

Análise da percepção dos alunos sobre as práticas docentes por meio da mineração de dados educacionais

Analysis of the perception of students about the teacher's practices through educational data mining

Fábio de Paula SANTOS 1; Ismar FRANGO Silveira 2; Carla Pineda LECHUGO 3

Recibido: 29/08/16 • Aprobado: 02/09/2016

Conteúdo

Introdução

1. Avaliação Institucional Docente
 2. Mineração de Dados, de Textos e Mineração de dados Educacionais
 3. Análise de Sentimentos
 4. Desenvolvimento da Pesquisa
 5. Considerações Finais
- Referências Bibliográficas

RESUMO:

Este artigo propõe um modelo computacional conceitual que aplica técnicas de Mineração de Dados Educacionais para a análise de respostas de alunos obtidas durante uma Avaliação Institucional. Através destas técnicas, procura-se encontrar novas categorias de análise no discurso dos alunos e ampliar a análise dos dados e provocar a reflexão sobre as práticas docentes. Para tanto, os autores propõem um modelo de avaliação que aplica, entre outras, a Análise de Sentimentos para identificar quais práticas pedagógicas são positivas ou negativas sob a ótica dos alunos.

Palavras-chave: Avaliação Institucional Docente; Mineração de Dados Educacionais; Análise de Sentimentos

ABSTRACT:

This paper aims the proposal of model and its subsequent application of data mining techniques in a set of students answers collected during an Institutional Evaluation. Using Educational Data Mining, the proposed model looks for new categories of analysis in the student's discourse and contribute to Teachers Evaluation effectiveness and promotes reflections about teaching practices. Therefore, the authors propose an Institutional Teaching Evaluation model that applies, among others, the Sentiment Analysis to identify which teaching practices are positive or negative from the perspective of students from Higher Education Institution.

Key words: Institutional Evaluation. Educational Data Mining. Sentiment Analysis

Introdução

O termo *Big Data* é aplicado à toda informação que não pode ser processada ou analisada

usando os processos ou ferramentas tradicionais (Zikopoulos, Eaton, & DeRoos, 2012). Os desafios do *Big Data* estão presentes também nas Instituições de Ensino Superior - que cada vez mais dispõe de mais dados sobre seus alunos, seus professores e suas práticas em sala de aula e estes dados se revestem nos diferentes formatos: desde os mais estruturados (armazenados em tabelas dos bancos de dados relacionais dos Sistemas Acadêmicos de Gestão) até os menos estruturados (como um vídeo de uma aula ou o áudio de uma palestra). Os sistemas de informação tradicionais são muito bons para analisar os dados estruturados, porém falham na análise dos dados apresentados em outros formatos, sendo que estes dados não estruturados são os mais volumosos na atualidade. Técnicas de *Big Data* tornam possível a análise de todos os dados disponíveis em uma Instituição de Ensino Superior, não importando o domínio de uma aplicação, formato ou seu volume com uma grande velocidade de análise. Conforme podemos observar na Figura 1, no universo de *Big Data* é possível extrair também informação de dados não estruturados, tais como a identificação de alunos em uma foto, por exemplo. A utilização dos bancos de dados analíticos e banco de dados *no-sql* permitem a exploração de grandes volumes de dados - na casa dos *zettabytes* - e, por último, uma maior velocidade na análise de dados é obtida através do processamento de grande volume de dados em tempo real (*streaming*) ao invés de ser realizado, por exemplo, somente durante a madrugada (em lote).

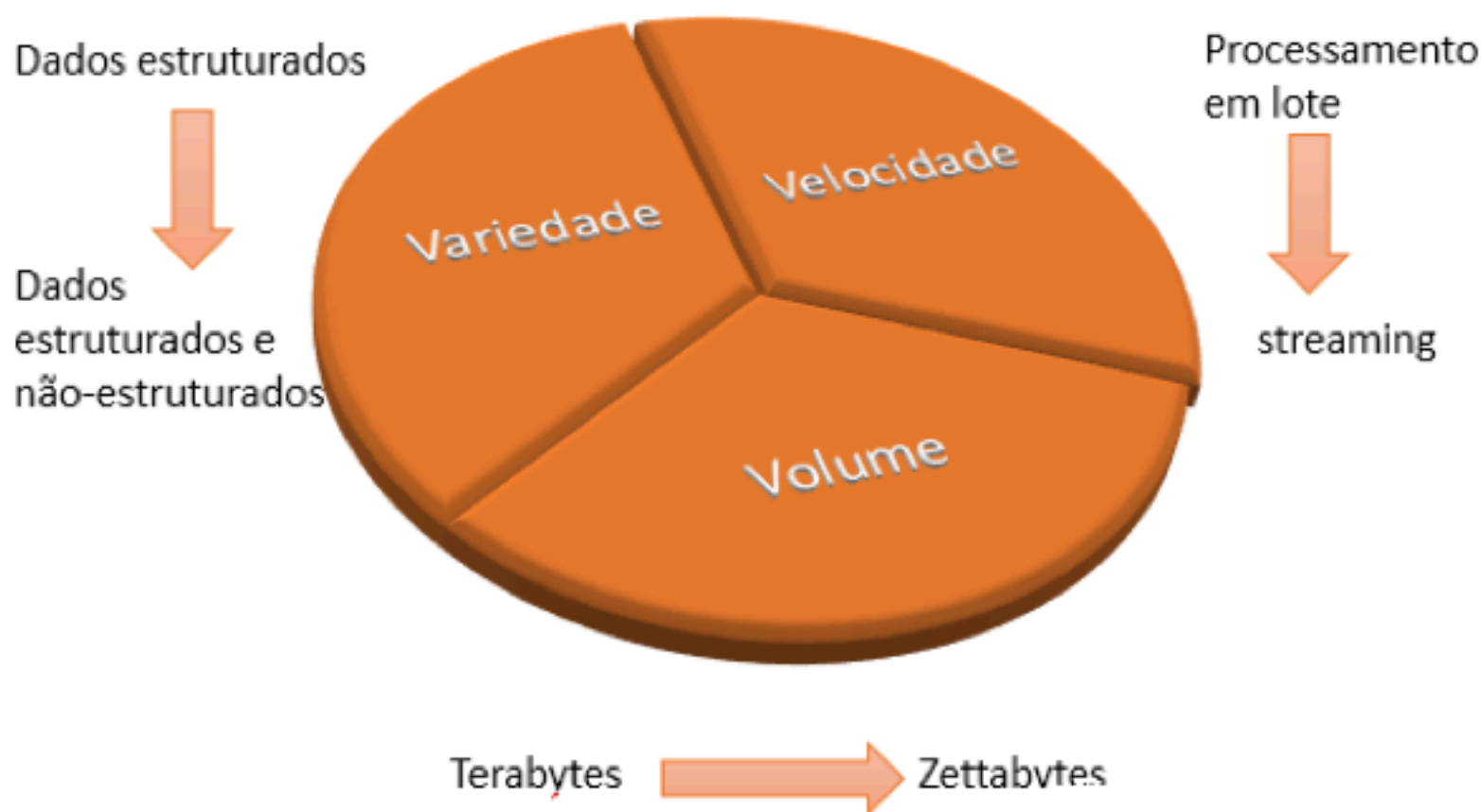


Figura 1 - Características do *Big Data*. Adaptado de(Zikopoulos et al., 2012).

No passado, o uso de dados amostrais (ou seja, coletas em pequenos volumes) em pesquisas ou avaliações era algo natural, pois coletando poucos dados era possível garantir que os números registrados fossem os mais precisos possíveis. Neste mundo de amostragem, a obsessão pela exatidão era mais crítica. Analisar apenas uma limitada quantidade de pontos de dados significa que os erros podem ser ampliados, o que potencialmente reduz a precisão do resultado final (Mayer-Schonberger & Cukier, 2013).

No mundo do *Big Data*, o afrouxamento da precisão pode ser considerado positivo, pois a possibilidade de falha é minimizada pela quantidade de dados obtida. Ao medir a temperatura de uma plantação podemos colocar apenas um sensor para toda a estufa e desta maneira, garantir a devida precisão. Por outro lado, se tivermos cem sensores a cada cem plantas, possivelmente usaremos sensores mais baratos e que talvez apresentem falhas na medição, registrando valores não exatos; mas a soma das várias leituras garantirá um resultado mais

abrangente.

Devido a este problema de escala, as avaliações docentes - em sua maioria - são formuladas ainda através de perguntas fechadas (de natureza quantitativa) de modo a permitir um processamento rápido dos questionários enviados, e por esta natureza, partem de premissas do avaliador, ou seja, o avaliador sabe de antemão o que deseja avaliar. Uma tentativa de cobrir esta falha está na criação de campos do tipo "observações", que permitem que o aluno se expresse livremente, porém limita e segmenta a resposta baseadas no tema daquela página ou formulário para reduzir a quantidade de questionamentos, pois teoricamente, eles deveriam ser lidos, processados e respondidos, o que, para uma instituição de ensino com um grande volume de alunos, demandaria um tempo de resposta muito grande, o que prejudicaria a tomada de decisão.

Com a utilização de atividades de Mineração de Dados Educacionais e de *Big Data* tornou-se possível fazer o mesmo através de perguntas abertas, tais como "Fale Bem ou "Reclame do professor" e desta forma, emergir visões das práticas pedagógicas vindas dos próprios alunos, ou ainda, obter (*on-line*) se uma prática pedagógica é considerada boa ou ruim sob o ponto de vista dos alunos.

Ora, se falamos em "alunos como coautores de um ambiente de aprendizagem"(Johnson et al., 2014), não seria óbvio ouvi-los?

E a importância de ouvi-los está, não somente em saber quais práticas pedagógicas são mais bem avaliadas pelos alunos, mas também na possível melhoria dos índices de aprendizagem e, em última instância, em uma contribuição para a redução da evasão discente no Ensino Superior. Possivelmente, alunos que durante o processo de ensino-aprendizagem vivenciaram situações das quais não se sentiram satisfeitos (ou frustrados) são os que potencialmente desistem da escola, do professor e também de aprender (Lechugo, 2016).

Neste sentido, cabe ressaltar que não é propósito deste trabalho afirmar que as práticas pedagógicas necessariamente devam satisfazer os alunos, pois, nem sempre aprender é fácil e prazeroso. Busca-se, ao ouvir os alunos, compreender suas expectativas, limitações e frustrações e, a partir deste entendimento, atribuir à Instituição e ao professor o papel de ajustar o que é possível e adequado para garantir o aprendizado de todos (professores e alunos). E uma coisa é certa: não ouvir os alunos, implica em não construir com eles, com os professores e com a escola, um caminho que seja efetivo e recompensador para todos.

1. Avaliação Institucional Docente

O trabalho de Grohmann e Ramos (Grohmann & Ramos, 2012), um dos aspectos carentes de estudos é a forma como os alunos avaliam o trabalho dos seus professores de pós-graduação e na graduação esta realidade não é diferente. A partir desta constatação, pergunta-se: não estará a Educação Superior defasada com relação à necessidade de ouvir os alunos, que são os principais protagonistas do processo de aprendizagem?

Como anteriormente adiantado, na maioria das Instituições de Ensino Superior, as Avaliações Institucionais Docentes são apresentadas aos alunos na forma de questões fechadas e isto deve-se à facilidade que este modelo oferece em sua tabulação: respostas fechadas permitem a tabulação dos dados em tempo real, ou seja, assim que o entrevistado responde, o avaliador pode fazer a análise do conjunto das respostas no mesmo instante.

Para exemplificar, pode-se citar o SAI - Sistema de Avaliação Institucional do CEETPS - Centro Estadual de Educação Tecnológica Paula Souza (CENTRO Paula Souza, 2014), que destina-se a avaliar anualmente o desempenho de todas as Etecs (Escolas Técnicas de nível médio) e Fatecs (Instituições de nível superior) que, entre outros objetivos estratégicos podemos elencar os itens: a) "obter a satisfação dos públicos que se relacionam com o Centro Paula Souza" e b) "alcançar e manter o grau de excelência diante do mercado em seus processos de ensino e aprendizagem" que estão ligados diretamente com a avaliação docente. Neste sistema de avaliação, o desempenho das unidades de ensino (Etecs e Fatecs) é expresso por indicadores

agregados, construídos a partir de um sistema de pontuação aplicado às respostas da comunidade escolar a questões de modelos específicos para cada segmento (alunos, professores, pais, etc.). As categorias destes indicadores são integradas em dimensões e estas por áreas (Tabela 1).

Categoria	Dimensão	Área
Insumo	Condições de ensino	Adequação do espaço físico e instalações
		Materiais didático-pedagógicos
Processo	Desempenho escolar	Rendimento escolar
		Acompanhamento e avaliação da aprendizagem
		Fluxo escolar
		Frequência escolar
	Gestão escolar	Gestão democrática da escola
		Equipe gestora e liderança
		Habilitação e formação dos profissionais
	Gestão pedagógica	Planejamento pedagógico
		Práticas pedagógicas
	Ambiente educativo	Clima escolar
		Disciplina e normas de convivência
	Resultado	Desempenho escolar
Gestão pedagógica		Nível de satisfação quanto às práticas pedagógicas
Ambiente educativo		Nível de satisfação quanto ao clima escolar
Geral		Nível de satisfação em relação ao curso
Indicadores objetivos		Índice de produtividade
		TCC - Taxa de conclusão do curso

Tabela 1 - Indicadores SAI - Categorias, Dimensões e Áreas. Fonte:(CENTRO Paula Souza, 2014).

As questões apresentadas aos alunos são categorizadas através destas dimensões. Questões que envolvem ações pedagógicas e de gestão que se relacionam diretamente à aprendizagem dos alunos recebem uma pontuação maior para refletir a missão do Centro Paula Souza que privilegia ações com maior impacto na aprendizagem, para isso, cada questão apresentada aos alunos é pontuada conforme este critério.

Todas as questões são fechadas e pré-determinadas pela Instituição. Por exemplo, dentro da área "acompanhamento e avaliação de aprendizagem" da dimensão "desempenho escolar" ao aluno é perguntado: "Em sala de aula, os professores dão atenção individual quando você e seus colegas apresentam dificuldades para acompanhar os conteúdos que estão sendo tratados?". E para esta questão é permitida somente as seguintes alternativas: "sim, todos; sim, grande parte deles; sim, alguns; não". Também pode-se verificar que esta mesma prática é utilizada pelo IFSP-Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

O IFSP é uma autarquia federal de ensino - fundada em 1909 e possui aproximadamente 24 mil alunos distribuídos em 39 *campi* (Instituto Federal de Educação Ciência e Tecnologia de São Paulo, 2016) e, diferentemente do CPS, segue as normas do SINAES (Sistema Nacional de Avaliação do Ensino Superior) baseadas em 10 dimensões. Pelas normas do SINAES, é criada uma Comissão Própria de Avaliação (CPA) que é responsável pela elaboração e execução de

uma Auto avaliação Institucional, um dos componentes do sistema.

Para o IFSP, em seu relatório de Auto avaliação Institucional de 2015, podemos retirar as seguintes questões do eixo "3 – Políticas Acadêmicas" dirigidas aos alunos:

3.4 - Você percebe que nos itens abaixo as ações estão implantadas de acordo com as políticas estabelecidas?

	não conheço	insuficiente	suficiente	muito boa	excelente	Sem resposta
3.4 a - pesquisa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.4 b - iniciação científica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.4 c - tecnológica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.4 d - artística	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.4 e - cultural	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3.5 - Você percebe que o apoio à realização dos itens abaixo é:

	não conheço	insuficiente	suficiente	muito boa	excelente	Sem resposta
3.5 a - programas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.5 b - projetos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.5 c - atividades	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.5 d - ações de extensão	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 2 - Exemplo do formulário de avaliação institucional do IFSP. Fonte: (IFSP,2016).

Ao percorrer este relatório do IFSP, nota-se que todas as perguntas formuladas para os alunos são apresentadas conforme a Figura 2 e que não há espaço para o aluno expor suas percepções subjetivas sobre as práticas docentes. Inclusive pode-se destacar que, neste formulário, não há nenhuma menção às práticas pedagógicas dos docentes. Isto reforça a hipótese de que, para Instituições de Ensino de grande porte, tais como o CEETPS e o IFSP, as Avaliações Docentes abertas - onde o aluno pode discorrer sobre suas impressões, inquietações ou elogios – são de difícil operacionalização, apesar da existência de recursos tecnológicos em ambas as Instituições.

2. Mineração de Dados, de Textos e Mineração de dados Educacionais

Dados são símbolos ou signos não estruturados, sem significado, como valores numéricos depositados em um banco de dados. Informação é o que está contida em uma descrição e que agrega significado e utilidade a estes dados. Conhecimento é algo que permite uma tomada de decisão e a Mineração de Dados é a ferramenta utilizada para este fim, tornando-se um elemento central responsável pela parte analítica do *Big Data* (Castro, 2016).

Segundo Silva (2015), a Mineração de Dados pode ser definida como um processo automático ou semiautomático para explorar analiticamente grandes bases de dados, com a finalidade de descobrir padrões relevantes e importantes para embasar o conhecimento do usuário final (geralmente tomadores de decisão). Denominadas também de técnicas de *Data Science* ou de *Data Mining*, elas procuram encontrar padrões inesperados nos dados (padrões não óbvios) e desta forma, ir além do que já é conhecido pelo usuário.

Dentro do contexto educacional, é possível minerar dados de alunos para verificar a relação entre uma abordagem pedagógica e o aprendizado do aluno. Através desta informação o professor poderia compreender se sua abordagem realmente está ajudando o aluno e desenvolver novos métodos de ensino mais eficazes. Para alcançar tais objetivos, as ferramentas de Mineração de Dados aplicam uma ou mais técnicas de Inteligência Artificial e de

Estatística (entre elas a Análise de Regressão, Clusterização e Análise de Séries Temporais). Inmon, Terdeman e Imhoff (2001) fazem uma interessante analogia com a mineração de pedras, onde com um trabalho árduo e uma preparação apropriada, podem ser obtidos diamantes e, sem uma metodologia e o trabalho apropriado, será bem provável conseguir um ralo financeiro e muita frustração (Inmon et al., 2001, p. 141).

Como as respostas dos alunos obtidas pelas Avaliações Institucionais Docente são predominantemente extraídas através de textos (obtidos através de formulários eletrônicos ou em papel), faz-se útil a aplicação de uma disciplina da Mineração de Dados: a Mineração de Textos.

A Mineração de Textos envolve a contribuição de muitos componentes e conhecimentos, dentre elas a Estatística, a Inteligência Artificial e a Linguística. Estas contribuições estão inseridas em todos os círculos de atividades, desde a recuperação de informação, passando pela extração de informação, a classificação e até a mineração na *web*. Termos como *web mining*, *web analytics*, *text analytics* estão todos relacionados à Mineração de Textos, que busca prover informações ricas sobre os textos para uma futura tomada de decisão. As atividades que compõe o processo de Mineração de Textos estão representadas na Figura 3.



Figura 3. - Análise Textual e seus processos. Fonte: Adaptado de (Miner et al., 2012).

Outra especialização da Mineração de Dados é a Mineração de Dados Educacionais (MDE), um campo emergente que aplica técnicas de Mineração de Dados em informações obtidas de ambientes educacionais a fim de descobrir padrões ou evidências científicas sobre alunos e também sobre as formas de aprendizagem.

“Assim, é possível compreender de forma mais eficaz e adequada os alunos, como eles aprendem, o papel do contexto na qual a aprendizagem ocorre, além de outros fatores que influenciam a aprendizagem. Por exemplo, é possível identificar em que situação um tipo de abordagem instrucional (e.g. aprendizagem individual ou colaborativa) proporciona melhores benefícios educacionais ao aluno. Também é possível verificar se o aluno está desmotivado ou confuso e, assim, personalizar o ambiente e os métodos

de ensino para oferecer melhores condições de aprendizagem (R. Baker, Isotani, & Carvalho, 2011).”

O crescimento do volume de dados educacionais em diferentes e heterogêneas origens de dados levou a necessidade de pesquisas com o foco e objetivos específicos ao ambiente educacional, que foram elencados por Jindal e Borah (2013) e divididos em objetivos acadêmicos e administrativos. Os objetivos acadêmicos da MDE envolvem informações orientadas para pessoas (estudantes, professores), orientada para os departamentos e instituições (novos cursos, adequação às novas demandas) e orientadas ao domínio (relativo a uma determinada área de atuação). Os objetivos administrativos são relacionados à gestão, como por exemplo, ações de infraestrutura escolar.

O trabalho de Hegazi e Abugroon (2016) também faz um agrupamento das áreas mais relevantes da MDE dentro das Instituições de Ensino Superior: pesquisas na área de gerenciamento de cursos (gestão de cursos e como melhorar o apoio ao aluno), pesquisas sobre o comportamento do aluno (prever a resposta do aluno para determinadas atividades propostas), sistemas de apoio à decisão (análise do processo educacional, tais como matrículas e seleção de cursos) e aplicações para evitar a evasão e reter o aluno, tal como prever qual aluno irá abandonar um curso ou prever quanto tempo um aluno irá retornar ao curso após trancá-lo (Hegazi & Abugroon, 2016).

Os componentes chave da Mineração de dados educacionais estão apresentados na Figura 4.



Figura 4 – Componentes da Mineração de dados educacionais. Adaptado de (Jindal & Borah, 2013).

Apontado na Figura 4 como um dos componentes da MDE, os *stakeholders* podem ser divididos em três grupos:

1. Primário: Personagens que estão diretamente envolvidos com o processo de ensino aprendizagem, como por exemplo, os alunos e professores;
2. Secundário: Atores que tem um envolvimento indireto com o crescimento da instituição de ensino, tais como os pais e ex-alunos;
3. Híbrido: Grupo de pessoas envolvidas no processo administrativo, ou seja, os administradores, funcionários e especialistas.

O ambiente, outro elemento que pode ser visto na Figura 4, envolve a sala de aula, os ambientes de aprendizagem virtual e modelos como a aprendizagem colaborativa e os sistemas adaptativos. Os dados para a tomada de decisão podem ser classificados como *off-line* (atendimento estudantil, informações de cursos) e *online* (*weblogs*, *e-mails*, planilhas, etc.). As tarefas podem ser de tomada de decisão para os *stakeholders* híbridos e baseadas na

aprendizagem para os *stakeholders* primários.

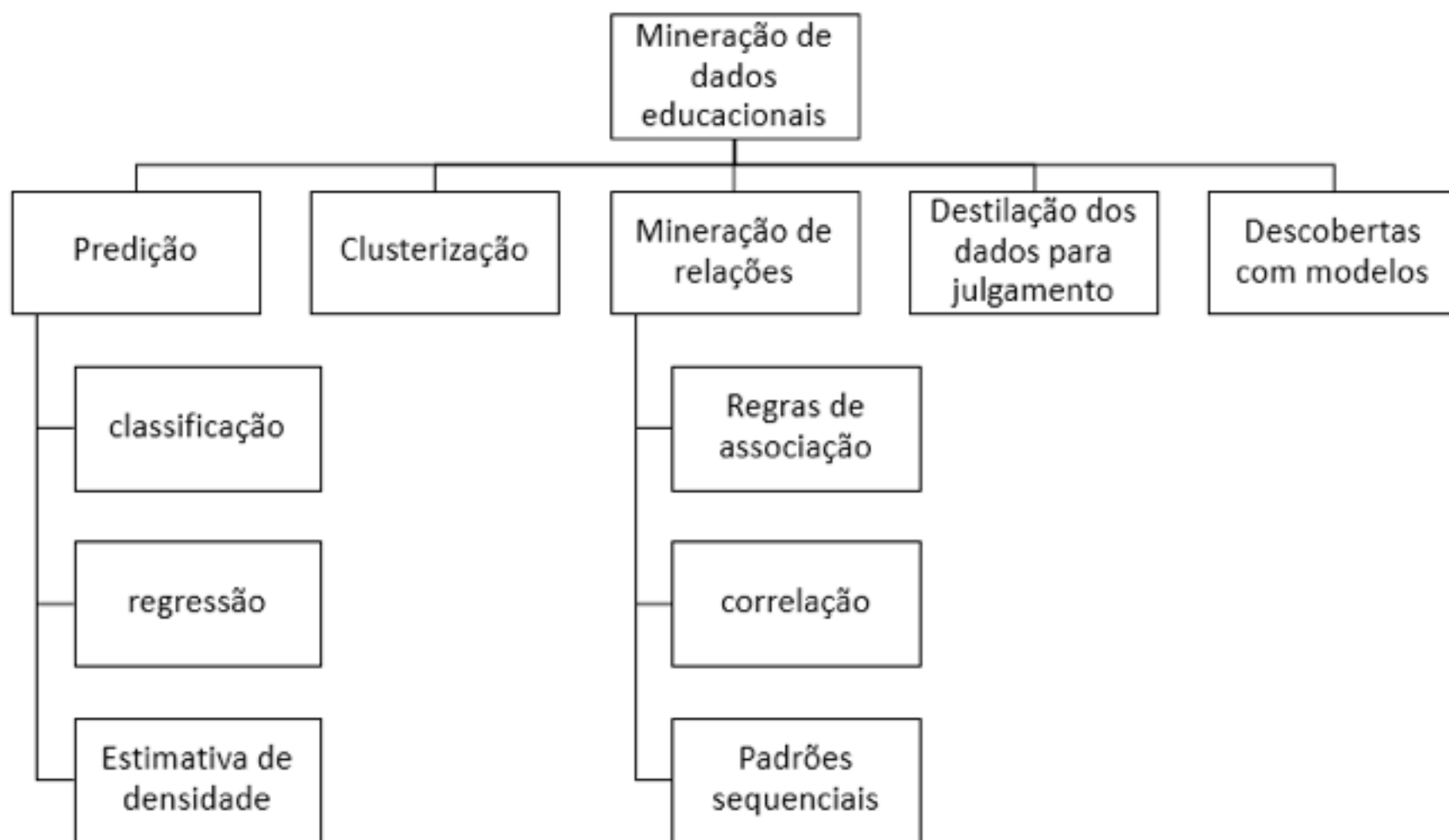


Figura 5 – Taxonomia das principais subáreas de pesquisa em MDE (Baker, 2010).

Os métodos de MDE podem ser divididos em dois grupos: os orientados para verificação (estatísticas, teste de hipótese, análise de variância, etc.) e os orientados à descoberta (predições, classificação, análise de clusterização, redes neurais, *web mining*, etc.), no entanto, segundo Baker (R. S. J. D. Baker, 2010), as técnicas diferem frequentemente das técnicas da mineração de dados tradicional, pois devem explorar níveis de hierarquia e organização dos dados educacionais. A figura abaixo apresenta uma taxonomia das principais subáreas de pesquisa em MDE.

Questões como tempo, a sequência de captura dos dados e o contexto em que esta coleta foi aplicada, tem um papel importante no estudo de dados educacionais e que serão respeitadas na aplicação das práticas requisitadas no modelo proposto adiante.

3. Análise de Sentimentos

Em termos da compreensão da linguagem natural, a Análise de Sentimentos é considerada como uma importante subárea da Análise Semântica com o objetivo de identificar não somente os assuntos que as pessoas falam e seu sentimento sobre esses assuntos, mas também se propõe a identificar características de um texto de acordo com um assunto específico, analisando e classificando-os em subgrupos pré-definidos, como sentimentos e emoções (feliz, triste), polaridade (positivo, negativo) ou em qualquer outro contexto previamente definido (Liu, 2015).

A popularidade do tema se inicia a partir do advento das redes sociais, onde várias *startups* (pequenas empresas de tecnologia) surgiram para o desenvolvimento de ferramentas e estratégias para a extração de sentimentos em textos.

As técnicas de Análise de Sentimentos podem ser resumidas na taxonomia da Figura 6 (Medhat, Hassan, & Korashy, 2014).

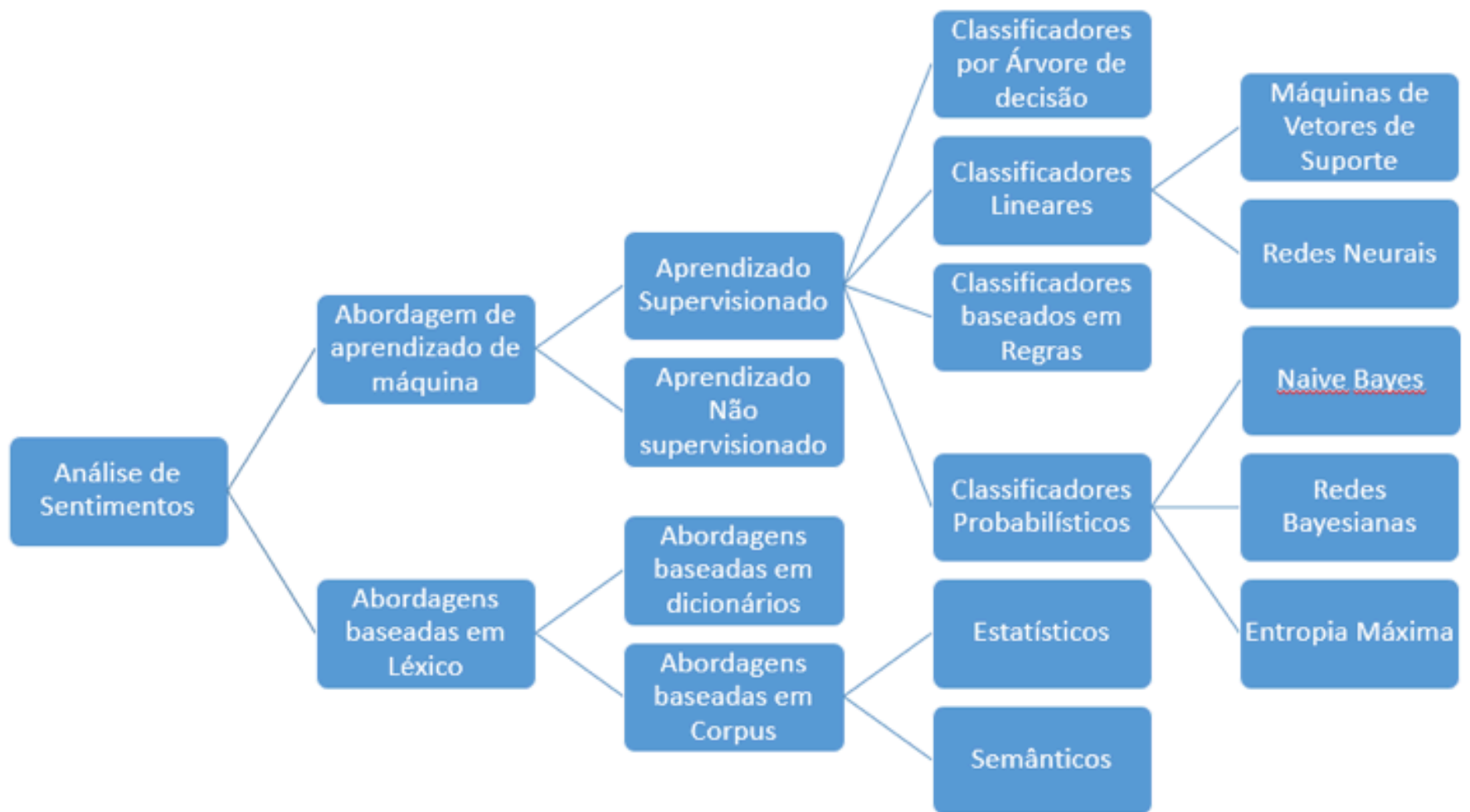


Figura 6 - Técnicas de Análise de Sentimentos. Adaptado de (MEDHAT; HASSAN; KORASHY, 2014).

A Análise de Sentimentos pode ser realizada através de diversas abordagens e que podem ser subdivididas em duas grandes áreas: técnicas de Aprendizagem de Máquina e das abordagens baseadas em Léxico (Figura 6).

A área de pesquisa chamada de Aprendizagem de Máquina visa desenvolver programas ou modelos computacionais capazes de - automaticamente - melhorar seu desempenho por meio de experiências. Modelos de Aprendizagem de Máquina podem ser supervisionados ou não-supervisionados. No primeiro modelo, os algoritmos ajustam parâmetros de um modelo a partir do erro medido entre respostas obtidas e esperadas. No segundo, os parâmetros de um modelo são ajustados com base na maximização de medidas de qualidade das respostas obtidas (Brunialti, Freire, Peres, & Lima, 2015), ou seja, o modelo deve aprender a rotular os dados.

Nos métodos baseados em léxico é criado um dicionário de palavras com sua orientação semântica conhecida (positiva, negativa ou neutra sobre o assunto) para realizar a classificação do documento, opinião ou frase. Nos métodos baseados em corpus, faz-se uso de técnicas sintáticas ou estatísticas, como a verificação da ocorrência concomitante de palavras com outras que já possuam sua orientação semântica conhecida.

A Análise de Sentimentos estuda principalmente as opiniões que expressam ou implicam sentimento positivo ou negativo. Opinião é um conceito abrangente que cobre sentimentos, avaliação, estima, ou ainda, a atitude e as informações associadas a um tópico ou uma opinião alvo e também à pessoa que detém uma determinada opinião. Para a Análise de Sentimentos, o termo "sentimento" significa apenas o sentimento positivo ou negativo subjacente implícito nesta opinião (LIU, 2015, p. 17). Tome-se a avaliação de um produto (Tabela 2):

Avaliação A: postada por John Smith	Data: 10 de setembro de 2011
(1). Eu comprei uma câmera Canon G12 há seis meses atrás. (2). Eu simplesmente adorei. (3) A qualidade das	

fotos é incrível. (4) A bateria dura muito tempo. (5). No entanto, a minha esposa acha ela muito pesada.

Tabela 2 - Exemplo de uma avaliação de produto. Fonte: (Liu, 2015).

A tabela acima foi dividida e numerada para exemplificar como a Análise de Sentimentos trabalha, identificando:

1. Opinião, sentimento e sentimento alvo (ou Tópico): A opinião contém várias opiniões com possíveis polaridades positivas ou negativas sobre a Cannon G12. A sentença 2, generaliza o produto como positivo. A sentença 3 positiva a qualidade das fotos. A sentença 4 mostra um sentimento positivo em relação à vida da bateria. Finalmente, a sentença 5, expressa uma polaridade negativa sobre a qualidade "peso" da Cannon G12. Este exemplo possui dois componentes chaves para a Análise de Sentimentos: um tópico ou alvo g e um sentimento s onde g pode ser qualquer entidade ou aspecto de uma entidade na qual uma opinião foi emitida, e s pode ser positivo, negativo ou neutro ou ainda estimado de forma numérica. Sentimento Positivo, Sentimento Negativo ou Sentimento Neutro são denominados Sentimentos ou Polaridade da Opinião. O tópico ou sentimento alvo na sentença 2 é a Cannon G12 e na sentença 3, um dos atributos da Cannon G12, a sua qualidade de fotos.
2. Detentor da opinião: A avaliação acima contém a opinião de duas pessoas, que podemos chamar de Detentor da opinião (*opinion holder*) ou fontes de opiniões (*opinion sources*) (Wiebe & Ellen, 2005). O detentor da opinião surge nas sentenças 2, 3 e 4, porém na sentença 5 quem opina é a esposa do autor.
3. Data da opinião: A data desta avaliação é 10 de setembro de 2011. A data é uma informação importante pois frequentemente, a evolução de uma opinião - ou como estas opiniões mudam ao longo de um período de tempo - é de interesse de uma Análise de Sentimentos.

Liu (2015) apresenta uma simplificação matemática da estrutura de dados para a Análise de Sentimentos através da seguinte quintupla: (ej, ajk, sijkl, hi, tl) onde:

- ej representa a entidade que está sendo analisada (um político, uma empresa, um celular);
- ajk, é o aspecto em questão (confiança, pontualidade, beleza);
- sijkl representa o sentimento da entidade em relação à questão, que pode assumir positivo (+), negativo (-) ou neutro;
- hl permite classifica o detentor da opinião (cliente, crítico);
- tl armazena a data da pesquisa de opinião.

Trata-se de uma estrutura de dados essencial para a Análise de Sentimentos e a falta de um destes elementos pode tornar a análise problemática. A ausência de tl (data) pode causar uma falha na análise pois as opiniões mudam ao longo do tempo. A falta de hl (detentor) pode fazer muita diferença na análise de um discurso (esta opinião é do presidente dos EUA ou do presidente do Congresso Nacional Brasileiro?).

Nos últimos anos, surgiu um considerável interesse no desenvolvimento de iniciativas de ensino e aprendizagem para aprimorar a retenção e progressão dos alunos (Glenda et al apud NOVAK; COWLING, 2011) e tais iniciativas passam pelo interesse em saber o que o aluno pensa destas práticas.

O trabalho de Altrabsheh, Cocea e Fallahkhair (2014) descreve um modelo de um sistema em tempo real para analisar o feedback de alunos durante aulas e em seguida apresentar seus resultados ao professor usando Análise de Sentimentos. Também realiza um comparativo entre as diferentes técnicas de aprendizagem de máquina, diferentes combinações de n-gramas, níveis de pré-processamento dos dados e do uso da classe neutra.

Kim e Calvo (2010) trabalharam com dois objetivos: o primeiro foi avaliar a viabilidade do uso da Análise de Sentimentos no estudo das respostas de 909 estudantes no USE (*Unit of Study Evaluation*) e o segundo, avaliar o mérito dos modelos categóricos (usando a *WordNet-Affect*) e dimensionais (onde as emoções são apresentadas em duas ou três dimensões) para a conceitualização das emoções.

Na Finlândia, país reconhecido pelas suas inovações no campo educacional, os professores são encorajados a encontrar novas soluções para melhorar o processo ensino-aprendizagem. O trabalho de Munzero e outros (2013) aplica técnicas de Análise de Sentimentos nos Diário de Aprendizagem dos alunos. Neste trabalho é possível observar que os obstáculos emocionais são conhecidos por atrapalharem o processo de aprendizagem: os alunos aprendem mais quando se sentem bem e confortáveis com determinada disciplina ou assunto e estes diários ajudam nesta direção. Entender como o aluno (de forma personalizada) aprende e quais os seus sentimentos sobre aquela disciplina, permite ao docente melhorar a qualidade de sua aula.

4. Desenvolvimento da Pesquisa

Na tentativa de ampliar o leque de dimensões a serem consideradas por uma Avaliação Institucional, em especial, a Avaliação Docente e mais precisamente, das práticas pedagógicas dos professores, esta pesquisa propõe um modelo computacional para utilizar a Mineração de Dados Educacionais, em especial, a Análise de Sentimentos para obter novas (ou confirmar as já existentes) dimensões de análise das práticas docentes a partir das percepções dos alunos.

Neste modelo, a captação dos dados é realizada por um formulário que é apresentado de forma aberta através de dois quadros onde o aluno deve falar “bem” ou “reclamar” de seus professores, e, onde as questões norteadoras do processo de avaliação deverão emergir das próprias respostas e opiniões dos alunos (Figura 8).

A Web Page

← → × ↕ http://

Prezado aluno(a), estamos realizando uma pesquisa, com o objetivo de compreender a percepção que os estudantes desta Faculdade têm sobre a prática pedagógica dos professores e, para tanto, gostaríamos da sua colaboração voluntária, através do preenchimento deste formulário. Agradecemos sua participação. Não é necessário identificar-se.

FALE BEM DOS PROFESSORES

RECLAME DOS PROFESSORES

ENVIAR

⌵

Figura 7 - Protótipo da interface do formulário. Baseado em (LECHUGO, 2016).

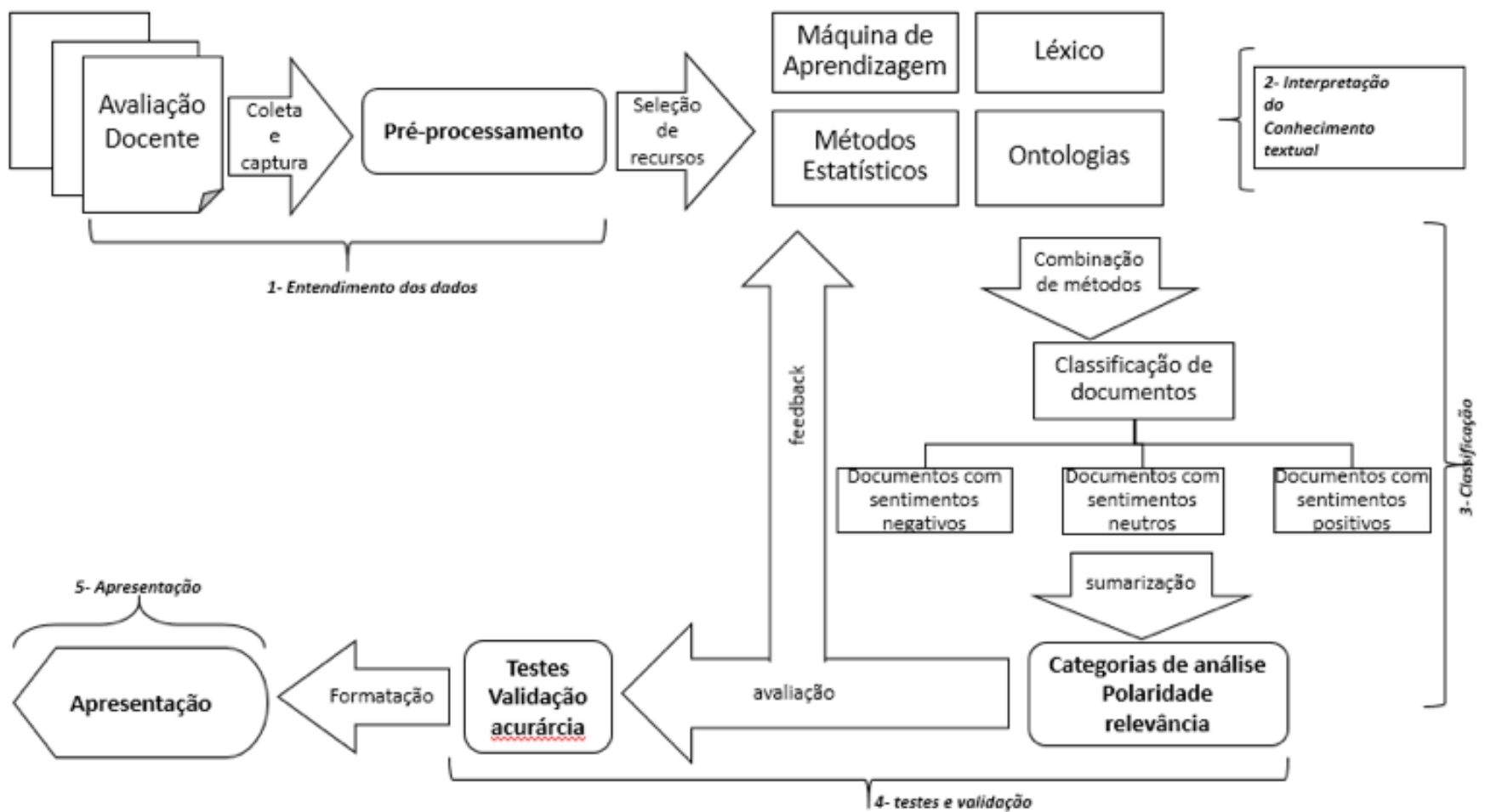


Figura 8 - Modelo Computacional para aplicação de Mineração de Dados Educacionais para Avaliação Institucional Docente. Fonte: Autoria própria.

A partir das respostas obtidas através da interface proposta, inicia-se o processo de Mineração destes dados, com o objetivo de identificar categorias de análise que possam “saltar” das afirmações dos alunos, das práticas apontadas como mais relevantes nas práticas docentes e também qual a orientação semântica das respostas dos alunos em relação ao seu corpo docente, ou seja, se elas são positivas ou negativas (Figura 7).

No modelo proposto na Figura 8, pode-se identificar cinco etapas principais:

1. Etapa de entendimento dos dados: Extraídos das ferramentas de coleta de dados (em formulários impressos ou eletrônicos), as opiniões dos alunos passam por algumas tarefas de pré-processamento e de limpeza de dados. Dentre as tarefas desta fase, podemos citar as atividades de tokenização, remoção de palavras não discriminantes, stemização, entre outras.
2. Etapa do conhecimento textual: Nesta fase os *tokens* obtidos no pré-processamento serão transformados em vetores de entrada para os algoritmos de Análise de Sentimento e diversas abordagens serão utilizadas: máquinas de aprendizagem, léxicos ou dicionários, ontologias e abordagens estatísticas.
3. Classificação: As opiniões dos alunos serão agrupadas por categorias de análise, que serão levantadas através de métodos estatísticos, o que permitirá o surgimento de novas categorias de análise das práticas docentes que não são contempladas sob a ótica atual (a do avaliador). A sumarização será responsável por identificar as categorias mais abordadas pelos alunos.
4. Testes e validação: Onde será definida a acurácia do modelo. Não atingido um certo nível de acurácia, o processo é refeito.
5. Apresentação: Finalmente, se os dados de resultados atingirem o nível de acurácia satisfatório, eles serão exibidos da forma mais amigável através de ferramentas de visualização de dados.

Para apurar a acurácia deste modelo computacional, os resultados obtidos serão confrontados com os resultados obtidos em uma pesquisa de Avaliação Institucional Docente realizada em uma Instituição de Ensino Superior Tecnológico (Lechugo, 2016), onde não foram usados modelos estatísticos e de Mineração de Dados Educacionais. Através desta análise comparativa, espera-se que novas categorias de análise das práticas pedagógicas possam surgir, assim

como, de forma automática, as mesmas categorias que foram levantadas de forma “manual” sejam identificadas, de forma a corroborar a utilidade deste formato de Avaliação Docente e ampliar os parâmetros de qualidade das Instituições, além de fomentar as discussões das práticas docentes mais eficientes (ou não) no processo de aprendizagem do aluno.

5. Considerações Finais

Cada vez mais as Instituições de Ensino Superior são pressionadas para aumentar seus índices de qualidade e diminuir a evasão escolar. Testes padronizados para a verificação da aprendizagem de concluintes, tal como o ENADE (Exame Nacional de Desempenho do Estudante) tornam-se cada dia mais importantes para verificar a qualidade da Educação. Com relação à evasão escolar, não há nas Instituições de Educação Superior nenhum instrumento que detecte formalmente as razões pelas quais o aluno abandona a escola, e esses motivos não identificados, acabam por não contribuir com o processo de melhoria da aprendizagem de alunos e professores. Portanto, a identificação e a organização destes dados com relação às satisfações e as insatisfações dos alunos se tornam tão necessária. Além disso, o gerenciamento destes dados de forma *on-line*, permite que professores e gestores tomem decisões rápidas e efetivas para corrigir os erros do processo.

Através deste trabalho, espera-se divulgar o uso de técnicas de mineração de dados educacionais, em especial a aplicação da Análise de Sentimentos em Avaliações Institucionais Docentes. Com isso, busca-se demonstrar a riqueza das análises não-amostrais que, a partir de todas as respostas dos alunos, e sem nenhum formato pré-concebido por parte do avaliador, permite identificar quais práticas pedagógicas são consideradas boas e quais não são sob a ótica dos alunos. Também, pretende-se contribuir com possíveis novas categorias de análises importantes para a avaliação do trabalho docente. Porém, as experiências positivas e negativas reportadas pelos alunos, não retira do professor e da Instituição a autoridade e responsabilidade da análise e tomada de decisão, até porque o objetivo da MDE em uma Avaliação Institucional Docente, não objetiva “entender” os alunos para “atendê-los”, mas sim, de compreender a realidade que *temos*, para refletirmos sobre que (alunos – professores – Instituição) *queremos* na Educação Superior.

Ao realizar a revisão da literatura para esta pesquisa, percebe-se que existem diversas contribuições da MDE na identificação do estado emocional do aluno e esta relação com o seu próprio aprendizado (principalmente aplicados ao ambiente de Ensino à Distância), mas pouco sobre o aluno e sua relação com a prática docente. A relevância deste trabalho está em ouvir principalmente os alunos e fazer com que eles tenham sua participação ampliada nas discussões sobre os rumos da Educação Superior a partir das mudanças nas práticas pedagógicas dos professores. Desta forma os alunos deixam de ser meros espectadores para se tornarem protagonistas na reconstrução dos modelos e dos papéis do professor e também dele próprio dentro da Instituição.

Referências Bibliográficas

- Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2014). Sentiment Analysis: Towards a Tool for Analysing Real-Time Students Feedback. *2014 IEEE 26th International Conference on Tools with Artificial Intelligence*, 419–423. <http://doi.org/10.1109/ICTAI.2014.70>
- Baker, R., Isotani, S., & Carvalho, A. (2011). Mineração de Dados Educacionais: Oportunidades para o Brasil. *Revista Brasileira de Informática na Educação*, 19(02), 03. <http://doi.org/10.5753/RBIE.2011.19.02.03>
- Baker, R. S. J. D. (2010). Data mining for education. *International Encyclopedia of Education*, 7, 112–118. <http://doi.org/10.4018/978-1-59140-557-3>
- Brunialti, L. F., Freire, V., Peres, S. M., & Lima, C. A. M. (2015). aprendizado de máquina em sistemas de recomendação baseados em conteúdo textual uma revisão sistematica. In *XI Brazilian Symposium on Information System*. Goiania, GO. Recuperado de

<http://www.lbd.dcc.ufmg.br/colecoes/sbsi/2015/029.pdf>

Castro, L. N. de. (2016). *Introdução à Mineração de dados*. São Paulo: Saraiva.

CENTRO Paula Souza. (2014). *SAI – Sistema de Avaliação Institucional Centro Paula Souza: Fundamentos Teórico-Methodológicos*. São Paulo.

Grohmann, M. Z., & Ramos, M. S. (2012). Competências docentes como antecedentes da Avaliação De Desempenho do Professor: Percepção de mestrandos de administração. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 17, 65–86.

Hegazi, M. O., & Abugroon, M. A. (2016). The State of the Art on Educational Data Mining in Higher Education. *International Journal of Computer Trends and Technology*, 31(1), 46–56. <http://doi.org/10.14445/22312803/IJCTT-V31P109>

Inmon, W. H., Terdeman, R. H., & Imhoff, C. (2001). *Data Warehousing*. São Paulo: Berkeley.

Instituto Federal de Educação Ciência e Tecnologia de São Paulo. (2016). Página da Comissão Própria de Avaliação Institucional do IFSP. Recuperado de <http://www.ifsp.edu.br/cpa/pesq2015.php>

Jindal, R., & Borah, M. D. (2013). A Survey on Educational Data Mining and Research Trends. *International Journal of Database Management Systems (IJDMS)*, 5(3), 53–73. <http://doi.org/10.5121/ijdms.2013.5304>

Johnson, L., Adams Becker, S., Estrada, V., Freeman, A., Karpylis, P., Vuorikari, R., & Punie, Y. (2014). *Horizon Report Europe: 2014 Schools Edition*. Brussels, Belgium. Recuperado de https://ec.europa.eu/jrc/sites/default/files/2014-nmc-horizon-report-eu-en_online.pdf

Kim, S. Mac, & Calvo, R. A. (2010). Sentiment Analysis in Student Experiences of Learning. *Third International Conference on Educational Data Mining (EDM2010)*, 111–120.

Lechugo, C. P. (2016). *A Educação Tecnológica: sua história, seus professores e a percepção dos alunos sobre as práticas pedagógicas*. Universidade de Sorocaba.

Liu, B. (2015). *Sentiment Analysis*. New York City, USA: Cambridge University Press. Recuperado de www.cambridge.org/9781107017894

Mayer-Schonberger, V., & Cukier, K. (2013). *Big Data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana* (1o ed). São Paulo: Elsevier Ltd.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <http://doi.org/10.1016/j.asej.2014.04.011>

Miner, G., Delen, D., Elder, J., Fast, A., Hill, T., & Nisbet, B. (2012). *Practical Text Mining and Statistical Analysis for non-structured text data applications*. Oxford: Elsevier.

Munezero, M., Montero, C. S., Mozgovoy, M., & Sutinen, E. (2013). Exploiting sentiment analysis to track emotions in students' learning diaries. *Proceedings of the 13th Koli Calling International Conference on Computing Education Research - Koli Calling '13*, 145–152. <http://doi.org/10.1145/2526968.2526984>

Novak, J., & Cowling, M. (2011). The implementation of social networking as a tool for improving student participation in the classroom. *ISANA International Academy Association Conference*, 1–10. Recuperado de [https://aicportal.cqu.edu.au/CMS/HumanResources/UploadFiles/Research/50/Cowling Novak ISANA 2011 Paper.pdf](https://aicportal.cqu.edu.au/CMS/HumanResources/UploadFiles/Research/50/Cowling%20Novak%20ISANA%202011%20Paper.pdf)

Silva, L. A. da. (2015). *Mineração de dados: uma abordagem introdutória e ilustrada*. São Paulo: Editora Mackenzie.

Wiebe, J. M., & Ellen, R. (2005). Creating Subjective and Objective Sentence Classifiers from Unannotated Texts. *Computational Linguistics and Intelligent Text Processing*, 3406, 486–497. <http://doi.org/10.1007/b105772>

1. Doutorando em Computação pela Universidade Mackenzie (SP), Mestre em Educação pela Universidade de Sorocaba (SP), graduado em Processamento de Dados pela Fatec de Sorocaba (SP) com 5 especializações: *Vocational Teacher Education* pela HAMK University of Applied Science (Finlândia), Gestão de Ambiente EAD pela UFF (RJ), Administração de Empresas pela FGV (RJ), Engenharia de Software pela Univ. São Judas Tadeu (SP) e em Docência no Ensino Superior pela Academia de Ensino Superior (SP). Atualmente é professor do curso de Análise e desenvolvimento de sistemas do Instituto Federal de São Paulo (SP) e do Centro Paula Souza (SP). E-mail: fabio.santos@ifsp.edu.br

2. Possui Graduação em Matemática-Informática pela Universidade Federal de Juiz de Fora (MG), Mestrado em Ciências pelo Instituto Tecnológico de Aeronáutica (SP) e Doutorado em Engenharia Elétrica pela Universidade de São Paulo (SP). Atualmente é Professor Adjunto I da Universidade Presbiteriana Mackenzie (SP), atuando nos cursos de Graduação da Faculdade de Computação e Informática e no Programa de Mestrado e Doutorado em Engenharia Elétrica e Computação; é Professor Titular II da Universidade Cruzeiro do Sul (SP), onde coordena o Bacharelado em Ciência da Computação e atua nos Programas de Mestrado (Profissional e Acadêmico) e Doutorado em Ensino de Ciências e Matemática. E-mail: ismar@mackenzie.br

3. Doutora e Mestre em Educação pela Universidade de Sorocaba, bacharel em Administração e pós-graduada em Metodologia do Ensino Superior. Atualmente é professora e coordenadora de curso no Centro Paula Souza (SP) e de cursos de pós-graduação nas áreas de Gestão de Negócios. E-mail: carlalechugo@gmail.com

Revista ESPACIOS. ISSN 0798 1015
Vol. 38 (Nº 05) Año 2017

[Índice]

[En caso de encontrar algún error en este website favor enviar email a webmaster]