



MINERAÇÃO DE DADOS PARA ANÁLISE DO DESEMPENHO DO CURSO DE SISTEMAS DE INFORMAÇÃO DA UEG NO ENADE

Gustavo Ribeiro de Mizael (IC) gustavomizaelsi@gmail.com, **José Leonardo de Oliveira Lima (PQ)**

Universidade Estadual de Goiás - Anápolis/GO.

Resumo: O trabalho teve como objetivo analisar como se encontra o curso de Sistemas de Informação (SI) do Câmpus de Ciências Exatas e Tecnológicas (CCET) da Universidade Estadual de Goiás (UEG) em relação aos cursos de SI do país, correlacionando a situação socioeconômica com o bom desempenho (notas maiores ou igual a 50) no Exame Nacional de Desempenho dos Estudantes (ENADE). Foi aplicada a mineração de dados com a metodologia *Cross-Industry standard process for data mining* (CRISP-DM), usando a tarefa de classificação e técnica de árvore de decisão, correlacionando variáveis dispersas e gerando, assim, o perfil dos alunos com bom desempenho dos cursos de SI no cenário nacional e local (CCET Anápolis). Após descobrir o perfil do aluno com bom desempenho, foi feita uma análise comparativa entre os perfis. Dentre as conclusões, constatou-se que o perfil de bom desempenho traçado na mineração, apresenta variáveis diferentes para cada cenário (nacional e local), sendo que o curso de SI do CCET não possui os mesmos problemas que os cursos de SI do Brasil, mostrando algumas peculiaridades e, com isso, um desempenho no ENADE acima da média dos cursos de SI do Brasil.

Palavras-chave: Mineração de dados. Descoberta do Conhecimento. Classificação. ENADE. Sistemas de Informação. Perfil.

Introdução

Nos últimos 10 anos, foi possível observar um grande crescimento da quantidade de vagas em instituições de educação superior, porém a qualidade de ensino não acompanhou o crescimento quantitativo dessas vagas (LIMA, 2016). Saber se o fator socioeconômico influencia no desempenho acadêmico do discente é uma informação importante para poder melhorar a qualidade do ensino, percepção essa também corroborada por Silva (2017). A pesquisa realizada teve como objetivo descobrir a relação da situação socioeconômica com o desempenho do ENADE 2014, obtido por alunos dos cursos de SI do Brasil e na UEG CCET – Anápolis, verificando também se a UEG – CCET apresentava os mesmos problemas socioeconômicos que ocorrem na maioria dos cursos de SI no Brasil.

Materiais e Métodos

O ambiente de estudo de caso da pesquisa foi o curso de SI da UEG no CCET, na turma de alunos que realizaram o ENADE 2014. Os dados utilizados foram os dados públicos disponibilizados no sítio do Ministério da Educação referentes ao resultado da prova e relatório socioeconômico que foram aplicados no ENADE 2014.

Utilizou-se o processo de Descoberta do Conhecimento em Bases de Dados (DCBD), que é realizado em cinco etapas e funciona em forma de ciclos interativos e



iterativos (FAYYAD et al., 1996). Para a mineração, foi aplicada a metodologia CRISP-DM (CHAPMAN et al, 2000; HEKIMA, 2016) com a tarefa de classificação e técnica de árvore de decisão para análise dos dados do ENADE 2014 dos cursos de SI de todo o Brasil.

Ao definir quais atributos fariam parte da análise, a nota bruta foi definida como variável meta. Sobre a variável meta, foram definidos dois parâmetros, sendo eles: nota azul (notas maiores ou iguais a 50) e nota vermelha (notas menores do que 50). Os softwares Excel e Excel2Converter foram utilizados para tratamento dos dados. A Mineração de dados foi realizada usando o software WEKA. Foi feito um comparativo entre os resultados obtidos no curso de SI da UEG CCET Anápolis e todos os cursos de SI do Brasil, utilizando os mesmos parâmetros.

Resultados e Discussão

O perfil do aluno com bom desempenho no curso de SI da UEG CCET, é o do aluno que estuda de 4 a 7 horas semanais extraclasse, ou também o aluno que estuda de 1 a 3 horas semanais e finalizou o ensino médio antes ou no ano de 2005. O perfil do aluno de SI com bom desempenho em âmbito nacional é formado pelo sexo masculino, que não recebeu nenhum incentivo financeiro, pois estuda em instituição superior pública e gratuita e tem renda total da família de R\$ 4.344,01 a R\$ 7.240,00 mensal.

Foi possível perceber que o fato da universidade ser gratuita se mostrou expressivo para o bom desempenho do aluno, pois no cenário nacional, apesar de os alunos provenientes de instituição de educação superior pública representar apenas 20% do total de alunos avaliados no ENADE 2014, esses alunos contemplam o perfil com bom desempenho, demonstrando influência do fator ambiental no desempenho do discente.

A informação que sugere a correlação entre as questões socioeconômicos e o desempenho no ENADE é que 60% dos alunos com bom desempenho da CCET possui uma renda total da família de R\$ 4.344,01 a R\$ 7.240,00 mensal, demonstrando que alunos com bom desempenho estão em uma faixa salarial mensal maior.

Percebeu-se que o curso da CCET tem mais variedade de alunos provenientes de colégio público ou particular com bolsa (por ação afirmativa ou de inclusão social), representando 30% dos alunos, além de possuir 10% dos alunos que participaram de atividades curriculares no exterior. Os alunos da CCET também apresentam uma maior quantidade de alunos que foram incentivados pelos pais a fazerem o curso,



representando 80% e possuindo também, 62% dos alunos que indicam a gratuidade da instituição como motivo da escolha da universidade, sendo que esse perfil apresenta bom desempenho no ENADE 2014.

No cenário nacional, alunos que ingressaram por ação afirmativa ou inclusão social representam apenas 17,3% dos alunos, possuindo também apenas 3% de discentes participantes de atividades curriculares no exterior. Os alunos no cenário nacional, apesar de, em sua maioria, serem incentivados pelos pais a fazerem o curso, essa quantidade representa 63,8% dos alunos e o principal motivo da escolha da instituição de ensino superior é a qualidade/reputação da instituição, representando 40% do total dos discentes.

A partir dos dados encontrados, infere-se que a participação, apoio familiar e estudo extraclasse cumprem papel importante na vida acadêmica do estudante e no seu desempenho. A renda é outro fator que contribui para um bom rendimento, sendo que as políticas públicas e de permanência precisam garantir aos estudantes o mínimo de estabilidade financeira e material para que possam permanecer dedicar aos estudos e ter qualidade de vida, para si e para a família.

O fato da instituição ser pública e ter os melhores resultados no âmbito nacional (e também local, pois o resultado da UEG, que também é pública, foi acima da média nacional) sugere a possibilidade de ter relação com compromisso maior das universidades públicas com a qualidade da formação.

Considerações Finais

A pesquisa conseguiu responder como se encontra o curso de SI da UEG - CCET Anápolis em relação ao cenário nacional, correlacionando a situação socioeconômica com o desempenho no ENADE 2014, atingindo, assim, o objetivo proposto. Concluiu-se que a quantidade de horas de estudo e a situação socioeconômica do aluno apresenta influencia no desempenho do ENADE. A situação socioeconômica dos alunos de SI da UEG CCET é diferente do cenário nacional, e pode ter correlação com o melhor desempenho do curso de SI da CCET no ENADE em relação ao cenário nacional. Pelas análises feitas, o curso de SI da CCET se encontra acima da média em relação aos cursos de SI no contexto Brasil como um todo e apresenta uma maior quantidade de



notas acima de 50 (notas azuis), proporcionalmente. Porém, mais estudos e análises precisariam ser feitos para afirmações conclusiva relacionadas a causa e efeito.

Referências

CHAPMAN, P et al. CRISP-DM 1.0: Step-by-step data mining guide. **Cross Industry Standard Process for Data Mining**, 2000. Disponível em: <<https://www.the-modeling-agency.com/crisp-dm.pdf>>. Acesso em: 5 set. 2017

FAYYAD, U. M., PIATETSKY-SHAPIO, G., SMUTH, P., et al. **Advances in Knowledge Discovery and Data Mining**. Menlo Park: AAAI Press, 1996. Disponível em: <<https://www.csd.uwo.ca/faculty/ling/cs435/fayyad.pdf>>. Acesso em: 21 abr. 2017.

HEKIMA. **Se você se interessa por Big Data, precisa entender o CRISP-DM**. 2016. Disponível em: <<http://www.bigdatabusiness.com.br/se-voce-se-interessa-por-big-data-precisa-entender-o-crisp-dm/>>. Acesso em: 19 maio 2017.

LIMA, José Leonardo Oliveira. **Avaliação discente em cursos de graduação a distância mediados por ambientes virtuais de aprendizagem: Necessidade de informações dos docentes na visão de especialistas europeus e brasileiros**. 2016. 298f., Tese (Doutorado em Ciência da Informação) - Universidade de Brasília, Brasília, 2016. Disponível em:<<http://repositorio.unb.br/handle/10482/22333>>. Acesso em: 25 out. 2017.

SILVA, Flávia Guimarães Menezes. **Desempenho dos estudantes no ENADE: influenciadores ambientais e organizacionais sob a luz da sociologia da educação de Bourdieu**. 2017. Disponível em: <<http://tede.unifacs.br/tede/handle/tede/617>>. Acesso em: 3 nov. 2017.