



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Evasão do curso de Licenciatura em Matemática (Noturno) da Universidade de Brasília

Lavousier Ferreira Diniz de Souza

Monografia apresentada como requisito parcial
para conclusão do Curso de Computação — Licenciatura

Orientadora

Prof.^a Dr.^a Maria Emilia Machado Telles Walter

Coorientador

Prof. Dr. Jan Mendonça Corrêa

Brasília
2016

Dedicatória

Dedico esse trabalho a minha mãe e irmã, pelo apoio incondicional e aos professores pelo conhecimento.

Agradecimentos

Agradeço a minha mãe e irmã que sempre me deram apoio e suporte para que pudesse chegar até aqui. A professora Maria Emília e ao professor Jan Mendonça Correia que me deram seus votos de confiança para que pudesse realizar esse trabalho.

Resumo

A evasão universitária representa um grande desafio para as instituições de ensino superior em todo o mundo, particularmente em países em desenvolvimento como o Brasil. Existem diferentes conceitos para evasão, mas o adotado neste trabalho é a evasão de curso, ou seja, consideramos evadido o aluno que ingressou no curso e saiu sem concluir esse curso. Na Universidade de Brasília (UnB), existem estudos sobre evasão, no Decanato de Ensino de Graduação, e em diferentes Departamentos. Em particular, no Departamento de Matemática da UnB, tanto quanto sabemos, não existem análises específicas sobre as causas da evasão no Curso de Licenciatura (Noturno). Assim, o objetivo principal deste trabalho é apresentar informações sobre a relação entre reprovações em disciplinas obrigatórias e os altos índices de evasão nesse curso. Para isso, inicialmente, foram feitas análises estatísticas para identificar os índices de reprovação nas disciplinas obrigatórias do Curso. Em seguida, técnicas de mineração de dados, com os softwares WEKA e ORANGE, foram aplicadas em todas as disciplinas obrigatórias, o que permitiu identificar conjuntos de disciplinas que estão mais relacionadas à evasão. Este trabalho visa colaborar para elaboração de políticas de combate à evasão no Curso de Licenciatura em Matemática (Noturno).

Palavras-chave: evasão, mineração de dados, Licenciatura em Matemática, Universidade de Brasília, UnB, análises estatísticas

Abstract

Higher education dropout represents a major challenge for institutions around the world, including in developing countries such as Brazil. There are different concepts for dropout, but the adopted in this work is a course dropout, that is, we consider evaded the student who entered the course and did not graduate. At the Universidade de Brasília (UnB), there are studies on evasion, in the Deanship of Undergraduate Studies, and in different Departments. In particular, in the Department of Mathematics at UnB, as far as we know, there are no analyzes on the causes of evasion in the Course of Deanship of Undergraduate Studies in Mathematics (Nocturnal). Thus, the main objective of this work is to present information about a relation between the reprobations in compulsory subjects and the high level of dropout rate in this course. For this, initially, statistical analyzes were performed to identify the reprobations rates in the compulsory subjects of the Course. Next, data mining techniques with WEKA and ORANGE software were applied in all compulsory subjects, which allowed to identify sets of disciplines that are more related to evasion. This work aims to collaborate to elaborate policies to combat evasion in Deanship of Undergraduate Studies in Mathematics (Nocturnal).

Keywords: *dropout rate, data mining, Teaching Degree in Mathematics, statistical analysis, Universidade de Brasília, UnB*

Sumário

1	Introdução	1
1.1	Justificativa	2
1.2	Problema	2
1.3	Objetivos	2
1.4	Descrição dos Capítulos	3
2	A Evasão no Ensino Superior	4
2.1	Definição de Evasão	4
2.2	Fatores para Evasão do Aluno	5
2.3	Evasão em Nível Internacional	8
2.4	Evasão no Brasil	13
2.5	Evasão nas Licenciaturas	15
2.6	Evasão na Universidade de Brasília	17
3	Mineração de Dados	20
3.1	Dado, Informação e Conhecimento	20
3.2	Definição de Mineração de Dados	20
3.3	Processos de Extração de Conhecimento de Dados	21
3.4	Tarefas de Mineração	22
3.5	Software WEKA	23
3.6	Software Orange	28
4	A Licenciatura em Matemática (Noturno) e Análises Estatísticas	30
4.1	Licenciatura em Matemática (Noturno)	30
4.1.1	Fluxo de Disciplinas e Disciplinas Obrigatórias	31
4.2	Dados dos Alunos	32
4.3	Análises Estatísticas	34
4.3.1	Formas de Ingresso e saída do curso	34
4.3.2	Disciplinas por Semestre	36

5	Disciplinas Obrigatórias Relacionadas à Evasão	55
5.1	Mineração de Dados com o WEKA	55
5.2	Mineração de Dados com o Orange	57
5.3	Consolidação dos resultados obtidos	64
6	Conclusão e Trabalhos Futuros	65
6.1	Contribuições	65
6.2	Trabalhos futuros	66
	Referências	67

Lista de Figuras

2.1	Taxa de evasão no ensino superior, países da OECD, em 2011 (adaptado de Shimosegawa [30]).	8
2.2	Porcentagem de estudantes com e sem empréstimo para pagar o ensino superior, por situação, Estados Unidos, 2001 (adaptado de Gladieux e Perna [18]).	11
2.3	Porcentagem de estudantes por categoria de constituição de empréstimo, evasão ou conclusão de curso e início tardio de curso superior, estudo em meio período e trabalho em período integral, Estados Unidos, 2001 (adaptado de Gladieux e Perna [18]).	11
2.4	Porcentagem de estudantes que constituíram empréstimo, que completaram o bacharelado ou sofreram evasão (adaptado de Gladieux e Perna [18]).	12
2.5	Porcentagem de estudantes por categoria de constituição de empréstimo ou não e evasão e início tardio de curso superior, estudo em meio período e trabalho em período integral, Estados Unidos, 2001 (adaptado de Gladieux e Perna [18]).	12
2.6	G1 - Desligamento por abandono do curso. Fonte: Gomes et al [19].	14
2.7	G2 – Desligamento voluntário. Fonte: Gomes et al [19].	14
2.8	G3 – Desligamento por não cumprimento de condição. Fonte: Gomes et al [19].	15
2.9	Evolução das matrículas de Educação Superior de Graduação, por Grau Acadêmico - Brasil 2003 - 2013. Fonte: INEP [12].	15
2.10	Quantidade de evadidos no Curso de Licenciatura em Informática na IFRN. Fonte: Souza et al [31].	16
2.11	Alunos evadidos e não evadidos por rendimento no curso em 2004 na UnB. Fonte: Cardoso [7].	18
2.12	Alunos evadidos e não evadidos por rendimento no curso em 2005 na UnB. Fonte: Cardoso [7].	18
2.13	Taxa de evasão aparente, total de alunos e grupos de curso por prestígio social (%), na UnB em 2004 e 2005. Fonte: Cardoso [7].	19

3.1	Tela inicial do software WEKA.	23
3.2	Tela pré-processamento do software WEKA.	24
3.3	Tela classificação do software WEKA.	25
3.4	Tela clusterização do software WEKA.	25
3.5	Tela associação do software WEKA.	26
3.6	Tela de seleção de atributos do software WEKA.	27
3.7	Tela de visualização de atributos do software WEKA.	27
3.8	Tela inicial do software Orange.	29
4.1	Fluxo de disciplinas do Curso de Licenciatura em Matemática (Noturno) da Universidade de Brasília.	31
4.2	Quantidade de alunos por forma de ingresso.	34
4.3	Porcentagem das diferentes formas de evasão.	35
4.4	Porcentagem de reprovação na disciplina Organização da Educação Brasileira.	36
4.5	Porcentagem de reprovação na disciplina Cálculo 1.	37
4.6	Porcentagem de reprovação na disciplina Introdução a Ciência da Compu- tação.	38
4.7	Porcentagem de reprovação na disciplina a Introdução a Álgebra Linear.	39
4.8	Porcentagem de reprovação na disciplina Cálculo 2.	40
4.9	Porcentagem de reprovação na disciplina Física 1.	41
4.10	Porcentagem de reprovação na disciplina Física 1 Experimental.	42
4.11	Porcentagem de reprovação na disciplina Desenvolvimento Psicológico e Ensino.	43
4.12	Porcentagem de reprovação na disciplina Física 2.	44
4.13	Porcentagem de reprovação na disciplina Cálculo 3.	45
4.14	Porcentagem de reprovação na disciplina Teoria dos Números.	46
4.15	Porcentagem de reprovação na disciplina Álgebra 1.	47
4.16	Porcentagem de reprovação na disciplina de Aprendizagem no Ensino.	48
4.17	Porcentagem de reprovação na disciplina Geometria 1.	49
4.18	Porcentagem de reprovação na disciplina Cálculo de Probabilidade 1.	50
4.19	Porcentagem de reprovação na disciplina Geometria 2.	51
4.20	Porcentagem de reprovação na disciplina Análise 1.	52
4.21	Porcentagem de reprovações na disciplina Variável Complexa 1	53
5.1	Dados gerados com o algoritmo REPTree no software WEKA.	56
5.2	Dados gerados com o algoritmo j48graft no software WEKA.	57
5.3	Dados gerados com o algoritmo j48graft no software WEKA.	57
5.4	Dados gerados com o algoritmo CN2 Rule no Orange.	58

5.5	Dados gerados com o algoritmo CN2 Rule no Orange.	59
5.6	Dados gerados com o algoritmo CN2 Rule no Orange.	60
5.7	Dados gerados com o algoritmo CN2 Rule no Orange.	61
5.8	Dados gerados com o algoritmo CN2 Rule no Orange.	62
5.9	Dados gerados com o algoritmo CN2 Rule no Orange.	63

Lista de Tabelas

4.1	Currículo da Licenciatura em Matemática - Disciplinas Obrigatórias [10]. . .	32
4.2	Taxas de reprovação das disciplinas obrigatórias.	54

Capítulo 1

Introdução

A evasão no ensino superior é um problema mundial e consiste na saída do aluno antes da conclusão do curso. No ensino superior, ainda não existe um consenso entre os autores sobre os conceitos e as diversas formas de evasão do aluno. Santos e Giraffa [28] relatam a dificuldade para encontrar uma definição única para a evasão. A Comissão Especial de Estudos sobre Evasão nas Universidades Públicas Brasileiras definiu evasão como a saída definitiva do aluno do seu curso de origem sem que o tenha concluído. A evasão universitária tem sido estudada em todo o mundo, sendo inclusive uma das estratégias da União Europeia até 2020 a redução do abandono e o aumento da conclusão dos alunos no ensino superior.

Um dos grandes problemas que afetam a qualidade do ensino superior no nosso país é a evasão de alunos. Tal problema possui diversas origens e afeta tanto a Instituição de Ensino Superior (IES) quanto o aluno e gera graves consequências sociais e financeiras. De acordo com Lobo [22] a obtenção dos fatores que causam a evasão de alunos no curso podem colaborar para elaboração de políticas de combate à evasão. Como principais motivos para a evasão no ensino superior, Santos e Giraffa [28] citaram: falta de motivação; falta de hábito e técnicas de estudo individualizado; dificuldade de organizar o tempo disponível; conciliar estudo e trabalho; e formação escolar anterior precária.

Em particular, foram realizados vários estudos sobre evasão nos cursos de Licenciatura no Brasil podendo-se citar como exemplo os realizados por Felicetti e Fossatti [15], Souza et al [31] e Gomes et al [19]. Felicetti e Fossatti [15] realizaram um estudo em uma IES do Rio Grande do Sul para comparar a evasão nos cursos de licenciatura de alunos que faziam parte do programa de bolsas do ProUni e os que não faziam parte. A partir do estudo feito por Souza et al [31] foram obtidas mais informações acerca das evasões no Curso de Licenciatura em Informática do Instituto Federal do Rio Grande do Norte (IFRN). Gomes et al [19] realizaram um estudo sobre a evasão dos Cursos do Centro de Ciências de Saúde da Universidade Federal do Espírito Santo (UFES) e trouxe mais informações

sobre quais as áreas tinham maior e menor evasão.

Entre os estudos sobre evasão na Universidade de Brasília pode-se citar como exemplo os de Azevedo e Santos [2] e Palmeira e Santos [24]. O estudo de Azevedo e Santos [2] utilizaram a mineração de dados aplicada ao estudo da evasão e desempenho dos alunos do Bacharelado em Ciência da Computação. O estudo de Palmeira e Santos [24] sobre evasão no Bacharelado em Ciência da Computação da UnB revelou as principais variáveis relacionadas com o perfil de evasão ou de formatura de um aluno.

1.1 Justificativa

Considerando a alta taxa de evasão no Curso de Licenciatura em Matemática (Noturno), torna-se necessário um estudo que identifique os alunos com maior propensão de evadir. Nesse contexto, atualmente não se tem uma visão clara do problema de evasão no Curso de Licenciatura em Matemática (Noturno) da Universidade de Brasília (UnB).

1.2 Problema

Não há estudos detalhados sobre causas da evasão de alunos no Curso de Licenciatura em Matemática (Noturno) da UnB, em particular, se reprovações em disciplinas obrigatórias podem indicar alunos com risco de evasão.

1.3 Objetivos

O objetivo geral deste trabalho é fazer análises estatísticas sobre a evasão no Curso de Licenciatura em Matemática (Noturno) da UnB.

Os objetivos específicos são:

- Realizar análises estatísticas das menções em disciplinas obrigatórias do Curso de Licenciatura Matemática (Noturno);
- Identificar disciplinas obrigatórias que estão relacionadas à evasão dos alunos usando técnicas de mineração de dados;
- Consolidar as disciplinas obrigatórias ligadas à evasão dos alunos, a partir das análises estatísticas e da mineração de dados.

1.4 Descrição dos Capítulos

No capítulo 2, é abordado o conceito de evasão, destacando os impactos da evasão no ensino superior, também são apresentados trabalhos sobre evasão em nível internacional, nacional, nas licenciaturas no Brasil e alguns casos de estudo de evasão da UnB.

No capítulo 3, são abordados os conceitos de dado, informação, conhecimento e mineração de dados. São apresentados os softwares de mineração de dados: WEKA e Orange e algumas de suas funcionalidades.

No capítulo 4, são apresentados inicialmente informações do Curso de Licenciatura em Matemática (Noturno) e os dados dos alunos, que foram utilizados nas análises deste trabalho. Em seguida, são discutidas formas de ingresso e saída do curso e as porcentagens de reprovações nas disciplinas obrigatórias em cada semestre.

No capítulo 5, são apresentados os resultados da mineração de dados, realizados com o WEKA e o Orange.

No capítulo 6, são apresentadas as conclusões obtidas e sugestões para trabalhos futuros.

Capítulo 2

A Evasão no Ensino Superior

Este capítulo visa apresentar conceitos gerais sobre a evasão no ensino superior. A Seção 2.1 apresenta definições de evasão, que variam entre os autores. A Seção 2.2 apresenta os principais fatores que levam à evasão. A Seção 2.3 aborda a evasão em nível internacional. A Seção 2.4 aborda a evasão no contexto brasileiro. A Seção 2.5 apresenta a evasão nos cursos de licenciaturas. A Seção 2.6 apresenta alguns casos de estudo de evasão da UnB.

2.1 Definição de Evasão

A evasão no ensino superior é um problema mundial. No ensino superior, ainda não existe um consenso entre os autores sobre os conceitos e as diversas formas de evasão do aluno. Santos e Giraffa [28] relatam a dificuldade para encontrar uma definição única para a evasão.

O aumento das pesquisas no Brasil sobre evasão se deu a partir de 1995, quando o Ministério da Educação (MEC) criou a Comissão Especial de Estudos sobre Evasão nas Universidades Públicas Brasileiras [13]. Esta comissão definiu evasão como a saída definitiva do aluno do seu curso de origem sem que o tenha concluído e divide a evasão no ensino superior em três tipos: evasão de curso; evasão de instituição; e evasão do sistema de ensino.

Ristoff [27] também considera esta definição mas ainda separa evasão de mobilidade, definida como a migração para outro curso. Pereira [25] define que evasão seria o abandono da universidade sem concluir nenhum curso, excluindo assim a opção de mobilidade. Ribeiro [26] lembra que ainda existe a situação de transferência externa, quando um aluno abandona a universidade para ingressar em outra, definindo assim uma evasão de uma dada universidade, mas não o desligamento do ensino superior. Nesta pesquisa considerou-se como sendo a evasão do curso.

De acordo Ribeiro [26], a divisão da evasão real e formal foi definida no Relatório sobre Evasão desenvolvido em 2004 pelo Núcleo de Apoio aos Estudos de Graduação (NAEG) e Centro de Estudos e Pesquisas em Políticas Públicas de Educação (CEPPPE) da Faculdade de Educação da Universidade de São Paulo:

- Evasão real: ruptura dos vínculos do aluno com uma universidade e/ou curso, imediatamente após o último semestre em que concluiu algum crédito com sucesso, quando ocorreu seu último contato com a universidade;
- Evasão formal: desligamento oficial da universidade que ocorre um tempo depois da evasão real.

Esse relatório indicou que a dificuldade de escolha (seja por pressão familiar ou por falta de informação) é o principal motivo para a evasão (44,5%), sucedido pela estrutura do curso (30,7%), insatisfação com o mercado de trabalho para o curso escolhido (13,4%) e razões pessoais (11,4%).

Segundo Baggi e Lopes [3], vários conceitos de evasão não podem ser aplicados aos dados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) porque a forma como os dados de evasão são organizados é diferente em cada IES dificultando associar quando um aluno sai de um e se matricula em outro curso ou IES. Assim os dados do INEP podem tratar como evasão um aluno que saiu de um curso e se matriculou em outro. A evasão no ensino superior pode ser vista como: o abandono do curso superior; a evasão da instituição (transferência), em que o estudante muda de instituição de nível superior mas não necessariamente de curso; abandono ou exclusão pela IES com a saída definitiva ou temporária do sistema.

2.2 Fatores para Evasão do Aluno

Tigrinho [32] explica que os alunos ficam indecisos sobre qual curso ingressar no Ensino Superior e, depois da escolha, vários fatores podem influenciar na evasão do aluno:

- Repetência: reprovações em disciplinas consideradas difíceis. A forma de avaliação de algumas IES pode influenciar na desistência do curso ou de uma área de estudo;
- Baixo rendimento durante o Ensino Médio: está relacionado ao item anterior, devido à falta de conhecimentos básicos. Esse déficit, acumulado durante o curso, pode levar à evasão do aluno;
- Escolha do curso e o conhecimento sobre a carreira profissional: a falta de informações sobre a profissão pode influenciar o aluno a escolher o curso sem conhecer em

detalhes quais dificuldades serão encontradas durante o curso e a carreira profissional;

- Mudança de curso: vários motivos podem influenciar na mudança de curso, como a dificuldade em disciplinas, a forma como a IES desenvolve o curso, os requisitos e quando o aluno não se identifica com o curso;
- Dificuldades Financeiras: programas como o Fundo de Financiamento Estudantil (FIES) e o Programa Universidade para Todos (Prouni) possuem restrições e também outras despesas, além da mensalidade, que podem influenciar na permanência do aluno mesmo em universidades gratuitas, pois estas dificultam a conciliação de trabalho devido aos horários das disciplinas;
- Desmotivação: a motivação influencia na permanência do aluno no curso e na percepção da carreira a ser seguida;
- Docentes despreparados para o ensino: mesmo com muito conhecimento, o professor pode não ter preparo didático para transmitir o conteúdo de forma clara e objetiva em consonância com a futura vida profissional do aluno.

De acordo com Morosini et al [23], as causas pessoais são predominantes:

- Escolha inadequada da carreira acadêmica: falta de maturidade do estudante na escolha do curso a partir de conclusões errôneas sobre o curso e a profissão;
- Falta de orientação vocacional e expectativas irrealistas sobre a carreira: muitos estudantes baseiam-se no status social da profissão, e não procuram a devida orientação sobre o curso e com profissionais da área;
- Definição de curso de ingresso: diversos fatores podem influenciar no critério de desempate na escolha do curso, quando um candidato tem mais de uma opção e se baseia em poucas informações para fazer sua escolha de curso;
- Falta de perspectivas de trabalho: o aluno não conhece em detalhes as atividades diárias da profissão e se baseia somente em poucas informações na escolha do curso;
- Dificuldades pessoais na adaptação: muitos estudantes sentem dificuldade com a área escolhida, seja pela perspectiva erroneamente construída ou envolvimento com o curso escolhido.

Morosini et al [23] também indicam causas institucionais que influenciam:

- Localização da instituição: a distância entre a IES e a residência do aluno pode influenciar a evasão, devido a falta de transporte de qualidade, podendo ser determinante para o desempenho do aluno;

- Problemas estruturais no curso: a definição do currículo pode ser incoerente com as demandas do mercado de trabalho.

Ainda segundo Morosini et al [23], causas gerais também podem influenciar na evasão do aluno:

- Deficiências acumuladas na educação básica que levam a resultados ruins e repetidas reprovações em disciplinas;
- Dificuldades em acompanhar o curso;
- Opção por outros rumos;
- Desmotivação;
- Rebaixamento da autoestima;
- Razões econômicas (condições relacionadas ao trabalho e às condições financeiras).

2.3 Evasão em Nível Internacional

A evasão universitária tem sido estudada em todo o mundo. De acordo com a Organization for Economic Co-operation and Development (OECD) a média da taxa de evasão mundial é de aproximadamente 30%, o Japão tem uma das menores taxas de evasão do mundo, seguido pela Dinamarca, com aproximadamente o dobro da taxa de evasão do Japão. De acordo com Shimosegawa [30], por outro lado, quarenta e três mil estudantes evadiram das universidades japonesas em 2011, o que não deve ser ignorado, pois pouco se sabe sobre evasão no ensino superior japonês e aqueles que evadem acabam sendo considerados como minoria. A Figura 2.1 mostra as evasões em nível internacional em 2011.

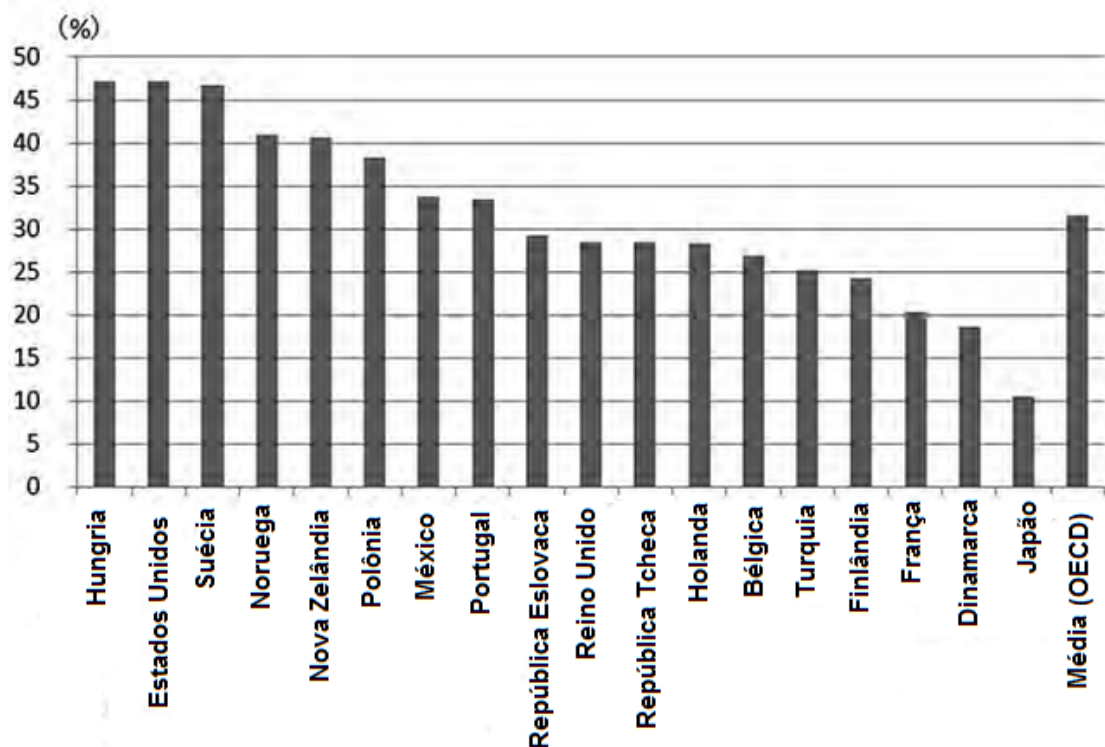


Figura 2.1: Taxa de evasão no ensino superior, países da OECD, em 2011 (adaptado de Shimosegawa [30]).

A Comissão Europeia publicou um artigo em dezembro de 2015 intitulado *Dropout and Completion on Higher Education in Europe*. Neste artigo foi feito um estudo comparativo de evasão e realização de ensino superior nos diferentes países europeus, para verificar se suas práticas governamentais estão sendo eficazes para melhorar a taxa de evasão. De acordo esse artigo as taxas das pessoas que terminaram o ensino superior na Dinamarca é

de 81%, na Noruega é de 71%, na Holanda de 76%, no Reino Unido tem 82% e na França 80%.

De acordo Vossensteyn et al [33], uma das estratégias da União Europeia até 2020 inclui a meta de ter ao menos 40% das pessoas de 30-34 anos com uma qualificação terciária, pois a educação é vital para o emprego, a justiça social e o crescimento econômico. Além disso, o aumento da escolaridade está ganhando grande importância no âmbito da União Europeia e a maior preocupação é que muitos alunos abandonam antes de obter o diploma, sendo que reduzir o abandono e aumentar a conclusão no ensino superior são estratégias privilegiadas.

Lobo [22] explica o problema de evasão baseado em sua teoria da troca: “o ser humano evita comportamentos muito custosos e procura status, relacionamentos, interações e estados emocionais compensadores. O estudante aplica a teoria da troca para determinar sua integração acadêmica e social. Se ele perceber que os benefícios são maiores que os custos, ele permanece. Caso contrário, ele se vai.”

Yancy [34], em seu estudo sobre a retenção no primeiro ano de estudantes homens negros em universidades públicas e privadas, onde predominam estudantes brancos, informa que homens negros tem as maiores taxas de evasão no ensino superior. Os dados do Departamento Americano de Educação revelam que em 2008 a porcentagem dos negros que concluíram o bacharelado é 34,3% para homens negros e 65,3% para mulheres negras, e aqueles que concluíram o mestrado é de 28,2% entre homens negros e de 71,8% entre mulheres negras. Yancy [34] relata que os homens negros estão despreparados para lidar com os pré-requisitos da universidade, apresentando problemas para conciliar a universidade e o trabalho. Racismo e dificuldades econômicas também dificultam a continuidade na universidade.

O estudo de Gladieux e Perna [18] tentaram estimar o impacto de bolsas públicas no ensino superior da Dinamarca. O sistema dinamarquês de ensino é um dos mais generosos do mundo, com bolsas de estudos que correspondem à 60% da seguridade social. Segundo o autor, uma grande preocupação da Europa Ocidental é que os alunos da pós graduação são velhos, o que impacta na oferta de trabalho e na arrecadação de tributos. Mas, por outro lado, se um limite de idade for estabelecido, a quantidade de alunos evadidos será maior. O autor conclui que a redução do horário de trabalho resultou na diminuição dos evadidos.

Em sua avaliação sobre Universidades Internacionais, Lobo [22] defende sete pontos para baixar o índice de evasão em uma instituição de ensino superior:

- Estabelecimento de um grupo de trabalho para reduzir a evasão: levantar níveis de satisfação dos alunos e estabelecer programas acadêmicos de integração e recuperação dos novos alunos. Exemplos bem sucedidos: *Gateway Program*, da Universidade

do Texas, em Austin, que recuperou alunos com problemas colocando-os no nível dos demais quanto à evasão e aprovação; e *Freshman Year College*, da Universidade da Cidade de Nova York, em Brooklyn, que passou de uma evasão de 50% para 23%, em 7 anos;

- Avaliações das estatísticas de evasão: identificar períodos críticos para a evasão e criar ações de combate a evasão. Exemplo bem sucedido: *Risk Point Intervention Program*, da Universidade do Texas, em San Antonio, reduzindo de 38% para 31% a taxa de evasão, em um ano;
- Identificar as causas da evasão: comparar prioridades dos alunos com avaliação dos serviços educacionais, administrativos e comunitários. Exemplo bem sucedido: *Student Satisfaction Inventory*, do Huntington College, em Indiana, reduzindo taxas de evasão de 50% para 25%, em sete anos;
- Concentrar a visão da IES no aluno: envolver coordenadores, professores e funcionários com o intuito de focar no sucesso e bem estar do aluno;
- Criar condições para atender aos objetivos que atraíram os alunos: não decepcionar os calouros é essencial;
- Tornar o ambiente na IES mais agradável para o aluno: campus limpo e arrumado, com boas condições de trabalho e climáticas. Processos de Avaliação Institucional demonstram o grau de insatisfação dos alunos com as condições de muitas IES;
- Criação de programas de aconselhamento e orientação dos alunos: tem que ser proativo e permanente. Exemplos bem sucedidos: *Summer Institute for Academic Achievement* e o *Tutorial Service* da Universidade de Pensilvânia, que teve uma taxa de evasão para os alunos participantes do programa de 20%, até o final do segundo ano, comparados com 69% de outros alunos com dificuldades, que não aderiram ao programa.

Gladieux e Perna [18] afirmaram que qualquer investimento é um risco, e empréstimo para pagar a faculdade também. Empréstimo para pagar a faculdade é um investimento mas, se o aluno abandona, é um investimento perdido. Em 2001, nos Estados Unidos, 350 mil ex-alunos que iniciaram os estudos seis anos antes, com empréstimos, não tinham concluído o ensino superior e tinham uma dívida a pagar.

A Figura 2.2 mostra os estudantes que constituíram empréstimo ou não para pagar o ensino superior, sendo caracterizados como: evadido; ainda cursando; completou o grau de certificado do curso superior de 2 anos; grau de associado, o aluno completou um conjunto de disciplinas de uma área; e o grau de bacharel correspondente a pelo menos 4 anos de curso.

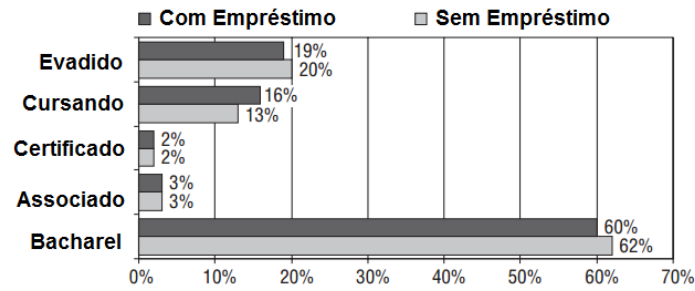


Figura 2.2: Porcentagem de estudantes com e sem empréstimo para pagar o ensino superior, por situação, Estados Unidos, 2001 (adaptado de Gladieux e Perna [18]).

A Figura 2.3 relaciona os alunos que evadiram ou completaram o bacharelado tendo constituído empréstimo ou não com: estudo em meio período; período integral de trabalho; e se iniciou tardiamente o ensino superior.

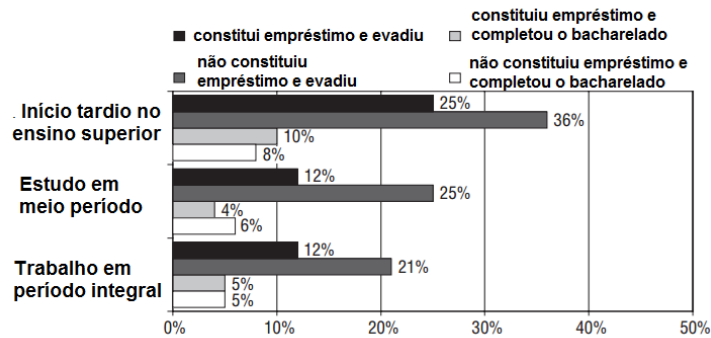


Figura 2.3: Porcentagem de estudantes por categoria de constituição de empréstimo, evasão ou conclusão de curso e início tardio de curso superior, estudo em meio período e trabalho em período integral, Estados Unidos, 2001 (adaptado de Gladieux e Perna [18]).

O questionamento de Sarker et al [29] sobre a forma de obter os dados para a análise de evasão por meio de questionários, onde somente alguns alunos colaboram, demonstra que é necessário o uso de métodos mais efetivos. O uso de técnicas de mineração de dados tem ganhado mais adeptos internacionalmente para análise de evasão universitária.

A Figura 2.4 relaciona os alunos que constituíram empréstimo e concluíram ou não o bacharelado com: valor médio dos salários; valor médio dos empréstimos; inadimplentes no empréstimo; e desempregados.

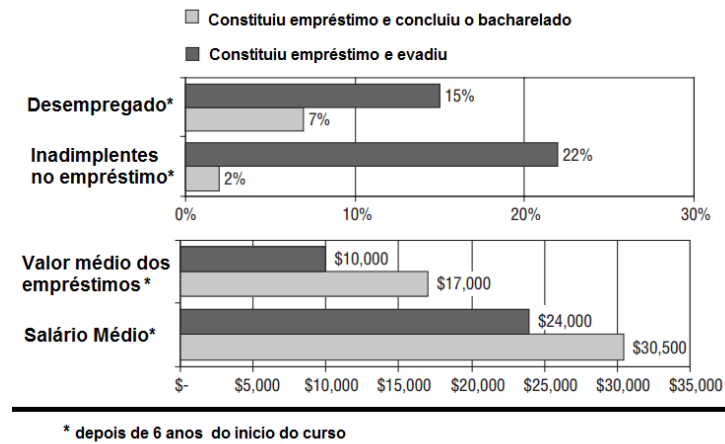


Figura 2.4: Porcentagem de estudantes que constituíram empréstimo, que completaram o bacharelado ou sofreram evasão (adaptado de Gladieux e Perna [18]).

Na Figura 2.5 o fato de evadir e constituir empréstimo ou não é relacionado com: início tardio no ensino superior; estudo em meio período; e trabalho em período integral. Segundo o autor, e de acordo com a figura, os alunos que não constituíram empréstimo são mais propensos a evadirem.

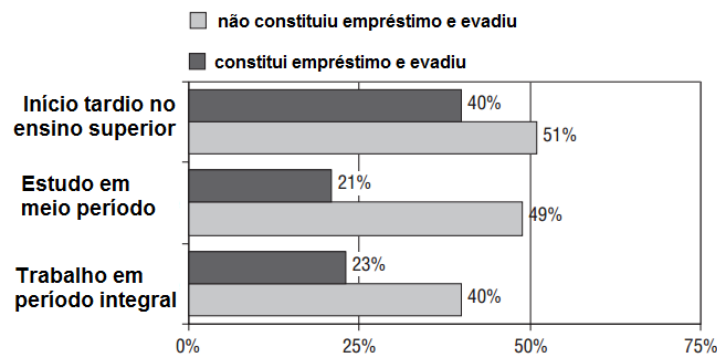


Figura 2.5: Porcentagem de estudantes por categoria de constituição de empréstimo ou não e evasão e início tardio de curso superior, estudo em meio período e trabalho em período integral, Estados Unidos, 2001 (adaptado de Gladieux e Perna [18]).

2.4 Evasão no Brasil

Um dos grandes problemas que afetam a qualidade do ensino superior no nosso país é a evasão de alunos. Tal problema possui diversas origens e afeta tanto a Instituição de Ensino Superior (IES) quanto o aluno e gera graves consequências sociais e financeiras. De acordo com Lobo [22] a obtenção dos fatores que causam a evasão de alunos no curso podem colaborar para elaboração de políticas de combate à evasão.

Cunha et al [8] fizeram um estudo descritivo sobre as razões e influências para a evasão universitária. Em particular, realizou um estudo com estudantes ingressantes nos Cursos de Ciências Contábeis de instituições públicas federais da Região Sudeste. Foi aplicado um questionário dividido em três partes: a realidade socioeconômica do estudante; fatores na escolha do curso e forma de ingresso; e a evasão do curso.

O autor relata que a média de idade está de acordo com as divulgadas pelo MEC¹ em 2011. Ainda, o autor esperava que o número de países com ensino superior fosse maior do que os 30% para os pais e 33% para as mães comparando com outros estudos.

Baseado no fato de 75% dos entrevistados possuírem renda familiar acima de 2 mil reais, o autor relata que isso está relacionado ao fato de 43% terem estudado em escola particular no ensino médio e 71% não exercerem nenhuma outra atividade.

Dos entrevistados, 91% não tiveram orientação vocacional na escolha do curso e 43% não escolheram o curso como primeira opção. Segundo o autor, a falta de orientação vocacional corrobora para não proporcionar satisfação ao aluno, prejudicando sua aprendizagem e sua vida profissional.

A lei 12711/2012 determinava que, em 2013, 12,5% das vagas deveriam reservadas a alunos de escola pública e os dados revelaram que, no curso estudado, 56% dos que responderam ao questionário estudaram no ensino público, sendo que 46% ingressaram pelo vestibular em ampla concorrência.

Cunha et al [8] concluíram que, como 65% dos entrevistados foram aprovados no primeiro processo seletivo e sem ajuda de um profissional, pode-se ter criado uma imagem irrealista do curso e da vida profissional.

Santos e Giraffa [28] fizeram uma análise qualitativa nas teses publicadas na CAPES² e nas reuniões da Associação Nacional de Pós-Graduação e Pesquisa em Educação (ANPEd). Como principais motivos para a evasão no ensino superior, os autores citaram: falta de motivação; falta de hábito e técnicas de estudo individualizado; dificuldade de organizar o tempo disponível; conciliar estudo e trabalho; e formação escolar anterior precária.

¹Ministério da Educação

²Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Na Universidade Federal do Espírito Santo (UFES) foi realizada uma pesquisa documental por Gomes et al [19], que analisaram 132 processos de desligamento, sendo 63 do Curso de Farmácia, 29 de Enfermagem, 15 do Curso de Medicina e 25 de Odontologia entre o período letivo entre 2002/1 e 2007/1. Os desligamentos foram divididos da seguinte forma:

- G1: Desligamento por abandono do curso – Desligamento automático do aluno que em dois períodos letivos consecutivos não se matriculou em disciplinas;
- G2: Desligamento voluntário – Desligamento concedido ao aluno que desistiu do vínculo com o curso na UFES;
- G3: Desligamento por não cumprimento de condição – Desligamento do aluno identificado anteriormente como provável desligado por rendimento e não cumpriu condição imposta pelos órgãos colegiados ou por tempo máximo de permanência.

Os cursos apresentaram as seguintes médias percentuais de evasão: Farmácia - 13,7%, Enfermagem - 2,0%, Odontologia - 2,0% e Medicina - 0,35%. Os gráficos abaixo representam os dados analisados no estudo de acordo com o curso e os tipos de desligamento:

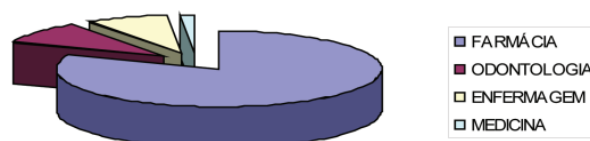


Figura 2.6: G1 - Desligamento por abandono do curso. Fonte: Gomes et al [19].

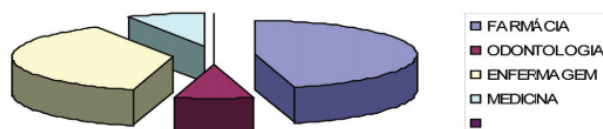


Figura 2.7: G2 – Desligamento voluntário. Fonte: Gomes et al [19].

Gomes et al [19] concluíram que o índice total de evasão dos Cursos do Centro de Ciências de Saúde da UFES é baixo. O Curso de Farmácia foi o que apresentou a maior média percentual, seguido do Curso de Enfermagem e Odontologia, que apresentaram a mesma média e, por último, Medicina.

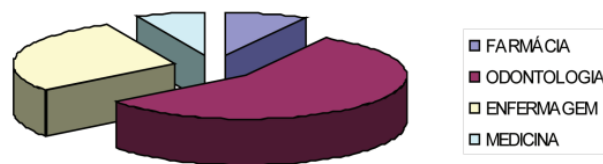


Figura 2.8: G3 – Desligamento por não cumprimento de condição. Fonte: Gomes et al [19].

2.5 Evasão nas Licenciaturas

Segundo os dados do INEP [12], no período de 2012-2013, as matrículas em licenciaturas cresceram apenas 0,6% enquanto nos Cursos de bacharelado e tecnólogos cresceram 4,4% e 5,4% respectivamente, conforme a Figura 2.9.

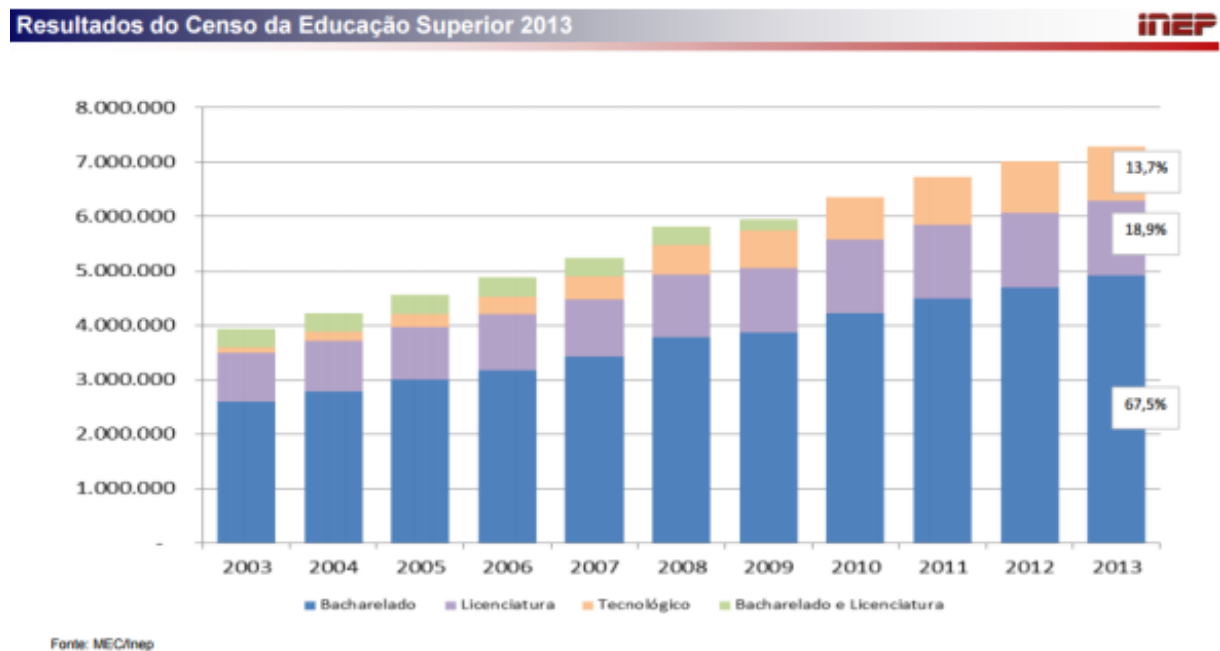


Figura 2.9: Evolução das matrículas de Educação Superior de Graduação, por Grau Acadêmico - Brasil 2003 - 2013. Fonte: INEP [12].

O estudo exploratório de Felicetti e Fossatti [15] realizado em uma IES do Rio Grande do Sul compara a evasão nos Cursos de licenciatura de alunos que fazem parte do programa de bolsa do ProUni³ e os que não são do ProUni, considerando a situação dos alunos no segundo semestre de 2012 que ingressaram em 2007, 2008 e 2009. Do total de 1.253

³Programa Universidade para Todos

estudantes, 60 eram bolsistas do ProUni. Em relação ao gênero e faixa etária, alunos ProUni e não ProUni foram semelhantes. Os cursos que tiveram menor evasão foram de Matemática e Computação e os de maior evasão, os de Letras e História, para os alunos que não são do ProUni. Para os alunos do ProUni, os maiores foram Matemática e Física e os menores, Química e Educação Física. Felicetti e Fossatti [15] conclui que o ProUni contribui para fidelizar o aluno nas licenciaturas.

De acordo com Souza et al [31], o Curso de Licenciatura em Informática do Instituto Federal do Rio Grande do Norte(IFRN) teve um total de 242 matrículas do primeiro semestre letivo de 2010 até o segundo semestre letivo de 2013. A Figura 2.10 mostra os números relacionados a evasão do curso.

Situações	Nº
Evasões (Ev)	74
Matrículas canceladas (MC)	57
Jubilados (J)	11
Transferidos entre turmas (TT)	10
Transferidos entre cursos (TC)	4
TOTAL	156

Figura 2.10: Quantidade de evadidos no Curso de Licenciatura em Informática na IFRN. Fonte: Souza et al [31].

O número total de evasões considerado foi de 135, tendo sido excluídos os jubilados e transferidos entre turmas, o que representou uma taxa de evasão de 55,78%. Quando ingressaram no curso, os alunos responderam a um questionário: 42% responderam que faziam o curso para se qualificar para o mercado de trabalho; 30% responderam que queriam aprimorar os conhecimentos em informática; 6% responderam que já davam aulas e queriam um diploma de licenciado; 18% responderam que queriam dar continuidade aos estudos; e 4% escolheram como resposta a opção ‘outros’.

Segundo Souza et al [31], uma formação básica deficitária é um empecilho para um desempenho satisfatório, pois os cursos na área de Computação exigem pré-requisitos da Matemática Básica. Outro caso importante é de alguns alunos que, mesmo com esforços para continuarem matriculados no curso, tem de interromper seus sonhos e buscar alternativas, por motivos financeiros, acadêmicos, de transporte e de segurança.

2.6 Evasão na Universidade de Brasília

O estudo de Azevedo e Santos [2] sobre mineração de dados aplicada ao estudo da evasão e desempenho dos alunos do Bacharelado em Ciência da Computação da UnB revelou que as disciplinas de Cálculo 1, Cálculo 3, Física 1, Física 3 e Programação Sistemática podem ser utilizadas como parâmetros para identificar um aluno com risco de evasão.

O estudo de Palmeira e Santos [24] sobre evasão no Bacharelado em Ciência da Computação da UnB revelou que a análise estatística dos dados foi fundamental para identificação das principais variáveis relacionadas com o perfil de evasão ou de formatura de um aluno. Foi verificado que a taxa de reprovação nas disciplinas obrigatórias do curso está de certa forma relacionada com a decisão por evasão do aluno, especialmente quando a reprovação ocorre por trancamento ou por falta de rendimento.

O estudo de Cardoso [7] na UnB sobre a evasão de alunos cotistas e não cotistas observou que 14% já eram alunos da universidade e fizeram o vestibular novamente pelo sistema de cotas devido a mudança de perspectiva profissional. De acordo com a análise das tabelas de evasão média nos cursos de 12%, chegando a mais de 80%. Os menores índices ocorrem na Medicina e Odontologia e os maiores foram: Artes plásticas (70%); Engenharia Mecânica (71%); Filosofia (75%); Bacharelado em Matemática (76%); Bacharelado em Física (76%) e Licenciatura em Química (86%). A justificativa para índices altos é a evasão aparente, ou seja, a mobilidade de um curso para outro dentro da mesma IES, enquanto a evasão real corresponde ao abandono definitivo ou mudança para outra IES.

A Figura 2.11 referente ao ano de 2004 na UnB revela que a evasão aparente por nível de rendimento no curso é evidente: entre os alunos com boas notas, a taxa de evasão é 1,9% enquanto o índice de evasão dos que tiveram notas baixas é 21,1%. Entre os alunos cotistas, o índice de evasão dos que tiveram alto rendimento foi 0,7% e, entre os de baixo rendimento foi de 16,7%, e os índices entre os não cotistas de alto rendimento foi 2,1% e os de baixo, 22,4% em 2004, já em 2005, entre os alunos cotistas, o índice de evasão dos que tiveram alto rendimento foi 0,0% e baixo rendimento foi de 11,2%; e os índices entre os não cotistas de alto rendimento foi 0,7% e os de baixo, 20,9%. A autora relata que já esperava que o rendimento dos alunos de alto rendimento fosse menor que a taxa dos que tiveram baixo rendimento porque, na UnB, quando o aluno não cumpre os requisitos fica em condição (o aluno deverá ser aprovado em um número mínimo de disciplinas) e se não for aprovado em todas as disciplinas da condição será desligado da universidade.

Alunos que trabalham tem maiores chances de evasão, pois o tempo de estudo diminui e influencia no rendimento acadêmico. Quando efetua a matrícula na UnB, o aluno responde um questionário informando se trabalha. Afirmaram que não 5% em 2004 e 3,7% em 2005, enquanto, 40,2% em 2004 e 26,9% em 2005 afirmaram que trabalhavam.

Alunos	Rendimento no curso	Evasão aparente	Mobilidade na UnB	Evasão total	Não-evadidos	Total
Não-cotista	Alto rendimento	2,1	4,0	6,1	93,9	100,0
	Baixo rendimento	22,4	7,6	30,0	70,0	100,0
	Total de alunos	11,8	5,7	17,5	82,5	100,0
Cotista	Alto rendimento	0,7	2,0	2,7	97,3	100,0
	Baixo rendimento	16,7	7,5	24,2	75,8	100,0
	Total de alunos	10,4	5,3	15,7	84,3	100,0
UnB	Alto rendimento	1,9	3,7	5,6	94,4	100,0
	Baixo rendimento	21,1	7,6	28,7	71,3	100,0
	Total de alunos	11,5	5,7	17,2	82,8	100,0

Fonte: Cespe/UnB e SAA/UnB, elaboração da autora.

Figura 2.11: Alunos evadidos e não evadidos por rendimento no curso em 2004 na UnB.
Fonte: Cardoso [7].

Alunos	Rendimento no curso	Evasão aparente	Mobilidade na UnB	Evasão total	Não-evadidos	Total
Não-cotista	Alto rendimento	0,7	3,3	4,0	96,0	100,0
	Baixo rendimento	20,9	7,5	28,4	71,6	100,0
	Total alunos	10,7	5,3	16,0	84,0	100,0
Cotista	Alto rendimento	0,0	3,0	3,0	97,0	100,0
	Baixo rendimento	11,3	4,7	16,0	84,0	100,0
	Total alunos	5,8	3,9	9,7	90,3	100,0
UnB	Alto rendimento	0,5	3,2	3,7	96,3	100,0
	Baixo rendimento	18,8	6,9	25,7	74,3	100,0
	Total alunos	9,7	5,0	14,7	85,3	100,0

Fonte: Cespe/UnB e SAA/UnB, elaboração da autora.

Figura 2.12: Alunos evadidos e não evadidos por rendimento no curso em 2005 na UnB.
Fonte: Cardoso [7].

Área	Grupo de Curso	Evasão aparente	
		Coorte 2004	Coorte 2005
Humanidades	Maior prestígio	8,6	6,9
	Menor prestígio	15,4	13,4
	Subtotal	7,7	6,7
Ciências	Maior prestígio	16,8	15,1
	Menor prestígio	2,4	2,4
	Subtotal	7,5	3,8
Saúde	Maior prestígio	11,5	9,7
	Menor prestígio	8,6	6,9
	Subtotal	15,4	13,4
UnB		11,9	7,7

Fonte: Cespe/UnB e SAA/UnB, elaboração da autora.

Figura 2.13: Taxa de evasão aparente, total de alunos e grupos de curso por prestígio social (%), na UnB em 2004 e 2005. Fonte: Cardoso [7].

Capítulo 3

Mineração de Dados

Este capítulo apresenta os conceitos gerais sobre mineração de dados. A Seção 3.1 apresenta as definições de dado, informação e conhecimento. A Seção 3.2 apresenta várias definições de mineração de dados. A Seção 3.3 apresenta os principais processos de extração de conhecimento pela mineração de dados. A Seção 3.4 define tarefas de mineração de dados. A Seção 3.5 mostra o software WEKA e algumas de suas funcionalidades. A Seção 3.6 mostra o software Orange e algumas de suas funcionalidades.

3.1 Dado, Informação e Conhecimento

De acordo com Angeloni [1], os conceitos de dado, informação e conhecimento são elementos fundamentais para tomadas de decisão e comunicação em organizações. Dados são elementos brutos desvinculados da realidade, sem um significado, são matéria prima para as informações. Informações são dados com significados dotados de relevância e propósito. Para Davenport e Prusak [9], conhecimento é a informação mais valiosa que foi processada pelos indivíduos. Seu valor agregado depende dos conhecimentos anteriores desses indivíduos, estando estritamente relacionado à sua percepção que codifica, distorce e decodifica, usando as informações de acordo com suas características pessoais.

3.2 Definição de Mineração de Dados

O armazenamento de dados tem tido uma grande importância desde a geração dos sistemas computacionais e, com sua capacidade crescendo diariamente através da evolução tecnológica, a dificuldade em processar as informações também aumentou. A essa situação Han e Kamber [20] definiram como “rico em dados, pobre em informação”. Neste sentido, a mineração de dados surge como uma tecnologia promissora que, de acordo com Berry e Linoff [4], mineração de dados é a exploração e análise de forma automática ou

semi-automática, de grandes quantidades de dados, com o intuito de descobrir padrões e regras significativos.

Existem diversas definições de Mineração de Dados sob diferentes perspectivas (estatística, banco de dados e aprendizado de máquina), respectivamente:

Segundo Hand et al [21], Mineração de Dados é a análise de grandes conjuntos de dados a fim de encontrar relacionamentos inesperados e de resumi-los os dados de uma forma que eles sejam tanto úteis quanto compreensíveis a quem está trabalhando com os dados;

De acordo com Cabena et al [6], Mineração de Dados é um campo interdisciplinar que junta técnicas de máquinas de conhecimentos, reconhecimento de padrões, estatísticas, banco de dados e visualização, para conseguir extrair informações de grandes bases de dados;

Para Fayyad et al [14], Mineração de Dados é um passo no processo de Descoberta de Conhecimento que consiste na realização de análises pela aplicação de algoritmos de descoberta que, sob certas limitações computacionais, produzem um conjunto de padrões.

3.3 Processos de Extração de Conhecimento de Dados

De acordo com Fayyad et al [14], KDD (*Knowledge Discovery in Database*) é um processo não trivial de identificação de novos padrões válidos, úteis e compreensíveis. Refere-se a todo processo de descoberta de conhecimento dividido nas seguintes fases: seleção; pre-processamento; transformação; mineração e interpretação. Mineração de dados é uma das atividades do KDD e consiste na aplicação da análise de dados e descoberta de algoritmos que produz uma enumeração de padrões a partir do dado.

Brachman e Anand [5] deram uma visão prática do KDD enfatizando alguns processos, como:

- Seleção de dados: visa a familiarização com o banco de dados, depois de definido o domínio sobre o qual será executado o processo de descoberta, sendo o próximo passo selecionar e coletar o conjunto de dados ou variáveis necessárias.
- Limpeza dos dados: tem como objetivo o tratamento de ruídos, dados estranhos ou inconsistentes.
- Redução dos dados: encontrar características úteis para representar o dado dependendo do objetivo da atividade. Usando a redução de dimensionalidade ou métodos de transformação para reduzir o número efetivo de variáveis ou encontrar representações invariáveis para o dado.

- Escolha do algoritmo de mineração: selecionar os métodos que serão utilizados para procurar os padrões do dado. Isto inclui decidir quais são os modelos e parâmetros que podem ser apropriados.
- Mineração dos dados: procurar pelos padrões de interesse numa forma de representação particular como classificação de árvores, regressão, agrupamento, entre outras.
- Interpretação dos modelos da interpretação: possibilita retornar aos passos anteriores para futura iteração. Este passo envolve a visualização dos modelos/padrões.
- Consolidação do conhecimento adquirido: incorporação do conhecimento dentro de algum sistema para uma nova ação ou simplesmente documentá-la e reportar às partes interessadas.

3.4 Tarefas de Mineração

De acordo com Galvao e Marin [16], na mineração de dados devem ser definidos os algoritmos e tarefas que serão utilizados de acordo com o objetivo para que se possa obter uma resposta para o problema. Podem ser agrupados em atividades descritivas e preditivas.

De acordo com Fayyad et al [14], para predição, o sistema encontra padrões com o propósito de realizar predições de um comportamento futuro de alguns grupos. A seguir, os tipos de tarefas:

- Classificação: predição de uma variável categórica, mediante a construção de uma função que mapeia ou classifica um dado dentro de várias outras classes. Constrói-se um modelo de algum tipo que possa ser aplicado a dados não classificados a fim de categorizá-los em classes. Algoritmos que mais se destacam: Redes Neurais, Back Propagation, Classificadores Bayesianos e Algoritmos Genéticos;
- Estimativa (regressão): Busca de funções lineares ou não, na qual a variável predita consiste em um atributo numérico (contínuo) presente no banco de dados com valores reais. Algoritmos: Métodos de estatística e Redes Neurais;
- Associação: Identificação e descrição de associações entre variáveis no mesmo item ou entre itens diferentes que tendem a co-ocorrerem (serem adquiridos juntos) em uma mesma transação. Algoritmos: Apriori e GSP;
- Clusterização: processo de partição de uma população heterogênea em vários sub grupos ou grupos mais homogêneos. Diferentemente da classificação, que tem variáveis predefinidas, na clusterização, automaticamente são identificados os grupos de

dados aos quais o pesquisador deverá atribuir as variáveis. Algoritmos: K-means, k-modes, k-prototypes entre outros;

- Sumarização: métodos para identificar e indicar características comuns em um conjunto de dados.

3.5 Software WEKA

O *Waikato Environment for Knowledge Analysis* (WEKA) é um software livre que reúne um conjunto de algoritmos de aprendizagem de máquina e ferramentas de pré-processamento de dados. De acordo com Cabena et al [6], os algoritmos de predição dividem os dados em N-1 partes aleatoriamente onde, por exemplo, 9 são usadas como treinamento e uma como teste. Essas subpartes são chamadas de pastas (*folds*). A Figura 3.1 mostra a tela inicial do WEKA:



Figura 3.1: Tela inicial do software WEKA.

O menu *Explorer* apresenta as opções de pré-processamento, classificador, clusterização, associação, selecionar atributos e visualizar. A seguir segue a definição de cada uma:

- Pré-processamento: adequação dos dados para os algoritmos. A Figura 3.2 mostra que é possível escolher os atributos pertinentes como por exemplo excluir sexo, nascimento, UF (Unidade da Federação) e tipo de escola para diminuir o escopo para a mineração de dados.

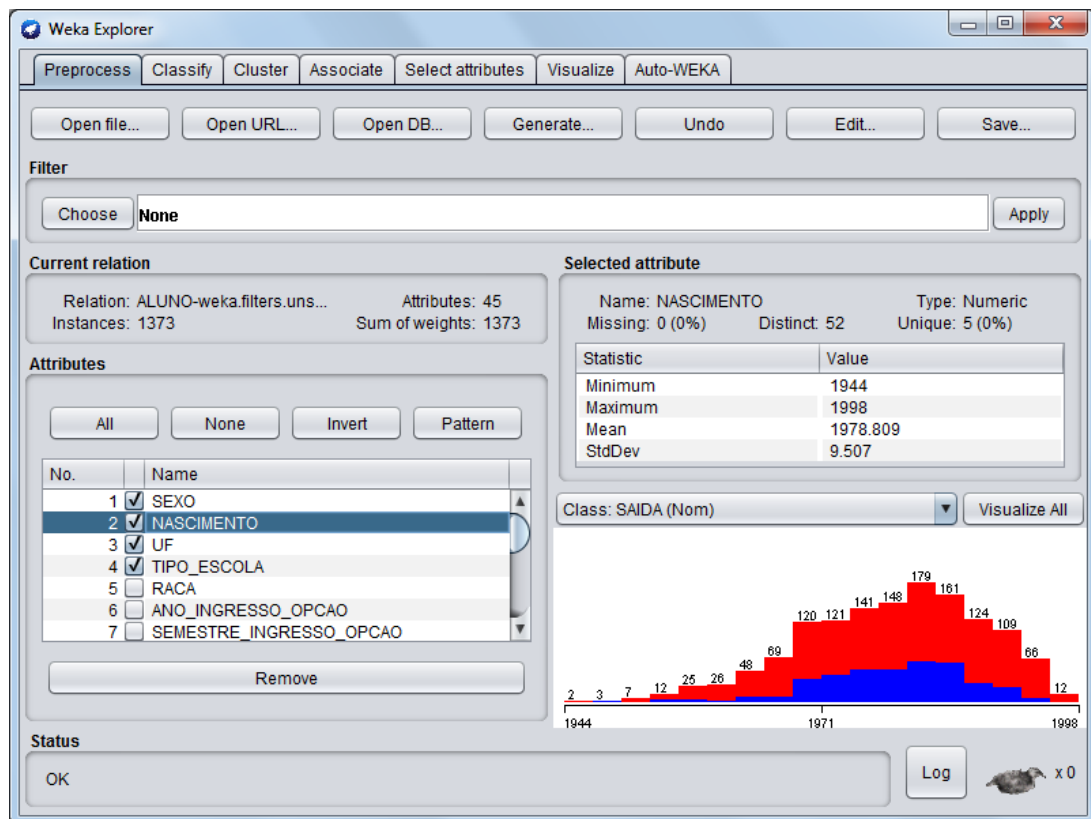


Figura 3.2: Tela pré-processamento do software WEKA.

- Classificador: algoritmo para classificar dados dentro de outras classes. A Figura 3.3 mostra a tela dos algoritmos de classificação na qual são escolhidos o algoritmo, a quantidade de pastas de treinamento para o algoritmo de classificação treinar sobre os dados e, posteriormente, gerar as classes;
- Clusterização: algoritmo para agrupar subgrupos homogêneos. A Figura 3.4 mostra a tela dos algoritmos de classificação na qual são escolhidos o algoritmo, a quantidade de pastas de treinamento para o algoritmo de classificação treinar sobre os dados e, posteriormente, gerar as regras de mineração de dados;

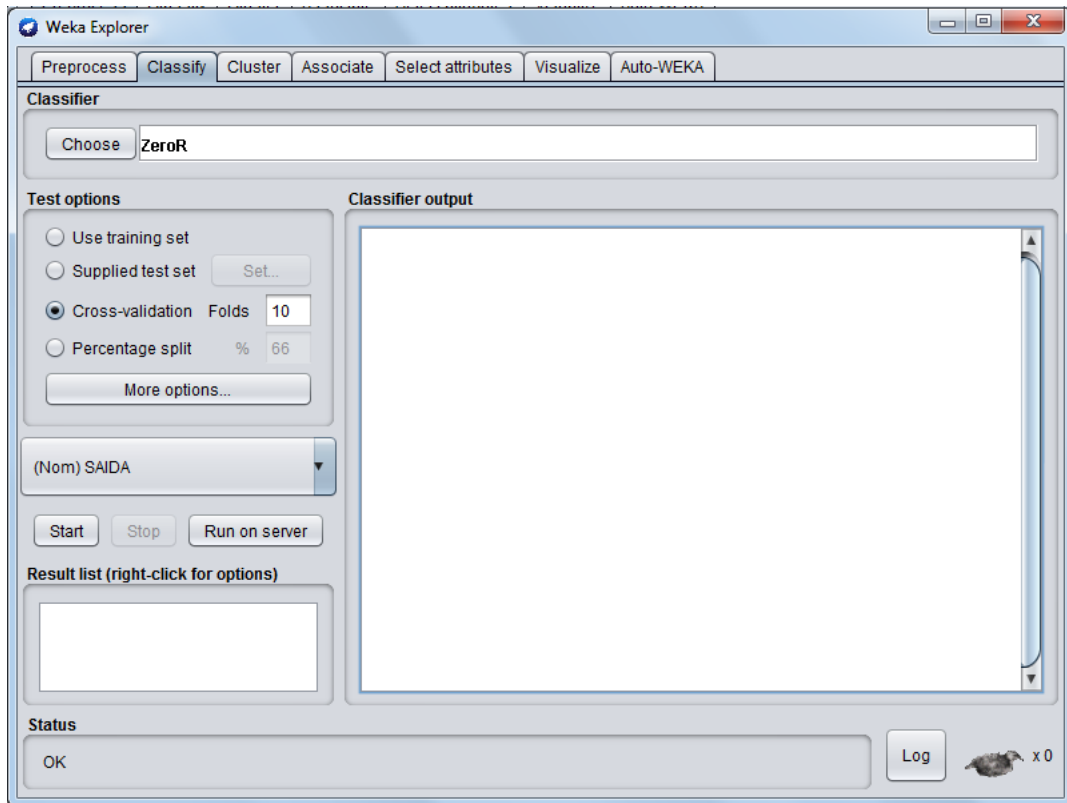


Figura 3.3: Tela classificação do software WEKA.

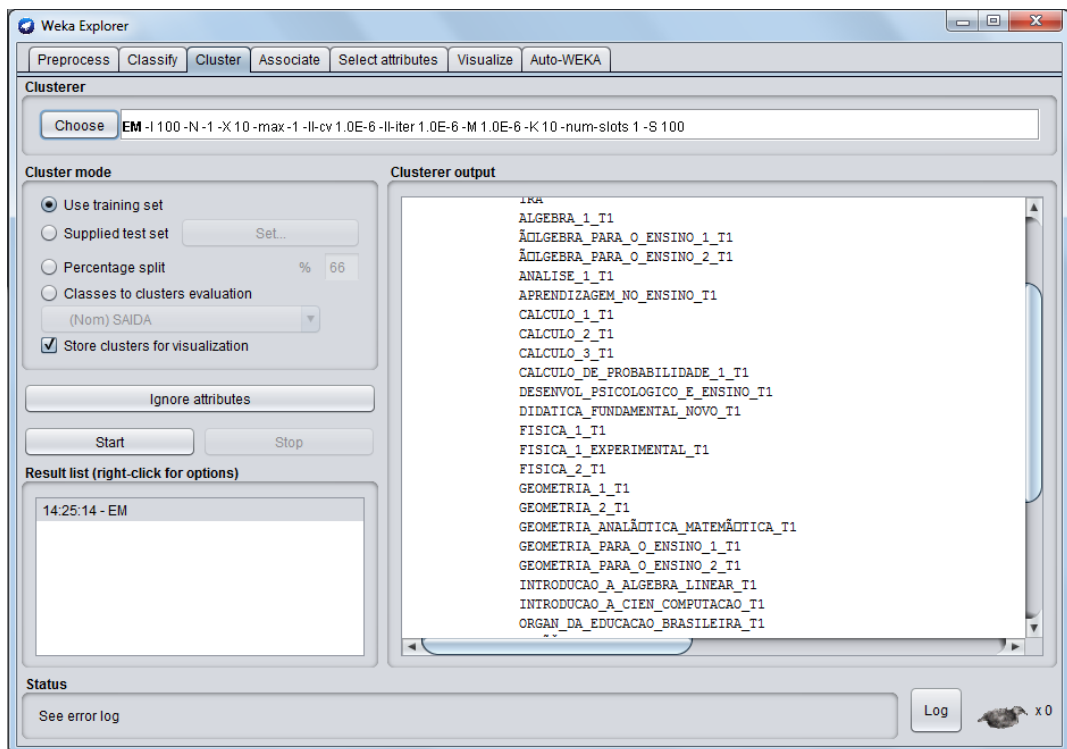


Figura 3.4: Tela clusterização do software WEKA.

- Associação: utilização do conjunto de dados para aprender e avaliar regras de associação. A Figura 3.4 mostra a tela de associação;
- Selecionar atributos: escolha dos atributos mais apropriados. A Figura 3.6 mostra a tela de seleção de atributos mais apropriados para serem usados na mineração de dados;
- Visualizar : demonstração em um gráfico cartesiano da distribuição de dados selecionados. A Figura 3.7 mostra se os dados são homogêneos;

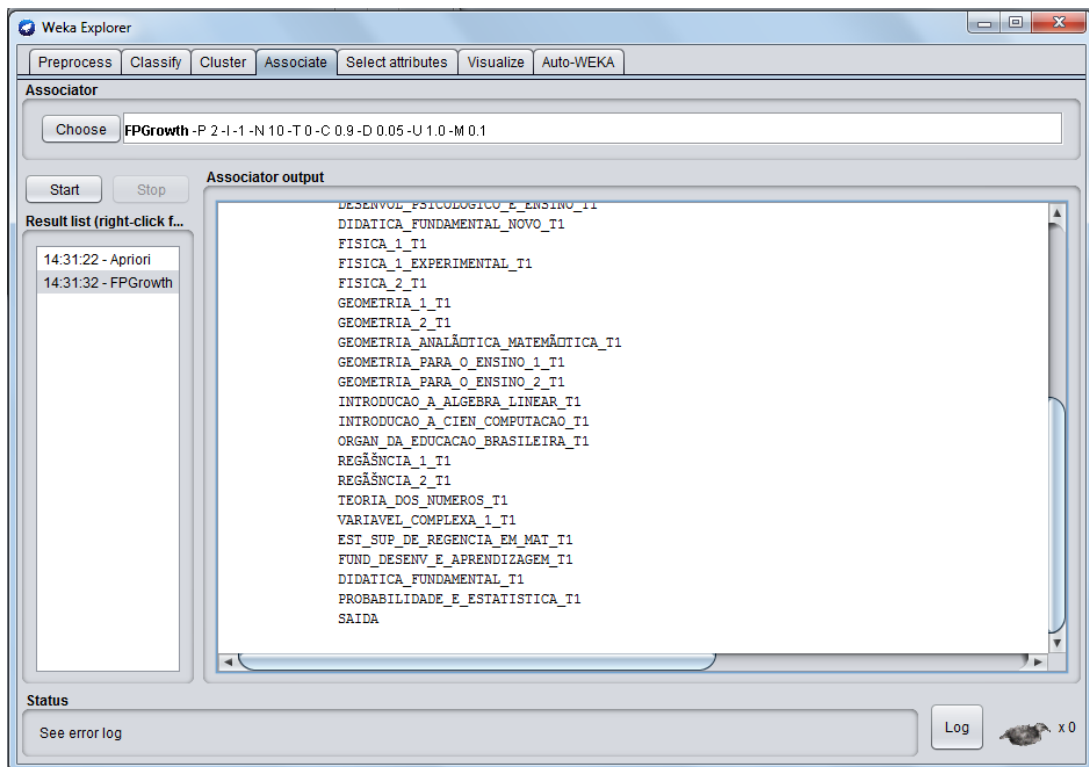


Figura 3.5: Tela associação do software WEKA.

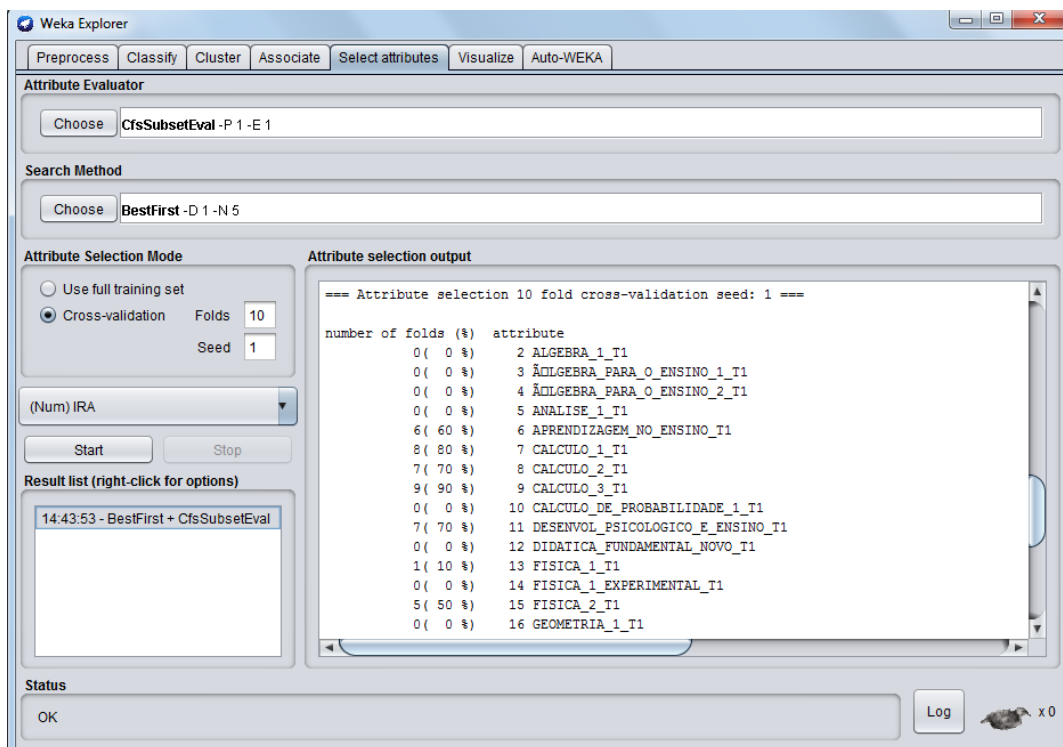


Figura 3.6: Tela de seleção de atributos do software WEKA.



Figura 3.7: Tela de visualização de atributos do software WEKA.

3.6 Software Orange

O Orange é um software livre que reúne um conjunto de algoritmos de aprendizagem de classificação, de regressão e não supervisionados. A Figura 3.8 mostra a tela inicial do Orange com os itens que representam as seguintes funcionalidades:

- *Data* - opções para abrir um arquivo ou uma tabela e opções de seleção de dados: aleatória; escolher colunas e escolher linhas, criar uma classificação e opções de pre-processamento e utilização de programas na linguagem de programação Python para filtrar os dados e a opção para salvar os dados.
- *Visualize* - opções para visualizar os resultados: plotagem por caixas; distribuição; plotagem espalhada; diagrama de peneira; mosaico; mapa de pilha; diagrama de Venn; mapa de silhueta e visualizador para o algoritmo CN2 rule.
- *Classify* - algoritmos de classificação: maioria; CN2 Rule; vizinhos próximos; árvore; floresta aleatória; SVM; redes bayesianas; Adaboost e classificação por carga.
- *Regression* - algoritmos de regressão: aprendizagem de significados; vizinhos próximos regressivo; árvore regressivo; floresta aleatória regressivo; SVM regressivo; Adaboost regressivo e gradiente estocástico.
- *Evaluate* - algoritmos de avaliação; teste; predição; matriz de confusão; curva de elevador e calibração.
- *Unsupervised* - algoritmos não supervisionados: distância no arquivo; distância na matriz; distância no mapa; agrupamento de hierarquia; K-means; PCA; análise de correspondência; distância; distância transformada e MDS.

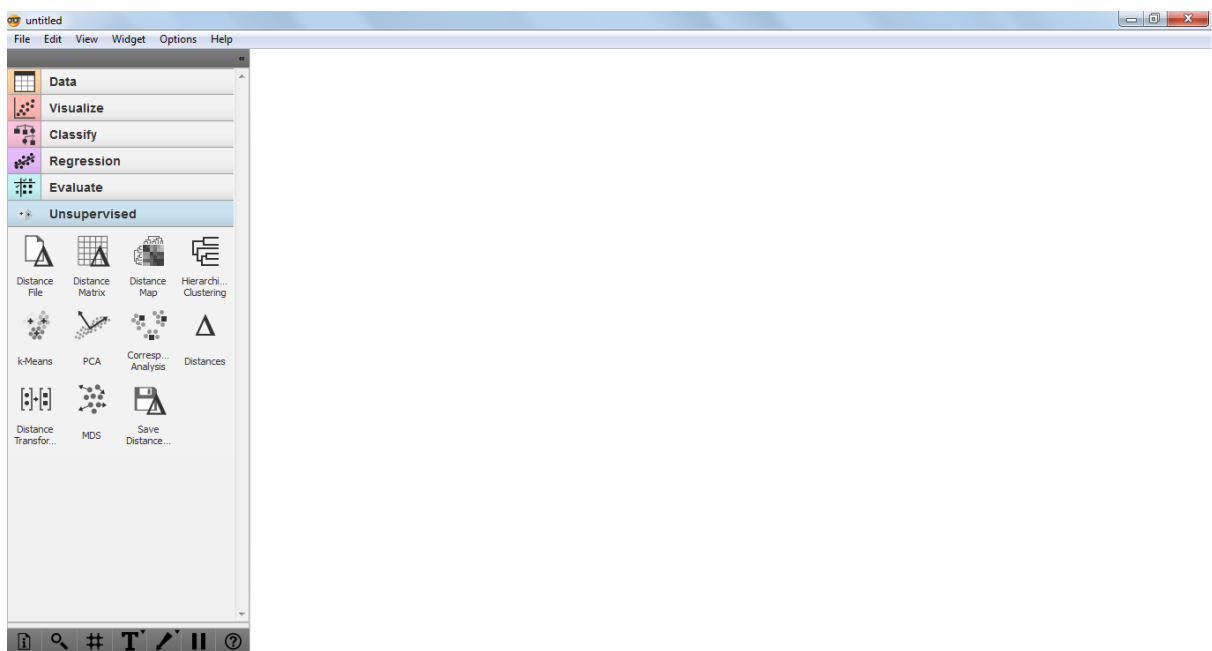


Figura 3.8: Tela inicial do software Orange.

Capítulo 4

A Licenciatura em Matemática (Noturno) e Análises Estatísticas

O objetivo deste capítulo é produzir análises sobre como as menções de reprovações nas disciplinas obrigatórias influenciaram na evasão dos alunos.

A Seção 4.1 apresenta informações do curso de Licenciatura em Matemática (Noturno) da UnB. A Seção 4.2 apresenta os dados dos alunos que foram utilizados na mineração de dados. A Seção 4.3 apresenta as análises estatísticas sobre as formas de ingresso e as reprovações de disciplinas obrigatórias divididas por semestre.

4.1 Licenciatura em Matemática (Noturno)

Em seguida, apresentamos informações gerais do Curso de Licenciatura em Matemática (Noturno), importantes para a realização deste trabalho:

- Habilitação - Licenciatura em Matemática;
- Turno - Noturno; reconhecido pelo Ministério da Educação (MEC) - sim;
- Currículo vigente em - 2014/1;
- Mínimo de créditos por período - 16;
- Máximo de créditos por período - 26;
- Limite mínimo de permanência - 7;
- Limite máximo de permanência - 14;
- Créditos exigidos para formatura - 188.

4.1.1 Fluxo de Disciplinas e Disciplinas Obrigatórias

A Figura 4.1 representa o fluxo de disciplinas do Curso de Licenciatura em Matemática (Noturno) da Universidade de Brasília.

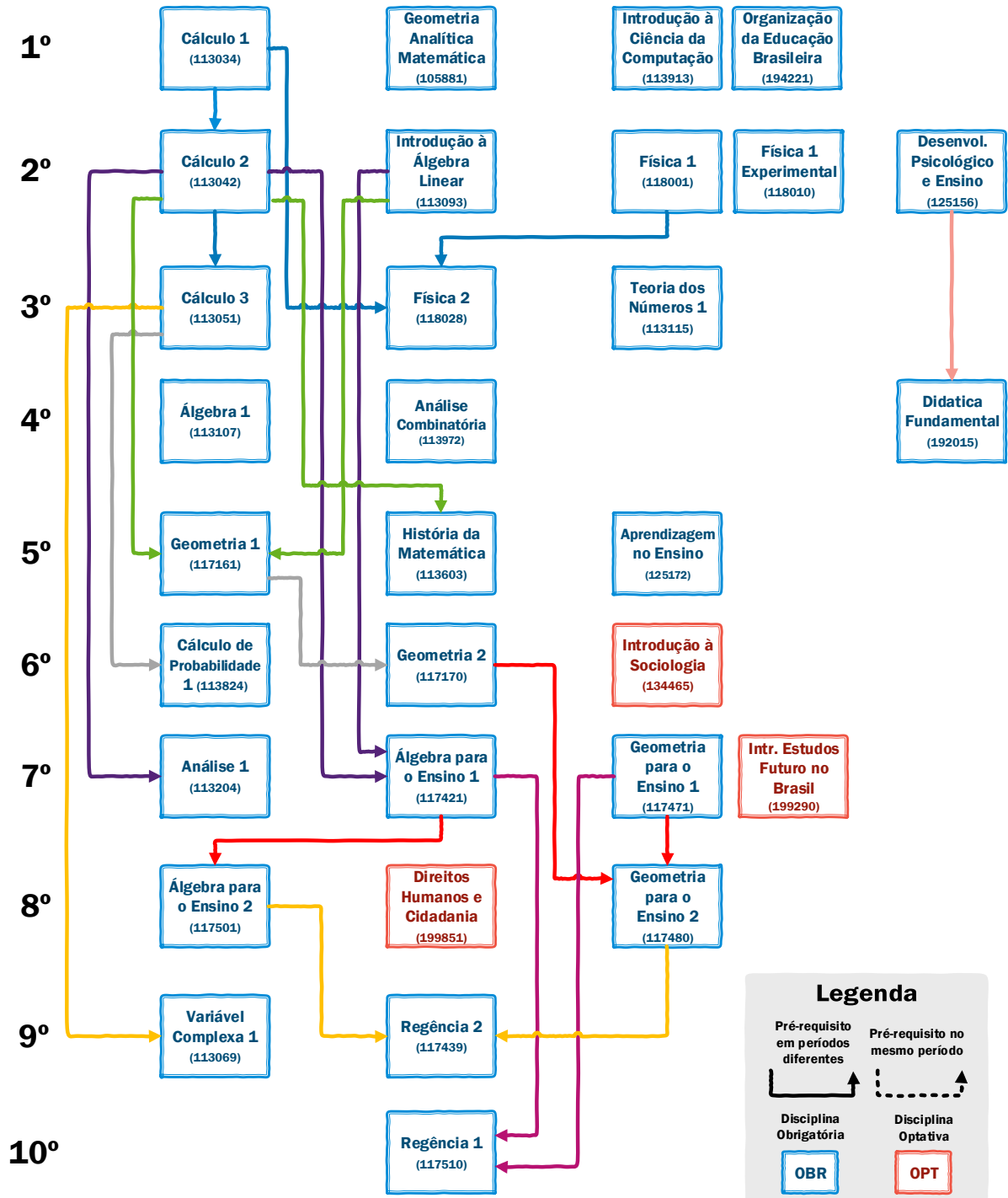


Figura 4.1: Fluxo de disciplinas do Curso de Licenciatura em Matemática (Noturno) da Universidade de Brasília.

Tabela 4.1: Currículo da Licenciatura em Matemática - Disciplinas Obrigatórias [10].

Código - Disciplina	Quantidade de Créditos
113107 - ÁLGEBRA 1	4
117421 - ÁLGEBRA PARA O ENSINO 1	6
117501 - ÁLGEBRA PARA O ENSINO 2	6
113204 - ANÁLISE 1	4
125172 - APRENDIZAGEM NO ENSINO	4
113034 - CÁLCULO 1	6
113042 - CÁLCULO 2	6
113051 - CÁLCULO 3	6
113824 - CÁLCULO DE PROBABILIDADE 1	6
125156 - DESENVOLVIMENTO PSICOLÓGICO E ENSINO	4
194531 - DIDÁTICA FUNDAMENTAL	4
118001 - FÍSICA 1	4
118010 - FÍSICA 1 EXPERIMENTAL	2
118028 - FÍSICA 2	4
117161 - GEOMETRIA 1	4
117170 - GEOMETRIA 2	4
105881 - GEOMETRIA ANALÍTICA MATEMÁTICA	4
117471 - GEOMETRIA PARA O ENSINO 1	6
117480 - GEOMETRIA PARA O ENSINO 2	6
113093 - INTRODUÇÃO A ÁLGEBRA LINEAR	4
113913 - INTRODUÇÃO A CIÊNCIA DA COMPUTAÇÃO	4
194221 - ORGANIZAÇÃO DA EDUCAÇÃO BRASILEIRA	4
117510 - REGÊNCIA 1	8
117439 - REGÊNCIA 2	8
113115 - TEORIA DOS NÚMEROS	4
113069 - VARIÁVEL COMPLEXA 1	6
Total de créditos obrigatórios	128
Total de disciplinas obrigatórias	26

O aluno deverá ser aprovado nas disciplinas obrigatórias descritas na Tabela 4.1, além de precisar cursar disciplinas optativas e de módulo livre, para integralizar 188 créditos, tal que possa concluir o Curso de Licenciatura em Matemática (Noturno).

4.2 Dados dos Alunos

Os dados dos alunos foram cedidos pelo Decanato de Planejamento e Orçamento (DPO) da Universidade de Brasília, com base na história de cada aluno que ingressou no Curso de Licenciatura em Matemática (Noturno), desde 1993 até 2016. Os dados utilizados dos alunos foram as menções das disciplinas obrigatórias e o Índice de Rendimento Acadêmico

(IRA). Conforme o Decanato de Ensino de Graduação (DEG) [11] as menções na UnB são mensuradas da seguinte forma:

- SR: de 0 até menor que 1 ou não cumpriu o mínimo de 75% de presença durante o semestre;
- II: de 1 até menor que 3;
- MI de 3 até menor que 5;
- MM: de 5 até menor que 7;
- MS: de 7 até menor que 9;
- SS: de 9 até 10.

De acordo com o DEG [11] o Índice de Rendimento Acadêmico (IRA) é calculado considerando os seguintes pesos nas menções:

- SR: peso 0;
- II: peso 1;
- MI: peso 2;
- MM: peso 3;
- MS: peso 4;
- SS: peso 5.

O processo de filtragem foi feito através do *script em Python* criado por Garcia [17] que foi realizado no Curso de Licenciatura em Matemática (diurno) da UnB. Os atributos a seguir foram removidos: código da UF, unidade, campus, grau do curso, forma do curso, área de conhecimento, código do departamento, código do instituto, nome do curso, semestre final do curso, ano final do curso, semestre inicial do curso, horas em extensão, data da resolução, duração plena, código do grau, tempo máximo de permanência no curso, créditos para formatura, semestre final opção, ano final da opção, semestre inicial da opção, ano inicial da opção, nome da opção, aluno registrado. Assim, ficaram os seguintes dados: sexo, nascimento, Unidade da Federação (UF), cotista, Portador de Necessidades Especiais (PNE), tipo de escola, raça, local de nascimento, ano de ingresso, semestre de ingresso, forma de ingresso, ano de saída, semestre de saída, forma de saída, período curricular, número de créditos da disciplina, Índice de Rendimento Acadêmico (IRA), código do curso, código da disciplina, menção, nome da disciplina.

4.3 Análises Estatísticas

4.3.1 Formas de Ingresso e saída do curso

A Figura 4.2 mostra a quantidade de alunos por forma de ingresso no período de 1993 a 2016:

- 70,2% ingressaram pelo vestibular;
- 15,3% ingressaram pelo Programa de Avaliação Seriada (PAS);
- 7,1% ingressaram por transferência;
- 3,1% ingressaram pelo Sistema de Seleção Unificada (SISU);
- 2,4% ingressaram pelo Exame Nacional do Ensino Médio (ENEM).

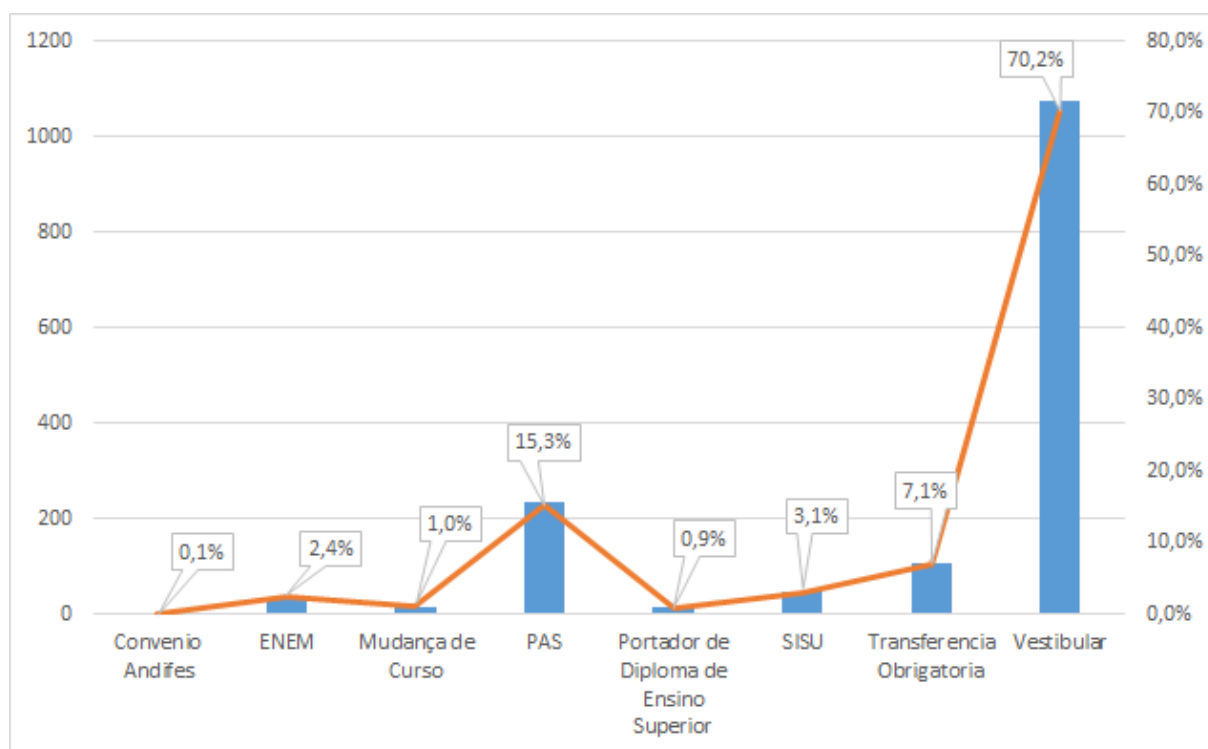


Figura 4.2: Quantidade de alunos por forma de ingresso.

A Figura 4.3 mostra a porcentagem pela forma de saída no período de 1993 a 2016:

- 28,7% foram desligados por não cumprir condição;
- 24,4% formaram;
- 16% ainda não concluíram;
- 11,3% foram desligados por abandono;
- 2,2% foram desligados por reprovaram três vezes a mesma disciplina obrigatória.

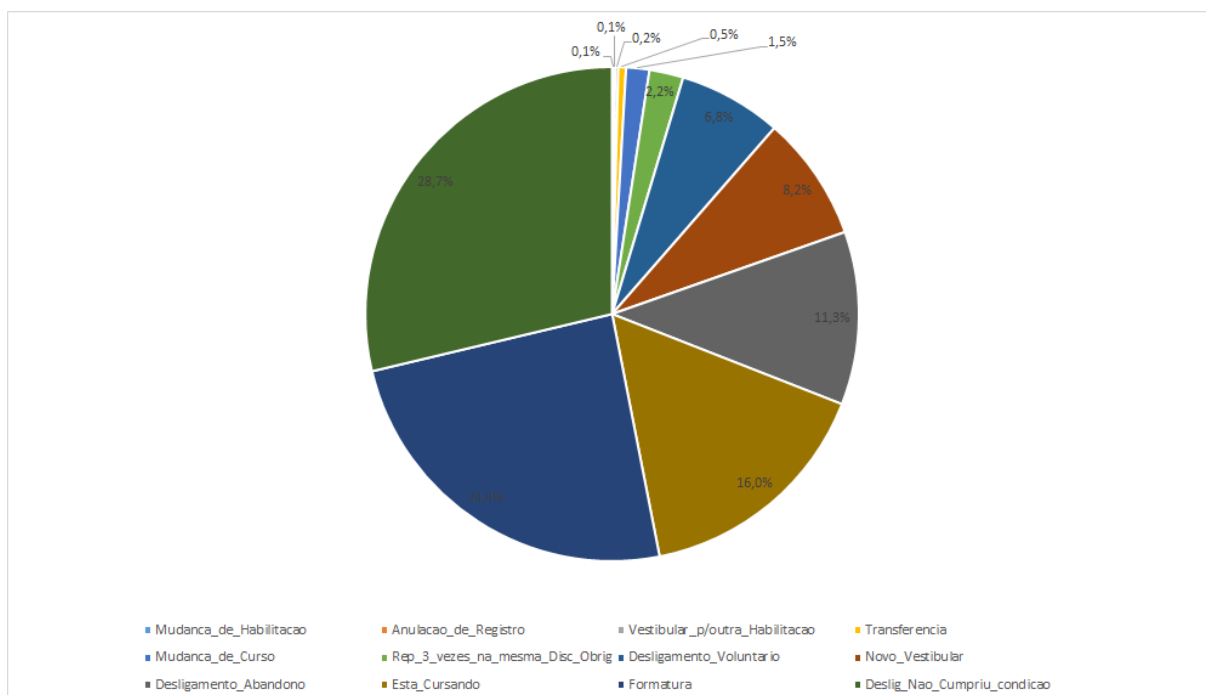


Figura 4.3: Porcentagem das diferentes formas de evasão.

4.3.2 Disciplinas por Semestre

Primeiro Semestre

As disciplinas obrigatórias do primeiro semestre são: Organização da Educação Brasileira; Cálculo 1 e Introdução a Ciência da Computação.

A Figura 4.4 refere-se à disciplina de Organização da Educação Brasileira.

As maiores taxas de reprovação foram:

- 1º semestre de 1996 com 44,4%;
- 1º semestre de 1998 com 40%;
- 2º semestre de 2015 com 46,9%.

As menores taxas de reprovação foram:

- 2º semestre de 2004 com 4%;
- 2º semestre de 2005 com 4,2%;
- 2º semestre de 2014 com 8,5%.

O percentual médio de reprovação de todos os semestres foi de 21,6%.

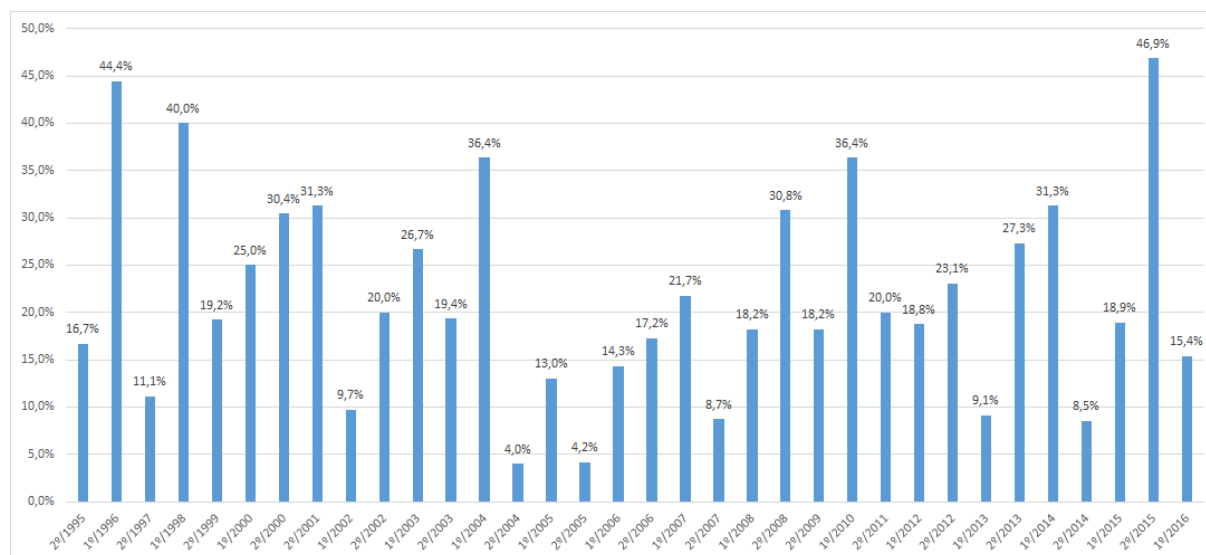


Figura 4.4: Porcentagem de reprovação na disciplina Organização da Educação Brasileira.

A Figura 4.5 refere-se à disciplina de Cálculo 1.

As maiores taxas de reprovação foram:

- 1º semestre de 1998 com 76%;
- 1º semestre de 2011 com 84,8%;
- 1º semestre de 2016 com 85,3%.

As menores taxas de reprovação foram:

- 2º semestre de 2002 com 10,8%;
- 2º semestre de 2004 com 15,4%;
- 2º semestre de 2005 com 35,5%.

O percentual médio de reprovação de todos os semestres foi de 53,9%.

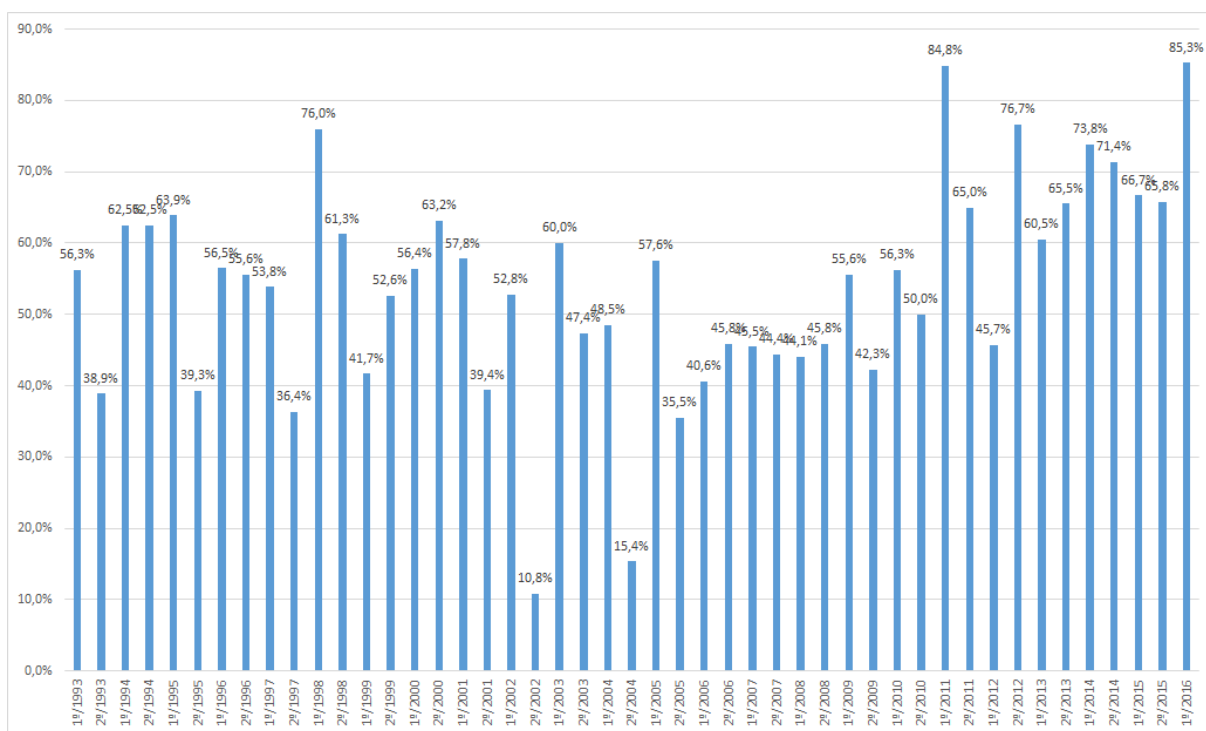


Figura 4.5: Porcentagem de reprovação na disciplina Cálculo 1.

A Figura 4.6 refere-se à disciplina de Introdução a Ciência da Computação.

As maiores taxas de reprovação foram:

- 1º semestre de 1994 com 76,9%;
- 1º semestre de 2008 com 75,0%;
- 1º semestre de 2011 com 70,6%.

As menores taxas de reprovação foram:

- 1º semestre de 2004 com 15,2%;
- 1º semestre de 2005 com 15,6%;
- 1º semestre de 2006 com 13,3%.

O percentual médio de reprovação de todos os semestres foi de 41,8%.

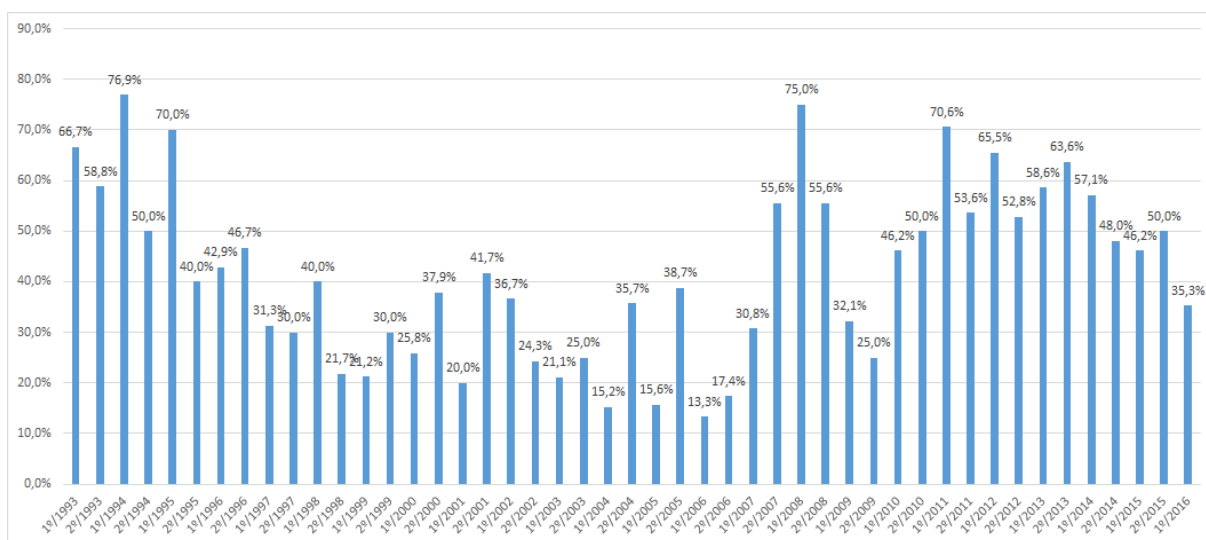


Figura 4.6: Porcentagem de reprovação na disciplina Introdução a Ciência da Computação.

Segundo Semestre

As disciplinas obrigatórias do segundo semestre são: Introdução a Álgebra Linear; Cálculo 2; Física 1; Física 1 Experimental e Desenvolvimento Psicológico e Ensino.

A Figura 4.7 refere-se à disciplina de Introdução a Álgebra Linear.

As maiores taxas de reprovação foram:

- 1º semestre de 1997 com 73,7%;
- 2º semestre de 2013 com 86,4%;
- 1º semestre de 2016 com 75%.

As menores taxas de reprovação foram:

- 1º semestre de 1997 com 20%;
- 1º semestre de 1998 com 14,3%;
- 2º semestre de 2003 com 25%.

O percentual médio de reprovação de todos os semestres foi de 52%.

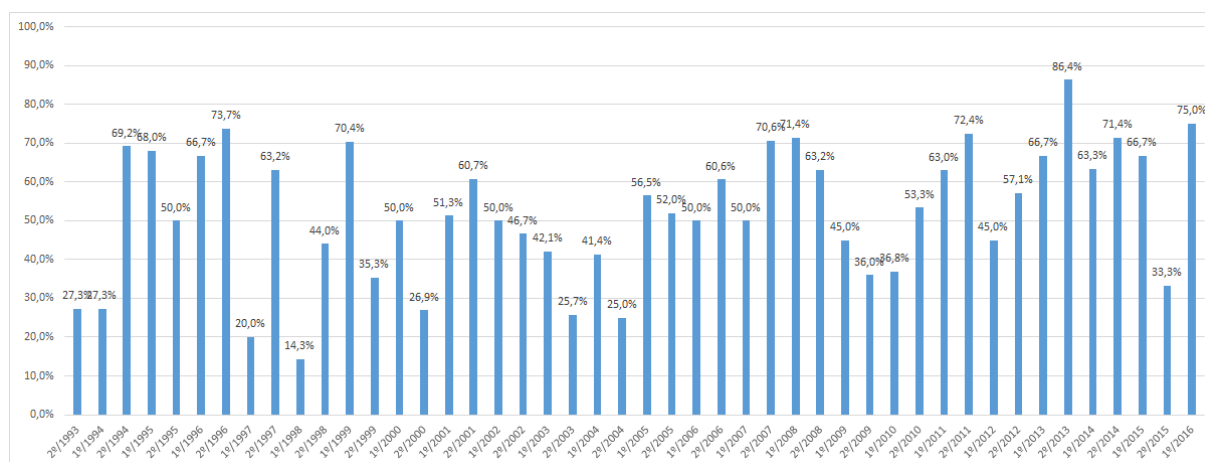


Figura 4.7: Porcentagem de reprovação na disciplina a Introdução a Álgebra Linear.

A Figura 4.8 refere-se à disciplina de Cálculo 2.

As maiores taxas de reprovação foram:

- 2º semestre de 1994 com 75,0%;
- 1º semestre de 1997 com 78,9%;
- 2º semestre de 2010 com 82,4%.

As menores taxas de reprovação foram:

- 2º semestre de 1995 com 16,7%;
- 1º semestre de 2010 com 16,7%;
- 1º semestre de 2010 com 18,8%.

O percentual médio de reprovação de todos os semestres foi de 42,4%.

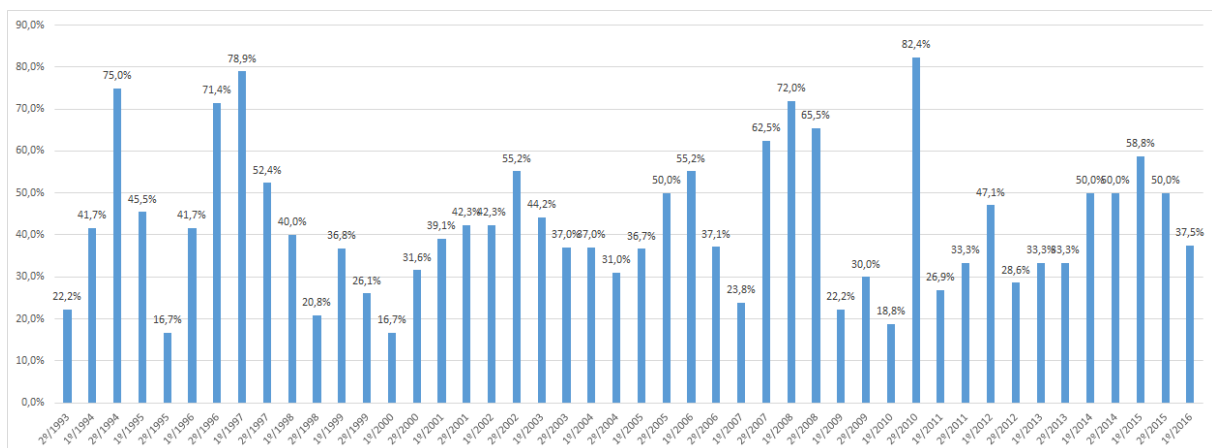


Figura 4.8: Porcentagem de reprovação na disciplina Cálculo 2.

A Figura 4.9 refere-se à disciplina de Física 1.

As maiores taxas de reprovação foram:

- 1º semestre de 2007 com 83,3%;
- 1º semestre de 2015 com 86,4%;
- 1º semestre de 2016 com 81%.

As menores taxas de reprovação foram:

- 2º semestre de 2006 com 25%;
- 1º semestre de 2011 com 22,2%;
- 2º semestre de 2011 com 28,6%.

O percentual médio de reprovação de todos os semestres foi de 56,7%.

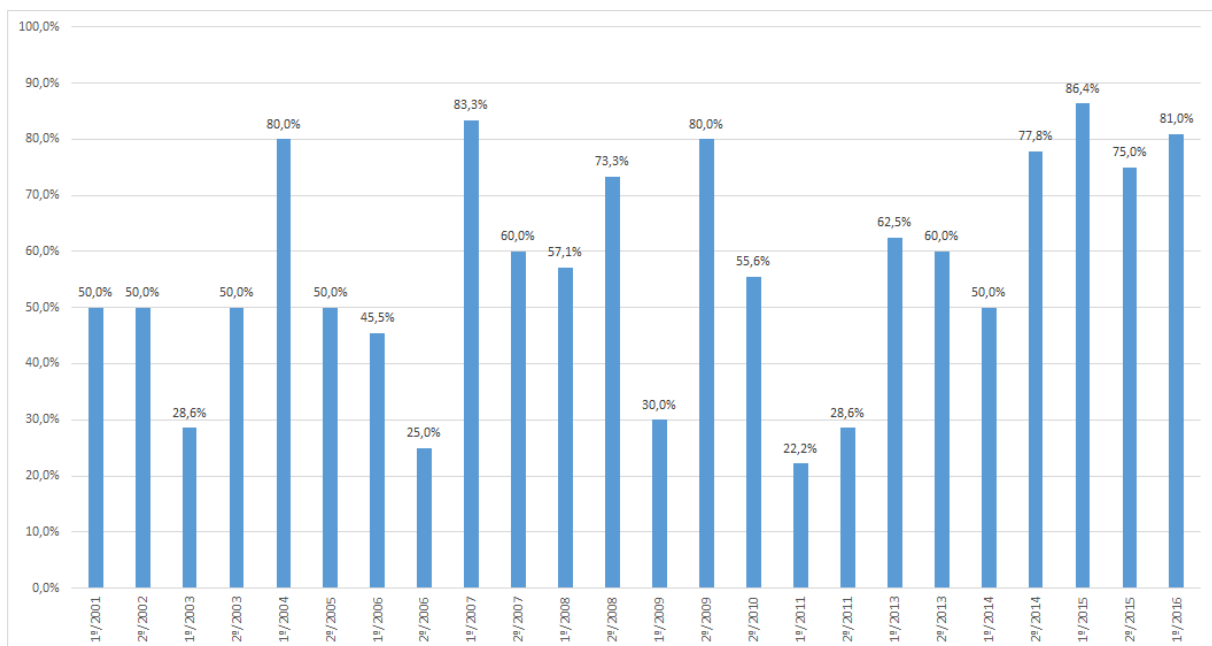


Figura 4.9: Porcentagem de reprovação na disciplina Física 1.

A Figura 4.10 refere-se à disciplina de Física 1 Experimental.

As maiores taxas de reprovação foram:

- 1º semestre de 2011 com 75%;
- 2º semestre de 2013 com 75%;
- 1º semestre de 2015 com 76,5%.

As menores taxas de reprovação foram:

- 1º semestre de 2007 com 16,7%;
- 1º semestre de 2013 com 20%;
- 1º semestre de 2014 com 14,3%.

O percentual médio de reprovação de todos os semestres foi de 45,4%.

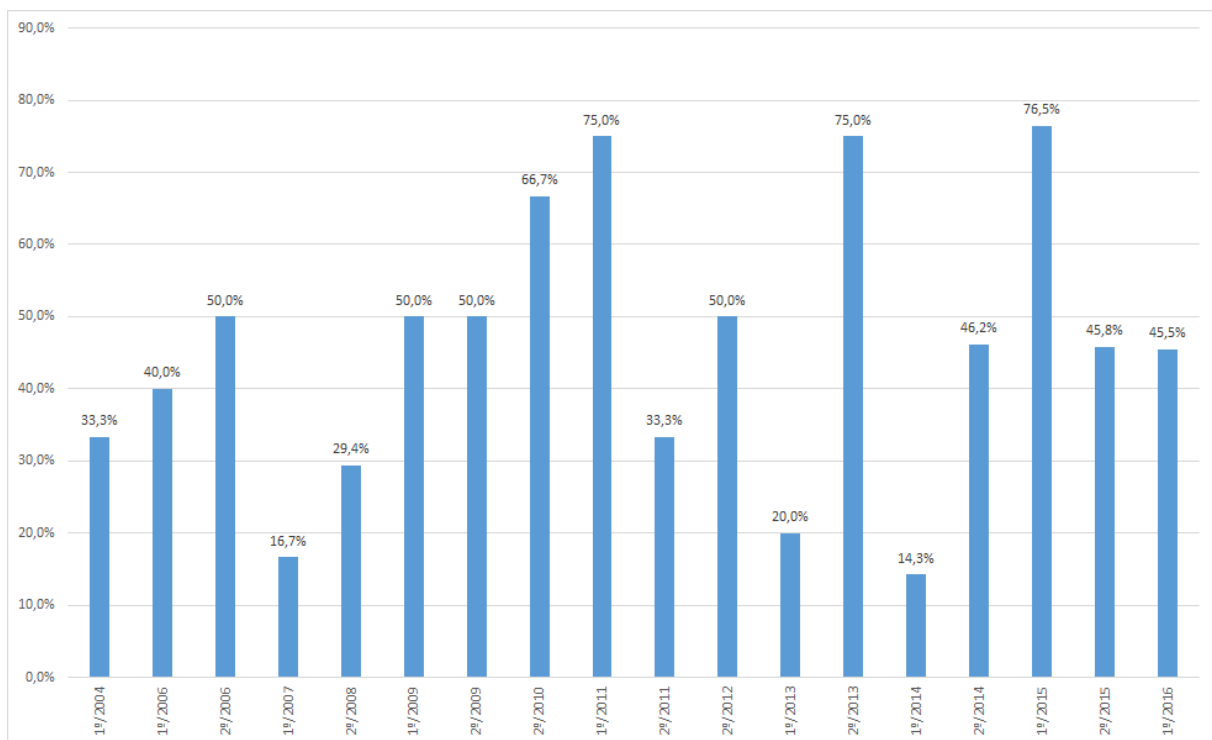


Figura 4.10: Porcentagem de reprovação na disciplina Física 1 Experimental.

A Figura 4.11 refere-se à disciplina de Desenvolvimento Psicológico e Ensino.

As maiores taxas de reprovação foram:

- 2º semestre de 1994 com 52,2%;
- 2º semestre de 2010 com 50%;
- 2º semestre de 2012 com 40,7%.

As menores taxas de reprovação foram:

- 1º semestre de 2003 com 9,1%;
- 1º semestre de 2007 com 10%;
- 2º semestre de 2007 com 8,7%.

O percentual médio de reprovação de todos os semestres foi de 24,4%.

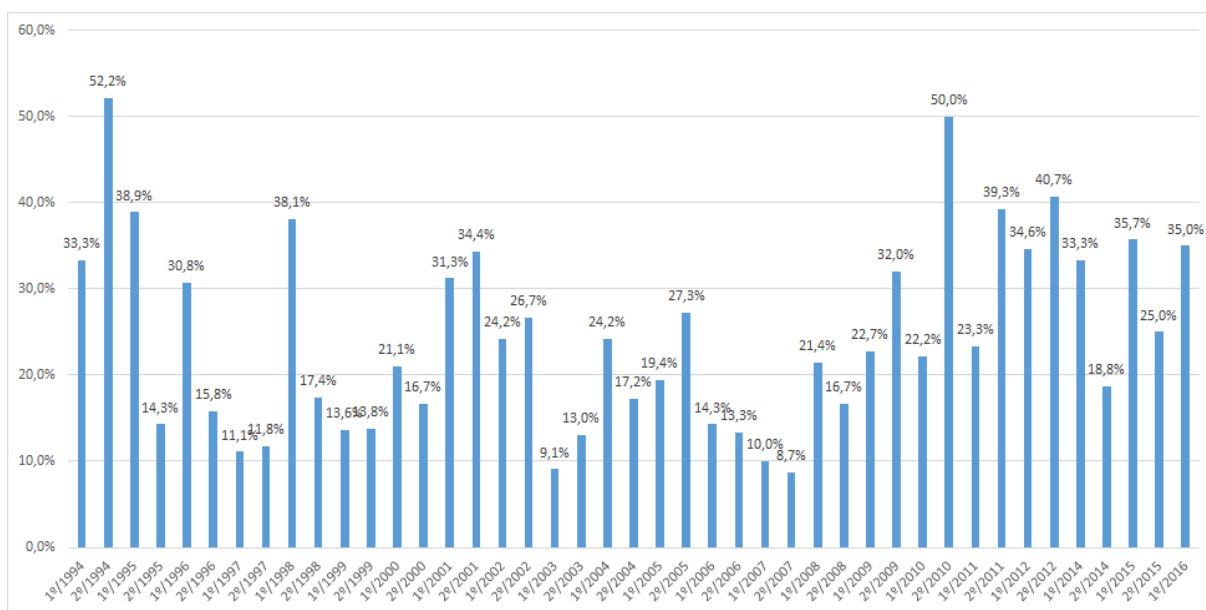


Figura 4.11: Porcentagem de reprovação na disciplina Desenvolvimento Psicológico e Ensino.

Terceiro Semestre

As disciplinas obrigatórias do terceiro semestre são: Física 2; Cálculo 3 e Teoria dos Números.

A Figura 4.12 refere-se à disciplina de Física 2.

As maiores taxas de reprovação foram:

- 2º semestre de 2009 com 75,0%;
- 1º semestre de 2014 com 72,7%;
- 2º semestre de 2015 com 83,3%.

As menores taxas de reprovação foram:

- 2º semestre de 2005 com 11,1%;
- 1º semestre de 2011 com 16,7%;
- 2º semestre de 2012 com 25,0%.

O percentual médio de reprovação de todos os semestres foi de 46%.

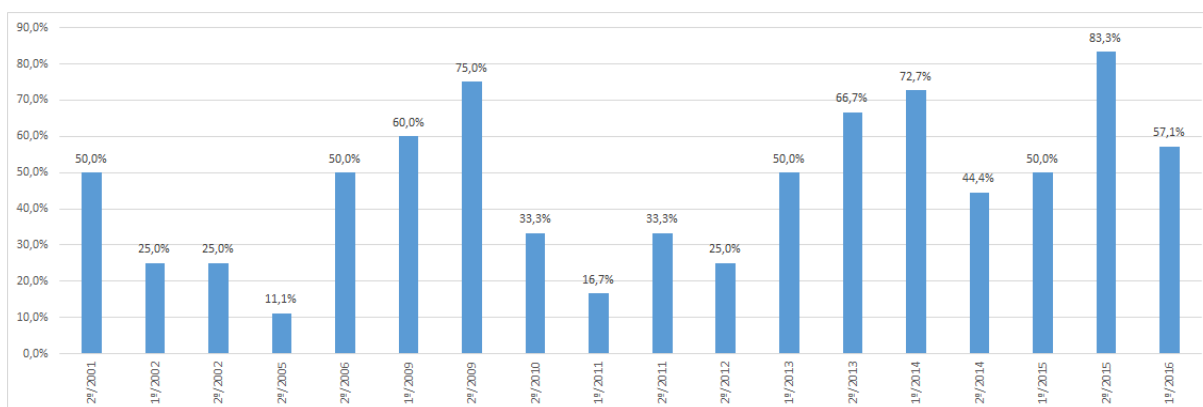


Figura 4.12: Porcentagem de reprovação na disciplina Física 2.

A Figura 4.13 refere-se à disciplina de Cálculo 3.

As maiores taxas de reprovação foram:

- 1º semestre de 1995 com 80%;
- 2º semestre de 1996 com 75%;
- 2º semestre de 2000 com 81,5%.

As menores taxas de reprovação foram:

- 1º semestre de 1999 com 7,7%;
- 2º semestre de 2007 com 13%;
- 1º semestre de 2011 com 9,1%.

O percentual médio de reprovação de todos os semestres foi de 43,9%.

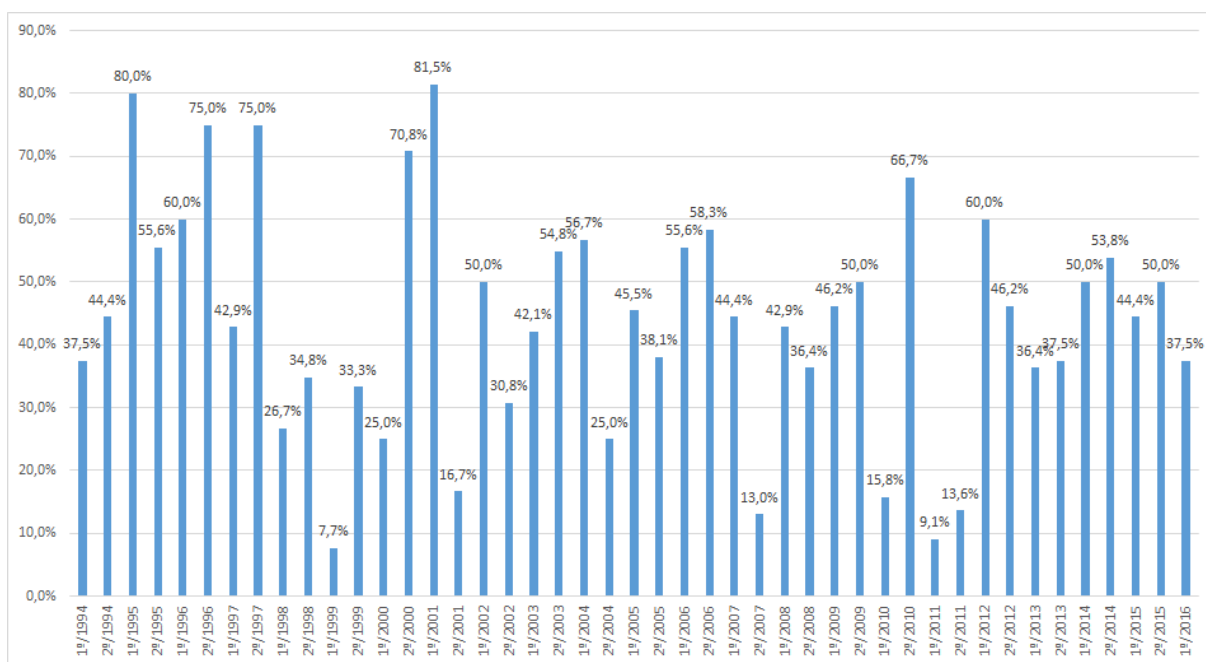


Figura 4.13: Porcentagem de reprovação na disciplina Cálculo 3.

A Figura 4.14 refere-se à disciplina de Teoria dos Números.

As maiores taxas de reprovação foram:

- 2º semestre de 2006 com 64%;
- 1º semestre de 2014 com 71,4%;
- 2º semestre de 2014 com 66,7%.

As menores taxas de reprovação foram:

- 2º semestre de 2000 com 10,7%;
- 1º semestre de 2002 com 10,7%;
- 1º semestre de 2013 com 7,7%.

O percentual médio de reprovação de todos os semestres foi de 36,9%.

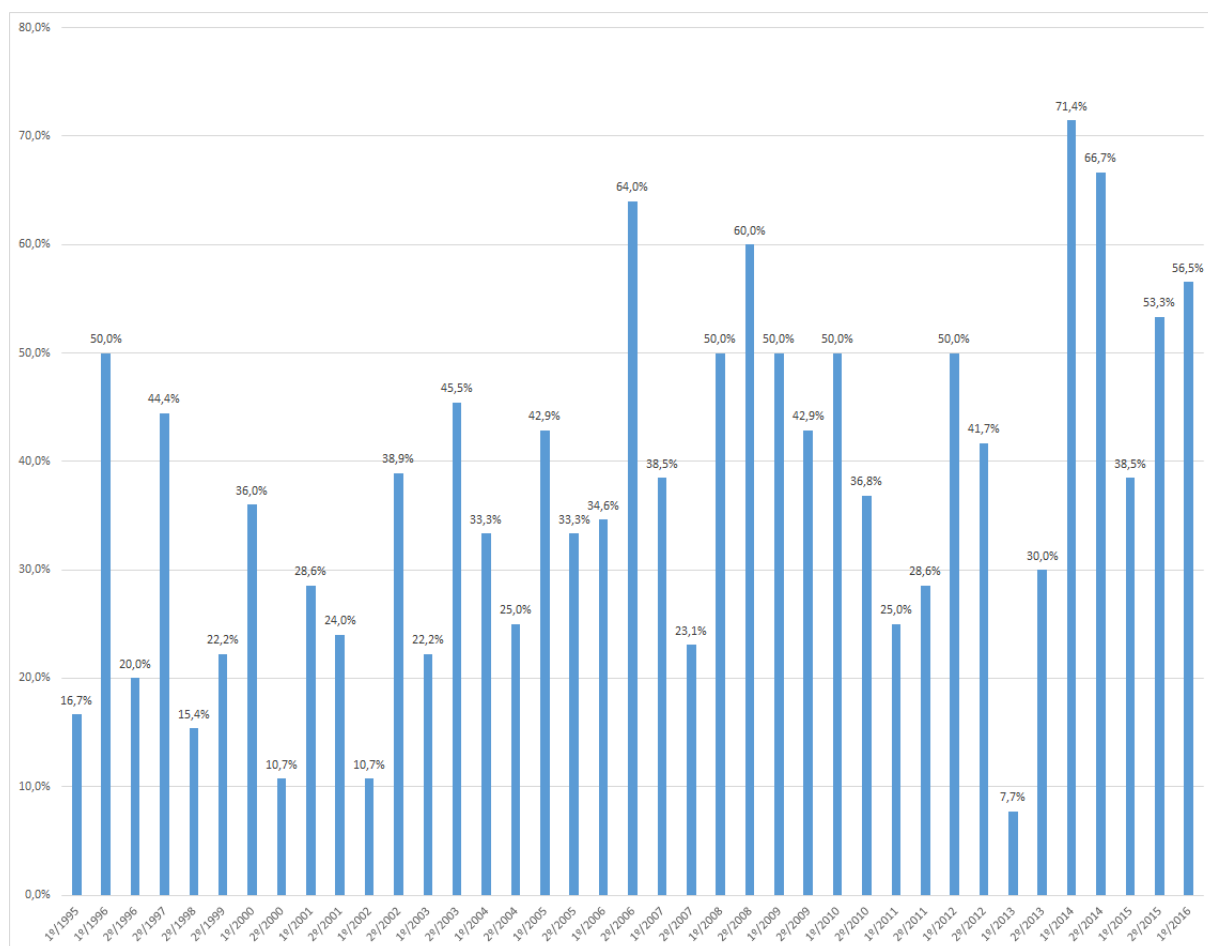


Figura 4.14: Porcentagem de reprovação na disciplina Teoria dos Números.

Quarto Semestre

A disciplina obrigatória do quinto semestre é Álgebra 1.

A Figura 4.15 refere-se à disciplina de Álgebra 1.

As maiores taxas de reprovação foram:

- 2º semestre de 2007 com 70%;
- 2º semestre de 2008 com 70,4%;
- 2º semestre de 2015 com 66,7%.

As menores taxas de reprovação foram:

- 1º semestre de 2000 com 8,3%;
- 2º semestre de 2005 com 10%;
- 1º semestre de 2011 com 12,5%.

O percentual médio de reprovação de todos os semestres foi de 37,6%.

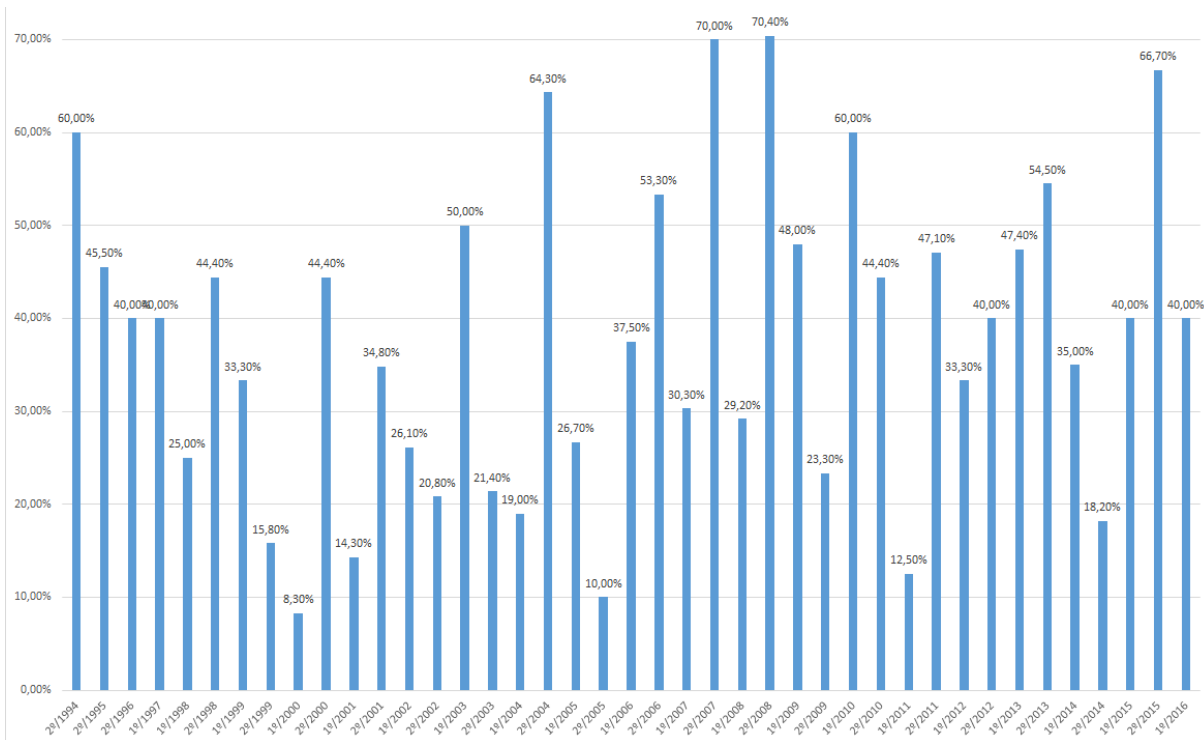


Figura 4.15: Porcentagem de reprovação na disciplina Álgebra 1.

Quinto Semestre

As disciplinas obrigatórias do quinto semestre são: Aprendizagem no Ensino e Geometria 1.

A Figura 4.16 refere-se à disciplina de Aprendizagem no Ensino.

As maiores taxas de reprovação foram:

- 1º semestre de 2008 com 51,9%;
- 2º semestre de 2014 com 60%;
- 2º semestre de 2015 com 66,7%.

As menores taxas de reprovação foram:

- 2º semestre de 1996 com 9,1%;
- 1º semestre de 2001 com 9,4%;
- 2º semestre de 2011 com 5,9%.

O percentual médio de reprovação de todos os semestres foi de 27,1%.

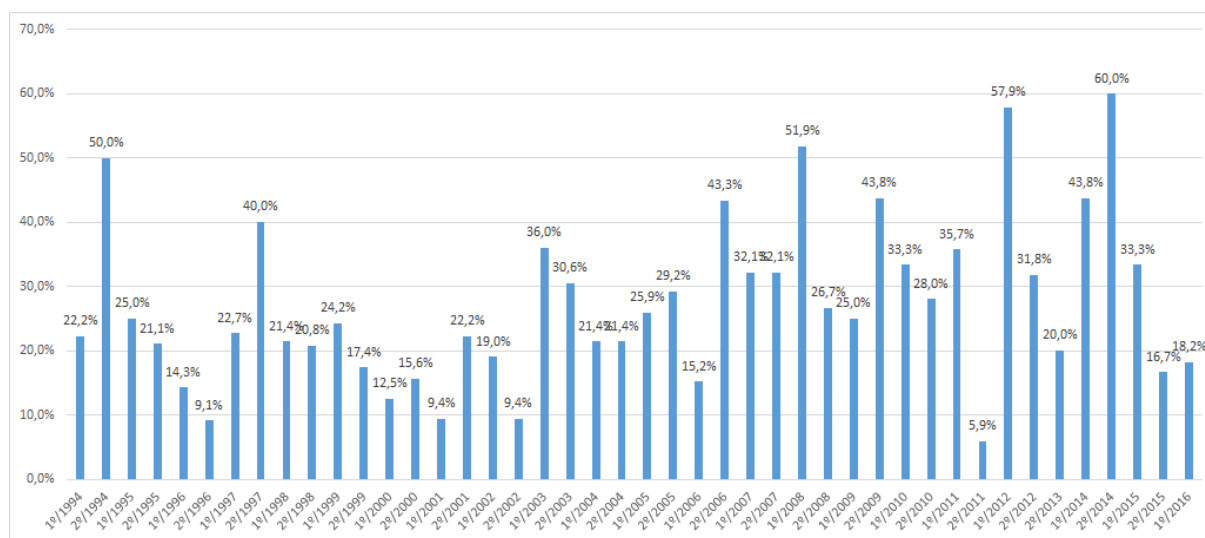


Figura 4.16: Porcentagem de reprovação na disciplina de Aprendizagem no Ensino.

A Figura 4.17 refere-se à disciplina de Geometria 1.

As maiores taxas de reprovação foram:

- 1º semestre de 2003 com 75%;
- 1º semestre de 2008 com 66,7%;
- 2º semestre de 2014 com 86,7%.

As menores taxas de reprovação foram:

- 2º semestre de 1998 com 5%;
- 2º semestre de 2000 com 5,1%;
- 2º semestre de 2003 com 7,1%.

O percentual médio de reprovação de todos os semestres foi de 32,6%.

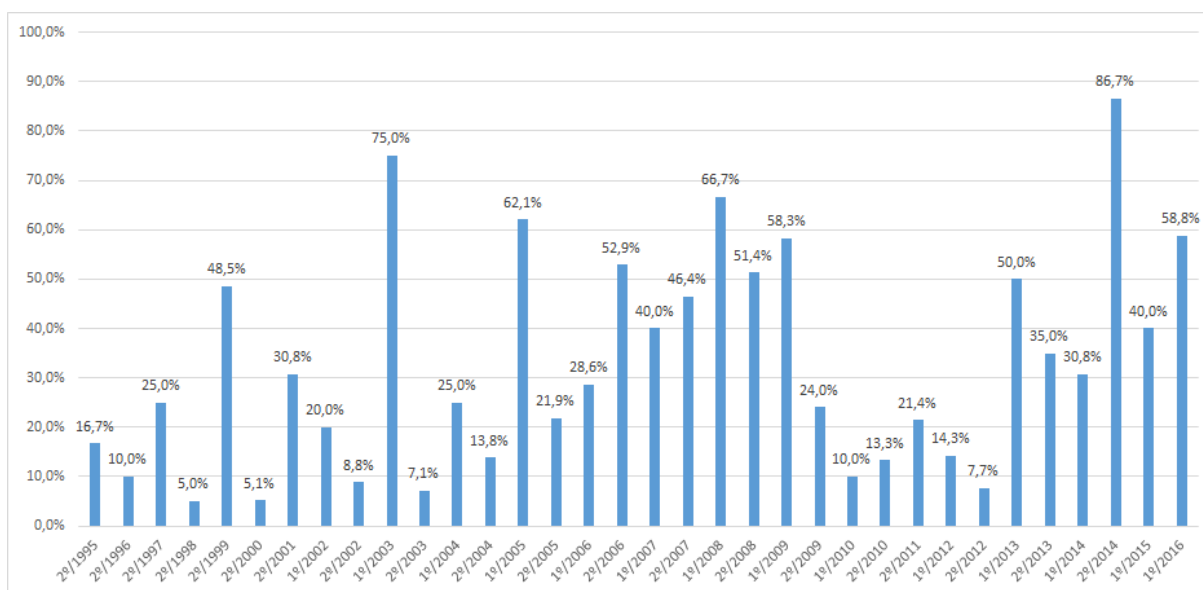


Figura 4.17: Porcentagem de reprovação na disciplina Geometria 1.

Sexto Semestre

As disciplinas obrigatórias do sexto semestre são: Cálculo de Probabilidade 1 e Geometria 2.

A Figura 4.18 refere-se à disciplina de Cálculo de Probabilidade 1.

As maiores taxas de reprovação foram:

- 2º semestre de 2006 com 87,5%;
- 2º semestre de 1994 com 85,7%;
- 1º semestre de 1994 com 86,7%.

As menores taxas de reprovação foram:

- 1º semestre de 1997 com 11,1%;
- 1º semestre de 1999 com 15,8%;
- 1º semestre de 2013 com 13,3%.

O percentual médio de reprovação de todos os semestres foi de 43%.

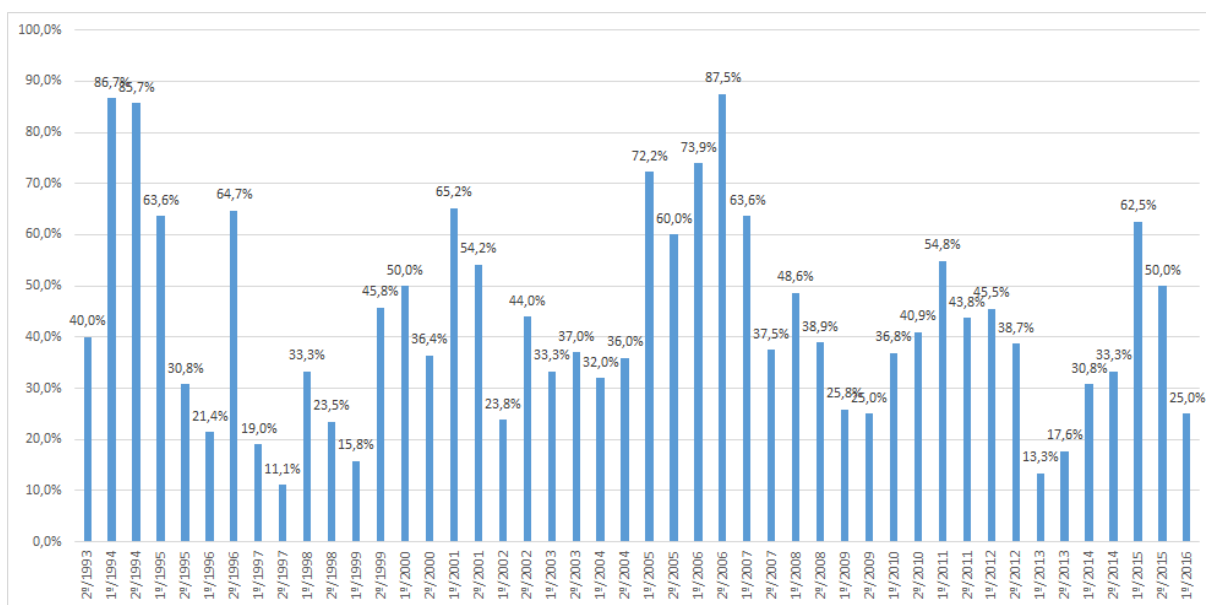


Figura 4.18: Porcentagem de reprovação na disciplina Cálculo de Probabilidade 1.

A Figura 4.19 refere-se à disciplina de Geometria 2.

As maiores taxas de reprovação foram:

- 1º semestre de 1996 com 50%;
- 2º semestre de 2002 com 66,7%;
- 2º semestre de 2003 com 50%.

As menores taxas de reprovação foram:

- 1º semestre de 2003 com 4,2%;
- 1º semestre de 2004 com 3,8%;
- 1º semestre de 2006 com 3,3%.

O percentual médio de reprovação de todos os semestres foi de 21,6%.

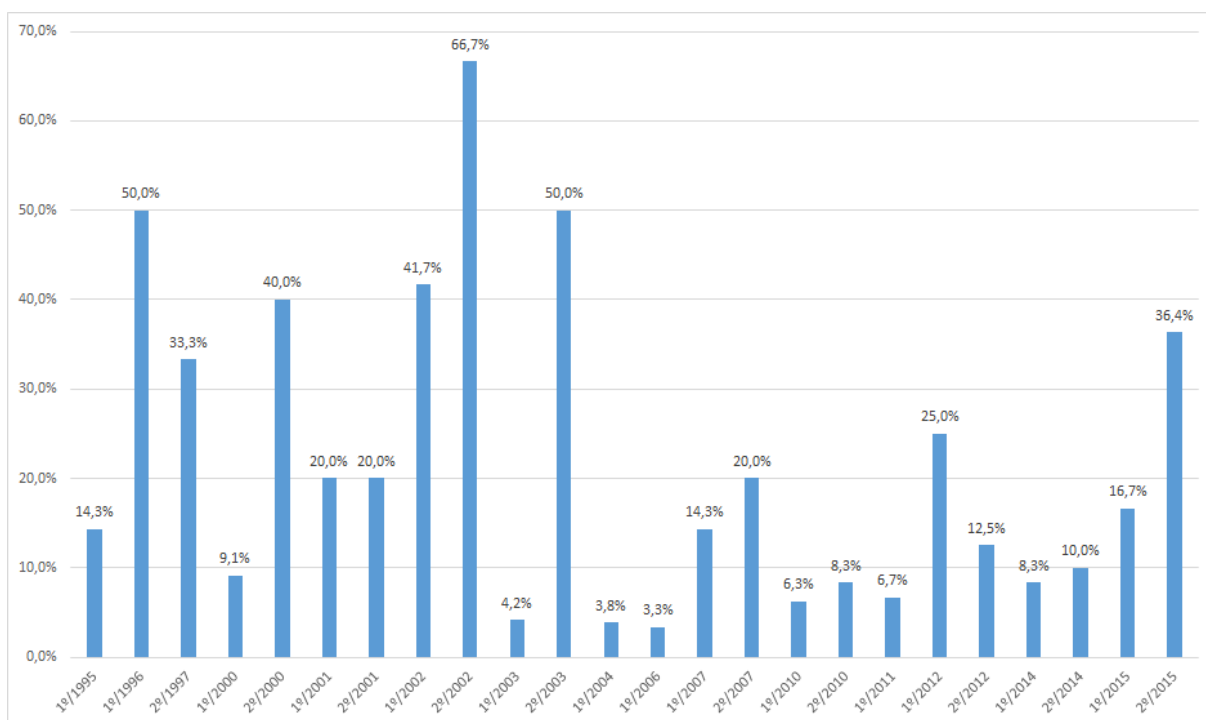


Figura 4.19: Porcentagem de reprovação na disciplina Geometria 2.

Sétimo Semestre

A disciplina obrigatória do sétimo semestre é Análise 1.

A Figura 4.20 refere-se à disciplina de Análise 1.

As maiores taxas de reprovação foram:

- 1º semestre de 2001 com 87,5%;
- 1º semestre de 2008 com 88,9%;
- 2º semestre de 2015 com 91,7%.

As menores taxas de reprovação foram:

- 2º semestre de 1996 com 16,7%;
- 2º semestre de 2007 com 16,1%;
- 1º semestre de 2010 com 14,3%.

O percentual médio de reprovação de todos os semestres foi de 47,6%.

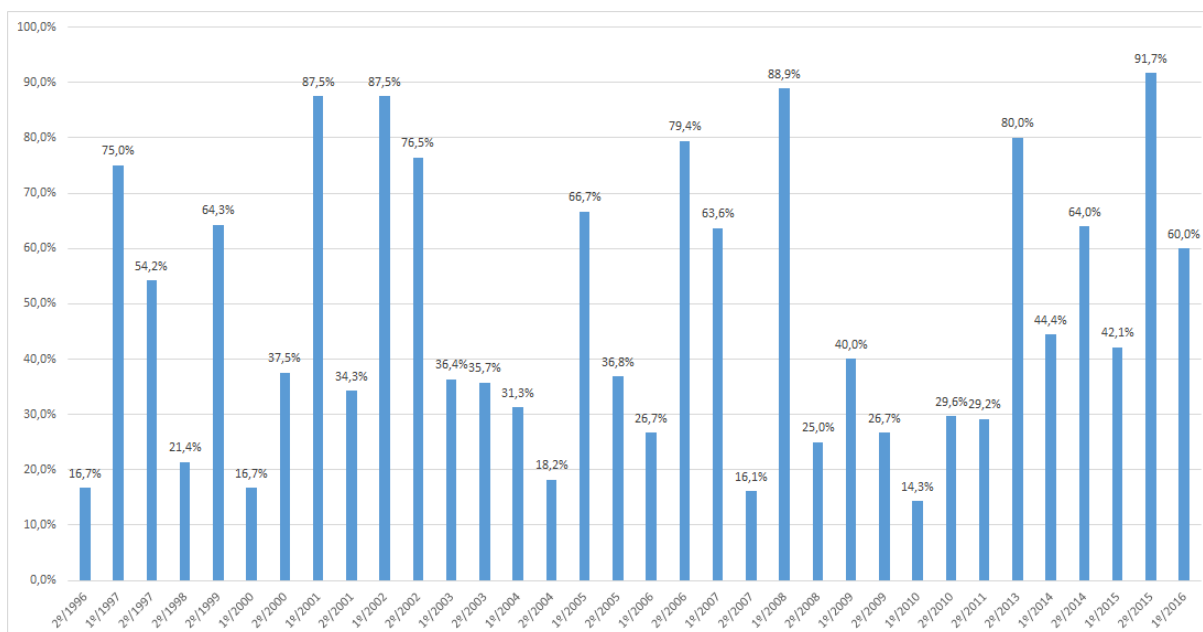


Figura 4.20: Porcentagem de reprovação na disciplina Análise 1.

Nono Semestre

A disciplina obrigatória do nono semestre é Variável Complexa 1.

A Figura 4.21 refere-se à disciplina de Variável Complexa 1.

As maiores taxas de reprovação foram:

- 1º semestre de 2001 com 87,5%;
- 1º semestre de 2008 com 88,9%;
- 2º semestre de 2011 com 75,0%.

As menores taxas de reprovação foram:

- 1º semestre de 2006 com 20,7%;
- 2º semestre de 2015 com 12,5%;
- 1º semestre de 2011 com 17,6%.

O percentual médio de reprovação de todos os semestres foi de 44,2%.

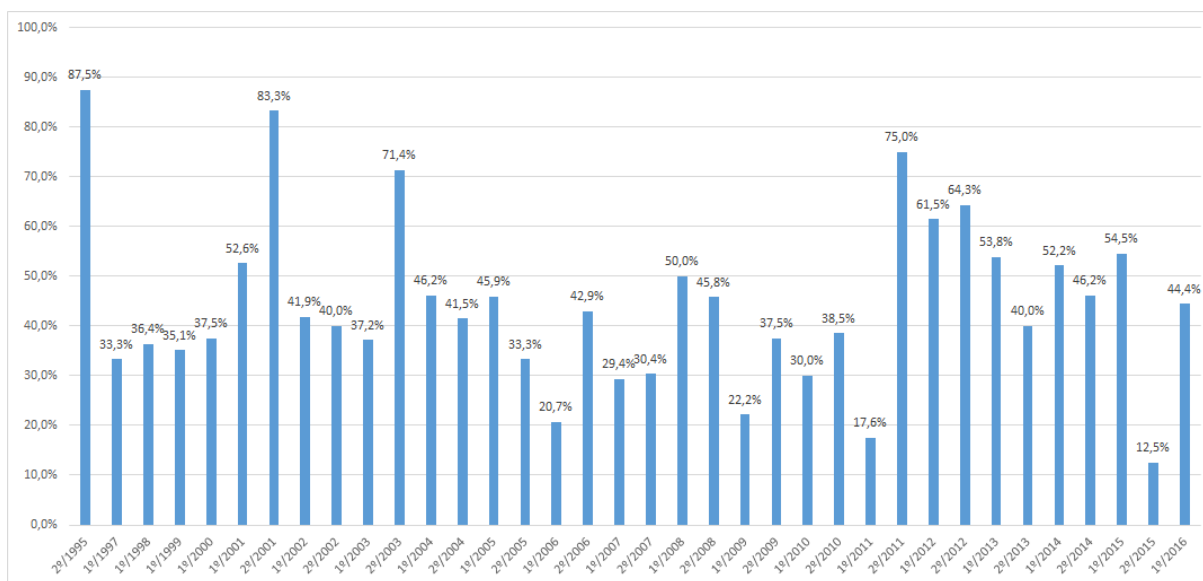


Figura 4.21: Percentagem de reprovações na disciplina Variável Complexa 1 .

Consolidação dos Dados Estatísticos

A Tabela 4.2 mostra as taxas de reprovação de todas as disciplinas obrigatórias do Curso de Licenciatura em Matemática (Noturno).

As três disciplinas com as maiores taxas de reprovações são de disciplinas dos dois primeiros semestres. Uma observação interessante é que, apesar de ser uma disciplina do último semestre, a disciplina de Variável Complexa, possui uma das maiores taxas de reprovação. A disciplina Física 1 possui a maior taxa de evasão, 56,7%, e juntamente das disciplinas Física 2 e Física 1 Experimental pertencem ao Departamento de Física. Conforme Tigrinho [32] como se destacou na Seção 2.2, a evasão do aluno pode estar relacionada ao baixo rendimento durante o Ensino Médio causado pela falta de conhecimentos básicos. Tal fato pode ser observado na Tabela 4.2, considerando que as maiores taxas de reprovação ocorrem principalmente nas disciplinas dos semestres iniciais. Esse déficit, acumulado durante o curso, pode levar à evasão do aluno.

Tabela 4.2: Taxas de reprovação das disciplinas obrigatórias.

Disciplinas Obrigatórias	Taxa Média de Reprovação	Semestre
FÍSICA 1	56,7%	2º
INTRODUÇÃO A ÁLGEBRA LINEAR	52%	2º
CÁLCULO 1	51,1%	1º
ANÁLISE 1	47,6%	7º
FÍSICA 2	46%	3º
FÍSICA 1 EXPERIMENTAL	45,4%	2º
VARIÁVEL COMPLEXA 1	44,2%	9º
CÁLCULO DE PROBABILIDADE 1	43%	6º
CÁLCULO 2	42,4%	2º
CÁLCULO 3	41,9%	3º
INTRODUÇÃO A CIÊNCIA COMPUTAÇÃO	41,8%	1º
ÁLGEBRA 1	37,6%	4º
TEORIA DOS NÚMEROS	36,9%	3º
GEOMETRIA 1	32,6%	5º
APRENDIZAGEM NO ENSINO	27,1%	5º
DESENVOLVIMENTO PSICOLÓGICO E ENSINO	24,4%	2º
ORGANIZAÇÃO DA EDUCAÇÃO BRASILEIRA	21,6%	1º
GEOMETRIA 2	21,6%	6º

Capítulo 5

Disciplinas Obrigatórias Relacionadas à Evasão

Neste capítulo, discutimos os resultados obtidos de mineração de dados realizado com os softwares: WEKA e Orange.

Na Seção 5.1 discutimos os resultados obtidos com o software WEKA. Na Seção 5.2 discutimos os resultados obtidos com o software Orange. Na Seção 5.3 consolidamos os resultados obtidos dos softwares WEKA e Orange, e os relacionamos às análises estatísticas do capítulo 4.

5.1 Mineração de Dados com o WEKA

As análises dessa Seção foram geradas no software WEKA com os algoritmos j48graft e REPTree. Resultados acima de 30 classificações corretas são consideradas como válidas, ou seja, pelo menos 30 alunos foram classificados pelo algoritmo de forma correta dentro do agrupamento específico (conjunto de disciplinas com determinadas menções). Este valor corresponde a 2% do total de alunos dos dados obtidos neste trabalho após a limpeza dos dados. Os atributos escolhidos foram: a forma de saída (evadiu ou formou) e as disciplinas obrigatórias.

A Figura 5.1 mostra que os alunos com o IRA menor que 2,4 foram caracterizados como prováveis evadidos com 502 classificações corretas. Portanto, é possível inferir que IRA menor que 2,4 está ligado à evasão dos alunos.

Ainda, alunos com IRA maior ou igual a 2,4 e menor que 2,86 e que reprovaram em Cálculo 1 foram caracterizados como prováveis evadidos com 37 classificações corretas. Portanto, é possível inferir que essa disciplina com esse intervalo de IRA está ligada à evasão dos alunos.

```

REPTree
=====
IRA < 2.4 : Evadiu (502/10) [270/4]
IRA >= 2.4
| IRA < 2.86
| | CALCULO_1_T1 = SR : Evadiu (7.25/2.34) [2.06/0.56]
| | CALCULO_1_T1 = II : Evadiu (15.96/6.94) [6.32/1.24]
| | CALCULO_1_T1 = MI : Evadiu (17.41/3.21) [7.54/3.35]
| | CALCULO_1_T1 = MM
| | | IRA < 2.54
| | | | PROBABILIDADE_F_ESTATISTICA_T1 = SR : Evadiu (0/0) [0/0]
| | | | PROBABILIDADE_F_ESTATISTICA_T1 = II : Evadiu (3.12/0.5) [0.47/0.08]
| | | | PROBABILIDADE_F_ESTATISTICA_T1 = MI : Evadiu (2.67/0.14) [1.41/0.07]
| | | | PROBABILIDADE_F_ESTATISTICA_T1 = MM : Evadiu (4/0.2) [3.61/1.1]
| | | | PROBABILIDADE_F_ESTATISTICA_T1 = MS : Evadiu (3.12/0.5) [1.47/0.08]
| | | | PROBABILIDADE_F_ESTATISTICA_T1 = SS : Formou (0.45/0.09) [0.41/0.06]
| | | | IRA >= 2.54
| | | | | CALCULO_2_T1 = SR : Evadiu (1.49/0.28) [0.16/0.12]
| | | | | CALCULO_2_T1 = II
| | | | | | IRA < 2.74 : Formou (4.46/0.44) [0.31/0.1]
| | | | | | IRA >= 2.74 : Evadiu (2.02/0.2) [2.41/0.31]
| | | | | | CALCULO_2_T1 = MI : Formou (2/0.28) [1.22/0.06]
| | | | | | CALCULO_2_T1 = MM : Formou (5.49/2.45) [5.94/3.17]
| | | | | | CALCULO_2_T1 = MS : Evadiu (3.01/0.91) [1.33/0.24]
| | | | | | CALCULO_2_T1 = SS : Evadiu (2.99/0.56) [0.33/0.24]
| | | | | CALCULO_1_T1 = MS : Evadiu (20.31/6.75) [10.96/6.58]
| | | | | CALCULO_1_T1 = SS

```

Figura 5.1: Dados gerados com o algoritmo REPTree no software WEKA.

A Figura 5.2 mostra que os alunos com menção SR em Introdução a Álgebra Linear e Álgebra 1 foram caracterizados como prováveis evadidos com apenas 1,8 classificações corretas. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, os alunos com menção SR em Introdução a Álgebra Linear e MI em Álgebra 1 foram caracterizados como prováveis evadidos com apenas 1,9 classificações corretas. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Além disso, os alunos com menção MI em Introdução a Álgebra Linear e II ou MI em Probabilidade e Estatística foram caracterizados como prováveis evadidos com apenas 7 classificações corretas. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

A Figura 5.3 mostra que os alunos com menção SR em Cálculo 1 e que reprovaram em Cálculo 3 foram caracterizados como prováveis evadidos com 3,4 classificações corretas. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, os alunos com menção MI em Cálculo 1 e que reprovaram em Cálculo 2 foram caracterizados como prováveis evadidos com 17,4 classificações corretas. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Por outro lado, os alunos com menção II em Cálculo 1 foram caracterizados como prováveis evadidos com 57,7 classificações corretas. Portanto, é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

```

INTRODUCAO_A_ALGEBRA_LINEAR_T1 = SR
| ALGEBRA_1_T1 = SR: Evadiu (1.8)
| ALGEBRA_1_T1 = II: Formou (2.3/0.3)
| ALGEBRA_1_T1 = MI: Evadiu (1.9)
| ALGEBRA_1_T1 = MM: Formou (3.3/0.3)
| ALGEBRA_1_T1 = MS: Formou (1.3/0.3)
| ALGEBRA_1_T1 = SS: Formou (0.0)
INTRODUCAO_A_ALGEBRA_LINEAR_T1 = II: Evadiu (57.7/13.0)
INTRODUCAO_A_ALGEBRA_LINEAR_T1 = MI
| PROBABILIDADE_E_ESTADISTICA_T1 = SR
| | CALCULO_1_T1 = SR: Evadiu (0.5)
| | CALCULO_1_T1 = II: Evadiu (0.2)
| | CALCULO_1_T1 = MI: Formou (0.0)
| | CALCULO_1_T1 = MM: Formou (2.3/0.3)
| | CALCULO_1_T1 = MS: Evadiu (0.7)
| | CALCULO_1_T1 = SS: Formou (0.0)
| PROBABILIDADE_E_ESTADISTICA_T1 = II: Evadiu (4.2)
| PROBABILIDADE_E_ESTADISTICA_T1 = MI: Evadiu (2.8/1.0)
| PROBABILIDADE_E_ESTADISTICA_T1 = MM: Formou (10.6/2.6)
| PROBABILIDADE_E_ESTADISTICA_T1 = MS: Formou (5.9/0.9)
| PROBABILIDADE_E_ESTADISTICA_T1 = SS: Formou (1.2/0.2)
-----

```

Figura 5.2: Dados gerados com o algoritmo j48graft no software WEKA.

```

CALCULO_1_T1 = SR
| CALCULO_3_T1 = SR: Evadiu (1.1)
| CALCULO_3_T1 = II: Evadiu (2.3)
| CALCULO_3_T1 = MI: Evadiu (0.0)
| CALCULO_3_T1 = MM: Formou (3.7/0.7)
| CALCULO_3_T1 = MS: Evadiu (0.6)
| CALCULO_3_T1 = SS: Evadiu (0.0)
CALCULO_1_T1 = II: Evadiu (57.7/12.0)
CALCULO_1_T1 = MI
| CALCULO_2_T1 = SR: Evadiu (5.7/1.0)
| CALCULO_2_T1 = II: Evadiu (3.3/1.0)
| CALCULO_2_T1 = MI: Evadiu (8.4/3.0)
| CALCULO_2_T1 = MM: Formou (17.5/5.5)
| CALCULO_2_T1 = MS: Formou (7.8/0.8)
| CALCULO_2_T1 = SS: Formou (2.0)

```

Figura 5.3: Dados gerados com o algoritmo j48graft no software WEKA.

5.2 Mineração de Dados com o Orange

As análises desta Seção foram geradas com o software Orange com o algoritmo CN2 Ruleg. A primeira coluna representa os nomes e as menções das disciplinas, a segunda a forma de evasão, a terceira coluna mostra as classificações corretas à esquerda e as incorretas à direita, a quarta coluna mostra as probabilidades de evadir à esquerda e as probabilidades de não evadir à direita. Resultados acima de 30 classificações corretas são consideradas como válidas, ou seja, pelo menos 30 alunos foram classificados pelo algoritmo de forma correta dentro do agrupamento específico (conjunto de disciplinas com determinadas menções) associada a determinada probabilidade (quarta coluna).

FISICA_1_EXPERIMENTAL_T1=II	→ SAIDA=Evadiu	[8, 0]	90 : 10	-0.00	1
FISICA_1_EXPERIMENTAL_T1=SR	→ SAIDA=Evadiu	[25.0, 0.0]	96 : 4	-0.00	1
FISICA_1_EXPERIMENTAL_T1=SR AND FISICA_1_T1=II	→ SAIDA=Evadiu	[9.1, 0.0]	91 : 9	-0.00	2
FISICA_1_EXPERIMENTAL_T1=SR AND FISICA_1_T1≠II	→ SAIDA=Evadiu	[8.4, 0.0]	90 : 10	-0.00	2
FISICA_1_EXPERIMENTAL_T1=SR AND FISICA_1_T1≠MM	→ SAIDA=Evadiu	[11.8, 0.0]	93 : 7	-0.00	2
FISICA_1_EXPERIMENTAL_T1=SR AND CALCULO_1_T1=II	→ SAIDA=Evadiu	[7.0, 0.0]	89 : 11	-0.00	2
FISICA_1_EXPERIMENTAL_T1=SR AND CALCULO_1_T1=MI	→ SAIDA=Evadiu	[5.7, 0.0]	87 : 13	-0.00	2
CALCULO_1_T1=SR AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR	→ SAIDA=Evadiu	[54.2, 0.0]	98 : 2	-0.00	2
CALCULO_1_T1=SR AND CALCULO_3_T1=MM	→ SAIDA=Evadiu	[8.1, 0.0]	90 : 10	-0.00	2
CALCULO_1_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=II	→ SAIDA=Evadiu	[37.0, 0.0]	97 : 3	-0.00	2
CALCULO_1_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MI	→ SAIDA=Evadiu	[18.2, 0.0]	95 : 5	-0.00	2
CALCULO_1_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS	→ SAIDA=Evadiu	[30.2, 0.0]	97 : 3	-0.00	2
CALCULO_1_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM	→ SAIDA=Evadiu	[158.8, 0.0]	99 : 1	-0.00	2
CALCULO_1_T1=SR AND PROBABILIDADE_E_ESTADISTICA_T1=MM	→ SAIDA=Evadiu	[106.9, 0.0]	99 : 1	-0.00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND CALCULO_1_T1=II	→ SAIDA=Evadiu	[13.2, 0.0]	93 : 7	-0.00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND CALCULO_1_T1=MI	→ SAIDA=Evadiu	[7.0, 0.0]	89 : 11	-0.00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR	→ SAIDA=Evadiu	[24.6, 0.0]	96 : 4	-0.00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND CALCULO_3_T1=SR	→ SAIDA=Evadiu	[5.7, 0.0]	87 : 13	-0.00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND CALCULO_2_T1=SR	→ SAIDA=Evadiu	[19.6, 0.0]	95 : 5	-0.00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND FISICA_1_T1=II	→ SAIDA=Evadiu	[10.4, 0.0]	92 : 8	-0.00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM	→ SAIDA=Evadiu	[86.1, 0.0]	99 : 1	-0.00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND CALCULO_1_T1=MS	→ SAIDA=Evadiu	[12.5, 0.0]	93 : 7	-0.00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND CALCULO_3_T1=SR	→ SAIDA=Evadiu	[12.8, 0.0]	93 : 7	-0.00	2
CALCULO_2_T1=SR AND CALCULO_1_T1=MM	→ SAIDA=Evadiu	[62.0, 0.0]	98 : 2	-0.00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=II	→ SAIDA=Evadiu	[18.9, 0.0]	95 : 5	-0.00	2

Figura 5.4: Dados gerados com o algoritmo CN2 Rule no Orange.

Pode-se observar na Figura 5.4 que, quem teve a menção II em Física Experimental 1 evadiu. Apesar de ter uma probabilidade de 90% apenas 8 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, quem teve a menção SR em Física Experimental 1, e com a menção II em Física 1 evadiu. Apesar de ter uma probabilidade de 91% apenas 8 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Por outro lado, quem teve a menção SR em Cálculo 1 e com a menção II em Introdução a Álgebra Linear evadiu, com uma probabilidade de 98% teve 54,2 classificações corretas. Portanto, é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=II AND CALCULO_1_T1#MI	→ SAIDA=Evadiu	[6,4, 0,0]	88 : 12	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=II AND CALCULO_1_T1#MS	→ SAIDA=Evadiu	[5,0, 0,0]	86 : 14	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND FISICA_1_T1=II	→ SAIDA=Evadiu	[5,5, 0,0]	87 : 13	-0,00	3
CALCULO_2_T1=SR AND CALCULO_1_T1=MM AND FISICA_1_T1=II	→ SAIDA=Evadiu	[31,1, 0,0]	97 : 3	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND CALCULO_1_T1=II AND FISICA_1_T1#SR	→ SAIDA=Evadiu	[5,0, 0,0]	86 : 14	-0,00	3
CALCULO_1_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1#II	→ SAIDA=Evadiu	[28,4, 0,0]	97 : 3	-0,00	3
CALCULO_2_T1=SR AND CALCULO_1_T1=MM AND FISICA_1_T1#MI	→ SAIDA=Evadiu	[22,2, 0,0]	96 : 4	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND FISICA_1_T1=II	→ SAIDA=Evadiu	[33,0, 0,0]	97 : 3	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND FISICA_1_T1=SR	→ SAIDA=Evadiu	[5,2, 0,0]	86 : 14	-0,00	2
CALCULO_1_T1=II AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=SS	→ SAIDA=Evadiu	[9,1, 0,0]	91 : 9	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND CALCULO_3_T1=SR	→ SAIDA=Evadiu	[6,4, 0,0]	88 : 12	-0,00	2
CALCULO_1_T1=II AND CALCULO_2_T1=MI	→ SAIDA=Evadiu	[10,4, 0,0]	92 : 8	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MI	→ SAIDA=Evadiu	[8,3, 0,0]	90 : 10	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=SS	→ SAIDA=Evadiu	[7,8, 0,0]	90 : 10	-0,00	2
ALGEBRA_1_T1=MS AND CALCULO_1_T1=MI	→ SAIDA=Formou	[0,0, 7,0]	11 : 89	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND TEORIA_DOS_NUMEROS_T1=II	→ SAIDA=Evadiu	[5,9, 0,0]	87 : 13	-0,00	2
CALCULO_2_T1=SR AND CALCULO_1_T1=MM AND FISICA_1_T1#SR	→ SAIDA=Evadiu	[15,3, 0,0]	94 : 6	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND CALCULO_1_T1=SS AND FISICA_1_T1=II	→ SAIDA=Formou	[0,0, 5,5]	13 : 87	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=SS AND CALCULO_1_T1#MI	→ SAIDA=Formou	[0,0, 5,1]	14 : 86	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MM AND FISICA_1_T1#MI	→ SAIDA=Formou	[0,0, 5,1]	14 : 86	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND CALCULO_2_T1=MS AND FISICA_1_T1=II	→ SAIDA=Formou	[0,0, 9,4]	9 : 91	-0,00	3
CALCULO_2_T1=SR AND CALCULO_1_T1=MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1#II	→ SAIDA=Evadiu	[9,5, 0,0]	91 : 9	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND CALCULO_2_T1=MS AND FISICA_1_T1#MI	→ SAIDA=Formou	[0,0, 7,7]	10 : 90	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MI AND FISICA_1_T1=II	→ SAIDA=Evadiu	[5,1, 0,0]	86 : 14	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND CALCULO_2_T1=SR AND CALCULO_1_T1#MI	→ SAIDA=Evadiu	[5,1, 0,0]	86 : 14	-0,00	3
CALCULO_1_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1#MI	→ SAIDA=Evadiu	[18,3, 0,0]	95 : 5	-0,00	3

Figura 5.5: Dados gerados com o algoritmo CN2 Rule no Orange.

Pode-se observar na Figura 5.5 que, quem teve a menção SR em Introdução a Álgebra Linear, e com a menção II em Introdução a Ciência da Computação, e menção diferente de MI em Cálculo 1 evadiu. Apesar de ter uma probabilidade de 88% apenas 6,4 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, quem teve a menção SR em Introdução a Álgebra Linear, e com a menção II em Introdução a Ciência da Computação, e menção diferente de MS em Cálculo 1 evadiu. Apesar de ter uma probabilidade de 86% somente 5 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Por outro lado, quem teve a menção SR em Cálculo 2, e com a menção MM em Cálculo 1, e II em Física 1 evadiu, com uma probabilidade de 97% teve 31,1 classificações feitas corretamente. Portanto, é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

INTRODUCAO_A_CIEN_COMPUTACAO_T1=SS AND ALGEBRA_1_T1≠MI	→ SAIDA=Evadiu	[5,2, 0,0]	86 : 14	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND CALCULO_2_T1=SR AND CALCULO_1_T1≠SR	→ SAIDA=Evadiu	[5,2, 0,0]	86 : 14	-0,00	3
ALGEBRA_1_T1=MS AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MM	→ SAIDA=Formou	[0,0, 8,5]	9 : 91	-0,00	2
CALCULO_1_T1=II AND CALCULO_2_T1=MI AND FISICA_1_T1=II	→ SAIDA=Evadiu	[6,0, 0,0]	87 : 13	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND CALCULO_2_T1=MS AND FISICA_1_T1≠MM	→ SAIDA=Formou	[0,0, 5,6]	13 : 87	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=II AND CALCULO_2_T1=II AND CALCULO_1_T1≠SR	→ SAIDA=Evadiu	[5,1, 0,0]	86 : 14	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND CALCULO_1_T1=II	→ SAIDA=Evadiu	[17,4, 0,0]	95 : 5	-0,00	3
ALGEBRA_1_T1=MS AND CALCULO_2_T1=MS	→ SAIDA=Formou	[0,0, 14,2]	6 : 94	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND CALCULO_1_T1≠MI	→ SAIDA=Evadiu	[37,9, 0,0]	97 : 3	-0,00	3
ALGEBRA_1_T1=MS AND FISICA_2_T1≠MM	→ SAIDA=Formou	[0,0, 7,8]	10 : 90	-0,00	2
INTRODUCAO_A_CIEN_COMPUTACAO_T1=II AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=MM AND FISICA_1_T1=II	→ SAIDA=Evadiu	[14,3, 0,0]	94 : 6	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND FISICA_1_T1≠MI	→ SAIDA=Evadiu	[20,3, 0,0]	96 : 4	-0,00	3
ALGEBRA_1_T1=MS AND TEORIA_DOS_NUMEROS_T1=SS	→ SAIDA=Formou	[0,0, 9,7]	9 : 91	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND CALCULO_1_T1=MM	→ SAIDA=Evadiu	[12,9, 0,0]	93 : 7	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND PROBABILIDADE_E_ESTADISTICA_T1=SR AND CALCULO_1_T1≠MI	→ SAIDA=Evadiu	[11,2, 0,0]	92 : 8	-0,00	3

Figura 5.6: Dados gerados com o algoritmo CN2 Rule no Orange.

Pode-se observar na Figura 5.6 que, quem teve a menção II em Introdução a Álgebra Linear, e com a menção SR em Cálculo 2, e menção diferente de SR em Cálculo 1 evadiu. Apesar de ter uma probabilidade de 86% apenas 5,2 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, quem teve a menção II em Cálculo 1, e com a menção MI em Cálculo 2, e menção de II em Física 1 evadiu. Apesar de ter uma probabilidade de 87% apenas 6 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Além disso, quem teve a menção II em Introdução a Álgebra Linear, e com a menção SR em Probabilidade e Estatística, e menção diferente de MI em Cálculo 1 evadiu. Apesar de ter uma probabilidade de 92% apenas 11,2 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND CALCULO_3_T1≠MI	→ SAIDA=Evadiu	[29,4, 0,0]	97 : 3	-0,00	3
ALGEBRA_1_T1=MS AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MM A...	→ SAIDA=Formou	[0,0, 5,1]	14 : 86	-0,00	3
ALGEBRA_1_T1=MS AND CALCULO_2_T1=MS AND FISICA_1_T1=II	→ SAIDA=Formou	[0,0, 8,1]	10 : 90	-0,00	3
ALGEBRA_1_T1=MS AND CALCULO_2_T1=MS AND CALCULO_1_T1≠MI	→ SAIDA=Formou	[0,0, 6,3]	12 : 88	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=II AND FISICA_1_T1=II	→ SAIDA=Evadiu	[5,6, 0,0]	87 : 13	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND ALGEBRA_1_T1=MM AND CALCULO_1_T1=II	→ SAIDA=Evadiu	[16,2, 0,0]	95 : 5	-0,00	3
ALGEBRA_1_T1=MS AND TEORIA_DOS_NUMEROS_T1=SS AND FISICA_1_T1=II	→ SAIDA=Formou	[0,0, 5,8]	13 : 87	-0,00	3
CALCULO_1_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=MM	→ SAIDA=Evadiu	[8,0, 0,0]	90 : 10	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND CALCULO_1_T1=MI AND ALGEBRA_1_T1=MM	→ SAIDA=Evadiu	[10,6, 0,0]	92 : 8	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND PROBABILIDADE_E_ESTADISTICA_T1=SR AND ALGEBRA_1_T1=MM	→ SAIDA=Evadiu	[8,7, 0,0]	91 : 9	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND CALCULO_3_T1=MM	→ SAIDA=Evadiu	[18,2, 0,0]	95 : 5	-0,00	3
ALGEBRA_1_T1=MS AND CALCULO_3_T1=II AND INTRODUCAO_A_ALGEBRA_LINEAR_T1≠SR	→ SAIDA=Formou	[0,0, 5,2]	14 : 86	-0,00	3

Figura 5.7: Dados gerados com o algoritmo CN2 Rule no Orange.

Pode-se observar na Figura 5.7 que, quem teve a menção SR em Introdução a Álgebra Linear, e com a menção MM em Variável Complexa, e menção diferente de MI em Cálculo 3 evadiu. Apesar de ter uma probabilidade de 97% apenas 29,4 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, quem teve a menção II em Introdução a Álgebra Linear, e com a menção MM em Álgebra 1, e menção diferente de II em Cálculo 1 evadiu. Apesar de ter uma probabilidade de 95% apenas 16,2 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Além disso, quem teve a menção II em Introdução a Álgebra Linear, e com a menção MI em Cálculo 1, e MM em Álgebra 1 evadiu. Apesar de ter uma probabilidade de 92% apenas 10,6 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND ALGEBRA_1_T1=MM AND CALCULO_3_T1=MM	→ SAIDA=Evadiu	[32,9, 0,0]	97 : 3	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND TEORIA_DOS_NUMEROS_T1=MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=MM	→ SAIDA=Formou	[0,0, 8,5]	9 : 91	-0,00	3
CALCULO_3_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS AND INTRODUCAO_A_ALGEBRA_LINEAR_T1≠MM	→ SAIDA=Evadiu	[9,4, 0,0]	91 : 9	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND CALCULO_2_T1≠MM AND CALCULO_1_T1≠MM	→ SAIDA=Formou	[0,0, 8,6]	9 : 91	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND TEORIA_DOS_NUMEROS_T1=MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1≠MS	→ SAIDA=Formou	[0,0, 9,0]	9 : 91	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=II AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=MM AND FISICA_1_T1≠MI	→ SAIDA=Evadiu	[10,1, 0,0]	92 : 8	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND ALGEBRA_1_T1=MM AND CALCULO_2_T1=II	→ SAIDA=Evadiu	[6,7, 0,0]	89 : 11	-0,00	3
CALCULO_3_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS AND TEORIA_DOS_NUMEROS_T1≠MM	→ SAIDA=Evadiu	[9,1, 0,0]	91 : 9	-0,00	3
ALGEBRA_1_T1=MS AND VARIAVEL_COMPLEXA_1_T1=II AND FISICA_1_T1=II	→ SAIDA=Formou	[0,0, 7,6]	10 : 90	-0,00	3
CALCULO_2_T1=SR AND PROBABILIDADE_E_ESTADISTICA_T1=MM AND FISICA_1_T1=II	→ SAIDA=Evadiu	[5,3, 0,0]	86 : 14	-0,00	3
TEORIA_DOS_NUMEROS_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS AND CALCULO_1_T1≠MM	→ SAIDA=Evadiu	[6,8, 0,0]	89 : 11	-0,00	3
TEORIA_DOS_NUMEROS_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS AND CALCULO_3_T1≠MM	→ SAIDA=Evadiu	[8,9, 0,0]	91 : 9	-0,00	3

Figura 5.8: Dados gerados com o algoritmo CN2 Rule no Orange.

Pode-se observar na Figura 5.8 que, quem teve a menção II em Introdução a Álgebra Linear, e com a menção MM em Álgebra 1, e MM em Cálculo 3 evadiu, com uma probabilidade de 97% teve 32,9 classificações feitas corretamente. Portanto, é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Por outro lado, quem teve a menção SR em Teoria dos Números, e com a menção MS em Introdução a Ciência da Computação, e menção diferente de MM em Cálculo 1 evadiu. Apesar de ter uma probabilidade de 89% apenas 6,8 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, quem teve a menção II em Introdução a Ciência da Computação, e com a menção MM em Introdução a Álgebra Linear, e menção MM em Física 1 evadiu. Apesar de ter uma probabilidade de 92% apenas 10,1 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

ALGEBRA_1_T1=MS AND TEORIA_DOS_NUMEROS_T1=MM AND PROBABILIDADE_E_ESTADISTICA_T1=MS	→ SAIDA=Formou	[0,0, 5,1]	14 : 86	-0,00	3
CALCULO_2_T1=SR AND ALGEBRA_1_T1=MM AND FISICA_1_T1=II	→ SAIDA=Evadiu	[13,8, 0,0]	94 : 6	-0,00	3
TEORIA_DOS_NUMEROS_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS AND CALCULO_2_T1≠MM	→ SAIDA=Evadiu	[5,7, 0,0]	87 : 13	-0,00	3
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND PROBABILIDADE_E_ESTADISTICA_T1=SR AND ALGEBRA_1_T1≠MS	→ SAIDA=Evadiu	[6,0, 0,0]	88 : 12	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND CALCULO_2_T1≠MM AND INTRODUCAO_A_ALGEBRA_LINEAR_T1≠MS	→ SAIDA=Formou	[0,0, 5,6]	13 : 87	-0,00	3
CALCULO_3_T1=SR AND CALCULO_2_T1=MM AND TEORIA_DOS_NUMEROS_T1≠MM	→ SAIDA=Evadiu	[6,6, 0,0]	88 : 12	-0,00	3
TEORIA_DOS_NUMEROS_T1=SR AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MS AND ALGEBRA_1_T1=SR	→ SAIDA=Evadiu	[8,3, 0,0]	90 : 10	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=SR AND VARIAVEL_COMPLEXA_1_T1=MM AND FISICA_1_T1≠SR	→ SAIDA=Evadiu	[11,2, 0,0]	92 : 8	-0,00	3
INTRODUCAO_A_CIEN_COMPUTACAO_T1=MI AND CALCULO_1_T1=MS	→ SAIDA=Evadiu	[6,3, 0,0]	88 : 12	-0,00	2
INTRODUCAO_A_ALGEBRA_LINEAR_T1=II AND CALCULO_1_T1=MI AND ALGEBRA_1_T1≠MS	→ SAIDA=Evadiu	[5,2, 0,0]	86 : 14	-0,00	3
CALCULO_1_T1=II AND INTRODUCAO_A_ALGEBRA_LINEAR_T1=MM AND INTRODUCAO_A_CIEN_COMPUTACAO_T1=MI	→ SAIDA=Evadiu	[8,7, 0,0]	91 : 9	-0,00	3
VARIAVEL_COMPLEXA_1_T1=MS AND TEORIA_DOS_NUMEROS_T1=MM AND CALCULO_3_T1=MM	→ SAIDA=Formou	[0,0, 5,8]	13 : 87	-0,00	3

Figura 5.9: Dados gerados com o algoritmo CN2 Rule no Orange.

Pode-se observar na Figura 5.9 que, quem teve a menção SR em Cálculo 2, e com a menção MM em Álgebra 1, e II em Física 1 evadiu. Apesar de ter uma probabilidade de 94% apenas 13,8 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Ainda, quem teve a menção SR em Teoria dos Números, e com a menção MS em Introdução a Ciência da Computação, e SR em Álgebra 1 evadiu. Apesar de ter uma probabilidade de 90% apenas 8,3 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

Além disso, quem teve a menção II em Cálculo 1, e com a menção MM em Introdução a Álgebra Linear, e de MI Introdução a Ciência da Computação evadiu. Apesar de ter uma probabilidade de 91% apenas 8,7 classificações foram feitas corretamente. Portanto, não é possível inferir que essas disciplinas e menções estão ligadas à evasão dos alunos.

5.3 Consolidação dos resultados obtidos

Os resultados obtidos com o WEKA e com o Orange mostraram que reprovações nas disciplinas: Cálculo 1, Cálculo 2 e Introdução a Álgebra Linear são as que mais provavelmente influenciam para o aluno evadir. Apesar de Física 1 e Física 1 Experimental estarem entre as disciplinas com maiores taxas de reprovação, conforme a classificação da Tabela 4.2, estas disciplinas não influenciam tanto na evasão do aluno se comparadas com as disciplinas do próprio Departamento de Matemática dos três primeiros semestres do curso. Os alunos com o IRA menor que 2,4 foram caracterizados como prováveis evadidos.

Capítulo 6

Conclusão e Trabalhos Futuros

Neste trabalho foram realizadas análises de evasão do Curso Licenciatura em Matemática (Noturno) da Universidade de Brasília, a partir de menções finais de alunos desse curso. Em particular, foram realizadas análises estatísticas das menções finais em disciplinas obrigatórias desse curso e em seguida foram identificadas disciplinas obrigatórias que possivelmente estão relacionadas à evasão dos alunos usando técnicas de mineração de dados, com os softwares WEKA e ORANGE.

As análises feitas com o WEKA permitiram inferir as seguintes informações. Primeiro, os alunos com o IRA menor que 2,4 foram caracterizados como prováveis evadidos.

As análises realizadas a partir dos resultados obtidos utilizando os softwares WEKA e Orange mostraram que reprovações nas disciplinas: Cálculo 1, Cálculo 2 e Introdução a Álgebra Linear estão ligadas à evasão de alunos. Na tabela de disciplinas com altas taxas de reprovações, as três disciplinas que devem ser cursadas nos dois primeiros semestres aparecem entre as que mais reprovam. Nota-se ainda que as disciplinas Variável Complexa 1, Cálculo de Probabilidade 1, Cálculo 3 e Álgebra 1 tem altas taxas de reprovação, mas não foram identificadas como ligadas à evasão. Isso pode indicar que os alunos têm a tendência de evadirem do curso nos primeiros semestres com reprovações em disciplinas obrigatórias do Departamento de Matemática.

6.1 Contribuições

As contribuições mais relevantes deste projeto são:

- Identificação das disciplinas obrigatórias com os maiores índices de reprovação do Curso de Licenciatura em Matemática (Noturno);
- Identificação de grupos de disciplinas e de menções de reprovação que podem ser usadas para identificar alunos com perfil de evasão.

6.2 Trabalhos futuros

Como trabalhos futuros que poderiam aprofundar o entendimento da evasão na UnB, sugerimos:

- Verificar para que outros cursos da UnB foram os alunos que evadiram do Curso de Licenciatura em Matemática (Noturno);
- Identificar as disciplinas do Departamento de Matemática da UnB que mais influenciam na evasão dos alunos de outros departamentos.

Referências

- [1] M. T. Angeloni. Elementos intervenientes na tomada de decisão. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/view/120/101>, acesso em 12 de Março de 2016, 04 2003. 20
- [2] L. A. Azevedo e Y. da S. Santos. Mineração de dados aplicada ao estudo da evasão e desempenho dos alunos do Bacharelado em Ciência da Computação da Universidade de Brasília. Disponível em: http://bdm.unb.br/bitstream/10483/11140/1/2015_LuizaAlencarAzevedo_YagodaSilvaSantos.pdf, 2015. 2, 17
- [3] C. A. S. Baggi e D. A. Lopes. Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 16:355 – 374, 07 2011. 5
- [4] M. J. Berry e G. Linoff. *Data Mining Techniques: for marketing, sales, and customer support*. John Wiley & Sons, Inc., New York, NY, USA, 1997. 20
- [5] R. J. Brachman e T. Anand. Advances in Knowledge Discovery and Data Mining. chapter The Process of Knowledge Discovery in Databases, pages 37–57. American Association for Artificial Intelligence, Menlo Park, CA, USA, 1996. 21
- [6] P. Cabena, P. Hadjinian, R. Stadler, J. Verhees, e A. Zanasi. *Discovering Data Mining: from concept to implementation*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1998. 21, 23
- [7] C. B. Cardoso. Efeitos da política de cotas na Universidade de Brasília: uma análise do rendimento e da evasão. Disponível em: <http://repositorio.unb.br/handle/10482/1891>, acesso em 11 de Dezembro de 2015, 2008. ix, 17, 18, 19
- [8] J. V. A. Cunha, E. M. Nascimento, e S. O. Durso. Razões e influências para a evasão universitária: um estudo com estudantes ingressantes nos Cursos de Ciências Contábeis de instituições públicas federais da Região Sudeste. Disponível em: <http://www.congressosp.fipecafi.org/web/artigos142014/403.pdf>, acesso em 05 de Dezembro de 2015, 2011. 13
- [9] T. H. Davenport e L. Prusak. *Information ecology: mastering the information and knowledge environment*. Oxford University Press, 1st edition, 1997. 20
- [10] Universidade de Brasília (2016) Acessado em: 09 de outubro de 2016. Currículo do Curso de Licenciatura em Matemática (Noturno). xii, 32

- [11] Decanato de Ensino de Graduação UnB. Decanato de Ensino de Graduação, 2016. Disponível em: http://www.unb.br/administracao/decanatos/deg/downloads/index/guia_coordenador.pdf, acesso em 22 de novembro de 2016. 33
- [12] INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Sínteses Da Educação Básica, 2014. Disponível em: <http://portal.inep.gov.br/basica-censo-escolar-sinopse-sinopse>, acesso em 01 de janeiro de 2016. ix, 15
- [13] Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras. Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas, 1996. Disponível em: <http://www.dominiopublico.gov.br/download/texto/me001613.pdf>, acesso em 12 de Dezembro de 2015, 1996. 4
- [14] U. M. Fayyad, G. Piatetsky-Shapiro, e P. Smyth. Advances in knowledge discovery and data mining. chapter From Data Mining to Knowledge Discovery: An Overview, pages 1–34. American Association for Artificial Intelligence, Menlo Park, CA, USA, 1996. 21, 22
- [15] L. Felicetti e P. Fossatti. Alunos ProUni e não ProUni nos cursos de licenciatura: evasão em foco. Disponível em: <http://ojs.c3s1.ufpr.br/ojs/index.php/educar/article/download/32842/22084>, acesso em 31 de Dezembro de 2015, 2014. 1, 15, 16
- [16] N. Galvão e H. Marin. Técnica de Mineração de Dados: uma revisão da literatura. *Acta Paul Enferm*, 22:686 – 690, 10 2009. 22
- [17] D. Garcia. Evasão no Curso de Licenciatura em Matemática (Diurno) da Universidade de Brasília, 2017. 33
- [18] P. Gladieux e L. Perna. Borrowers who drop out a neglected aspect of the college student loan trend. Disponível em: <http://www.highereducation.org/reports/borrowing/borrowers.pdf>, pages = 2-5, acesso em 04 de janeiro de 2016, 2005. ix, 9, 10, 11, 12
- [19] M. Gomes, M. Monteiro, A. Damasceno, T. Almeida, e R. Carvalho. Evasão acadêmica no ensino superior: Estudo na Área da Saúde. Disponível em: www.congressosp.fipecafi.org/web/artigos102010/419.pdf, pages = 6-13, acesso em 01 de janeiro de 2016, 2010. ix, 1, 14, 15
- [20] J. Han e M. Kamber. *Data Mining: concepts and techniques*. Morgan Kaufmann Publishers Inc. 3rd edition, San Francisco, CA, USA, 2012. 20
- [21] D. J. Hand, P. Smyth, e H. Mannila. *Principles of Data Mining*. MIT Press, Cambridge, MA, USA, 2001. 21
- [22] M. B. C. M. Lobo. Panorama da evasão do ensino superior brasileiro: aspectos gerais das causas e soluções, 2011. Disponível em: http://www.institutolobo.org.br/imagens/pdf/artigos/art_087.pdf, acesso em 05 de Dezembro de 2015, 2011. 1, 9, 13

- [23] M. C. Morosini, A. O. Casartelli, A. C. B. Silva, B. S. Santos, R. E. Schmitt, e R. M. Gessinger. A evasão na educação superior no Brasil: uma análise da produção de conhecimento nos periódicos Qualis entre 2000-2011. Disponível em: http://www.alfaguia.org/www-alfa/images/ponencias/clabesI/ST_1_Abandono/12_MorosiniM_Abandono_ESBrasil.pdf, acesso em 05 de Dezembro de 2015, 2011. 6, 7
- [24] L. B. Palmeira e M. da P. Santos. Evasão no Bacharelado em Ciência da Computação da Universidade de Brasília: análise e Mineração de Dados. 2, 17
- [25] J. T. V. Pereira. Uma contribuição para o entendimento da evasão: um estudo de caso. *Revista Avaliação*, 2, 1996. 4
- [26] M. A. Ribeiro. O projeto profissional familiar como determinante da evasão universitária: um estudo preliminar. *Revista Brasileira de Orientação Profissional*, 6:55 – 70, 12 2005. 4, 5
- [27] D. I. Ristoff. *Universidade em Foco: reflexões sobre a educação superior*. Insular, 1999. 4
- [28] P. K. Santos e L. M. M. Giraffa. Evasão na educação superior: Um estudo sobre o censo da educação superior no brasil, 2013. Disponível em: http://www.alfaguia.org/www-alfa/images/ponencias/clabesIII/LT_1/ponencia_completa_200.pdf, acesso em 10 de Dezembro de 2015, 2013. 1, 4, 13
- [29] F. Sarker, T. Tiropanis, e H. C. Davis. [IEEE 2014 International Conference on Control, Decision and Information Technologies (CoDIT) - Metz, France (2014.11.3-2014.11.5)] - Linked Data, Data Mining and external open data for better prediction of at-risk students. 2014. 11
- [30] S. Shimosegawa, M. Miwa. Dropout from higher education and social stratification in Japan. 1:18, 2015. ix, 8
- [31] O. Souza, P. Morais, e F. Silva. Um estudo sobre a evasão no Curso de Licenciatura em Informática do IFRN – Campus Natal – Zona Norte. Disponível em: www.1bd.dcc.ufmg.br/colecoes/wei/2015/023.pdf, acesso em 25 de Dezembro de 2015, 2014. ix, 1, 16
- [32] L. M. V. Tigrinho. Evasão escolar nas instituições de ensino superior, 2008. Disponível em: http://www.institutolobo.org.br/images/pdf/artigos/art_087.pdf, acesso em 05 de Dezembro de 2015, 2011. 5, 54
- [33] Hans Vossensteyn, Andrea Kottmann, Ben Jongbloed, Frans Kaiser, Leon Cremonini, Bjorn Stensaker, Elisabeth Hovdhaugen, e Sabine Wollscheid. Dropout and completion in higher education in Europe: main report, 2015. Contract no EAC-2014-0182. 9
- [34] S. W. Yancy. *The retention of first year black male students at predominately white private and public universities and colleges*. Outskirts Press, 2015. 9