

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
COORDENAÇÃO DE LICENCIATURA EM INFORMÁTICA
DESENVOLVIMENTO DE SISTEMAS PARA INTERNET E DISPOSITIVOS
MÓVEIS**

PEDRO HENRIQUE DE ALENCAR MACHADO

**UM PROTÓTIPO DE DATA MINING APLICADO AO PROCESSO DE
GESTÃO EDUCACIONAL PARA DETECÇÃO DE EVASÃO ESCOLAR**

MONOGRAFIA DE ESPECIALIZAÇÃO

FRANCISCO BELTRÃO

2014

PEDRO HENRIQUE DE ALENCAR MACHADO

**UM PROTÓTIPO DE DATA MINING APLICADO AO PROCESSO DE
GESTÃO EDUCACIONAL PARA DETECÇÃO DE EVASÃO ESCOLAR**

Monografia de Especialização apresentada a Coordenação de Licenciatura em Informática, da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Especialista em Desenvolvimento de Sistemas para Internet e Dispositivos Móveis”.

Orientador: Ms. Paulo Júnior Varela

FRANCISCO BELTRÃO

2014



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Francisco Beltrão
Diretoria de Pesquisa e Pós-Graduação
Curso de Especialização em Desenvolvimento de Sistemas para Internet
e Dispositivos Móveis

TERMO DE APROVAÇÃO

Dia 11 do mês de dezembro de 2014 às 16 horas, no mini auditório do Câmpus Francisco Beltrão, realizou-se a apresentação pública da monografia pelo estudante Pedro Henrique de Alencar Machado intitulada “Um protótipo de data mining aplicado ao processo de gestão educacional para detecção de evasão escolar”. Finalizada a apresentação e arguição, a Banca Examinadora declarou **aprovada** a monografia do estudante, como requisito parcial para obtenção do título de Especialização em Desenvolvimento e Sistemas para Internet e Dispositivo Móveis.

Professor Paulo Júnior Varela - UTFPR
(Orientador)

Professor Flávio de Almeida e Silva - UTFPR
(Convidado)

Professor Dr. Ademir Roberto Freddo - UTFPR
(Convidado)

Professor Dr. Ademir Roberto Freddo - UTFPR
(Coordenação)

Ao meu pai
Miguel Angelo Machado
(*in memoriam*)

AGRADECIMENTOS

Não há dúvidas que as palavras contidas neste tópico, não serão suficientes para expressar a minha total gratidão quanto a todos os envolvidos nesta “conquista”, mas uma coisa cabe muito a este item, esta conquista é de cada um descrito nos parágrafos que se segue, sem mais delongas.

Aos meus pais, pela educação, pelo amor incondicional, apoio e confiança, sem as quais jamais idealizaria nada que foi um dia almejado.

A todos meus amigos distantes, em especial meu irmão José (Obina), pelos momentos de descontração, desordem e ciladas.

Aos meus amigos Mauro e Cassio, que permitiram que este trabalho se tornasse realidade, me apoiando até nos momentos de descanso e relaxo.

Ao meu professor orientador Sr. Ms. Paulo Júnior Varela, por quem demonstrei enorme admiração desde as primeiras aulas ministradas. Pela confiança no tema abordado, disposição em sempre ajudar e principalmente pelos conhecimentos transmitidos durante todo o processo.

A minha amiga e noiva Juliéli, pelos momentos de refúgio, conversas e silêncio. Por sua conta aprendi o quanto tudo isso vale a pena.

À Deus, por ter me permitido a vida, e pessoas tão especiais. Graças a ti, estou aqui.

*“Não sabendo que era impossível, ele foi
lá e fez”.*
Jean Cocteau.

RESUMO

A necessidade de gestão da aprendizagem é um objetivo cada vez mais importante, tendo em vista os problemas encontrados em nosso país em prover uma educação eficiente e de qualidade. Dentre esses problemas podemos mencionar alguns, como por exemplo: Pouco investimento em tecnologia de gestão, sobrecarga de trabalho enfrentada pelos professores, falta de interesse dos alunos sobre temas abordados, entre tantos outros. Tendo em vista a gama de informações que as Instituições de Ensino Superior têm em suas bases de dados, uma análise dos dados pode prover mecanismos que auxiliem docentes e gestores em tomadas de decisões mais assertivas. Este trabalho descreve a criação de um protótipo *web* de mineração de dados que visa detectar a evasão escolar de alunos do curso de Licenciatura em Informática da Universidade Tecnológica Federal do Paraná – Campus Francisco Beltrão. Uma aplicação *web* foi criada, cuja foi preparada para receber parâmetros de um aluno e calcular a probabilidade do referente tornar-se um aluno evasivo.

Palavras-chave: Mineração de dados. Educação. Protótipo.

ABSTRACT

The need for learning management is an increasingly effulgent goal, in view of the problems encountered in our country to provide an efficient and quality education. Among these we can mention some problems, such as: Little investment in technology management, workload faced by teachers, lack of interest from students about topics covered, among many others. Considering the range of information that higher education institutions have in their databases, an analysis of data can provide mechanisms to assist teachers and administrators in taken more assertive decisions. This paper describes the creation of a prototype web data mining that aims to detect the school dropout students of Bachelor in Informatics of the Federal Technological University of Paraná - Campus Francisco Beltrão. A web application was created, which was prepared to receive parameters of a student and to calculate the probability of referring to become an elusive.

Keywords: Data mining. Education. Prototype.

LISTA DE ILUSTRAÇÕES

Figura 1 - Etapas do KDD	7
Figura 2 - Diagrama de Caso de Uso	16
Figura 3 - Diagrama de Classe.....	17
Figura 4 - Analisar probabilidade.....	19
Figura 5 - DER Alunos	25
Figura 6 - Situação dos Alunos	26
Figura 7 - Código para recuperar planilha xls.....	27
Figura 8 - Código média de acontecimento.....	27
Figura 9 - Código classificação dos alunos	28
Figura 10 - Design responsivo do protótipo.....	33
Figura 11 - Formulário	33
Figura 12 - Método preparaAluno()	34
Figura 14 - Criação da sessão para apresentação do resultado.....	35
Figura 15 - Gráfico de probabilidade	35

LISTA DE TABELAS

Tabela 1 - Métodos implementados na WEKA.....	11
Tabela 2 - Resultado da classificação questionário sócio educacional.	30

LISTA DE SIGLAS E ABREVIATURAS

API	<i>Application Programming Interface</i>
ARFF	<i>Attribute Relation File Format</i>
CSS	<i>Cascading Style Sheets</i>
DCU	Diagrama de Caso de Uso
DER	Diagrama de Entidade Relacionamento
GPL	<i>General Public Licence</i>
HTML	<i>HyperText Markup Language</i>
IDE	<i>Integrated Development Environment</i>
IES	Instituição de Ensino Superior
JDBC	<i>Java DataBase Connectivity</i>
JSF	<i>JavaServer Faces</i>
KDD	<i>Knowledge Discovery in Databases</i>
MD	Mineração de Dados
MVC	<i>Model View Controller</i>
PIB	Produto Interno Bruto
SGBD	Sistema Gerenciador de Banco de Dados
SQL	<i>Structured Query Language</i>
TI	Tecnologia da Informação
UML	<i>Unified Modeling Language</i>
UTFPR	Universidade Tecnológica Federal do Paraná
WEKA	<i>Waikato Environment for Knowledge Analysis</i>

SUMÁRIO

1. INTRODUÇÃO	1
1.1 OBJETIVOS.....	2
1.2 JUSTIFICATIVA.....	3
1.3 ESTRUTURA DO TRABALHO.....	4
2. REFERENCIAL TEÓRICO	5
2.1 EVASÃO ESCOLAR NAS IES.....	5
2.2 MINERAÇÃO DE DADOS.....	6
2.3 TAREFAS DE MINERAÇÃO DE DADOS.....	7
2.4 WEKA.....	9
3. MATERIAIS E MÉTODOS	12
3.1 MATERIAIS.....	12
3.2 MÉTODOS.....	15
4. RESULTADOS	21
4.1 REQUISITOS.....	21
4.2 ANÁLISE.....	22
4.3 EXPERIMENTOS.....	23
4.4 DESENVOLVIMENTO.....	25
5. CONCLUSÃO	37
6. REFERÊNCIAS	39
ANEXO A	44

1. INTRODUÇÃO

Uma pesquisa realizada pela UNESCO (2014) destacou uma problemática relacionada ao número de analfabetos adultos no Brasil. A pesquisa coloca o país como 8º (dentre 150 países avaliados) com o maior índice de analfabetos adultos. A pesquisa destaca como uma das prováveis causas do número anteriormente citado, a má qualidade da educação, falta de preparação dos docentes, entre outras causas.

A pesquisa ainda destaca que menos de 10% dos professores brasileiros, na época do exame, estavam fazendo algum curso de preparação. Das 150 nações participantes da pesquisa, somente 25% delas aplicaram no ano anterior 6% do PIB (Produto Interno Bruto) na área da educação, cuja é uma intenção já anteriormente apresentada pela própria UNESCO.

Para a obtenção de sucesso nos desafios do analfabetismo, da evasão escolar, entre outras problemáticas ligadas à educação brasileira, o uso da tecnologia da informação pode ser usado como mecanismo de auxílio, como por exemplo, nas seguintes atividades: Agilidade na análise dos dados, armazenamento e centralização das informações, interatividade dos alunos com os educadores, automação de tarefas e processos, dentre tantas outras.

Como ferramenta de auxílio na análise de dados está a utilização de técnicas de mineração de dados para extrair as informações implícitas existentes em grandes bases de dados das organizações em questão. Lucas (2002) acredita que:

“A transformação da informação da informação em conhecimento pode fazer com que as grandes organizações sobrevivam neste mercado globalizado, pois esta transformação fornecerá informações que após serem analisadas de forma correta passam a ser utilizadas para tomadas de decisões mais seguras.”

Um dos atuais desafios das Instituições de Ensino Superior é organizar os dados e transformá-los em informações úteis. O uso de tecnologias computacionais podem auxiliar gestores e educadores a obterem uma vantagem competitiva sustentável.

A extração das informações previamente destacadas, segundo Fayyad (1996) denomina-se *Knowledge Discovery in Databases* (KDD) ou então descoberta de conhecimento em base de dados. Ainda segundo o autor, por mais comum que seja usar os termos KDD e Mineração de Dados com o mesmo significado, o mesmo define KDD como sendo o processo de extração de conhecimento dos dados como

um todo, e Mineração de Dados, como apenas uma etapa particular do KDD, sendo que nesta etapa a extração de padrões dos dados é realizada através do uso de algoritmos específicos.

Descobrir o conhecimento nas mais diversas bases de dados de diferentes instituições é o objetivo da tarefa de Mineração de Dados, além de permitir uma maior agilidade no processo de tomadas de decisões por parte dos gestores responsáveis. O armazenamento de dados por si só não traz nenhuma contribuição, é necessário realizar uma análise sobre os elementos previamente armazenados e a partir daí criar informações relevantes que podem auxiliar em tomadas de decisões. Chiara (2003) destaca que: “Para a aplicação de técnicas de mineração de dados é necessário que se tenha uma coleção de dados disponível. Entretanto, o problema é conseguir dados relevantes para extrair deles conhecimento potencialmente útil”.

Tendo em vista todas as informações anteriormente apresentadas, o presente trabalho propõe o desenvolvimento e aplicação de técnicas de mineração de dados sobre informações, previamente coletadas, de alunos de diferentes períodos e situações do curso de Licenciatura em Informática da Universidade Tecnológica Federal do Paraná – Campus Francisco Beltrão, a fim de promover dados e possíveis estratégias para auxílio nas tomadas de decisões por parte dos seus gestores.

1.1 OBJETIVOS

A seguir serão apresentados os objetivos gerais e específicos do trabalho em questão.

1.1.1 OBJETIVOS GERAIS

Aplicar técnicas de mineração de dados sobre uma base de dados fornecida pela IES UTFPR – Campus Francisco Beltrão e outra criada com base em um questionário aplicado em alunos do curso de Licenciatura em a fim de classificar e prever a evasão.

1.1.2 OBJETIVOS ESPECÍFICOS

Para a concretização do objetivo geral, elencam-se a seguir os seguintes objetivos específicos:

- a. Aplicação de um questionário sócio educacional (Anexo A) sobre alunos que estão cursando Licenciatura em Informática na IES UTFPR – Campus Francisco Beltrão;
- b. Levantar junto a IES informações de alunos que estão cursando e que estão em situação desistentes do curso de Licenciatura em Informática na IES UTFPR – Campus Francisco Beltrão;
- c. Estudar e configurar a ferramenta WEKA para utilizá-la na aplicabilidade de técnicas de mineração de dados nas bases de informações previamente criadas;
- d. Tratar as informações coletadas e criar uma base de dados que será utilizada nos testes;
- e. Estudar e aplicar o algoritmo Naive Bayes sobre as bases de dados criadas, a fim de identificar e classificar os registros armazenados;
- f. Construir uma aplicação capaz de receber uma nova instância de um aluno e classificá-lo nas bases de dados criadas;

1.2 JUSTIFICATIVA

A fim de permanecerem competitivas no mercado moderno, as IES devem considerar suas atividades como quaisquer outras exercidas dentro do âmbito empresarial. Segundo Furtado (2004), desta forma, as instituições podem ser impactadas pelas mesmas situações que afligem os mercados:

“[...]o setor educacional vem atraindo um número crescente de novos atores e o mercado educacional de novos integrantes, que passam a disputar o domínio deste mercado com as instituições tradicionais. Estas, por sua vez, vêem-se forçadas a rever suas práticas e métodos até então utilizados, como condição para que possam continuar tendo relevância em seus serviços prestados e que sobrevivam em um cenário altamente competitivo”.

Frente às argumentações anteriormente apresentadas, justifica-se da seguinte forma a relevância ao tema proposto: a crescente importância da gestão da informação em todas as organizações e a pouca disponibilidade de estudos e pesquisas voltadas para esta área tendo como foco a aplicação deste conceito dentro

do âmbito de uma IES. Além disto, é importante ressaltar o que fora previamente lícitado no item 1 deste trabalho quanto a problemática do analfabetismo no Brasil e o que está sendo planejado para resolução e amenização desta, torna de suma importância prover técnicas que auxiliem os gestores em tomadas de decisões a fim de trabalharem de maneira mais específica e minimizar o número de possíveis egressos.

1.3 ESTRUTURA DO TRABALHO

O trabalho está dividido em 5 capítulos. O primeiro apresenta a introdução, objetivos e justificativas para a realização deste trabalho. O Capítulo 2 apresenta uma contextualização teórica sobre os conceitos e características das técnicas de mineração de dados. O Capítulo 3 apresenta os materiais e métodos utilizados para a realização deste trabalho. O Capítulo 4 apresenta os resultados obtidos, que se resume aos resultados obtidos com a aplicação do algoritmo e a aplicação de manipulação e apresentação de novas instâncias de alunos ingressos, dentre outros recursos. Finalizando, no capítulo 5 são apresentadas as conclusões, dificuldades encontradas e as sugestões de trabalhos futuros.

2. REFERENCIAL TEÓRICO

Este capítulo apresenta o referencial teórico do trabalho em questão, todas as técnicas e tecnologias utilizadas como parâmetro para a construção deste trabalho. A fundamentação teórica serve como base para a fundamentação da pesquisa, neste item apresenta-se características sobre a Mineração de Dados com suas técnicas e tarefas, ferramentas para *data mining*, a API WEKA, utilizado no sistema web desenvolvido para a detecção da evasão dos alunos do curso de Licenciatura em Informática da IES UTFPR – Campus Francisco Beltrão.

2.1 A EVASÃO ESCOLAR NAS IES

De acordo com Gaioso (2005), a evasão escolar é um problema social que causa, além de outros, uma interrupção no ciclo de estudos. “É um problema que vem preocupando as instituições de ensino em geral, pois a evasão de alunos provoca graves consequências sociais, acadêmicas e econômicas”. (BAGGI, 2010).

Alunos desistentes fazem com que certos investimentos sejam aplicados de maneira errada, levando em consideração que o discente não proverá nenhum dos retornos esperados. Dentre tantas, esta é uma das mais agravantes problemáticas da evasão escolar (LEAL, 2007).

Hipólito (2009) descreve que, em 2009, somente 8% da população brasileira adulta tem formação superior, em controversa com outros países: Coreia, 32%, Espanha, 28%, Rússia, 55% e Chile, 13%. (HIPÓLITO, 2009).

De acordo com uma pesquisa realizada pelo Inep em 2006, o percentual de evasão anual de alunos de cursos de Ciências, Matemáticas e Computação tem uma taxa média de evasão em torno dos 28%, consideravelmente acima da média nacional abaixo apenas da evasão anual média registrada pelos cursos de Serviços (LEAL, 2007).

2.2 MINERAÇÃO DE DADOS

A mineração de dados provê maneiras e métodos automatizados para facilitar a descoberta de padrões em dados, sem a tendenciosidade e a limitação de uma análise baseada na instituição humana. (BRAGA, 2005).

Devido ao avanço da tecnologia da informação e o crescimento da Internet, as organizações são capazes de coletar e armazenar enorme quantidade de dados. Pessoas, gradualmente, estão percebendo que os dados não são iguais à informação, os devem ser analisados e extraídos para torná-lo informação. A alta capacidade de armazenamento de dados têm dificultado a capacidade de interpretá-los e convertê-los em informações de apoio a tomada de decisões (PASTA, 2011).

A relação entre grandes massas de dados e, posteriormente a descoberta de informações e novos conhecimentos, podem ser úteis para realizar atividades de tomada de decisão. (BAKER, 2011).

Graças ao avanço tecnológico e a alta oferta de recursos tecnológicos, corporações estão armazenando em banco de dados milhares de informações, que aliado às aplicações, serve de base para tomadas de decisões (SHIBA, 2008).

Somente dados armazenados em grandes bases de dados, por si só, não representam nenhum tipo de informação, a atividade de descobrir conhecimento em banco de dados denomina-se *Knowledge Discovery in Databases* – KDD (Descoberta de Conhecimento em Bases de Dados) (SILVA, 2004).

Segundo Fayyad (1996) KDD é a extração de conhecimento dos dados previamente armazenados, onde uma das etapas é a mineração de dados.

Um dos principais desígnios da disciplina de Mineração de Dados é descobrir os conhecimentos em grandes bases de dados de diferentes corporações por meio da aplicação de algoritmos específicos.

Segundo Oliveira e Rosa (2007):

“A descoberta de conhecimento em bases de dados é multidisciplinar e, historicamente, se origina de diversas áreas, dentre as quais podem ser destacadas a estatística, inteligência computacional, reconhecimento de padrões e banco de dados”.

Os autores Goldschmidt e Passos (2005) agrupam as atividades do KDD em três distintos grupos, cujos são ilustrados na Figura 1:

- Desenvolvimento tecnológico;

- Execução de KDD;
- Aplicação dos resultados.

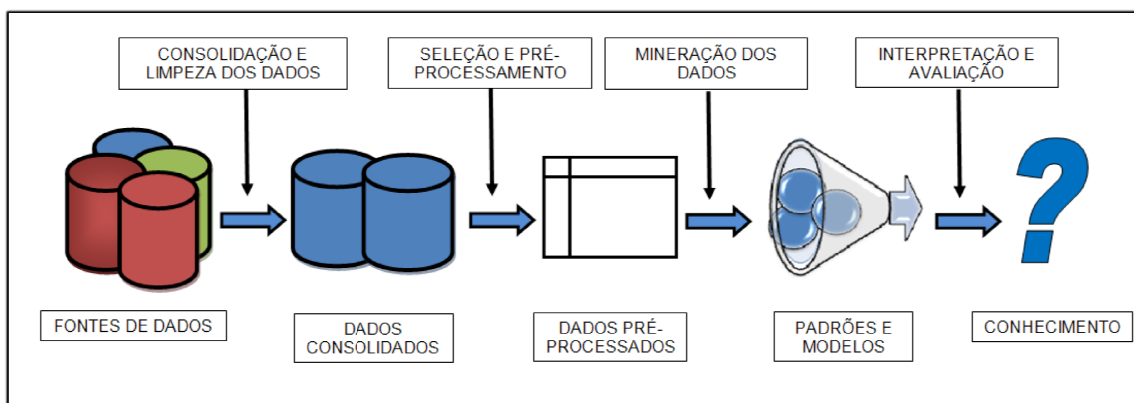


Figura 1 - Etapas do KDD Fonte: Pasta apud Figueira (1988)

Cada etapa do processo de Mineração de Dados é entendida como um sub processo do KDD. “Eles consistem em uma preparação mais refinada dos dados provenientes das etapas anteriores” (HORST E MONARD, 2000):

- **Consolidação dos dados:** coleta e consolidação dos dados que dará início ao processo de extração do conhecimento.
- **Seleção e pré-processamento:** melhora na qualidade e transformação dos dados, a fim de evitar possíveis distorções na extração do conhecimento.
- **Mineração de dados:** definição de quais algoritmos serão utilizados na extração do conhecimento.
- **Interpretação e avaliação:** análise dos resultados obtidos para o julgamento do modelo obtido da fase anterior. Nesta etapa também se busca criar uma forma de interpretar os resultados visando à leitura direta dos mesmos.

2.3 TAREFAS DE MINERAÇÃO DE DADOS

De acordo com o objetivo esperado, diferentes tarefas de MD podem ser realizadas. Tarefa de Mineração de Dados é o modo como as informações serão mineradas (PASTA, 2011).

As tarefas possíveis de um algoritmo de extração de padrões podem ser agrupadas em atividades descritivas e preditivas. Classificação e Regressão são consideradas tarefas de atividade preditiva, enquanto as atividades de Associação, Clusterização e Sumarização são as principais atividades descritivas. (GALVÃO, 2009).

2.3.1 CLASSIFICAÇÃO

A tarefa de classificação consiste na criação de classes previamente definidas de acordo com suas semelhanças, preconiza Pasta (2011).

Essa tarefa determina um conjunto de classes (padrões) que podem ser usadas para classificar novos objetos. “A tarefa de classificação busca uma função que permite associar corretamente cada registro de um banco de dados a um único rótulo categórico de chamado de classe” (RABELO, 2007).

Os mais conhecidos algoritmos utilizados na tarefa de classificação são: Redes Neurais, *Back-Propagation*, Classificadores Bayesianos e Algoritmos Genéticos.

2.3.2 REGRESSÃO

Por lidar com resultados contínuos, esta tarefa pode ser utilizada como uma tarefa de classificação, estabelecendo-se que diferentes faixas de valores correspondem a diferentes classes (PASTA, 2011).

2.3.3 ASSOCIAÇÃO

A tarefa descritiva de Associação consiste em identificar a relação entre atributos. Por exemplo, um cliente que adquire um produto X, em N% de vezes, adquire o produto Y também (CAMILO, SILVA, 2009).

“Os seus algoritmos procuram por relações entre os itens das transações, analisando os que ocorrem simultaneamente, dando possibilidade de entendimento de novos modelos” (PASTA, 2011).

2.3.4 CLUSTERIZAÇÃO OU SEGMENTAÇÃO

A tarefa descritiva de clusterização ou segmentação procura identificar a proximidade e similaridade entre registros. “Um *cluster* é uma coleção de registros similares entre si.” (CAMILO, SILVA, 2009).

Seu objetivo principal é fazer a partição da base de dados em um número determinado de *clusters*, nos quais as instâncias destes *clusteres* sejam similares. (PASTA, 2011).

2.3.5 SUMARIZAÇÃO

A tarefa descritiva de sumarização tem por objetivo a identificação e apresentação das principais características dos dados, de forma concisa e compreensível (PASTA, 2011).

A sumarização visa identificar e indicar as características comuns entre um conjunto de dados. Esta tarefa é aplicada nos clusters obtidos na tarefa de clusterização ou segmentação. As técnicas de sumarização são, na maior parte das vezes, aplicadas à análise exploratória de dados e à geração automática de relatórios (GALVÃO, 2009).

2.4 WEKA

Waikato Environment for Knowledge Analysis WEKA é uma ferramenta que contempla algoritmos de preparação de dados, mineração e validação de resultados (SILVA, 2004).

O *software* WEKA tem sido bastante utilizado no meio acadêmico em pesquisas que envolvam a área de MD. Sua escolha se justifica por causa de sua ampla aplicabilidade – já que lida com atributos numéricos (reais e inteiros), nominais e caracteres (*string*) (PASTA, 2011).

O WEKA foi desenvolvido por universitários da Universidade de *Waikato*, na Nova Zelândia, no ano de 1999 e sua licença é *General Public Licence* (GPL), o que significa que é um programa de distribuição livre (SILVA, 2004).

Este software é formado por um conjunto de algoritmos que implementam várias técnicas que são utilizadas para a resolução de problemas reais de MD. O WEKA foi desenvolvido na linguagem Java, cuja principal característica é sua portabilidade, assim sendo podendo ser executado em diversas plataformas, dentre as quais, Windows, MAC Os X e Linux. O único requisito é que o computador possua a máquina virtual Java instalada (MORATE, 2010).

O *software* WEKA é composto por dois pacotes: um pacote autônomo, para manipulação direta dos algoritmos, usando o formato de dados próprio, e um pacote de classes em Java que implementam estes algoritmos. Nessa segunda forma, é possível desenvolver uma aplicação em linguagem Java que faça uso destes algoritmos e aplicá-los em quaisquer bancos de dados através de uma conexão JDBC (*Java DataBase Connectivity*) (PASTA, 2011).

Alguns métodos implementados no WEKA, são:

Classificação	Predição numérica	Agrupamento	Associação
Árvore de decisão induzida	Regressão linear	EM	<i>Apriori</i>
Regras de aprendizagem	Geradores de arvores modelo	<i>Cobweb</i>	FPGrowth
<i>Naive Bayes</i>	Regressão local de pesos	<i>SimpleKMeans</i>	<i>PredictiveApriori</i>
Tabelas de decisão	Aprendizado baseado em instância	<i>DBScan</i>	<i>Tertius</i>
Regressão local de pesos	Tabela de decisão	CLOPE	
Aprendizado baseado	<i>Perceptron</i>		

em instância	multicamadas		
Regressão logica			
<i>Perceptron</i>			
<i>Perceptron</i> multicamada			
Comitê de <i>perceptrons</i>			
SVM			

Tabela 1 - Métodos implementados na WEKA
Fonte: Adaptado de Pasta (2011)

3 MATERIAIS E MÉTODOS

Este capítulo apresenta as ferramentas, técnicas e tecnologias utilizadas para a elaboração e construção deste trabalho. Neste escopo é descrita a metodologia utilizada para atingir o resultado final bem como artefatos de modelagem construídos para auxiliar na construção da ferramenta.

3.1 MATERIAIS

Os materiais e ferramentas usadas para desenvolver a aplicação *data mining* aplicado ao processo de gestão educacional para detecção de evasão escolar foram:

a) **JavaServer Faces (JSF):** De acordo com Melo (2010), JSF é um *framework* de componentes para desenvolvimento web em Java. Essa definição foi definida por meio do *Java Community Process (JSP)*, o que significa que grandes empresas como Apache, BEA Systems, IBM, Oracle, entre outras, aprovam as suas especificações, isso torna o JSF imediatamente um padrão de mercado. A utilização do JSF, entre outras, permite, também, a separação entre a lógica, a regra de negócio e a navegação entre os componentes da interface, seguindo o modelo MVC (*Model View Controller*).

No JSF a *controller* fica por conta de um *servlet*, chamado *Faces Servlet*, cujo é destinado a receber as requisições, encaminhá-las a *model* e enviar respectivas respostas. A camada *model* é constituída por classes de entidade e de negócio e que sendo assim recebem dados da camada *view* e executam as regras pertinentes ao negócio. A *view* é composta por uma hierarquia de componentes visuais organizados.

b) **IDE Net Beans:** O *Net Beans* é um ambiente de desenvolvimento integrado, gratuito e de código fonte aberto. A IDE foi popularizada pela empresa Sun Microsystems em 1999, cuja nomeou o projeto como *NetBeans Developer X2*. Em junho de 2000 a Sun disponibilizou o código fonte da aplicação, tornando-a uma plataforma *OpenSource*. A versão utilizada neste trabalho foi a 7.3, cuja proporciona ao desenvolvedor ferramentas úteis de formatação de código, marcação de pontos de parada (*breakpoints*) e outras ferramentas úteis que facilitam o desenvolvimento.

- c) **Bootstrap:** É um framework para a estruturação, diagramação e desenvolvimento da camada *front-end* de uma aplicação web. Construído e distribuído de maneira *OpenSource* em Agosto de 2010 na empresa *Twitter* pelos engenheiros Mark Otto e Jacob, o *Bootstrap* foi o projeto que mais teve acesso desde a sua divulgação à Outubro de 2014. O *framework* é compatível com a maior parte dos navegadores web, e desde a versão 2.0 também suporta e aplica o *design* responsivo.
- d) **HTML 5:** Segundo Boneli (2011), o HTML 5 é a próxima geração da linguagem de marcação HTML, sucedendo HTML 4.01, XHTML 1.0 e XHTML 1.1. Oferece novas funções que são necessárias para aplicativos móveis modernos. Ele padroniza diferentes características de aplicações *web* que desenvolvedores utilizam a anos, porém, que nunca foram documentados por um comitê de padronização. Um dos principais objetivos é facilitar a manipulação dos elementos de interface, possibilitando o desenvolvedor modificar as características dos objetos de forma não intrusiva e de maneira transparente ao usuário final.
- e) **JQUERY:** Segundo Silva (2008), Jquery é uma maneira fácil e simples de escrever JavaScript e colocada ao alcance não somente de programadores experientes, mas também de designers e desenvolvedores com pouco conhecimento de programação. Em 22 de agosto de 2005, John Resig, um desenvolvedor americano de profundo conhecimento em JavaScript, escreveu em seu blog um artigo relatando sua frustração com a maneira verbosa de se escrever código JavaScript para obter os resultados pretendidos. Neste artigo publicou alguns exemplos no quais propunha o uso de seletores CSS com o principal objetivo de simplificar e dar mais versatilidade ao código. O nome ainda não existia, mas nessa ocasião foi lançada a ideia que traria como resultado a biblioteca Jquery.
- f) **PrimeFaces:** Segundo Melo (2010), o *PrimeFaces* é uma biblioteca de componentes para *JavaServerFaces* com mais de 90 componentes registrados. É uma das ferramentas mais completas e uma das primeiras a estar totalmente convertida para JSF 2.0. Para a aplicação em questão, foi utilizado a API em

questão para a apresentação em forma de gráficos, dos resultados do cálculo de probabilidade evasiva de novas instâncias de alunos.

g) **WEKA:** WEKA é uma suíte de mineração de dados muito popular no meio acadêmico, desenvolvido utilizando a linguagem Java. Criada nas dependências da Universidade de Waikato, Nova Zelândia. Atualmente é mantida por uma comunidade de entusiastas por ser um software livre disponível sobre a licença GPL. A ferramenta WEKA oferece uma série de recursos, dentre os quais uma API que permite realizar a integração dos algoritmos e técnicas de mineração de dados à uma aplicação JAVA. (SILVA, 2004).

h) **Astah:** É uma ferramenta de diagramação UML, com ela é possível a criação de diagramas de caso de uso, diagramas de classe, além de outros diagramas, tais como diagrama de fluxo de dados, diagrama de entidade-relacionamento e outras funcionalidades úteis na fase de especificação e projeto de um sistema.

Segundo o *Astah Community* (2014), este ambiente permite criar diagramas rapidamente possuindo inúmeros recursos. Apesar de o Astah possuir versões pagas, o *Astah Community* é de uso gratuito.

i) **PostgreSQL:** O PostgreSQL é um poderoso sistema gerenciador de banco de dados objeto-relacional de código aberto. Tem cerca de 15 anos de desenvolvimento ativo e uma arquitetura de grande reputação e confiabilidade na comunidade *OpenSource*. Funciona em todos os grandes sistemas operacionais, incluindo GNU/Linux, Unix e MS *Windows*. É totalmente compatível com ACID, tem suporte completo a chaves estrangeiras, junções, visões, gatilhos e procedimentos armazenados. Suporta também o armazenamento de objetos binários, incluindo figuras, sons ou vídeos.

Devido às características apresentadas e a sua alta compatibilidade com a API Weka, para o desenvolvimento da aplicação de *data mining*, cuja é objeto principal deste trabalho, foi utilizado com sistema gerenciador de banco de dados o PostgreSQL versão 9.3.

3.2 MÉTODOS

O desenvolvimento da aplicação data mining aplicado ao processo de gestão educacional para detecção de evasão escolar foi dividido em etapas de acordo com o modelo sequencial linear de Pressman (2002). As etapas são:

- a) **Requisitos:** A definição do escopo e requisitos foi realizada a partir da comparação e dos dados coletados no questionário sócio educacional aplicado sobre os alunos do curso de Licenciatura em Informática da Universidade Tecnológica Federal do Paraná – Campus Francisco Beltrão, e as informações coletadas junto à secretaria da instituição a respeito dos alunos desistentes do mesmo curso.
- b) **Análise:** Nesta disciplina foram feitos estudos sobre os seguintes classificadores (a fim de detectar qual seria escolhido para aplicação do trabalho em questão): Naive Bayes, IBK.
- c) **Projeto:** Nesta disciplina foi realizada a preparação dos dados para aplicação das técnicas de mineração de dados. Ainda aqui foi construído o modelo relacional do banco de dados.
- d) **Desenvolvimento:** Nesta disciplina foi aplicado o algoritmo Naive Bayes sobre a base de dados previamente criada nos alunos que responderam o questionário sócio educacional além da construção do protótipo web para a análise de probabilidade da evasão escolar em novas instâncias de alunos, cujas ferramentas e tecnologias utilizadas estão listadas na seção 3.1.

3.2.1 MODELAGEM DA APLICAÇÃO

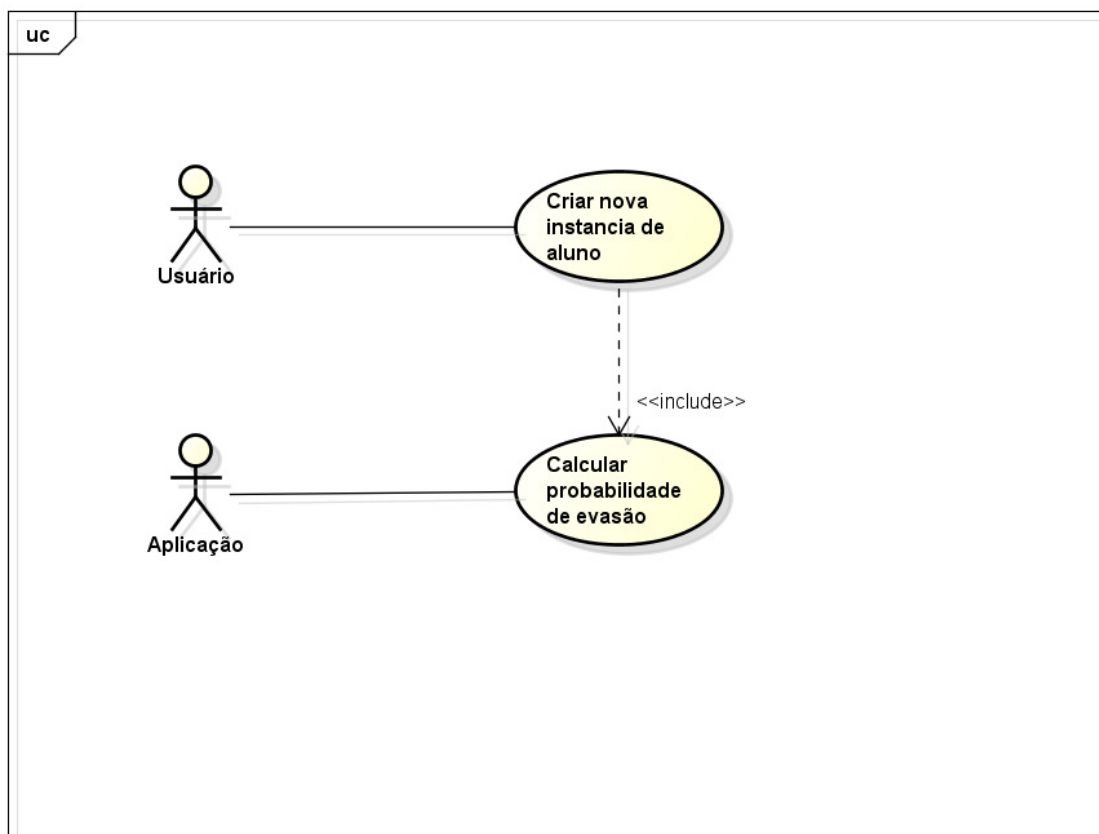
Nesta seção, são apresentados artefatos gerados que auxiliaram no desenvolvimento da aplicação, são eles, diagramas de casos de uso, diagrama de classe e diagrama de sequencia.

3.2.2 CASO DE USO

Na visão de Bezerra (2007, p. 54), “Um caso de uso representa um relato de uso de certa funcionalidade do sistema em questão, sem revelar a estrutura e o comportamento internos desse sistema”.

O diagrama de casos de uso (DCU) corresponde a uma visão externa do sistema e representa graficamente os atores, casos de uso e relacionamentos entre esses elementos. O diagrama de caso de uso tem o objetivo de ilustrar em um nível alto de abstração quais elementos externos interagem com que funcionalidades do sistema. (BEZERRA, 2007, p. 57)

Neste tópico é apresentado o diagrama de caso de uso da aplicação proposta, conforme Figura 2.



powered by Astah

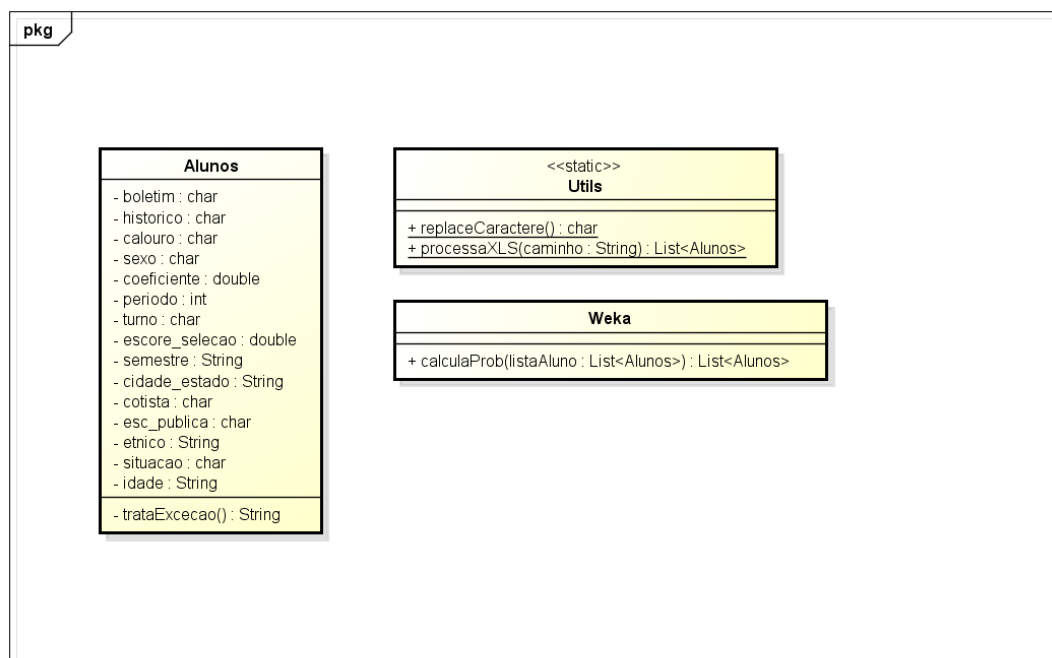
Figura 2 - Diagrama de Caso de Uso

O diagrama de caso de uso apresenta as principais funcionalidades da aplicação, a função “Criar nova instância de aluno” compreende da ação do usuário em preencher as respectivas informações de um novo aluno, criar um novo, e posteriormente solicitar o cálculo de uma nova previsão, cuja é compreendida na funcionalidade “Calcular probabilidade de evasão” ação executada pela aplicação.

Assim que o usuário executar a ação de instanciar novos aluno, automaticamente, a aplicação realiza o cálculo da probabilidade de evasão do mesmo, e retorna os números para a aplicação.

3.2.3 DIAGRAMA DE CLASSE

O diagrama de classe é do tipo estrutural e especifica as Classes, abstrações e suas responsabilidades, dentro do escopo de um projeto. Define as estruturas, exibindo a colaboração entre classes e mostra a estruturação conceitual do banco de dados (modelo lógico). Na figura 3 é ilustrado o diagrama de classe que fora utilizado para criação da aplicação e posterior para a criação do modelo lógico do banco de dados.



powered by Astah

Figura 3 - Diagrama de Classe

O diagrama acima ilustra as três classes com seus respectivos métodos e atributos. A primeira classe “Alunos” elucida o modelo criado a partir dos dados que foram transformados e preparados na planilha de Relação de Alunos (Vide Anexo A). O método trataExcecao() prepara uma instância de Aluno para posteriormente realizar o cálculo de probabilidade de evasão. A classe *Utils* tem um estereótipo do tipo *static*, o que implica em dizer que a referente classe será do tipo Estática. As classes

estáticas são aquelas que contêm apenas membros estáticos, além de que elas não podem ser instanciadas. O método `replaceCharacter()` tem a função de tratar qualquer tipo de caractere especial que um atributo de Aluno possa conter, tendo em vista que a API Weka não trabalha com caracteres especiais. O método `processaXMS()` recebe o caminho da planilha com o resultado do questionário sócio educacional (Vide anexo A), processa-os e os retorna com uma Lista de Alunos. Por último a classe Weka, é a responsável por receber a Lista de Alunos já preparadas, aplicar a classificação utilizando o algoritmo Naive Bayes e calcular a probabilidade de evasão de cada um, atribuindo o resultado em uma mesma lista de Alunos.

3.2.4 DIAGRAMA DE SEQUENCIA

O diagrama do tipo comportamental que especifica determinado trecho de interação, formada por um conjunto de objetos e seus relacionamentos, incluindo as mensagens que poderão ser enviadas entre eles. Exibindo a ordem temporal de determinada ação no sistema, ou seja, a ordem que os eventos ocorrem.

Para o desenvolvedor, o diagrama de sequencia oferece informações detalhadas de como determinada funcionalidade do sistema é executada. Possibilita o entendimento do curso da aplicação ao ser efetuada determinada interação. É utilizado principalmente para especificar o fluxo básico de algum caso de uso, e suas variações.

Para melhor descrever o fluxo das ações que a aplicação deve executar no momento de realizar a análise de probabilidade para cada instância de alunos, foi criado um diagrama de sequencia, cujo é ilustrado na figura 4.

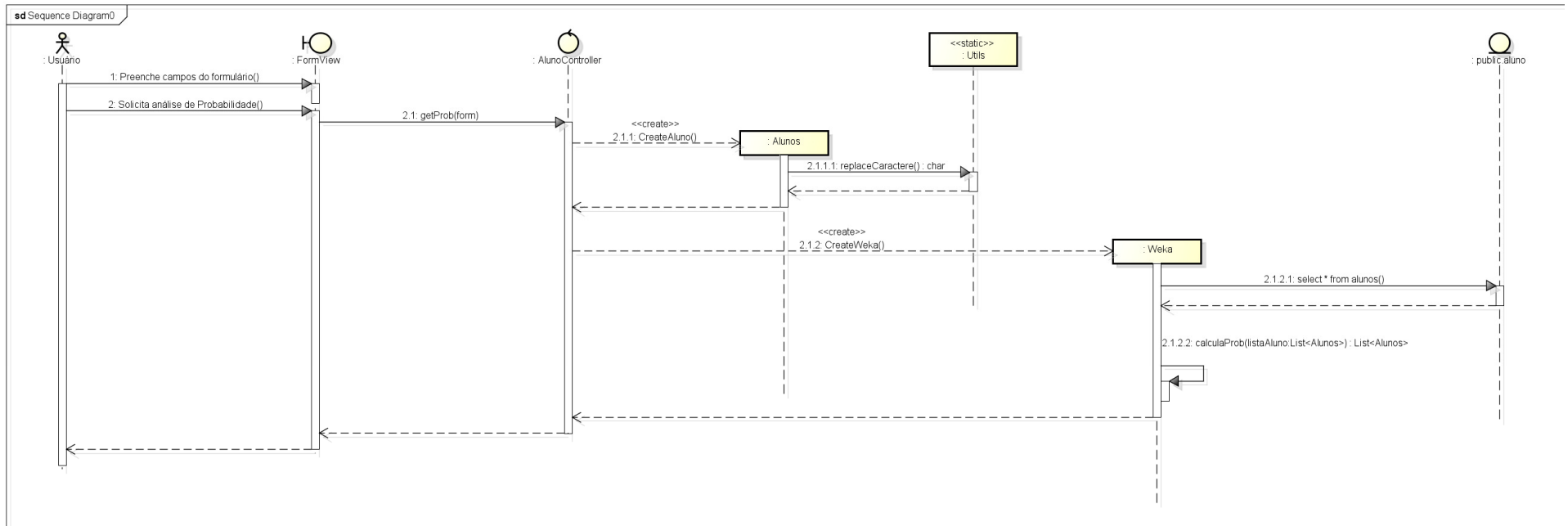


Figura 4 - Analisar probabilidade

O respectivo diagrama ilustra a ação de requerer a probabilidade de evasão de uma nova instância de aluno. Para tanto a ação é iniciada assim que o usuário preenche o formulário com os dados do novo aluno, em seguida é solicitado a probabilidade de evasão do mesmo. A partir daí a aplicação instancia um novo objeto do tipo Aluno com os parâmetros que fora previamente informado e prepara seus dados. Na sequencia é instânciado um objeto da classe WEKA que recupera as informações na base de dados de alunos e calcula a probabilidade evasiva desse novo aluno. Feito isso é retornado ao usuário da aplicação o resultado da análise.

4 RESULTADOS

Este capítulo apresenta uma visão geral, da realização de cada uma das disciplinas destacadas no modelo sequencial linear de Pressman (2002) uma vez que o objetivo deste trabalho foi o estudo da matéria de mineração de dados para possibilitar a criação de um protótipo de *data mining* aplicado ao processo de gestão educacional para detecção de evasão escolar.

4.1 REQUISITOS

Conforme o objetivo geral e os específicos propostos no início deste trabalho e alinhados com a metodologia empregada, num primeiro momento buscou-se compreender o problema proposto, a evasão escolar dos alunos do curso de Licenciatura em Informática da IES UTFPR – Campus Francisco Beltrão.

O primeiro passo para se descobrir conhecimento em bases de dados é uma correta definição do problema a ser tratado. Desta forma iniciou-se contato com a IES para obtenção de dados a serem analisados, e posteriormente um questionário foi criado e aplicado aos alunos que estavam cursando Licenciatura em Informática, a fim de definir o tipo de informação que seria interessante de ser descoberta na base de dados e iniciar o processo de KDD.

4.1.1 INFORMAÇÕES DOS ALUNOS DE LICENCIATURA EM INFORMÁTICA

Nesta etapa entrou-se em contato com a secretaria de Licenciatura em Informática da IES em questão solicitando acesso às informações dos alunos que estão cursando ou já cursaram, para que fossem realizadas as respectivas etapas da mineração dos dados. Sendo assim, a secretaria forneceu uma planilha com 247 registros de alunos que cursam ou já cursaram Licenciatura em Informática. (Vide Anexo B).

4.1.2 QUESTIONÁRIO SÓCIO-EDUCACIONAL

Feito uma análise sob os atributos dos registros da planilha de relação dos alunos do curso de Licenciatura em Informática, foi elaborado um questionário com perguntas sócio educacionais a fim de, além de outros, obtermos uma maior

variedade de atributos dos registros a serem testados e instanciados. Aplicado o questionário, foram obtidas 36 respostas. (Vide Anexo A).

4.2 ANÁLISE

A partir das informações coletadas, e das principais funcionalidades definidas foi então dado início à análise da aplicação proposta, abordando um estudo em alguns dos algoritmos de classificação de dados a fim de identificar qual se adequaria melhor ao trabalho proposto.

4.2.1 ANÁLISE DOS CLASSIFICADORES

Ainda durante a disciplina de Análise, foi realizado um estudo sobre dois diferentes algoritmos de classificação, sendo eles: IBK e Naive Bayes. O estudo foi realizado a fim de definir qual o classificador seria mais bem aplicado à solução em questão. Neste tópico, será descrito o estudo realizado com cada um, e as razões escolhidas para escolha do algoritmo proposto.

4.2.1.1 IBK

O algoritmo IBK é uma versão do algoritmo de clusterização k-NN (*k-Nearest Neighbor*) utilizado em tarefas de clusterização.

Atribui-se para o respectivo algoritmo (IBK), como ponto fraco, que o resultado atribuído pela ferramenta API WEKA, é a probabilidade da proximidade com seus respectivos vizinhos mais próximos.

4.2.1.2 NAIVE BAYES

Uma característica atraente desse classificador é a sua capacidade de produzir estimativas de probabilidade ao invés de simples classificações. Isto significa que, para cada rótulo de classe, o classificador pode gerar uma estimativa de o novo objeto pertencer à mesma.

4.2.1.3 CONCLUSÃO DA ANÁLISE

Realizado os respectivos estudos e testes sobre cada um dos classificadores anteriormente citados, determinou-se pela utilização do classificador Naive Bayes para sequência do trabalho, pelas seguintes razões:

- Probabilidades nos resultados das classificações;
- Amplo número de materiais didáticos para pesquisa;
- Possibilidade de trabalhar com mais que dois resultados nas classificações, por exemplo, diferentes tipos de desistência de alunos: desistência por motivo de trabalho, dificuldade na absorção de disciplinas, entre outros.

4.3 EXPERIMENTOS

Nesta seção são apresentados os artefatos e estudos realizados durante a disciplina de Projeto, cuja já fora previamente apresentada anteriormente. Durante a respectiva fase foi realizado a seleção dos dados conforme o processo de KDD, com o objetivo de analisar as informações de real interesse. Junto à seleção, limpeza e transformação dos dados, foi construído o diagrama de classe e o modelo de entidade relacionamento correspondente às informações posteriormente utilizadas.

4.3.1 Seleção, limpeza e transformação dos dados.

Após um estudo sobre os dados contidos na planilha de Relação de Alunos (Vide Anexo B), entendeu-se que alguns dados seriam desnecessários, como o nome, número do RG, CPF, nome da mãe e e-mail, que foram eliminados durante o processo de seleção dos dados.

Efetuada a seleção dos dados, foi hora de realizar uma limpeza e transformação sobre os dados contidos na planilha, conforme ainda processo de KDD. Na planilha de relação de alunos foi criado um padrão para a coluna “Calouro”. Nas informações fornecidas pela secretaria os resultados para o campo em questão eram apresentados como “Sim” ou “Não”, a transformação nos dados substituiu o valor “Sim” para o caractere “S” e para “Não” o caractere “N”. Os resultados para a coluna “Cotista” foram também alterados. Anteriormente existiam variações quanto a alunos do tipo cotistas, a regra foi alterada a fim de armazenar exclusivamente o

valor “S” para alunos cotistas e “N” para alunos não cotistas, o mesmo ocorreu para a coluna “Estudou em Escola Publica”.

Como não havia a coluna idade, a mesma foi criada e os resultados foram criados baseado nos valores atribuídos para o atributo “Data de Nascimento”.

A transformação mais relevante nos dados foi na coluna “Situação”, cuja anteriormente tratava diferentes resultados, Regular, Desistente, Trancado e Sem Cursar. Devido à referente tratar-se do atributo classe, os resultados foram exclusivamente tratados como “R” (Regular) para alunos regulares e “D” (Desistente) para registros com situação diferente de regular.

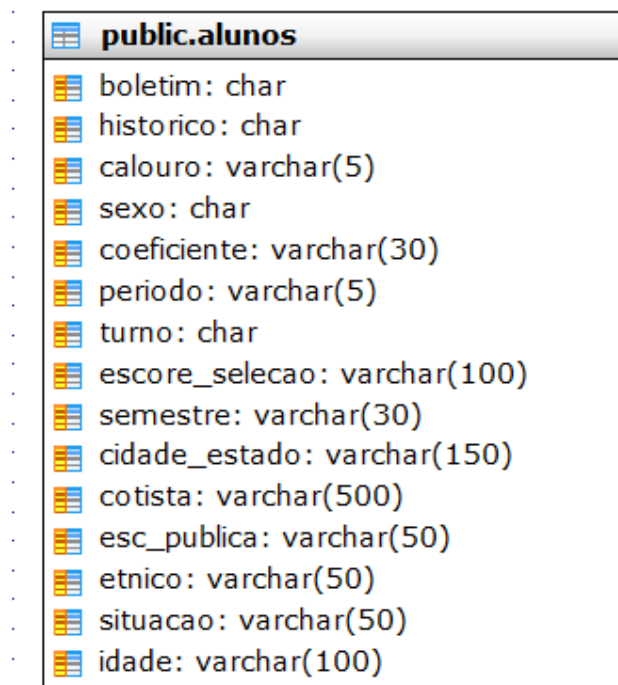
As seguintes técnicas apresentadas corresponderam à seleção, limpeza e transformação dos dados para posteriormente serem utilizados para a criação do banco de dados utilizados nas aplicações.

4.3.2 DIAGRAMA DE ENTIDADE RELACIONAMENTO

De acordo com Rezende (2002, p.201) “O principal propósito do DER é representar os objetos de dados e suas relações, sendo que cada entidade termina representada pelo menos uma tabela de dado”.

O DER representa toda a estrutura lógica do banco de dados é expressa graficamente. A sua simplicidade e a clareza gráfica podem explicar, em grande parte, a ampla disseminação do uso do modelo E-R.

Na Figura 5 será apresentado o diagrama de entidade relacionamento da aplicação proposta.



public.alunos	
boletim:	char
historico:	char
calouro:	varchar(5)
sexo:	char
coeficiente:	varchar(30)
periodo:	varchar(5)
turno:	char
escore_selecao:	varchar(100)
semestre:	varchar(30)
cidade_estado:	varchar(150)
cotista:	varchar(500)
esc_publica:	varchar(50)
etnico:	varchar(50)
situacao:	varchar(50)
idade:	varchar(100)

Figura 5 - DER Alunos

A figura acima ilustra a tabela Alunos que compõe o banco de dados que será utilizado para aplicação das técnicas de mineração de dados.

4.4 DESENVOLVIMENTO

Nesta seção serão apresentados os artefatos e estudos realizados durante a disciplina de Desenvolvimento, cuja já fora previamente apresentada. Durante a respectiva fase foi realizado a criação da base de dados (conforme descrito na seção 4.3.3) e a população da tabela com os dados que foram previamente transformados (conforme de descrito na seção 4.3.1). Em seguida foi realizada a aplicação do algoritmo Naive Bayes sobre a base de dados previamente criada nos alunos que responderam o questionário sócio educacional (Vide Anexo A). Por ultimo foi construído o protótipo web para a análise de probabilidade de evasão escolar de novas instâncias de alunos.

4.4.1 CRIAÇÃO DA BASE DE DADOS

A primeira atividade realizada na disciplina de desenvolvimento foi a criação da base de dados e suas respectivas tabelas, bem como a inserção dos dados, conforme informações que foram previamente limpas e transformadas.

Conforme descrito na seção 4.3.3 o banco de dados construído é composto de uma única tabela denominada Alunos. Com os dados previamente preparados e transformados, foi feito a importação da planilha xls Relação Alunos para a base de dados.

Utilizando a funcionalidade de análise de informações dispostas na ferramenta WEKA, observa-se que, dos registros persistidos na tabela Alunos, 137 estão com situação regular, enquanto 109 estão com situação desistente, conforme ilustrado na figura 6.

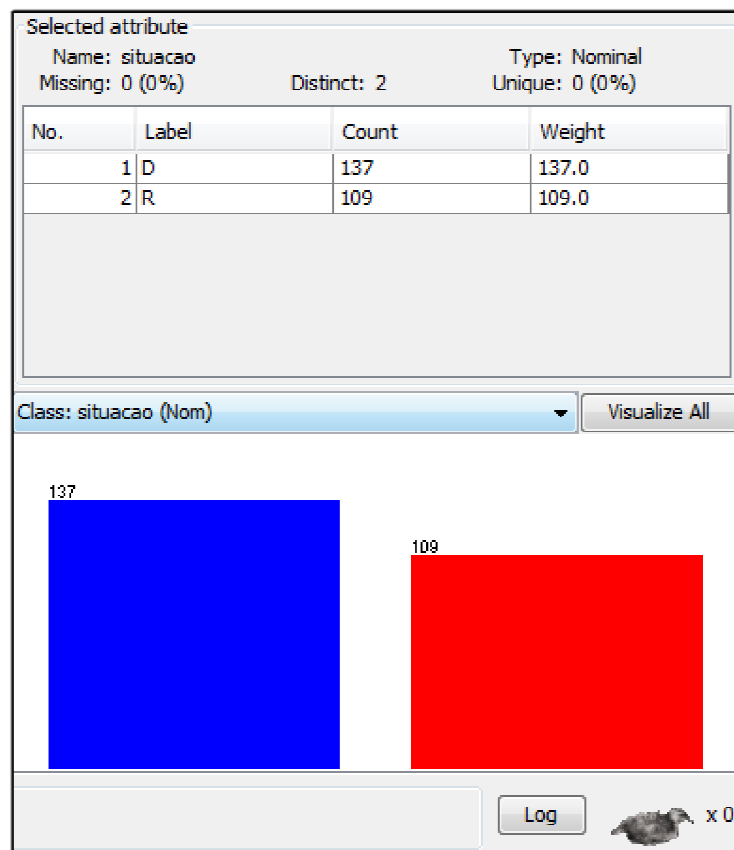


Figura 6 - Situação dos Alunos

4.4.2 CLASSIFICAÇÃO DOS ALUNOS QUE RESPONDERAM O QUESTIONÁRIO SÓCIO EDUCACIONAL

Assim que realizado a criação e inserção dos valores na base de dados, podem-se criar as primeiras classificações, e assim treinarmos se nossas informações, previamente coletadas, seriam suficientes para a análise proposta.

Para tanto, usou-se dos alunos que responderam o questionário sócio educacional (Vide Anexo A) para o treinamento. Cada resposta do questionário, foi exposta ao algoritmo de classificação Naive Bayes encontrado na API WEKA, pacote classifier.bayes.NaiveBayes. Para tanto, foi criada uma aplicação que recuperasse todas as informações contidas na planilha .xls, e as sujeitasse à classificação do procedimento em questão.

O que se esperava ao sujeitar um registro no algoritmo em questão era obter a probabilidade do referente se encaixar em uma das situações, Regular, ou Desistente, ou seja, a probabilidade do respectivo ser um aluno desistente.

Para a importação da planilha .xls foi usado a API *JExcelAPI* versão 2.6.10, conforme ilustrado na figura 7.

```
WorkbookSettings workbookSettings = new WorkbookSettings();
workbookSettings.setEncoding("ISO-8859-1");
Workbook workbook = Workbook.getWorkbook(new File("C:\\mining\\response.xls"), workbookSettings);
Sheet sheet = workbook.getSheet(0);
int linhas = sheet.getRows();
```

Figura 7 - Código para recuperar planilha xls

Como algumas informações contidas na base de dados de alunos não foram captadas no questionário sócio educacional foi usado a sua média de acontecimento como valor padrão. Isso aconteceu para os seguintes atributos: boletim, histórico, coeficiente, turno, escore seleção e étnico, conforme ilustrado na figura 8.

```
for (int i = 1; i < linhas; i++) {
    Aluno aluno = new Aluno();

    aluno.setBoletim("B");
    aluno.setHistorico("H");
    aluno.setCoeficiente("0.32");
    aluno.setTurno("N");
    aluno.setEscore_selecao("505,6");
    aluno.setEtnico("Não informado");
}
```

Figura 8 - Código média de acontecimento

Criado a lista de alunos com base na planilha sócio educacional foi aplicada a classificação em cada um dos registros a fim e obter a probabilidade de desistência de cada um dos alunos que responderam o questionário. A figura 9 ilustra o código que realiza a classificação em cada um.

```

for (Aluno aluno : listaAluno) {
    DenseInstance X = new DenseInstance(15);
    X.setDataset(data);

    X.setValue(0, aluno.getBoletim());
    X.setValue(1, aluno.getHistorico());
    X.setValue(2, aluno.getCalouro());
    X.setValue(3, aluno.getSexo());
    X.setValue(4, aluno.getCoefficiente());
    X.setValue(5, aluno.getPeriodo());
    X.setValue(6, aluno.getTurno());
    X.setValue(7, aluno.getEscore_selecao());
    X.setValue(8, aluno.getSemestre());
    X.setValue(9, aluno.getCidade_estado());
    X.setValue(10, aluno.getCotista());
    X.setValue(11, aluno.getEsc_publica());
    X.setValue(12, aluno.getEtnico());
    X.setValue(14, aluno.getIdade());

    double probs[] = nbc.distributionForInstance(X);

    System.out.println("\tProbabilidade de Regular : " + probs[0]);
    System.out.println("\tProbabilidade de Desistente : " + probs[1]);
    //System.out.println("\t" + excessao);
    System.out.println("\n_____ \n");
    cont++;
}

```

Figura 9 - Código classificação dos alunos

Conforme ilustrado acima, foi criado um laço de repetição sobre a lista de alunos previamente capturada da planilha de respostas do questionário sócio educacional. Para cada iteração foi criada uma instância da classe `DenseInstance`, atribuído para cada índice do objeto o referente atributo da classe `Aluno` e por fim aplicado o cálculo de probabilidade. Como resposta, o método retorna um vetor do tipo `double`, em que a primeira posição contém a probabilidade associada ao valor “R” (Regular) e na segunda posição a probabilidade associada ao valor “D” (Desistente). Na sequência foi solicitado a impressão das informações no `console` da IDE a fim de criarmos um relatório para melhor análise dos resultados. A tabela 2 exibe o resultado da execução da aplicação em questão

ID	Probabilidade Regular	Probabilidade Desistente
1	0.020514922357273382	0.9794850776427266
2	0.4859663677383961	0.5140336322616039
3	0.007536764357115227	0.9924632356428847
4	0.01251988535769687	0.9874801146423031
5	0.018866141685067752	0.9811338583149323
6	0.009160183099642444	0.9908398169003575
7	0.006915387519420173	0.9930846124805798
8	0.019759670742309273	0.9802403292576907
9	0.010473725496815755	0.9895262745031842
10	0.012694828837316197	0.9873051711626838
11	0.003589107407910066	0.9964108925920898
12	0.012694828837316197	0.9873051711626838
13	0.12257654987099788	0.877423450129002
14	0.006650561629692196	0.9933494383703079
15	0.01703655345072511	0.9829634465492749
16	0.01547977961487763	0.9845202203851223
17	0.03088511243009884	0.9691148875699012
18	0.08848517139467055	0.9115148286053295
19	0.17249166393715445	0.8275083360628455
20	0.04437069920959597	0.9556293007904041
21	0.036022331451894936	0.9639776685481051
22	0.28700227310117776	0.7129977268988222
23	0.012694828837316197	0.9873051711626838
24	0.013844869866956279	0.9861551301330438

ID	Probabilidade Regular	Probabilidade Desistente
25	0.022109455041123216	0.9778905449588767
26	0.02953200435279171	0.9704679956472082
27	0.019077016725296308	0.9809229832747037
28	0.01965290449715235	0.9803470955028477
29	0.01207155018819024	0.9879284498118098
30	0.01844530570966318	0.9815546942903368
31	0.03363512817733323	0.9663648718226667
32	0.005333598591456509	0.9946664014085435
33	0.019077016725296308	0.9809229832747037
34	0.014590963766227025	0.985409036233773
35	0.01502230874995348	0.9849776912500464
Total	0	35

**Tabela 2 - Resultado da classificação questionário sócio educacional.
Fonte: Autoria própria**

Conforme descrito na tabela de resultados da classificação do questionário sócio educacional, 35 dos 35 alunos que responderam o questionário, ou seja, 100% dos alunos tem maior probabilidade de se tornarem desistentes. Com base nos dados apresentados concluímos que, umas das possíveis causas referentes aos resultados apresentados são que as informações contidas na nossa base de dados de alunos não são suficientes para análise e classificação das referentes predições de evasão, levando em consideração que a maior parte dos registros da nossa base de dados (56%) se trata de alunos com situação de desistência.

A fim de validarmos o cálculo da probabilidade de evasão, foi criada uma tabela de testes, onde inserimos os mesmos 246 registros, porém alteramos a situação de alguns registros totalizando 220 alunos com situação Regular (89%) e 26 (11%) com situação Desistente. Aplicado o mesmo procedimento sobre a referente tabela, o resultado apresentado foi mais satisfatório: 25 dos 35 registros, ou seja, 71,42% tem maior probabilidade de continuarem como alunos Regulares,

em contra partida, 28,58% dos alunos teriam maior probabilidade de tornarem-se desistentes.

Para os cálculos das probabilidades de evasão acima destacados, os seguintes atributos foram utilizados como parâmetros para os cálculos: idade, sexo, coeficiente, período atual, turno, semestre que entrou na faculdade, cidade e estado de origem, aluno oriundo de escola pública ou não, aluno cotista ou não e se o aluno é calouro ou não.

Sendo assim, cabe como trabalho futuro, a coleta de maiores informações dos alunos que estão cursando e alunos que desistiram do curso em questão, totalizando um maior número de registros, e uma maior variedade de informações.

4.4.3 PROTÓTIPO WEB PARA A ANÁLISE DE PROBABILIDADE DE EVASÃO ESCOLAR DE NOVAS INSTÂNCIAS DE ALUNOS

Assim que criado a base de dados e a treinada com os resultados da planilha do questionário sócio educacional, foi realizado a implementação de um protótipo web, cujo objetivo é receber a instância de um novo aluno, aplicar o algoritmo sobre a base de dados previamente criada e obter a probabilidade do respectivo se enquadrar em uma das situações: “Regular” ou “Desistente”. As técnicas e ferramentas utilizadas para a construção da ferramenta em questão foram previamente descritas na seção 3 deste trabalho.

4.4.3.1 COMUNICAÇÃO COM O MODELO RELACIONAL

A primeira atividade realizada para a preparação e criação do protótipo *web*, foi criar a comunicação entre a API WEKA com o banco de dados relacional. A ferramenta WEKA trabalha preferencialmente com bases de dados no formato texto. Por esta razão, quase todos os tutoriais e apostilas sobre a ferramenta disponibilizados na Internet mostram como utilizar a WEKA para minerar bases de dados estruturadas nos formatos ARFF ou CSV. No entanto, é possível, também, utilizar a WEKA para minerar dados diretamente de tabelas de SGBD's relacionais. Esta forma de utilização da WEKA é bastante interessante já que, na maior parte das aplicações práticas, os dados a serem minerados se encontram originalmente

armazenados em algum banco relacional (*MySQL, Oracle, PostgreSQL, SQL Server, etc.*). Desta forma, quando se efetua a conexão da WEKA com um SGBD, deixa de existir a necessidade de exportar os dados do banco para texto e posteriormente formatá-los como ARFF (ou CSV) - atividades que, em geral, consomem um tempo considerável.

Para realizar a comunicação entre a ferramenta WEKA e o banco de dados relacional, foi utilizado o driver JDBC do SGBD *Postgresql*: “*postgresql-9.3-1102.jdbc3.jar*”. Feito isso, foi necessário configurar a API WEKA para comunicar-se com o banco de dados, previamente criado. Para tanto, a API oferece um arquivo nomeado “*DatabaseUtils.props*”, cujo pode ser extraído de dentro do arquivo “*weka.jar*”. No arquivo devem ser informados os seguintes parâmetros: Driver JDBC, URL do banco de dados e informações de *login*. Feito isso, a ferramenta está pronta para conectar-se ao nosso banco *Postgresql*.

4.4.3.2 CRIAÇÃO DA INTERFACE WEB

Preparado e testado a comunicação entre o banco de dados relacional e a ferramenta WEKA, o próximo passo foi a construção da interface web, cuja é o meio de interação do usuário com a aplicação. Para tanto foi utilizado um *template bootstrap* (apresentado na seção 3.1) que se adequa a diferentes resoluções, conceito conhecido como *design* responsivo, conforme ilustrado na figura 18. Segundo Garone (2012), *design* responsivo é: “... uma técnica de estruturação HTML e CSS, em que o site se adapta ao *browser* do usuário sem precisar definir diversas folhas de estilos para cada resolução.”.

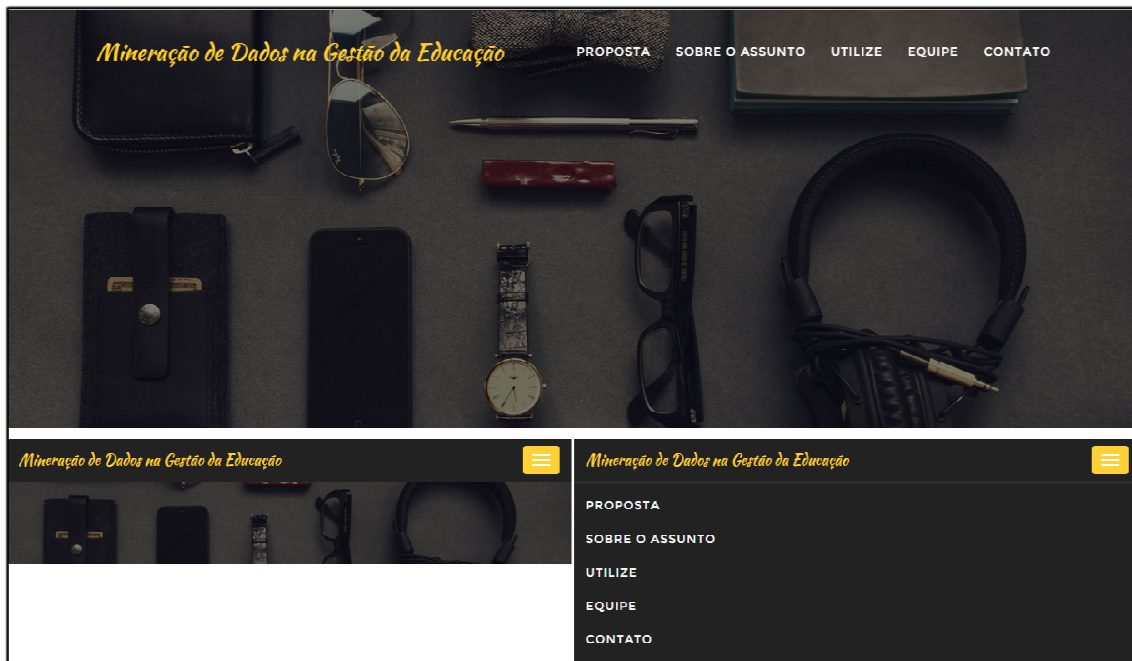


Figura 10 - Design responsivo do protótipo

A figura 10 ilustra as adequações que a página *web* realiza, baseado em diferentes resoluções de *interface*.

O protótipo *web* possui uma única página *xhtml*, cuja realiza uma breve apresentação do tema proposto, apresenta os participantes do projeto e oferece um breve formulário, ilustrado na figura 11, cujo será objeto principal desta seção.

Figura 11 - Formulário

A figura 11 ilustra o formulário que capturará os respectivos dados de uma nova instância de aluno, aplicará as técnicas de mineração de dados e retornará a

probabilidade desse novo aluno se enquadrar em uma das duas situações: “Regular” ou “Desistente”. O formulário dispõe dos seguintes campos: Nome, Código de Matrícula, Coeficiente, Idade, Sexo, É Calouro, Turno, Estudou em escola pública, Período que cursa atualmente, Aluno Cotista, Cidade, Estado, Mês e Ano de ingresso na universidade. Com exceção dos campos Nome e Código de Matrícula, os demais campos são utilizados para o cálculo da probabilidade evasiva do novo aluno.

4.4.3.3 CRIAÇÃO DA REGRA DE NEGÓCIO

Com a interface de interação com o usuário pronta, o próximo passo foi o desenvolvimento da regra de negócio, cuja captura as informações digitadas pelo usuário, aplica técnicas de mineração de dados e apresenta as probabilidades da nova instância de aluno. O primeiro método da aplicação é responsável por recuperar os dados digitados na interface e criar, com base nessas informações, uma nova instância de aluno, conforme ilustrado na figura 12.

```

public void preparaAluno() throws Exception {
    HttpServletRequest request = (HttpServletRequest) FacesContext.getCurrentInstance().getExternalContext().getRequest();

    String nome = request.getParameter("nome");
    String codigo = request.getParameter("codigo");
    String coeficiente = request.getParameter("coeficiente");
    String idade = request.getParameter("idade");
    String cidade = request.getParameter("cidade");
    String estado = request.getParameter("estado");
    String semestre = request.getParameter("semestre");
    String sexo = request.getParameter("sexo");
    String isCalouro = request.getParameter("calouro");
    String turno = request.getParameter("diurno");
    String escPublica = request.getParameter("escolaPublica");
    String periodo = request.getParameter("periodo");
    String cotista = request.getParameter("cotista");

    Aluno aluno = new Aluno(nome, codigo, isCalouro, sexo, coeficiente.replace(",", "."), periodo, turno,
        semestre, cidade.toUpperCase() + "/" + estado.toUpperCase(), cotista, escPublica, idade);

    this.conectaWeka(aluno);
}

```

Figura 12 - Método preparaAluno()

Criado a instância do aluno, o próximo método aplica o algoritmo naive bayes, sobre o novo objeto. O processo em questão é o mesmo que fora previamente apresentado na Figura 13 - Código classificação dos alunos, com a diferença de que o cálculo será executado uma única vez sobre o aluno que fora instânciado na imagem acima.

Executado o cálculo de probabilidade evasiva do aluno em questão, o próximo método armazena as informações calculadas em uma sessão que será posteriormente recuperada para apresentação do resultado, conforme ilustrado na figura 14.

```
public void createSessao(String prob1, String prob2, String nome, String codigo) throws IOException{
    FacesContext facesContext = FacesContext.getCurrentInstance();
    HttpSession sessaoHttp = (HttpSession) facesContext.getExternalContext().getSession(true);
    sessaoHttp.setAttribute("probReg", prob1);
    sessaoHttp.setAttribute("probDes", prob2);
    sessaoHttp.setAttribute("saudacao", "Probabilidade de evasão do aluno " + nome + " código de matrícula: " + codigo);
    FacesContext.getCurrentInstance().getExternalContext().redirect("charExemplo"+".xhtml");
}
```

Figura 14 - Criação da sessão para apresentação do resultado

4.4.3.4 APRESENTAÇÃO DO RESULTADO

Assim que executado o método ilustrado na figura 12, as informações previamente calculadas são exibidas graficamente em uma nova página, cuja ilustra a probabilidade do aluno se enquadrar em cada uma das situações: “Regular” e “Desistente”, conforme ilustrado na figura 15.

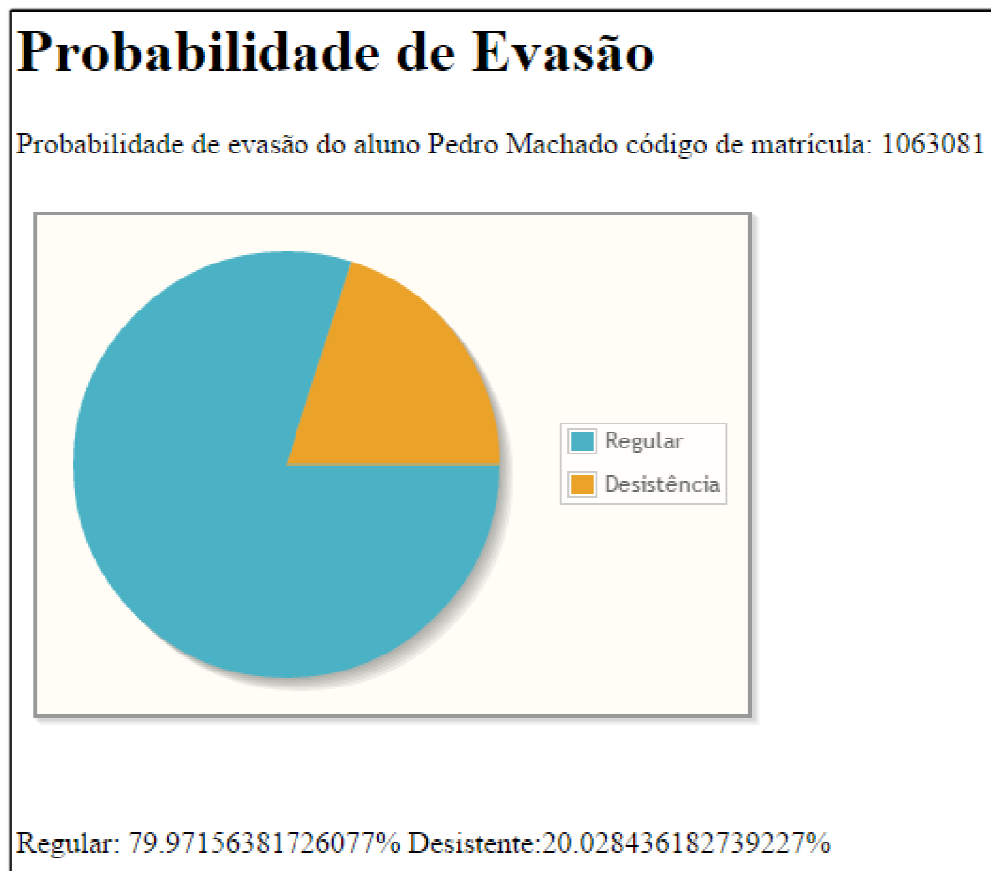


Figura 15 - Gráfico de probabilidade

A figura acima ilustra um teste fictício efetuado sobre uma nova instância do aluno Pedro Machado, código de matrícula 1063081, coeficiente: 89,6, idade 26 a 29 anos, sexo masculino, calouro, estuda no período noturno, estudou em escola pública, cursa atualmente o 6º período, ingressou na universidade por meio de alguma cota, é de Dois Vizinhos – PR e ingressou na universidade no segundo semestre de 2012. O resultado para a nova instância do aluno é de 79,97% de chance de ser regular contra 20,03% de chance de ser desistente.

5 CONCLUSÃO

A principal utilidade da pesquisa e da aplicação desenvolvida é auxiliar os docentes em possíveis tomadas de decisões quanto aos alunos que ministra. Assim que sujeitado alunos ao cálculo da probabilidade evasiva, depois que adquirido novas instâncias e novos treinamentos no banco de dados, o professor poderá tomar decisões mais assertivas.

A utilização das técnicas aqui aplicadas no ambiente proposto permitiu a viabilidade e a utilidade prática da metodologia em um estudo de caso real, sendo que os resultados poderão auxiliar os gestores a elucidar características relevantes em relação a diversas situações observadas neste estudo.

Devido ao baixo número de informações contidas na nossa base de dados, pode-se observar a baixa confiabilidade nos cálculos aplicados sobre os alunos que responderam o questionário sócio educacional. Essa dificuldade somente será amenizada com o aprendizado da base de dados bem como o crescimento horizontal e vertical da base de dados.

Essa dificuldade pode ser observada quando se aplica o algoritmo de classificação em nossa primeira base de dados (fornecida pela secretaria), cujo resultado foi que 100% dos alunos testados têm maior probabilidade de serem evasivos. O resultado deu-se devido aos registros na nossa base de dados possuírem resultados semelhantes quando analisado a situação dos mesmos. 137 estão com situação regular, enquanto 109 estão com situação desistente. Uma base de dados fictícia foi criada cuja possuía 220 alunos com situação Regular (89%) e 26 (11%) com situação Desistente, cuja se refere a um ambiente de ensino mais próximo do real (onde há mais alunos Regulares que alunos desistentes). O mesmo procedimento foi aplicado, e o resultado apresentado foi mais satisfatório: 71,42% dos alunos têm maior chance de continuarem como alunos Regulares, em contra partida, 28,58% dos alunos teriam maior probabilidade de tornarem-se desistentes.

Alguns desafios e dificuldades foram encontrados durante o processo de desenvolvimento do trabalho proposto, entre eles:

- a) A dificuldade inicial neste trabalho foi da liberação da base de dados os alunos do curso de Licenciatura em Informática da Universidade

Tecnológica Federal do Paraná – Campus Francisco Beltrão, haja vista tratar-se de dados pessoais dos acadêmicos ali inscritos e matriculados;

- b) A limitação das informações contidas na base de dados, que não contempla informações a respeito de alunos aprovados e matriculados, bem como maiores informações sociais, cujas foram capturadas no questionário sócio educacional aplicado.

Após o estudo abordado neste trabalho, estabelecem-se algumas recomendações para pesquisas de mesmo cunho. Alguns assuntos merecem aprofundamento em pesquisas ou trabalhos futuros. Os principais são:

- a) Utilização de outras técnicas de Mineração de Dados não contempladas neste estudo, como por exemplo, Redes Neurais e Algoritmos Genéticos;
- b) Aplicação de um questionário sócio educacional sobre alunos recém-matriculados, bem como em alunos desistentes do curso, a fim de aprimorar e colher um maior número de registros na nossa base de dados;
- c) Melhorias no protótipo web de cálculo de evasão escolar, a fim de torná-lo uma ferramenta aos gestores da instituição, cujos poderão estabelecer períodos que serão aplicados às respectivas técnicas aqui apresentadas e planejar as tomadas de decisões;

6 REFERÊNCIAS

BAGGI, Cristiane Aparecida dos Santos. **Evasão e Avaliação Institucional no Ensino Superior: Uma discussão bibliográfica**. 2010. Dissertação Mestrado - PUC Campinas, SP.

BAKER, Ryan Shaun Joazeiro de. **Mineração de Dados Educacionais: Oportunidades para o Brasil**. 2011.

BATISTA P., SILVA M.J. **Mining Web Access Logs of an On-line Newspaper.**, Departamento de Informatica, Faculdade de Ciencias . Universidade de Lisboa. Disponível em:<http://xldb.fc.ul.pt/data/Publications_attach/rpec02.pdf>. Acesso em: 10 ago. 2014.

BEZERRA, Eduardo. **Princípios de análise e Projeto de Sistemas com UML**. 2ª Edição. Rio de Janeiro: Campus, 2007.

BOENTE, A. N. P. ; OLIVEIRA, F. S. G. ; ROSA, J. L. A.. **Utilização de Ferramenta de KDD para Integração de Aprendizagem e Tecnologia em Busca da Gestão Estratégica do Conhecimento na Empresa**. Anais do Simpósio de Excelência em Gestão e Tecnologia, v. 1, p. 123-132, 2007.

BRAGA, Luis Paulo Vieira. **Introdução à Mineração de Dados**. 2ª edição, 2005. 212 p. 19 – 20.

BRASIL. **Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais. Sinopses do ensino superior. Censos do ensino superior. Comunicações pessoais**. Disponível em: www.inep.gov.br. Acesso em: 10 nov. 2014.

CAMILO, Cássio Oliveira; SILVA João Carlos da. **Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas**. Relatório Técnico. 2009

CHIARA, Ramon. **Aplicação de Técnicas de Data Mining em Logs de Servidores Web**. 2003. Dissertação (Mestrado). Instituto de Ciências Matemáticas e de Computação - ICMC-USP. 2003.

CRISP-DM. **Cross Industry Standart Process for Data Mining**. Disponível em:<<http://www.crisp-dm.org/>>. Acesso em: 10 out. 2014.

CRUZ, Armando J. R. da. **Data Mining via Redes Neurais Artificiais e Máquinas de Vetores de Suporte**. 2007. 123 f. Dissertação (Mestrado em Sistemas de Informação) – Universidade do Minho, Lisboa, 2007.

DIAS, Maria M. **Um Modelo de Formalização do Processo de Desenvolvimento de Sistemas de Descoberta de Conhecimento em Banco de Dados**. 2001. 212 f. Tese (Doutorado em Engenharia da Produção) – Universidade Federal de Santa Catarina – UFSC, Florianópolis, 2001.

FAYYAD; PIATETSKY-SHAPIRO, G.; SMYTH, P. From **Data Mining to Knowledge Discovery: An Overview**. Menlo Park, CA: AAAI Press/The MIT Press, 1996

FIGUEIRA, Rafael. **Mineração de dados e banco de dados orientados a objetos**. 1988. Dissertação (Mestrado em Ciências da Computação) – Universidade Federal de Santa Catarina, Florianópolis, 2011.

FURTADO, M. I. V. **Inteligência competitiva para o ensino superior privado: Uma abordagem através da mineração de textos**. 2004. Tese (Doutorado). COPPE/UFRJ. Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2004.

GAIOSO, Natalicia Pacheco de Lacerda. **O fenômeno da evasão escolar na educação superior no Brasil**. 2005. Dissertação Mestrado – Programa de Pós Graduação em Educação da Universidade Católica de Brasília, Brasília, 2005.

GALVÃO, Noemi Dreyer. **Técnica de mineração de dados: uma revisão da literatura**. Cuiabá – MT. 2009.

GOLDSCHMIDT, R.R.; PASSOS, E. **Data Mining: Um guia prático, conceitos, técnicas, ferramentas, orientações e aplicações**. São Paulo: Elsevier 2005.

HIPÓLITO, Oscar. **Financiamento e expansão do ensino superior**. Disponível em: <http://www.jornaldaciencia.org.br/Detail.jsp?id=62770>. Acesso em 10 jan. 2015.

HIRAGI, GILBERTO de O.. **Mineração de Dados em Base de Germoplasma**. 2008, 108 f. Dissertação (Mestrado em Informática) - Universidade de Brasília- UnB, Brasília, 2008.

LEAL, Roberto. **A evasão no ensino superior brasileiro. Artigo. Instituto Lobo para o Desenvolvimento da Educação, da Ciência e da Tecnologia, 2007**

LUCAS, Anelise de Macedo. **Utilização de Técnicas de Mineração de Dados considerando os Aspectos Temporais**. 2002. Dissertação (Mestrado). Porto Alegre: PPGC da UFRGS, 2002.

MELO, Alexandre. **Programação JAVA para Web**. NOVATEC; 2014

MORATE, Diego G. **Manual de WEKA**. Valladolid, 2010. Disponível em: <<http://www.metaemotion.com/diego.garcia.morate/>>. Acesso em: 12 nov 2014

PASTA, Arquela. **Aplicação da Técnica de Data Mining na Base de Dados do Ambiente de Gestão Educacional: Um estudo de caso de uma Instituição de Ensino Superior de Blumenau – SC**. 2011. Dissertação (Mestrado). Univali.

PRESSMAN, R. **Engenharia de software**, 5ª ed., Rio de Janeiro: McGraw-Hill, 2002.

RABELO, Emerson. **Avaliação de Técnicas de Visualização para Mineração de Dados**. 2007. 103 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual de Maringá. Maringá, 2007.

REZENDE, Denis Alcides. **Engenharia de Software e Sistemas de Informação**. 2ª Edição. Rio de Janeiro: Brasport, 2002.

SCOSS, Anne M.. **A Clusterização e Classificação no Processo De Data Mining para Análise do Desempenho Docente no Ensino de Graduação**. 2006. 86 f. Trabalho de Conclusão de Curso (Especialização) - Universidade do Extremo Sul Catarinense - UNESC, Criciúma, 2006.

SHIBA, Sonia Kaoru. **Modelagem de processo de extração de conhecimento em banco de dados para sistemas de suporte à decisão**. 2008. Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. 2008

SILVA, Mauricio. **JQUERY - A biblioteca do programador javascript**. 3ª Edição; NOVATEC; 2013

SILVA, Marcelino Pereira dos Santos. **Mineração de Dados – Conceitos, Aplicações e Exeperimentos com Weka**. 2004

UNESCO. **Relatório sobre educação para todos**. Disponível em: <<http://www.onu.org.br/unesco-lanca-relatorio-sobre-educacao-para-todos/>>

Acessado em 10 set. 2014

ANEXO A – Questionário Sócio Educacional

Questionário Sócio Educacional

Questionário destinado aos alunos do curso de Licenciatura em Informática da Universidade Tecnológica Federal do Paraná, Campus Francisco Beltrão.

Os itens abaixo fazem parte de um questionário sócio educacional aplicado aos alunos de Licenciatura em Informática da UTFPR – Campus Francisco Beltrão. As respostas serão utilizadas única e exclusivamente como material didático de pesquisa, portanto sinta-se a vontade em respondê-las.

Com o objetivo de traçar um perfil dos alunos do curso de Licenciatura em Informática da UTFPR – Campus Francisco Beltrão e avaliar a eficácia das disciplinas ministradas no referente curso de graduação, solicitamos a sua atenção para responder o questionário a seguir.

* Required

1. Sexo *

Mark only one oval.

Masculino

Feminino

2. Idade *

Mark only one oval.

Até 20 anos

21 a 25 anos

26 a 29 anos

31 a 40 anos

Mais de 40 anos

3. Qual sua cidade de origem *

Ex.: Curitiba

4. Ocupa atualmente *

Escolha dentre as opções abaixo, qual a sua atual situação profissional.

Mark only one oval.

Empregado em empresa privada

Funcionário público

Tem negócio próprio

Administra negócio familiar

Não trabalha

Pesquisas científicas dentro do âmbito universitário

Other:

5. Estado civil *

Mark only one oval.

Solteiro(a)

Casado(a)

Separado(a)/Divorciado(a)
Viúvo(a)
Other:

6. Reside atualmente *

Mark only one oval.

Com os pais
Com esposo(a) e/ou filhos
Com amigos
Sozinho(a)
Other:

7. Qual a faixa de renda mensal da sua família *

Escolha dentre as opções, qual a faixa mensal de renda total da sua família. A soma total da renda mensal de cada um dos membros

Mark only one oval.

Até R\$2500,00
De R\$2500,00 a R\$4000,00
De R\$4000,00 a R\$6000,00
De R\$6000,00 a R\$9000,00
Mais de R\$9000,00

8. Qual o meio de transporte utilizado para ir à faculdade *

Mark only one oval.

Veículo Próprio
Ônibus
Carona
Other:

9. Assinale a melhor opção quanto ao seu ponto de vista financeiro *

Mark only one oval.

Não trabalho e meus gastos são financiados pelos meus pais
Trabalho e tenho ajuda da minha família
Trabalho e me sustento
Trabalho e contribuo com o sustento da minha família
Trabalho e sou o principal responsável pelo sustento da minha família

10. Em que tipo de escola você cursou o ensino médio *

Mark only one oval.

Toda em escola pública
Toda em escola particular
A maior parte do tempo em escola pública
A maior parte do tempo em escola particular
Metade em particular e outra em pública

11. Há quanto tempo concluiu o Ensino Médio *

Mark only one oval.

Menos de 1 ano
Entre 1 e 3 anos
Entre 4 e 6 anos

Entre 7 e 10 anos
Mais de 10 anos

12. O que você mais utiliza para se manter atualizado? *

Mark only one oval.

Jornais
Revistas
TV
Rádio
Internet, Redes Sociais.

13. Qual (is) meio(s) você utiliza para acessar a Internet? *

Mark only one oval.

Internet móvel (2/3/4G)
Em casa
Na universidade
Em meios públicos
Mais de uma das alternativas

14. Quais redes sociais você possui conta? *

Mark only one oval.

Facebook
Twitter
Google +
LinkedIn
WhatsApp
Mais de uma alternativa
Nenhuma
Other:

15. Com que frequência você acessa suas redes sociais?

Mark only one oval.

Diariamente
Semanalmente
Mensalmente
Esporadicamente (Semestralmente, Anualmente, ...)
Nunca

16. Por que razão escolheu o curso de Licenciatura em Informática? *

Mark only one oval.

Adequação às minhas aptidões pessoais
Prestígio da profissão
Bom mercado de trabalho
Perspectiva de Boa Remuneração
Other:

17. Quem tomou a decisão de escolha do seu curso? *

Mark only one oval.

Eu mesmo(a)
Meus pais
Companheiro(a)

Eu e meus pais
Eu e meu companheiro(a)
Other:

18. Qual período do está cursando atualmente *

Mark only one oval.

- 1º
- 2º
- 3º
- 4º
- 5º
- 6º
- 7º
- 8º

19. Qual o semestre e ano de ingresso no curso *

Ex.: 2/2010

20. Qual sua intenção após concluir o curso *

Mark only one oval.

- Atuar como empregado de empresa privada
- Realizar concurso público
- Administrar Negócios da Família
- Criar meu próprio negócio
- Dar sequencia nos estudos e pesquisa (Mestrado / Doutorado)
- Other:

21. Qual disciplina você tem maior aptidão *

Caso tenha aptidão por mais de uma área abaixo, poderá selecionar mais de uma opção.

Check all that apply.

- Algoritmo / Lógica de Programação
- Empreendedorismo, Administração e/ou Comunicação
- Filosofia e/ou Psicologia
- Pedagogia, Gestão de Pessoas e/ou Educação
- Matemática, Cálculos e/ou Fundamentos Matemáticos
- Engenharia de Software (UML, IFML, entre outras)
- Linguagem de Programação Estruturada (C, Pascal, entre outras)
- Linguagem de Programação Orientada a Objetos (C#, JAVA, C++, entre outras)

22. Quantas horas semanais (em média) você disponibiliza para estudo das disciplinas fora do ambiente das aulas. Não conte as horas em sala de aula *

Mark only one oval.

- 0
- 1 a 3
- 4 a 6
- 7 a 9
- 10 a 15
- 16 ou mais

23. Quais as maiores dificuldades encontradas no curso até o presente momento *

Check all that apply.

Conciliação com atividades extracurriculares

Problema de relacionamento com professores e/ou alunos

Dificuldade no aprendizado do conteúdo das aulas

Concentrar-se nas referentes disciplinas

Problema com disciplinas de lógica e/ou programação

Problema com disciplinas de gestão educacional

Falta de interesse nos assuntos abordados

Dificuldade com alguma disciplina em específico

Nenhuma

Other:

24. Comentários e/ou sugestões

Por favor, caso haja necessário, descreva abaixo possíveis novos questionamentos ou remoção de outros. Sinta-se a vontade para críticas, sugestões e/ou elogios.