

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**EDMILSON HOMMA JUNIOR**

**ANÁLISE DO PERFIL DE ESTUDANTES DE GRADUAÇÃO BASEADA NAS  
SUAS EXPERIÊNCIAS E PERCEPÇÕES SOBRE A UNIVERSIDADE: UMA  
ABORDAGEM POR TÉCNICAS DE CLUSTERIZAÇÃO**

**LONDRINA**

**2022**

**EDMILSON HOMMA JUNIOR**

**ANÁLISE DO PERFIL DE ESTUDANTES DE GRADUAÇÃO BASEADA NAS  
SUAS EXPERIÊNCIAS E PERCEPÇÕES SOBRE A UNIVERSIDADE: UMA  
ABORDAGEM POR TÉCNICAS DE CLUSTERIZAÇÃO**

**Profile analysis of undergraduate students based on their experiences and  
perceptions about university: an approach by clustering techniques**

Trabalho de conclusão de curso de graduação  
apresentada como requisito para obtenção do título de  
Bacharel em Engenharia de Produção da Universidade  
Tecnológica Federal do Paraná (UTFPR).  
Orientador: Bruno Samways dos Santos.

**LONDRINA**

**2022**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**EDMILSON HOMMA JUNIOR**

**ANÁLISE DO PERFIL DE ESTUDANTES DE GRADUAÇÃO BASEADA NAS  
SUAS EXPERIÊNCIAS E PERCEPÇÕES SOBRE A UNIVERSIDADE: UMA  
ABORDAGEM POR TÉCNICAS DE CLUSTERIZAÇÃO**

Trabalho de Conclusão de Curso de Graduação  
apresentado como requisito para obtenção do título de  
Bacharel em Engenharia de Produção da Universidade  
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 21 de novembro de 2022

---

Bruno Samways dos Santos  
Doutorado em Engenharia de Produção e Sistemas  
Universidade Tecnológica Federal do Paraná

---

Pedro Rochavetz de Lara Andrade  
Doutorado em Engenharia de Produção  
Universidade Tecnológica Federal do Paraná

---

Rafael Henrique Palma Lima  
Doutorado em Engenharia de Produção  
Universidade Tecnológica Federal do Paraná

**LONDRINA**

**2022**

## **AGRADECIMENTOS**

Agradeço, primeiramente, à minha família e aos meus amigos que estão comigo independente de distâncias e condições. São o meu norte e a força que me faz seguir, sempre.

Agradeço, também, ao meu orientador Prof. Dr. Bruno Samways dos Santos pelo tempo e esforço dedicado não somente a este trabalho, mas também aos diversos outros em que tivemos a oportunidade de atuar juntos. Têm sido um aprendizado e tanto.

À UTFPR, ficam meus agradecimentos pelo apoio e suporte ao longo da graduação e à Fundação Araucária pelo financiamento para a execução do projeto de iniciação científica (plano de trabalho PIC2021040000224) do edital 2021 PROPPG 02 - PIBIC/PIBIC-Af.

Por fim, à todos que de alguma forma participaram da minha jornada até aqui e que agora fazem parte do que sou e do que serei.

## RESUMO

Com o desafio de entender perfis de alunos dentro do contexto universitário, a mineração de dados têm sido utilizada cada vez mais como meio para gerar informações úteis aos tomadores de decisão. Nesse sentido, o presente trabalho buscou aplicar técnicas de clusterização sobre dados obtidos através de um questionário aplicado nos estudantes de uma universidade pública. Para tal, foi elaborado um questionário com 39 questões quantitativas e qualitativas, contendo perguntas sobre o aluno e sobre suas percepções em relação à universidade, no qual avaliam a satisfação e fidelidade, de acordo com a metodologia CSAT (*Customer Satisfaction*) e NPS (*Net Promoter Score*), respectivamente. Através dele, foram aplicadas três técnicas de agrupamento, sendo elas: *K-Means*; *Fuzzy C-Means*; e Método Hierárquico. O conjunto de dados foi dividido em dois, um para questões em que o aluno avalia a experiência universitária e outro de cunho pessoal. Os resultados foram avaliados pelo método da silhueta, o qual retornou os melhores valores para o *K-Means* em ambos os casos, com notas 0.181 (com dois *clusters*) e 0.498 (com três *clusters*), respectivamente. Porém, para o caso de três *clusters*, o *Fuzzy C-Means* conseguiu atingir um balanceamento de instâncias melhor entre os grupos, trazendo a oportunidade de analisar de forma paralela ao *K-Means*. Com isso, foram analisadas as diferenças entre os grupos, resultando em pelo menos cinco perfis distintos. Dois deles se diferenciaram, principalmente, por serem “promotores” ou “neutro/detratores” em relação à universidade, sendo que estes se caracterizaram por enfrentar maiores responsabilidades e uma situação de vida mais estressante. Os demais perfis se diferenciaram, em especial, pela relação que possui com a família, por frequentar psicólogo ou psiquiatra, e pela quantidade de situações de estresse, ansiedade ou falta de motivação que enfrentam durante a graduação.

Palavras-chave: clusterização; universidade; EDM; NPS.

## ABSTRACT

With the challenge of understanding student profiles within the university context, data mining has been increasingly used as a means to generate useful information for decision makers. In this sense, the present work sought to apply clustering techniques on data obtained through a questionnaire applied to students at a public university. To this end, a questionnaire was prepared with 39 quantitative and qualitative questions, containing questions about the student and their perceptions regarding the university, in which they assess satisfaction and loyalty, according to the CSAT (Customer Satisfaction) and NPS (Net Promoter Score), respectively. Through it, three clustering techniques were applied, namely: K-Means; Fuzzy C-Means; and Hierarchical Method. The data set was divided in two, one for questions in which the student evaluates the university experience and another for a personal question. The results were evaluated by the silhouette method, which returned the best values for K-Means in both cases, with scores of 0.181 (with two clusters) and 0.498 (with three clusters), respectively. However, for the case of three clusters, Fuzzy C-Means managed to achieve a better balance of instances between the groups, bringing the opportunity to analyze in parallel with K-Means. Thus, the differences between the groups were analyzed, resulting in at least five distinct profiles. Two of them differed, mainly, for being “promoters” or “neutral/detractors” in relation to the university, and they were characterized by facing greater responsibilities and a more stressful life situation. The other profiles differed, in particular, by the relationship they have with the family, by attending a psychologist or psychiatrist, and by the amount of situations of stress, anxiety or lack of motivation they face during graduation.

Keywords: clustering; university; EDM; NPS.

## LISTA DE ILUSTRAÇÕES

Figura 1 - As etapas do processo de KDD .....	19
Figura 2 - As etapas da pesquisa.....	25
Figura 3 - Boxplot comparativo da pergunta de NPS em relação ao curso. ....	33
Figura 4 - Boxplot comparativo da pergunta de NPS em relação à universidade. .....	34
Figura 5 - Boxplot comparativo para a pergunta NPS em relação ao curso( <i>K-Means</i> ).....	36
Figura 6 - Boxplot comparativo para a pergunta NPS em relação à universidade ( <i>Fuzzy C-Means</i> ).....	37
Quadro 1 - Ideias de ações extraídas do estudo. ....	47

## LISTA DE TABELAS

Tabela 1 - Perguntas disponíveis no questionário elaborado.....	25
Tabela 2 - Variáveis que foram selecionadas para a aplicação.....	27
Tabela 3 - Comportamento das respostas qualitativas.....	29
Tabela 4 - Comportamento das respostas quantitativas .....	30
Tabela 5 - CSAT das perguntas de satisfação.....	30
Tabela 6 - CSAT por dimensão.....	31
Tabela 7 - NPS em relação ao curso e universidade.....	31
Tabela 8 - Resultados pelo método da silhueta.....	32
Tabela 9 - Distribuição das instâncias dentro dos <i>clusters</i> , para cada técnica.....	32
Tabela 10 - Médias das questões avaliativas por <i>cluster</i> . .....	34
Tabela 11 - Resultados pelo método da silhueta.....	35
Tabela 12 - Distribuição das instâncias dentro dos <i>clusters</i> , para cada técnica. .....	36
Tabela 13 - Médias das questões avaliativas por <i>cluster</i> ( <i>K-Means</i> ). .....	37
Tabela 14 - Médias das questões avaliativas por <i>cluster</i> ( <i>Fuzzy C-Means</i> ). .....	38
Tabela 15 - Principais destaques por perfil (intracluster).....	38
Tabela 16 - Principais destaques por perfil - <i>K-Means</i> (intracluster).....	41
Tabela 17 - Principais destaques por perfil – <i>Fuzzy C-Means</i> (intracluster).....	44



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>13</b>
<b>1.1</b>	<b>Objetivo geral.....</b>	<b>15</b>
<b>1.2</b>	<b>Objetivos específicos.....</b>	<b>15</b>
<b>1.3</b>	<b>Justificativa.....</b>	<b>15</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>17</b>
<b>2.1</b>	<b>Experiência de aprendizado .....</b>	<b>17</b>
<b>2.2</b>	<b>Métricas de experiências .....</b>	<b>17</b>
<b>2.3</b>	<b>Mineração de dados .....</b>	<b>19</b>
<b>2.4</b>	<b>Tipos de aprendizado de máquina.....</b>	<b>19</b>
<b>2.5</b>	<b>Técnicas de clusterização .....</b>	<b>21</b>
<b>2.5.1</b>	<b>Métodos para clusterização.....</b>	<b>21</b>
<b>2.5.2</b>	<b>K-Means .....</b>	<b>23</b>
<b>2.5.3</b>	<b>Fuzzy C-Means .....</b>	<b>23</b>
<b>2.6</b>	<b>Método da Silhueta.....</b>	<b>24</b>
<b>3</b>	<b>METODOLOGIA .....</b>	<b>25</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÕES .....</b>	<b>29</b>
<b>4.1</b>	<b>Análise descritiva das respostas .....</b>	<b>29</b>
<b>4.2</b>	<b>Análise das clusterizações .....</b>	<b>32</b>
<b>4.2.1</b>	<b>Variáveis avaliativas .....</b>	<b>32</b>
<b>4.2.2</b>	<b>Variáveis pessoais.....</b>	<b>35</b>
<b>4.3</b>	<b>Análise dos perfis encontrados .....</b>	<b>38</b>
<b>4.3.1</b>	<b>Com dois grupos .....</b>	<b>38</b>
<b>4.3.2</b>	<b>Com três grupos .....</b>	<b>41</b>
<b>4.4</b>	<b>Implicações práticas .....</b>	<b>46</b>
<b>5</b>	<b>CONCLUSÃO .....</b>	<b>51</b>
	<b>REFERÊNCIAS.....</b>	<b>53</b>
	<b>APÊNDICE A - Parecer do Comitê de Ética em Pesquisa (CEP) .....</b>	<b>56</b>
	<b>APÊNDICE B - Questionário de pesquisa .....</b>	<b>59</b>

## 1 INTRODUÇÃO

No Brasil as políticas educacionais ganharam força após a década de 90, principalmente na educação superior, devido ao contexto de expansão do nível de educação (SANTOS, 2017). Porém, segundo Maiochi (2012), além de enfrentar um desaceleramento nessa expansão, as Instituições de Ensino Superior (IES) têm enfrentado altos índices de evasão de alunos durante a graduação, levando-as a buscar uma gestão voltada para clientes, processos e resultados.

Segundo o Censo da Educação Superior, dos ingressantes em IES em 2010, apenas 40% concluíram o seu curso de ingresso ao final de 10 anos de acompanhamento de sua trajetória em 2020, ou seja. A evasão de alunos pode ser tanto por causas internas, sendo referentes a aspectos didático-pedagógicos e à infraestrutura, quanto por causa externas, ligadas a aspectos sócio-político-econômicos, e também as causas relacionadas aos alunos que se referem à vocação e a outros problemas de ordem mais pessoal (BIAZUS, 2004). Todas estas variáveis podem ter características diferentes dependendo do perfil de diferentes grupos de alunos.

Dalen (2010) destaca, ainda, que uma taxa de evasão elevada resulta em prejuízos para a instituição, tanto financeiros quanto de reputação. Para agir sobre este problema, é necessário identificar e entender quais são os estudantes que têm um maior risco de desistência e seus motivos por trás do atrito com a universidade. Nesse sentido, a gestão universitária se torna um desafio para oferecer uma experiência melhor aos alunos e aumentar a sua retenção ao longo do curso.

Novas tecnologias têm surgido para apoiar tomadores de decisão a gerenciar melhor seus negócios, e, sob essa ótica, o uso de mineração de dados na educação, ou *Educational Data Mining* (EDM), é um campo de estudo que vem se fortalecendo cada vez mais (ROMERO, 2007), sendo possível citar trabalhos como o de Bonde (2018) e Dol (2021). O EDM busca analisar bases de dados com o propósito de resolver problemas no âmbito educacional, utilizando-se de conhecimentos de diversas áreas, como a de ciências computacionais, e explorando características e comportamentos do estudante.

Em relação à mineração de dados (MD) Witten e Frank (2005) explica que é a extração de informações implícitas, sem conhecimento prévio e potencialmente úteis dos dados. E, para isso, constrói-se programas computacionais que vasculhem

bancos de dados automaticamente, procurando padrões e regularidades. Assim, é possível utilizá-la nas mais diversas áreas, bastando apenas que existam dados passíveis de manipulação, como é no caso da educação. Em conjunto com este processo, é possível, ainda, aplicar técnicas de aprendizado de máquina, ou *machine learning*, cujos algoritmos podem ser aplicados em quatro tarefas principais da mineração de dados: Clusterização, Classificação, Regressão e Associação (AGGARWAL, 2015).

Sob essa ótica, estudos vêm demonstrando resultados expressivos, como, por exemplo, identificando, por meio de algoritmos de classificação, alunos que possuem uma aprendizagem mais lenta em relação à média geral, para que, assim, possam ser tomadas as devidas providências buscando diminuir esta diferença entre eles (KAUR, 2015). Ou ainda, prevendo o desempenho final de um aluno de engenharia ao final do curso de acordo com seus resultados nos três primeiros anos (ADEKITAN; SALAU, 2019).

Além desses, o trabalho de Dalen (2010) realizou a predição da permanência de calouros com um conjunto com mais de 16 mil estudantes de uma universidade pública americana e concluiu que, com uma base de dados da própria universidade, é possível identificar com acurácia alunos em risco de abandono e ainda auxiliar na alocação de recursos para aumentar a sua retenção. Já com uma abordagem de clusterização, o estudo realizado por Campagni (2014) mostrou a correlação entre evasão de alunos com base em uma avaliação de diferentes aspectos do curso, realizada por alunos que estão perto de se graduar, além de concluir que em cursos bem avaliados os alunos possuem um desempenho melhor.

Dessa forma, uma vez que sejam disponibilizados dados sobre a universidade e/ou alunos, é possível utilizar a mineração de dados para extrair informações relevantes, tanto para o meio acadêmico como para tomadores de decisões, para que, assim, sejam estabelecidas estratégias específicas e direcionadas a determinados públicos de alunos.

A partir desta perspectiva, surgiu o questionamento principal deste trabalho: como identificar e entender grupos de alunos a partir de dados de comportamentos e percepções de experiências universitárias utilizando técnicas de *machine learning*?

## 1.1 Objetivo geral

Utilizar técnicas de clusterização para a análise de perfis de alunos formados com base nas suas experiências vividas e percepções sobre o contexto universitário, tanto da estrutura geral do campus quanto do corpo docente.

## 1.2 Objetivos específicos

Foram definidos os seguintes objetivos específicos:

- Elaborar o instrumento de coleta de dados sobre a percepção do aluno sobre a faculdade e experiência universitária;
- Coletar os dados dos alunos da unidade em análise por meio de questionário digital, com aplicação online;
- Aplicar técnicas de agrupamento nos dados obtidos;
- Analisar os resultados e perfis obtidos pelos métodos;
- Discutir as implicações práticas destes resultados.

## 1.3 Justificativa

A má experiência do aluno no contexto universitário pode gerar um desempenho ruim e uma percepção negativa do estudante sobre a instituição, resultando em altas taxas de evasão e má reputação da universidade. Portanto, deve-se buscar entender os motivos causadores dos atritos, para que seja possível, então, atuar sobre e melhorar o ambiente acadêmico. Porém, segundo Delen (2010), apesar de um estudo qualitativo gerar bons resultados por meio de testes e teorias, ela não fornece um instrumento necessário para prever, com algum nível de acurácia, as dificuldades do estudante durante a sua graduação. Nesse sentido, surge a oportunidade de utilizar de um estudo quantitativo através do uso de mineração de dados, a fim de entender os aspectos que geram tais experiências e relacioná-las com os perfis dos estudantes.

Meyer e Schwager (2007) afirmam que a satisfação do cliente é o resultado de uma série de experiências que ocorrem quando a lacuna entre a expectativa e as experiências oferecidas são eliminadas. Sob essa ótica, Kindlein & Schwaiger (2015) afirma que as instituições educacionais devem colocar estudantes empoderados no centro do processo educacional e adotar uma postura orientada para o cliente, podendo, assim, considerar a instituição como uma provedora de serviço educacional.

Uma análise que traz esta perspectiva é o estudo de Delen (2010), onde compara técnicas de machine learning para análise da gestão da evasão de alunos, e mostra que as variáveis mais impactantes da experiência do cliente-aluno foram: a interação social do estudante; a expectativa prévia do ingressante; e a condição financeira dos pais.

Assim, com variáveis semelhantes a estas, técnicas de agrupamento podem formar grupos distintos de alunos para identificar as principais diferenças entre eles e utilizá-las para direcionar ações que busquem a melhorias no serviço prestado e na experiência vivida pelo cliente-aluno enquanto graduando.

Portanto, esta pesquisa justifica-se pela contribuição para a ciência e o ambiente acadêmico, já que ela traz uma forma de analisar, através de dados, o comportamento humano e relacioná-la com as ações externas a ela. Por fim, este estudo poderá trazer contribuições para a IES e seus alunos, pois abrirá espaço para melhorias que impactariam ambos.

## **2 REFERENCIAL TEÓRICO**

### **2.1 Experiência de aprendizado**

Em um contexto geral, alunos que entram para a graduação criam expectativas que podem ou não serem correspondidas, gerando uma experiência positiva ou negativa que afeta sua motivação (HIROTA 2018). Isso ocorre em diversos aspectos, como em relação aos professores, cursos, estrutura das salas de aula, laboratório, restaurante e até mesmo amizades que serão criadas nesse espaço. Já em um contexto de aula, Chung e McLarney (2000) destacam que a experiência educacional é construída através de interações entre professores e estudantes, os quais se envolvem na aula trazendo suas experiências e perspectivas, e são elas que formam a experiência educacional. Nesse sentido, é papel dos educadores olhar para os perfis dos estudantes, identificar as particularidades e traçar estratégias para que a interação professor-aluno se torne mais dinâmica e personalizada (SILVEIRA *et al.*, 2015).

Em outra análise, o estudo de Veloso e Almeida (2001) realizado na Universidade Federal de Mato Grosso, campus Cuiabá, conclui que a evasão de alunos é mais um fenômeno institucional do que um processo dependente do aluno. Nele, é destacado que quando o aluno não encontra condições para sua permanência no ambiente universitário, acaba optando pelo abandono, e sugere que as universidades desenvolvam ações de acompanhamento e integração daqueles que ingressam à vida universitária. Portanto, entender a experiência que o estudante está passando e agir proativamente para melhorá-la pode ser um dos pontos chave para o sucesso de todos os envolvidos nesse contexto.

### **2.2 Métricas de experiências**

Quando se trata de métricas para medir a experiência, existem metodologias que ganharam espaço no mercado como o proposto por Reichheld (2003), que avalia a possibilidade de um cliente recomendar uma empresa a um amigo ou colega, medido através do indicador NPS (*Net Promoter Score*), o qual indica a relação direta entre o crescimento de uma empresa e o entusiasmo do cliente em recomendá-la. Tal métrica está fortemente ligada à experiência positiva que a companhia é capaz de proporcionar ao cliente (SITUMORANG *et al.*, 2016).

Assim, da mesma forma como é utilizado em um contexto de mercado, também é possível aplicar em um cenário acadêmico, avaliando a instituição provedora do serviço de educação. Sob essa ótica, Almeida et al. (2020) realizou uma pesquisa aplicando essa métrica na Universidade Federal de Itajubá, para o curso de engenharia de produção, e concluiu que o NPS permite fazer análises visuais e simplificadas sobre a opinião dos graduandos, o que facilita observar os pontos de melhorias para a instituição, além de servir para comparação entre elas.

Na metodologia NPS, cada entrevistado responde se recomendaria um determinado produto ou serviço para um amigo ou colega, em uma escala de 0 a 10. Após a coleta, a amostra é separada entre promotores (nota 9 e 10), neutros (nota 7 e 8) e detratores (nota menor que 7). Assim, calcula-se a nota final subtraindo a proporção de promotores pela proporção de detratores, como ilustrado abaixo.

$$NPS = \frac{\text{Total de Promotores} - \text{Total de Detratores}}{\text{Total de respondentes}} \quad (1)$$

Com isso, é possível criar uma noção ampla e estratégica sobre como os alunos estão se sentindo em relação à universidade ou ao curso, e, a partir disso, criar e analisar hipóteses junto com demais variáveis.

Outra métrica utilizada é o *Customer Satisfaction* (CSAT), proposto por Hsieh e Essex (2006), o qual busca avaliar a satisfação do cliente em relação aos principais elementos de um serviço através de uma escala *Likert*. Ela utiliza o conceito de bottom-up, no qual um conjunto de avaliações se juntam para criar a uma noção final de nota, já que eles compartilham um mesmo tópico. Vale observar que é uma abordagem distinta do NPS que faz uma avaliação genérica sobre o serviço, ainda que seja possível aprofundar as análises posteriormente.

Neste caso, a escala *Likert* vai de “1” a “5”, sendo “1” muito insatisfeito e “5” muito satisfeito. Já para calcular a nota CSAT, apesar da literatura trazer algumas alternativas, aqui é adotada a equação que traz a proporção de respondentes que estão satisfeitos, ou seja, que responderam “Satisfeitos” ou “Muito Satisfeitos” dividido pelo total de respondentes (2).

$$CSAT = \frac{\text{Total de Satisfeitos (4 e 5)}}{\text{Total de respondentes}} \quad (2)$$

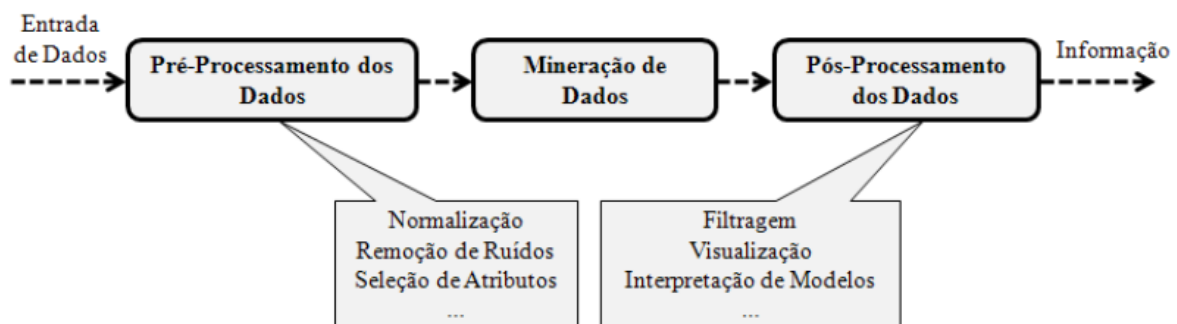
Assim, a nota individual de cada aspecto pode ser utilizada para se ter uma visão direcionada sobre aspectos do serviço que estão satisfatórios, ou não, na perspectiva do cliente. Além disso, ainda é possível calcular uma nota final com a média de todos os aspectos para se ter uma visão geral e simplificar a análise.

### 2.3 Mineração de dados

De acordo com Aggarwal (2015), nos últimos anos criou-se um grande desafio por conta do crescente volume de dados. A MD é o estudo de coletar, limpar, processar, analisar e gerar ideias relevantes a partir destes dados. Além disso, esse termo é utilizado para descrever diferentes aspectos de processamentos de dados.

A MD é parte processo conhecido como *Knowledge Discovery in Databases* (KDD), de acordo com Tan et al. (2006), o qual busca converter dados brutos em informações úteis por meio de processos como transformação, processamento e pós-processamento dos resultados da MD. A Figura 1 ilustra o fluxo deste processo.

Figura 1 - As etapas do processo de KDD



Fonte: Tan et al. (2006)

Além disso, segundo Han (2012), uma das áreas de trabalho dentro da MD é o aprendizado de máquinas, ou machine learning, assim como estatística, tecnologias de banco de dados, visualização de dados etc.

### 2.4 Tipos de aprendizado de máquina

O aprendizado de máquina vem se destacando muito na sociedade e tornando-se um dos pilares da tecnologia da informação (SMOLA, 2008). Ele é considerado um subcampo da inteligência artificial e, essencialmente, faz a máquina aprender com dados e entregar uma solução. Além disso, vale destacar que “o aprendizado de máquina busca resolver problemas do mundo real que apresentem



relevância e possuam bases de dados contendo informações que possibilitem alcançar a solução” (ROZA, 2016. p. 10).

Para isso, é utilizada a tecnologia computacional para criar algoritmos capazes de responder a questões consideradas importantes, com o intuito de extrair algum valor. Assim, ela busca ajudar as pessoas a tomar decisões mais inteligentes, interpretando algo além da capacidade humana e tornando aquilo que seria inviável, possível.

Segundo Han (2012), o aprendizado de máquina é dividido em quatro classes principais de aprendizagem.

- a) **Supervisionado**: utilizam-se dados rotulados na fase de treinamento da base, e esta tarefa busca prever uma resposta certa para cada dado. Um exemplo é a classificação de e-mails que são ou não spam.
- b) **Não supervisionado**: neste caso, são usados dados não rotulados para o processo de aprendizado do modelo, e tem como objetivo encontrar grupos que compartilham semelhanças dentro destes dados. Poderia ser utilizado, por exemplo, para identificar grupos de clientes a fim de direcionar estratégias de propagandas.
- c) **Semi-supervisionado**: utiliza tanto dados rotulados como dados não rotulados para ensinar o modelo. Uma forma de abordar isso é utilizando dados rotulados para ensinar o modelo e os dados não rotulados para ajustar os limites entre classes. Uma situação em que é utilizado é na análise de fala, no qual se rotula apenas partes de um áudio, já que rotular tudo seria muito trabalhoso.
- d) **Aprendizagem ativa**: é uma forma de permitir que o usuário tenha uma participação ativa no processo de aprendizagem, por exemplo, rotulando alguns dos dados. Esse tipo de abordagem é utilizado, por exemplo, na análise de avaliações feitas por clientes, em que há a necessidade de uma interpretação do comentário.

Quando se trata de aprendizado não supervisionado, Bramer (2016) explica que existem diversas áreas que podem extrair benefícios pelo agrupamento de objetos similares, como: na economia, encontrando países com sistemas econômicos similares; no financeiro, para encontrar empresas com performances financeiras semelhantes; no marketing, ao formar *clusters* consumidores com comportamentos de compras semelhantes; na medicina, agrupando pacientes com sintomas

semelhantes; em análises de crimes, no qual há a possibilidade de agrupar regiões por incidências de roubos, ou até por crimes mais raros, como assassinatos. Portanto, existem diversas possibilidades e depende do objetivo e profundidade de investigação por parte do autor para extrair informações relevantes e assertivas, independentemente da área de aplicação.

## 2.5 Técnicas de clusterização

Segundo Han (2012), as técnicas de clusterização particionam os objetos em grupos de tal forma que eles sejam similares dentro do mesmo grupo (intracluster), mas diferente entre grupos (intercluster). Essa similaridade é definida, geralmente, de acordo com a distância entre os objetos em um espaço, baseado em uma função de distância.

Uma análise de *cluster*, de acordo com Tan *et al.* (2006), são divididos em grupos que são significativos, úteis, ou ambos, sendo que a meta é encontrar um grupo significativo. Porém, isso nem sempre é possível e acaba sendo utilizado como uma forma de iniciar um estudo.

### 2.5.1 Métodos para clusterização

As técnicas fundamentais de clusterização, segundo Han (2012), podem ser separadas, de maneira geral, de acordo com as seguintes categorias.

- **Métodos de particionamento:** dado um conjunto de  $n$  objetos, são construídas  $k$  partições de dados, onde cada partição representa um *cluster* e  $k \leq n$ , sendo que cada partição precisa conter pelo menos um objeto. Além disso, esse método é baseado em distância, no qual utilizam-se técnicas iterativas de realocação que tentam melhorar o particionamento, movendo os objetos de um grupo para outro de acordo com o  $k$  estabelecido. Uma das formas de avaliar a qualidade desse método é medindo quão próximos ou relacionados estão os objetos dentro do mesmo *cluster*, e quão distantes estão os outros objetos de *cluster* diferentes. Alguns exemplos para este método é o *K-Means*, técnica baseada em centróides, o qual será aprofundado mais a frente, e o *K-Medoids* que se baseia em objetos.

- **Métodos hierárquicos:** estes métodos criam uma decomposição de um conjunto de dados, e de acordo como isso é formado, podem ser classificados tanto como aglomerativos, como divisivos. O primeiro caso, também denominado de “*bottom-up*”, começa com cada objeto formando um grupo separado, os quais vão sendo agrupados até que se formem um único *cluster*, ou até que chegue em uma condição de parada. Já o segundo caso, conhecido como “*top-down*”, começa com todos os objetos em um único *cluster* e vão sendo separados, a cada iteração, até cada objeto se tornar um grupo próprio, ou que também chegue em uma condição de parada. Ambos podem ser baseados em distância, densidade ou continuidade.
- **Métodos baseados em densidade:** um problema do método de particionamento é que ele só é bom em encontrar grupos em formatos esféricos, e o método baseado em densidade resolve esse problema. De forma geral, ele continua agrupando os objetos enquanto a densidade da vizinhança estiver dentro de um ponto limite, ou seja, para cada ponto de objeto dentro de um *cluster*, a vizinhança de um determinado raio deve conter um número mínimo de pontos. Com isso, ele consegue filtrar dados ruidosos ou *outliers* e encontrar *clusters* de formatos arbitrários. O DBSCAN (*Density-Based Clustering Based on Connected*) é uma técnica bastante utilizada dentro desse contexto, mas pode-se citar ainda o OPTICS (*Ordering Points to Identify the Clustering Structure*) e o DENCLUE (*Clustering Based on Density Distribution Functions*).
- **Métodos baseados em grade:** ele quantifica os objetos em um número finito de células que formam uma estrutura de grade e, a partir disso, analisa a clusterização. Essa abordagem possui uma alta velocidade de processamento, já que não é impactada pela quantidade de objetos, apenas pela quantidade de células em cada dimensão do espaço quantificado. Além disso, ela pode ser integrada com o método baseado em densidade e método hierárquicos. Dois exemplos para esse método são o STING (*Statistical Information Grid*) e o CLIQUE (*Clustering In Quest*).

Apesar dessas divisões, Han (2012) ainda cita a existência de outros métodos que não se encaixam em nenhuma delas, como é o caso da técnica *Fuzzy C-Means*, a qual utiliza uma teoria probabilística aplicando “lógica *fuzzy*” em que a variável pode assumir qualquer valor entre 0 e 1, seguindo uma função de pertinência, diferentemente de uma lógica booleana. Este método é especialmente bom quando há problemas em que o objeto não necessariamente pertence a um grupo, mas sim possui um grau de pertinência a vários grupos.

### 2.5.2 *K-Means*

O *K-Means* é uma técnica de particionamento que utiliza, de acordo com Bramer (2016), algoritmo de clusterização exclusivo, no qual cada objeto é alocado para apenas um dos conjuntos de grupos. A quantidade de grupos é denominada de  $k$ , sendo este um número inteiro, e, a partir disso, são selecionados os  $k$  pontos iniciais que serão utilizados como centróides. Após isso, cada objeto próximo ao centróide é agrupado de acordo com suas distâncias e um novo ponto de centróide é recalculado, repetindo esse procedimento até que o centróide não mude mais.

De acordo com Kubat (2017), é necessário, ainda, avaliar e comparar soluções encontradas, e para isso, utiliza-se a análise de qualidade dos *clusters*, de acordo com a distância média entre os objetos e os centróides que eles pertencem. Assim, é utilizada uma função que soma todas as distâncias de cada objeto com os centróides dos *clusters*, como segue:

$$SD = \sum_{j=1}^K \sum_{i=1}^{n_j} d(x_i^{(j)}, c_j) \quad (3)$$

Na Equação 3,  $SD$  é distância somada,  $d(x_i, c_j)$  é a distância entre cada objeto  $x_i$  e o centróide  $c_j$ ,  $K$  é o número de *clusters* e  $n_j$  é o número de objetos.

### 2.5.3 *Fuzzy C-Means*

Nesta técnica, diferente do *K-Means*, cada objeto não será alocado a apenas um *cluster*, e sim a vários. Segundo Tan (2007), o algoritmo do *Fuzzy C-Means* começa, assim como no *K-Means*, estabelecendo arbitrariamente os pontos iniciais dos  $C$  *clusters*, chamado de pseudo-partições, e atribui valores de pertencimento (*membership weights*), o qual é uma probabilidade de cada objeto estar em um *cluster*.

Com isso, são estabelecidos os centróides que serão realocados conforme cada iteração, com base na distância entre os objetos e os centróides. Assim, tal processo se repete até que os centróides não sofram mais alteração de posição, ou seja, até que a soma das distâncias alcance o menor valor possível.

## 2.6 Método da Silhueta

Diferentemente dos métodos citados anteriormente, o método da silhueta é voltado para avaliação dos resultados retornados pelas técnicas de clusterização através de um coeficiente. Para isso, utiliza-se de uma abordagem que combina coesão e separação (TAN, 2007), ou seja, faz uma análise através da distância e proximidade entre os *clusters*, e utiliza a distância como padrão para realizar os seguintes cálculos:

- Calcula a distância média entre cada objeto e os demais objetos do mesmo *cluster* ( $a_i$ );
- Calcula a distância média entre cada objeto e os demais objetos que não são do mesmo *cluster*, e seleciona o menor valor entre eles ( $b_i$ );
- Calcula o coeficiente de silhueta ( $S_i$ ) para cada objeto, que varia de -1 a 1, com a seguinte conta:

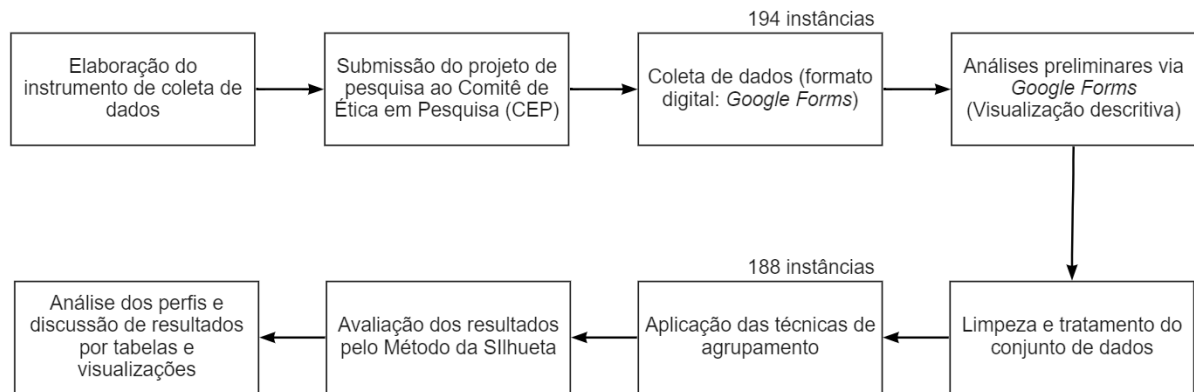
$$S_i = \frac{(b_i - a_i)}{\max(a_i, b_i)} \quad (4)$$

O objetivo é que o coeficiente de silhueta seja positivo, mostrando que existem mais pontos próximos do *cluster* do que distantes, e que  $a_i$  seja o mais perto de 0 quanto possível, o que indicaria que os *clusters* estão bem separados (IGUAL; SEGUÍ, 2017). Assim, para se ter uma análise geral, é tirado a média entre todos os coeficientes de silhuetas individuais.

### 3 METODOLOGIA

O presente trabalho passou por algumas etapas, desde a construção do instrumento da pesquisa até a aplicação dos métodos e análises dos resultados. A Figura 2 ilustra cada uma delas e a quantidade de instâncias de acordo com a etapa.

**Figura 2 - As etapas da pesquisa.**



**Fonte: Autoria própria (2022)**

Para a elaboração do instrumento de pesquisa, foram estruturadas cinco seções principais, sendo elas: o Termo de Consentimento Livre e Esclarecido (TCLE), onde o aluno realiza o aceite de participar; informações gerais que contém a situação do aluno na faculdade e forma de ingresso; perfil sociodemográfico, trazendo perguntas sobre as características do aluno e suas condições de vida; informações sobre rotina de estudo e bem-estar físico e mental, onde explora-se seus hábitos e questões de saúde, e por fim; avaliação da estrutura e experiência de aprendizagem, para captar a percepção do aluno sobre a universidade. Na Tabela 1, é possível observar as perguntas e a natureza delas.

**Tabela 1 - Perguntas disponíveis no questionário elaborado.**

**(continua)**

#	Informações gerais	Tipo	Descritivo
1	Curso de Graduação	Nominal	Curso em que o aluno está matriculado
2	Ano/Semestre de ingresso	Discreta	Ano e semestre em que foi matriculado
3	Período	Ordinal	Período atual que consta no sistema
4	Ingresso	Nominal	Se ingressou por transferência ou por SISU
5	Cotas	Nominal	Se ingressou via cotas
Perfil sociodemográfico			
6	Idade	Discreta	Idade do aluno
7	Moradia em Londrina	Discreta	Se mora sozinho, com a família, colegas ou em república

**Tabela 1 - Perguntas disponíveis no questionário elaborado.****(conclusão)**

#	Perfil sociodemográfico		
8	Trabalho	Nominal	Se trabalha e se depende do trabalho para se sustentar
9	Núcleo familiar	Discreta	Quantidade de pessoas que compõe o núcleo familiar
10	Renda familiar mensal	Ordinal	Intervalo de renda com base em salários mínimos (2022)
11	Filhos	Discreta	Quantidade de filhos
12	Sexo	Nominal	Sexo masculino ou feminino
13	Gênero	Nominal	Gênero com qual se identifica
14	Cor/Raça	Nominal	Cor ou raça com qual se identifica
15	Estudo em escola pública	Nominal	Se estudou escola pública
<b>Rotinas de estudo e bem-estar físico social</b>			
16	Exercícios físicos	Discreta	Quantidade de tempo por semana
17	Meditação	Discreta	Frequência por semana
18	Psicólogo ou psiquiatra	Nominal	Se frequenta ou não
19	Relação familiar	Nominal	Se considera saudável ou não a relação com a família
20	Ansiedade e Estresse	Nominal	Escolha múltipla de situações de ansiedade e estresse de acordo com o DASS ( <i>Depression Anxiety and Stress Scale</i> )
21	Estudo fora do horário de aula	Discreta	Quantidade de tempo por semana
22	Materiais de estudo	Nominal	Escolha múltipla de materiais utilizados para estudo
23	Local de estudo	Discreta	Local onde estudo com maior frequência
24	Modo de estudo	Nominal	Se prefere estudar sozinho ou com colegas
<b>Avaliação da estrutura e experiências</b>			
25	Sala de aula	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
26	Laboratórios	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
27	Didática dos professores	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
28	Disponibilidade dos professores	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
29	Métodos de avaliação dos professores	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
30	Conhecimento do conteúdo dos professores	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
31	Restaurante universitário (frequência)	Ordinal	Se utiliza com alta ou baixa frequência, ou nenhuma
32	Restaurante universitário	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
33	Biblioteca (frequência)	Ordinal	Se utiliza com alta ou baixa frequência, ou nenhuma
34	Biblioteca	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
35	Área de convivência	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
36	Estacionamento	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
37	Quadra esportiva	Discreta	Escala de 1 (Muito ruim) a 5 (Muito bom)
38	Recomendação do Curso	Discreta	Escala de 0 a 10 (NPS)
39	Recomendação da Universidade	Discreta	Escala de 0 a 10 (NPS)

**Fonte: Autoria própria (2022)**

Além disso, foi necessário elaborar o projeto de pesquisa para o CEP (Comitê de Ética e Pesquisa), a fim de obter a aprovação para aplicação do questionário na universidade. Portanto, este projeto foi aprovado dentro das diretrizes do CEP sob o CAAE: 57956421.6.0000.0177 e número do parecer: 5.458.136. Após validado, foi possível aplicá-lo, através do *Google Forms*, para os alunos de todos os cursos do campus através do e-mail institucional, totalizando mais de 2.000 convites. Ao todo, com dois disparos, um em cada semestre, foram obtidas 194 respostas.

Para a pergunta 20 sobre estresse e ansiedade, teve-se o cuidado de se pautar no modelo de perguntas DASS-21, assim como aplicado por Patias *et al.* (2016) em adolescentes brasileiro, sendo elas perguntas que abordam questões de natureza sensível de uma forma assertiva. Porém, foi decidido em não realizar perguntas relacionadas a depressão pois não faz parte do escopo do presente trabalho.

Com as respostas obtidas, foi possível analisar a distribuição de cada resposta pela própria ferramenta do formulário, o qual mostra gráficos de setores ou histogramas para cada pergunta. Após esta análise prévia, passou-se para a etapa de pré-processamento da base de dados, no qual foi realizada a remoção de instâncias com dados faltantes, discretização de algumas variáveis e a seleção das variáveis a serem utilizadas na próxima etapa. Tais variáveis foram, então, normalizadas pelo método de *MinMaxScaler* da biblioteca *sklearn*, concentrando os valores das variáveis no intervalo  $[0,1]$ . Na Tabela 2 estão descritas as variáveis selecionadas e o formato final após realizar a discretização.

**Tabela 2 - Variáveis que foram selecionadas para a clusterização.**

(continua)

#	Variáveis selecionadas	Descritivo (após discretização)
10	Renda familiar mensal	Escala de 0 a 7
11	Filhos	Binária de 0 ou 1
16	Exercícios físicos	Escala de 0 a 24
17	Meditação	Escala de 0 a 3
18	Psicólogo ou psiquiatra	Binário de 0 ou 1
19	Relação familiar	Escala de -2 a 1
20	Ansiedade e Estresse	Escala de 0 a 10
25	Sala de aula	Escala de 1 a 5
26	Laboratórios	Escala de 1 a 5
27	Didática dos professores	Escala de 1 a 5
28	Disponibilidade dos professores	Escala de 1 a 5
29	Métodos de avaliação dos professores	Escala de 1 a 5
30	Conhecimento do conteúdo dos professores	Escala de 1 a 5



**Tabela 2 - Variáveis que foram selecionadas para a aplicação.****(conclusão)**

#	Variáveis selecionadas	Descritivo (após discretização)
32	Restaurante universitário	Escala de 1 a 5
34	Biblioteca	Escala de 1 a 5
35	Área de convivência	Escala de 1 a 5
36	Estacionamento	Escala de 1 a 5
37	Quadra esportiva	Escala de 1 a 5
38	Recomendação do Curso	Escala de 0 a 10 (NPS)
39	Recomendação da Universidade	Escala de 0 a 10 (NPS)

**Fonte: Autoria própria (2022)**

Para a discretização, foi considerado como valor mais positivo conforme se traduzem em uma experiência boa, e mais negativo para aqueles considerados uma experiência ruim, como por exemplo “ter uma boa relação com a família” sendo “1”, “ter uma má relação com a família” sendo “-1”, e zero para uma situação neutra.

Em uma próxima etapa, seguiu-se com a aplicação dos algoritmos de clusterização, sendo utilizado para este trabalho as técnicas *K-Means*, *Fuzzy C-Means* e Agrupamento Hierárquico, a fim de abranger uma variedade de metodologias e encontrar aquela que melhor se adeque à base de dados utilizada. Assim, para cada técnica, analisou-se uma variedade de quantidades de *clusters* (de 2 a 5) e avaliado o resultado por meio do método da silhueta, comparando qual seria a quantidade ideal de grupos. Além disso, foram rodados dois conjuntos diferentes de variáveis, sendo a primeira com as perguntas avaliativas, da 26 ao 39, e a segunda com as questões pessoais do aluno (discretizadas), como ilustrado na Tabela 2.

Por fim, realizou-se uma etapa de exploração dos grupos obtidos por meio de tabelas de contingências e comparação de medidas de posição, cruzando com as demais variáveis que ficaram de fora do algoritmo e avaliando as diferenças e particularidades de cada um, traçando-se, assim, perfis de alunos.

## 4 RESULTADOS E DISCUSSÕES

### 4.1 Análise descritiva das respostas

Com as respostas coletadas através do formulário, foi analisada a distribuição geral das perguntas e, na sequência, foi tabulada uma versão resumida delas. Por praticidade, foi dividido em duas tabelas, uma para atributos qualitativos, trazendo a resposta mais frequente (Tabela 3), e outra para atributos quantitativos, trazendo a média e outras medidas de dispersão (Tabela 4).

**Tabela 3 - Comportamento das respostas qualitativas.**

#	Contagem	Únicos	Resposta mais frequente	Representatividade
1	194	7	Engenharia de Produção	31%
2	194	18	2020/1	11%
3	194	10	10º	14%
4	192	4	SISU (Sem estar cursando uma graduação anteriormente)	76%
5	194	3	Não ingressou via contas	59%
6	194	20	23	13%
7	194	17	Moro com a minha família.	46%
8	194	4	Não, e não dependo do trabalho para me manter.	35%
9	187	6	4	32%
10	194	8	De 1 a 3 salários mínimos (de R\$ 1.212,00 até R\$ 3.636).	39%
11	194	4	Nenhum filho	97%
12	194	2	Masculino	53%
13	194	5	Homem Cisgênero	49%
14	194	5	Branco	74%
15	194	5	Sim, estudei INTEGRALMENTE em escola pública	46%
16	194	17	Não pratico	38%
17	194	4	Não pratico meditação	87%
18	194	2	Não frequenta psicólogo ou psiquiatra	78%
19	194	4	Relação saudável	84%
20	178	106	Não sinto motivação pra estudar	4%
21	194	5	Até 2 (duas) horas por semana	27%
22	194	35	Caderno de anotações, Livros físicos, Livros Digitais, Textos encontrados na internet, Vídeos	21%
23	193	3	Casa	63%
24	193	2	Estuda sozinho	77%
31	194	3	Alta frequência	51%
33	194	3	Baixa frequência	64%

**Fonte: A autoria própria (2022)**

Em termos de distribuição, é possível observar que algumas perguntas tiveram predominância de uma única resposta, como a 11 com 98% dos respondentes não tendo filhos, a 17 com 84% do público não praticando meditação, a 19 com 88% tendo uma relação saudável com a família e a 24, em que 81% dos alunos responderam estudar sozinhos em casa.

**Tabela 4 - Comportamento das respostas quantitativas**

#	Contagem	Média	D.P.	Min	25%	50%	75%	Max
25	194	4,05	0,83	1	4	4	5	5
26	194	4,07	0,90	1	4	4	5	5
27	194	3,51	0,88	1	3	4	4	5
28	194	4,08	0,98	1	4	4	5	5
29	194	3,31	1,02	1	3	3	4	5
30	194	4,34	0,73	2	4	4	5	5
32	193	2,97	1,11	1	2	3	4	5
34	193	4,09	0,93	1	4	4	5	5
35	193	3,13	1,16	1	2	3	4	5
36	192	3,29	1,10	1	3	3	4	5
36	192	3,29	1,10	1	3	3	4	5
37	191	3,48	1,10	1	3	3	4	5
38	194	8,08	2,35	0	7	8	10	10
39	194	8,53	2,04	0	8	9	10	10

**Nota: D.P. – Desvio Padrão.  
Fonte: Autoria própria (2022)**

Já olhando para as variáveis quantitativas, destacam-se algumas perguntas com médias menores, sendo que a questão 32 sobre o restaurante universitário foi a que teve a maior quantidade de alunos insatisfeitos, gerando uma nota média de 2,89. A biblioteca (35), a área de convivência (34) e os métodos de avaliação dos professores (29) seguem em termos de nota, marcando as notas inferiores a 3,2. Outra forma de visualizá-la é por meio da métrica de experiência CSAT, ilustrada na Tabela 5.

**Tabela 5 - CSAT das perguntas de satisfação.**

(continua)

Pergunta	Satisfeitos
Conhecimento do professor	88%
Disponibilidade do professor	79%
Sala de Aula	78%
Laboratório	78%

**Tabela 5 - CSAT das perguntas de satisfação.****(conclusão)**

Pergunta	Satisfeitos
Biblioteca	75%
Didática do professor	53%
Quadra de esportes	46%
Métodos de avaliação	42%
Áreas de convivência	40%
Estacionamento	38%
Restaurante universitário	30%

**Fonte: Autoria própria (2022)**

De acordo com a Tabela 5, é possível observar com uma clareza maior a proporção de alunos que estão satisfeitos com sua jornada universitária em diversos aspectos, com destaque para o conhecimento do professor que foi muito bem pontuada e atingiu um índice de satisfação de 88% entre os pesquisados. Para simplificar ainda mais a visualização, foi dividido em duas dimensões, a “estrutural” e “docência”, trazido na Tabela 6, onde nota-se que o aluno está mais satisfeito, de forma geral, com a docência do que com a parte estrutural da instituição.

**Tabela 6 - CSAT por dimensão.**

Dimensão	Satisfeitos
Docência	66%
Estrutural	55%

**Fonte: Autoria própria (2022)**

Ainda sobre métricas de experiência, foi calculado também o NPS, para as perguntas 38 e 39, trazendo a proporção de promotores, neutros e detratores da universidade ou curso (Tabela 7).

**Tabela 7 - NPS em relação ao curso e universidade.**

Pergunta	NPS	Promotores	Neutros	Detratores
Recomendaria o curso?	36	52%	32%	16%
Recomendaria a universidade?	51	61%	29%	10%

**Fonte: Autoria própria (2022)**

De forma geral, os alunos que responderam à pesquisa são mais promotores da universidade do que do próprio curso, atingindo uma nota de 51 para aquela e 36 para esta. É possível observar ainda, que a proporção de neutros é muito próxima entre ambas, concentrando mais as diferenças nas extremidades. O fato de existir

poucos detratores é um grande indício de que a instituição tem mais chance de ser recomendada do que detratada, mas ainda há uma oportunidade em converter a proporção de neutros em promotores, já que há uma quantia significativa de alunos nessa faixa (em torno de 60 alunos). Nesse sentido, é importante aprofundar esta pesquisa para entender os motivos de alunos neutros e detratores estarem nestas categorias, sendo que os motivos de satisfação já citados anteriormente é um início promissor.

## 4.2 Análise das clusterizações

Nesta etapa, foram avaliados os resultados obtidos pelas técnicas, tanto para as variáveis avaliativas quanto para as variáveis de cunho pessoal.

### 4.2.1 Variáveis avaliativas

Para as variáveis avaliativas, o melhor resultado obtido foi com  $k = 2$ , pelo método da silhueta, no qual observou-se uma variação baixa entre as técnicas. Ainda assim, todos tiveram resultados positivos, mostrado na Tabela 8.

**Tabela 8 - Resultados pelo método da silhueta.**

Técnica	Valor da silhueta ( $k = 2$ )
<i>K-Means</i>	0.181
<i>Fuzzy C-Means</i>	0.178
Método Hierárquico	0.118

**Fonte: Aatoria própria (2022)**

A nota baixa pode ter se dado devido à quantidade de instâncias ser baixa e a quantidade de variáveis ser alta, o que pode tornar mais difícil que o algoritmo separe os grupos de maneira clara.

Em seguida, foi avaliado também a distribuição entre os grupos para cada *cluster*, a fim de verificar o equilíbrio entre eles. Segue a comparação na Tabela 9.

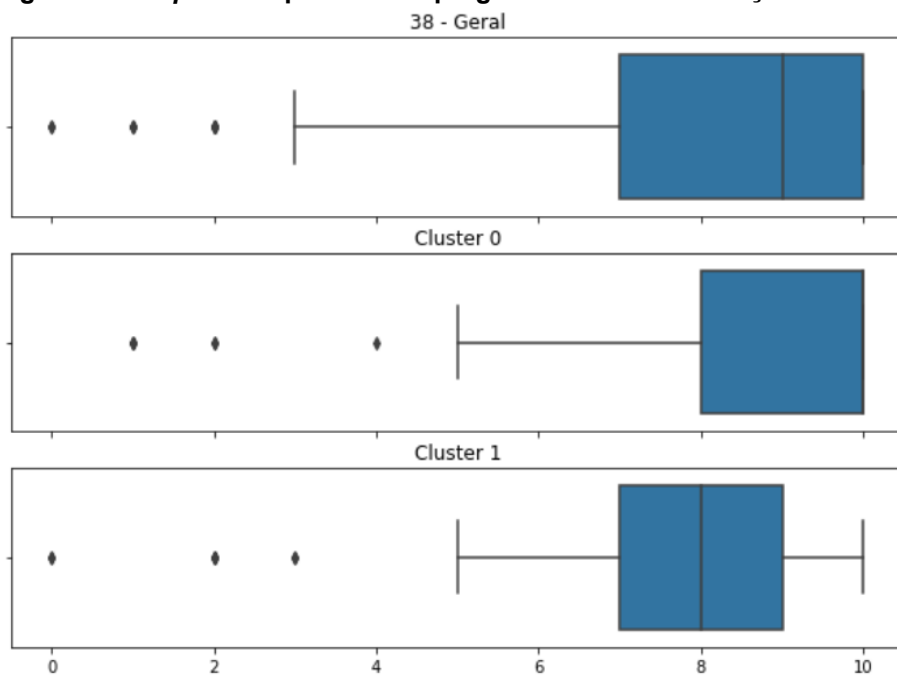
**Tabela 9 - Distribuição das instâncias dentro dos *clusters*, para cada técnica.**

Técnica	<i>Cluster 0</i>	<i>Cluster 1</i>
<i>K-Means</i>	103	85
<i>Fuzzy C-Means</i>	88	100
Método Hierárquico	126	62

**Fonte: Aatoria própria (2022)**

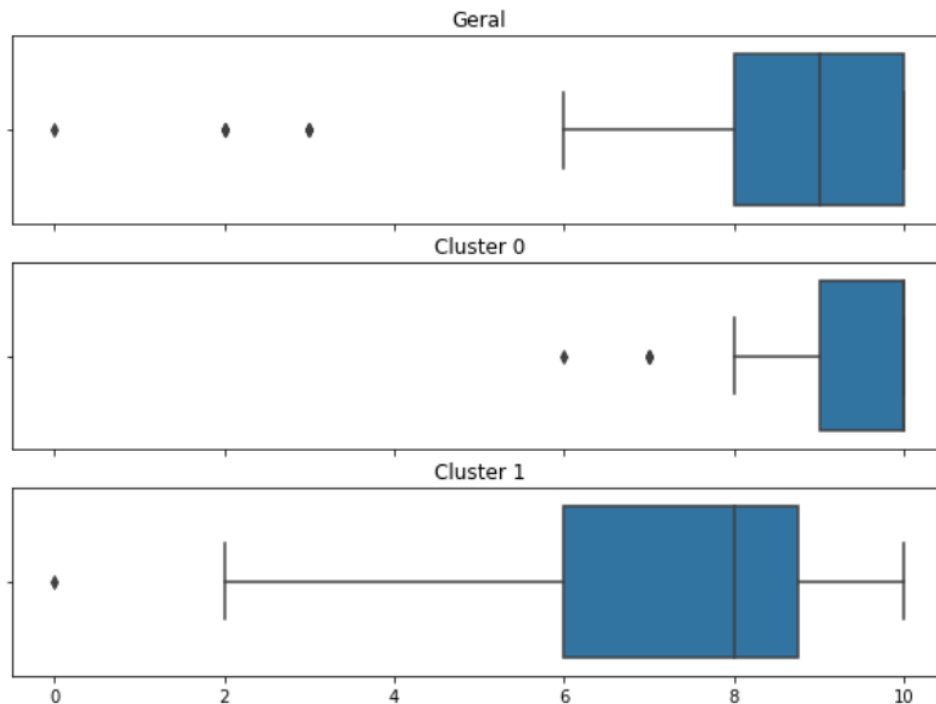
Como foi possível observar, o *K-Means* teve o melhor desempenho pelo método da silhueta e teve um balanceamento significativo entre os grupos, portanto se tornou objeto de estudo nas próximas análises. Assim, foi realizada uma exploração das medidas de posição da pergunta de NPS, já que é interesse deste trabalho entender se há uma separação entre alunos promotores e detratores em relação à universidade. Nas Figuras 3 e 4 é possível observar três *boxplots* que trazem esta diferença.

**Figura 3 - *Boxplot* comparativo da pergunta de NPS em relação ao curso.**



Fonte: Autoria própria (2022)

**Figura 4 - Boxplot comparativo da pergunta de NPS em relação à universidade.**



**Fonte: Autoria própria (2022)**

Com isso é possível concluir que há, de fato, uma predominância de promotores no *Cluster 0*, enquanto que o *Cluster 1* se concentra mais em notas inferiores a 9, ou seja, neutros e detratores. Porém, essa segregação é maior na pergunta 39 do que na pergunta 38, que fala sobre o curso. Como observado anteriormente, quando se trata do curso, comparado à universidade, há uma proporção maior de alunos que detratam ou são neutros em termos de recomendação. Apesar de não aparecer nos *boxplots* a mediana para o *cluster 0*, em ambas as perguntas o valor é 10, reforçando mais ainda a predominância de promotores.

Nesse mesmo propósito, foi analisada a diferença das médias das perguntas de 25 a 37, as quais perguntam sobre a satisfação em relação a aspectos da universidade, como mostrado na Tabela 10.

**Tabela 10 - Médias das questões avaliativas por *cluster*.**

(continua)

#	<i>Cluster 0</i>	<i>Cluster 1</i>	Diferença
25	4,40	3,64	0,76
26	4,50	3,58	0,92
27	3,81	3,12	0,69
28	4,48	3,67	0,81
29	3,77	2,78	0,99
30	4,53	4,08	0,45

**Tabela 10 - Médias das questões avaliativas por *cluster*.****(conclusão)**

#	<i>Cluster</i> 0	<i>Cluster</i> 1	Diferença
32	3,32	2,54	0,78
34	4,50	3,59	0,92
35	3,68	2,47	1,21
36	3,80	2,68	1,11
37	3,98	2,86	1,12

**Fonte: Aatoria própria (2022)**

É possível observar que todos possuem uma diferença acima de 0,4 pontos e algumas perguntas chegam a ultrapassar 1 ponto de diferença. Entende-se que há uma separação entre um grupo mais satisfeito e promotor, de forma geral, e outro menos satisfeito com as características da universidade, composto por neutros e detratores. Nota-se, ainda, que a pergunta 32 sobre o restaurante é a que tem a segunda menor média dentro do *cluster* 1 e a menor dentro do *cluster* 2, mostrando que é uma insatisfação maior para ambos os grupos. Já a didática do professor (pergunta 30), é a que tem a menor diferença de nota entre os grupos, e além disso, este item possui as melhores notas, indicando que a variável não foi determinante na formação dos *clusters*.

#### 4.2.2 Variáveis pessoais

Já para as variáveis pessoais dos alunos, os melhores resultados foram obtidos com três *clusters*, sendo estes ilustrados logo abaixo na Tabela 11.

**Tabela 11 - Resultados pelo método da silhueta.**

Técnica	Valor da silhueta (k = 3)
<i>K-Means</i>	0.498
<i>Fuzzy C-Means</i>	0.266
Método Hierárquico	0.496

**Fonte: Aatoria própria (2022)**

Como é possível observar, os valores retornados pelo método da silhueta aqui são melhores do que os de dois *clusters*, e isso pode ser explicado, em parte, pelo fato de serem utilizadas menos variáveis e tornar mais claro para o modelo distinguir os grupos. A Tabela 12 mostra a distribuição das instâncias entre os *clusters*, para cada técnica, sendo que os *clusters* foram harmonizadas para que grupos semelhantes fossem comparados na mesma coluna.



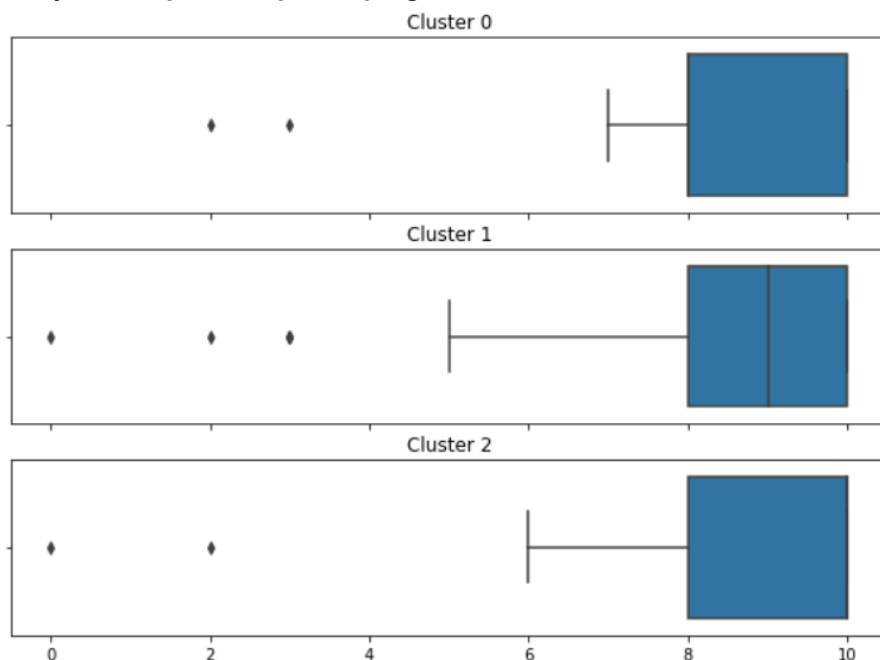
**Tabela 12 - Distribuição das instâncias dentro dos *clusters*, para cada técnica.**

Técnica	Cluster 0	Cluster 1	Cluster 2
<i>K-Means</i>	22	131	35
<i>Fuzzy C-Means</i>	64	83	41
Método Hierárquico	21	132	35

Fonte: Autoria própria (2022)

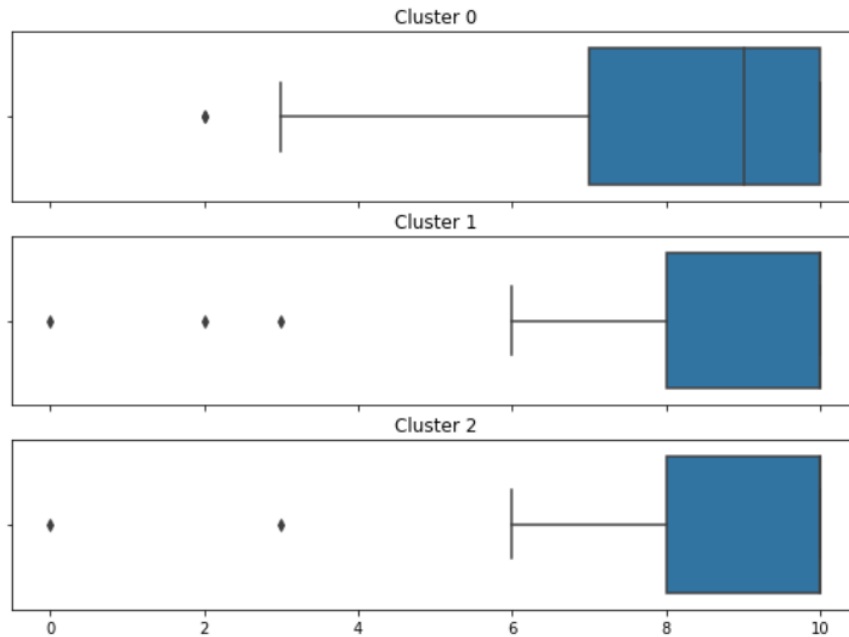
Nota-se aqui que o *K-Means* e o Método Hierárquico tiveram uma distribuição semelhante, além de terem tido um valor muito próximo no método da silhueta, indicando que formaram basicamente os mesmos grupos. Já o *Fuzzy C-Means* teve uma distribuição mais equilibrada entre os três grupos e se distinguiu das demais técnicas. Por esse motivo, as análises dos perfis referentes aos três grupos seguem de forma paralela para o *K-Means* e o *Fuzzy C-Means*, a fim de explorar diferenças entre eles.

Ao analisar em termos de recomendação do curso, as diferenças que antes eram marcantes agora se tornam mais sutis, como ilustrado na Figura 5 e 6.

**Figura 5 - *Boxplot* comparativo para a pergunta NPS em relação à universidade (*K-Means*).**

Fonte: Autoria própria (2022)

**Figura 6 - Boxplot comparativo para a pergunta NPS em relação à universidade (Fuzzy C-Means).**



Fonte: Autoria própria (2022)

Assim, percebe-se que, como esperado, o foco do algoritmo não é mais separar entre alunos promotores e detratores ou satisfeitos e insatisfeitos, já que as variáveis utilizadas neste caso não envolvem as questões avaliativas que trazia esta possibilidade, e sim questões voltadas para as características do aluno. Porém, ainda assim dá para perceber uma diferença, mesmo que pequena, quando comparadas as médias das questões de satisfação em cada *cluster*, evidenciado nas Tabelas 13 e 14.

**Tabela 13 - Médias das questões avaliativas por *cluster* (K-Means).**

#	Cluster 0	Cluster 1	Cluster 2
25	3,68	4,11	4,09
26	3,73	4,18	3,94
27	3,14	3,58	3,40
28	3,73	4,21	4,00
29	2,86	3,39	3,34
30	4,09	4,32	4,51
32	2,82	3,02	2,86
34	4,05	4,11	4,06
35	2,73	3,24	3,00
36	2,68	3,44	3,11
37	3,27	3,52	3,43
Média final	3,34	3,74	3,61

Fonte: Autoria própria (2022)

**Tabela 14 - Médias das questões avaliativas por cluster (Fuzzy C-Means).**

#	Cluster 0	Cluster 1	Cluster 2
25	3,97	4,10	4,10
26	4,02	3,93	4,20
27	3,33	3,27	3,73
28	3,95	3,88	4,35
29	3,11	3,20	3,54
30	4,16	4,49	4,39
32	2,94	2,90	3,02
34	3,98	4,02	4,20
35	2,97	3,02	3,31
36	3,22	3,12	3,43
37	3,44	3,39	3,54
Média final	3,55	3,57	3,80

Fonte: Autoria própria (2022)

Vale ressaltar que o fato dessa diferença ter diminuído não desqualifica o resultado, apenas muda a abordagem da análise, e destaca a necessidade de encontrar as novas variáveis que se tornaram importantes para esta técnica.

### 4.3 Análise dos perfis encontrados

Para avaliar os perfis, foram divididas as análises para o caso em que obteve-se dois grupos (*K-Means*) e para os casos de três grupos (*K-Means* e *Fuzzy C-Means*).

#### 4.3.1 Com dois grupos

Nesta etapa, foram exploradas as características dos dois grupos por meio de tabelas com as proporções que comparam a distribuição das respostas com maior destaque dentro do grupo, para cada pergunta. Vale ressaltar que a soma entre as porcentagens não dão 100% pois não é uma proporção entre *clusters*, e sim intracluster. Dessa forma, na Tabela 15 são apresentadas as principais perguntas que tiveram destaque por terem diferenças significativas em tais análises.

**Tabela 15 - Principais destaques por perfil (intracluster).**

(continua)

#	Pergunta	Principal diferença	Promotor	Neutro/Detrator
1	Curso de Graduação	Engenharia de Produção	27%	35%
2	Ano/Semestre de ingresso	-		
3	Período	6º período para frente	30%	58%
4	Ingresso	Pelo SISU sem estar graduando anteriormente	85%	67%

**Tabela 15 - Principais destaques por perfil (intracluster).****(conclusão)**

#	Pergunta	Principal diferença	Promotor	Neutro/Detrator
5	Cotas	-		
6	Idade	Acima de 23 anos	29%	40%
7	Moradia em Londrina	Mora sozinho	22%	12%
8	Trabalho	Trabalha e depende dele para se manter	28%	40%
9	Núcleo familiar	-		
10	Renda familiar mensal	Renda até 6 salários mínimos	48%	41%
11	Filhos	-		
12	Sexo	Masculino	47%	61%
13	Gênero	Homem cisgênero	43%	58%
14	Cor/Raça	Pardo	22%	9%
15	Estudou em escola pública	-		
16	Exercícios físicos	Não pratica	42%	34%
17	Meditação	-		
18	Psicólogo ou psiquiatra	Frequenta psicólogo ou psiquiatra	19%	25%
19	Relação familiar	Considera não saudável	6%	18%
20	Ansiedade e Estresse	Assinalou 3 ou mais situações	59%	69%
21	Estudo fora do horário de aula	-		
22	Materiais de estudo	-		
23	Local de estudo	Estuda em casa	61%	67%
24	Modo de estudo	-		
	Satisfação geral	Média do CSAT das questões de satisfação	75%	40%
	Recomendação geral	Média do NPS da 38 e 39	71	8

**Fonte: Autoria própria (2022)**

A partir dela, é possível observar características que se destacam dentro do grupo “neutro/detrator”, sendo as principais que tiveram maior diferença:

- Está cursando o 6º período para frente e entrou no curso já estando cursado uma graduação anteriormente;
- É mais velho, trabalha e depende dele para se manter;
- É do sexo masculino e se identifica com o gênero “Homem Cisgênero”;
- Menor proporção de pardos;
- Possui uma relação não saudável com a família e passa por mais de três situações de estresse, ansiedade ou falta de motivação.

Além deste público estar mais insatisfeito com a universidade a ponto de não recomendá-la, ainda é possível destacar outras diferenças menores, mas que se destacam, como:

- Proporção maior de alunos cursando Engenharia de Produção;
- Possui renda acima de seis salários mínimos;
- Praticam mais atividades físicas;
- Mais alunos que frequentam psicólogo ou psiquiatra;
- Mais alunos que estudam em casa.

Dado esse cenário, entende-se que o perfil “neutros/detratores” são alunos mais velhos que já possuem maiores responsabilidades por precisarem se sustentar e estão mais avançados no curso. Apesar de praticarem mais exercícios físicos, passam por mais situações de estresse, ansiedade ou falta de motivação e possuem uma relação menos saudável com a família. Conseqüentemente, o aluno “promotor” é o oposto deste perfil, concentrando mais alunos jovens e novos no curso, os quais possuem menores responsabilidades e acabam passando menos por situações de estresse emocional.

O estudo de Hewson (2018) aplicado em um curso à distância mostra que estudantes que se sentem parte de uma comunidade tendem a avaliar melhor as experiências acadêmicas, além de concluir que o suporte em questões do ensino seria o aspecto mais valorizado pelos alunos, entre diversos outros avaliados. Além disso, também traz uma avaliação em três perfis de estudantes: o aluno motivado e engajado que tem o foco de trazer benefícios à carreira no longo prazo e não precisa de muita ajuda; o aluno motivado mas distraído que, apesar de gostar do curso, tem dificuldade de acompanhar e precisa de uma atenção maior no momento inicial do semestre para se adaptar; e por fim o aluno pressionado e desengajado que está ali para conseguir uma promoção no emprego e que precisa de compreensão por parte dos professores e que os materiais sejam mais envolventes.

Nesse sentido, é possível traçar um paralelo com os resultados aqui apresentados em que, entre neutros e detratores, há uma predominância maior de alunos que trabalham e dependem dele para se manter, gerando uma pressão maior tanto na sua jornada como estudante como na sua jornada de trabalho. Além disso, alunos que entram por transferência podem enfrentar uma dificuldade maior de criar grupos e se sentir parte de uma comunidade. Com isso, este perfil pode estar enfrentando uma experiência mais negativa em relação à graduação e, conseqüentemente, se tornando mais crítico na hora de avaliar os aspectos da universidade.

#### 4.3.2 Com três grupos

Para avaliar os resultados obtidos com três *clusters*, assim como na análise feita anteriormente, foi construída a mesma visualização no formato de tabela comparando as principais diferenças encontradas entre cada um dos três grupos. Primeiramente serão discutidos os perfis obtidos pela técnica de *K-Means* (Tabela 16).

**Tabela 16 - Principais destaques por perfil - *K-Means* (intracluster).**

(continua)

#	Pergunta	Principal diferença	Perfil 1	Perfil 2	Perfil 3
1	Curso de Graduação	-			
2	Ano/Semestre de ingresso	-			
3	Período	6º período para frente	55%	46%	40%
4	Ingresso	-			
5	Cotas	-			
6	Idade	Acima de 23 anos	82%	54%	58%
7	Moradia em Londrina	Mora sozinho(a)	14%	26%	16%
8	Trabalho	Não trabalha e não depende dele para se manter	36%	40%	34%
9	Núcleo familiar	Menos de 2 pessoas	67%	41%	72%
10	Renda familiar mensal	Renda até 6 salários mínimos	64%	34%	44%
11	Filhos	-			
12	Sexo	Masculino	59%	46%	54%
13	Gênero	-			
14	Cor/Raça	Branco	82%	80%	72%
15	Estudo em escola pública	Estudou integralmente ou parcialmente em escola pública	64%	51%	66%
16	Exercícios físicos	-			
17	Meditação	-			
18	<b>Psicólogo ou psiquiatra</b>	<b>Frequenta psicólogo ou psiquiatra</b>	<b>27%</b>	<b>100%</b>	<b>0%</b>
19	<b>Relação familiar</b>	<b>Considera não saudável</b>	<b>95%</b>	<b>0%</b>	<b>0%</b>
20	Ansiedade e Estresse	Assinalou 4 ou mais situações	64%	63%	41%
21	Estudo fora do horário de aula	Estuda acima de 4 horas	27%	17%	27%
22	Materiais de estudo	-			

**Tabela 16 - Principais destaques por perfil - K-Means (intracluster). (conclusão)**

#	Pergunta	Principal diferença	Perfil 1	Perfil 2	Perfil 3
23	Local de estudo	Estuda em casa	50%	57%	68%
24	Modo de estudo	Sozinho	73%	65%	81%
	Satisfação geral	Média do CSAT das questões de satisfação	49%	56%	61%
	Recomendação	Média do NPS da 38 e 39	34	57	40

**Fonte: Autoria própria (2022)**

Nesse cenário é possível observar que as variáveis 18 e 19 foram determinantes na formação do perfil, já que houve uma concentração grande de alguma dessas características em cada um deles. Porém, outras características também se destacaram em cada perfil e serão discutidas logo em seguida.

No perfil 1 é possível observar a predominância das seguintes características:

- Possui uma relação não saudável com a família;
- Alguns frequentam psicólogo ou psiquiatra;
- Passam por quatro ou mais situações de ansiedade, estresse ou falta de motivação;
- É o perfil que menos estuda em casa e boa parte deles estudam sozinhos;
- É mais velho e mais avançados no curso;
- Maior proporção de renda familiar de até seis salários mínimos;
- A maioria veio de escolas públicas;
- É o perfil menos satisfeito com a universidade e que menos a recomenda.

É possível entender que este perfil possui uma certa similaridade com o perfil neutro/detrator observado no caso de dois *clusters*, sendo ambos perfis insatisfeitos com a universidade, de maneira geral. Ademais, além de enfrentar mais situações de estresse e ansiedade, também possui uma relação não saudável com a família. Já o perfil 2 é marcado pelas seguintes características:

- Frequenta psicólogo ou psiquiatra;
- Possui uma relação saudável com a família;
- Muitos passam por quatro ou mais situações de estresse, ansiedade ou falta de motivação;

- Poucos estudam acima de 4 horas e é o público que menos estuda sozinho;
- Menor proporção de quem veio de escolas públicas;
- Maior concentração de mulheres;
- Perfil mais jovem e que mora sozinho;
- Renda familiar acima de seis salários mínimos e núcleo familiar acima de 2 pessoas;
- Maior concentração de quem não trabalha e não depende dele para se manter;
- É o perfil que mais recomenda a universidade e/ou curso.

Neste perfil, é observado que, apesar de enfrentar muitas situações de estresse, ansiedade ou falta de motivação, tem uma condição de vida mais estável e com menores responsabilidades. Além disso, os hábitos de estudos mostram que ele busca mais momentos de interação social.

Por fim, o perfil 3 é destacado pela maior proporção dos seguintes aspectos:

- Não frequenta psicólogo ou psiquiatra e possui uma relação saudável com a família;
- Menor presença de ansiedade, estresse ou falta de motivação;
- É quem mais estuda sozinho e em casa;
- É mais novo no curso e vieram de escolas públicas;
- Pessoas declaradas brancas;
- É o perfil mais satisfeito com a universidade, de forma geral.

Aqui é destacado o fato de serem pessoas que têm uma vida mais saudável, dado o fato de enfrentarem menos situações de ansiedade, estresse ou falta de motivação. Isso pode ser um dos fatores que as tornem menos exigentes na hora de avaliar a universidade, sendo mais satisfeitos.

Da mesma forma, foram analisados os perfis obtidos pela técnica de *Fuzzy C-Means*, trazido na Tabela 17 a seguir.



**Tabela 17 - Principais destaques por perfil – Fuzzy C-Means (intracluster).**

#	Pergunta	Principal diferença	Perfil 1	Perfil 2	Perfil 3
1	Curso de Graduação	-			
2	Ano/Semestre de ingresso	-			
3	Período	6º período para frente	41%	51%	40%
4	Ingresso	-			
5	Cotas	-			
6	Idade	Acima de 23 anos	50%	61%	67%
7	Moradia em Londrina	Mora sozinho(a)	16%	24%	16%
8	Trabalho	-			
9	Núcleo familiar	Menos de 2 pessoas	79%	43%	68%
10	Renda familiar mensal	Renda até 6 salários mínimos	39%	41%	51%
11	Filhos	-			
12	Sexo	Masculino	48%	49%	59%
13	Gênero	Homem cisgênero	44%	49%	54%
14	Cor/Raça	Branco	80%	78%	69%
15	Estudo em escola pública	Estudou integralmente ou parcialmente em escola pública	55%	59%	72%
16	Exercícios físicos	Pratica exercícios físicos	59%	61%	64%
17	Meditação	-			
18	<b>Psicólogo ou psiquiatra</b>	<b>Frequenta psicólogo ou psiquiatra</b>	<b>0%</b>	<b>100%</b>	<b>0%</b>
19	<b>Relação familiar</b>	<b>Considera não saudável</b>	<b>23%</b>	15%	0%
20	<b>Ansiedade e Estresse</b>	<b>Assinalou 4 ou mais situações</b>	<b>83%</b>	<b>66%</b>	<b>12%</b>
21	Estudo fora do horário de aula	Estuda acima de 4 horas	22%	20%	30%
22	Materiais de estudo	-			
23	Local de estudo	Estuda em casa	67%	56%	65%
24	Modo de estudo	Sozinho	81%	68%	78%
	Satisfação geral	Média do CSAT das questões de satisfação	55%	55%	64%
	Recomendação	Média do NPS da 38 e 39	27	52	49

**Fonte: Autoria própria (2022)**

Diferente da anterior, a relação familiar deixa de ser tão determinante e esse papel passa a ser estabelecido pela pergunta 20, de ansiedade, estresse ou falta de motivação. Por outro lado, a pergunta 18 permanece sendo incisiva na formação dos grupos. Vale destacar que a distribuição entre os perfis muda aqui, sendo mais equilibrada do que a anterior, com 64, 41 e 83 alunos para cada perfil,

respectivamente. A seguir serão destacados os pontos mais distintos de cada perfil, a começar pelo perfil 1, o qual era composto por 22 alunos antes e agora passa a ser composto por 64 alunos:

- A grande maioria passa por quatro ou mais situações de estresse, ansiedade ou falta de motivação;
- Não frequenta psicólogo ou psiquiatra e possui concentra uma maior proporção de alunos que possuem uma relação não saudável com a família;
- É o público mais jovem e mais novo no curso;
- Possui um núcleo familiar menor, mas renda familiar maior;
- Menor proporção do sexo masculino e homens cisgêneros;
- Maior proporção de pessoas declaradas brancas;
- Menos pessoas que vieram de escolas públicas;
- Estuda sozinho e em casa;
- Público que menos recomenda a universidade e/ou curso.

Como é possível observar pela proporção de alunos que assinalaram mais de quatro situações para a pergunta 20, é um perfil marcado pela presença muito forte de ansiedade, estresse ou falta de motivação. Porém, ao mesmo tempo é um perfil que não frequenta psiquiatra ou psicólogo, o que dá a entender que são pessoas que estão passando por dificuldades emocionais sem conseguir buscar ajuda. Além disso, é um perfil com núcleo familiar menor, que estuda mais em casa e sozinho, dando uma noção de isolamento social mais forte do que os demais perfis. Com isso, é possível relacionar o fato de ser o perfil que menos recomenda o curso e que avalia pior as experiências acadêmicas com o mesmo estudo apresentado para o perfil neutro/detrator anteriormente, relacionado à falta de senso de comunidade.

Para o perfil 2 se destacaram as seguintes características:

- Frequenta psicólogo ou psiquiatra;
- Enfrenta bastante situações de estresse, ansiedade ou falta de motivação;
- Estuda menos, fora de casa e acompanhado;
- É o perfil mais avançado no curso;
- Núcleo familiar e renda familiar maior;
- Maior proporção de quem mora sozinho;
- É quem mais recomenda a universidade e/ou curso.

Este perfil foi o que menos oscilou em relação ao obtido pelo *K-Means*, sendo muito semelhante ao perfil 2 anterior. A maior mudança foi no aumento de 15% na proporção de alunos que consideram possuir uma relação não saudável com a família, característica que era zerada no modelo anterior. Na prática, seis alunos com esta resposta e que frequentavam psicólogo no perfil 1 do *K-Means* migrou para o perfil 2 do *Fuzzy C-Means*, monopolizando aqui os alunos que frequentam profissionais da saúde mental.

Por último, o perfil 3 desta técnica concentrou os pontos destacados a seguir:

- Não frequenta psicólogo e possui uma relação saudável com a família;
- Maior proporção de pessoas que vieram de escolas públicas;
- É do sexo masculino e homem cisgênero;
- Menor índice de pessoas declaradas brancas;
- São mais velhos e com a menor renda entre os perfis;
- Praticam mais exercícios físicos;
- Também estudam sozinhos e em casa;
- Perfil mais satisfeito com a universidade de forma geral.

Assim como o perfil 3 do *K-Means*, este perfil se destaca por ter uma saúde psicológica/emocional e uma ótima relação familiar, com uma proporção menor ainda de alunos que enfrentam quatro ou mais situações de estresse, ansiedade ou falta de motivação. Outras mudanças que ocorreram nesse sentido foram: quantidade de alunos acima de 23 anos (+9%); renda de até seis salários-mínimos (+7%); vieram de escolas públicas (+6%); e a nota de NPS geral aumentando em 9 pontos.

#### **4.4 Implicações práticas**

Com as informações de perfis obtidas, o próximo passo importante a ser tomado é entender como elas podem ser usadas na prática pelas partes interessadas, como os tomadores de decisão, se tratando de gestão educacional, e professores. Portanto, é preciso entender qual abordagem faz mais sentido para o uso final, se é a técnica que traz destaque para a distinção entre alunos que estão satisfeitos e leais à universidade ou a técnica que se caracteriza mais pelas questões ligadas à saúde do aluno.

Vale destacar que essa pesquisa foi realizada a partir da percepção subjetiva dos alunos, podendo haver diferenças na percepção e de como julgá-la, já que muitas vezes o comportamento pode ser diferente da fala. Portanto, é importante

complementá-la com outros tipos de pesquisas que buscam responder o porquê das hipóteses levantadas, com entrevistas em profundidade, por exemplo. Além disso, por conter questões avaliativas da universidade, uma aplicação constante da mesma contribuiria para acompanhar as evoluções ou involuções da universidade ao longo do tempo, correlacionando-as com as ações implementadas e buscando medir suas eficácias.

Dado este contexto, a seguir são discutidas algumas possibilidades de ações que foram possíveis extrair da pesquisa e dos perfis observados. Para isso, foi montado um quadro (Quadro 1) trazendo algumas ideias de ações e uma breve descrição.

**Quadro 1 - Ideias de ações extraídas do estudo.**

(continua)

<b>Ações</b>	<b>Descrição</b>
Suporte por parte dos professores	Ao entender quais alunos precisam de uma maior atenção, é possível focalizar esforços para dar um suporte maior a eles, como é o caso do aluno pressionado e desengajado (perfil detrator), discutido anteriormente. Uma sugestão seria se aproximar mais destes alunos no início do semestre para mantê-lo engajado.
Suporte por parte da instituição	Cabe a instituição facilitar e incentivar ações que gerem valor para melhorar a experiência do aluno enquanto graduando. Um exemplo seria facilitar, incentivar e até coordenar a implementação das melhorias nos pontos de atritosa ressaltados pela pesquisa.
Incentivo ao tratamento psicológico	Dado a existência de um perfil que passa por situações de problemas emocionais e/ou psicológicos, que não consegue ou busca algum acompanhamento especializado (perfil 1 do <i>Fuzzy c-means</i> ), cabe à universidade incentivar o tratamento por meio de profissionais da área, tanto os que são fornecidos gratuitamente pela instituição, quanto os particulares.
Incentivo ao tratamento psicológico	Dado a existência de um perfil que passa por situações de problemas emocionais e/ou psicológicos, que não consegue ou busca algum acompanhamento especializado (perfil 1 do <i>Fuzzy c-means</i> ), cabe à universidade incentivar o tratamento por meio de profissionais da área, tanto os que são fornecidos gratuitamente pela instituição, quanto os particulares.

Quadro 2 - Ideias de ações extraídas do estudo.

(conclusão)

Ações	Descrição
Auxílio de bolsas direcionadas	Ao identificar perfis que possuem dificuldades financeiras maiores e passam por situações difíceis emocionalmente, como é o caso do perfil 1 ( <i>K-Means</i> ), priorizar a eles bolsas de auxílio pode agregar mais valor do que o usual.
Tornar a pesquisa constante	Uma forma de permitir que haja evoluções constantes, é implementando esse tipo de pesquisa de maneira regular no campus. Uma opção seria colocá-la no momento antes da matrícula semestral, no qual já acontece uma pesquisa atualmente. Se não, ainda seria possível manter a pesquisa no formato de formulário online.
Melhorar pontos de insatisfação mostrados pelo CSAT	Ao analisar a satisfação dos alunos em relação aos diversos aspectos da universidade, é possível entender onde mais há uma insatisfação maior. Com isso, é possível buscar os porquês de tais atritos e atuar sobre eles para aumentar o nível de satisfação para cada serviço oferecido.
Pontos de atritos para neutros e detratores	Pela grande quantidade de alunos que se enquadraram no perfil de neutros e detratores, projetos que busquem atacar seus pontos de atritos podem ajudar a minimizar a quantidade de detratores, converter os neutros e aumentar os promotores, melhorando experiência dos acadêmicos e a imagem da universidade.

Fonte: Autoria própria (2022)

Além de melhorar a experiência dos graduandos, tais ações são fundamentais para diminuir a taxa de evasão que várias universidades enfrentam, já que elas ajudam a concentrar os esforços nos pontos de atritos de cada perfil e diminuir as partes negativas das situações que eles enfrentam. Nesse sentido, é possível relacionar características dos perfis observados com outros estudos que trazem relação com a evasão dos alunos. Um exemplo seria com o estudo de Homma Junior (2021), o qual aplica técnicas de clusterização sobre dados de sistema dos alunos de engenharia de uma universidade pública paranaense, para obter grupos distintos de acordo com a permanência ou abandono no curso. Entre eles, há o grupo caracterizado por alunos que abandonaram o curso ou que estão atrasados, sendo maior a proporção, quando comparado aos outros dois perfis, de alunos do sexo masculino (83%) e de pessoas que vieram de escolas públicas (62%). Alguns perfis

encontrados neste trabalho também destacam tais características, como é o caso do perfil 1 do *K-Means*, o que abre espaço para questionar se, para estes perfis, o abandono ou atraso no curso também estariam presentes.

Nesse caso, para que seja possível fazer tal cruzamento, seria necessário a coleta do registro do aluno (RA) para relacionar cada respondente com seus dados do sistema da universidade, assim como realizado no estudo mencionado. Com isso, seria possível fazer a relação adequada e entender se realmente existe uma conexão entre os dois trabalhos e, então, elaborar medidas de melhorias cabíveis.

Como mostra o estudo de Lee (2018), o qual utilizou o NPS para medir a satisfação ao suporte de TI (Tecnologia da Informação) da Universidade de Illinois, a métrica se torna uma ferramenta utilizada desde a linha de frente de serviços de *service desk* até às lideranças da universidade. A métrica passou a ser publicada anualmente e utilizada como parte do plano anual para buscar melhores níveis de serviços do departamento, sendo que em 2017, foi obtido uma nota de 77 pontos. Da mesma forma, é possível utilizar as métricas de satisfação abordadas neste estudo e combinar com características dos perfis obtidos, como é o caso do perfil neutro/detrator, para elaborar ações assertivas e monitorar os resultados, a fim de criar uma gestão universitária que esteja melhorando constantemente.

Com o perfil neutro/detrator, é possível entender que para melhorar seu nível de satisfação e convertê-los a promotor, é necessário criar um ambiente de ensino onde há um suporte maior dos professores, como já mostrado pelo estudo de Hewson (2018), pois são alunos mais atarefados e pressionados que têm menos tempo para se dedicar ao contexto acadêmico do que os demais. É possível, também, que haja turmas com uma maior concentração deste perfil por concentrarem pessoas que trabalham de dia e estudam de noite, como é o caso do curso de Engenharia de Produção do campus Londrina. Neste caso, a compreensão por parte do professor é essencial para manter os alunos engajados e torná-los promotores do curso.

Já se tratando da qualidade de vida do aluno, Nogueira-Martins (2018) destaca que olhar para a saúde mental do aluno não deve ser apenas uma preocupação dos profissionais da área da saúde, mas também dos gestores, docentes, pedagogos e assistentes sociais da universidade. Em seu estudo é levantado, ainda, três situações causadores de estresse para os alunos, sendo eles relacionados ao:

- Processo de profissionalização, enfrentando a construção de uma carreira profissional em um mercado altamente competitivo;

- Situações específicas de ensino e dificuldades nos relacionamentos com colegas e professores;
- Características pessoais como traços de personalidades, história pessoal e familiar de saúde, situações econômicas e sociais, problemas familiares e outros eventos estressantes da vida.

Como é possível perceber, são características já mencionadas em alguns perfis formados neste estudo, principalmente entre os que se destacaram pela presença maior de estresse, ansiedade e falta de motivação. Com isso, entender potenciais alunos que precisem de suporte para lidar com tais situações pode abrir espaço para evitar que isso se prolongue ao longo da sua jornada profissional, caso sejam tomadas ações para tal.

Portanto, uma vez que estudos como este avançam, espera-se que fique cada vez mais claro para os tomadores de decisão da instituição o que pode ser feito para que melhore o ambiente acadêmico, a qualidade de vida dos alunos e diminua a taxa da desistência nos cursos. Isso pode trazer não só resultados imediatos e diretos, com a melhoria na experiência dos alunos, mas também a longo prazo e indiretos, com o aumento de profissionais qualificados para o mercado, impactando positivamente nas demandas locais.

## 5 CONCLUSÃO

A partir das três técnicas de clusterização utilizadas neste trabalho, foi possível encontrar pelo menos cinco perfis de alunos graduandos diferentes, sendo dois a partir de variáveis em que os alunos avaliaram a universidade, e três pelas variáveis de cunho pessoal do aluno. Em ambas as situações, a técnica que obteve as melhores notas pelo método da silhueta foi o *K-Means*, atingindo 0.181 para a primeira situação e 0.498 para a segunda. Além disso, foi explorado, ainda, o *Fuzzy C-Means* para as variáveis de cunho pessoal devido ao seu bom desempenho no balanceamento dos grupos, mas os três perfis obtidos seguiram uma lógica semelhante para ambas as técnicas, apesar de demonstrar particularidades em alguns deles.

Entre os perfis obtidos, pode-se destacar o “neutro/detrator” que avalia pior a universidade e tende a recomendar menos o curso ou a instituição. É um perfil que se caracteriza, principalmente, pela presença maior de alunos mais velhos que possuem maiores responsabilidades por precisarem se sustentar e que também estão mais avançados no curso, além de enfrentarem mais situações de estresse, ansiedade ou falta de motivação e possuem uma relação menos saudável com sua família. Nesse sentido, o perfil promotor possui características opostas a este.

Além desses, outros perfis encontrados foram: perfil 1, marcado pela relação não saudável com a família e presença de alunos que enfrentam situações de estresse, ansiedade ou falta de motivação; perfil 2 com predominância de estudantes que frequentam psicólogos ou psiquiatras, além de passarem por situações de desequilíbrio emocional, ainda que com uma proporção abaixo do perfil anterior; e o perfil 3 cujo alunos possuem uma relação saudável com a família, não frequentam psicólogo e passam por bem menos situações de estresse, ansiedade ou falta de motivação.

Por fim, foi possível relacionar cada perfil com as demais características que não foram utilizadas no modelo, permitindo uma análise aprofundada de cada perfil. Com isso, fica o desafio de entender qual deles seria o ideal para ser utilizado na gestão universitária e aprimorar o estudo para que seja cada vez mais rico em informações, como, por exemplo, correlacionando com dados de evasão dos alunos da instituição.



Como limitações deste trabalho, destaca-se a abrangência de perguntas no questionário, pois há aspectos não abordados, como uma avaliação dos departamentos do curso e suporte por parte da instituição, por exemplo. Além disso, poderia ser utilizado técnicas de pré-processamento para buscar um desempenho melhor no algoritmo, como o PCA que trabalha a redução de dimensionalidade das variáveis (ou atributos) para simplificar o modelo, já que foram utilizadas 21 variáveis neste trabalho.

Para trabalhos futuros, recomenda-se uma revisão na estrutura do questionário para buscar contemplar melhor a experiência do universitário enquanto graduando, e tornar a pesquisa constante para que possa ser mensurado a evolução ao longo do tempo. Ainda sugere-se uma aplicação de técnicas mais robustas como testes de hipóteses para validar as diferenças entre os perfis. E, por fim, entende-se como oportunidade cruzar o questionário utilizado neste trabalho com a pesquisa que acontece na universidade semestralmente, de forma que se complementem ou se unam para formar uma pesquisa mais assertiva.

## REFERÊNCIAS

- ADEKITAN, A. I.; SALAU O. The impact of engineering students' performance in the first three years on their graduation result using educational data mining. **Heliyon**, [S.L.], v. 5, n. 2, fev. 2019. Elsevier BV. DOI: <http://dx.doi.org/10.1016/j.heliyon.2019.e0125>
- AGGARWAL, C. C. **Data mining**: The textobook. New York, EUA: Springer, 2015.
- ALMEIDA, D. S.; BARROS, M. A. C. F.; SILVA, C. E. S.; MELLO, C. H. P. Utilização do Net Promoter Score para avaliação do curso de Engenharia de Produção de uma Universidade Federal de Ensino Superior. **Anais do Encontro Nacional de Engenharia de Produção**. 2020.
- BIAZUS, C. A. **Sistemas de fatores que influenciam o aluno a evadir-se dos cursos de graduação na UFSM e na UFSC**: um estudo no curso de ciências contábeis. Tese. (Doutorado em Engenharia de Produção). Universidade Federal de Santa Catarina, Florianópolis, 2004.
- BONDE, S. N.; KIRANGE, D. K. Survey on Evaluation of Student's Performance in Educational Data Mining. **2018 Second International Conference On Inventive Communication And Computational Technologies (Icicct)**, [S.L.], abr. 2018. DOI: <http://dx.doi.org/10.1109/icicct.2018.8473228>.
- BRAMER, M. **Principles of Data Mining**. 3. ed. Undergraduate topics in computer science, 2016.
- CAMPAGNI, R.; MERLINI, D.; VERRI, M. Finding regularities in courses evaluation with k-means clustering. **CSEDU**, 2014.
- CHUNG, E.; MCLARNEY, C. The classroom as a service encounter: suggestions for value creation. **Journal of Management Education**, vol. 24, n. 4, 2000.
- DELEN, D. A comparative analysis of machine learning techniques for student retention management. **Decision Support Systems**, v. 49, n. 4, p. 498-506, 2010. DOI: <https://doi.org/10.1016/j.dss.2010.06.003>
- DOL, S. M. Use of Classification Technique in Educational Data Mining. **2021 4Th Biennial International Conference On Nascent Technologies In Engineering (IcnTE)**, [S.L.], 15 jan. 2021. DOI: <http://dx.doi.org/10.1109/icnte51185.2021.9487739>.
- HAN, J.; KAMBER, M.; PEI, J. **Data Mining**: concepts and techniques. 3 ed. Waltham, MA, EUA: Morgan Kaufmann, 2012.
- HEWSON, E. R. F. Students' emotional engagement, motivation and behaviour over the life of an online course: reflections on two market research case studies. **Journal of Interactive Media in Education**, v. 2018, n. 1, 2018.

HIROTA, L. C. B. **O perfil motivacional de alunos de graduação em física**. 2018. Dissertação (Mestrado em Física Aplicada) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2018. DOI:10.11606/D.76.2018.tde-29102018-171213.

HOMMA JUNIOR, E. *et al.* Análise de perfil dos graduando por técnicas de clusterização. **Anais do Encontro Nacional de Engenharia de Produção**. out, 2022. ENEGEP 2022. DOI: 10.14488/ENEPEP2022\_TN\_ST\_384\_1906\_43651

HUANG, Z. Extensions to the K-Means algorithm for clustering large data sets with categorical values. **Data Mining and Knowledge Discovery 2**, p. 283-304, 1998.

HSIEH, C. M.; ESSEX, E. Measuring client satisfaction among older adults and families. In B. Berkman (Ed.) **Handbook of social work in health and aging**, p. 1009-1017. New York: Oxford University Press, 2006.

IGUAL, L.; SEGUÍ, S. **Introduction to Data Science**: a python approach to concepts, techniques and applications. Introduction to Data Science. Undergraduate Topics in Computer Science. [S. l.]: Springer International Publishing, 2017. DOI: [https://doi.org/https://doi.org/10.1007/978-3-319-50017-1\\_1](https://doi.org/https://doi.org/10.1007/978-3-319-50017-1_1).

KAUR, P.; SINGH, M.; JOSAN, G. S. Classification and prediction based data mining algorithms to predict slow learners in education sector. **Procedia Computer Science**, v. 57, p. 500-508, 2015.

KINDLEIN, K.; SCHWAIGER, M. Empowered students – service-oriented universities: how to manage higher education in the 21st century. **Developments In Marketing Science**: Proceedings of the Academy of Marketing Science, [S.L.], p. 323-323, 2015. Springer International Publishing. DOI:10.1007/978-3-319-10912-1\_107.

KUBAT, M. An introduction to Machine Learning. 2. ed. **Springer**. 2017. DOI: 10.1007/978-3-319-63913-0

LEE, S. Net Promoter Score. **Proceedings Of The 2018 Acm Siguccs Annual Conference**, [S.L.], v. 10, p. 63-64, 2018. ACM. DOI: 10.1145/3235715.3235752.

NOGUEIRA-MARTINS, L. A.; NOGUEIRA-MARTINS, M. C. F. Saúde mental e qualidade de vida de estudantes universitários. **Revista Psicologia, Diversidade e Saúde**, 7(3), p. 334-337, 2018. doi: 10.17267/2317-3394rpds.v7i3.2086

MAIOCHI, A. C. B.; BARBATO, C. N. Estratégia para fidelização de alunos de graduação: estudo de caso com alunos de Instituição de ensino superior particular da região metropolitana de Campinas. **Revista Intellectus**, 8. ed. n 20, 2012.

MEYER, C.; SCHWAGER, A. Understanding customer experience. **Harvard Business Review**, 2007.

PATIAS, N D. *et al.* Depression Anxiety and Stress Scale (DASS-21) - Short Form: adaptação e validação para adolescentes brasileiros. **Psico-USf**, [S.L.], v. 21, n. 3, p.

459-469, dez. 2016. FapUNIFESP (SciELO). DOI: <http://dx.doi.org/10.1590/1413-82712016210302>.

RÄTSCH, G. A brief introduction into Machine Learning. In: 21st **Chaos Communication Congress**. Germany: Friedrich Miescher Laboratory of the Max Planck Society [S.l.: s.n.], p. 16, 2004.

REICHHELD, F. F. The Ultimate Question 2.0: how net promoter companies thrive in a customer-driven world. **Harvard Business Review**, 2003.

ROMERO, C., VENTURA, S. Educational Data Mining: A survey from 1995 to 2005. **ScienceDirect**, v. 33, p. 135-146, 2007.

ROZA, F. S. **Aprendizagem de máquina para apoio à tomada de decisão em vendas do varejo utilizando registros de vendas**, Florianópolis, 2016.

SANTOS JUNIOR, J; da S.; REAL, G. C. M. A evasão na educação superior: o estado da arte das pesquisas no brasil a partir de 1990. **Avaliação: Revista da Avaliação da Educação Superior** [online], Campinas, v. 22, n. 2, p. 385-402, ago. 2017.

SILVEIRA, A. L. F. *et al.* Congresso Internacional, ABED de educação a distância (21º CIAED). **Análise de Perfil e Necessidades dos Alunos**: Fatores Determinantes para a Mediação da Aprendizagem na Tutoria da EAD Sebrae, Florianópolis, 2015.

SITUMORANG, S. H.; RINI, E. S.; MUDA, I. Customer experience, net emotional value and net promoter score on Muslim middle class women in Medan. **International Journal of Economic Research**. v. 14, n. 20, 2017.

SMOLA, A.; VISHWANATHAN, S. Introduction to machine learning. **Cambridge University**, UK, v. 32, p. 10, 2008.

TAN, Pang-Ning; STEINBACH, M.; KUMAR, V. Introduction to Data Mining. **Pearson Education**, Inc. 2006.

VELOSO, T. C. M. A; ALMEIDA, E. P. Evasão nos cursos de graduação da Universidade Federal de Mato Grosso, campus universitário de Cuiabá – um processo de exclusão. **Série-Estudos - Periódico do Programa de Pós-Graduação em Educação da UCDB**, 2013.

WITTEN, I. H.; FRANK, E. **Data Mining**: practical Machine Learning tools and techniques. 2. ed. San Francisco, 2005.

## **APÊNDICE A - Parecer do Comitê de Ética em Pesquisa (CEP)**



## PARECER CONSUBSTANCIADO DO CEP

### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Análise do perfil de estudantes de graduação baseada nas suas rotinas e percepções sobre a universidade

**Pesquisador:** BRUNO SAMWAYS DOS SANTOS

**Área Temática:**

**Versão:** 2

**CAAE:** 57956421.6.0000.0177

**Instituição Proponente:** Universidade Tecnológica Federal do Paraná Campus Dois Vizinhos

**Patrocinador Principal:** Financiamento Próprio

### DADOS DO PARECER

**Número do Parecer:** 5.458.136

#### Apresentação do Projeto:

**Título:** Análise do perfil de estudantes de graduação baseada nas suas rotinas e percepções sobre a universidade.

**Objetivo da pesquisa:**

Aplicar técnicas de aprendizagem de máquina (supervisionada e não supervisionada) para encontrar padrões ocultos e informações não triviais em conjuntos de dados de estudantes, analisando o seu perfil e as percepções sobre a experiência universitária.

#### Introdução:

Com o aumento crescente de tecnologias no mundo moderno, o sistema educacional está acompanhando este processo e se inovando cada vez mais. Uma forma que vem ganhando destaque, nesse sentido, é a aplicação de mineração de dados no âmbito da educação, conhecido como Educational Data Mining, ou EDM (ROMERO; VENTURA, 2010). O intuito é encontrar possíveis melhorias para o sistema educacional por meio da mineração de dados, entendendo os estilos de aprendizado dos estudantes para criar oportunidades de melhoria do desempenho acadêmico, descobrindo novos padrões com esta abordagem (FERNANDES et al., 2019).

A mineração de dados é uma das etapas mais importantes do processo de descoberta de conhecimento em bancos de dados, ou Knowledge Discovery in Databases (KDD). Este processo

**Endereço:** Estrada para Boa Esperança, km 04 - Zona Rural - Bloco G 10, sala 675  
**Bairro:** Área Rural **CEP:** 85.660-000  
**UF:** PR **Município:** DOIS VIZINHOS  
**Telefone:** (46)3536-8215 **E-mail:** coep-dv@utfpr.edu.br



Continuação do Parecer: 5.458.136

Declaração de concordância	Termo_Autorizacao_Instituicao_Assinado.pdf	23/03/2022 10:04:41	BRUNO SAMWAYS DOS SANTOS	Aceito
----------------------------	--	------------------------	-----------------------------	--------

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

DOIS VIZINHOS, 09 de Junho de 2022

---

**Assinado por:**  
**Edival Sebastião Teixeira**  
**(Coordenador(a))**

**Endereço:** Estrada para Boa Esperança, km 04 - Zona Rural - Bloco G 10, sala 675  
**Bairro:** Área Rural **CEP:** 85.660-000  
**UF:** PR **Município:** DOIS VIZINHOS  
**Telefone:** (46)3536-8215 **E-mail:** coep-dv@utfpr.edu.br

**APÊNDICE B - Questionário de pesquisa**



## Análise do perfil de estudantes de graduação baseada nas suas rotinas e percepções sobre a universidade

Seção 1 de 6

### Análise do perfil de estudantes de graduação baseada nas suas rotinas e percepções sobre a universidade

O presente questionário busca entender a sua percepção sobre a UTFPR, desde a sua estrutura geral, como também o departamento do curso de forma específica. Também gostaríamos de saber um pouco mais sobre o seu cotidiano de forma a entender a sua rotina com estudos e saúde que possam impactar no seu entendimento sobre a UTFPR, buscando identificar e entender melhor as suas demandas para uma experiência universitária mais prazerosa.

Todos os dados aqui coletados serão analisados para fins acadêmicos e pesquisa. O instrumento possui questões elaboradas pelos autores e adaptadas de outros questionários da literatura, devidamente identificados.

As questões são direcionadas às suas experiências como aluno(a) no modo PRESENCIAL, e o tempo estimado de preenchimento do questionário completo é de 10 minutos.

Por favor, antes de prosseguir para o questionário, leia com atenção as informações contidas no Termo de Consentimento Livre e Esclarecido (TCLE), disponível no link:

<https://drive.google.com/file/d/1hbw-C4XyA-rBH0sVpZbgcftT6loshirn/view?usp=sharing>

Seção 2 de 6

Informações gerais

Gostaríamos de saber um pouco sobre você e a sua situação aqui na UTFPR - Londrina

1. Curso de Graduação \*

- Engenharia de Produção
- Engenharia Mecânica
- Engenharia Química
- Engenharia Ambiental
- Engenharia de Materiais
- Licenciatura em Química
- Tecnologia em Alimentos

2. Qual foi o ANO e SEMESTRE em que você ingressou no curso? Por favor, responda da seguinte forma: ANO/SEMESTRE (por exemplo, 2017/1) \*

Texto de resposta curta

---

3. Em qual período do curso você está de acordo com o sistema? \*

- 1°
- 2°
- 3°
- 4°
- 5°
- 6°
- 7°
- 8°
- 9°
- 10°

4. Qual foi o seu modo de ingresso no curso?

- Transferência interna
- Transferência externa
- SISU (Sem estar cursando uma graduação anteriormente)
- SISU (Estando cursando uma graduação anteriormente)

5. Você entrou na UTFPR via cotas? \*

- Sim
- Não
- Prefiro não informar

Seção 3 de 6

Perfil sociodemográfico



Agora, gostaríamos de saber um pouco mais sobre você, como a sua idade, renda, entre outros.

6. Qual é a sua idade? \*

1. 16

2. 17

3. 18

.  
.  
.

33. 48

34. 49

35. 50

36. Mais que 50

7. Situação de moradia em Londrina \*

Moro com a minha família.

Moro sozinho.

Moro com alguns colegas.

Moro em república (muitos colegas).

Outros...

8. Você trabalha? \*

Não, e não dependo do trabalho para me manter.

Não, mas dependo do trabalho para me manter.

Sim, e dependo do trabalho para me manter.

Sim, mas não dependo do trabalho para me manter.

9. Quantas pessoas compõem o seu núcleo familiar? (Considerar aqueles com quem divide rendas ou recebe auxílio). \*

- Só eu
- 2
- 3
- 4
- 5
- Mais que 5

10. Qual é a sua renda familiar mensal? (Considere como o valor total do grupo da questão anterior). \*

- Nenhuma renda.
- Até 1 salário mínimo (até R\$ 1.212,00).
- De 1 a 3 salários mínimos (de R\$ 1.212,00 até R\$ 3.636).
- De 3 a 6 salários mínimos (de R\$ 3.636,00 até R\$ 7.272,00).
- De 6 a 9 salários mínimos (de R\$ 7.272,00 até R\$ 10.908,00).
- De 9 a 12 salários mínimos (de R\$ 10.908,00 até R\$ 14.544,00).
- De 12 a 15 salários mínimos (de R\$ 14.544,00 até R\$ 18.180,00).
- Mais de 15 salários mínimos (mais de R\$ 18.180,00).

11. Quantos filhos você tem? \*

- Nenhum
- 1
- 2
- 3
- 4
- 5 ou mais

12. Sexo: (Esta informação é necessária para fins estatísticos, uma vez que no Brasil os órgãos oficiais trabalham com este dado) \*

- Feminino
- Masculino

13. Com qual gênero você se identifica? (Cisgênero: pessoa que se identifica com o gênero que lhe foi atribuído no nascimento. Transgênero: pessoa que NÃO se identifica com o gênero que lhe foi atribuído no nascimento. Não Binário: pessoa que NÃO se identifica com as categorias binárias de gênero (homem/mulher) e/ou com representações relacionadas ao "masculino" e/ou "feminino"). \*

- Mulher Cisgênero
- Homem Cisgênero
- Homem Transgênero
- Mulher Transgênero/Travesti
- Não binário
- Outro
- Não Sei
- Prefiro não informar

14. Dentre as seguintes alternativas, com qual você se reconhece ou se identifica como de cor ou raça: \*

- Indígena
- Preto
- Pardo
- Branco
- Amarelo
- Prefiro não informar

15. Você já estudou em escola pública antes de entrar na UTFPR? \*

- Sim, estudei INTEGRALMENTE em escola pública
- Sim, estudei menos da metade do meu ensino em escola pública
- Sim, estudei mais da metade do meu ensino em escola pública
- Não, estudei INTEGRALMENTE em escola particular
- Prefiro não informar

Seção 4 de 6

Informações sobre rotinas de estudos e bem-estar físico e mental



Esta seção busca verificar o que você faz/pratica no seu cotidiano para cuidar do corpo e da mente.

16. Quantas horas por semana (em média) você pratica exercícios físicos? Obs.: incluindo exercícios de alongamento e mobilidade. \*

- 1. Não pratico
- 2. 30 minutos
- 3. 1 hora
- 4. 2 horas
- .
- .

20. 18 horas

21. 19 horas

22. 20 horas

17. Com qual frequência você medita? \*

- Não pratico meditação
- Até duas vezes por semana
- De três a seis vezes por semana
- Medito todos os dias

18. Você frequenta algum psicólogo ou algum psiquiatra?

- Sim
- Não

19. Como você considera a sua relação com a sua família?

- Não saudável
- Saudável
- Não tenho contato
- Prefiro não informar



### DASS 21 - Depression Anxiety and Stress Scale

Para a questão 20, todas as alternativas aqui foram adaptadas e traduzidas livremente do questionário DASS 21 - Depression Anxiety and Stress Scale. (Fonte: Psychology Foundation of Australia, disponível em: <http://www2.psy.unsw.edu.au/dass/>). O questionário DASS 21 é um conjunto de três escalas de autorrelato para medir os estados emocionais negativos de depressão, ansiedade e ansiedade. As adaptações aqui foram feitas para identificar os momentos difíceis em situações possíveis para o estudante.

20. Assinale quais itens a seguir se enquadra(m) nas suas experiências vividas durante a graduação. Pode assinalar quantas alternativas quiser.

- A minha mão sua na hora de fazer prova ou apresentar trabalho.
- Fico ansioso momentos antes de uma avaliação
- Durmo mal no dia antes de uma avaliação
- Sinto-me próximo de entrar em pânico durante uma avaliação
- Sinto dificuldades em respirar durante uma avaliação
- Tenho dificuldades em relaxar em uma avaliação
- Não sinto motivação pra estudar
- Não me sinto feliz na faculdade/universidade
- Não faço perguntas durante a aula por vergonha
- Não consigo estudar porque fico ansioso

21. Quantas horas por semana (em MÉDIA) você estuda FORA dos horários de aula?

- Não estudo fora dos horários de aula
- Até 2 (duas) horas por semana
- Até 4 (quatro) horas por semana
- Até 6 (seis) horas por semana
- Acima de 6 (seis) horas por semana

22. Quais tipos de materiais você costuma utilizar para estudar?

- Caderno de anotações
- Livros físicos
- Livros Digitais
- Textos encontrados na internet
- Vídeos
- Outros...

23. Onde você estuda com maior frequência?

- Universidade
- Casa
- Outro local

24. Como você prefere estudar?

- Sozinho
- Com colegas

Seção 5 de 6

Avaliação da Estrutura e Experiência de Aprendizagem (PRESENCIAL)



Nesta seção, faremos algumas perguntas acerca da estrutura que a UTFPR oferece para você durante o seu tempo dedicado para os estudos. Aqui, queremos identificar a sua percepção quanto aos itens que avaliam esta estrutura.

Aqui, a maioria das questões serão baseadas em escala, sendo dividida da seguinte forma:

- 1 → Muito ruim
- 2 → Ruim
- 3 → Regular
- 4 → Boa
- 5 → Muito boa

25. Como você avalia as salas de aula da UTFPR - Londrina? \*

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito boa

26. Como você avalia os laboratórios da UTFPR - Londrina? \*

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom

27. Como você avalia a didática dos professores da UTFPR - Londrina? \*

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom

28. Como você avalia a disponibilidade dos professores da UTFPR - Londrina? \*

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom

29. Como você avalia os métodos de avaliação dos professores da UTFPR - Londrina? \*

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom

30. Como você avalia o conhecimento do conteúdo ministrado pelos professores da UTFPR - Londrina \*

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom



37. Como você avalia a quadra de esportes da UTFPR - Londrina?

	1	2	3	4	5	
Muito ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom

38. Você recomendaria o seu curso na UTFPR - Londrina a um amigo(a) ou colega? \*

	0	1	2	3	4	5	6	7	8	9	10	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Com certeza

39. Você recomendaria a UTFPR - Londrina a um amigo(a) ou colega? \*

	0	1	2	3	4	5	6	7	8	9	10	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Com certeza

Caso você tenha interesse em receber os resultados desta pesquisa, por favor, informe o seu e-mail.

Texto de resposta curta

Seção 6 de 6

Muito obrigado por sua participação! ✕ ⋮

Descrição (opcional)