São José dos Campos - SP
LV SBPO
LV Simpósio Brasileiro de Pesquisa Operacional
São José dos Campos, SP - 6 a 9 de novembro de 2023

Clustering Heterogêneo como Ferramenta para a Análise da Percepção Docente sobre Estruturas Curriculares

Vinícius Costa Bulhões, Éverton Santi, Sérgio Queiroz de Medeiros
Escola de Ciências e Tecnologia - ECT/UFRN
Campus Universitário - Lagoa Nova, Natal - RN, 59078-970
viniciuscbulhoes@gmail.com, everton.santi@ufrn.br,
sergio.medeiros@ufrn.br

Miriam Karla Rocha
Programa de Pós-Graduação em Administração - PPGA/UFERSA
Centro de Engenharias - CE/UFERSA
R. Francisco Mota, 572 - Pres. Costa e Silva, Mossoró - RN, 59625-900
miriam.rocha@ufersa.edu.br

Luciano Ferreira
Escola de Administração da UFRGS
Rua Washington Luiz, 855 - Centro Histórico, Porto Alegre - RS, 90010-460
ferreira.luciano@ufrgs.br

RESUMO

Este trabalho apresenta uma aplicação para problemas de *clustering* baseados em múltiplas matrizes de dissimilaridades, cujos dados de entrada são obtidos pelo processo de *Card Sorting*. Este tipo de análise por agrupamentos visa (a) identificar segmentos de indivíduos com percepção similar em relação a como um conjunto de objetos deve ser categorizado e (b) qual estrutura de categorias melhor representa as opiniões dos indivíduos de um mesmo segmento. Para o estudo de caso apresentado, analisamos como o corpo docente de um curso de graduação percebe os relacionamentos entre suas diferentes componentes curriculares obrigatórias. A partir deste processo, alguns direcionamentos são fornecidos no sentido de contribuir com o processo de discussão e melhoria desta grade curricular à luz das Novas Diretrizes Curriculares das Engenharias.

PALAVRAS CHAVE. Análise por agrupamentos, Programação Matemática, Educação. PO na Administração Pública e Educação, Outras aplicações em PO

ABSTRACT

This work presents an application for clustering problems based on multiple dissimilarity matrices, where the input data is obtained through the Card Sorting process. This type of clustering analysis aims to (a) identify segments of individuals with similar perception regarding how a set of objects should be categorized and (b) determine the category structure that best represents the opinions of individuals within the same segment. For the presented case study, we analyze how the faculty of an undergraduate course perceives the relationships between its different mandatory curriculum components. Based on this process, some directions are provided to contribute to the discussion and improvement of the curriculum in light of the New Curricular Guidelines for Engineering.

KEYWORDS. Clustering. Mathematical Programming. Education. OR in Public Administration and Education, Other Applications in OR

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

1. Introdução

O suporte à tomada de decisão baseada em dados se apresenta como uma tarefa desafiadora ao se considerarem os avanços tecnológicos e o aumento que estes avanços trouxeram ao volume de dados disponíveis [Liu et al., 2021]. Cada vez mais as organizações se voltam para o uso de dados, inteligência artificial e *Analytics* como forma de manter ou melhorar seu desempenho [Korherr et al., 2022]. Neste contexto, a área de *Machine Learning* (*ML*), ou Aprendizagem de Máquina, ganha destaque como mecanismo para agregação de valor aos dados. *ML* é baseada no aprendizado de um modelo, podendo ser utilizada para classificação, regressão ou outras tarefas [Gambella et al., 2021]. *ML* pode ser considerada sob duas perspectivas distintas a partir de seu objetivo principal:

- Aprendizagem supervisionada: ajustar uma função que aproxime com margem de erro tolerável o relacionamento entre entradas e saídas, a fim de que se possa generalizar tal relacionamento;
- Aprendizagem não-supervisionada: compreender características e possíveis associações entre as amostras observadas, sem que para tal seja necessário uma resposta conhecida para o problema.

A formulação baseada em problemas de otimização, abordagem diretamente ligada à Pesquisa Operacional (PO), é uma das formas mais clássicas para a modelagem de problemas de aprendizagem não-supervisionada, como é o caso de problemas de *Clustering* ou Análise por Agrupamentos. Esta é uma ferramenta que pode ser utilizada para a categorização, sendo útil para a identificação de padrões ou regras que descrevem objetos ou fenômenos [Xu e Wunsch, 2005].

Clustering tem como objetivo identificar subconjuntos dentre um conjunto de amostras, de tal forma que os objetos mais similares entre si sejam designados ao mesmo agrupamento (*cluster*), enquanto que objetos dissimilares, por sua vez, são designados a *clusters* distintos [Hansen e Jaumard., 1997]. Aplicações em *clutering* são vislumbradas em diversas áreas do conhecimento, como agropecuária [Mallioris et al., 2023], saúde [Firouraghi et al., 2023], comportamento [Citerne et al., 2023], genética [Carrier et al., 2023], estudos sociais [Hmimssa, 2023], dentre outras.

Tradicionalmente, métodos de *clustering*, como por exemplo k-*means* [MacQueen, 1967], consideram como entrada uma única matriz de distâncias D entre os objetos a serem agrupados. Logo, caso haja mais de uma matriz de dissimilaridades, obtidas a partir das opiniões de m indivíduos, estes dados devem ser agregados para que tais métodos sejam utilizados. Este processo de agregação, por sua vez, gera a perda da informação originalmente contida nos dados. Este problema, ou limitação dos métodos tradicionais, levou à concepção de métodos que suportam o uso de múltiplas matrizes de distâncias, ou dissimilaridades, ao agrupar um conjunto de itens. Os trabalhos de Blanchard et al. [2012] e Santi et al. [2016] apresentam dois novos problemas de *clustering* com duplo propósito:

- Identificar segmentos de indivíduos com percepções similares sobre como objetos de um conjunto se relacionam e devem ser agrupados, e;
- Quais, dentre os objetos avaliados pelos indivíduos, estão relacionados, obtendo uma estrutura de categorias entre estes para cada segmento de indivíduos.

Blanchard et al. [2012] demonstraram seu modelo ao analisar como consumidores percebem o posicionamento de varejistas no mercado, enquanto que Santi et al. [2016] ilustraram a

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

aplicação de seu modelo avaliando como consumidores percebem o posicionamento de diferentes marcas de doces em relação aos produtos da concorrência. O modelo de Santi et al. [2016] demanda menor ajuste de parâmetros para ser utilizado e apresenta uma maior capacidade de recuperação de informação.

Ambos os métodos foram concebidos com base no processo de *Card Sorting*, ou Tarefa de Triagem (TT), para a coleta de dados. *Card Sorting* é um método simples e efetivo, que reflete naturalmente a forma como os participantes percebem a organização de um conjunto de objetos em categorias, podendo ser executada com menor esforço cognitivo se comparado à avaliação par-a-par ou outras técnicas disponíveis [Blanchard e Banerji, 2016; Blanchard et al., 2012].

O objetivo deste trabalho é demonstrar uma nova abordagem para analisar a percepção docente acerca da estrutura curricular de um curso de graduação, em especial a percepção sobre como as componentes curriculares estão relacionadas entre si. Para tanto é apresentada uma nova aplicação de *clustering* baseada em múltiplas fontes de dados oriundos do processo de TT, na qual o método proposto por Santi et al. [2016] é utilizado para analisar a percepção docente acerca da estrutura curricular atual do Bacharelado em Ciências e Tecnologia da Universidade Federal do Rio Grande do Norte, sendo este um curso de primeiro ciclo para quem pretende cursar engenharia. Desta forma, além de fornecer uma proposta inovadora na área de ensino, as informações obtidas podem ser úteis no processo de reformulação da grade curricular do curso. Tal reformulação ocorre para fins de acomodar novas exigências listadas pelas Diretrizes Curriculares Nacionais das Engenharias (DCNs das Engenharias) [MEC, 2019].

O restante deste texto está organizado como segue: a Seção 2 contextualiza a aplicação proposta, enquanto que a Seção 3 descreve os materiais e métodos utilizados. A Seção 4 apresenta os resultados desta pesquisa e sua discussão, enquanto que a Seção 5 discorre sobre as delimitações e direções futuras desta pesquisa.

2. As novas DCNs das Engenharias e a redução da evasão escolar

A Resolução $N^{\underline{o}}$ 2, publicada em 24 de Abril de 2019, pelo Ministério da Educação instituiu as Diretrizes Curriculares Nacionais do Cursos de Graduação em Engenharia (DCNs das Engenharias). De acordo com este documento, estas diretrizes deverão ser consideradas pelas Instituições de Educação Superior (IES) ao desenvolver e avaliar cursos de Engenharia no âmbito dos Sistemas de Educação Superior do país [MEC, 2019].

Dentre as demandas geradas a partir desta resolução, está a reformulação do Projeto Pedagógico dos Cursos (PPC), havendo necessidade de adaptar as estruturas curriculares às novas diretrizes em um prazo de três anos. Pretende-se remodelar os cursos de Engenharia de tal forma que haja maior flexibilidade em relação à grade curricular, com foco em competências e interdisciplinaridade.

O documento lista as competências a serem desenvolvidas para que o perfil dos egressos das Engenharias possa estar em consonância não só com os conteúdos necessários, mas também com as demandas da sociedade em suas múltiplas dimensões, além de formar um profissional com capacidade para absorver e utilizar novas tecnologias. As competências descritas no documento englobam: (a) visão holística e humanista; (b) atuação em inovação e empreendedorismo; (c) criatividade; (d) inter e transdisciplinaridade; (e) multiplicidade de dimensões em sua atuação e, (f) atuação isenta e comprometida com responsabilidade social e desenvolvimento sustentável.

Uma possível interpretação a partir da leitura das novas DCNs é o objetivo de tornar os cursos mais atrativos aos estudantes sem prejuízos à formação. As DCNs permitirão uma maior flexibilidade em termos de currículo, proporcionando ao discente contato mais amplo com problemas e tecnologias atuais, o que poderá melhorar a forma como o aluno percebe o curso. Neste sentido

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

é possível sugerir que tais mudanças na estrutura dos cursos poderão contribuir para melhorar os índices de evasão escolar caso estas sejam percebidas positivamente pelos discentes, estreitando seu vínculo com o curso e instituição. Baggi e Lopes [2011] destacam que além das desigualdades sociais, fatores que geram desinteresse pelo curso também podem estar por trás dos altos índices de evasão.

A evasão escolar é um problema recorrente no Brasil, especialmente em diferentes cursos de Exatas ou Engenharias. O estudo de Davok e Bernard [2016] analisou dados da Universidade Estadual de Santa Catarina no 2008 a 2010 relativos aos cursos de Ciências Exatas e da Terra. Como resultado, uma taxa de evasão média de 58,8% foi observada. A Universidade Federal de Juiz de Fora (UFJF), por sua vez, apresentou uma taxa de aproximadamente 42% em seu Bacharelado Interdisciplinar em Ciências Exatas em 2009 e 54% em 2010 [Ciribelli e Ferreira., 2014].

Dados incluídos no Plano Quadrienal da Escola de Ciências e Tecnologia da Universidade Federal do Rio Grande Norte (UFRN) [ECT, 2023] sugerem que a taxa de evasão do Bacharelado em Ciências e Tecnologia (BCT) apresentou uma média de 21% a partir de 2015. A série histórica disposta no documento, no entanto, contém taxas de evasão que superam este valor em outros anos, como por exemplo para o segundo semestre de 2012, com uma taxa de aproximadamente 63%.

O BCT figura como um dos maiores cursos da UFRN, tendo aproximadamente 100 docentes em seu quadro funcional e uma média de aproximadamente 590 novos alunos a cada semestre [ECT, 2023]. O Bacharelado é responsável pela formação discente nos anos iniciais de diversos cursos de Engenharia da UFRN, como Engenharia Ambiental, Biomédica, Computação, Materiais, Mecatrônica, Petróleo e Telecomunicações.

3. Materiais e Métodos

Para este estudo foram consideradas as etapas descritas a seguir, desde a coleta dos dados até a interpretação dos resultados.

3.1. Etapa 1 - Coleta de dados

Os docentes do BCT foram convidados via *email* institucional a responder a esta pesquisa a partir da ferramenta *kardSort* [Balachandran, 2023]. Esta ferramenta permitiu:

1. Aplicar um questionário prévio para a caracterização do público-alvo do estudo;

2. Aplicar uma Tarefa de Triagem aberta (ver Blanchard e Banerji [2016]), na qual os participantes deveriam estabelecer categorias a partir do conjunto das disciplinas obrigatórias (objetos) da grade curricular atual do BCT.

As disciplinas consideradas neste estudo são listadas a seguir, compreendendo as componentes obrigatórias para a formação generalista do curso. As disciplinas do primeiro semestre incluem Pré-Cálculo, Vetores e Geometria Analítica, Cálculo I, Química Geral, Práticas de Leitura e Escrita I e Ciência Tecnologia e Sociedade. As disciplinas do segundo semestre incluem Cálculo II, Álgebra Linear, Lógica de Programação, Introdução à Física Clássica I, Práticas de Leitura e Escrita II, Gestão e Economia da Ciência, Tecnologia e Inovação e Probabilidade e Estatística;

O terceiro semestre do curso contempla Cálculo III, Metodologia Científica e Tecnológica, Linguagem de Programação, Introdução à Física Clássica II, Práticas de Leitura e Escrita em Inglês, Meio Ambiente e Desenvolvimento Urbano e Física Experimental I. Por fim, o quarto semestre inclui Computação Numérica, Introdução à Física Clássica III e Física Experimental II.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

3.2. Etapa 2 - Aplicação de *clustering* heterogêneo para a segmentação do corpo docente

A partir da conclusão do processo de *Card Sorting*, foram calculadas as matrizes de dissimilaridades D^k $(k = 1, 2, \ldots, m)$ para cada um dos m indivíduos (docentes) participantes por meio do método de Takane [1980]. Cada uma destas matrizes contém os valores d_{ij}^k, os quais representam o grau de dissimilaridade para cada par de objetos (i, j) de acordo com a percepção de cada docente k. Por fim, c^k corresponde ao número de categorias criadas por cada docente k durante este processo.

A formulação do Problema de *Clustering* Heterogêneo (PCH) [Santi et al., 2016] é baseada nas seguintes variáveis de decisão:

- p^{kg}, cujo valor deverá ser 1 se o indivíduo k é designado ao segmento g, para todo $k = 1, 2, ..., m$ e para todo $g = 1, 2, ..., G$, e zero caso contrário, e;
- e_{ij}^g cujo valor deverá ser 1 se o objeto i for associado à mediana representada pelo objeto j, dentro da estrutura de categorias identificadas para o segmento g, para todo $g = 1, 2, ..., G$ e para todo par $i, j = 1, 2, ..., n$, e zero caso contrário.

Logo, tem-se o Problema de Programação Não-Linear Inteira:

$$\text{minimize } Z = \sum_{g=1}^{G} \sum_{k=1}^{m} p^{kg} \left[\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^k e_{ij}^g \right] \tag{1}$$

sujeito a:

$$\sum_{g=1}^{G} p^{kg} = 1, \ \ \forall\, k = 1, 2, ..., m \tag{2}$$

$$\sum_{j=1}^{n} e_{ij}^g = 1, \ \ \forall\, g = 1, 2, ..., G; \ \forall\, i = 1, 2, ..., n \tag{3}$$

$$e_{ij}^g \leq e_{jj}^g, \ \ \forall\, g = 1, 2, ..., G; \ \forall\, i, j = 1, 2, ..., n \tag{4}$$

$$\sum_{j=1}^{n} e_{jj}^g = \left\lfloor \frac{\sum_{k=1}^{m} c^k p^{kg}}{\sum_{k=1}^{m} p^{kg}} \right\rfloor, \ \ \forall\, g = 1, 2, ..., G \tag{5}$$

$$e_{ij}^g \in \{0, 1\}, \ \ \forall\, g = 1, 2, ..., G; \ \forall\, i, j = 1, 2, ..., n \tag{6}$$

$$p^{kg} \in \{0, 1\}, \ \ \forall\, k = 1, 2, ..., m; \ \forall\, g = 1, 2, ..., G. \tag{7}$$

Em que (1) minimiza a soma das dissimilaridades para as associações entre os n objetos, condicional à pertinência dos indivíduos em cada segmento g. O conjunto de restrições em (2) garante que cada indivíduo k será designado a um segmento g, enquanto que (3) determina que todo o objeto i será associado a outro objeto j para todo o segmento g. As desigualdades em (4) garantem que as associações entre dois objetos i e j somente serão possíveis quando o objeto j for uma mediana. As igualdades em (5) impõem que o número de categorias recuperado para um segmento g deverá ser igual ao piso do número médio de categorias criado pelos indivíduos designados a aquele segmento. Por fim, (6-7) são as restrições de domínio sobre os valores das variáveis de decisão.

Para resolver o problema (1-7), o algoritmo de Busca em Vizinhança Variável Generalizado (*GVNS*), também proposto por Santi et al. [2016], foi aplicado.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

3.3. Etapa 3 - Análise e interpretação dos resultados

Para esta etapa, os resultados gerados pelo algoritmo foram tabulados utilizando planilhas eletrônicas para fins de visualização e leitura. Os resultados gerados pelo modelo não possuem uma resposta correta a ser recuperada, portanto a interpretação dos resultados foi conduzida de forma subjetiva considerando o conhecimento dos autores sobre o curso e a rotulagem fornecida pelos participantes na TT. Nesta etapa, analisaram-se as soluções geradas e os demais dados coletadas para:

1. Compreender quantas e quais formas de percepção (padrões) existem em relação ao conjunto de disciplinas categorizadas pelos docentes, e;

2. Se há características relativas aos respondentes que sugerem o porquê das diferentes formas de percepção;

4. Resultados e Discussão

O BCT possui um quadro docente com aproximadamente 100 professores. Destes, 29 responderam à pesquisa, sendo uma destas observações descartada. Os resultados gerados pelo tratamento dos dados são apresentados e discutidos nas seções que seguem.

4.1. Resolução do PCH

O algoritmo *GVNS* utilizado foi executado com tempo limite de 10 minutos, dado que os autores deste demonstraram sua eficiência para instâncias 10 vezes maiores que a aqui considerada neste mesmo tempo de execução. O computador utilizado possui 6 processadores Intel(R) Xeon(R) Gold 5120 CPU @ 2.20GHz e 6GB de memória. Foram conduzidas 10 rodadas de resolução para cada um dos valores possíveis para G, variando de 1 a 10. A Tabela 1 apresenta o resumo destes resultados.

G	Custo Médio	Desvio Padrão	Custo Mínimo
1	512,96	0,00	512,96
2	478,12	0,00	478,12
3	473,50	0,00	473,50
4	471,45	0,05	471,43
5	469,66	0,07	469,62
6	467,86	0,30	467,56
7	466,36	0,31	465,99
8	465,18	0,30	464,60
9	463,46	0,33	463,07
10	462,28	0,37	461,74

Tabela 1: Custo das soluções geradas. Elaboração própria.

Os dados presentes na Tabela 1 mostram que o algoritmo gerou soluções de forma estável, pois o desvio-padrão em relação ao custo da função objetivo é pequeno para todas as instâncias resolvidas. Quando da proposição do algoritmo, Santi et al. [2016] demonstraram que este era robusto mesmo para instâncias até 12 vezes maiores que a considerada neste estudo. Logo, pôde-se utilizar a melhor solução obtida para cada segmento G para fins de interpretação, com alguma segurança quanto à qualidade do resultado.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

A Figura 1 ilustra o comportamento da função objetivo em relação a cada mudanças no número de segmentos G testada. Sabe-se que quanto maior o número de segmentos G menor será o custo, pois estes segmentos serão formados com menor heterogeneidade entre os indivíduos. Sabe-se também que para valores de G superiores ao número de padrões que os dados contém, o custo da função objetivo não reduzirá de forma expressiva, pois haverá a formação de segmentos com poucos ou único indivíduo.

Verifica-se, portanto, que a solução mais promissora é a solução para $G = 2$, dado que o custo da solução não diminui de forma significativa a partir deste ponto (regra do "cotovelo"). A redução do custo observada é de 6,79% em relação a $G = 1$. Isto é, o modelo sugere que há duas formas distintas de perceber os relacionamentos entre os itens agrupados pelo participantes.

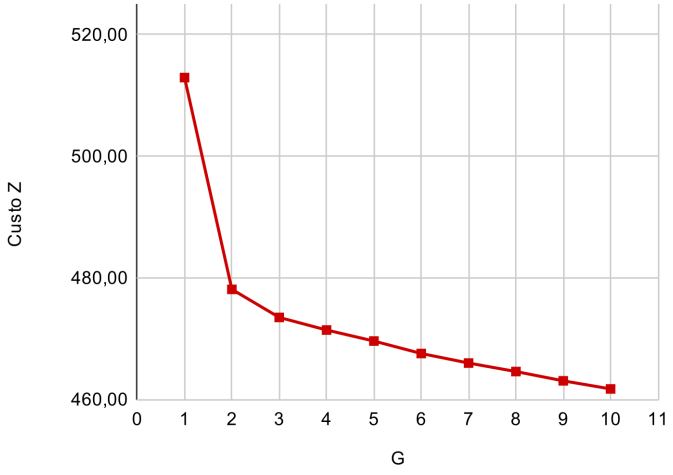

Figura 1: Número de segmentos *versus* custo da solução. Elaboração própria.

4.2. Discussão

Os pontos discutidos nesta seção provêm de uma análise subjetiva das estruturas de categorias recuperadas pelo modelo e com base nos dados do questionário aplicado. Há de se considerar que são possíveis diferentes interpretações para estes mesmos resultados, portanto há espaço para discussão em direções divergentes.

A estrutura de categorias recuperada pelo modelo para dois segmentos é mostrada na Tabela 2. Observa-se que as duas estruturas de categorias geradas diferem quanto à forma de organizar as disciplinas. Para o Segmento 1, 12 docentes foram identificados com percepção similar à visão representada, enquanto que para o Segmento 2, 16 docentes foram identificados. Para o Segmento 1 há uma categorização baseada sobretudo na área em que as disciplinas se encaixam. Para o segundo segmento, as disciplinas foram agrupadas considerando a ideia de sequência, ou nível de complexidade. Cada segmento é analisado a seguir.

4.3. Segmento 1: percepção orientada por Câmara Temática

Ao todo 12 docentes foram associados à percepção identificada neste segmento. Como sugerido anteriormente, a estrutura de categorias recuperada pelo modelo mostra que as disciplinas estão organizadas de acordo com a área. Esta categorização faz sentido ao considerar a forma como o BCT está organizado, pois cada docente faz parte de uma Câmara Temática (CT), que por sua vez é responsável por um conjunto restrito de disciplinas. Dentre as CTs estão: Ciência, Tecnologia e

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

Segmento 1 (12 docentes) - Percepção por Câmara Temática

Categoria	Mediana	Componentes
Informática	*Linguagem de Programação*	Computação Numérica Lógica de Programação
Formação Complementar / Aplicações em C&T	*Meio Ambiente e Desenvolvimento Urbano*	Gestão da Economia, Ciência, Tecnologia e Inovação Química Geral Ciência Tecnologia e Sociedade
Física	*Física Experimental I*	Introdução à Física Clássica III Introdução à Física Clássica II Introdução à Física Clássica I Física Experimental II
Complementar / Desenvolvimento Humano	*Práticas de Leitura e Escrita II*	Práticas de Leitura e Escrita em Inglês Metodologia Científica e Tecnológica Práticas de Leitura e Escrita I
Matemática	*Vetores e Geometria Analítica*	Probabilidade e Estatística Álgebra Linear Cálculo II Cálculo III Cálculo I Pré-Cálculo

Segmento 2 (16 docentes) - Percepção por Complexidade

Categoria	Mediana	Componentes
Formação avançada	*Introdução à Física Clássica III*	Computação Numérica Física Experimental II Introdução à Física Clássica II Física Experimental I Cálculo III Probabilidade e Estatística Cálculo II Química Geral
Formação complementar	*Meio Ambiente e Desenvolvimento Urbano*	Metodologia Científica e Tecnológica Gestão da Economia, Ciência e Inovação Práticas de Leitura e Escrita em Inglês Práticas de Leitura e Escrita II Práticas de Leitura e Escrita I
Formação básica	*Cálculo I*	Linguagem de Programação Introdução à Física Clássica I Vetores e Geometria Analítica Lógica de Programação Pré-Cálculo

Tabela 2: Estruturas de categorias recuperadas. Elaboração própria

Sociedade (3 docentes), Física (12 docentes), Informática (19 docentes), Matemática (24 docentes), Meio Ambiente (4 docentes), Negócios Tecnológicos (6 docentes), Práticas de Leitura e Escrita (6 docentes), Química (5 docentes) e Tecnologia (13 docentes).

Ao considerar a composição das categorias é possível observar que as disciplinas sob a responsabilidade das CTs de Informática, Física e Matemática são percebidas de forma muito clara em relação às suas câmaras. De forma distinta, nas categorias em que as medianas são Práticas de Leitura e Escrita II e Meio Ambiente e Desenvolvimento Urbano, a visão dos docentes sugere uma interação de áreas.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

Observando os dados brutos da TT, especialmente os rótulos atribuídos pelos participantes às categorias, pode-se sugerir que Física, Matemática e Informática são percebidas como a base para a formação discente, especialmente relacionadas a conteúdos. As demais categorias são vistas como uma espécie de formação complementar. Estas seriam responsáveis por expandir a visão discente em relação às aplicações de sua formação básica.

Esta observação é interessante ao se considerar as mudanças demandadas pelas novas DCNs das Engenharias. A formação interdisciplinar e atenta às demandas sociais e econômicas, bem como ao desenvolvimento das diversas competências, e não só o aprendizado de conteúdos, é uma das maiores mudanças a ser implementada. Neste sentido, estas disciplinas complementares já existentes no curso desenvolvem um trabalho que converge em direção à esta nova proposta.

Pode-se considerar a expansão desta formação complementar, ou mesmo uma maior integração desta formação com as CTs de Matemática, Física e Informática. Estas três CTs tendem a apresentar maiores taxas de retenção em suas disciplinas. A formação complementar poderia atuar de forma a auxiliar os alunos na percepção do quão importante estas três áreas são, no intuito de aumentar seu interesse por elas a partir da exploração de problemas e projetos cuja implementação dependam de conhecimentos mais aprofundados nestas áreas.

Em relação ao questionário aplicado, pode-se sugerir que os docentes deste segmento em sua maioria julgam que a grade curricular atual é pouco adequada. Estes também gostariam de poder ministrar mais disciplinas optativas e sugerem que o curso não deveria aumentar sua quantidade de disciplinas obrigatórias. Este pensamento, de certa forma, converge em relação às novas DCNs, pois daria maior flexibilidade para o aluno direcionar sua formação com base em seus interesses.

4.4. Segmento 2: percepção orientada por nível de complexidade

A estruturação em três categorias para as componentes curriculares obrigatórias que é percebida nesse segmento fornece uma direção interessante em termos de reformulação do curso, especialmente em relação à formação generalista que este permite. Ao considerar a categoria rotulada como "Formação Básica", pode-se vislumbrar um possível desenho do que seriam os semestres iniciais do curso.

Disciplinas como Pré-Cálculo, por exemplo, tem o propósito de nivelar os estudantes e consolidar conhecimentos básicos do Ensino Médio para que haja um maior proveito dos discentes ao cursar as demais componentes, especialmente aquelas relacionadas à Matemática. Cálculo I é a mediana desta categoria e sua taxa de sucesso está diretamente relacionada ao êxito em Pré-Cálculo. É possível sugerir que as disciplinas desta categoria permaneçam como obrigatórias em uma nova proposta de grade curricular, pois teriam um propósito de formação mínima em termos de conteúdos, habilitando o estudante a se aprofundar em tópicos mais avançados e de seu interesse.

Discussões acerca da reformulação do curso estão em concordância com esta visão, pois em uma nova proposta que será implementada as disciplinas de Lógica de Programação e Linguagem de Programação, por exemplo, estão sendo novamente inseridas no primeiro e segundo semestre do curso, respectivamente.

De maneira similar ao Segmento 1, pode-se perceber uma categoria com disciplinas de "Formação Complementar", com foco em expandir a visão discente em relação aos assuntos que relacionam sua atuação aos conteúdos. Esta categoria, como se pode observar, compreende a fusão das categorias de formação complementar do Segmento 1. Estas componentes são de fato pertinentes para estreitar o vínculo discente com o curso, pois nelas há a possibilidade de discutir temas e problemas relevantes da sociedade de forma mais flexível em termos de conteúdos se comparado às disciplinas de formação básica.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

Quanto à categoria rotulada como "Formação Avançada", tem-se disciplinas que permitem ao discente um aprofundamento maior dos conteúdos das diversas áreas. Como o bacharelado tem uma formação generalista e também permite que o aluno direcione sua formação para outras áreas específicas, pode-se flexibilizar a obrigatoriedade de algumas destas disciplinas a depender da ênfase pretendida, algo que vai de encontro às novas DCNs das Engenharias. Discussões preliminares entre os colegiados dos cursos relacionados a este currículo apontam para esta flexibilização. Um exemplo disto seria a flexibilização da obrigatoriedade da disciplina de Computação Numérica. Caso o aluno queira se formar como Bacharel em Ciências e Tecnologia (generalista) esta disciplina não faria parte de seu currículo obrigatório. Porém, para o aluno do BCT que pretende cursar Engenharia da Computação, esta disciplina ainda poderia permanecer como obrigatória dado que fosse compreendida como essencial para permitir ao aluno uma formação específica.

Em sua maior parte os docentes designados a este segmento julgam que a grade curricular atual é adequada. Em comparação com o Segmento 1, estes docentes atuam com menor frequência em disciplinas optativas do curso. Porém consideram, assim como o Segmento 1, que deveria haver uma menor quantidade de disciplinas obrigatórias, expressando também o desejo de atuar em optativas com mais regularidade. Uma boa parte dos docentes deste segmento destaca que atua em turmas grandes, com mais de 90 alunos, com certa regularidade. Isto ajuda a explicar o desejo de reduzir as componentes obrigatórias, pois isto possibilitaria aos docentes também direcionarem esforços em áreas mais sintonizadas com seu perfil acadêmico. Isto seria benéfico no sentido de qualificar mão de obra para atuação em pesquisa, podendo acomodar áreas não cobertas por disciplinas obrigatórias.

5. Considerações finais

Este trabalho demonstrou o uso de *Card Sorting* e métodos de *clustering* baseados em múltiplas matrizes de dissimilaridades como ferramenta de suporte ao processo de análise da percepção da grade curricular de um curso de graduação. A aplicação mostrada difere dos usos comumente associados a estas abordagens, especialmente ao *Card Sorting*, que tradicionalmente é aplicado em problemas relacionado a *Marketing*, Psicologia ou mesmo em projetos de experiência de usuário (UX).

A análise apresentada neste trabalho, por sua vez, é limitada. O número de docentes que responderam à pesquisa é pequeno quando comparado ao tamanho do corpo docente que o Bacharelado em Ciências e Tecnologia possui. Adicionalmente, dada a natureza subjetiva da análise apresentada, deve-se ter cautela ao generalizar quaisquer observações, pois outras formas de interpretar as estruturas de categorias geradas seriam possíveis com base em conhecimentos e experiências distintas.

No entanto, a abordagem utilizada para a análise apresentada, com base em um método de *clustering* alternativo aos métodos tradicionais, permite verificar diferentes pontos de vista sobre como o curso está organizado ou mesmo como poderia ser repensado. Não necessariamente o resultado do modelo direciona sobre o que deveria ser feito para que o curso seja adequado às novas DCNs. O resultado aqui produzido pode auxiliar no processo de entender como a grade curricular é percebida atualmente pelos indivíduos que fazem parte da comunidade acadêmica.

Considerando que um dos objetivos das novas DCNs é tornar os cursos de Engenharia mais atrativos aos discentes, melhorando a maneira como estes percebem a relação entre sua formação, o mercado e a sociedade, analisar como este grupo percebe a grade curricular é fundamental. No entanto, uma tentativa de coleta de dados junto a este público não obteve sucesso.

Embora seja complexo determinar o motivo pelo qual os discentes não se dispuseram a responder à pesquisa, algumas possibilidades podem ser levantadas, como o período de transição

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

entre o regime de Ensino Remoto e o Regime Presencial, ocasionado pela pandemia da Covid 19. Uma nova tentativa de coleta de dados poderá ser conduzida futuramente.

Referências

Baggi, C. A. S. e Lopes, D. A. (2011). Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica. *Avaliação (Campinas)*, 16(2):355–374.

Balachandran, K. (2023). kardsort. https://kardsort.com/. Acessado: 25/05/2023.

Blanchard, S. J. e Banerji, I. (2016). Evidence-based recommendations for designing free- sorting experiments. *Behavior Research Methods*, 48(4):1318–1336.

Blanchard, S. J., Aloise, D., e DeSarbo, W. S. (2012). The heterogeneous p-median problem for categorization based clustering. *Psychometrika*, 77(4):741–762.

Carrier, A., Gilbert, I., Leclerc, P., Duchesne, M., e Robert, C. (2023). Characterization of the genetic pool of the canadienne dairy cattle breed. *Genetics Selection Evolution*, 55(1). Cited by: 0.

Ciribelli, B. C. d. N. e Ferreira., C. d. S. (2014). Retenção e evasão escolares no bacharelado interdisciplinar em ciências exatas da universidade federal de juiz de fora. *Gestão e avaliação*, 4 (2).

Citerne, A., Roda, C., Rancière, F., e Momas, I. (2023). Subgroups of perceptions and related behaviors during the covid-19 lockdown: experience of adolescents in the paris birth cohort. *Child and Adolescent Psychiatry and Mental Health*, 17(1). Cited by: 0; All Open Access, Gold Open Access.

Davok, D. F. e Bernard, R. P. (2016). Avaliação dos índices de evasão nos cursos de graduação da universidade do estado de santa catarina - udesc. *Avaliação (Campinas)*, 21(2):503–521.

ECT (2023). Plano quadrienal da escola de ciências e tecnologia 2023-2027. https://www.ect.ufrn.br/. Acessado: 25/05/2023.

Firouraghi, N., Bergquist, R., Fatima, M., Mohammadi, A., Hamer, D. H., Shirzadi, M. R., e Kiani, B. (2023). High-risk spatiotemporal patterns of cutaneous leishmaniasis: a nationwide study in iran from 2011 to 2020. *Infectious Diseases of Poverty*, 12(1). Cited by: 0.

Gambella, C., Ghaddar, B., e Naoum-Sawaya, J. (2021). Optimization problems for machine learning: A survey. *European Journal of Operational Research*, 290(3):807–828. ISSN 0377-2217.

Hansen, P. e Jaumard., B. (1997). Cluster analysis and mathematical programming. *Mathematical Programming*, 79(1):191–215.

Hmimssa, A. (2023). Citizenship under pressure: Mapping debates over religious freedom in the second decade in quebec. *Journal of Canadian Studies*, 57(1):114 – 135. Cited by: 0.

Korherr, P., Kanbach, D. K., Kraus, S., e Mikalef, P. (2022). From intuitive to data-driven decision-making in digital transformation: A framework of prevalent managerial archetypes. *Digital Business*, 2(2):100045. ISSN 2666-9544.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891

Liu, A., Miller, W., Cholette, M. E., Ledwich, G., Crompton, G., e Li, Y. (2021). A multi-dimension clustering-based method for renewable energy investment planning. *Renewable Energy*, 172: 651–666. ISSN 0960-1481.

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1:281–297.

Mallioris, P., Teunis, G., Lagerweij, G., Joosten, P., Dewulf, J., Wagenaar, J. A., Stegeman, A., e Mughini-Gras, L. (2023). Biosecurity and antimicrobial use in broiler farms across nine european countries: Toward identifying farm-specific options for reducing antimicrobial usage. *Epidemiology and Infection*, 151.

MEC (2019). Resolução 2, de 24 de abril de 2019. http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=112681-rces002-19&category_slug=abril-2019-pdf&Itemid=30192. Acessado: 25/05/2023.

Santi, E., Aloise, D., e Blanchard, S. J. (2016). A model for clustering data from heterogeneous dissimilarities. *European Journal of Operational Research*, 253(3):659–672.

Takane, Y. (1980). Analysis of categorizing behavior by a quantification method. *Behaviormetrika*, 7:75–86.

Xu, R. e Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678.

DOI: 10.59254/sbpo-2023-174891
http://live-proceedings.pantheonsite.io/p/174891