



MINERAÇÃO DE DADOS APLICADA AO ESTUDO DA RELAÇÃO ENTRE AS REPROVAÇÕES NAS DISCIPLINAS DE LÓGICA DE PROGRAMAÇÃO E A EVASÃO NO CURSO DE SISTEMA DE INFORMAÇÃO - CCET

Leonardo Ribeiro da Silva Junior¹, Noeli Antônia Pimentel Vaz¹

leojunior1326@gmail.com, noeli@ueg.br

¹Universidade Estadual de Goiás – CCET – Sistemas de Informação -
Anápolis – GO

RESUMO

A evasão é um problema grave enfrentado por Instituições de Ensino Superior, pois com a saída dos alunos das instituições de ensino são desperdiçados recursos sociais, acadêmicos e econômicos. A busca de suas causas tem sido objeto de pesquisas educacionais desenvolvidas ao longo dos anos. Com o objetivo de investigar a evasão gerada no curso de Sistemas de Informação - SI do Campus Anápolis de Ciências Exatas e Tecnológicas Henrique Santillo - CCET da Universidade Estadual de Goiás - UEG o presente trabalho visa investigar se existe relação entre as reprovações nas disciplinas de Lógica de Programação I e II e a evasão identificada no período de 2003 a 2017. Será utilizado o processo de Descoberta de Conhecimento em Base de Dados (DCBD) em todas as suas etapas.

Palavras chave: Evasão, Sistemas de Informação, Lógica de programação, Mineração de dados, Descoberta de Conhecimento em Base de Dados.

DATA MINING APPLIED TO THE STUDY OF THE RELATIONSHIP BETWEEN REPROGRAMS IN PROGRAMMING LOGIC DISCIPLINES AND EVASION IN THE INFORMATION SYSTEM COURSE - CCET

ABSTRACT

Evasion is a serious problem faced by Higher Education Institutions, because with the exit of students from educational institutions, social, academic and economic resources are wasted. The search for its causes has been the object of educational research developed over the years. With the objective of investigating the evasion generated in the course of Information Systems of the Anápolis Campus of Exact and Technological Sciences - Henrique Santillo of the State University of Goiás the present work aims to investigate if there is a relation between the reprograms in the disciplines of Programming Logic I and II and the evasion identified in the period from 2003 to 2017. The process of Knowledge Discovery in Database (DCBD) will be used in all its stages.

Keywords: Evasion in Information Systems, Programming Logic, Data Mining, Knowledge Discovery in Databases.

I. INTRODUÇÃO

A evasão é um dos principais problemas que atingem Instituições de ensino Superior no Brasil. A evasão caracteriza-se pelo ato do discente não concluir o curso, sendo causada pelos mais diversos motivos. A evasão está presente em todas as áreas do Ensino superior, principalmente em cursos da área de computação. De acordo com o Censo da educação superior brasileira, realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - INEP (2016), em cursos de Sistemas de Informação, a taxa de alunos que evadem chega a ultrapassar 2/3 do total de alunos matriculados. Observa-se, empiricamente, no curso de SI – CCET, que trancamentos, desistências e abandonos ocorrem em períodos de oferta das disciplinas de Lógica e Programação I e II.

O principal objetivo do ensino de Lógica de Programação em um curso superior, não é ensinar uma linguagem de programação específica, mas sim ensinar sequências lógicas que serão utilizadas na resolução de problemas, independente da linguagem de programação utilizada. O estudo da relação entre a evasão e as reprovações no curso de SI – CCET será apoiado pelo processo de Descoberta de Conhecimento em Bancos de Dados.

A Descoberta de Conhecimento em Bancos de Dados - DCBD (do inglês *Knowledge Discovery in Databases – KDD*) é o processo que tem como objetivo transformar dados comuns, armazenados em base de dados em conhecimento útil, ou seja, transformá-los em informação nova que possa trazer algum benefício ao usuário.

As informações a serem utilizadas nesta pesquisa foram obtidas da base de dados do sistema acadêmico da Universidade Estadual de Goiás. Durante a etapa de Mineração de Dados do processo de DCBD, será utilizado o software livre Weka para a aplicação do algoritmo de regras de associação, com o objetivo de encontrar as possíveis relações entre a evasão dos discentes e as reprovações nas disciplinas de Lógica de Programação.

II. METODOLOGIA

A metodologia utilizada para o desenvolvimento do trabalho é norteada pelas cinco etapas propostas por Fayyad et al. (1996) no processo de DCBD, sendo elas: seleção dos dados, pré-processamento, transformação dos dados, mineração de dados e análise dos resultados.

Até o presente momento, foram selecionados dados do sistema acadêmico da UEG do curso de SI – CCET, obtendo um total de 13.325 registros correspondentes as disciplinas ofertadas entre 2003 e 2017 com informações que apresentam, dentre outras informações, o desempenho dos discentes que cursaram as disciplinas. É importante ressaltar que os dados analisados não possuem campos que identifiquem os discentes. Após a etapa de seleção, realizou-se a limpeza dos dados e foram retiradas informações não aplicáveis à pesquisa. Atualmente, os dados estão sendo transformados para o formato adequado para a aplicação do algoritmo *apriori*, para que posteriormente seja realizada etapa da Mineração de Dados e a análise dos resultados.

Para a etapa de Mineração de Dados, foi selecionado o algoritmo *apriori* para obtenção de regras das regras de associação. Durante o processo de Mineração de

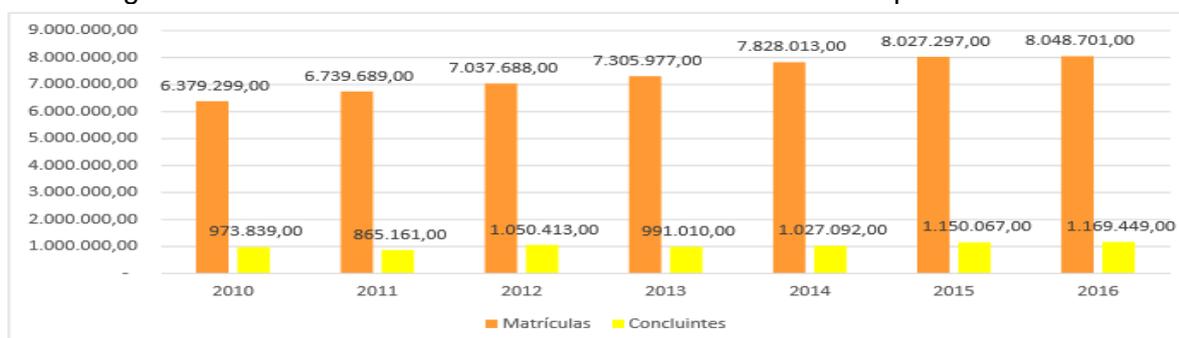
24 e 25 de agosto de 2018

Dados, será utilizado o software livre Weka para a aplicação do algoritmo de regras de associação para encontrar as possíveis relações entre a evasão dos discentes e as reprovações nas disciplinas de Lógica de Programação.

FUNDAMENTAÇÃO TEÓRICA

Com a expansão do ensino superior no Brasil, houve o surgimento de novas instituições de ensino públicas e privadas. Dentre outras dificuldades evidentes, a evasão é um problema na educação superior brasileira, que vem ocupando um papel relevante em pesquisas e discussões. Na Figura 1, é possível verificar a evolução do número de matrículas e o número de concluintes no Ensino Superior brasileiro entre o período de 2010 a 2016.

Figura 01: Número de matrículas e concluintes no Ensino Superior brasileiro.



Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2016)

Como pode ser visto na figura anterior, apesar do aumento significativo do número de matrículas ao longo dos anos, também pode ser visualizada grande discrepância em relação ao número de concluintes.

O cenário para os cursos de Sistemas de Informação no Brasil, também é bastante impactante, quando é comparado o número de ingressantes e concluintes, conforme exposto na Figura 02.

Figura 02: Relação entre matrículas e concluintes em Sistemas de Informação.



Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2016)

Estudos realizados no Brasil por autores que buscam investigar as causas e fatores que influenciam o alto índice de evasão dos cursos relacionados a área de Computação, incluindo o curso de Sistemas de Informação, relatam que uma possível causa que contribui para o alto índice de evasão relacionado a este curso, seria o de o curso demandar um raciocínio lógico elevado em relação às disciplinas de lógica de

programação e afins. O curso pode demandar um alto nível de esforço, logo nos primeiros semestres, assim exigindo dos discentes a obtenção de novos conhecimentos e de um raciocínio lógico maduro e complexo, algo que não é normalmente exigido nas disciplinas do ensino médio. Gomes et al.(2008, p.02) relata que “[...]as dificuldades encontradas nas disciplinas iniciais de algoritmos contribuem para a elevada taxa de evasão em cursos de computação”. Como geralmente as disciplinas de Lógica de Programação são pré-requisitos para diversas outras disciplinas ao longo do curso, a reprovação em tais disciplinas pode acabar atrapalhando toda a vida acadêmica do discente, causando certo desânimo em relação ao curso.

A Descoberta de Conhecimento em Bancos de Dados – DCBD (do inglês *Knowledge Discovery in Databases – KDD*) é o processo que tem como objetivo transformar dados comuns, armazenados em base de dados em conhecimento útil, ou seja, transformá-los em informação nova que possa trazer algum benefício ao usuário. Fayyad et al. (1996) define o processo KDD como sendo um processo não trivial, de busca e extração de conhecimento útil e previamente desconhecido a partir dos dados armazenados em um banco de dados. Com a aplicação da Mineração de dados e do algoritmo *Apriori* identificar se as reprovações nas disciplinas de Lógica de Programação I e II tem um alto índice de associação com as evasões dos discentes no curso.

III. CONSIDERAÇÕES FINAIS

O estágio atual do desenvolvimento da pesquisa e com análises dos resultados obtidos até o presente momento, pode se comprovar um alto índice de reprovações em disciplinas de Lógica de Programação, sendo uma média que se aproxima dos 50% dos discentes que se matriculam nessas disciplinas (em alguns anos superando essa média).

Com a aplicação do algoritmo *Apriori* e a finalização desta pesquisa, espera-se identificar uma possível relação entre o alto índice reprovação nas matérias relacionadas a lógica de programação e a evasão no curso, procurando evidenciar alunos que reprovam nas disciplinas de Lógica de Programação I e/ou Lógica de programação II, pelos dados históricos, evadiram do curso.

REFERÊNCIAS

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. **From Data Mining to Knowledge Discovery in Databases**. AI Magazine, Providence, Rhode Island, v. 17, n. 3, p.37-54, jan. 1996. Disponível em: <<https://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1230/1131>>. Acesso em: 22 abr. 2018.

GOMES, A., HENRIQUES, J., MENDES, A. J. **Uma proposta para ajudar alunos com dificuldades na aprendizagem inicial de programação de computadores**. In Educação, Formação & Tecnologias; vol.1(1), pp. 93- 103. 2008. Disponível em <http://eft.educom.pt>. Acesso em: 22 abr. 2018.