



MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
BAIANO-CAMPUS GUANAMBI

ALEX SOUZA CARVALHO

**MINERAÇÃO DE DADOS EDUCACIONAIS PARA
PREVISÃO DE DESEMPENHO E EVASÃO DE
ESTUDANTES NAS DISCIPLINAS INTRODUTÓRIAS DE
ALGORITMOS**

GUANAMBI

2018



ALEX SOUZA CARVALHO

**MINERAÇÃO DE DADOS EDUCACIONAIS PARA
PREVISÃO DE DESEMPENHO E EVASÃO DE
ESTUDANTES NAS DISCIPLINAS INTRODUTÓRIAS DE
ALGORITMOS**

Projeto do Trabalho de Conclusão de Curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia Baiano-Campus Guanambi ligado ao Ministério da Educação, como requisito parcial da disciplina Tópicos de Pesquisa em Computação para a obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Área de concentração: Análise e Desenvolvimento de Sistemas

Linha de pesquisa: Informática na Educação

Orientador: Msc. Tiago do Carmo Nogueira
Instituto Federal Baiano

GUANAMBI
2018

RESUMO

CARVALHO, Alex. Mineração de Dados Educacionais para Previsão de Desempenho e Evasão de Estudantes nas Disciplinas Introdutórias de Algoritmos. 2018. 9 f. Projeto do Trabalho de Conclusão de Curso – Tecnologia em Análise e Desenvolvimento de Sistemas, Instituto Federal de Educação, Ciência e Tecnologia Baiano-Campus Guanambi. Guanambi, 2018.

Este trabalho apresenta um estudo com a utilização de mineração de dados voltada ao contexto educacional, em cursos superiores de informática, onde há um elevado número de alunos reprovados nos componentes curriculares de programação. A necessidade de sua realização surgiu a partir da observação de uma performance insatisfatória de estudantes com relação às disciplinas de algoritmo e estrutura de dados. Para tanto, escolheu-se a utilização de dados de alunos, matriculados em uma instituição pública de ensino superior, a serem coletados e analisados utilizando-se de algoritmos de Máquina de Vetor de Suporte e processados com a ferramenta WEKA, a fim de se criar modelos de regressão capazes de tanto apontar riscos de evasão quanto predizer as possíveis causas do elevado número de reprovações, em tais disciplinas.

Palavras-chave: Mineração de Dados Educacionais. Evasão Escolar. Máquina de Vetor de Suporte.

ABSTRACT

CARVALHO, Alex. Educational Data Mining for Performance Forecasting and Student Evasion in Introductory Algorithm Disciplines. 2018. 9 f. Projeto do Trabalho de Conclusão de Curso – Tecnologia em Análise e Desenvolvimento de Sistemas, Instituto Federal de Educação, Ciência e Tecnologia Baiano-Campus Guanambi. Guanambi, 2018.

This work presents a study with the use of data mining focused on the educational context, in higher education courses, where there is a high number of students who fail in the curricular components of programming. The need for its realization arose from the observation of an unsatisfactory student performance in relation to the disciplines of algorithm and data structure. In order to do so, we chose the use of data from students enrolled in a public institution of higher learning, to be collected and analyzed using Support Vector Machine algorithms and processed with the WEKA tool, in order to create regression models capable of both pointing out evasion risks and predicting the possible causes of the high number of failures in such disciplines.

Keywords:Educational Data Mining. School Evasion. Support Vector Machine.

LISTA DE FIGURAS

LISTA DE TABELAS

Tabela 1 – Cronograma de Atividades - Semestre II-2018 e I-2019	7
---	---

SUMÁRIO

1 – INTRODUÇÃO	1
2 – PROBLEMA E JUSTIFICATIVA	3
3 – OBJETIVOS	5
3.1 Objetivo Geral	5
3.2 Objetivos Específicos	5
4 – CRONOGRAMA	7
Referências	9

1 INTRODUÇÃO

Nos cursos superiores voltados às áreas da informática há um constante problema de evasão e baixo desempenho nas disciplinas introdutórias de programação. Gomes, Henriques e Mendes (2008) e Manhães et al. (2011), entre outros autores, descrevem que o problema vem da incapacidade de compreensão, por parte dos alunos, dos conceitos ligados às disciplinas e sua aplicação prática aos problemas cotidianos ou, até mesmo, do desbalanceamento das turmas principalmente quando envolvem grande quantidade de alunos por sala. No entanto, é impossível especular as causas que geram tal problema sem que antes haja uma análise aplicada diretamente às suas possíveis fontes.

Para que seja possível que se tenha uma noção das principais causas de tal obstáculo na educação superior em determinadas instituições brasileiras é necessário que um estudo seja realizado possuindo um foco local, uma vez que estas representam, acima de tudo, um problema relacionado ao ambiente de convívio dos próprios estudantes e sua urgência de resolução demonstra-se indispensável. Este trabalho busca obter algumas respostas ao problema de baixo desempenho nos níveis educacionais, bem como prever os riscos de evasão utilizando-se de técnicas de mineração de dados educacionais.

O objetivo deste trabalho, além de examinar as principais fontes causadoras da evasão e reprovação escolar, objetiva-se, também, à análise dos principais fatores que determinam o desempenho de alunos nas disciplinas introdutórias de programação, como algoritmos e estruturas de dados, em cursos superiores de informática para que seja possível identificar suas origens no ambiente pedagógico. Espera-se, com isso, a obtenção de resultados que, além demonstrar as causas das questões propostas, levantem outras questões de como nosso sistema educacional precisa ser trabalhado para que o potencial de inúmeros estudantes não deixe de ser aproveitado pela própria sociedade num futuro próximo.

A metodologia encontra-se dividida em três fases onde, na primeira fase, são apontados a instituição de ensino superior onde os dados serão trabalhados. Na segunda fase são descritos os procedimentos relacionados à coleta de dados, o local de onde as informações serão extraídas e os principais indicadores com os quais serão trabalhados e, por fim, na terceira fase, são demonstradas as abordagens e ferramenta que serão utilizadas durante o processamento dos dados coletados na segunda fase.

A execução da abordagem metodológica, apresentada na seção 5, servirá como forma de trabalhar os conceitos estudados no referencial teórico (seção 4), bem como responder às questões levantadas durante a apresentação do problema e justificativa da pesquisa (seção 2), além de alcançar os objetivos do trabalho, apresentados na seção 3.

2 PROBLEMA E JUSTIFICATIVA

No Brasil, segundo Gomes, Henriques e Mendes (2008), Fernandes e Junior (2016) e Júnior et al. (2005), as disciplinas de algoritmo e programação são responsáveis tanto por um grande número de reprovações como também representam uma das causas da evasão de estudantes em cursos de computação no ensino superior e médio/técnico em instituições públicas. Tal fato pode estar relacionado ao nível teórico do material estudado uma vez que a programação possui conceitos próprios e, de certa forma, pouco compartilhados com as demais disciplinas, sua especificidade nos módulos iniciais bem como a não difusão do contexto nas escolas públicas de ensino básico nos anos anteriores.

De acordo com Manhães et al. (2011), diversos fatores podem ser apontados como possíveis causas de evasão de alunos, entre eles é possível reconhecer aspectos socioeconômicos regionais, localização geográfica, tempo de duração e adequação ao mercado de trabalho. Isso significa que suas principais causas podem variar entre instituições e regiões específicas e, por se tratar de um componente curricular de significância maior dentro de um curso de computação, as causas de seu elevado número de reprovações e evasões devem ser melhor estudadas e analisadas para que se possa criar indicadores capazes de determinar as principais fontes do problema.

Para tal, discussão mais aprofundada se torna necessária a fim de demonstrar como isso pode vir a atingir o ensino, não apenas em determinada instituição mas, também, dentro de um contexto social. É necessário que se discuta suas causas e consequências e como elas se relacionam para que, até mesmo, seus professores tenham melhor ideia de como lidar com esse quesito, sem que haja perda no desempenho dos alunos ou até dos próprios professores, beneficiando a instituição e sociedade ao final desse processo. Por outro lado, negar a existência do problema em questão, ignorando sua origem e efeitos, bem como a importância da discussão do tema poderá representar um esgotamento no curso de computação com rapidez proporcional ao nível de evasão dos estudantes.

Portanto, um estudo aprofundado no quesito é necessário para que sejam analisadas as dificuldades envolvidas no ensino de programação em certas instituições, por se tratar de um problema de interesse comum, o que poderá, até mesmo, contribuir com processos que englobam temas socioeconômicos regionais inseridos na realidade de cada instituição, além de apresentar relevante valor para o meio acadêmico.

3 OBJETIVOS

3.1 Objetivo Geral

Identificar, através da mineração de dados, as possíveis causas de reprovação e evasão, bem como analisar os principais fatores que determinam o desempenho de alunos em disciplinas de algoritmo e programação nos cursos superiores e técnicos de informática, para que seja possível examinar suas principais origens dentro do contexto educacional.

3.2 Objetivos Específicos

- Indicar métodos automatizados de coleta e análise de dados que possibilitem a determinação de fatores que corroboram com o abandono escolar;
- Identificar as causas mais comuns causadoras da queda na produtividade do aluno na sala de aula;
- Levantar questões relacionadas ao cenário socio-cultural, envolvendo tanto as causas quanto as consequências de um baixo desempenho acadêmico.

4 CRONOGRAMA

Para execução das atividades a serem realizadas até a defesa do Trabalho de Conclusão do Curso teremos as seguintes etapas, conforme Tabela 1 abaixo.

Tabela 1 – Cronograma de Atividades - Semestre II-2018 e I-2019

Atividade	Período											
	II Semestre de 2018						I Semestre de 2019					
	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	
Definição do Tema, Introdução, Objetivos e Problema	•	•	•									
Elaboração da Justificativa e Hipóteses	•	•	•									
Revisão Bibliográfica / Teórica da pesquisa	•	•	•	•								
Definição da Metodologia			•	•								
Elaboração dos Instrumentos de Coleta de Dados			•	•								
Revisão do Pré-TCC				•								
Entrega do Pré-TCC				•								
Defesa do Pré-TCC					•							
Realização da Pesquisa						•						
Coleta dos dados da pesquisa							•					
Análise dos dados da pesquisa								•				
Revisão do Pré-TCC e Elaboração do TCC									•			
Entrega do TCC										•		
Defesa do TCC											•	

Fonte: P rópria (2018).

Referências

FERNANDES, V. dos S.; JUNIOR, V. F. Linguagem de programação: Evasão e reprovação no instituto federal catarinense, campus avançado sombrio. 2016. Citado na página 3.

GOMES, A.; HENRIQUES, J.; MENDES, A. Uma proposta para ajudar alunos com dificuldades na aprendizagem inicial de programação de computadores. **Educação, Formação & Tecnologias-ISSN 1646-933X**, v. 1, n. 1, p. 93–103, 2008. Citado 2 vezes nas páginas 1 e 3.

JÚNIOR, J. et al. Ensino de algoritmos e programação: uma experiência no nível médio. In: **XIII Workshop de Educação em Computação (WEI'2005)**. São Leopoldo, RS, Brasil. [S.l.: s.n.], 2005. Citado na página 3.

MANHÃES, L. M. B. et al. Previsão de estudantes com risco de evasão utilizando técnicas de mineração de dados. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2011. v. 1, n. 1. Citado 2 vezes nas páginas 1 e 3.