

MINERAÇÃO DE DADOS APLICADA NA IDENTIFICAÇÃO DE ALUNOS COM PERFIL DE EVASÃO NO ENSINO SUPERIOR

Danielle Emanuelle Aparecida Ribeiro (FHO - Fundação Hermínio Ometto)

daniribeiro@alunos.fho.edu.br

Mauricio Acconcia Dias (FHO - Fundação Hermínio Ometto) macdiaspae@gmail.com

Resumo

Atualmente o perfil dos alunos de cursos superiores apresenta uma mudança que torna complexa a tarefa de motivação independente dos estágios do curso. A falta de interesse dos alunos acarreta num aumento significativo da evasão e esforços são necessários para a redução destes índices. Uma possível solução para este problema é traçar um perfil do aluno que evadiu do curso e, desta forma, tentar identificar possíveis evasões e tratar o problema antes que aconteça. Portanto o objetivo deste trabalho é o uso de mineração de dados educacionais a fim de traçar o perfil do aluno que teria fortes chances de evadir o curso. Para isso realizou-se um estudo de dados, não identificados, sobre alunos em seu primeiro semestre no curso de Sistemas de Informação da Unicamp. Os resultados obtidos mostram que uma das formas de identificar um aluno que tende a evadir o curso é analisando seu desempenho em algumas disciplinas pontuais do curso. Apesar de relativamente óbvio, o resultado indica quais disciplinas devem ser modificadas caso seja necessária uma ação para diminuição da evasão.

Palavras-chave: mineração de dados; índices educacionais; evasão no ensino superior; rapidminer

1. Introdução

As instituições de ensino superior têm apresentado altos índices de evasão em seus cursos¹. Este problema atinge tanto as instituições públicas como as privadas, e seus efeitos são diversos, causando principalmente apreensão aos responsáveis pela gestão acadêmica, que buscam melhorar os indicadores de retenção e conclusão.

Apesar de ter registrado avanços nos últimos anos, a educação no Brasil ainda apresenta dados insatisfatórios. A necessidade de expansão da educação superior no país é visível, pois a média nacional é de que apenas 24% dos jovens brasileiros, com idade entre 18 e 24 anos, têm acesso

¹ <http://agenciabrasil.ebc.com.br/educacao/noticia/2018-09/pesquisa-mostra-evacao-de-30-em-cursos-superiores-privados>

ao ensino superior. (INEP, 2014). Segundo o resumo técnico do censo da educação superior, realizado pelo MEC em 2016, no qual participaram 2.407 instituições de ensino superior, registrou-se 34.366 cursos, 8.048.701 matrículas, 2.985.644 ingressos e 1.169.449 concluintes de graduação. As Instituições de Ensino Superior (IES) privados têm uma participação de 75,3% (6.058.623) no total de matrículas de graduação. Já a rede pública, participa com 24,7% (1.990.078) (INEP, 2016).

A cada ano é oferecido pelas instituições um aumento no número de vagas para os alunos ingressarem na graduação. Porém, parte dos alunos que entram nas universidades não concluem o curso, apesar dos incentivos dos programas de financiamento governamental para os alunos do setor privado, em especial o FIES e o PROUNI².

O alto número de alunos evadidos cria um círculo vicioso: eles abandonam a educação ou até a rejeitam, sentem-se perdedores, o mercado os desvaloriza, o conhecimento não é passado a frente, a economia e sociedade perdem e, ao final, a civilização estanca uma boa parcela da evolução. (Revista Unicamp, 2017). A utilização de índices para medição de evasão, varia entre medição de desistência em uma determinada disciplina de um curso, ou a desistência em um curso de graduação, desistência de um curso, mas ingresso em outro.

Considerando a importância deste cenário para a educação no país, este projeto de pesquisa tem como objetivo a utilização de mineração de dados para obtenção de um modelo de aluno que tende a evadir o curso e, a partir deste modelo, propor soluções para o problema que sejam mais eficazes. Neste trabalho, apresentaremos um estudo utilizando dados, não identificados, oriundos do sistema acadêmico da Unicamp, de alunos de graduação do curso de Sistema de Informação, correspondente ao período de 2013 a 2018.

Este trabalho está estruturado da seguinte forma: na seção seguinte é apresentada a fundamentação teórica, seguida pelos trabalhos relacionados; na quarta seção é apresentado a metodologia empregada para atingir os resultados esperados, em seguida é exposto os resultados encontrados e por fim, a última seção contempla a conclusão desta pesquisa.

2. Referencial Teórico

A evasão universitária é um fenômeno cada vez mais presente e preocupante. Trata-se de um problema difícil de combater, pois as causas são diversas e variam de um contexto para o outro

² <https://www.prouni.com.br/prouni-e-fies/>

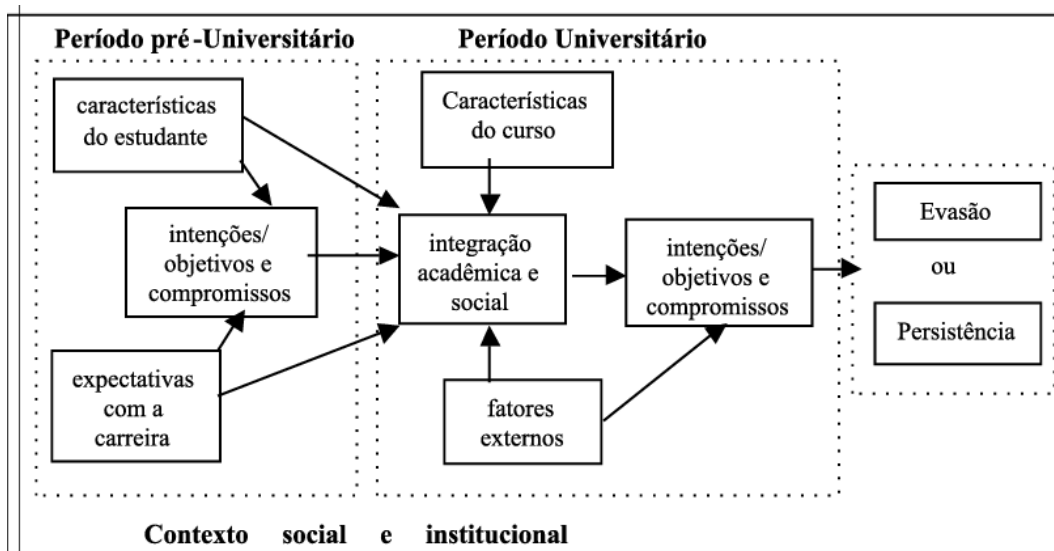
(Lobo e Silva et al.; 2007). O desempenho acadêmico é uma problemática que tem levado muitos estudiosos a pesquisarem as suas causas, a fim de que possam ser encontradas alternativas para ajudar no processo de gestão.

No Brasil, segundo Almeida e Veloso (2002), em alguns casos, o baixo desempenho no ensino médio reflete no desempenho das primeiras disciplinas do curso superior, resultando em abandono do curso pelas reprovações nos primeiros semestres. Outro caso seria o fato de o aluno buscar o curso de baixa procura com o objetivo de, após ter ingressado, tentar a transferência interna para o curso de seu real interesse.

Por outro lado, Silva (2006) indica que a magnitude da evasão está diretamente relacionada com a qualificação do corpo docente e não somente com o status socioeconômico dos estudantes. Para ele, os jovens amadurecem mais tarde, fazendo algumas vezes escolhas prematuras no que diz respeito ao seu futuro profissional.

O modelo da relação aprendizagem/permanência do Tinto (1997) destaca a influência de aspectos pessoais e sociais existentes antes do ingresso no curso universitário (Figura 1), como aspectos relacionados com o contato acadêmico, metodologia de aprendizagem e integração institucional.

Figura 1 - Modelo da relação aprendizagem/permanência



Fonte: Tinto, 1997 apud Rigo, 2016

Com o aumento do uso de sistemas informatizados nas universidades, cresce cada dia mais o volume de dados gerado e armazenado em bases de dados. Este grande volume de dados tem impulsionado o interesse na sua utilização. A mineração de dados educacionais é um campo de investigação ainda não consolidado, está relacionada à aplicação de técnicas da mineração de

dados junto aos mais diversos domínios de dados obtidos em diversos contextos educacionais, em sua grande maioria provenientes de ensino a distância (MANHÃES, 2011).

Para Scheuer e McLaren (2012), a Mineração de Dados Educacionais concentra-se em elaborar, estudar e executar procedimentos computadorizados para descobrir padrões em grandes conjuntos de dados educacionais que, de outra maneira, seriam difíceis ou improváveis de explorar devido ao vasto volume de dados que existem. A mineração de dados educacionais e seus recursos, possibilitam desenvolver ou adaptar métodos e algoritmos de mineração existentes, para que esses possam apoiar efetivamente processos de detecção de comportamentos ligados à evasão escolar, de tal modo que seja possível compreender melhor os dados em contextos educacionais, produzidos principalmente por alunos e professores, considerando os ambientes nos quais eles interagem (RIGO, 2012).

A mineração de dados ocorre, normalmente, com alguns objetivos básicos:

- Predição: a predição é um dos objetivos fundamentais da mineração de dados, utiliza algumas variáveis que se encontram no banco de dados, com a finalidade de prever valores desconhecidos. (JUNIOR, 2015 apud WITTEN; FRANK; HALL, 2011). Exemplo: Predizer o percentual que será aumentado de tráfego na rede se a velocidade aumentar; (CAMILO, 2009);
- Classificação: o objetivo de um algoritmo de classificação é encontrar alguma correlação entre os atributos e uma classe, de modo que o processo de classificação possa usá-lo para prever a classe de um exemplo novo e desconhecido. (COSTA et al., 2013). Exemplo: Diagnosticar onde uma determinada doença pode estar presente; (CAMILO, 2009);
- Associação: a associação consiste na descoberta de regras que mostram quais atributos estão relacionados. Seu principal objetivo é encontrar conjuntos de itens ou eventos que ocorram junto. Exemplo: Determinar os casos onde um novo medicamento pode apresentar efeitos colaterais; (CAMILO, 2009);
- Agrupamento: a tarefa de agrupamento, também conhecida como “clustering”, é uma técnica onde os algoritmos de agrupamento possuem aprendizado não supervisionado. Com esta técnica se espera conhecer novos atributos alvos a partir de conjunto de dados, sem ter classificação prévia (JUNIOR, 2015 apud COSTA et al, 2013).

Considerando as áreas de estudo e os conceitos discutidos nesta seção é possível analisar alguns trabalhos relacionados.

3. Trabalhos relacionados

No trabalho de Tinto (1975), foi definido um modelo teórico para explicar as causas da evasão discente, considerando o processo de desgaste do aluno como uma interação sócio psicológico entre as características do aluno universitário.

Já em Santos et. al (2016), os autores apresentam um modelo computacional conceitual sobre a aplicação de técnicas de Mineração de Dados com foco na Análise de Sentimentos de forma a despertar a percepção dos alunos acerca das práticas docentes, além disso buscam compreender as expectativas e limitações dos estudantes.

No tocante a aplicação de Mineração de Dados Educacionais para o estudo da evasão, Manhães et al. (2011) realizaram um estudo para identificar a eficiência de dez algoritmos de classificação na previsão de estudantes com risco de evasão utilizando dados acadêmicos das primeiras notas semestrais referentes a alunos de graduação da universidade brasileira UFRJ. Os resultados mostraram que é possível identificar com precisão de 80% a situação final do aluno no curso.

Junior (2015) utilizou os dados de uma universidade comunitária brasileira - UNISIC, referente a um determinado período, para identificar alunos com perfil de evasão no ensino superior. A tarefa de classificação foi a base do trabalho, juntamente com a técnica de árvore de decisão e o algoritmo C4.5 (J48). Os experimentos foram feitos na plataforma WEKA³.

Digiampietri (2016), propõe uma metodologia baseada em mineração de dados para o acompanhamento e identificação precoce dos estudantes com grande potencial de desistência ou desligamento compulsório. Nascimento et. al (2018), propôs um trabalho utilizando bases de dados educacionais fornecidas pelo INEP e aplicou técnicas de mineração de dados com a finalidade de melhor explicar indicadores como a evasão e a reprovação escolar. Foi utilizado os modelos de regressões linear e robusta.

Os trabalhos relacionados aqui apresentados abordam o tema deste projeto de pesquisa de diversas formas diferentes e contribuíram para a definição das técnicas e ferramentas utilizadas que são apresentadas a seguir.

³ <https://www.cs.waikato.ac.nz/ml/weka/>

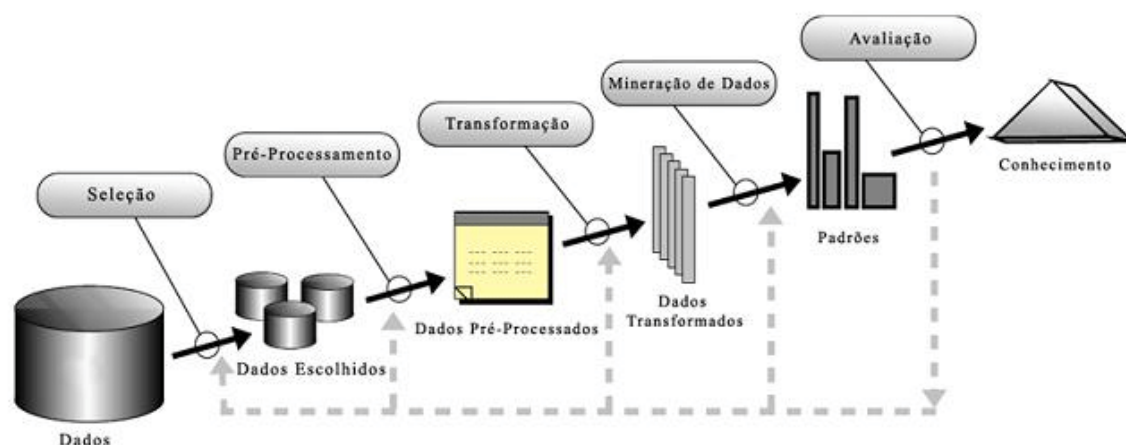
4. Ferramentas e Método

O Curso de Sistema de Informação da Faculdade de Tecnologia da Unicamp possui um currículo pleno com 42 disciplinas, divididas em 08 semestres, totalizando uma carga horária de 3060 horas. O que permite uma formação ampla e sólida para o mercado de trabalho. (Faculdade de Tecnologia, 2019).

A fonte de dados para esse trabalho é oriunda de sistemas acadêmicos da Unicamp, relativo aos alunos matriculados no curso em questão, preservando os dados pessoais por não serem importantes ao traçar padrões de evasão. Os dados estão organizados principalmente em duas planilhas: a planilha Vida_Academica que apresenta dados dos alunos como: id, ano e período de ingresso, ano de egresso, nota do vestibular, nível do curso (se é um curso de tecnologia, graduação, pós-graduação e outros) e idade; e a planilha Historico_Escolar que possui informações como id, cod_disc, nota, sit_disc, cr.

Este trabalho optou por utilizar o método KDD em seu desenvolvimento.

Figura 2 - Processo de KDD



Fonte: FAYYAD et al. 1996

Segundo Castanheira (2008), KDD (Knowledge Discovery in Databases) é um processo de descoberta de conhecimento em bases de dados que tem como objetivo principal extrair conhecimento a partir de grandes bases de dados. O processo de KDD consiste em uma sequência de etapas que devem ser executadas sequencialmente, pois ao final de cada etapa, o resultado obtido serve de auxílio para a etapa seguinte, podendo repetir etapas anteriores sempre que necessário.

Inicialmente os dados são tratados para que a fase de mineração de dados seja executada somente em relação aos dados importantes da base de dados. Desta maneira, como apresentado

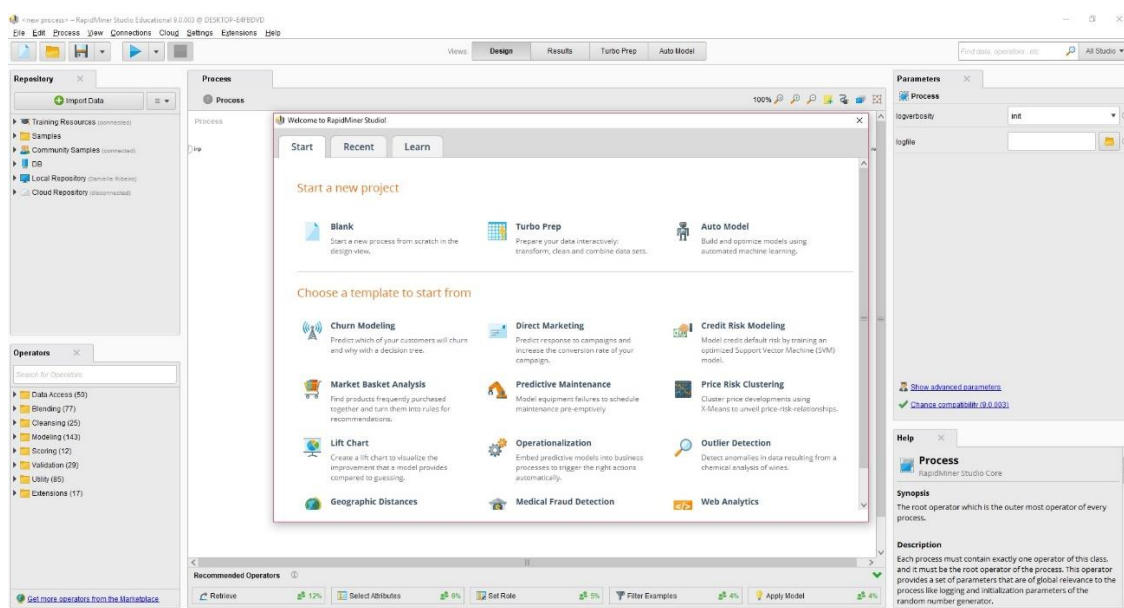
na Figura 2, os dados são selecionados, pré-processados e transformados de acordo com os algoritmos escolhidos. Em seguida as técnicas de mineração de dados são aplicadas para obtenção de padrões que são avaliados para a obtenção da informação.

Para o presente trabalho foi escolhido a técnica de agrupamento, ou “clustering”. Essa técnica tem a função de unir dados em grupos com conteúdo semelhantes e dados diferentes em grupos distintos gerando uma diferenciação de conteúdo em cada grupo (Castanheiras, 2008). Neste caso os dados devem ser agrupados de forma a indicar os alunos que evadiram por suas características em comum.

Atualmente a análise de dados deste tipo é realizada com auxílio de ferramentas computacionais. Para este projeto a ferramenta escolhida foi o RapidMiner®. Esta ferramenta é um sistema para análise de dados que utiliza a aprendizagem de máquina, podendo ser considerado uma alternativa para a ferramenta WEKA, que é uma das mais utilizadas no mundo. Inclusive a opção por utilizar esta ferramenta e não o framework WEKA é uma tentativa de diversificação nas abordagens encontradas nos trabalhos relacionados.

Pelo fato desta ferramenta ser desenvolvida na linguagem de programação Java, permite a sua utilização versátil em qualquer sistema operativo e ambiente de trabalho. O projeto RapidMiner® começou em 2001 por Ralf Klinkenberg, Ingo Mierswa e Simon Fischer na Unidade de Inteligência Artificial da Universidade de Dortmund (Alemanha). Em 2006, Ingo Mierswa e Ralf Klinkenberg fundaram a empresa Rapid-I. (RapidMiner, 2019)

Figura 3 - Tela inicial RAPIDMINER



O RapidMiner® oferece um manual bem elaborado além de uma introdução aos conceitos fundamentais sobre mineração de dados. É possível encontrar vasto material em formato de vídeos e tutoriais sobre como utilizar a ferramenta.

Com base no que foi apresentado foram realizadas as etapas do KDD e os resultados deste trabalho são apresentados a seguir.

5. Resultados

O estudo realizado neste trabalho foi feito em duas etapas: primeiramente foram analisados diretamente os dados, com o objetivo de encontrar atributos desnecessários e melhorar organizar os dados. Em um segundo momento, tratando os dados diretamente na ferramenta RAPIDMINER, foram realizadas análises sobre os resultados dos algoritmos de agrupamento (ou “clusterização”) pelos quais os dados foram processados.

Após a seleção dos dados, seguindo as etapas do processo de KDD, o passo seguinte foi a limpeza da base criada, exclusão dos atributos não considerados importantes para a mineração, como o número do cr, nota do vestibular, nível do curso (se é um curso de tecnologia, graduação, pós-graduação e outros).

Foi criada então uma tabela contendo apenas as informações relevantes das duas tabelas iniciais. O objetivo dessa ação foi armazenar o maior número de informações relevantes referentes ao perfil do aluno no decorrer do curso como, por exemplo, quantas disciplinas cursou em cada semestre, quantas vezes abandonou disciplinas durante o semestre e o curso, qual o status do curso.

Com os dados selecionados e pré-processados, a próxima etapa no processo de KDD é a transformação dos dados, principalmente os atributos que apresentam valores muito abrangentes, o que dificulta o processamento e entendimento dos resultados. Nesta etapa foi realizada a substituição do status_curso de texto para números com intervalo de 10 unidades entre os possíveis status.

Tabela 1 - Alteração feita nos status_curso para concatenação das tabelas.

Status_Curso (antes da transformação)	Status_Curso (após a transformação)
Abandono	10
Ativo	20
Concluído	30
Desligamento	40
Ingresso em outro curso	50
Integralização Excedida por Projeção	60
Matrícula Cancelada	70

Fonte: Autor

A análise prévia dos dados permitiu um melhor entendimento da base de dados e do problema considerado no trabalho. Buscando a identificação do perfil do aluno evasor foram criados dois questionamentos para serem respondidos através do estudo:

- Quem reprova em disciplinas consideradas chave do curso, como lógica, algoritmo e arquitetura, desiste?
- Há alguma relação entre a idade e a evasão?

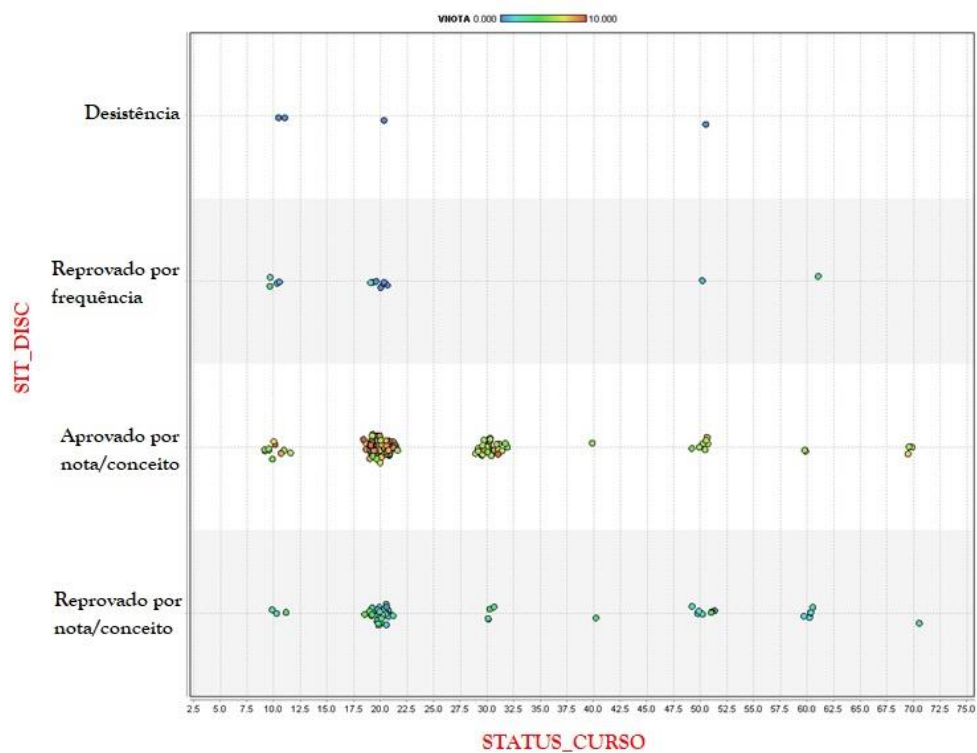
Inicialmente foram classificados os alunos que reprovaram nas principais disciplinas da grade curricular. As seguintes disciplinas foram consideradas disciplinas chave do curso com base em análises junto à coordenação de curso:

- SI100 – Algoritmos e Programação de Computadores I
- SI101 – Fundamentos de Sistema de Informação
- SI120 – Lógica Matemática
- TT106 – Organização e Arquitetura de Computadores

Neste experimento, foram criados seis arquivos, um para cada disciplina. Submetemos cada um dos arquivos ao algoritmo de agrupamento, utilizando as configurações padrões do RAPIDMINER, para tentar descobrir se existe um padrão que reprovam nestas disciplinas.

A figura 04 mostra o perfil dos alunos que reprovaram na disciplina SI100. Foram selecionados 317 registros de alunos que cursaram essa disciplina e foi observado que tantos 10% de alunos que reprovaram ou desistiram da matéria também evadiram do curso ou trocaram de curso.

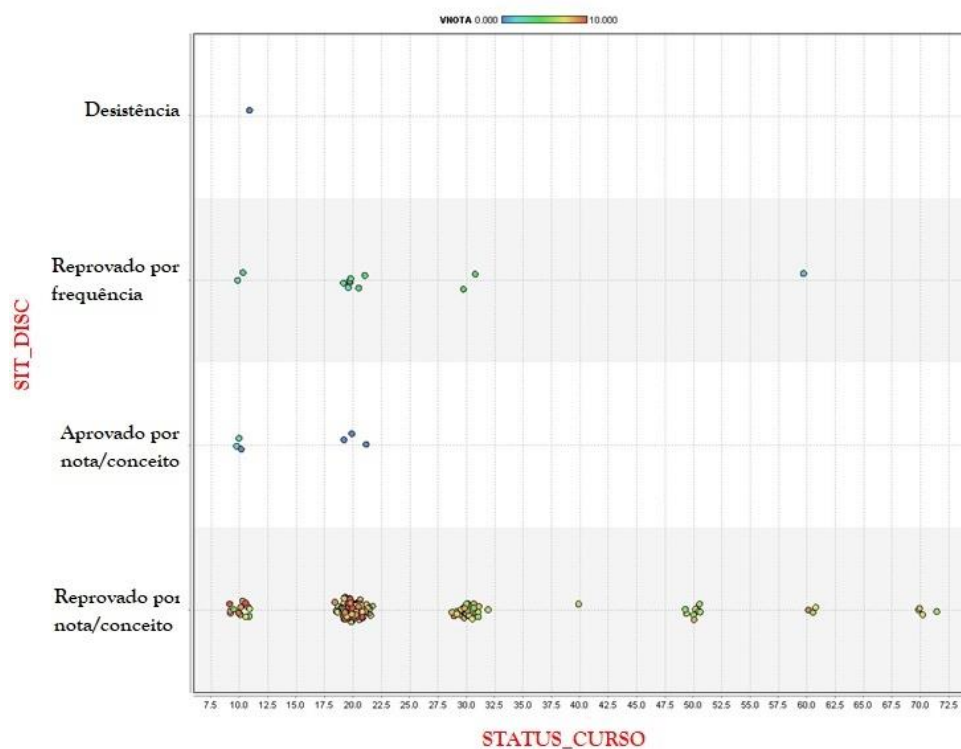
Figura 4 - Disciplina com maior número de alunos desistentes – SI100



Fonte: Autor

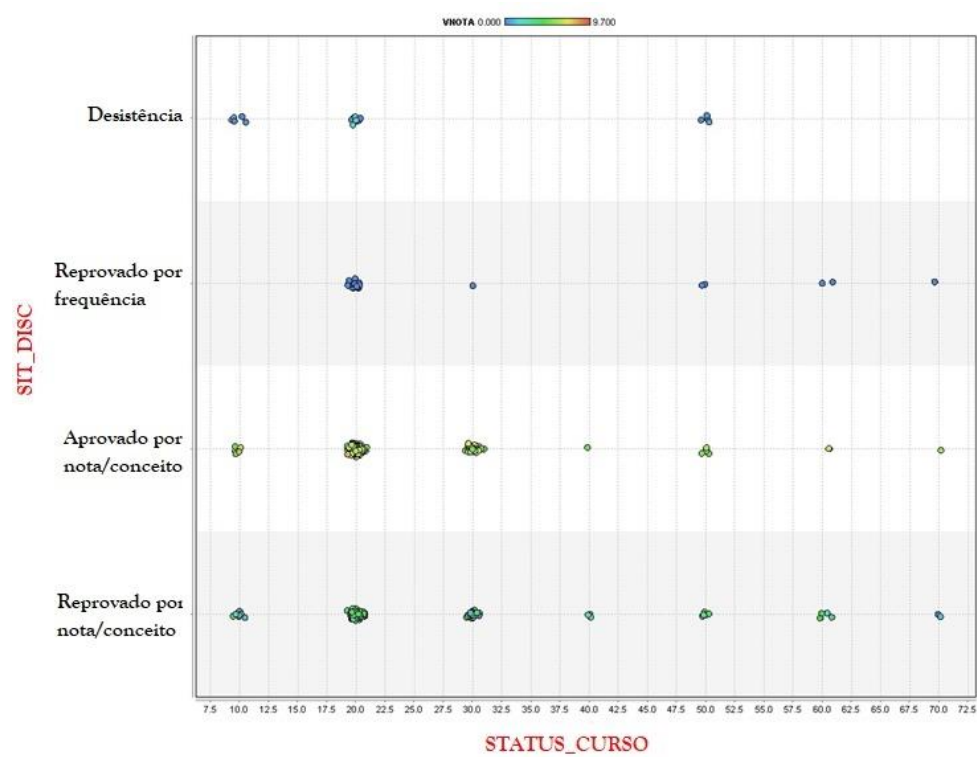
Foi feito a mesma análise para as outras disciplinas e obtivemos os seguintes resultados.

Figura 5 - Disciplina SI101 – 279 registros e 37 desistentes



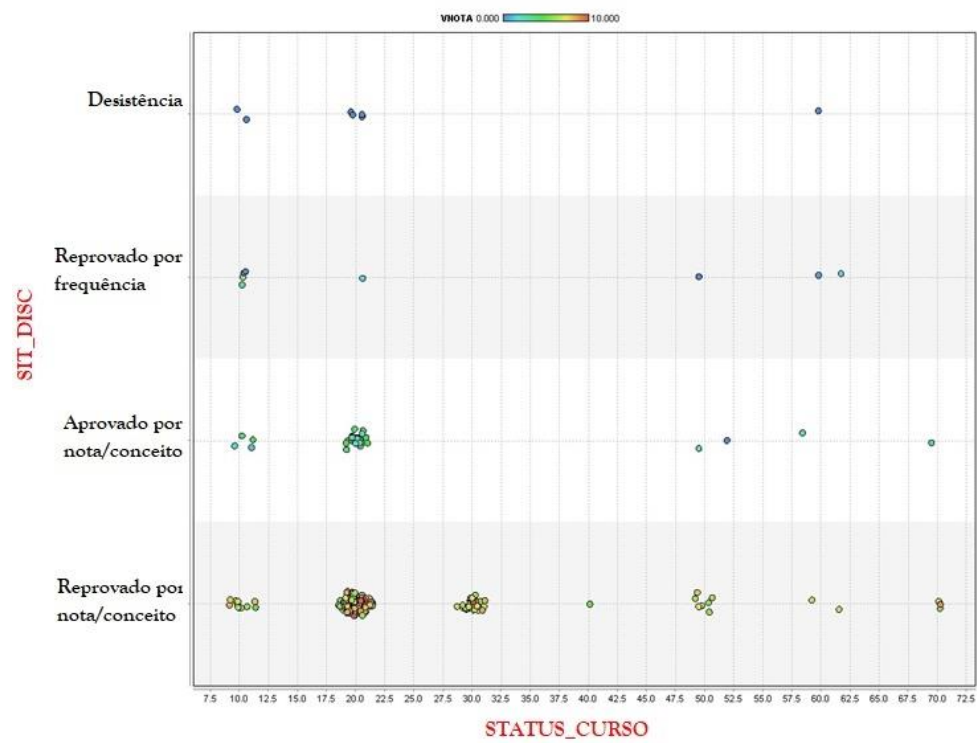
Fonte: Autor

Figura 6 - Disciplina SI120 – 409 inscritos e 52 desistentes



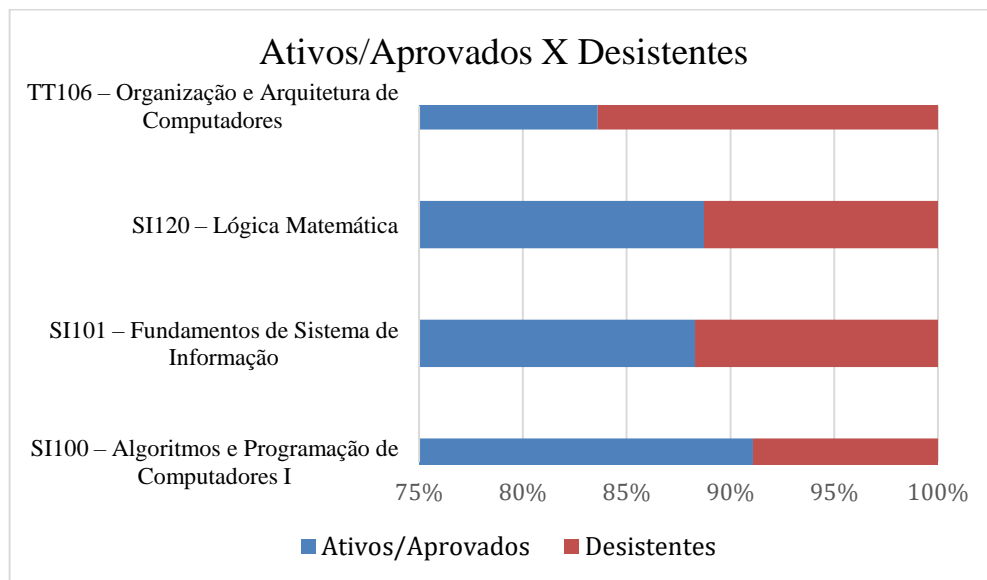
Fonte: Autor

Figura 7 - Disciplina TT106 – 290 inscritos e 41 desistentes



Fonte: Fonte

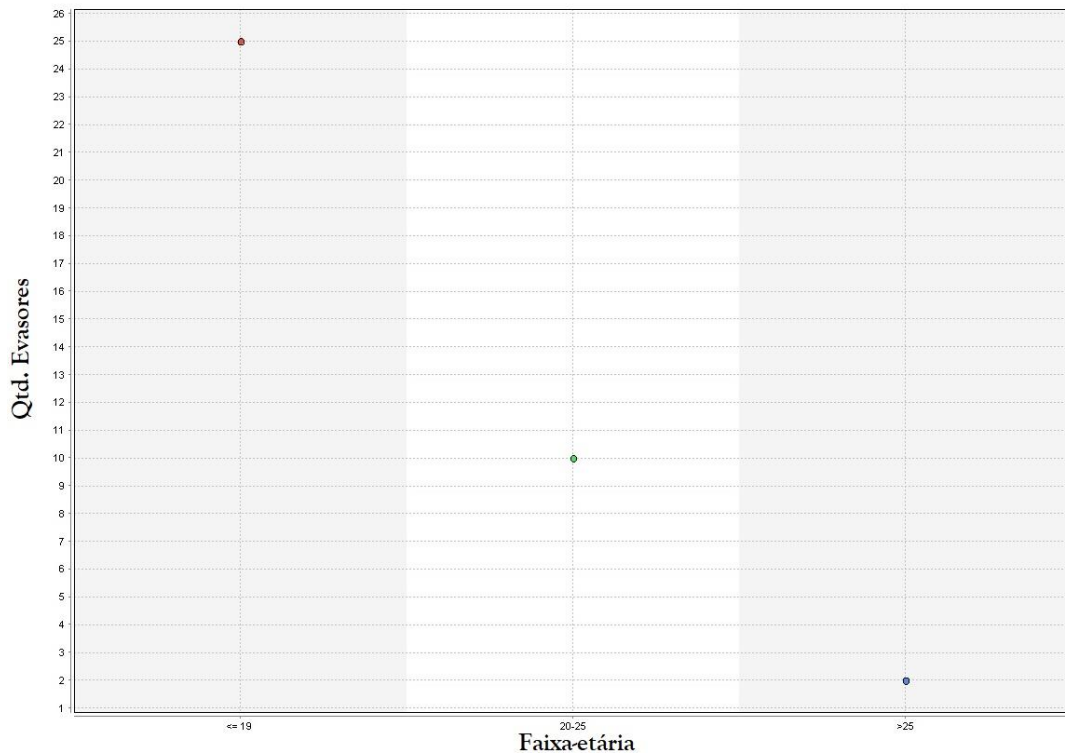
Os resultados mostraram que os alunos que reprovaram nas disciplinas citadas, dificilmente continuam no curso. Foi realizado um gráfico no Microsoft Excel para exemplificar a porcentagem do teste realizado.



Utilizou-se, em seguida, da tabela Vida_Academica e aplicou-se o algoritmo de clusterização a fim de poder observar grupos de alunos que evadiram e que não evadiram conforme a idade. Após a análise foi possível visualizar que a grande maioria dos alunos que ingressou no curso, o fazem com menos de 20 anos, outra parcela significativa está na faixa etária entre 20 e 25 anos.

O resultado foi a confirmação de diversos estudos realizados que comprovam que há uma grande tendência em evasão quando o ingressante é menor de 19 anos, pois há um questionamento quanto maturidade que a faculdade exige. Além da dúvida na escolha do curso certo. A figura 12 mostra a relação da quantidade de alunos evasores pela faixa-etária dos ingressantes.

Figura 8 - Faixa-etária x Evasões



Fonte: Autor

6. Conclusões

Como foi citado durante todo o trabalho, não foram utilizados os dados pessoais dos alunos em questão, nem tampouco dados de ordem socioeconômica. Sendo assim, o trabalho ficou limitado à vida curricular do aluno, permitindo apenas traçar o perfil evasor com base no ambiente interno da universidade e não levando em consideração o ambiente externo da instituição.

Considerando os estudos realizados, foram analisados aspectos do problema de evasão escolar, que constitui um elemento importante a ser tratado para que um número maior de estudantes finalize com sucesso o ensino superior.

É de extrema importância pesquisas de novas medidas para combater esta dificuldade enfrentada pelas instituições de ensino, uma vez que faz-se necessário averiguar as causas da evasão para nortear ações que venham a minimizar este problema.

Foi possível notar a importância da mineração de dados e dos estudos nessa área, pois essa ferramenta permite a realização de estudos detalhados do padrão de alunos com perfil de evasão, além de permitir a avaliação de medidas preventivas.

Sendo assim, este trabalho obteve sucesso e identificou alunos com perfis de evasão universitária, utilizando mineração de dados, tendo como público-alvo os alunos do curso de Sistema de Informação da Faculdade de Tecnologia da Unicamp.

Após os experimentos e estudos apresentados, conclui-se que há dois perfis de alunos que os gestores da unidade em questão devem se atentar: os que reprovaram nas disciplinas chaves do curso, principalmente, a disciplina Algoritmos e Programação de Computadores I, pois o risco de ele evadir é alto; e os alunos que ingressam no curso com menos de 19 anos, pois como foi apresentado, também é um perfil de possível evasor.

Com este estudo, pode-se estender trabalhos futuros de diversas formas. Uma delas é na criação de um módulo que automatize as etapas realizadas manualmente neste trabalho, além da criação de uma interface intuitiva e simples para auxiliar os gestores na resolução deste problema.

7. Referências Bibliográficas

ANDRIOLA, W. B.; ANDRIOLA, C. G.; MOURA, C. P. **Opiniões de docentes e de coordenadores acerca do fenômeno da evasão discente dos cursos de graduação da Universidade Federal do Ceará (UFC)**. Rio de Janeiro-RJ: Ensaio: Avaliação e Políticas Públicas em Educação, 2006.

BAKER, R., ISOTANI, S., CARVALHO, A. (2011), **Mineração de Dados Educacionais: Oportunidades para o Brasil**. **Revista Brasileira de Informática na Educação**. In: BRAZILIAN JOURNAL OF COMPUTERS IN EDUCATION (RBIE, 2011).

BRUM, Flávio de; MOZZAQUATRO, Patricia Mariotto; ZANATTA, Jocias Maier. ESTUDO SOBRE OS ALGORITMOS DE CLUSTERIZAÇÃO HIERARCHICAL CLUSTERER E SIMPLE K-MEANS APLICADOS NO AGRUPAMENTO DE PADRÕES SIMILARES. Revista da Universidade Vale do Rio Verde, [s.l.], v. 17, n. 1, p.1-9, 2019. Universidade Vale do Rio Verde (UninCor). <http://dx.doi.org/10.5892/ruvrd.v17i1.4957>.

CAMILO, C. O.; SILVA, J. C. **Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas**. Instituto de Informática, Universidade Federal de Goiás, Goiânia – GO, 2009.

CORNELIUS JUNIOR, R. **USO DA MINERAÇÃO DE DADOS NA IDENTIFICAÇÃO DE ALUNOS COM PERFIL DE EVASÃO DO ENSINO SUPERIOR**. 2015. 138 f. TCC (Graduação) - Curso de Ciência da Computação, Universidade de Santa Cruz do Sul, Santa

Cruz do Sul, 2015. Disponível em: <<https://repositorio.unisc.br/jspui/bitstream/11624/535/1/Romeu%20Cornelius%20Junior%20-%20TCC%20-%20Final.pdf>>. Acesso em: 15 de jan. 2019.

FAYYAD, U. M., Piatetsky-Shapiro, G., Smyth, P., et al. (1996). **Knowledge discovery and data mining: towards a unifying framework**. In KDD, volume 96, pages 82– 88.

FRIES, D.; BERNARD, D., R., P. **Avaliação dos índices de evasão nos cursos de graduação da Universidade do Estado de Santa Catarina – UDESC**. Sorocaba, SP, v. 21, n. 2, p. 503-521, jul. 2016.

GONÇALVES, E. J., de Oliveira, M. A., Junior, J. H. F., Feitosa, G. E., Mendes, D. H., Cortes, M. I., Feitosa, R. G., and Lopes, Y. S. (2014). **Uma abordagem baseada em agentes de apoio ao ensino a distância utilizando técnicas de engenharia de software**.

GONÇALVES, Tayná Costa; SILVA, Josenildo Costa da; CORTES, Omar Andres Carmona. **Técnicas de mineração de dados: um estudo de caso da evasão no ensino superior do Instituto Federal do Maranhão**. Revista Brasileira de Computação Aplicada, [s.l.], v. 10, n. 3, p.11-20, 21 set. 2018. UPF Editora. <http://dx.doi.org/10.5335/rbca.v10i3.8427>.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior**. Brasília: INEP, 2016. Disponível em: <http://download.inep.gov.br/educacao_superior/censo_superior/documentos/2016/notas_sobre_o_censo_da_educacao_superior_2016.pdf>. Acesso em: 15 ago. 2018.

KAMPFF, A. C. **Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem como Apoio à Prática Docente**, Tese de Doutorado, Universidade Federal do Rio Grande do Sul, Porto Alegre –RS, 2009.

LANES, Mariele; ALCÂNTARA, Cleber. **Predição de Alunos com Risco de Evasão: estudo de caso usando mineração de dados**. Anais do Xxix Simpósio Brasileiro de Informática na Educação (sbie 2018), [s.l.], p.1921- 1925, 28 out. 2018. Brazilian Computer Society (Sociedade Brasileira de Computação - SBC). <http://dx.doi.org/10.5753/cbie.sbie.2018.1921>.

LIMA, Franciele Santos de; ZAGO, Nadir. **Desafios conceituais e tendências da evasão no ensino superior: a realidade de uma universidade comunitária**. Revista Internacional de

Educação Superior, Campinas, SP, v. 4, n. 2, p. 366-386, abr. 2018. ISSN 2446-9424. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/riesup/article/view/8651587/17797>>. Acesso em: 21 ago. 2018.

LIRA, Kaynan & De Oliveira, Marcos & Gonçalves, Enyo & Pires Magalhaes, Regis. (2016). **Utilizando Mineração de Dados e Sistemas Multiagentes na Análise da Evasão em Educação a Distância por meio do Perfil dos Alunos.**

LOBO, Maria Beatriz de Carvalho Melo. **Panorama da Evasão no Ensino Superior Brasileiro: Aspectos gerais das causas e soluções.** Instituto Lobo para o Desenvolvimento da Educação, da Ciência e da Tecnologia. Mogi das Cruzes, SP: 2012. Disponível em: <http://www.institutolobo.org.br/imagens/pdf/artigos/art_087.pdf>. Acesso em: 16 ago. 2018.

LOBO, Roberto Leal et. al. **A evasão no Ensino Superior Brasileiro.** Cadernos de Pesquisa, v. 37, n. 132. Disponível em <http://www.scielo.br/scielo.php?pid=S0100-15742007000300007&script=sci_arttext#tab07>. Acesso em: 10 jul 2018.

MASCHIO, Pedro et al. **Um Panorama acerca da Mineração de Dados Educacionais no Brasil.** Anais do Xxix Simpósio Brasileiro de Informática na Educação (sbie 2018), [s.l.], p.1936-1940, 28 out. 2018. Brazilian Computer Society (Sociedade Brasileira de Computação - SBC). <http://dx.doi.org/10.5753/cbie.sbie.2018.1936>.

NASCIMENTO, Rafaella Leandra Souza do; CRUZ JUNIOR, Geraldo Gomes da; FAGUNDES, Roberta Andrade de Araújo. **Mineração de Dados Educacionais: Um Estudo Sobre Indicadores da Educação em Bases de Dados do INEP.** Renote, [s.l.], v. 16, n. 1, p.1-11, 21 ago. 2018. Universidade Federal do Rio Grande do Sul. <http://dx.doi.org/10.22456/1679-1916.85989>.

PAZ, F. J.; CAZELLA, S. C. **Identificando o perfil de evasão de alunos de graduação através da Mineração de Dados Educacionais: um estudo de caso de uma Universidade Comunitária.** In: Congresso Brasileiro de Informática na Educação (CBIE 2017), 2017, Recife.

RIGO, S. J.; CAZELLA, S. C.; CAMBRUZZI, W. **Minerando Dados Educacionais com foco na evasão escolar: oportunidades, desafios e necessidades.** In: Anais do Workshop de Desafios da Computação Aplicada à Educação. 2012.

RIGO, S. J.; CAMBRUZZI, W; BARBOSA, J. L. V.; CAZELLA, S. C. **Aplicações de Mineração de Dados Educacionais e Learning Analytics com foco na evasão escolar: oportunidades e desafios.** In: Revista Brasileira de Informática na Educação, Volume 22, Número 1, 2014.

RIGO, Sandro José; BARBOSA, Jorge; CAMBRUZZI, Wagner. **Educação em Engenharia e mineração de dados educacionais: oportunidades para o tratamento da evasão.** EaD & Tecnologias Digitais na Educação, Dourados, v. 2, n. 3, p. 30- 40, nov. 2014. ISSN 2318-4051. Disponível em: <<http://ojs.ufgd.edu.br/index.php/ead/article/view/3409/2106>>. Acesso em: 21 ago. 2018.

SILVA, Cristovão et al. **Avaliação dos níveis de expectativa e motivação de estudantes de graduação do ensino presencial através da Mineração de dados.** Anais dos Workshops do Vii Congresso Brasileiro de Informática na Educação (cbie 2018), [s.l.], p.409-418, 28 out. 2018. Brazilian Computer Society (Sociedade Brasileira de Computação - SBC). <http://dx.doi.org/10.5753/cbie.wcbie.2018.409>.

TAN, P.; STEINBACH, M.; KUMAR, V.; FERNANDES, A. **Introdução ao Datamining: mineração de dados.** Rio de Janeiro: Editora Ciência Moderna, 2009.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data Mining: Practical machine learning tools and techniques.** 3. ed.. Morgan Kaufmann, 2011.