foi o aplicativo: Brainly - Estude com a Gente, com um rating de 77,1. Seguindo o perfil de agrupamento: Bem Avaliado Novo.

## **6. Conclusões**

Neste estudo foram abordadas as seguintes questões: Primeiro, quais informações pode-se obter sobre aplicativos disponíveis no Storage da Google Play e como podemos utilizá-las para eleger perfis de aplicativos. Segundo, quais os diferentes padrões podem ser usados para eleger perfis de aplicativos. Por fim, como os perfis de aplicativos podem ser utilizados para identificar quais modelos de aplicativos tem melhor aceitação por parte dos usuários.

Nossos resultados mostram que os aplicativos abordados neste trabalho compartilham várias similaridade e diferenças em termos de Atenção e Avaliação dos usuário e tempo de disponibilidade no Storage. Ele foram nomeados como: Populares; Impopular; Bem Avaliado Desatualizado; Bem Avaliado Antigo; Bem Avaliado Novo. A respeito da distribuição dos aplicativos pelos perfis, a maior porcentagem de aplicativos encontra-se no perfil Bem Avaliado Novo, para ambos os aplicativos das matérias de Física e Matemática. Sobre o resultado da pesquisa de opinião dos usuários desses aplicativos, pode-se perceber que a maior aceitação dos aplicativos está sobre aplicativos que se encaixam no perfil de maior porcentagem de aplicativos, para ambos os grupos de matérias de aplicativos.



A análise dos perfis de aplicativos, baseada nos dados dos aplicativos disponíveis no storage do Google Play rendeu um bom volume de informações referentes aos modelos abordados pelos desenvolvedores e o modelo de melhor aceitação para os usuários alvos desses aplicativos. Foco foi dado no levantamento de perfis, considerando a resposta dos usuários a estes aplicativos e seu tempo de “vida” dentro do storage Google Play, assim, é possível gerar conclusões sobre os modelos que possuem uma aceitação imediata pelos usuários e os que têm uma grande rejeição.

## 6.1 Limitações

O estudo realizado apresentou limitações quanto a amostra escolhida como objeto de estudo. O número de dados foi limitado pela escolha de aplicativos para as matérias de Matemática e Física, apenas disponíveis na língua Portuguesa. Consequentemente, não é possível estender os resultados obtidos para todos os aplicativos disponíveis para o auxílio no processo de aprendizagem.

O estudo foi conduzido entre alunos de instituições pré-vestibular e do primeiro e segundo período do ensino superior da UFMG, logo, não é possível estender esta avaliação para todos os usuários das demais classes sociais. Entretanto este fato não mostra que os resultados sejam inválidos, mas sim que não é possível ter uma conclusão generalista.

## 6.2 Trabalhos Futuros

Como trabalhos futuros, em particular, seria relevante conduzir estudo que aborde: i) o modelo de desenvolvimento de um aplicativo, baseando-se no perfil de melhor aceitação; ii) propor um aplicativo modelo que se encaixa nas melhores respostas obtidas sobre o perfil eleito com melhor *rating* de avaliação, pelos usuários, utilizada neste trabalho; iii) analisar o ciclo de vida de um aplicativo entre os perfis encontrados. Esses estudos propostos podem revelar conhecimento novo acerca de informações que rodeiam o escopo abordado por este estudo.

## Referências Bibliográficas

- Alencar, G. A., Pessoa, M. dos S., Santos, A. K. de F. S., Carvalho, S. R. R. de, Lima, H. A. de B.. WhatsApp como ferramenta de apoio ao ensino. In: Latin American Conference of Learning Objects (LACLO), 2015. Maceió. Anais do LACLO 2015. p. 787-795
- Brooke, J. (1996). “*SUS-A quick and dirty usability scale. Usability evaluation in industry*”, 189(194), 4-7.
- Cruz, J. R., Ramos, R., Rodrigues, W. B. (2015). “*Guideline para desenvolvimento de aplicativos educacionais mobile*”. Workshop de Computação da Faccamp. doi: ISSN (pp. 2447-4703).
- Doni, M.V. "Análise de *cluster*: métodos hierárquicos e de particionamento." Universidade Presbiteriana Mackenzie (2004).
- Filho, N.F., Ellen, F.B. "Estudo e Definição de um conjunto de características e requisitos para ambientes de aprendizagem móvel." *Brazilian Symposium on*

- Computers in Education* (Simpósio Brasileiro de Informática na Educação-SBIE). Vol. 23. No. 1. (2012).
- Gomes, A.S., Filho, J.A.C, Gitirana, V., Spinillo, A. "Avaliação de software educativo para o ensino de matemática." WIE 2002 Workshop Brasileiro de Informática Educativa. Florianópolis: SBC. (2002).
- Hair, J. F. Jr.; Anderson, R. E.; Tatham, R. L.; Black, W. C. "Análise Multivariada de Dados". 5 ed. Porto Alegre: Bookman, 2005.
- Jeff, S. (2009). "Measuring Usability With The System Usability Scale (SUS)." Disponível em: <<https://measuringu.com/sus/>>. Acesso em: 31/10/2018.
- Laine, T. H., Sedano, C. A., Joy, M., Sutinen, E. "Critical Factors for Technology Integration in Game-Based Pervasive Learning Spaces". *IEEE Transactions on Learning Technologies*, Vol. 3(4) (2010).
- Lupion T.P., Paulo R.A., Irala E.A.F. "Grupos de consenso: uma proposta de aprendizagem colaborativa para o processo de ensino-aprendizagem." *Revista diálogo educacional* 4.13 (2004).
- Moore, P., Hu, B., Wan, J. (2008) "A Context Ontology for Pervasive Mobile Computing", *The Computer Journal*, Vol. 53, p. 191-207.
- Morais, C. M. M. D. "Proposição de indicadores para investigação de licitações por meio de técnicas de reconhecimento de padrões estatísticos e mineração de dados". 2016. xx, 126 f., il. Universidade de Brasília, Brasília, 2016.
- Moscato, P., Fernando J.V.Z. "Uma Visão Geral de Clusterização de Dados." DCA/FEEC/Unicamp (2002).
- Nascimento, E.G.A. "Avaliação do uso do software GeoGebra no ensino de geometria: reflexão da prática na escola." XII Encontro de Pós-Graduação e Pesquisa da Unifor, ISSN 8457 (1808): (2012).
- Oliveira, E. D. S. (2014) "Experiência de uso do WhatsApp como Ambiente Virtual de Aprendizagem em um curso a distância", In: Anais do XX Workshop de Informática na Escola, pages 179-183.
- Souza, A.A.N.; Schneider, H.N. "Aprendizagem colaborativa nas redes sociais: novos olhares sobre a prática pedagógica. " In: II Congresso Internacional TIC e Educação. Lisboa, Pt. 2012. p. 2100-2111.
- Valente, J.A. "Análise dos diferentes tipos de softwares usados na educação. " O computador na sociedade do conhecimento, p. 71-86, 1999.
- Vieira, F.M.S. "Avaliação de software educativo: reflexões para uma análise criteriosa. " Campinas: EDUTECCNET, 1999.