

UNIVERSIDADE FEDERAL DO PARANÁ

HERON HENRIQUE FONSACA  
MATHEUS ANTONIO SEVERO DOROCINSKI

**DEVIAS: MINERAÇÃO DE DADOS EDUCACIONAIS PARA O AUXÍLIO À  
TOMADA DE DECISÃO**

CURITIBA

2018

HERON HENRIQUE FONSACA  
MATHEUS ANTONIO SEVERO DOROCINSKI

**DEVIAS: MINERAÇÃO DE DADOS EDUCACIONAIS PARA O AUXÍLIO À  
TOMADA DE DECISÃO**

Trabalho apresentado como requisito parcial à obtenção de grau tecnológico em Análise e Desenvolvimento de Sistemas, Setor de Educação Profissional e Tecnológica da Universidade Federal do Paraná.

Orientadora: Prof.<sup>a</sup>. Dr.<sup>a</sup>. Rafaela Mantovani Fontana

Co-orientador: Prof. Dr. Alexander Robert Kutzke,

CURITIBA

2018

## RESUMO

Este trabalho é construído sobre os problemas das universidades brasileiras em relação aos seus índices de evasão e retenção. Sendo esses problemas amplamente encontrados nas mais diversas instituições de ensino no Brasil e causando várias consequências negativas à sociedade em geral, este projeto objetiva propor uma solução em formato de sistema para auxílio ao combate dos fenômenos da evasão e retenção escolar através da prevenção, a qual se dá na identificação antecipada dos possíveis casos. Para isso, o software foi desenvolvido de forma a tratar dados que são fornecidos pela Universidade Federal do Paraná (UFPR) relativos à performance dos alunos do curso de Tecnologia em Análise e Desenvolvimento de Sistemas (TADS) e gerar além de análises específicas que facilitam a compreensão do comportamento do aluno de TADS, possam prever quais alunos são mais propensos a desistir do curso ou concluir o curso apenas depois do tempo definido pela matriz curricular. As tecnologias implementadas neste projeto foram a linguagem de programação PHP, o SGBD PostgreSQL, a ferramenta de mineração de dados WEKA e a biblioteca Chart.js para criação de gráficos. O sistema apresentado também possui funcionalidades que ajudam professores e equipes que trabalham na universidade a gerenciar casos de jubramento e atendimento à alunos com problemas de desempenho. A esse sistema foi dado o nome de Devias.

## LISTA DE FIGURAS

FIGURA 1 – EXEMPLO DE UTILIZAÇÃO DO TRELLO .....	21
FIGURA 2 – ARQUITETURA: SISTEMA DEVIAS .....	32
FIGURA 3 – ARQUITETURA: PROCESSO DE MINERAÇÃO DE DADOS.....	33
FIGURA 4 – TELA DE ANÁLISE: GRÁFICO DA SITUAÇÃO GERAL DO CURSO .....	36
FIGURA 5 – TELA DE ANÁLISE: GRÁFICO DE CONCLUSÃO DO SEMESTRE.....	37
FIGURA 6 – TELA DE ANÁLISE: GRÁFICO DE RETENÇÃO.....	37
FIGURA 7 – TELA DE ANÁLISE: GRÁFICO DE PREDIÇÕES.....	38
FIGURA 8 – TELA DE ANÁLISE: GRÁFICO DE DESEMPENHO POR GRUPO DE DISCIPLINA .....	39
FIGURA 9 – TELA DE ANÁLISE: Listagem de área do gráfico.....	40
FIGURA 10 – TELA DE DADOS .....	41
FIGURA 11 – TELA DE ATENDIMENTOS.....	42
FIGURA 12 – TELA DE JUBILAMENTOS.....	42
FIGURA 13 – TELA DE USUÁRIOS .....	43
FIGURA 14 – TELA DE IMPORTAÇÃO DA PLANILHA.....	44
FIGURA 15 – TELA DE CONFIGURAÇÃO DE DISCIPLINAS.....	44
FIGURA 16 – TELA DE ANÁLISE DO ALUNO .....	45
FIGURA 17 – DIAGRAMA DE CASOS DE USO.....	52
FIGURA 18 – DIAGRAMA DE CLASSES .....	68
FIGURA 19 – DIAGRAMA DE SEQUÊNCIA: ATUALIZAR FONTE DE DADOS .....	69
FIGURA 20 – DIAGRAMA DE SEQUÊNCIA: VISUALIZAR ANÁLISES .....	70
FIGURA 21 – DIAGRAMA DE SEQUÊNCIA: VISUALIZAR ALUNOS .....	71
FIGURA 22 – DIAGRAMA DE SEQUÊNCIA: MANTER ATENDIMENTO.....	73
FIGURA 23 – DIAGRAMA DE SEQUÊNCIA: MANTER JUBILAMENTO.....	75
FIGURA 24 – DIAGRAMA DE SEQUÊNCIA: MANTER JUBILAMENTO.....	76
FIGURA 25 – DIAGRAMA DE SEQUÊNCIA: MANTER USUÁRIOS .....	77
FIGURA 26 – MODELO LÓGICO DOS DADOS.....	79

## LISTA DE TABELAS

TABELA 1 – DICIONÁRIO DE DADOS.....	18
TABELA 2 – TREINAMENTOS EVASÃO .....	26
TABELA 3 – TREINAMENTOS RETENÇÃO .....	27
TABELA 4 – ESPECIFICAÇÃO UC001: ATUALIZAR FONTE DE DADOS .....	53
TABELA 5 – ESPECIFICAÇÃO UC002: VISUALIZAR ANÁLISES .....	55
TABELA 6 – ESPECIFICAÇÃO UC003: VISUALIZAR ALUNOS.....	57
TABELA 7 – ESPECIFICAÇÃO UC004: MANTER ATENDIMENTO .....	59
TABELA 8 – ESPECIFICAÇÃO UC005: MANTER JUBILAMENTO .....	61
TABELA 9 – ESPECIFICAÇÃO UC006: VISUALIZAR ANÁLISE DO ALUNO .	64
TABELA 10 – ESPECIFICAÇÃO UC007: MANTER USUÁRIOS.....	66

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>6</b>
1.1	PROBLEMA.....	8
1.2	OBJETIVO GERAL.....	9
1.3.	OBJETIVOS ESPECÍFICOS.....	9
1.4	JUSTIFICATIVA.....	9
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>11</b>
2.1.	EVASÃO ESCOLAR.....	11
2.2.	RETENÇÃO ESCOLAR.....	12
2.3.	MINERAÇÃO DE DADOS .....	13
2.4.	APRENDIZADO DE MÁQUINA .....	15
2.5.	PLANILHA DE DADOS.....	17
2.6.	SOFTWARES SEMELHANTES .....	18
<b>3</b>	<b>MATERIAIS E MÉTODOS.....</b>	<b>20</b>
3.1.	SCRUM .....	20
3.2.	ADAPTAÇÕES DO MÉTODO À EQUIPE .....	21
3.3.	CRONOGRAMAS.....	22
3.2.1.	Sprint 1 .....	23
3.2.2.	Sprint 2 .....	23
3.2.3.	Sprint 3 .....	23
3.2.4.	Sprint 4 .....	23
3.2.5.	Sprint 5 .....	24
3.2.6.	Sprint 6 .....	24
3.2.7.	Sprint 7 .....	24
3.2.8.	Sprint 8 .....	24
3.2.9.	Sprint 9 .....	27
3.2.10.	Sprint 10 .....	28
3.2.10.	Sprint 11 .....	28
<b>4</b>	<b>APRESENTAÇÃO DO SISTEMA .....</b>	<b>31</b>
4.1.	ARQUITETURA DO SISTEMA.....	31
4.1.1	Processo de Mineração de Dados.....	33
4.2.	APRESENTAÇÃO DO DEVIAS.....	35
4.2.1.	Análises .....	35

4.2.2. Dados .....	40
4.2.3. Atendimentos.....	41
4.2.4. Usuários.....	43
4.2.5. Configurações.....	43
4.2.5. Análise do aluno .....	44
<b>5 CONSIDERAÇÕES FINAIS .....</b>	<b>46</b>
<b>REFERÊNCIAS.....</b>	<b>48</b>
<b>APÊNDICE A – HISTÓRIAS DE USUÁRIO .....</b>	<b>51</b>
<b>APÊNDICE B – DIAGRAMA DE CASOS DE USO .....</b>	<b>52</b>
1 – UC001: ATUALIZAR FONTE DE DADOS .....	53
2 – UC002: VISUALIZAR ANÁLISES .....	55
3 – UC003: VISUALIZAR ALUNOS.....	57
4 – UC004: MANTER ATENDIMENTO .....	59
5 – UC005: MANTER JUBILAMENTO .....	61
6 – UC006: VISUALIZAR ANÁLISE DO ALUNO.....	64
7 – UC007: MANTER USUÁRIOS.....	66
<b>APÊNDICE D – DIAGRAMA DE CLASSES.....</b>	<b>68</b>
<b>APÊNDICE E – DIAGRAMA DE SEQUÊNCIA.....</b>	<b>69</b>
1 – ATUALIZAR FONTE DE DADOS .....	69
2 – VISUALIZAR ANÁLISES .....	70
3 – VISUALIZAR ALUNOS.....	71
4 – MANTER ATENDIMENTO .....	72
5 – MANTER JUBILAMENTO .....	75
6 – VISUALIZAR ANÁLISE DO ALUNO.....	76
7 – MANTER USUÁRIOS.....	77
<b>APÊNDICE F – Modelo Lógico de Dados .....</b>	<b>79</b>

## 1 INTRODUÇÃO

O ensino superior é um grau de ciclo de educação posterior ao ensino secundário (ensino fundamental e médio no Brasil), que, se caracteriza pela oferta de cursos específicos em áreas do conhecimento normalmente realizado em faculdades e universidades.

A graduação no ensino superior, apresenta uma importante peça para o desenvolvimento de uma nação, pois além oferecer formações muito mais avançadas que as estâncias de ensino anteriores, é um ciclo de educação cada vez mais importante e valorizado pela sociedade. A educação, esta, que também tem papel fundamental no crescimento individual de uma pessoa. A respeito da educação, (GADOTTI, 2005, p. 1) afirma:

A educação é um dos requisitos fundamentais para que os indivíduos tenham acesso ao conjunto de bens e serviços disponíveis na sociedade. Ela é um direito de todo ser humano como condição necessária para ele usufruir de outros direitos constituídos numa sociedade democrática.

A graduação no Brasil é tida por grande parte da população como um caminho explícito para uma carreira satisfatória. Não é para menos, já que o mercado de trabalho também exige uma formação como um requisito mínimo para o exercício das principais profissões e daquelas profissões que podemos chamar de “emergentes”. Em 2016, houveram mais de 8 milhões de matrículas em cursos superiores presenciais e educação a distância (EAD), de acordo com a Sinopse Estatística da Educação Superior (INEP, 2017), número este, que nunca deixou de aumentar a cada estudo realizado.

Apesar do enorme número de inscritos nas instituições de ensino, uma parcela não tão expressiva conseguirá de fato alcançar a conclusão e término do ciclo de estudos até o ensino superior. Os motivos não são poucos, estudos têm apontado descontentamento com horários das disciplinas, falta de cursos noturnos, impossibilidade de trabalhar e estudar ao mesmo tempo, expectativas não-correspondidas e falta de informações sobre curso e profissão (LOTUFO, SOUZA JR, COVACIC & BRITO, 1998).

Segundo a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), 39% da população brasileira entre 25 e 34 anos não concluiu seus estudos até o ensino médio, (OCDE, 2015). Em muitos casos, o

indivíduo inicia o ciclo de educação, mas não termina o mesmo. Para compreender esses casos, é necessário analisar então a evasão e a retenção escolar.

A evasão escolar é um fenômeno definido como interrupção no ciclo de estudos (GAIOSO, 2005), geralmente em uma instituição de ensino, sendo um problema comum que afeta várias entidades daquele segmento pelo mundo, principalmente no Brasil.

Enquanto isso, a retenção escolar define-se como extensão do prazo esperado para conclusão dos estudos de um ciclo, ou seja, quando um aluno necessita de mais tempo que o previsto para término dos estudos, seja por reprovação ou afastamento. É importante ressaltar que a retenção de um aluno pode culminar com sua evasão de fato (CAMPELLO & LINS, 2008).

Tanto a evasão como a retenção de alunos são problemas que geram consequências negativas em vários âmbitos maiores: ocupação de vagas limitadas, gasto de recursos e até perda de verba governamental para as instituições de ensino.

Com o intuito de reduzir os problemas causados pela retenção e evasão acadêmica, a prevenção e diagnóstico precoce podem auxiliar os estabelecimentos interessados no combate à desistência e baixo desempenho, poupando assim esforços e recursos que seriam desperdiçados para reintegrar um discente, mantê-lo na instituição e/ou ajudá-los tardiamente (CAMPELLO & LINS, 2008).

Atualmente, em advento do uso de Ambientes Virtuais de Aprendizagem (AVAs) nas instituições de ensino, somado a outras tecnologias que auxiliam no processo de aprendizado e controle acadêmico, a todo momento é gerada uma enorme quantidade de dados educacionais. O tratamento, análise e extração de informações úteis destes dados caracteriza o que chamamos de Mineração de Dados (COSTA, BAKER, AMORIM, MAGALHÃES & MARINHO, 2013).

As aplicações existentes para a Mineração de Dados hoje, são diversas, e no âmbito educacional ganha até mesmo seu próprio nome: Educacional Data Mining (EDM). O título se deve ao fato dos dados educacionais não possuírem independência estatística, nem mesmo hierarquia entre os dados (COSTA, BAKER, AMORIM, MAGALHÃES, & MARINHO, 2013). Estas dificuldades

adicionais impactam diretamente na aplicação dos algoritmos utilizados na mineração de dados tradicional, exigindo desta forma, adaptações nos algoritmos existentes para a Mineração de Dados Educacionais (BAKER, 2010).

A Mineração de Dados, torna-se então, uma ferramenta interessante para contornar o problema da evasão e retenção nas instituições, dado que, é possível organizar os dados inerentes a ambos os casos e transformá-los em informações úteis, cabe ao analista escolher a melhor forma de evidenciá-las (BERSON, 2001; WITTEN & FRANK, 2000).

### 1.1 PROBLEMA

Considerando as grandes taxas de evasão e retenção presentes nos cursos de graduação, mantendo em mente a preocupação com os discentes obterem melhores resultados e não se desviarem do principal foco que é a diplomação, é possível evidenciar os impasses que motivaram à resolução deste trabalho.

1. O Governo Federal, através do Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais – REUNI (Governo Federal 2007) relata que “os índices de retenção e evasão de estudantes nos cursos de graduação atingem, em alguns casos, níveis alarmantes”.
2. Atualmente as instituições de ensino têm se mostrado extremamente eficientes em armazenar grandes quantidades de dados, obtidos de suas operações diárias, porém, a maioria ainda não usa adequadamente essa gigantesca quantidade de dados para transformá-la em conhecimentos que possam ser utilizados em suas próprias atividades (da Costa Côrtes, S., Porcaro, R. M., & Lifschitz, S. 2002).

Assim, propõe-se a criação de um software capaz de trazer à tona as situações de riscos dos alunos. Por meio da análise dos dados capturados, o sistema fará com que o professor esteja sempre ciente e apto a tomar uma providência que lhe é cabida.

## 1.2 OBJETIVO GERAL

O objetivo deste trabalho é desenvolver um *software* de mineração de dados que fornecerá ao corpo docente ou coordenação de um curso predições de evasão e retenção acadêmica, assim como indicadores de desempenho de seus alunos a partir dos dados fornecidos para o sistema, disponibilizando essas informações em formato de gráficos e relatórios.

## 1.3. OBJETIVOS ESPECÍFICOS

- Analisar a evasão, retenção e desempenho acadêmico, assim como suas causas e consequências;
- Implementar algoritmos de *machine learning* para a predição de desempenho acadêmico;
- Gerar relatórios com os resultados das predições realizadas;
- Produzir gráficos para interpretação dos dados dos alunos de maneira analítica
- Documentar o sistema

## 1.4 JUSTIFICATIVA

A aplicação dos modelos de Mineração de Dados se popularizaram tanto nos dias atuais porque eles provaram ser efetivos (QUONIAM, TARAPANOFF, ARAÚJO JÚNIOR & ALVARES, 2001). O estudo de caso “retenção e evasão em cursos superiores” se torna bastante promissor a partir do momento em que se imagina os resultados da aplicação de alguns modelos de mineração existentes, como por exemplo:

1. Agrupamento: identificar perfis de alunos, através de características academicamente semelhantes, obtendo assim, situações visíveis para um posicionamento.
2. Predição: Análise dos dados atuais para então realizar projeções, estas, que poderiam engatilhar ações preventivas por parte do corpo docente e reduzir os índices de retenção e evasão.
3. Associação: Descoberta de relações entre duas ou mais variáveis que impactam diretamente o aluno, e como a atenção dada para resolução de um único problema poderia acarretar na resolução de outros.

Estes modelos apresentados (DE AMO, 2004) evidenciam os benefícios da aplicação dos modelos de mineração, comprovando ainda mais sua eficácia.

A geração de conhecimento a partir de um processo de Mineração de Dados não possui um único caminho. Muitas vezes o retorno a atividades já concluídas se faz necessário, já que o descobrimento de uma nova regra pode exigir um novo dado, e talvez até uma reorganização dos já existentes, ou ainda, novas discussões podem levar a consideração de novas variáveis, e com isso aperfeiçoar ainda mais o modelo existente. (CAMPELLO & LINS, 2008).

A definição de um modelo robusto leva tempo, mas suas vantagens são inquestionáveis. Para o tomador de decisões que deixa de confiar apenas na sua experiência e em dados de fácil consulta, a utilização de uma ferramenta baseada em Mineração de Dados dá para suas escolhas uma maior segurança e um aumento considerável da sua taxa de acertos, tornando indispensável no seu âmbito profissional (DE AMO, 2004).

## 2 FUNDAMENTAÇÃO TEÓRICA

A evasão escolar, assim como a retenção, são conceitos que podem assumir diferentes definições, quase sempre com uma mesma origem, que acabam por corresponderem-se entre si. Este capítulo apresenta algumas das definições propostas na literatura, como parte da fundamentação do trabalho aqui apresentado.

A partir da compreensão da retenção e evasão escolar, e até suas consequências, se faz justificável a utilização de técnicas de mineração de dados com o objetivo de antecipar e prever o acontecimento dos fenômenos anteriormente descritos.

Para isso, será apresentado de maneira mais ampla o método pelo qual é possível extrair conhecimento e informação a partir dos dados concebidos das bases de dados da Universidade Federal do Paraná (UFPR), o data mining e suas especificidades, como forma de auxiliar a própria instituição na análise e previsão de desempenho dos alunos do curso de Tecnologia em Análise e Desenvolvimento de Sistemas (TADS).

Neste capítulo, serão abordados os temas da mineração de dados e principalmente o subtópico de Aprendizado de Máquina - a técnica que proporciona a predição de valores a partir de padrões identificados por algoritmos inteligentes – com o intuito de tornar explícito o que foi utilizado aqui como base ao desenvolvimento de um sistema de análise e predição de dados.

As informações fornecidas pela instituição de ensino da UFPR também são evidenciadas ao final do capítulo para a percepção do leitor sobre o que foi ou não manipulado por este projeto para sua construção. Um projeto de propósito semelhante a este também é descrito para efeito comparativo.

### 2.1. EVASÃO ESCOLAR

Como descrito por Favero (2006, p. 50), a evasão escolar é “o ato da desistência, incluindo os que nunca se apresentaram ou se manifestaram de alguma forma para os colegas e mediadores do curso, em qualquer momento”; e escrito por Abramovay & Castro (2003), refere-se ao aluno que deixa a escola, mas com a possibilidade de retorno à mesma.

Por efeito de estudo, a evasão escolar será caracterizada por este trabalho como a maioria dos autores a utilizaram e resumiram-na: um aluno que interrompeu seu ciclo de estudo na UFPR no curso de TADS e não se encontram registros no sistema que indicam sua volta ou reintegração.

A importância da evasão se dá por meio das consequências e impactos negativos causados à sociedade. O problema em questão desdobra-se e atinge inicialmente a universidade, a qual não obtém o retorno esperado dos seus investimentos, sendo que recursos materiais (espaço, equipamentos, estruturas) e humanos (funcionários, como professores) ficam ociosos e são desperdiçados – é importante ressaltar que na universidade pública o dinheiro governamental é então mal-empregado. Em seguida, logo se compreende as próximas consequências do problema em questão em âmbitos maiores, dada a importância da universidade, como parte fundamental para o desenvolvimento de uma nação, como retratado no artigo de (MACHADO, 2001, p. 333):

Na sociedade em que vivemos, o conhecimento transformou-se no principal fator de produção, no elemento fundamental para a produção de riquezas, explicitando-se com muita nitidez sua imediata vinculação com o universo do trabalho. De modo geral, a importância decisiva da Educação para uma justa "distribuição" desse "bem" tem sido reconhecida, e as Universidades, como centros de criação de conhecimento, desempenham, nesse cenário, um papel de destaque.

Dessa forma, é possível concluir que o fenômeno da evasão escolar gera uma série de adversidades à sociedade, das quais podem ser citadas, entre as principais, a redução de mão-de-obra especializada (profissionais capacitados ao mercado de trabalho) e assim, ao potencial de desenvolvimento direto de um país inteiro. A evasão escolar destaca-se no Brasil, frente a muitos outros países, e possui números alarmantes. Segundo pesquisa, entre 2000 e 2005, a evasão média nas universidades foi de 22% e atingiu 12% nas públicas e 26% nas particulares (SILVA FILHO, 2007).

## 2.2. RETENÇÃO ESCOLAR

A retenção, também reconhecida como permanência prolongada, trata-se do caso em que o estudante, para concluir seu curso requer um tempo superior ao previsto pela matriz curricular (SEVERINO PEREIRA et al., 2015).

Quando comparada à evasão escolar, a retenção apresenta consequências e efeitos muito semelhantes. Em ambos os casos o aluno que retém ou evade de seu curso, não está em conformidade com o plano base proposta à sua formação.

Assim como descrito na evasão, a retenção do aluno também gera prejuízos à sociedade e ao próprio aluno, pois há atraso na formação de profissionais qualificados e retorno de conhecimento para aquela; assim como demora para término de um importante ciclo ao estudante que é de grande importância à sua carreira.

A retenção acadêmica é um problema comum nas universidades brasileiras e afeta inclusive, um número muito maior de estudantes em relação à evasão. Segundo estudo de (SEVERINO PEREIRA et al, 2015), realizado sobre a Universidade Federal do Espírito Santo (UFES), 52,2% dos estudantes que deveriam se formar nos semestres de 2007/2 a 2012/1 ficaram retidos.

O fenômeno da retenção tem grande presença também no curso analisado por este trabalho, na UFPR o curso de TADS contabilizou 407 alunos que se formaram com retenção num total de 568 alunos formados entre 2005 e 2017. Logo, 71% desses formandos, precisaram de mais tempo além dos 6 períodos, propostos pela matriz curricular, para concluírem o curso de TADS na UFPR.

Tendo em vista os dois problemas que afetam comumente as universidades brasileiras, da evasão e retenção acadêmica, este trabalho propõe o uso das técnicas de mineração de dados como forma de prevenção aos fenômenos descritos.

Por meio da mineração de dados, ao compreender os padrões que descrevem a evasão e retenção dos discentes, seria possível antecipar o acontecimento desses problemas e tomar medidas para contorná-los.

### 2.3. MINERAÇÃO DE DADOS

A mineração de dados ou *data mining*, é reconhecida como parte do processo de Knowledge Discovery in Databases (KDD), esse definido por (FRAWLEY et al, 1992, p. 58) como “a extração não trivial de informação potencialmente útil, anteriormente desconhecida e implícita”. Esse processo,

crescentemente estudado e aplicado sobre fontes de dados, permite a compreensão de novas informações que não são encontradas de maneira tradicional.

Nesse sentido, essas novas informações obtidas podem também significar a revelação de relações entre os dados até então não identificadas, que podem produzir novos conhecimentos e descobertas científicas. Em muitos casos, o processo de mineração de dados pode ser de grande valor e permitir várias utilidades, como por exemplo o auxílio à tomada de decisão. E, dada sua natureza, a descoberta de informação cabe nas mais variadas áreas do conhecimento, tal como vendas, bioinformática e ações contraterrorismo (Baker et al., 2011).

O conceito de mineração de dados é amplo e pode agregar também diversos domínios de conhecimento em seus processos, como descrito por (DA COSTA CÔRTEZ, 2002, p. 2):

Frequentemente, mineração de dados tem sido considerada e classificada como uma mistura de pesquisas em estatística, inteligência artificial e bancos de dados. Até recentemente, não era reconhecido como um campo de interesse para os estatísticos, sendo mesmo considerado, nesta área, como uma área de pesquisa 'pouco relevante'. Devido à sua importância prática, entretanto, o campo tem emergido como uma área de crescimento acentuado e de elevada importância, destacando-se pelo surgimento de diversos congressos científicos e produtos comerciais.

A mineração de dados pode ser utilizada para inúmeros objetivos, e neste trabalho destaca-se a função dessa para auxílio na compreensão de dados educacionais. Como foco de crescente interesse, a educação logo passou a ser vista como um subdomínio da mineração de dados, conhecida como Educational Data Mining (EDM), sendo essa dedicada à análise de dados de origem educacional, como ambientes exclusivos ao tema.

Segundo a Educational Data Mining Organization, grupo que incentiva projetos sobre o assunto e até organiza eventos para discussão, a EDM é uma disciplina emergente dedicada a exploração dos singulares e crescentes dados produzidos por meios educacionais.

A EDM busca descobrir informações relevantes compreendendo dados sob a perspectiva educacional, tornando assim seus métodos mais eficientes e

adaptados ao estudo dos alunos, às formas de ensino e aprendizagem e também às suas respectivas técnicas. Exemplos das utilizações das competências da EDM podem incluir: análise de qual abordagem instrucional (individual ou coletiva) é mais adequada ao aluno, verificar se um aluno está desmotivado ou confuso, e até a busca de alternativas para a melhoria do desempenho do estudante (BAKER et al, 2011).

#### 2.4. APRENDIZADO DE MÁQUINA

*Machine Learning* ou Aprendizado de máquina (AM), é um dos campos da Inteligência Artificial (AI) que tem por enfoque transcrever dados obtidos de análises prévias ou puro armazenamento, em conhecimento computacional (HAN, J., KAMBER, M., & PEI, J., 2006). Baseados em experiências passadas, ou melhor, dados que representam instâncias antigas de um banco, os algoritmos de AM são capazes de reconhecer padrões complexos e tomar decisões inteligentes, ou ainda, gerar novas variáveis de interesse.

Para tanto, o aprendizado é dado através do treinamento da máquina, este, que não se difere tanto de um conhecimento obtido por indução. Indução é um recurso utilizado para derivar conhecimento novo, a sua utilização deve ser feita com cautela, pois se o número de exemplos for insuficiente, ou se os exemplos não forem bem aplicados, os resultados obtidos podem ser de pouco valor (MONARD & BARANAUSKAS, 2003).

O aprendizado indutivo pode ser feito de duas formas principais, sendo o primeiro – supervisionado - através do fornecimento ao processo de um conjunto de exemplos de treinamento que já possuem um rótulo ou classe definido. Isso é, o algoritmo de AM recebe uma lista de exemplos contendo valores de características ou atributos, associados a um resultado (classe). O objetivo do algoritmo é então classificar por si só novos exemplos que não receberam rótulos. Há ainda algoritmos que não recebem rótulos – os não-supervisionados – que geram seus próprios agrupamentos pelo processo (MONARD & BARANAUSKAS, 2003).

A partir do princípio do aprendizado indutivo, é possível aplicar sobre esse o problema de previsão da evasão ou retenção de um aluno usando como base

seus dados históricos. Para isso, uma das alternativas é utilizar algoritmos já desenvolvidos, testados e amplamente utilizados pela comunidade.

Com o objetivo de mensurar a qualidade dos testes realizados em algoritmos de aprendizado de máquina, é preciso também adotar medidas que apontem se é possível ou não confiar nas previsões dos mesmos. Para isso, podem ser utilizados a acurácia, definida como chance do algoritmo acertar sua previsão, tal como precisão; e a área da curva ROC. A curva *Receiver Operating Characteristic* (ROC) é uma forma de visualização da performance de classificadores através de indicadores (BRADLEY, 1997). Um método de análise da curva ROC de maneira simples para comparação do desempenho de classificadores é a utilização da área sob a curva ROC.

#### 2.4.2. Bayes

Os algoritmos de Bayes (Naive Bayes, por exemplo) baseiam-se no teorema de Bayes, e são simples classificadores probabilísticos que calculam um conjunto de probabilidades sobre a frequência e combinações de valores dentro de um dataset (PATIL & SHEREKAR, 2013).

#### 2.4.3. Redes Neurais Artificiais

As redes neurais artificiais (RNA) são algoritmos que tem como princípio o funcionamento dos neurônios biológicos. Os neurônios simulados, na RNA, são elementos processadores interligados que atuam paralelamente para concluir uma tarefa. Esses modelos apresentam técnicas estatísticas não-lineares que podem resolver muitos problemas diferentes de alta complexidade e são utilizados em casos onde não é possível definir claramente as regras de solução (VELLASCO, 2007).

#### 2.4.4. Árvore de Decisão

Os algoritmos de árvore de decisão (Decision Tree) criam árvores binárias para solução dos problemas. Esses algoritmos partem do princípio desenvolvido por J.R. QUINLAN, em 1986. Nesses algoritmos, uma árvore é gerada para modelar o processo de classificação e para cada tupla listada e seus resultados é aplicado sua lógica (PATIL & SHEREKAR, 2013).

## 2.5. PLANILHA DE DADOS

Para o desenvolvimento do sistema, foi utilizado uma planilha de dados dos alunos que passaram pelo curso de Tecnologia em Análise e Desenvolvimento de Sistemas da Universidade Federal do Paraná, com seu ingresso datado do primeiro semestre de 2009 (ano da estruturação anterior do curso) até o período mais recente. Estes registros foram retirados diretamente da base de dados da universidade (sistema SIE) e contam, basicamente, com o histórico escolar de cada aluno, demonstrando seu desempenho acadêmico dentro de toda disciplina cursada, por meio de sua frequência e notas desempenhadas.

A planilha de dados contém além da relação do desempenho do aluno em cada disciplina já cursada, outras informações sobre o aluno, como quando o mesmo ingressou no curso, ou quando e como ele deixou o curso (caso tenha se formado ou desistido, por exemplo); o arquivo fornecido também conta com informações da matéria cursada, como sua carga horária, e do curso, como o ano referência da versão do curso.

Foi vetada a disponibilização de dados sensíveis que fossem capazes de caracterizar o discente, isto, para de maneira alguma violar quaisquer direitos de informação sem o consentimento do mesmo.

Apesar da autenticidade da fonte, algumas informações não atingiam sua totalidade, isto é, necessitavam de seu entendimento antes de serem trabalhados. O resultado foi a elaboração de um pequeno dicionário de dados para melhor compreensão de casos como: códigos sem uma descrição apropriada, e intervalos não definidos. A Tabela 1 ilustra um resumo dos atributos mais importantes do artefato gerado.

TABELA 1 – DICIONÁRIO DE DADOS

<b>Coluna</b>	<b>Descrição</b>
COD_CURSO	Código do curso
ID_ALUNO	Número identificador único do aluno
FORMA_EVASAO_ITEM	Número indicador da forma de evasão do aluno (formou-se, desistiu, etc.)
ANO_EVASAO	Ano da evasão
PERIODO_EVA_ITEM	Período da evasão
NUM_VERSAO	Ano da versão do curso em vigência
COD_ATIV_CURRIC	Código da disciplina
NOME_ATIV_CURRIC	Nome da disciplina
ANO	Ano que o aluno cursou a disciplina
PERIODO_ITEM	Período que o aluno cursou a disciplina
MEDIA_FINAL	Média final do aluno na disciplina
NUM_FALTAS	Número de faltas do aluno na disciplina

Fonte: Os Autores (2018).

## 2.6. SOFTWARES SEMELHANTES

Pesquisando softwares destinados à análise de dados acadêmicos com o auxílio de *data mining* foi identificado o software *TADA-Ed for Educational Data Mining*, o qual atua como ferramenta facilitadora para a mineração de dados com foco educacional.

O TADA-Ed é uma plataforma desenvolvida em 2005 por Agathe Merceron e Kalina Yacef, dedicada a professores com o propósito de permitir a mineração de dados, os quais são produzidos através da resolução de exercícios online por seus estudantes em uma plataforma própria não especificada. (MERCERON & YACEFE, 2005)

O software, que tem como objetivo facilitar a mineração de dados educacionais, apresenta uma série de funcionalidades que permitem consumo, visualização, filtro e transformação dos dados originários de um banco de dados relacional. (MERCERON & YACEFE, 2005)

TADA-Ed é uma plataforma desenvolvida de forma genérica para a aplicação de técnicas de *data mining* em qualquer base de dados relacional, e preza pelo grande número de possibilidades fornecidas ao usuário para descobrir novas informações escondidas em seus dados. (MERCERON & YACEFE, 2005)

Em contraste ao sistema Devias, aquele apresenta uma interface muito mais complexa e dá ao usuário recursos genéricos para que o mesmo prepare seus dados, transforme-os da maneira adequada, aplique algoritmos ou correlacione os dados como desejar; enquanto que o software desenvolvido neste projeto é moldado para estruturas pré-definidas de entrada de dados, já acompanha todo o processo de transformação e pré-processamento dos dados sem necessidade de interação do usuário e fornece a esse análises já prontas para propostas de escopo muito mais definidas.

Neste capítulo foram apresentados os principais conceitos que circundam o projeto Devias, e em quais bases teóricas foi desenvolvido esse sistema. No próximo capítulo serão apresentados os materiais e métodos utilizados para a elaboração e desenvolvimento deste projeto.

### 3 MATERIAIS E MÉTODOS

Neste capítulo serão apresentados os métodos e tecnologias utilizados para o desenvolvimento do software Devias, desde a sua organização, definindo as fases de desenvolvimento; até a sua concepção, expondo as ferramentas empregadas.

#### 3.1. SCRUM

A metodologia ágil para desenvolvimento de software é um formato de trabalho para construção de aplicações que focam na simplicidade e nas iterações como principal meio para entrega de softwares com qualidade de forma rápida e flexível, tendo em sua essência a comunicação e a habilidade de fornecer valor ao mercado em curtos períodos de tempo (MARÇAL et al, 2007).

O manifesto ágil, documento escrito por especialistas em processos de desenvolvimento de software e largamente reconhecido pela comunidade, define os quatro valores do desenvolvimento ágil: indivíduos e iterações são mais importantes do que processos e ferramentas; software funcionando é mais importante do que documentação completa; colaboração com o cliente é mais importante do que negociação de contratos; adaptação a mudanças é mais importante do que seguir o plano inicial (BASSI FILHO, 2008).

Muitos métodos ágeis surgiram ao longo do tempo, seguindo então os princípios descritos pelo manifesto ágil, um deles foi o *Scrum*, método ágil que foi adaptado ao desenvolvimento do software, e na verdade tem foco na interação dos membros da equipe e na flexibilidade de alteração das variáveis do projeto (requisito, prazo, recurso ou tecnologia) (FADEL et al, 2010).

O *Scrum* pressupõe o uso de alguns artefatos, sendo esses o *Product Backlog*, o qual contém as funcionalidades esperadas ao sistema; além do *Sprint Backlog*, documento com a divisão do *Product Backlog* em porções menores para seu desenvolvimento de fato em um determinado prazo, o chamado *Sprint*. O prazo definido para uma *Sprint* geralmente está entre duas a quatro semanas, e ao seu final, o objeto gerado deverá ser um produto com funcionalidades entregáveis, de uso para o cliente (SILVA et al, 2009).

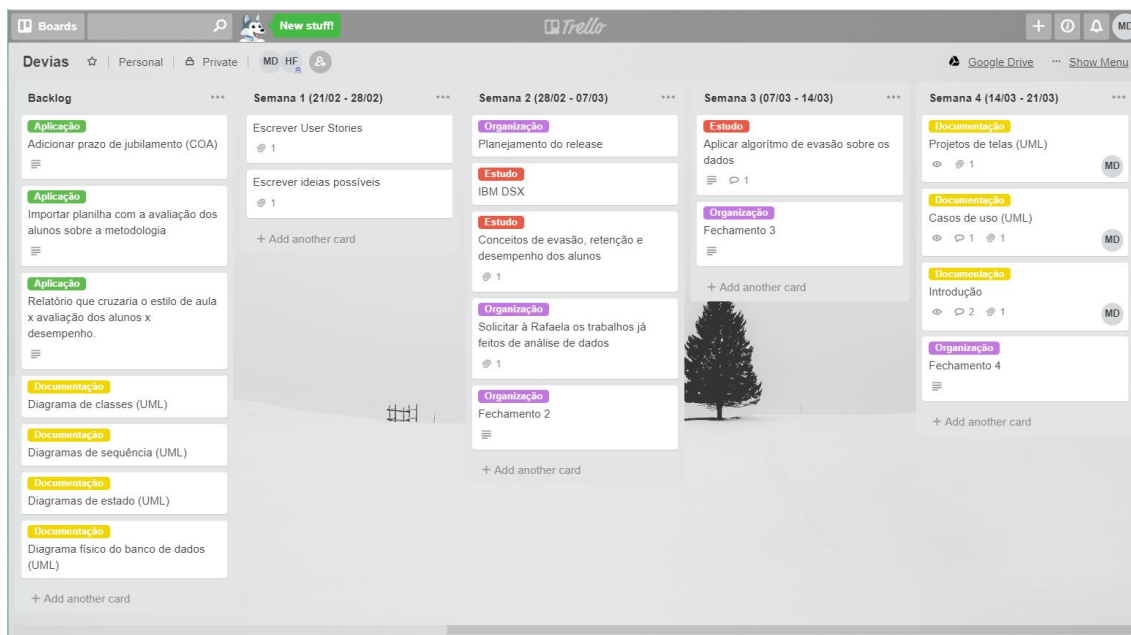
O andamento de cada *Sprint* deve ser reportado por todos os membros da equipe participante durante as reuniões diárias, que tem como principal objetivo identificar progressos e impedimentos dentro do desenvolvimento. Ao fim de cada *Sprint*, além da entrega do produto, é realizada uma nova reunião com os intuits de retrospectiva e definição do próximo *Sprint* (SILVA et al, 2009).

### 3.2. ADAPTAÇÕES DO MÉTODO À EQUIPE

As reuniões diárias pressupostas pelo método *Scrum* não foram realizadas, em substituição reuniões presenciais mais longas e detalhadas ocorriam semanalmente ou quinzenalmente dependendo do momento do projeto.

A organização das sprints e das atividades necessárias ao desenvolvimento do projeto ocorreu via Trello, ferramenta estilo Kanban na qual foram elaborados os *cards* com tarefas e entregas que ocorreram pela equipe organizados por *sprint*. O exemplo da utilização é demonstrado na Figura 1.

FIGURA 1 – EXEMPLO DE UTILIZAÇÃO DO TRELLO



Fonte: Os Autores (2018).

### 3.3. CRONOGRAMAS

O desenvolvimento do projeto se deu através de *sprints* e ao decorrer do mesmo definiam-se novas tarefas e como as próximas fases do projeto deveriam acontecer, de forma iterativa.

A seguir serão apresentadas em detalhes as *sprints* do projeto Devias.

TABELA 2 – CALENDÁRIO DE SPRINTS

<b>Sprint</b>	<b>Data</b>	<b>Resumo</b>
Sprint 1	21/02 – 28/02	<i>User stories;</i> <i>Brainstorm.</i>
Sprint 2	28/02 – 07/03	Testes com IBM DSX; Análise dos conceitos de evasão e retenção.
Sprint 3	07/03 – 14/03	Concepção da planilha de dados; Teste com o Weka.
Sprint 4	14/03 – 21/03	Casos de uso; <i>Mockups</i> das telas;
Sprint 5	21/03 – 28/03	Definição da linguagem de programação; Testes com Nodejs; Testes com PHP.
Sprint 6	28/03 – 04/04	Definição de todas as tecnologias utilizadas; Início da implementação.
Sprint 7	04/04 – 18/04	Testes com diferentes algoritmos de <i>machine learning</i> ; Processo de enriquecimento dos dados.
Sprint 8	18/04 – 02/05	Geração do modelo final de predição
Sprint 9	02/05 – 16/05	Desenvolvimento das telas de CRUD do sistema
Sprint 10	16/05 – 30/05	Geração dos gráficos de análise; Funcionalidades para listagem de alunos presentes nos gráficos.
Sprint 11	30/05 – 13/06	Tela de Análise do aluno; Listagem geral de alunos.

### 3.2.1. Sprint 1

A primeira *sprint* do projeto durou uma semana, e nela foram levantados os esboços de *user stories*, os quais serviriam como base para uma nova discussão entre o grupo para então desenvolvimento de fato.

Nesta fase também foi discutido possíveis ideias para o sistema, com o objetivo de listar o maior número de possibilidades que seriam analisadas em questão de viabilidade e prioridade posteriormente.

### 3.2.2. Sprint 2

Durante a segunda *sprint* do projeto, foi realizado um estudo para possível implementação da ferramenta IBM DSX no projeto, como meio para aplicação de algoritmos de Machine Learning, o qual não se mostrou tão eficiente quanto a equipe esperava e foi então substituído por tecnologias mais adequadas.

O início do intensivo estudo dos conceitos de retenção e evasão escolar também fizeram parte desta *sprint*.

### 3.2.3. Sprint 3

Na terceira *sprint* foram realizadas as primeiras análises aos dados da UFPR, e conseqüentemente o desenvolvimento das lógicas iniciais para obtenção das informações mais importantes do arquivo. Nessa fase já se iniciaram os processos de teste na ferramenta Weka para treinamento de algoritmos de aprendizado de máquina.

### 3.2.4. Sprint 4

A quarta *sprint* do projeto concebeu-se finalmente a introdução do projeto Devias em sua forma final, além dos mockups de tela que auxiliaram no esclarecimento das ideias do grupo em relação ao formato e organização visual do projeto.

Nesta fase também se deu a criação dos casos de uso possíveis ao sistema.

### 3.2.5. Sprint 5

Durante a quinta *sprint* a equipe voltou-se ao desenvolvimento do software Devias, e fez um levantamento para escolher qual seria a tecnologia utilizada para estruturar o back-end da aplicação, uma prova de conceito foi realizado com a linguagem NodeJS para a função, entretanto a mesma não atendeu a alguns requisitos levantados pela equipe.

### 3.2.6. Sprint 6

A sexta *sprint* do projeto foi de suma importância à seleção de tecnologias para o desenvolvimento do software, além de especificar como se dariam as visualizações sobre os dados apresentados e planejar como ocorreria todo o processo de mineração de dados a partir da entrada dos dados referentes à performance dos alunos de TADS da UFPR.

A seguir será descrito quais foram as ferramentas e tecnologias definidas finalmente nesta *sprint*.

### 3.2.7. Sprint 7

A sétima *sprint* do projeto foi o período dedicado a organização do desenvolvimento do software e estudo dos processos de aprendizado de máquina.

Nessa *sprint* estudou-se também como alterar o arquivo inicial fornecido pela UFPR para interpretação dos dados relativos à performance dos alunos de TADS, uma vez que o arquivo antigo não possuía todas as informações necessárias ao processo de treinamento dos algoritmos de aprendizado de máquina indutivo.

### 3.2.8. Sprint 8

A oitava *sprint* (18/04 - 02/05) caracterizou pela aplicação dos algoritmos de aprendizado de máquina sobre a base de dados do software Devias, a qual já disponibilizava, agora, dados enriquecidos e detalhadamente preparados ao processo de treinamento. A estruturação do código da aplicação para comportar esse processo de *machine learning* também já estava em implementação paralelamente.

A seguir constam os resultados obtidos através do treinamento de algoritmos na ferramenta WEKA.

### 3.2.8.1. Aplicação de algoritmos de aprendizado de máquina

Para realizar o treinamento dos algoritmos de *machine learning*, se faz necessário um processo de transformação dos dados até a obtenção de uma estrutura ideal para consumo da ferramenta WEKA.

Ao longo deste projeto, muitos algoritmos e diversas organizações de dados foram utilizadas para testar principalmente a variável de acurácia do treinamento. Enfim, foi concluído que a estrutura de dados mais adequada para aplicação de *machine learning* foi a das notas e faltas dos alunos de TADS da UFPR em seu primeiro período de curso. O aumento da quantidade de informação relativos ao desempenho do aluno não geraram aumento relevante nos níveis de acurácia dos modelos.

A estrutura do modelo utilizado pelo projeto é caracterizada por dois atributos para cada disciplina do primeiro período do curso de TADS de 2009 a 2017, e ao final desses atributos, uma classificação definida anteriormente, quando se quer treinar o algoritmo; ou uma classificação esperada, quando o algoritmo prediz o valor.

A tabela gerada para cada tipo (evasão ou retenção) pelo sistema é semelhante à estrutura apresentada na TABELA 1, na qual o código da disciplina mais “n” representa a nota final do aluno e “f” representa o número de faltas e “c” o classificador:

TABELA 2 – MODELO TABELA CLASSIFICAÇÃO

Coluna	Descrição
138n	Média final do aluno na disciplina Algoritmos e Lógica de Programação I
138f	Faltas do aluno na disciplina Algoritmos e Lógica de Programação I
152n	Média final do aluno na disciplina Técnicas de Comunicação

152f	Faltas do aluno na disciplina Técnicas de Comunicação
100n	Média final do aluno na disciplina Matemática Aplicada
100f	Faltas do aluno na disciplina Matemática Aplicada
165n	Média final do aluno na disciplina Noções de Hardware
165f	Faltas do aluno na disciplina Noções de Hardware
105n	Média final do aluno na disciplina Tecnologias de Informação
105f	Faltas do aluno na disciplina Tecnologias de Informação
137n	Média final do aluno na disciplina Fundamentos de Informática
137f	Número de faltas do aluno na disciplina Fundamentos de Informática
104n	Média final do aluno na disciplina Sistemas Operacionais
104f	Número de faltas do aluno na disciplina Sistemas Operacionais

Fonte: Os Autores (2018).

Após os treinamentos realizados pelos principais algoritmos do WEKA para registros que deveriam ser classificados como retentores e não retentores ou evasores e não evasores, obteve-se os seguintes resultados de acurácia e curva ROC, o registro em negrito representa o melhor em desempenho e consequentemente utilizado pela equipe neste projeto:

TABELA 2 – TREINAMENTOS EVASÃO

Algoritmo	Acurácia	Área curva ROC
<b>BayesNet</b>	<b>75.18</b>	<b>0.820</b>
NaiveBayes	72.65	0.823
Logistic	74.14	0.811

MultilayerPerceptron	72.51	0.765
SGD	74.14	0.727
SimpleLogistic	74.14	0.813
SMO	74.59	0.730
J48	69.39	0.692
RandomForest	74.88	0.812
RandomTree	67.31	0.673

Legenda: Treinamento de algoritmos para evasão de alunos, em destaque o algoritmo com melhor desempenho. Fonte: Os Autores (2018).

TABELA 3 – TREINAMENTOS RETENÇÃO

Algoritmo	Acurácia	Área curva ROC
BayesNet	83.59	0.724
NaiveBayes	45.85	0.692
Logistic	89.00	0.713
MultilayerPerceptron	87.81	0.730
SGD	89.00	0.500
SimpleLogistic	88.83	0.688
SMO	89.00	0.500
J48	87.47	0.549
<b>RandomForest</b>	<b>89.00</b>	<b>0.731</b>
RandomTree	80.54	0.527

Legenda: Treinamento de algoritmos para retenção de alunos em destaque o algoritmo com melhor desempenho. Fonte: Os Autores (2018).

### 3.2.9. Sprint 9

A nona *sprint* foi utilizada para o desenvolvimento das telas de parametrizações e CRUD's do sistema.

Pensando em como seria a interação do usuário com o sistema, optou-se por utilizar a mínima quantidade de telas possível, com o desafio de dispor o conteúdo de maneira enxuta, sem que seja passada a impressão de que há muitos componentes na tela.

A solução se deu através da criação de status para a interface em uso. O clique do botão “Novo” passa a tela para o status de inclusão, onde os inputs da tela tornam-se digitáveis. O clique do botão Editar, só está disponível após a seleção de um item na tabela de registros, após selecionado um registro e clicado o botão “Editar”, os campos se tornam editáveis com as informações do registro em questão.

Foram desenvolvidas quatro telas do sistema que seguem este padrão de consulta, inclusão, alteração e exclusão durante o período desta *sprint*. Suas implementações foram feitas de forma rápida e concisa, devido ao desenvolvimento de maneira dinâmica, onde alterando a fonte de dados e os campos da página, o código pode ser reaproveitado inúmeras vezes.

#### 3.2.10. Sprint 10

O resultado desta décima *sprint* é descrito pelas implementações das análises presentes no sistema. O grupo dedicou seus esforços a elaborar os gráficos conforme as histórias de usuário.

Neste momento foi explorada a biblioteca javascript Chart.js. O fluxo proposto para a geração da análise teve que ser adaptado ao seu uso, pois, forçando a passagem de um Json para a formação do gráfico da análise, as classes passaram por uma alteração do seu retorno.

Em primeira instância foram definidas as formas de geração de cada gráfico, isto é, seu modelo. A forma de demonstrar os dados é de suma importância para a interpretação dos dados, para isso foram testadas ao menos duas formas para cada análise.

Ao final da *sprint* a página de análises estava pronta e operacional, gerida por uma lista de análises, o clique sobre um dos itens gerava a requisição para o servidor, solicitando a construção do select na base de dados, retornando um json para a camada de visualização, que por sua vez, montava o gráfico no lado do cliente.

#### 3.2.10. Sprint 11

A última *sprint* do desenvolvimento do sistema Devias foi caracterizada pelos retoques finais nas telas já concluídas e desenvolvimento da tela de

análise do aluno. A tela representa grande importância para a aplicação, pois, seria a responsável pelo relatório individual de cada aluno, contendo seu histórico, as disciplinas que restam para a conclusão e a definição das características do aluno, como:

- Retentor;
- Evasor;
- Graduado;
- Riscos de retenção;
- Riscos de evasão;
- Concluirá em tempo hábil.

A complexidade da tela exigiu a delegação de atividades entre os membros da equipe para entrega no prazo estipulado.

### 3.4. TECNOLOGIAS UTILIZADAS

Para escolha das tecnologias para desenvolvimento do software Devias levou-se em consideração a conformidade entre os objetivos e pontos fortes das tecnologias em relação às necessidades da aplicação.

A linguagem de *back-end* responsável por articular as páginas dinâmicas da aplicação é o PHP, escolhido devido a facilidade de integração com páginas web uma vez que a linguagem é desenvolvida com esse propósito. Dentro das características que auxiliam na escolha do PHP pode-se citar: fácil uso em conjunto com HTML, gratuidade, código aberto, entre outros (NIDERAUER, 2004).

Sobre a linguagem PHP, também foi utilizado o *framework* de desenvolvimento Slim, o qual possibilita o uso de *routes*, url amigáveis, consumo de formulários entre outros de maneira mais fácil em relação à pura implementação do PHP.

O sistema gerenciador de banco de dados (SGBD) escolhido para consumo do software foi o PostgreSQL, o qual é gratuito e permite o uso de várias funções avançadas aos SGBDs comuns. Entre as vantagens desse sistema, segundo (GUTTOSKI, 2006, p. 21):

O PostgreSQL é um sistema gerenciador de banco de dados objeto-relacional, desenvolvido na University of Califórnia, Berkeley Computer Science Department, com versão estável desde 1996. Tem mais de

15 anos de desenvolvimento ativo e uma arquitetura com forte reputação devido à sua confiabilidade, integridade de dados e exatidão.

Como ferramenta para a aplicação de algoritmos de aprendizado de máquina, o WEKA foi a escolha realizada pela equipe uma vez que essa é de código aberto, possui mais de vinte algoritmos diferentes de mineração de dados. (DE MELLO et al, 2007)

O WEKA mostrou-se também mais adequado ao projeto Devias uma vez que permitiu fácil integração dos processos de *machine learning* ao resto do sistema.

As visualizações de gráficos dentro da aplicação foram desenvolvidas por meio da biblioteca Chart.js, o qual faz uso de HTML5, CSS e JavaScript para construir gráficos para sistemas web. A biblioteca é gratuita e de código aberto.

Outra biblioteca vastamente aplicada ao Devias foi jQuery, pelo fato dessa permitir uma fácil integração entre requisições JavaScript e facilitar o desenvolvimento de aplicações Web através da redução de linhas de código e maximização da produtividade (TANAKA, 2011).

Para melhorar a identidade visual do sistema web foi utilizado também um *template* pré-desenvolvido e disponibilizado gratuitamente na plataforma GitHub, o Carbon *Bootstrap 4 Admin Template*, desenvolvido por Mohammed Isa <<https://github.com/mohd-isa>>.

Este capítulo apresentou os materiais e métodos utilizado no desenvolvimento do projeto. O próximo capítulo apresenta o sistema, definindo sua arquitetura e demonstrando as telas de sua interface.

## 4 APRESENTAÇÃO DO SISTEMA

Neste capítulo são explicadas a arquitetura em que se baseia o sistema juntamente com as funcionalidades presentes em cada tela de maneira detalhada.

### 4.1. ARQUITETURA DO SISTEMA

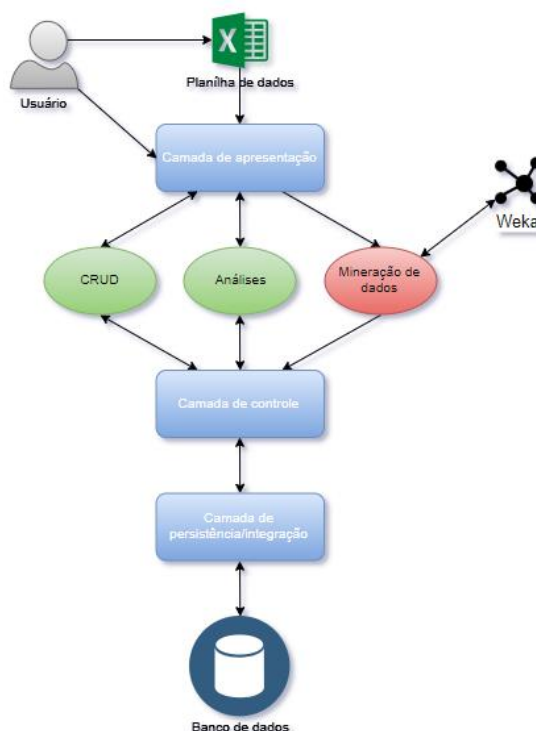
O sistema Devias estruturou-se a partir de uma arquitetura tipo MVC (Modelo, Visão e Controle) que tem por objetivo fragmentar o processo de desenvolvimento de softwares, buscando um maior rendimento e escalabilidade do sistema. A fragmentação dos componentes tem como finalidade inicial a separação da interface do usuário da lógica da aplicação (SILVA, 2012). Aprofundando ainda mais o conceito, conforme apontado por Silva (2012, p. 2) em descrição das funções de cada camada:

[...] as camadas Modelo, Visão e Controle exercem esta divisão de funcionalidades ao desenvolver e executar um *software*. No padrão MVC, o Modelo trabalha na manipulação dos dados internos de uma aplicação, e se comunica especialmente com o armazenamento de dados. A camada de Visão ou apresentação trabalha na interface do usuário, capturando as suas ações e enviando ao Controlador, acessa os dados do Modelo através do Controlador e aplica a apresentação desses dados conforme o evento. Por fim, a camada de Controle exerce funcionalidades que envolvem o comportamento da aplicação; controla os fluxos entre as camadas de Visão e Modelo, e gera a resposta ao usuário.

A Figura 2 retrata a arquitetura do sistema e como ela se comunica com as funcionalidades do sistema.

FIGURA 2 – ARQUITETURA: SISTEMA DEVIAS

Arquitetura sistema Devias



Fonte: Os Autores (2018).

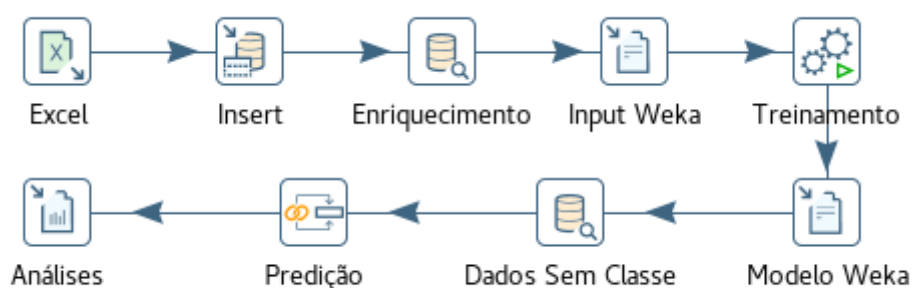
Os retângulos levemente azulados representam as camadas presentes em um modelo MVC, com suas devidas restrições. Os balões verdes representam operações básicas de maior impacto do usuário, onde envolvem a comunicação de todas as camadas para a conclusão da atividade, como por exemplo, as operações de consulta, inclusão, alteração e exclusão, ou ainda, as chamadas realizadas às análises. O balão em vermelho em destaque é o momento da comunicação com o software externo Weka, onde são realizadas as operações de data mining e machine learning, para o desenvolvimento do modelo de dados e levantamento das predições.

Em complemento à arquitetura, os processos para a mineração de dados representam uma grande parte da proposta do sistema Devias, com o intuito de descrever este procedimento, o próximo tópico retrata de maneira explicativa as fases e a ordem dos acontecimentos para o processo de importação.

#### 4.1.1 Processo de Mineração de Dados

A mineração de dados por parte do sistema Devias para a predição das variáveis dos alunos pode ser descrita por meio da arquitetura disposta na Figura 3.

FIGURA 3 – ARQUITETURA: PROCESSO DE MINERAÇÃO DE DADOS



Fonte: Os Autores (2018).

O processo inicia-se logo após a carga da nova base de dados, isto é, a leitura dos registros da planilha recém-importada e inserção das informações para a tabela base do sistema, que leva o nome de `aluno_disciplina`. Este procedimento é descrito nos fluxos de “Excel” e “Insert” da Figura 3.

A partir deste momento são realizadas operações de banco para criação e preenchimento das outras tabelas utilizadas pelo sistema, são os casos da tabela “principal” e tabela “aluno”. Em ambas as tabelas, além da distribuição de dados que ocorre durante a inserção, é aplicada uma funcionalidade da mineração de dados conhecida por Enriquecimento.

O enriquecimento dos dados, consiste basicamente na geração de novos valores de interesse, mas que não implicam no uso e processamento de algoritmos de *machine learning* para seu descobrimento (CÔRTEZ, C. S., PORCARO, R. M.; LIFSCHITZ, S., 2002), isto é, se eu subtrair um valor presente em uma coluna por um outro valor de uma coluna diferente e multiplicá-lo por uma constante, obtenho um novo dado que pode ser de meu interesse, com isto, estou enriquecendo minha fonte de dados.

Os dados providos do enriquecimento são de grande importância para o sistema Devias, como por exemplo, as variáveis de perfis dos alunos, presentes

na tabela “aluno” foram geradas durante este processo, e posteriormente servirão de auxílio para a geração do modelo de predição, são elas:

- “is\_evasor” – indicador se o aluno evadiu do curso;
- “is\_retentor” – indicador se o aluno é retentor ou reteve durante sua passagem pelo curso;
- “is\_formado” – indicador se o aluno se graduou pelo curso.

Suas origens são dadas de grandes consultas realizadas durante o pré-processamento no fluxo de Enriquecimento, presente na Figura 3 como terceiro item.

O término deste procedimento de “organização” da base de dados, marca o início do processo de chamada ao sistema Weka. Por meio do php, são preparados dois csv’s, um para o treinamento do modelo de retenção, e outro para o modelo de evasão. Ambos contendo os dados de alunos que já passaram por seu respectivo fenômeno.

Este arquivo que foi preparado, acrescido do cabeçalho apropriado e alterada sua extensão para .arff, torna-se então, o arquivo de input esperado pelo Weka para dar início ao treinamento.

Por meio da linha de comando, também chamada dentro do código php, são passados como parâmetros o arquivo arff e o algoritmo de treinamento desejado. Para o treinamento e obtenção do modelo de evasão foi utilizado o algoritmo Bayesian network, que obteve o melhor resultado com a utilização dos dados de evasão, enquanto que para o modelo de retenção, foi utilizado o algoritmo de Random Forest, conforme apontamento realizado no capítulo 3, Tabelas 1 e 2.

Realizada a execução da linha de comando, o Weka retorna nos dois casos um arquivo no formato .model. Este arquivo possui o classificador capaz de dizer se um aluno irá reter ou evadir ao realizar seu cruzamento com um outro arquivo arff que ainda não foi treinado.

Para o cruzamento com o classificador, é gerado então dois novos arff’s, agora, com os dados dos alunos que ainda não passaram pelos fenômenos da retenção ou evasão. A conclusão deste cruzamento, também executado por linha de comando, é convertido para um csv simples, e inserido dentro da base

de dados na tabela “aluno”, nos campos de “previsao\_evasao” e “previsao retenção”.

Com as informações prontas na base de dados, o processo de importação é encerrado, tornando o sistema apto a gerar as análises desenvolvidas e apresentar as interfaces em toda sua totalidade, conforme demonstrado nos tópicos a seguir.

#### 4.2. APRESENTAÇÃO DO DEVIAS

Dentro do sistema Devias encontramos seis interfaces principais que em sua totalidade são capazes de representar todas as funcionalidades do sistema, são elas:

- Análises, onde estão situados os gráficos das análises realizadas;
- Dados, onde encontra-se uma listagem personalizada dos alunos, que possibilita a parametrização de buscas e filtros;
- Atendimentos, onde encontram-se as operações de acompanhamento junto ao aluno, como os atendimentos prestados pelos professores aos alunos e os acompanhamentos de jubilações registrados;
- Usuários, área reservada para o controle de usuários do sistema;
- Configurações, onde é possível realizar a importação de uma nova planilha de dados e configuração das disciplinas do curso;
- Análise do aluno, uma ficha do aluno, que apresenta seu histórico e com as pendências para a conclusão do curso.

As subseções a seguir apresentam cada uma destas telas.

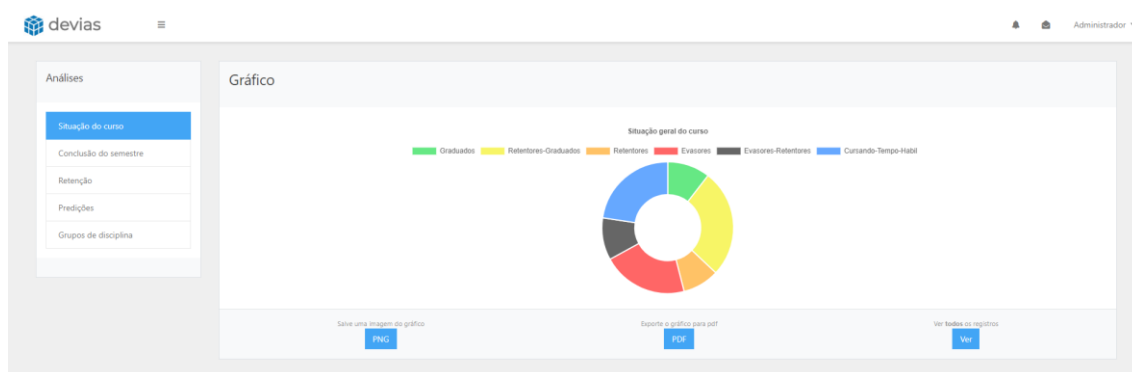
##### 4.2.1. Análises

Nesta tela, é possível realizar consultas às análises feitas sobre os dados importados, para cada opção de análise é listado seu respectivo gráfico. O sistema conta atualmente com cinco opções de análises desenvolvidas.

#### 4.2.1.1. Situação geral do curso

Para esta análise foi utilizado um gráfico em rosca, onde cada área do gráfico representa a quantidade de alunos que se encaixam em determinado perfil dentro do curso, conforme a Figura 4 abaixo.

FIGURA 4 – TELA DE ANÁLISE: GRÁFICO DA SITUAÇÃO GERAL DO CURSO



Fonte: Os Autores (2018).

Os rótulos para este gráfico correspondem ao perfil de cada aluno dentro do curso, são os perfis:

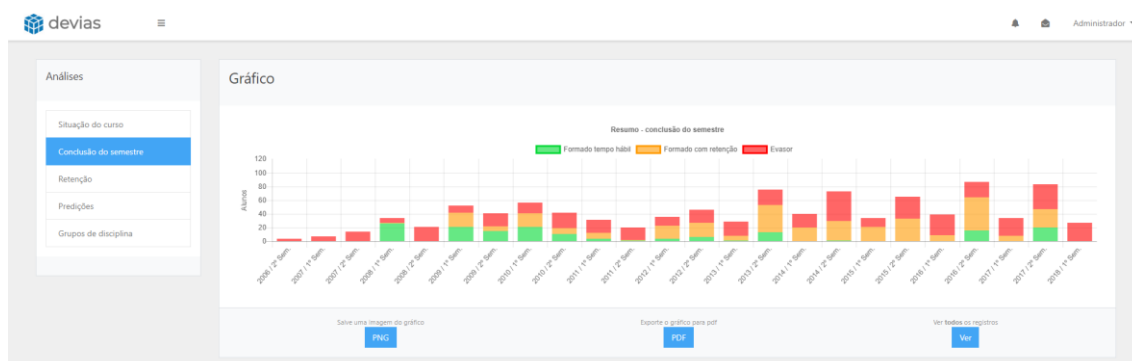
- a) Graduado, já formados no curso – representados pela cor verde;
- b) Graduado, porém, passando por retenção – representados pela cor amarela;
- c) Retentor, dentro do curso há mais tempo que o esperado para formação – representados pela cor laranja;
- d) Evasor, que por algum motivo deixou o curso – representados pela cor vermelha;
- e) Evasor, que ainda passou por um período de retenção antes de deixar o curso – representados pela cor preta;
- f) Cursando em tempo hábil, que é o aluno sem nenhuma irregularidade com vínculo ao curso – representados pela cor azul.

#### 4.2.1.2. Conclusão do semestre

Neste gráfico o modelo de barras em pilhas foi o escolhido para realizar um balanço do fim de cada semestre. Para seu entendimento, é importante ressaltar que o tamanho da pilha (eixo y) é influenciado pelo número de alunos,

e a separação das pilhas é feita por semestres (eixo x), conforme mostra a Figura 5.

FIGURA 5 – TELA DE ANÁLISE: GRÁFICO DE CONCLUSÃO DO SEMESTRE



Fonte: Os Autores (2018).

Os rótulos deste gráfico representam as ações dos alunos que impactam em um término de um semestre para o curso, são as ações:

- Concluíram o curso em tempo hábil – representados pela cor verde;
- Concluíram o curso, porém passando pelo fenômeno da retenção – representados pela cor laranja;
- Evadiram – representados pela cor vermelha.

#### 4.2.1.3. Retenção

Na análise de Retenção, o gráfico de pilhas agrupadas foi escolhido com o intuito de demonstrar a quantidade de períodos (eixo x) que os alunos em retenção (eixo y) cursam a mais, até que passem por algumas das medidas conclusivas (formação ou evasão) ou não (períodos que estão cursando a mais sem perder o vínculo com a instituição).

FIGURA 6 – TELA DE ANÁLISE: GRÁFICO DE RETENÇÃO



Fonte: Os Autores (2018).

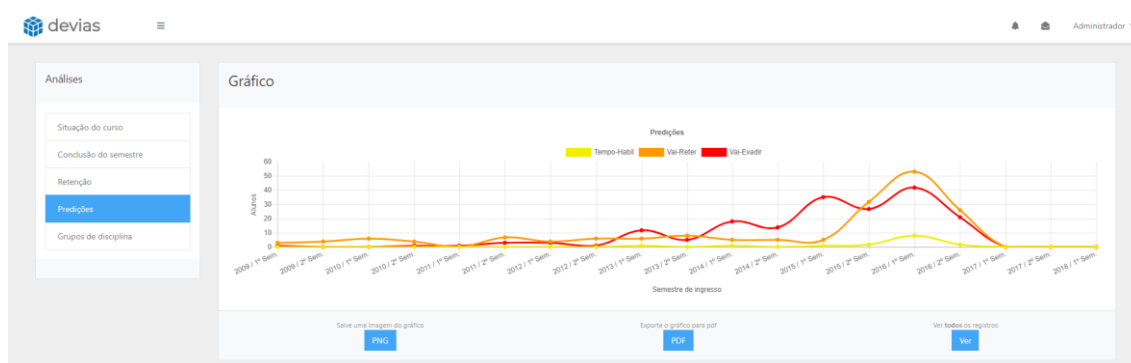
Os rótulos para este gráfico são as ações conclusivas somadas aos alunos que continuam cursando em retenção, representadas desta forma:

- Alunos que permanecem vinculados à instituição – representados pela cor amarela;
- Alunos que se formaram depois de completarem a quantidade especificada de semestres – representados pela cor laranja;
- Alunos que depois de reterem pela quantidade especificada de semestres evadiram do curso – representados pela cor vermelha.

#### 4.2.1.4. Predições

Para as predições do sistema foram escolhidas linhas de projeções, estas, com o objetivo de demonstrar a quantidade de alunos (eixo x) que segundo o modelo aplicado, enquadram-se na variável de interesse específica, separando os alunos por seu ano de ingresso (eixo y).

FIGURA 7 – TELA DE ANÁLISE: GRÁFICO DE PREDIÇÕES



Fonte: Os Autores (2018).

Os rótulos para o gráfico de predições são as variáveis de interesse que impactam diretamente nos índices de sucesso de um curso de graduação:

- Alunos que provavelmente irão concluir em tempo hábil – representados pela cor amarela;
- Alunos que possivelmente passarão pelo fenômeno da retenção – representados pela cor laranja;
- Alunos que correm o risco de evadir do curso – representados pela cor vermelha.

#### 4.2.1.4. Desempenho por grupo de disciplinas

Para esta análise foi utilizado novamente um gráfico de barras empilhadas, agora, com a finalidade de demonstrar o desempenho dos alunos (eixo x) do curso nos grupos de disciplinas (eixo y) definidos através da parametrização na tela de configuração de disciplinas.

FIGURA 8 – TELA DE ANÁLISE: GRÁFICO DE DESEMPENHO POR GRUPO DE DISCIPLINA



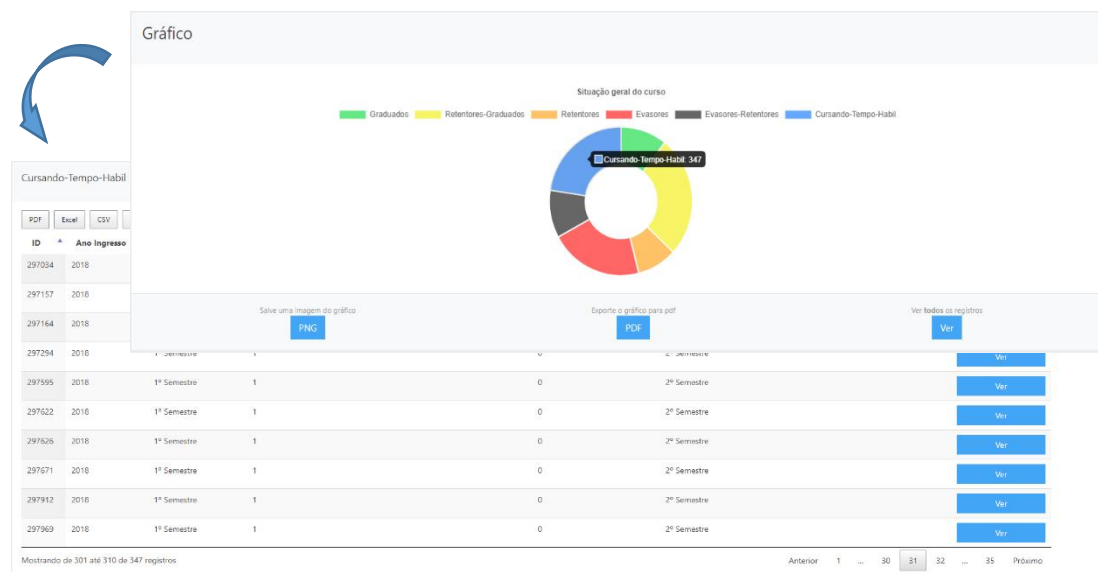
Fonte: Os Autores (2018).

Os rótulos para esta análise são as possíveis conclusões que o aluno define para a sua passagem pela disciplina:

- Aprovado – representado pela cor verde;
- Reprovado – representado pela cor laranja;
- Desistiu – representado pela cor vermelha.

Todos os gráficos apresentam a possibilidade de exportação nos formatos png e pdf, e ainda a funcionalidade de ao clicar sobre uma região específica, ou uma barra do gráfico, é gerada uma relação dos alunos que correspondem à informação daquela área clicada. Por exemplo, clicando sobre a região de alunos cursando em tempo hábil como mostra a Figura 9, é carregada uma lista dos alunos que possuem esta característica, podendo ainda gerar um relatório pdf, uma planilha excel, csv, cópia dos dados para a área de transferência com aquela lista gerada.

FIGURA 9 – TELA DE ANÁLISE: LISTAGEM DE ÁREA DO GRÁFICO



Fonte: Os Autores (2018).

A funcionalidade de relatório e exportações não é exclusiva desta tela, e podem ser feitas em qualquer tabela dentro do sistema, possibilitando uma liberdade bem grande sobre os dados ali presentes.

#### 4.2.2. Dados

A tela de dados é uma grande listagem dos alunos presentes na última importação, colocando em evidência alguns de seus dados, é possível realizar filtragens e pesquisar para atingir o resultado esperado como mostrado na Figura 10.

FIGURA 10 – TELA DE DADOS

GRR	Nome	Ano Ingresso	Período Ingresso	Períodos Cursados	Formado	Evasor	Retentor	Ano Evasão	Período Eva
466		2012	2º sem.	2	-	✖ Evasor	-	2013	2º sem.
724		2006	1º sem.	7	✔ Graduado	-	⚠ Retentor	2009	1º sem.
770		2009	2º sem.	6	-	✖ Evasor	-	2012	2º sem.
779		2008	1º sem.	6	✔ Graduado	-	-	2010	2º sem.
1975		2011	1º sem.	8	✔ Graduado	-	⚠ Retentor	2014	2º sem.
2235		2006	1º sem.	6	✔ Graduado	-	-	2008	1º sem.
3069		2015	1º sem.	7	-	-	⚠ Retentor		
3645		2011	1º sem.	5	✔ Graduado	-	-	2013	2º sem.
3924		2012	1º sem.	10	✔ Graduado	-	⚠ Retentor	2016	
6204		2007	1º sem.	1	-	✖ Evasor	-	2007	2º sem.

Mostrando de 1 até 10 de 1.535 registros

Fonte: Os Autores (2018).

Cada coluna possui sua pesquisa, e fazendo corrimento lateral da tabela é possível encontrar o botão de análise do aluno, onde poderá ser consultada sua ficha com dados mais específicos como as notas e pendências para conclusão do curso.

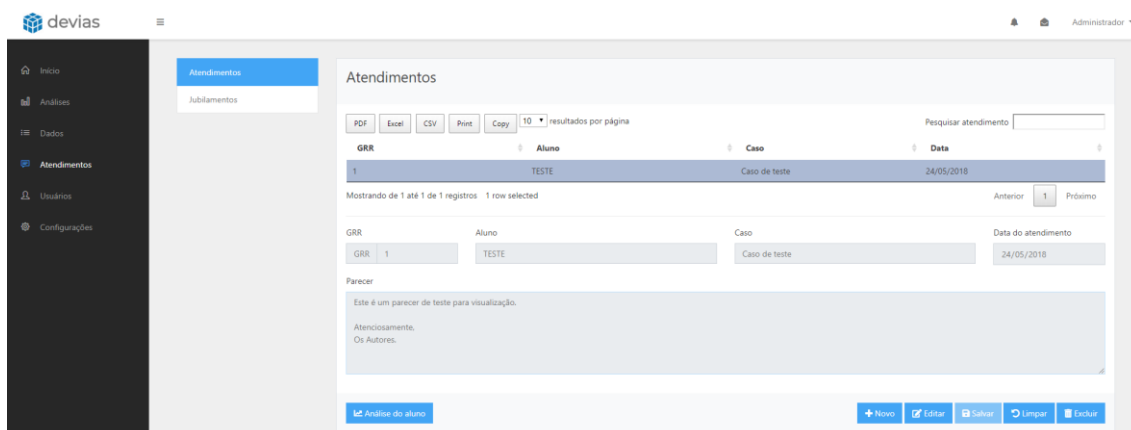
#### 4.2.3. Atendimentos

A tela de atendimentos é dividida em duas abas, Atendimentos aos alunos e Jubilamentos. Destinada aos professores que participam da Comissão de Orientação Acadêmica (COA), a tela possui a restrição de acesso, permitindo somente os usuários que atendem por aquela função o acesso.

##### 4.2.3.1. Atendimentos aos alunos

A tela conta com uma tabela dos atendimentos juntamente aos campos que caracterizam o atendimento, para que as operações de consulta, inserção, alteração e exclusão ocorram de forma rápida e sucinta, conforme Figura 11.

FIGURA 11 – TELA DE ATENDIMENTOS



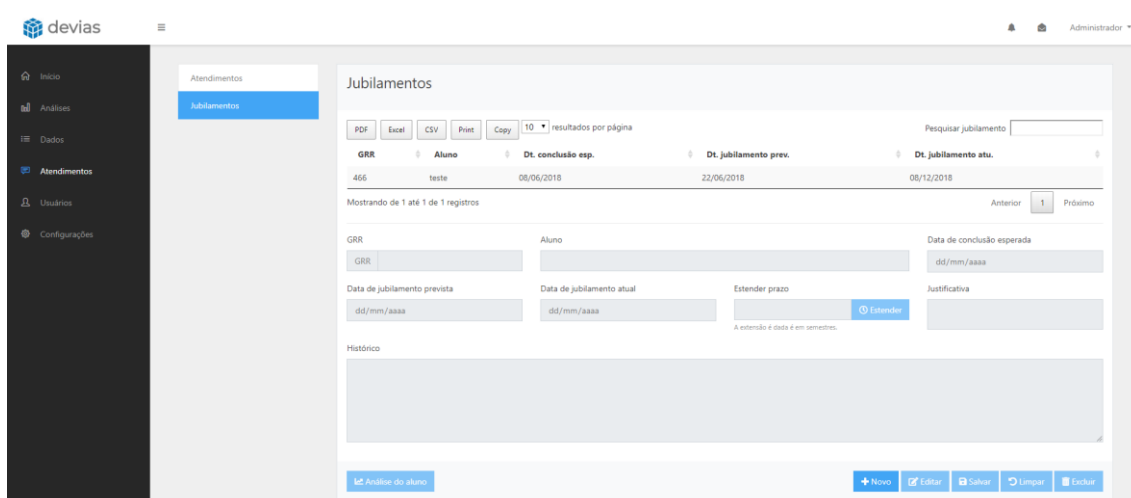
Fonte: Os Autores (2018).

As operações da tela se dão através da barra de ação localizada no canto inferior direito, e na mesma barra encontra-se o botão para a tela de análise do aluno, caso um registro de um atendimento de aluno esteja selecionado.

#### 4.2.3.2. Jubilamentos

Bastante similar a tela de atendimentos ao aluno, a tela de jubramento possui alguns poucos campos à mais, e apresenta uma funcionalidade fora as operações padrões de controle de registros, nela é possível fazer diretamente a extensão do prazo do jubramento apenas informando o número de semestres seguido de um motivo, vide Figura 12.

FIGURA 12 – TELA DE JUBILAMENTOS

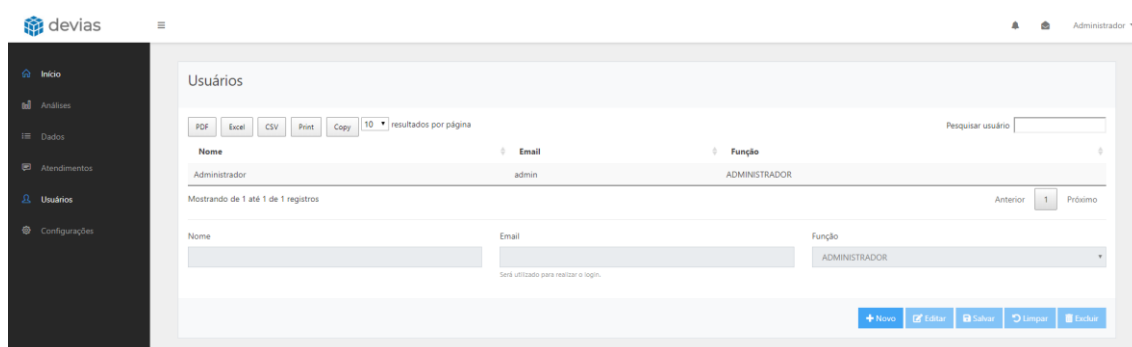


Fonte: Os Autores (2018).

#### 4.2.4. Usuários

Para a tela de usuários, as operações de controle de registros são dadas sobre as contas dos usuários do sistema, com suas devidas restrições. Não esquecendo que também há o controle de acesso para abri-la. O modelo da tela é apresentado na Figura 13.

FIGURA 13 – TELA DE USUÁRIOS



Fonte: Os Autores (2018).

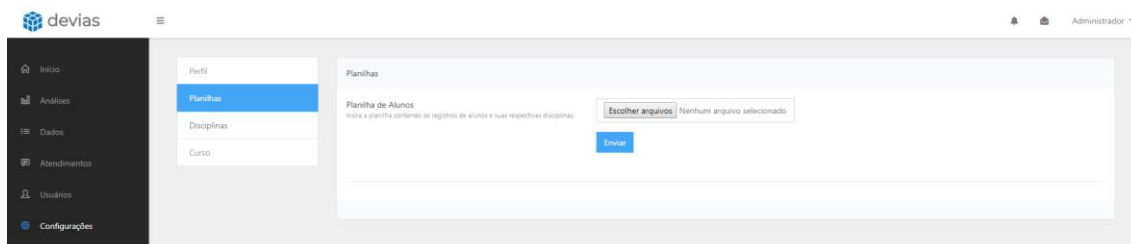
#### 4.2.5. Configurações

Na interface de configurações são disponibilizadas as parametrizações do sistema. Para este momento da aplicação é possível realizar a troca de bases de dados do sistema através da importação da planilha, e a configuração das disciplinas do curso.

##### 4.2.5.1. Planilha de dados

Na interface em questão, é o local que o sistema dispõe para o usuário inserir a planilha de dados no formato esperado, para então iniciar o processo de troca da base dados. A seleção do arquivo por meio do seletor, seguido do clique do botão “Enviar” dá início também aos processos de mineração de dados com o intuito de gerar as predições de evasão, retenção e conclusão em tempo hábil dos alunos recém-importados. A conclusão do procedimento marca a troca da base de dados e toda e qualquer análise apresentada no tópico Análises passará a ser gerada com base nos novos dados. A representação da tela é ilustrada na Figura 14.

FIGURA 14 – TELA DE IMPORTAÇÃO DA PLANILHA

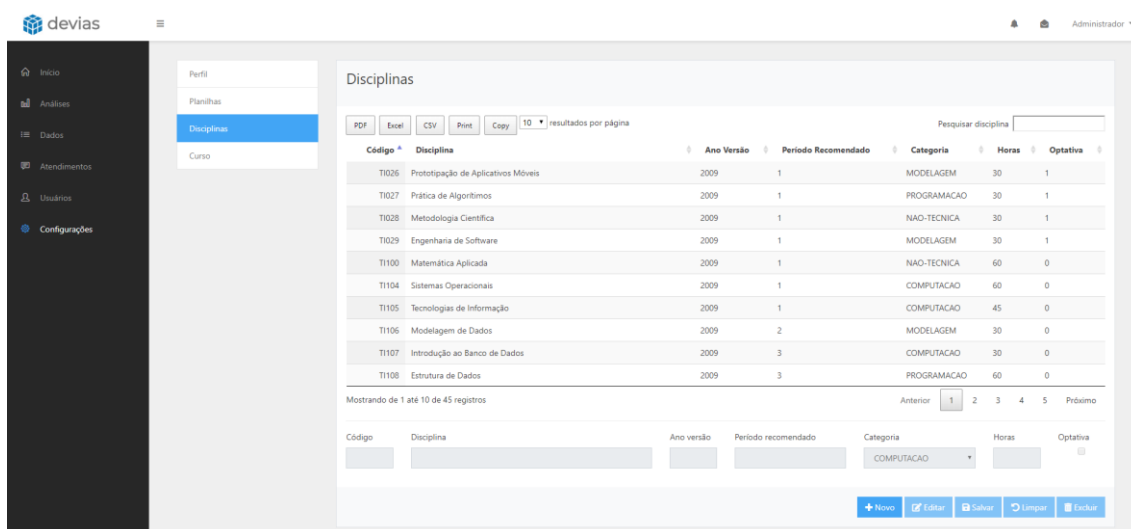


Fonte: Os Autores (2018).

#### 4.2.5.2. Disciplinas

Na tela de disciplinas são configuradas as matérias que são necessárias para o término do curso, uma vez que o aluno é capaz de fazer disciplinas inertes ao curso que pertencer, é percebida a necessidade de um tratamento específico dos registros que entrariam como um empecilho nas definições do que falta ou não para o aluno concluir o curso. Como parametrização da disciplina ainda, é possível classifica-la dentro de uma categoria. Este agrupamento é utilizado dentro da análise de desempenho por grupo de disciplinas, anteriormente apresentado. A Figura 15 demonstra o esquema da tela.

FIGURA 15 – TELA DE CONFIGURAÇÃO DE DISCIPLINAS



Fonte: Os Autores (2018).

#### 4.2.5. Análise do aluno

Nesta última interface do sistema é onde ocorre a apresentação da ficha do aluno. Acessada de diversas áreas do sistema, a tela recebe como parâmetro o código referente ao aluno para gerar o seu histórico dentro do curso e também

as disciplinas que faltam para a conclusão. Além disto, são dispostas algumas informações como as características que o aluno possui, segundo as análises previamente apresentadas, podendo apresentar o perfil de evasor, retentor, graduado, ou ainda, que apresenta riscos de evasão ou retenção. Vide Figura 16.

FIGURA 16 – TELA DE ANÁLISE DO ALUNO

**Informações do aluno**

GRR: 272615  
 Nome:  
 Curso:  
 Ano de Ingresso: 2015  
 Semestre atual: 6º  
 Características:  
 ★ terminará em tempo hábil

**Para terminar o curso**

Copy Excel CSV PDF Print resultados por página Faltam 120 horas de optativas

Disciplinas restantes:

CÓDIGO DISCIPLINA	DISCIPLINA	OPTATIVA
TI026	Prototipação de Aplicativos Móveis	1
TI027	Prática de Algoritmos	1
TI028	Metodologia Científica	1
TI029	Engenharia de Software	1
TI135	Inglês Técnico I	1

Mostrando de 1 até 5 de 14 registros

**Histórico**

GRR	CURSO	CÓDIGO DISCIPLINA	DISCIPLINA	ANO QUE CURSOU	SEMESTRE	FALTAS	MÉDIA	SITUAÇÃO	PERÍODO CURSADO
272615	4BA	D5630	DIREITO APLICADO	2018	1º Semestre	0	0.000000	Matricula	6
272615	4BA	TI026	Prototipação de Aplicativos Móveis	2018	1º Semestre	0	0.000000	Matricula	6
272615	4BA	TI027	Prática de Algoritmos	2018	1º Semestre	0	0.000000	Matricula	6
272615	4BA	TI028	Metodologia Científica	2018	1º Semestre	0	0.000000	Matricula	6
272615	4BA	TI100	Matemática Aplicada	2015	2º Semestre	0	85.000000	Aprovado	1
272615	4BA	TI104	Sistemas Operacionais	2015	2º Semestre	1	88.000000	Aprovado	1
272615	4BA	TI105	Tecnologias de Informação	2015	2º Semestre	0	88.000000	Aprovado	1
272615	4BA	TI106	Modelagem de Dados	2016	1º Semestre	0	73.000000	Aprovado	2
272615	4BA	TI107	Introdução ao Banco de Dados	2016	2º Semestre	0	82.000000	Aprovado	3
272615	4BA	TI108	Estrutura de Dados	2016	2º Semestre	2	80.000000	Aprovado	3

Mostrando de 1 até 10 de 41 registros

Fonte: Os Autores (2018).

Neste capítulo foram apresentadas a arquitetura em que se baseia o sistema e todas as suas funcionalidades desenvolvidas. O capítulo 5 a seguir, apresentará as considerações finais do grupo em relação ao desenvolvimento do projeto no decorrer deste percurso.

## 5 CONSIDERAÇÕES FINAIS

O projeto de desenvolvimento do sistema Devias surgiu como uma proposta da orientação da equipe, observando a situação atual do curso de Tecnologia em Análise e Desenvolvimento de Sistemas da UFPR, que há muitos anos sofre com duas situações extremamente opostas entre si. A primeira, a passagem efêmera dos alunos por entre as disciplinas, sem de fato, concluí-las, deixando a universidade. E a segunda, as passagens que dizem completamente o contrário, até deixam a impressão de serem eternas, onde o aluno se mantém vinculado à instituição, reprovando uma, duas, três vezes a mesma disciplina e prolongando sua estadia.

A principal questão envolvida na motivação do projeto foi: é possível identificar quem são estes alunos, antes mesmo deles se tornarem este indicador negativo?

Os testes dos algoritmos de predição foram então, realizados. Enviando para treinamento apenas o primeiro semestre dos alunos. Dentro dos melhores resultados foi possível obter uma acurácia de 89% com o algoritmo de Random Forest para os casos de retenção, e 75% com Bayes Network para os casos de evasão. Este número não aumentava não importando quantos semestres fossem adicionados ao treinamento. Foi constatado, portanto, que o primeiro período de um aluno diz muito mais sobre ele, do que qualquer outro.

O objetivo final, que era desenvolver uma aplicação capaz de mostrar para o usuário informações relevantes sobre o curso e seus alunos, de maneira a auxiliar a tomada de decisão, foi atingido com sucesso, dentro daquilo que lhe fora proposto.

De qualquer maneira, o sistema não atingiu seu potencial máximo, foram identificadas após sua conclusão possíveis melhorias, reduções de limitadores e novas funcionalidades que tornariam a ferramenta ainda mais completa, como:

- Desenvolvimento de outras análises, que trariam ainda mais certezas ao tomador de decisões.
- Integrações com o sistema de matrículas da universidade, possibilitando o consumo de alguns dados básicos, porém, interessantes para a aplicação, como por exemplo a foto do aluno.

- Utilização de uma ferramenta para geração de gráficos mais poderosa, visto que, em algumas situações foram identificadas limitações por parte daquela utilizada, como por exemplo: possibilidade de utilizar mais de um eixo x e a falta de dinamismo na passagem de parâmetros para a geração.
- Acesso irrestrito aos dados, descartando a necessidade de geração de uma planilha e a remontagem da estrutura do banco de dados em toda importação. Este mesmo acesso possibilitaria a exploração de novas variáveis, ampliando ainda mais a taxa de acurácia nas predições.

Futuramente, o sistema terá de lidar com o problema conhecido na mineração de dados como *cold start*. O *cold start* é uma situação pela qual o sistema passa quando há escassez ou falta de dados para realizar uma determinada análise ou levantamento (LÁZARO, 2010).

Os cursos de graduação de tempos em tempos passam por uma reformulação de sua grade de atividades, alterando muitas de suas disciplinas. O sistema Devias teria que dispor de uma nova funcionalidade, que permitiria fazer a equivalência entre disciplinas novas, ou ainda, uma maneira de relacioná-las às anteriores e ainda tornar possível a utilização do curso antigo para a análise do curso recém-criado.

A disponibilidade do sistema para outros cursos, até mesmo de diferentes áreas é uma visão bastante próxima da realidade, e seria um grande passo a ser dado para contribuir com a melhora da qualidade dos cursos de graduação da Universidade Federal do Paraná.

## REFERÊNCIAS

ABRAMOVAY, Miriam; CASTRO, Mary Garcia. **Ensino médio: múltiplas vozes**. Unesco, 2003.

BAKER, R., ISOTANI, S., & CARVALHO, A. **Mineração de dados educacionais**: Oportunidades para o Brasil. Brazilian Journal of Computers in Education. 2011. Disponível em: <<http://br-ie.org/pub/index.php/rbie/article/view/1301/1172>>. Acesso em: 10 maio 2018.

BASSI FILHO, Dairton Luiz. Experiências com desenvolvimento ágil. **São Paulo: IME-USP**, 2008.

BIAZUS, C. A. **Sistema de fatores que influenciam o aluno a evadir-se dos cursos de graduação na UFSM e na UFSC**: um estudo no curso de Ciências Contábeis. Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Produção. 2004. Disponível em: <<https://repositorio.ufsc.br/handle/123456789/87138>>. Acesso em: 10 maio 2018.

CAMPELLO, A. V. C.; LINS, L. N. **Metodologia de análise e tratamento da evasão e retenção em cursos de graduação de instituições federais de ensino superior**. Anais XXVIII Encontro Nacional de Engenharia de Produção, Rio de Janeiro, 2008. Disponível em: <[http://www.abepro.org.br/biblioteca/enegep2008\\_TN\\_STO\\_078\\_545\\_11614.pdf](http://www.abepro.org.br/biblioteca/enegep2008_TN_STO_078_545_11614.pdf)>. Acesso em: 10 maio 2018.

COSTA, Evandro et al. Mineração de dados educacionais: conceitos, técnicas, ferramentas e aplicações. **Jornada de Atualização em Informática na Educação**, v. 1, n. 1, p. 1-29, 2013.

DA COSTA CÔRTEZ, Sérgio; PORCARO, Rosa Maria; LIFSCHITZ, Sérgio. **Mineração de dados-funcionalidades, técnicas e abordagens**. PUC, 2002.

DA SILVA, Valéria Martins. **Revisão sistemática da evolução mvc na base acm**. 2012. Disponível em: [https://www.researchgate.net/profile/Valeria\\_Silva8/publication/264003410\\_Revision\\_sistemica\\_da\\_evolucao\\_MVC\\_na\\_base\\_ACM/links/00b4953c839fa9bdd2000000.pdf](https://www.researchgate.net/profile/Valeria_Silva8/publication/264003410_Revision_sistemica_da_evolucao_MVC_na_base_ACM/links/00b4953c839fa9bdd2000000.pdf) Acesso em: 19 jun 2018.

DE AMO, S. **Técnicas de mineração de dados**. Jornada de Atualização em Informática. 2004. Disponível em: <<http://www.deamo.prof.ufu.br/arquivos/JAI-cap5.pdf>>. Acesso em: 10 maio 2018.

DE MELLO, Carlos Eduardo R.; DA SILVA, Geraldo Zimbrão; DE SOUZA, Jano M. Extensão do WEKA para Métodos de Agrupamento com Restrição de Contigüidade. **Programa de Engenharia de Sistemas e Computação Universidade Federal do Rio de Janeiro (UFRJ)**. Acessado em, v. 10, 2007.

FADEL, Aline Cristine; SILVEIRA, Henrique da Mota. Metodologias ágeis no contexto de desenvolvimento de software: XP, Scrum e Lean. **Monografia do Curso de Mestrado FT-027-Gestão de Projetos e Qualidade da Faculdade de Tecnologia–UNICAMP**, v. 98, p. 101, 2010.

FAVERO, Rute Vera Maria. Dialogar ou evadir: Eis a questão!: um estudo sobre a permanência e a evasão na educação a distância. 2006.

FRANK, E., HALL, M., TRIGG, L., HOLMES, G., & WITTEN, I. H. **Data mining in bioinformatics using Weka**. *Bioinformatics*, p. 2479-2481. 2004.

GADOTTI, M. **A questão da educação formal/não-formal**. Sion: Institut Internacional des Droits de 1<sup>o</sup> Enfant, p. 1-11, 2005.

GAIOSO, NP de L. O fenômeno da evasão escolar na educação superior no Brasil. **Brasília, DF: Universidade Católica de Brasília**, 2005.

HAN, Jiawei; KAMBER, Micheline; PEI, Jian. Mining frequent patterns, associations, and correlations. **Data Mining: Concepts and Techniques (2nd ed., pp. 227-283)**. San Francisco, USA: Morgan Kaufmann Publishers, 2006.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse Estatística da Educação Superior 2016**. Brasília: Inep, 2017. Disponível em: <<http://portal.inep.gov.br/basica-censo-escolar-sinopse-sinopse>>. Acesso em: 8 maio 2018.

LÁZARO, A. D. S. **Análise e seleção de Algoritmos de filtragem de informação para solução do problema Cold-start Item**. Monografia de Graduação, Departamento de Ciência da Computação, Universidade Federal de Lavras. (2010) Disponível em: <[http://www.bsi.ufla.br/wp-content/uploads/2013/07/Monografia\\_Alexsandra.pdf](http://www.bsi.ufla.br/wp-content/uploads/2013/07/Monografia_Alexsandra.pdf)>. Acesso em: 19 jun 2018.

LOTUFO, A. D., SOUZA Jr, C., COVACIC, M., & BRITO, J. M. **Evasão e repetência na FEIS/UNESP: análise e resultados**. In Anais do XXVI Congresso Brasileiro de Ensino de Engenharia (COBENGE), p. 185-203, 1998.  
MACHADO, Nílson José. A Universidade e a organização do conhecimento: a rede, o tácito, a dádiva. **Estudos avançados**, v. 15, n. 42, p. 333-352, 2001.

MERCERON, Agathe; YACEF, Kalina. Tada-ed for educational data mining. **Interactive multimedia electronic journal of computer-enhanced learning**, v. 7, n. 1, p. 267-287, 2005.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, v. 1, n. 1, p. 32, 2003.

OCDE. **Education at a Glance Brazil 2015**. Disponível em: <<http://www.oecd.org/eco/surveys/Brasil-2015-resumo.pdf>>. Acesso em: 8 maio 2018.

PATIL, Tina R.; SHEREKAR, S. S. Performance analysis of Naive Bayes and J48 classification algorithm for data classification. **International Journal of Computer Science and Applications**, v. 6, n. 2, p. 256-261, 2013.

QUONIAM, L., et al. **Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil**. 2001. Disponível em: <<http://repositorio.unb.br/handle/10482/965>>. Acesso em: 10 maio 2018.

SEVERINO PEREIRA, Alexandre et al. Fatores relevantes no processo de permanência prolongada de discentes nos cursos de graduação presencial: um estudo na Universidade Federal do Espírito Santo. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 23, n. 89, 2015.

SILVA FILHO, Roberto Leal Lobo et al. A evasão no ensino superior brasileiro. **Cadernos de pesquisa**, v. 37, n. 132, p. 641-659, 2007.

SILVA, F. G.; HOENTSCH, Sandra CP; SILVA, Leila. Uma análise das Metodologias Ágeis FDD e Scrum sob a Perspectiva do Modelo de Qualidade MPS. BR. **Scientia Plena**, v. 5, n. 12, 2009.

TANAKA, Eduardo Hideki. Design de interfaces web com jQuery. In: **Proceedings of the Companion Proceedings of the 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction**. Brazilian Computer Society, 2011. p. 29-32.

VELLASCO, M. M. B. R. **Redes Neurais Artificiais**. Rio de Janeiro, Brasil, PUC, notas de Aula, Brasil, 2007.

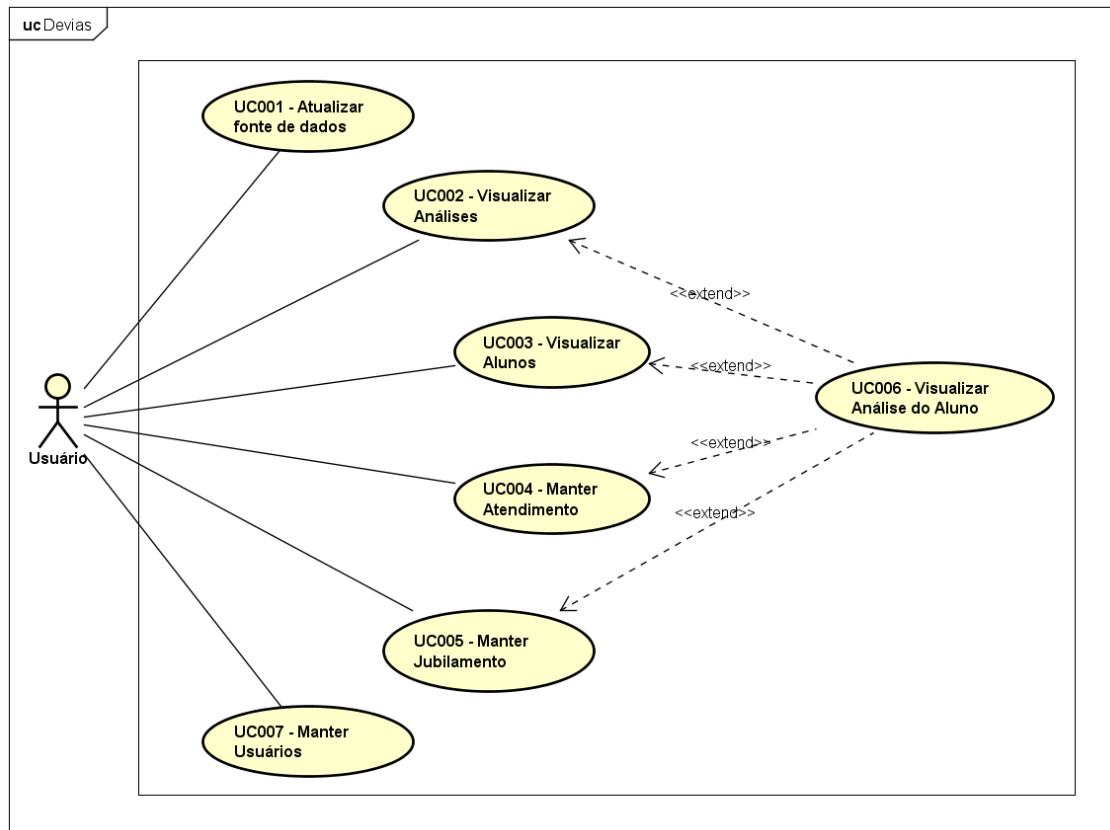
## APÊNDICE A – HISTÓRIAS DE USUÁRIO

- Como usuário gostaria de visualizar o desempenho de cada aluno (notas, faltas, sequência de matérias, matérias reprovadas) de forma individual para que ao consultar um aluno saiba exatamente seus pontos fracos, onde estão focadas suas maiores dificuldades.
- Como usuário gostaria de visualizar quais disciplinas faltam para o aluno concluir o curso para ter uma visão de quanto tempo ele ainda permaneceria vinculado à instituição.
- Como usuário gostaria de visualizar os alunos que correm riscos de evasão para que possam ser tomadas as devidas providências de auxílio e medidas cabíveis.
- Como usuário gostaria de visualizar os alunos que pendem para a retenção para que trabalhem em cima das dificuldades e entender seus motivos.
- Como usuário gostaria de visualizar a situação geral do curso, com o perfil de cada aluno para entender o atual momento do curso, e quais os problemas deveriam ter uma atenção especial.
- Como usuário gostaria de poder controlar os atendimentos prestados aos alunos para que seja facilitada sua consulta e seja mantida uma maior organização sobre o procedimento.
- Como usuário gostaria de poder realizar um acompanhamento do jubramento dos alunos.
- Como usuário gostaria de visualizar o desempenho dos alunos por grupos de disciplinas específicos (programação, análise, não-técnicas)
- Como usuário gostaria de visualizar o tempo em que os alunos costumam ficar em retenção até de fato, concluírem ou evadirem.
- Como usuário gostaria de visualizar a conclusão do semestre com os feitos dos alunos (evadiram, formaram-se em tempo hábil, formaram-se, porém, em retenção).

## APÊNDICE B – DIAGRAMA DE CASOS DE USO

O sistema Devias em sua totalidade, descreve suas funcionalidades com sete casos de uso inferidos pelo ator Usuário, como é mostrado na Figura 17.

FIGURA 17 – DIAGRAMA DE CASOS DE USO



Fonte: Os Autores. (2018)

## APÊNDICE C – ESPECIFICAÇÃO DOS CASOS DE USO

## 1 – UC001: ATUALIZAR FONTE DE DADOS

O caso de uso “Atualizar Fonte De Dados” pode ser observado detalhadamente na Tabela 1.

TABELA 4 – ESPECIFICAÇÃO UC001: ATUALIZAR FONTE DE DADOS

<b>Nome</b>	Atualizar Fonte de Dados
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para realizar a importação da planilha de dados para dentro do sistema.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema. Possuir um acesso de função como Coordenador ou Professor COA.
<b>Pós-condições</b>	O usuário passará a visualizar dentro do sistema os dados importados.

## PROTÓTIPO (DV 1)

A Web Page

http://dashtads.com.br

TADS

Dashboard

Graficos

Tabelas

Atendimento

Usuarios

Configurações

Atualizar fonte de dados

Selecione um arquivo .csv

Atualizar


<b>FLUXO PRINCIPAL</b>
<ol style="list-style-type: none"><li>1. O sistema apresenta a tela (DV 1).</li><li>2. O usuário pressiona o botão “...” e seleciona a planilha de dados que deseja importar.</li><li>3. O usuário pressiona o botão “Enviar”.</li><li>4. O sistema faz a leitura da planília e passa a inserir os registros no banco de dados.</li><li>5. O sistema realiza as operações de tratamento de dados e machine learning e insere os dados gerados nas suas respectivas tabelas.</li><li>6. O caso de uso é encerrado.</li></ol>
<b>FLUXO ALTERNATIVO - N/A</b>
<b>FLUXO DE EXCEÇÃO – N/A</b>

Fonte: Os Autores (2018).

## 2 – UC002: VISUALIZAR ANÁLISES

O caso de uso “Visualizar Análises” pode ser observado detalhadamente na Tabela 2.

TABELA 5 – ESPECIFICAÇÃO UC002: VISUALIZAR ANÁLISES

<b>Nome</b>	Visualizar Análises
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para visualizar as análises presentes no sistema.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema.
<b>Pós-condições</b>	O usuário poderá visualizar os gráficos, e os alunos que o compõem.
<b>PROTÓTIPO (DV 1)</b>	
	
<b>FLUXO PRINCIPAL</b>	
7. O sistema apresenta a tela (DV 1).	

8. O usuário seleciona uma análise.
9. O sistema busca na base de dados as informações e apresenta o gráfico em tela.
10. O usuário clica sobre uma área específica dentro do gráfico.
11. O sistema apresenta a lista de alunos que correspondem àquela região específica do gráfico. (A1)
12. O caso de uso é encerrado.

**FLUXO ALTERNATIVO – Visualizar aluno (A1)**

1. O usuário clica no registro do aluno.
2. O sistema inicia o caso de uso UC006 – Visualizar Análise do Aluno.

**FLUXO DE EXCEÇÃO – N/A**

## 3 – UC003: VISUALIZAR ALUNOS

O caso de uso “Visualizar Alunos” pode ser observado detalhadamente na Tabela 3.

TABELA 6 – ESPECIFICAÇÃO UC003: VISUALIZAR ALUNOS

<b>Nome</b>	Visualizar Alunos
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para visualizar os alunos presentes na última importação de dados realizada.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema.
<b>Pós-condições</b>	O usuário poderá visualizar os alunos e alguns de seus dados.
<b>PROTÓTIPO (DV 1)</b>	
<b>FLUXO PRINCIPAL</b>	
1. O sistema apresenta a tela. (DV1)	

2. O sistema realiza a busca dos alunos na base de dados e mostra as informações em tela.
3. O caso de uso é encerrado

**FLUXO ALTERNATIVO – Visualizar aluno (A1)**

1. O usuário clica no registro do aluno.
2. O sistema inicia o caso de uso UC006 – Visualizar Análise do Aluno.

**FLUXO DE EXCEÇÃO – N/A**

## 4 – UC004: MANTER ATENDIMENTO

O caso de uso “Manter Atendimento” pode ser observado detalhadamente na Tabela 4.

TABELA 7 – ESPECIFICAÇÃO UC004: MANTER ATENDIMENTO

<b>Nome</b>	Manter Atendimento
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para inserir um atendimento na base de dados.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema. Possuir um acesso de função como membro do COA ou Coordenador.
<b>Pós-condições</b>	O usuário terá inserido o atendimento realizado a um aluno na base de dados, permitindo seu controle para os devidos fins.

### PROTÓTIPO (DV 1)

The screenshot shows a web browser window with the URL 'http://dashtds.com.br'. The page title is 'A Web Page'. The application header includes a hamburger menu, the text 'TADS', and icons for mail, notifications, and user profile. A sidebar on the left contains navigation options: Dashboard, Graficos, Tabelas, Atendimento, Usuarios, and Configurações. The main content area is titled 'Atendimentos' and features a search bar. Below the search bar is a table with columns: Name (job title), Age, Nickname, and Employee. The table contains three rows of data. To the right of the table is a form with fields for 'Aluno:', 'Caso:', and 'Parecer:'. Below the form are buttons for 'Novo', 'Editar', 'Salvar', 'Cancelar', and 'Excluir'. At the bottom of the 'Atendimentos' section, there is an 'Importar atendimento' section with a file selection box and an 'Importar' button. Below this is a section titled 'Jubilamento' with another search bar.

Name (job title)	Age	Nickname	Employee
Giacomo Guilizzoni Founder & CEO	40	Peldi	<input type="radio"/>
Marco Botton Tuttofare	38		<input checked="" type="checkbox"/>
Mariah Maclachlan Better Half	41	Patata	<input type="checkbox"/>
Valerie Liberty Head Chef	:)	Val	<input checked="" type="checkbox"/>
<a href="#">Data Grid Docs</a>			<input type="checkbox"/>

### FLUXO PRINCIPAL

1. O sistema apresenta a tela. (DV1)

2. O sistema realiza a busca dos atendimentos na base de dados e mostra as informações em tela.
3. O usuário pressiona o botão “Novo”. (A1, A2)
4. O sistema alterna o comportamento da tela para o modo de inclusão.
5. O usuário insere as informações.
6. O usuário pressiona o botão “Salvar”.
7. O sistema consiste as informações. (E1)
8. O sistema faz a inclusão do atendimento na base de dados.
9. O caso de uso é encerrado.

#### **FLUXO ALTERNATIVO – Botão “Editar” (A1)**

1. O usuário seleciona um registro.
2. O usuário pressiona o botão “Editar”.
3. O sistema alterna o comportamento da tela para o modo de edição.
4. O usuário altera as informações.
5. O usuário pressiona o botão “Salvar”.
6. O sistema consiste as informações. (E1)
7. O sistema faz a alteração do atendimento na base de dados.
8. O caso de uso é encerrado.

#### **FLUXO ALTERNATIVO – Botão “Excluir” (A2)**

1. O usuário seleciona um registro.
2. O usuário pressiona o botão “Editar”.
3. O sistema faz a exclusão do atendimento na base de dados.
4. O caso de uso é encerrado.

#### **FLUXO DE EXCEÇÃO – Dados inválidos (E1)**

1. O sistema emite a mensagem “Por favor, insira corretamente os dados”.

## 5 – UC005: MANTER JUBILAMENTO

O caso de uso “Manter Jubilamento” pode ser observado detalhadamente na Tabela 5.

TABELA 8 – ESPECIFICAÇÃO UC005: MANTER JUBILAMENTO

<b>Nome</b>	Manter Atendimento
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para inserir um acompanhamento para jubilamento na base de dados.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema. Possuir um acesso de função como membro do COA ou Coordenador.
<b>Pós-condições</b>	O usuário terá iniciado o acompanhamento de jubilamento de um aluno na base de dados, permitindo seu controle para os devidos fins.

## PROTÓTIPO (DV 1)

The screenshot shows a web application interface for managing employee jubilees. The interface includes a navigation menu on the left with options like Dashboard, Graficos, Tabelas, Atendimento, Usuarios, and Configurações. The main content area shows a 'Jubilamento' section with a table of employees and a form for adding new records.

**Importar atendimento**

Selecione um arquivo .csv

**Jubilamento**

Aluno:

Ano:  Turno:  Tempo em retenção:

Jubila em:  Adiar prazo em:  semestres

Name (job title)	Age	Nickname	Employee
Giacomo Guilizzoni Founder & CEO	40	Peldi	<input checked="" type="checkbox"/>
Marco Botton Tuttofare	38		<input checked="" type="checkbox"/>
Mariah MacLachlan Better Half	41	Patata	<input type="checkbox"/>
Valerie Liberty Head Chef		;) Val	<input checked="" type="checkbox"/>
<a href="#">Data Grid Docs</a>			<input type="checkbox"/>

<b>FLUXO PRINCIPAL</b>
<ol style="list-style-type: none"><li>1. O sistema apresenta a tela. (DV1)</li><li>2. O sistema realiza a busca dos jubilamentos na base de dados e mostra as informações em tela.</li><li>3. O usuário pressiona o botão “Novo”. (A1, A2, A3)</li><li>4. O sistema alterna o comportamento da tela para o modo de inclusão.</li><li>5. O usuário insere as informações.</li><li>6. O usuário pressiona o botão “Salvar”.</li><li>7. O sistema consiste as informações. (E1)</li><li>8. O sistema faz a inclusão do acompanhamento na base de dados.</li><li>9. O caso de uso é encerrado.</li></ol>
<b>FLUXO ALTERNATIVO – Botão “Editar” (A1)</b>
<ol style="list-style-type: none"><li>1. O usuário seleciona um registro.</li><li>2. O usuário pressiona o botão “Editar”.</li><li>3. O sistema alterna o comportamento da tela para o modo de edição.</li><li>4. O usuário altera as informações.</li><li>5. O usuário pressiona o botão “Salvar”.</li><li>6. O sistema consiste as informações. (E1)</li><li>7. O sistema faz a alteração do atendimento na base de dados.</li><li>8. O caso de uso é encerrado.</li></ol>
<b>FLUXO ALTERNATIVO – Botão “Excluir” (A2)</b>
<ol style="list-style-type: none"><li>1. O usuário seleciona um registro.</li><li>2. O usuário pressiona o botão “Editar”.</li><li>3. O sistema faz a exclusão do atendimento na base de dados.</li><li>4. O caso de uso é encerrado.</li></ol>
<b>FLUXO ALTERNATIVO – Botão “Estender” (A3)</b>
<ol style="list-style-type: none"><li>1. O usuário seleciona um registro.</li><li>2. O usuário preenche o número de semestres à acrescentar e seleciona o(s) motivo(s) na caixa de seleção.</li><li>3. O usuário pressiona o botão “Estender”.</li><li>4. O sistema acresce o prazo para o jubilamento na base de dados do sistema.</li></ol>

5. O caso de uso é encerrado.

**FLUXO DE EXCEÇÃO – Dados inválidos (E1)**

1. O sistema emite a mensagem “Por favor, insira corretamente os dados”.

## 6 – UC006: VISUALIZAR ANÁLISE DO ALUNO

O caso de uso “Visualizar Análise Do Aluno” pode ser observado detalhadamente na Tabela 6.

TABELA 9 – ESPECIFICAÇÃO UC006: VISUALIZAR ANÁLISE DO ALUNO

<b>Nome</b>	Visualizar Análise do aluno
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para visualizar os alunos presentes na última importação de dados realizada.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema.
<b>Pós-condições</b>	O usuário poderá visualizar os alunos e alguns de seus dados.
<b>PROTÓTIPO (DV 1)</b>	
<b>FLUXO PRINCIPAL</b>	
1. O sistema apresenta a tela. (DV1)	

2. O sistema realiza a busca das informações do aluno na base de dados e mostra as informações em tela.
3. O caso de uso é encerrado

**FLUXO ALTERNATIVO – N/A**

**FLUXO DE EXCEÇÃO – N/A**

## 7 – UC007: MANTER USUÁRIOS

O caso de uso “Manter Usuários” pode ser observado detalhadamente na Tabela 7.

TABELA 10 – ESPECIFICAÇÃO UC007: MANTER USUÁRIOS

<b>Nome</b>	Manter Usuário
<b>Ator principal</b>	Usuário
<b>Resumo</b>	Neste caso de uso está descrito o fluxo necessário para inserir um usuário na base de dados.
<b>Pré-condições</b>	O usuário precisa estar com uma sessão ativa no sistema. Possuir um acesso de função como Coordenador ou Professor COA.
<b>Pós-condições</b>	O usuário terá inserido o atendimento realizado a um aluno na base de dados, permitindo seu controle para os devidos fins.

## PROTÓTIPO (DV 1)

A Web Page

http://dashads.com.br

TADS

Dashboard

Graficos

Tabelas

Atendimento

Usuarios

Configurações

localizar

Nome	Perfil	Ativo
Rafaela Mantovani	COA	<input checked="" type="checkbox"/>
L Luiz Antonio Neves	Coordenador	<input checked="" type="checkbox"/>
Alex Kutzke	COA	<input checked="" type="checkbox"/>
Jaiminho	Professor	<input checked="" type="checkbox"/>

Nome: Rafaela Mantovani Perfil: COA

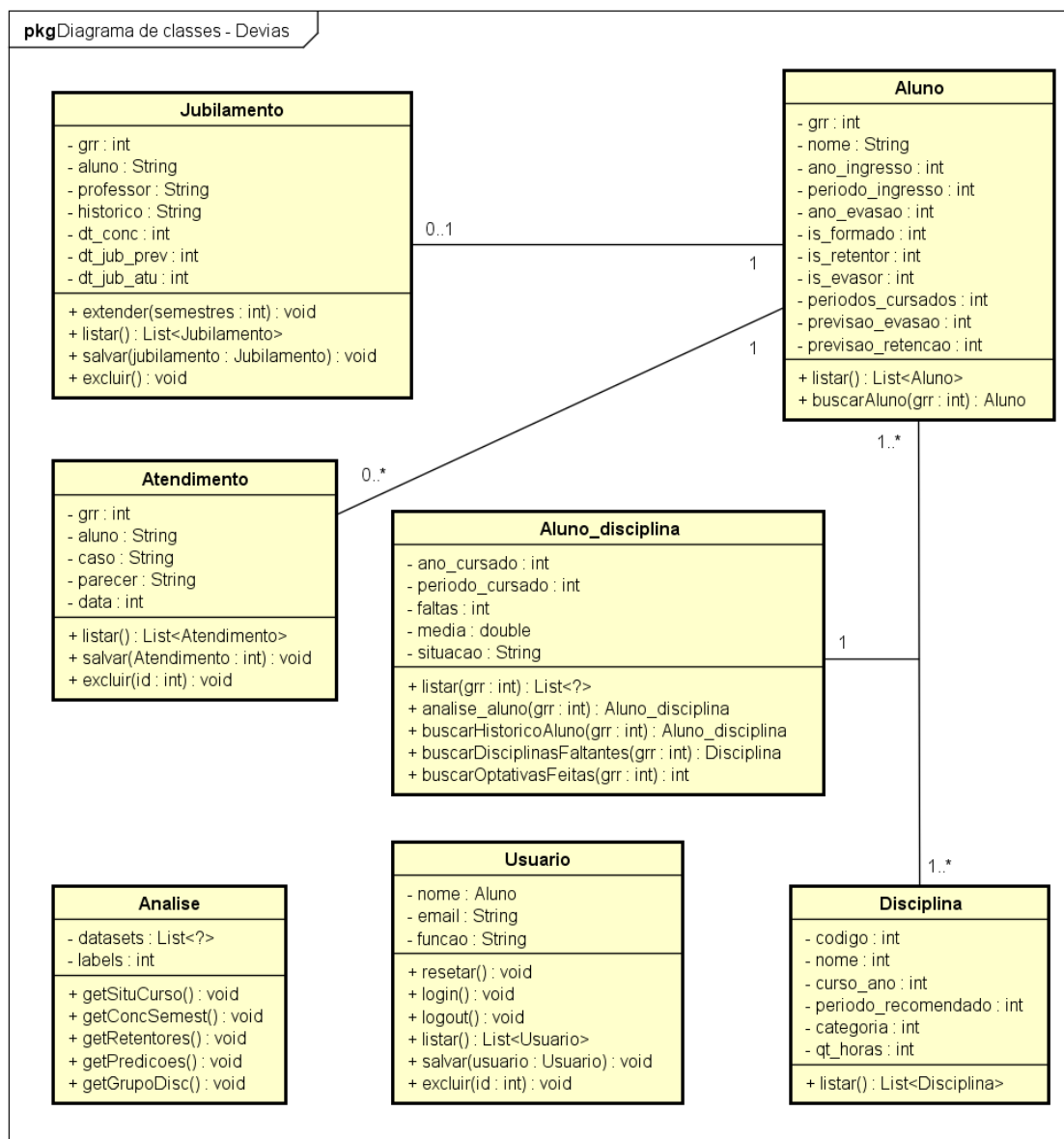
Novo Editar Salvar Cancelar Excluir

<b>FLUXO PRINCIPAL</b>
<ol style="list-style-type: none"> <li>1. O sistema apresenta a tela. (DV1)</li> <li>2. O sistema realiza a busca dos usuários na base de dados e mostra as informações em tela.</li> <li>3. O usuário pressiona o botão “Novo”. (A1, A2)</li> <li>4. O sistema alterna o comportamento da tela para o modo de inclusão.</li> <li>5. O usuário insere as informações.</li> <li>6. O usuário pressiona o botão “Salvar”.</li> <li>7. O sistema consiste as informações. (E1)</li> <li>8. O sistema faz a inclusão do usuário na base de dados.</li> <li>9. O caso de uso é encerrado.</li> </ol>
<b>FLUXO ALTERNATIVO – Botão “Editar” (A1)</b>
<ol style="list-style-type: none"> <li>1. O usuário seleciona um registro.</li> <li>2. O usuário pressiona o botão “Editar”.</li> <li>3. O sistema alterna o comportamento da tela para o modo de edição.</li> <li>4. O usuário altera as informações.</li> <li>5. O usuário pressiona o botão “Salvar”.</li> <li>6. O sistema consiste as informações. (E1)</li> <li>7. O sistema faz a alteração do usuário na base de dados.</li> <li>8. O caso de uso é encerrado.</li> </ol>
<b>FLUXO ALTERNATIVO – Botão “Excluir” (A2)</b>
<ol style="list-style-type: none"> <li>1. O usuário seleciona um registro.</li> <li>2. O usuário pressiona o botão “Editar”.</li> <li>3. O sistema faz a exclusão do atendimento na base de dados.</li> <li>4. O caso de uso é encerrado.</li> </ol>
<b>FLUXO DE EXCEÇÃO – Dados inválidos (E1)</b>
<ol style="list-style-type: none"> <li>1. O sistema emite a mensagem “Por favor, insira corretamente os dados”.</li> </ol>

## APÊNDICE D – DIAGRAMA DE CLASSES

A Figura 18 descreve o Diagrama de Classes de análise do sistema Devias.

FIGURA 18 – DIAGRAMA DE CLASSES



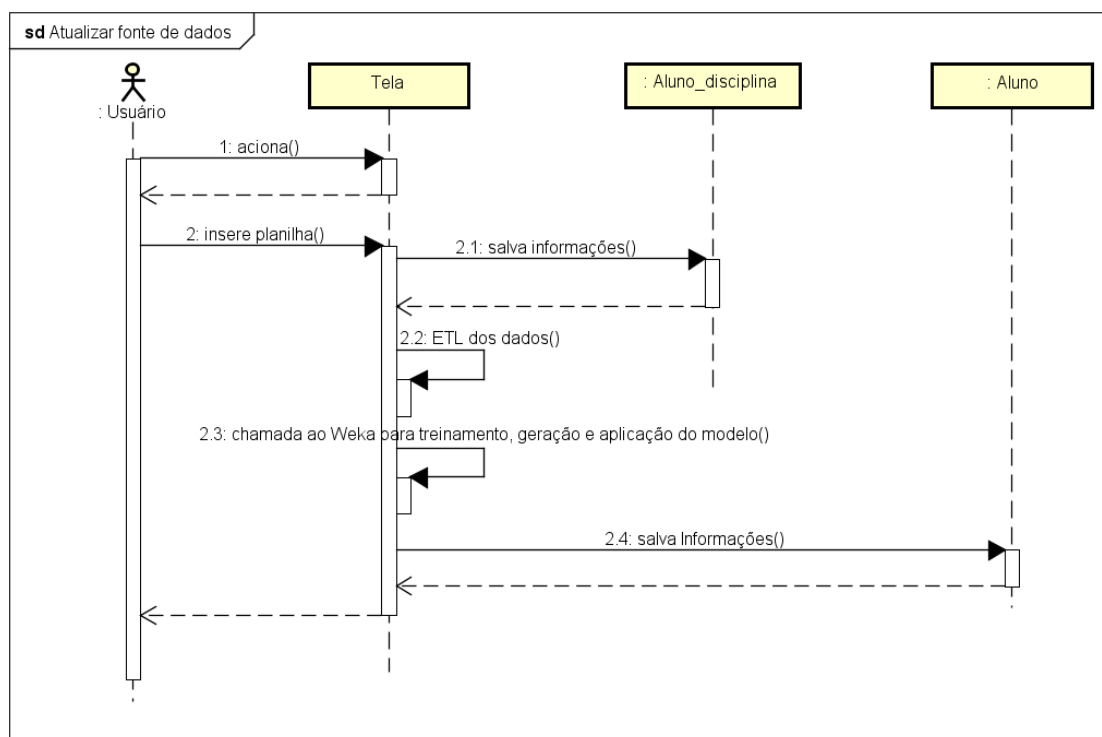
Fonte: Os autores (2018).

## APÊNDICE E – DIAGRAMA DE SEQUÊNCIA

### 1 – ATUALIZAR FONTE DE DADOS

A Figura 19 demonstra o diagrama de sequência do caso de uso “Atualizar Fonte de Dados”.

FIGURA 19 – DIAGRAMA DE SEQUÊNCIA: ATUALIZAR FONTE DE DADOS



Fonte: Os autores (2018).

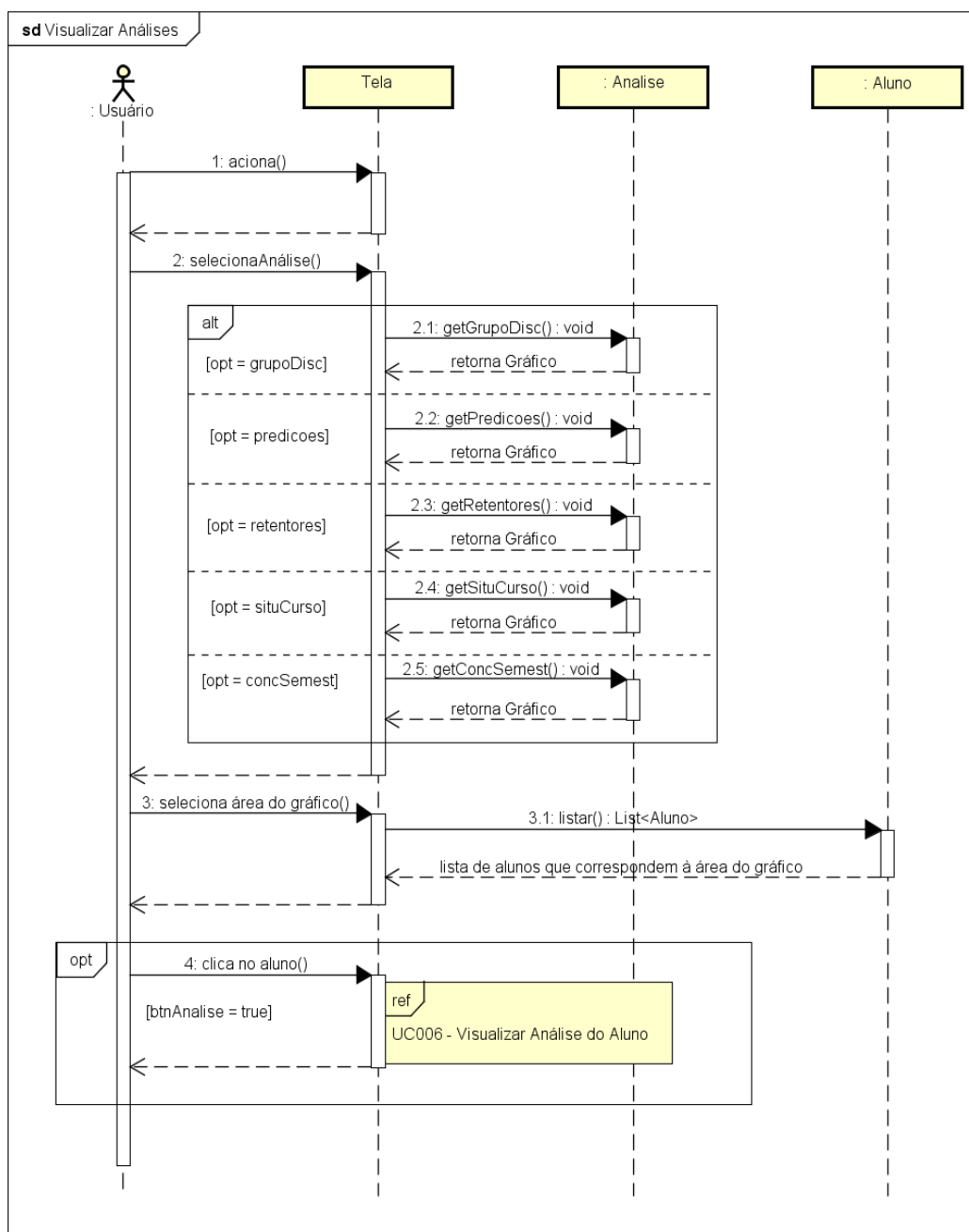
O diagrama inicia-se com o usuário acionando a tela de Configurações, logo como resposta o sistema devolve ao usuário o controle com a tela de Configurações em foco. O usuário pode a partir desse momento inserir o diretório da planilha de importação através do seletor de arquivos disposto na tela. Selecionado a planilha e pressionando o botão “Enviar”, o sistema inicia o processo de importação, primeiramente lendo os registros do csv e inserindo-os diretamente no banco de dados para a tabela correspondente à classe Aluno\_disciplina. Com a confirmação do sucesso da inserção, o sistema realiza algumas operações de banco para tratamento de alguns dados já esperados, como sumarização e enriquecimento, passando então para o momento de realizar a chamada externa ao Weka. O Weka por sua vez, recebe os dados e prontamente realiza o treinamento com o algoritmo de escolha, dando origem ao

modelo de predição. A aplicação do modelo seguida da inserção dos novos dados gerados na tabela de Alunos concluem o fluxo de importação com uma mensagem de sucesso.

## 2 – VISUALIZAR ANÁLISES

A Figura 20 demonstra o diagrama de sequência do caso de uso “Visualizar Análises”.

FIGURA 20 – DIAGRAMA DE SEQUÊNCIA: VISUALIZAR ANÁLISES



Fonte: Os autores (2018).

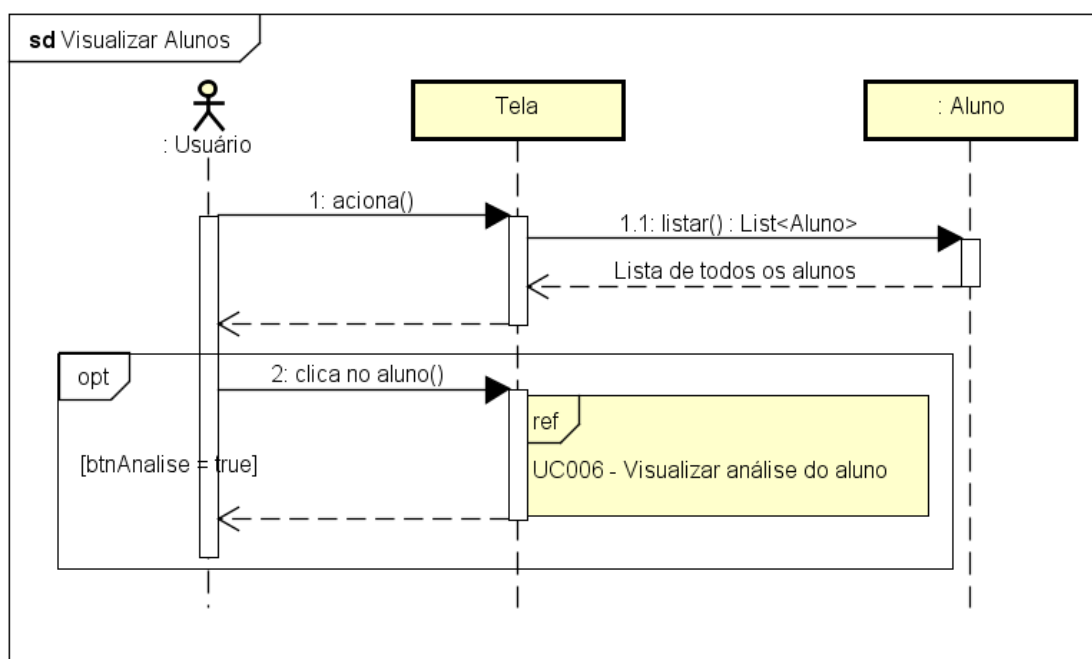
O diagrama inicia-se com o usuário acionando a tela de Configurações, logo como resposta o sistema devolve ao usuário o controle com a tela de Análises em foco. O usuário escolhe uma das análises dispostas na página e o sistema por sua vez, acessa o método da classe responsável por retornar as informações do gráfico daquela análise, para então o mostrar em tela.

O usuário então torna-se capaz de clicar sobre as informações do gráfico, o resultado deste feito é uma chamada à classe Aluno, que retorna uma lista de alunos que atendem pelo perfil proposto pela informação selecionada do gráfico.

### 3 – VISUALIZAR ALUNOS

A Figura 21 demonstra o diagrama de sequência do caso de uso “Visualizar Alunos”.

FIGURA 21 – DIAGRAMA DE SEQUÊNCIA: VISUALIZAR ALUNOS



Fonte: Os autores (2018).

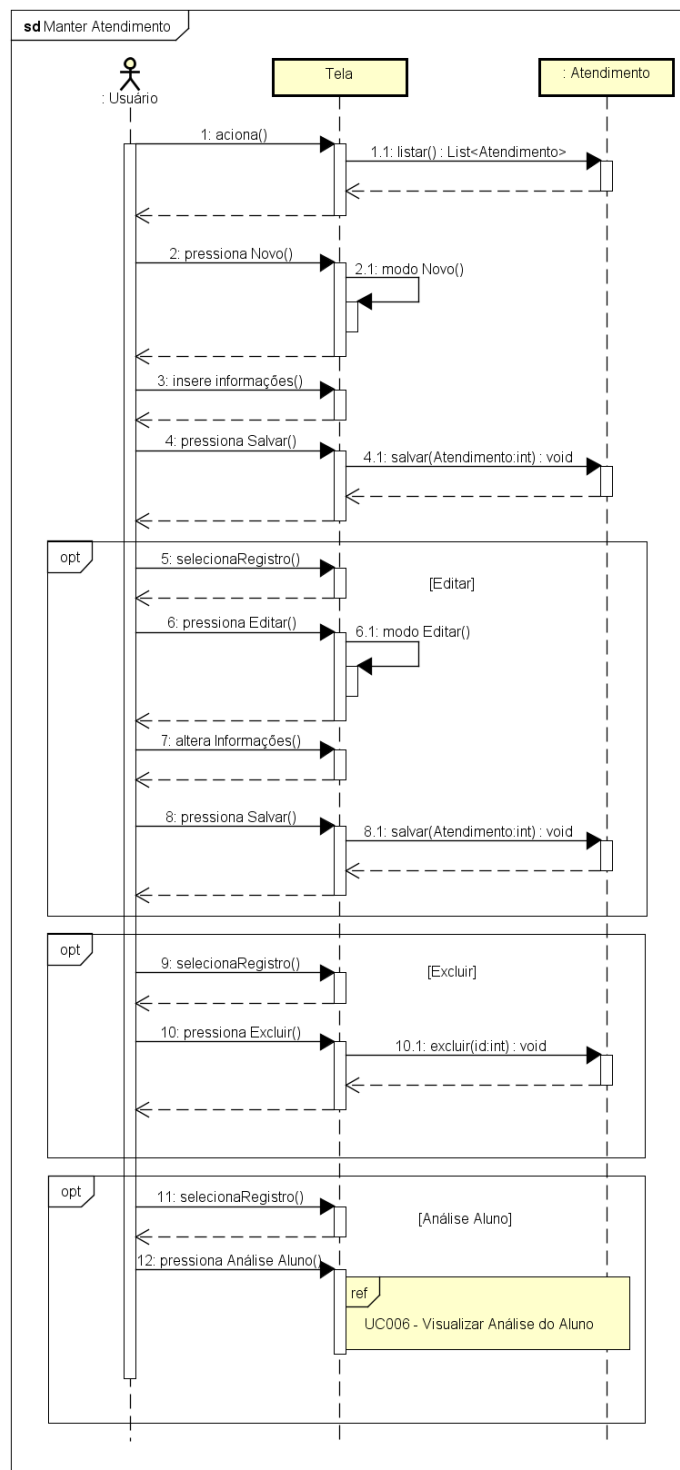
O diagrama inicia-se com o usuário acionando a tela de Configurações, o sistema então pede à classe Aluno para que retorne uma lista de todos os alunos e suas respectivas informações. Esta lista gerada é proposta em tela como uma tabela, podendo ser filtrada e pesquisada dentro de seus registros.

O usuário conta ainda com a opção de acessar a página de análise do aluno que conta com informações específicas do desempenho e situação do aluno.

#### 4 – MANTER ATENDIMENTO

A Figura 22 demonstra o diagrama de sequência do caso de uso “Manter Atendimento”.

FIGURA 22 – DIAGRAMA DE SEQUÊNCIA: MANTER ATENDIMENTO



Fonte: Os autores (2018).

O diagrama inicia-se com o usuário acionando a tela de Atendimentos, o sistema então pede à classe Atendimento para que retorne uma lista de todos os atendimentos e suas respectivas informações. Esta lista gerada é proposta

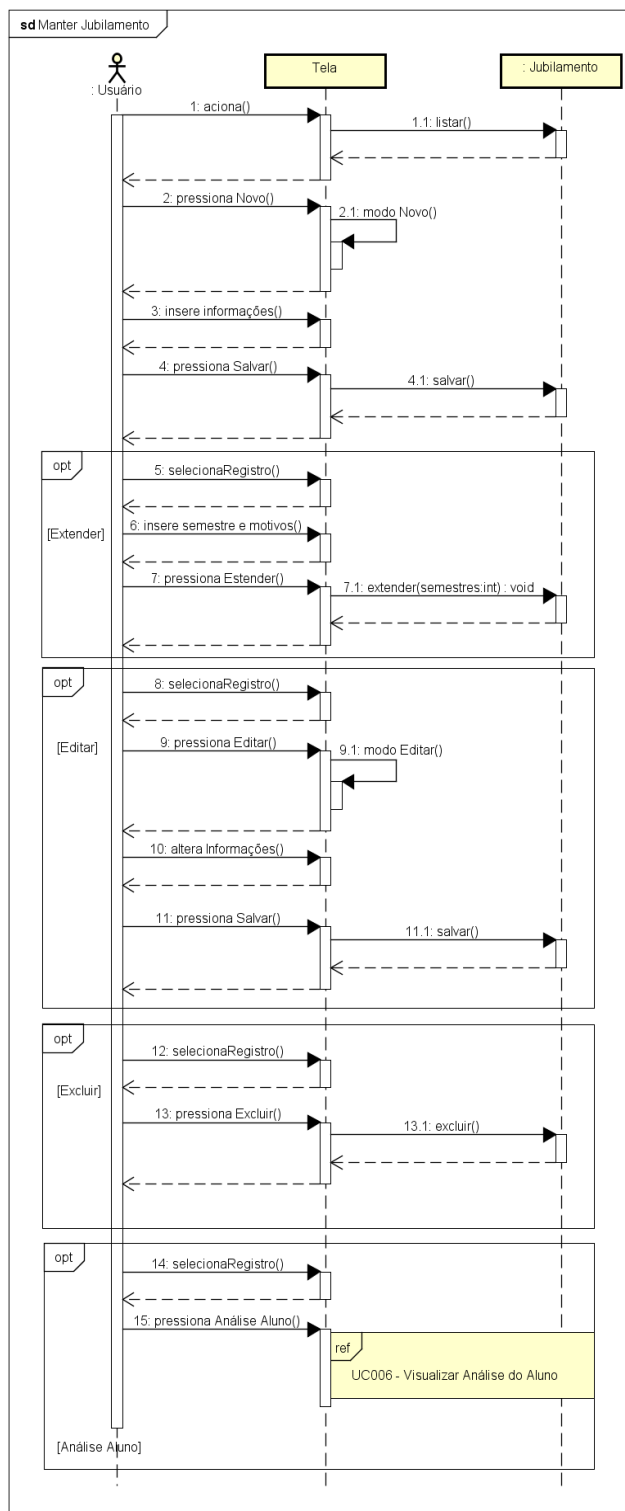
em tela como uma tabela, podendo ser filtrada e pesquisada dentro de seus registros.

O esquema proposto trata-se de um CRUD, com a pretensão de manter o controle sobre os atendimentos prestados aos alunos pelos professores. A tela ainda conta com a opção de acessar a página de análise do aluno para ter acesso às informações específicas do desempenho e situação do aluno do atendimento.

## 5 – MANTER JUBILAMENTO

A Figura 23 demonstra o diagrama de seqüência do caso de uso “Manter Jubilamento”.

FIGURA 23 – DIAGRAMA DE SEQUÊNCIA: MANTER JUBILAMENTO



Fonte: Os autores (2018).

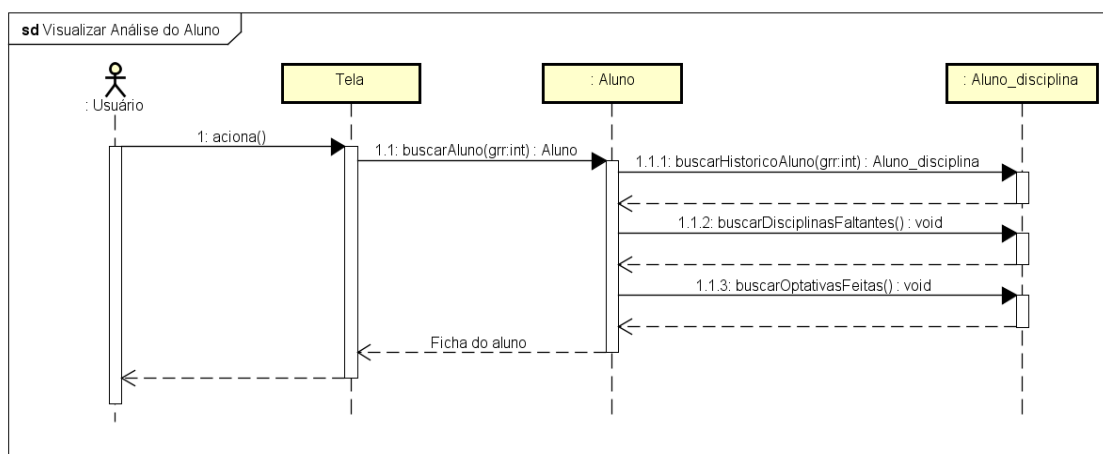
O diagrama inicia-se com o usuário acionando a tela de Jubilamentos, o sistema então pede à classe Atendimento para que retorne uma lista de todos os atendimentos e suas respectivas informações. Esta lista gerada é proposta em tela como uma tabela, podendo ser filtrada e pesquisada dentro de seus registros.

O esquema proposto trata-se de um CRUD, com a pretensão de manter o controle sobre os acompanhamentos de jubramento incididos sobre os alunos pelos professores. Além das funcionalidades padrões de consulta, inclusão, alteração e exclusão, é possível realizar a extensão do prazo de jubramento, propondo um motivo para o realizado. A tela ainda conta com a opção de acessar a página de análise do aluno para ter acesso às informações específicas do desempenho e situação do aluno do acompanhamento.

## 6 – VISUALIZAR ANÁLISE DO ALUNO

A Figura 24 demonstra o diagrama de sequência do caso de uso “Visualizar Análise Do Aluno”.

FIGURA 24 – DIAGRAMA DE SEQUÊNCIA: MANTER JUBILAMENTO



Fonte: Os autores (2018).

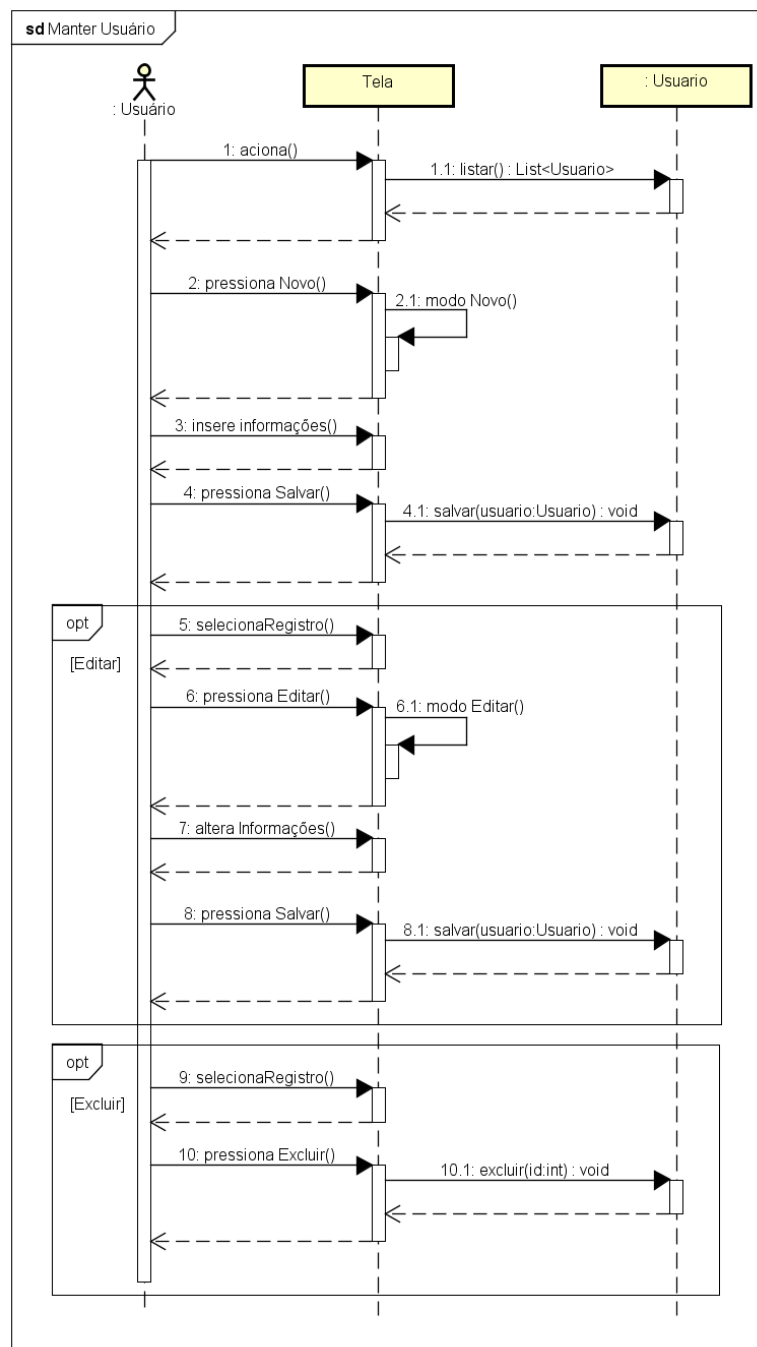
O diagrama inicia-se com o usuário acionando a tela de Análise do aluno, o sistema então pede à classe Aluno para que retorne o aluno correspondente ao GRR passado, juntamente com suas informações. Ainda não retornada a mensagem, o sistema faz ainda requisições à classe Aluno\_disciplina para que traga também o histórico, as disciplinas que faltam para o aluno concluir e as optativas que já concluiu. Com todas estas informações em memória, é montada

a ficha do aluno, e então retornada para a tela para preenchimento das lacunas e tabelas.

## 7 – MANTER USUÁRIOS

A Figura 25 demonstra o diagrama de sequência do caso de uso “Manter Usuários”.

FIGURA 25 – DIAGRAMA DE SEQUÊNCIA: MANTER USUÁRIOS



Fonte: Os autores (2018).

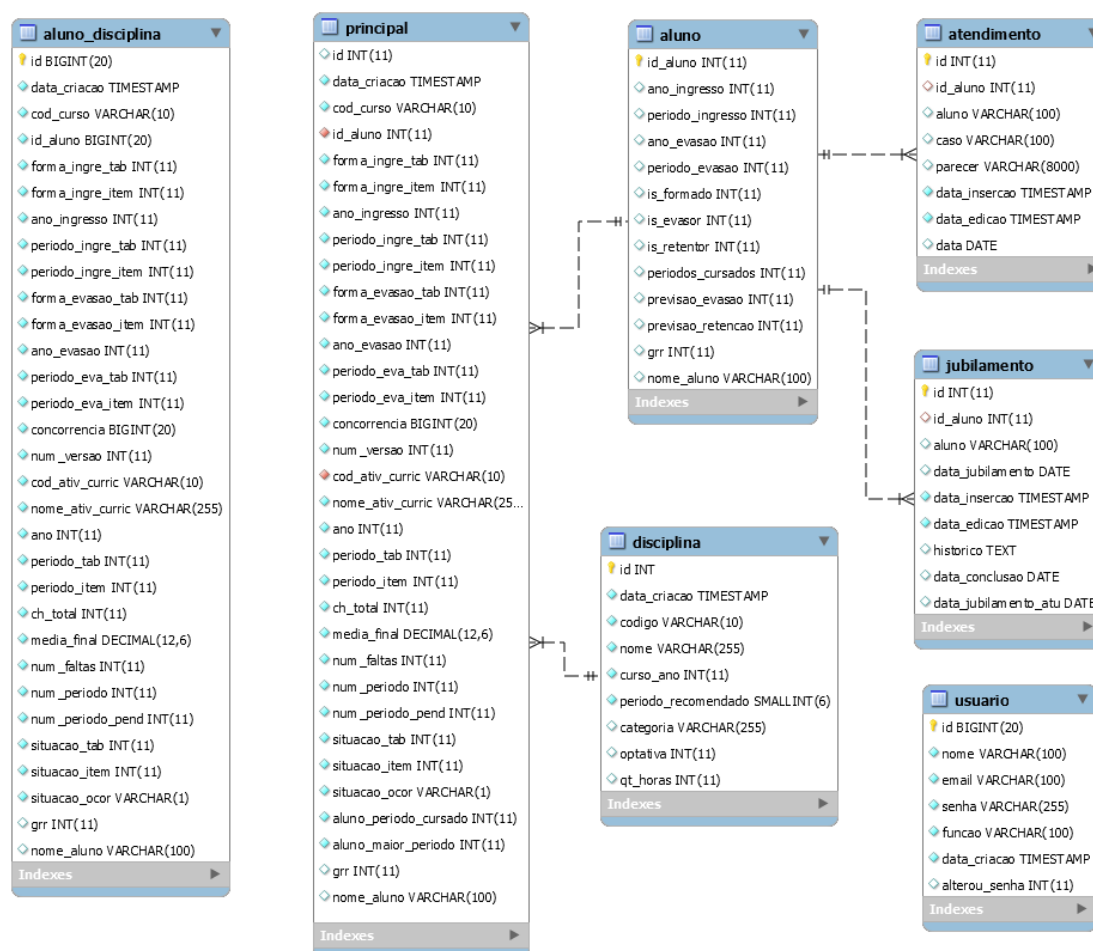
O diagrama inicia-se com o usuário acionando a tela de Usuários, o sistema então pede à classe Usuario para que retorne uma lista de todos os usuários e suas respectivas informações. Esta lista gerada é proposta em tela como uma tabela, podendo ser filtrada e pesquisada dentro de seus registros.

Realizado o devido controle de acesso sobre esta página, o esquema proposto trata-se de um CRUD, com a pretensão de controlar os usuários do sistema.

## APÊNDICE F – MODELO LÓGICO DE DADOS

Para a modelagem do banco de dados do sistema, foi optado por manter uma tabela com os dados brutos advindos da importação, a tabela “aluno\_diciplina”. Foram criadas novas tabelas para destrinchar as informações da tabela de importação, as tabelas “aluno”, “disciplina” e “principal”, de maneira que, pudessem resguardar também as novas informações geradas por meio do processo de enriquecimento, descrito no capítulo quatro. A Figura 26 demonstra o modelo lógico de dados do sistema Devias.

FIGURA 26 – MODELO LÓGICO DOS DADOS



Fonte: Os Autores (2018).