

DESCOBERTA DE CONHECIMENTOS SOBRE INTEGRAÇÃO CURRICULAR NOS ESTADOS DA REGIÃO SUL DO BRASIL POR MEIO DO ENADE 2012: UM ESTUDO UTILIZANDO MINERAÇÃO DE DADOS EDUCACIONAIS

Patricia. B. Estivalet*
Emerson. B. Estivalet**
Eliseo. B. Reategui***

RESUMO

O presente artigo tem como objetivo investigar a integração curricular em cursos de graduação, por meio de uma pesquisa nos microdados disponibilizados na base de dados do INEP, referente ao Enade 2012. As respostas dos estudantes foram classificadas conforme o algoritmo de Mineração de Dados J48, para geração da Árvore de Decisão, e interpretados segundo os princípios da Análise de Conteúdo. Foram identificadas 11 228 instâncias classificadas, sendo 10 919 (97,24%) na categoria A (currículo bem integrado), 309 (2,7%) na categoria B (relativamente integrado) e nas categorias C (pouco integrado) e D (não apresenta integração) não houve marcação. Pode-se inferir que a pergunta não foi entendida pelos estudantes, gerando dados que não correspondem às realidades dos cursos.

PALAVRAS-CHAVE: mineração de dados educacionais, integração curricular, ENADE.

DESCUBRIMIENTO DE CONOCIMIENTOS SOBRE LA INTEGRACIÓN CURRICULAR EN LOS ESTADOS DEL SUR DE BRASIL POR MEDIO DE ENADE 2012: UN ESTUDIO UTILIZANDO LA MINERÍA DE DATOS EDUCATIVOS

RESUMEN

Este artículo tiene como objetivo enterarse del conocimiento acerca de la integración curricular en las carreras de pregrado, por medio de una búsqueda de microdatos disponibles en la base de datos de INEP, referidos en Enade 2012. Las respuestas de los estudiantes fueron clasificadas según el algoritmo de Data Mining J48, para la generación del Árbol de Decisión e interpretadas de acuerdo a los principios del Análisis de Contenido. Se identificaron un total de 11 228 instancias clasificadas, con 10 919 (97,24%) en la categoría A (currículo bien integrado), 309 (2,7%) en la categoría B (relativamente integrado) y en las categorías C (poco integrado) y D (no presenta integración) no hubo acotación. Se puede inferir que la pregunta no fue entendida por los estudiantes, generando datos que no corresponden a las realidades de los cursos.

PALABRAS CLAVE: minería de datos educativos, integración curricular, ENADE.



1 INTRODUÇÃO

O Sistema Nacional de Avaliação do Ensino Superior (SINAES), criado pela Lei n. 10 861, de 14 de abril de 2004, está subordinado ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), cuja missão é promover estudos, pesquisas e avaliações sobre o Sistema Educacional Brasileiro com o objetivo de subsidiar a formulação e a implementação de políticas públicas para a área educacional.

Para isso, conta com instrumentos de aferição de qualidade, como o Exame Nacional de Desempenho dos Estudantes (Enade), que visa acompanhar o processo de aprendizagem e o desempenho acadêmico dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares dos seus respectivos cursos de graduação, suas habilidades e competências no âmbito da profissão que se ligam à realidade brasileira e mundial e a outras áreas do conhecimento. Vale destacar que tal exame, em observância à sua lei de criação, passou a fazer parte do currículo dos cursos superiores como componente obrigatório. Além dos conhecidos indicadores de qualidade dos cursos, publicados como resultado de cada processo de avaliação, o exame gera um conjunto de dados de cada estudante, dos cursos e das instituições de Educação Superior avaliados, isto é, os microdados.

Neste estudo, serão utilizados os microdados disponibilizados pelo INEP referente ao Enade 2012 (Inep, 2015), especificamente as respostas dos acadêmicos para a questão número 44, que trata sobre a integração do currículo do curso em que estavam se formando. Nesse ano, os cursos avaliados foram: bacharelados em Administração, Ciências Contábeis, Ciências Econômicas, Comunicação Social, Design, Direito, Psicologia, Relações Internacionais, Secretariado Executivo e Turismo; e tecnólogos em: Gestão Comercial, Gestão de Recursos Humanos, Gestão Financeira, Logística, Marketing e Processos Gerenciais.

Escolheu-se o exame de 2012 por ter sido este o ano mais recente no qual o tema integração curricular foi abordado por meio da seguinte pergunta: «Como você avalia o currículo do seu curso em relação à integração entre os conteúdos das diferentes disciplinas?». Vale ressaltar que não há um padrão de perguntas no Enade em todos os anos em que é realizado.

Para que seja possível compreender o que significa currículo integrado ou integração curricular, é necessário fazer aqui um breve marco teórico.

Desde a Antiguidade, diferentes civilizações vêm demonstrando a preocupação com as formas de se organizar o currículo para que se possa atender às demandas sociais que justificam sua existência. Assim, encontra-se em Santomé (1998) o relato dos primeiros modelos de organização curricular compostos por discipli-

* Universidade Federal do Rio Grande do Sul. UFRGS / Instituto Federal Catarinense. IFC, Rio do Sul. SC/BR, patricia.estivalete@ifc.edu.br.

** Instituto Federal Catarinense. IFC, Rio do Sul. SC/BR, emerson.estivalete@ifc.edu.br.

*** Universidade Federal do Rio Grande do Sul. UFRGS, Porto Alegre. RS/BR, eliseo-reategui@gmail.com.

nas: *Trivium* (gramática, retórica e dialética) e *Quadrivium* (aritmética, geometria, música e astronomia), percebendo-se a preocupação em buscar a melhor forma de organizar os conhecimentos que necessitavam ser aprendidos.

Os estudos de Goodson (1995) e de Pacheco (2005) demonstram que o conceito de currículo como sequência estruturada, com as características que se conhece hoje, passou a ganhar força e projeção política no contexto da Reforma Religiosa, especialmente no âmbito do Calvinismo europeu, em que passou a ser negociado, reconstruído e aprimorado com o propósito de atender aos interesses da burguesia emergente, preocupada em acumular capital. Esse processo se desdobrou tanto nos espaços educativos formais quanto fora deles. Assim, em 1663, o termo «currículo» foi dicionarizado. Goodson (1995, p. 31) esclarece que «a palavra *currículo* vem do latim *scurrere*, correr, e refere-se a curso/percurso. As implicações etimológicas são que, com isso, o currículo seja um curso a ser seguido, ou especificamente apresentado».

Nessa perspectiva, o currículo pode ser entendido como um processo dinâmico que se constrói lentamente na sociedade, no enfrentamento de conflitos e interações sociais. Nessa construção, o ontem e o hoje estão em constante confronto, formando o que Elias (2011) chama de teia social, em que se encontram as mudanças e as permanências do processo civilizador que vão provocando questionamentos e reformulações do conhecimento a ser utilizado no presente e no futuro.

Considerando essa perspectiva, o currículo deve ser entendido como um importante elemento que orienta o desempenho das atividades docentes e discentes, organizando os conhecimentos a serem trabalhados ao longo de todo o processo educativo, visto que o desenvolvimento curricular consiste em uma dinâmica de ações contínuas cujas fases requerem determinada articulação entre si.

Na mesma linha de raciocínio, pode-se encontrar em Santomé (1998) uma das preocupações centrais da organização curricular sob uma perspectiva integrada: a valorização das tomadas de decisão em conjunto. Essa forma articulada de ensinar e aprender pode possibilitar a humanização do conhecimento, tornando-o mais significativo. Isso porque reconhecemos que o processo pedagógico é proveniente de objetos concretos, em que as pessoas envolvidas aparecem como sujeitos participantes e desenvolvem capacidades para intervir na história umas das outras.

Considerando os conceitos até agora apresentados, é necessário que se reconheça o importante papel da universidade e suas interconexões. Ela tem a ciência e os saberes nas mais diversas áreas do conhecimento como mote de discussão, no sentido da transformação social que pode oferecer condições potencialmente capazes de gerar transformações que venham ao encontro das demandas do mundo do trabalho, tanto no âmbito das questões produtivas como sociais e ambientais.

Deve-se reconhecer que o currículo é algo muito mais rico que um rol de disciplinas a vencer, não sendo somente um instrumento que sistematiza finalidades, objetivos e estratégias de ensino. É, sem dúvida, um complexo sistema de conhecimento, compartilhado socialmente, que pode assumir diversas formas, inclusive com tecnologias digitais. É fruto da concepção de educação de uma sociedade, a qual seus representantes, de acordo com suas especificidades, elaboram, orientam e participam da sua existência, utilizando várias formas de se comunicar que podem acontecer para além do encontro presencial.



Isso posto, é relevante observar que o acadêmico precisa compreender seu percurso formativo como um caminho para ampliar seus sentidos e seus significados para a vida e para o mundo do trabalho. Da mesma forma, deve perceber pontos de articulação entre os saberes das disciplinas no intento de superar modelos tradicionais de interpretação de problemas e da própria realidade, acolhendo para si o desafio de estar em constante processo de aprendizagem, ora mais próximo da prática, ora buscando luz na teoria. Essa ideia será posteriormente retomada na discussão dos achados, tendo como base os dados minerados.

2 OBJETIVOS

Partindo da pergunta apresentada na introdução, definiram-se os seguintes objetivos: 1) investigar novos conhecimentos, com base nas respostas dos acadêmicos concluintes, acerca da integração dos conteúdos das diferentes disciplinas do curso, utilizando para isso a tarefa de classificação, por meio da técnica de Árvore de Decisão; 2) analisar se os saberes teóricos implicados na compreensão da pergunta têm relação com as respostas e com os perfis formativos dos concluintes.

3. A MINERAÇÃO DE DADOS E O PERCURSO METODOLÓGICO

Com o propósito de atender ao primeiro objetivo, recorreu-se à Mineração de Dados Educacionais (MDE), que tem a intenção de desenvolver métodos para investigar conjuntos de dados coletados em cenários educativos. Segundo Baker, Isotani e Carvalho (2011), por meio dos métodos de mineração de dados é possível compreender como se desenvolve o aprendizado do estudante e o papel do contexto no processo de aprendizagem. Ademais, os autores abordam que é possível identificar em que situação um tipo de método de ensino proporciona melhores benefícios para o estudante.

Para realizar a mineração, é preciso passar por várias etapas que compreendem o Processo de Descoberta de Conhecimento em Bases de Dados (KDD), como: seleção de dados, pré-processamento, transformação dos dados, mineração dos dados e interpretação (Tan, Steinbach, & Kumar, 2009). Na etapa de mineração, Baker (2010) apresenta uma taxonomia dos principais métodos em MDE, tais como: predição, agrupamento, mineração de relações, destilação de dados para facilitar decisões humanas e descobertas com modelos.

Na área de predição, foco deste estudo, o objetivo é desenvolver modelos que identifiquem variáveis preditivas, por meio da análise e combinação dos diversos aspectos encontrados nos dados. Para isso, dentre as várias tarefas existentes na predição, a classificação tem a intenção de identificar a classe de determinado registro. Nessa tarefa, um conjunto de dados é utilizado para gerar um modelo preditivo capaz de classificar novos registros (Camilo & Silva, 2009).

A tarefa de classificação por Árvore de Decisão funciona como um fluxograma em forma de árvore, em que cada nó representa um atributo da base de dados.



As ligações entre os nós representam diferenças nos valores dos atributos, e as folhas indicam a classe à qual o registro pertence. A Árvore de Decisão é uma técnica eficaz e bastante utilizada em problemas de classificação, visto que o conhecimento é representado por meio de regras escritas em linguagem natural com a intenção de facilitar sua compreensão (Zuben & Attux, 2010).

O estudo aqui apresentado empregou o processo de descoberta de conhecimento em bases de dados composto pelas seguintes etapas: seleção de dados, pré-processamento, transformação dos dados, mineração dos dados e interpretação.

Na fase de seleção, elegeu-se a base de dados do Enade 2012 disponibilizada no portal do Instituto Nacional de Estudos e Pesquisas Educacionais (Inep, 2015).

Na fase de pré-processamento, o atributo objetivo (*qe_i44*) que norteou o estudo referiu-se à seguinte questão: «como você avalia o currículo do seu curso em relação à integração entre os conteúdos das diferentes disciplinas?» Esta tinha as seguintes alternativas de respostas, que se tornaram classes para o minerador: A – é bem integrado, B – é relativamente integrado, C – é pouco integrado e D – não apresenta integração. Dentre os atributos que compõem a tabela original, foram escolhidos como atributos preditivos: *código da categoria administrativa da IES (cd_catad)*, *código da organização acadêmica da IES (cd_orgac)* e *código da UF de funcionamento do curso (cd_uf_curso)*.

Na fase de transformação, foi gerado um arquivo do tipo CSV (*Comma Separated Values*), resultando em 19 268 registros que se referiam aos estados da região Sul para serem usados na ferramenta de mineração de dados Weka (*Waikato Environment for Knowledge Analysis*), desenvolvida pela Universidade de Waikato (Waikato, 2008).

Por fim, na fase de mineração de dados, utilizou-se a tarefa de classificação baseada na construção da Árvore de Decisão (Zuben & Attux, 2010) com o algoritmo J48 (Quinlan, 1993). Finalizando o processo, o método *K-fold cross-validation*, proposto por Mosteller e Tukey (1968) para o algoritmo J48, foi escolhido e configurado com o valor de *K* igual a 10, como sugerido por Kohavi (1995).

4 RESULTADOS E DISCUSSÕES

Com a utilização da tarefa de classificação por Árvore de Decisão na ferramenta Weka, foi possível visualizar a lista de resultados da classificação referente aos microdados processados do Enade 2012, como mostra a figura 1, a seguir.

Com o propósito de atender ao segundo objetivo, fez-se uma análise detalhada das métricas da matriz de confusão, demonstrada na figura 1, para o teste realizado, os seguintes conhecimentos foram encontrados: a) acurácia em relação à precisão para a categoria A (é bem integrado) de 0,585, para B (é relativamente integrado) de 0,504, para C (é pouco integrado) e D (não apresenta integração) de 0,00; b) *recall* para A de 0,986, para B de 0,043, para C e D de 0,00; c) 11 228 instâncias classificadas corretamente, perfazendo 58,27%, e 8.040 instâncias classificadas incorretamente, atingindo 41,72%, para um total de 19.268 instâncias. Dessas 11.228 instâncias classificadas corretamente, 10.919 instâncias foram clas-



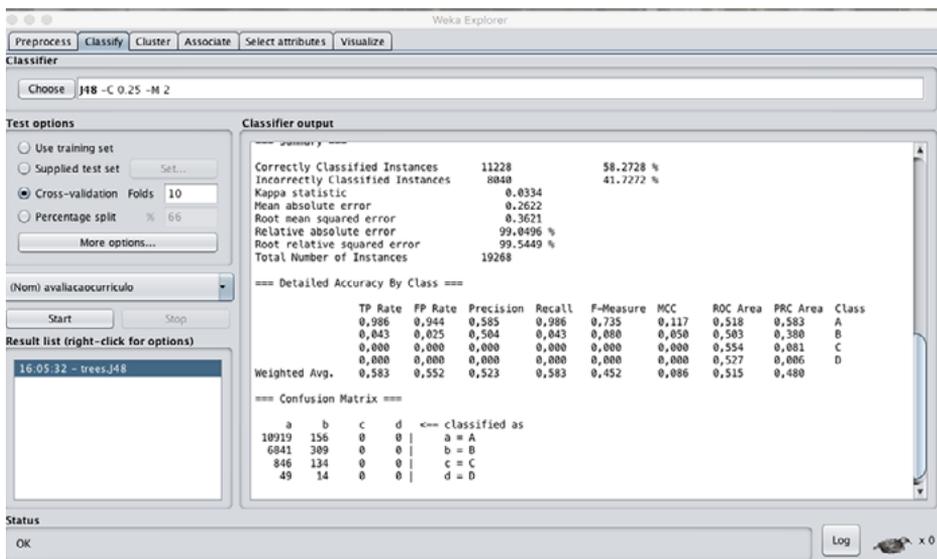


Figura 1. Métricas do classificador extraído do minerador Weka
 Fonte: Os autores (2019).

sificadas para a categoria A (é bem integrado) e 309 instâncias para a categoria B (é relativamente integrado).

De acordo com as instâncias classificadas corretamente para a classe A (bem integrado), obteve-se 10 919 marcações, ou seja, 97,24% das respostas. A classe B (relativamente integrado) obteve 309 marcações, ou seja, 2,7% das respostas. Já nas classes C (pouco integrado) e D (não apresenta integração) não houve nenhuma marcação.

Tais resultados sugerem que não houve, por parte dos estudantes, a compreensão do que estava sendo perguntado, considerando que a pergunta – «Como você avalia o currículo do seu curso em relação à integração entre os conteúdos das diferentes disciplinas?» – apresenta um conceito complexo, conforme foi abordado brevemente no marco teórico sobre o currículo, tornando-se mais complexo ainda quando se trata de integração curricular.

A fim de que os acadêmicos reunissem os saberes necessários para responder à pergunta em questão, para além de um senso comum, baseado em uma lógica dedutível de interesse em não prejudicar o curso que estava sendo avaliado, necessitariam ter estudado tais conceitos nos seus processos formativos, o que certamente não ocorreu. Tal afirmativa pode ser feita porque nenhum dos cursos avaliados era de licenciatura nem tinham alguma relação com a área da Educação, tendo em vista que os conceitos de currículo e currículo integrado pertencem, em grande medida, a essa área do conhecimento. Dezesete cursos foram avaliados no exame, sendo onze de bacharelado e seis de tecnólogo (modalidade intermediária entre o ensino

básico e a Educação Superior caracterizada pela formação prática em menos tempo que um curso de graduação).

Isso posto, infere-se que a pergunta não foi entendida pelos respondentes, gerando dados que não representam a realidade dos cursos avaliados, pois o questionamento não alcançou condições de compreensão para acessar a percepção do estudante quanto à integração do currículo do curso que estavam em processo de conclusão. Essa impossibilidade, provavelmente, resultou na marcação da alternativa que o estudante entendeu ser menos prejudicial ao curso que estava sendo avaliado.

5 CONSIDERAÇÕES

A mineração de dados educacionais, com a intenção de aplicar a tarefa de classificação por Árvore de Decisão pelo minerador Weka, pode ser uma importante aliada para a descoberta de determinados conhecimentos que sejam do interesse do pesquisador. Ela envolve a separação do que interessa da grande quantidade de dados que compõem um banco, onde essas informações podem ou não estar disponíveis para minerar.

A primeira ideia era a de analisar como os estudantes concluintes percebiam a integração curricular das disciplinas nos seus respectivos cursos. No entanto, ao apurar os dados, pôde-se verificar que as informações apontaram para a não compreensão da pergunta por parte dos respondentes, o que provavelmente resultou em respostas equivocadas. Daí a importância do cuidado necessário tanto na produção dos dados minerados quanto na sua interpretação, de modo que eles possam colaborar para a compreensão de determinada realidade que se quer investigar.

Por fim, cabe destacar que o presente trabalho, mais do que descobrir os conhecimentos apresentados, por meio da mineração de dados, busca divulgar essa importante tecnologia como possibilidade de novas pesquisas em grandes bases de dados, para encontrar informações sobre as diferentes realidades dos cursos superiores e das comunidades acadêmicas. Talvez esse possa ser um caminho para alargar formas de compreensão, no sentido de melhorar as condições de acesso, permanência e aproveitamento por parte dos estudantes nos cursos avaliados.

RECIBIDO: noviembre 2021; ACEPTADO: diciembre 2021



REFERÊNCIAS BIBLIOGRÁFICAS

- BARDIN, L. (2016). *Análise de conteúdo*. São Paulo: Edições 70.
- BAKER, R.S.J., ISOTANI, S. & CARVALHO, A.M.J.B. (2011). Mineração de dados educacionais: oportunidades para o Brasil. *Revista Brasileira de Informática na Educação*, 19(2). doi: <http://dx.doi.org/10.5753/rbie.2011.19.02.03>.
- CAMILO, C.O. & SILVA, J.C. (2009). *Mineração de dados: conceitos, tarefas, métodos e ferramentas*. Relatório técnico – INF. 2009. Recuperado de http://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-09.pdf.
- ELIAS, N. (2011). *O processo civilizador: uma história dos costumes* (2.ª ed., vol. 1, R. Jungmann, Trad.). Rio de Janeiro: Zahar.
- GOODSON, I.F. (1995). *Currículo: teoria e história* (7.ª ed.). Petrópolis, RJ: Vozes.
- INSTITUTO Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). (2015) *Microdados*. Recuperado de <http://portal.inep.gov.br/web/guest/microdados>.
- KOHAVI, R.A. (1995). *Study of cross-validation and bootstrap for accuracy estimation and model selection*. Recuperado de <http://robotics.stanford.edu/~ronnyk/accEst.pdf>.
- LEI n. 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior (SINAES) e dá outras providências. Recuperado de http://www.planalto.gov.br/ccivil_03/ato2004-2006/2004/lei/110.861.htm.
- MOSTELLER, F. & TUKEY, J. W. (1968). «Data analysis, including statistics», in G. Lindzey & E. Aronson (eds.). *Handbook of Social Psychology* (vol. 2). Massachusetts: Addison-Wesley.
- PACHECO, J.A. (2005). *Escritos curriculares*. São Paulo: Cortez.
- QUINLAN, J. R. (1993). *C4.5: programs for machine learning*. San Francisco: Morgan Kaufmann Publishers Inc.
- SANTOMÉ, J.T. (1998). *Globalização e interdisciplinaridade: o currículo integrado* (C. Shiling, Trad.). Porto Alegre: Artes Médicas Sul.
- TAN, P.-N., STEINBACH, M. & KUMAR, V. (2009). *Introdução ao DataMining Mineração de Dados*. Rio de Janeiro: Ciência Moderna.
- WAIKATO. (2008). *Waikato Environment for Knowledge Analysis (WEKA)*. *Weka 3—Data Mining with Open Source Machine Learning Software in Java*. Recuperado de <http://www.cs.waikato.ac.nz/ml/weka/>.
- ZUBEN, F.J.V. & ATTUX, R.R. de F. (2010). *Árvore de decisão*. Recuperado de ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia004_1s10/notas_de_aula/topico7_IA004_1s10.pdf.

