









REPTree, JRip e K\*, os autores construíram modelos para os estados afetivos tédio, confusão, concentração comprometida e frustração.

Os resultados destes trabalhos mostram que o uso das técnicas de MDE no estudo da afetividade do aluno é promissor; portanto, pertinente à realização de mais pesquisas na área. Nesse sentido, a próxima seção apresenta os métodos de mineração, bem como possíveis abordagens no que diz respeito aos dados afetivos em ambientes virtuais de aprendizagem.

### 3.1. Métodos de Mineração de Dados Educacionais

Métodos de mineração têm sido utilizados para prever o desempenho e detectar os comportamentos dos estudantes, tais como predição, classificação, regressão, estimativa de densidade, clusterização, detecção de outliers, análise de redes sociais, mineração de relacionamentos, processos e texto.

Este estudo investiga como a MDE pode auxiliar no processo de ensino e de aprendizagem em AVA, em especial sobre aspectos relacionados à afetividade. Para tanto, utilizou-se o método de clusterização a fim de obter os perfis motivacionais dos alunos.

De acordo com Han e Kamber (2001) a análise de *cluster* refere-se ao processo de agrupamento em classes similares de objetos, sejam estes físicos ou abstratos. Desse modo, define-se cluster como sendo uma coleção de objetos que compartilham aspectos comuns segundo critérios pré-estabelecidos. Tais aspectos são fatores que distinguem cada coleção de objetos. Em termos de algoritmo, a clusterização consiste em um tipo de aprendizado não-supervisionado, uma vez que os atributos não são definidos previamente.

Segundo De Amo (2004) na realização de tarefas de clusterização são usados vários algoritmos sobre o mesmo conjunto de dados. Após a mineração, os *clusters* obtidos são analisados. Dentre os algoritmos de aprendizagem não-supervisionada que tratam do problema de clusterização, o K-means é um dos mais difundidos.

Segundo Silva et al., (2016) o K-Means é considerado o principal algoritmo de clusterização. A abordagem por particionamento requer que o usuário defina a quantidade de *clusters* que deseja formar. A partir dessa informação, o algoritmo divide o conjunto em partes disjuntas dentro do parâmetro estabelecido satisfazendo o seguinte conjunto de regras: (1) os elementos numa mesma parte, obedecendo o critério fornecido, estão próximos e; (2) seguindo esse mesmo critério, os elementos de partes distintas estão distantes.

As regras operadas pelos métodos de particionamento para obter essa subdivisão são as seguintes: é criada uma partição inicial aleatória de k partes e na sequência, através de repetidas iterações, os elementos das partes vão sendo realocados para outras, de forma a refinar o particionamento a cada iteração. Isto significa que em uma nova iteração, cada parte passa a conter objetos que efetivamente estão próximos (SILVA et al., 2016).

Nesse contexto, ressalta-se que a proposta deste trabalho é usar a mineração de dados educacionais aplicando clusterização para inferir os fatores motivacionais captados das interações dos alunos em um AVA. Pretende-se identificar a relação da predominância dos traços de personalidade com o comportamento manifestado pelo aluno.

## 4. Metodologia

O objetivo da mineração de dados educacionais é o de identificar padrões de clusterização de alunos que compartilhem as mesmas características quanto aos fatores motivacionais e os de predominância dos traços de personalidade. Vislumbra-se obter um conjunto de clusters com distintas combinações do grau de motivação do aluno com o

fator de predominância dos traços de personalidade de modo a derivar os diferentes perfis que culminam em um aluno motivado, desmotivado ou neutro.

Nesse sentido, foram selecionados o total de 375 registros, no qual cada um dos registros equivale a um aluno, extraídos do AVA ROODA, referente às interações dos estudantes em disciplinas de cursos de Pós-Graduação da Universidade Federal do Rio Grande do Sul. O processo adotado foi composto por sete etapas, conforme ilustra a Figura 2, as quais são detalhadas na sequência.



Figura 2 - Processo de mineração dos dados educacionais.

O conjunto de dados selecionados, que se refere a etapa de dados não tratados, contém 125 exemplares de combinações para cada tipo de fator de predominância de estados de ânimo (positivo, negativo e ambíguo), com os graus dos fatores motivacionais de confiança, esforço e independência.

Cada fator apresenta um estado que pode ser muito baixo, baixo, médio, alto ou muito alto, e a inferência do estado motivacional do aluno que pode ser motivado, desmotivado ou ainda neutro. Esses dados foram exportados em um arquivo do tipo planilha, na extensão *.Comma Separated Values* (.csv).

Na sequência, no pré-processamento dos dados, foi realizada a limpeza dos dados contidos na planilha, sendo descartados aqueles que eram desnecessários ou redundantes: duas colunas continham dados inferidos a partir das demais informações, o que caracteriza atributos dependentes, sendo portanto descartados.

Ao término da fase de pré-processamento, estavam definidos um conjunto de quatro atributos nominais (variáveis) referentes aos elementos: fator de predominância dos traços de personalidade (est\_animo\_fpred), confiança (confianca), esforço (esforco) e independência (independencia).

Na etapa subsequente, de transformação dos dados (dados modificados na Figura 2), os atributos acima referidos foram definidos de acordo com a sintaxe do minerador *Waikato Environment for Knowledge Analysis* (WEKA) e acrescidos de um atributo nominal do tipo classe, destinado a classificar cada agrupamento detectado pelo minerador em motivado, desmotivado ou neutro.

O WEKA (<https://www.cs.waikato.ac.nz/ml/weka/>) é uma suíte de software para aprendizagem de máquina, desenvolvido na linguagem de programação Java, que suporta várias tarefas de mineração de dados, tais como pré-processamento de dados, agrupamento, classificação, regressão, visualização e características de seleção. Suas técnicas são baseadas na suposição de que os dados estão disponíveis em um único arquivo ou relação. Ao término da fase de transformação, os dados foram convertidos em um arquivo em formato suportado pelo minerador, cuja extensão é do tipo *Attribute Relation File Format* (ARFF).

Na fase seguinte, envolvendo a MDE, o conjunto de dados com 5 atributos e 375 instâncias foi submetido à tarefa de clusterização usando o algoritmo K-means. Inicialmente, adotou-se a abordagem de explorar distintos tamanhos de *clusters* antes de optar por uma quantidade em definitivo.

Deste modo, foram executados três diferentes clusterizações com o algoritmo K-means, gerando na primeira execução do algoritmo oito *clusters*, na segunda nove *clusters*

e por fim dez *clusters*. O objetivo foi selecionar uma configuração de *clusters* que fornecesse a melhor quantidade de agrupamentos, porém sem que o número de instâncias do menor agrupamento fosse insignificante em relação aos agrupamentos menores, bem como se apresenta a menor soma de erro quadrático.

Tendo em vista esses critérios, a execução do K-means com solicitação de dez *clusters* foi o resultado selecionado, pois o erro quadrático apresentado pela execução com dez *clusters* foi de 666 contra 685 apresentado pela execução com nove *clusters*, e 688 da execução com a geração de oito *clusters*. Desta forma entende-se que a saída que gerou o menor erro quadrático configura-se como a mais interessante na exploração dos dados. Concluído este passo, deu-se andamento na interpretação dos resultados, abordada na próxima seção.

## 5. Discussão e Análise de Resultados

Como mencionado anteriormente, esta pesquisa de Mineração de Dados Educacionais (MDE) contou com uma amostra de 375 registros comportamentais de alunos das disciplinas de cursos de Pós-Graduação na Universidade Federal do Rio Grande do Sul. A maior parte das disciplinas foram na modalidade a distância com o suporte tecnológico do AVA ROODA para as interações entre os atores da Educação a Distância (EaD) nas temáticas desenvolvidas.

O conjunto é formado de três subconjuntos com 125 registros cada. O subconjunto contém combinações entre um fator de predominância dos traços de personalidade (positivo, negativo ou ambíguo) e os fatores motivacionais confiança, esforço e independência, em seus respectivos níveis (muito baixo (mbaixo), baixo, médio, alto ou muito alto (malto)).

A MDE foi realizada depois de finalizadas as etapas iniciais, que envolvem desde a seleção até a transformação dos dados. Neste trabalho, fez-se uso do algoritmo de particionamento K-Means, a fim de encontrar o menor erro quadrático. Assim, foram definidos dez *clusters* conforme resultado apresentado na Tabela 1.

Tabela 1 – Clusterização dos fatores de predominância e motivacionais dos alunos.

Atributo	Estados de ânimo	Confiança	Esforço	Independência	Categoria
Cluster 1	Ambíguo	mbaixo	medio	baixo	Desmotivado
Cluster 2	Positivo	medio	mbaixo	medio	Motivado
Cluster 3	Negativo	medio	malto	baixo	Desmotivado
Cluster 4	Ambíguo	alto	alto	medio	Motivado
Cluster 5	Negativo	baixo	mbaixo	mbaixo	Desmotivado
Cluster 6	Positivo	malto	baixo	alto	Motivado
Cluster 7	Negativo	mbaixo	medio	malto	Desmotivado
Cluster 8	Ambíguo	mbaixo	malto	medio	Desmotivado
Cluster 9	Ambíguo	malto	baixo	malto	Motivado
Cluster 10	Negativo	mbaixo	baixo	mbaixo	Desmotivado

Os resultados demonstram que os clusters 2, 4, 6 e 9 indicam que os alunos considerados neste agrupamento pertencem à categoria comportamental de ser motivado. Analisando os possíveis padrões, percebeu-se que quando o estado de ânimo é positivo, presente nos clusters 2 e 6, os fatores motivacionais de confiança e independência são médio, alto ou muito alto, e o esforço como muito baixo ou baixo, os alunos foram agrupados na categoria motivado. Nos clusters 4 e 9, o estado de ânimo é ambíguo, os fatores motivacionais de confiança e independência são médio, alto ou muito alto, e o esforço baixo ou alto, os estudantes ainda se mantêm na categoria motivado. Ou seja, observou-se que o esforço está na maior parte dos clusters (75%) com o nível muito baixo ou baixo, e que a confiança e a independência estão em níveis médio, alto ou muito alto, sugerindo que o aluno confiante e independente, mesmo sendo pouco esforçado está motivado.

Os clusters 1, 3, 5, 7, 8 e 10 estão na categoria de comportamento desmotivado. Averiguando os possíveis padrões, analisou-se que quando o estado de ânimo é negativo, como ocorre nos clusters 3, 5, 7 e 10, os fatores motivacionais de confiança são muito baixo, baixo ou médio, o esforço e a independência possuem níveis de muito baixo, baixo, médio ou muito alto. Percebeu-se que para a confiança a maior parte (75%) são muito baixo ou baixo, o esforço (75%) são muito baixo, baixo ou médio, e a independência (75%) são muito baixo ou baixo. Dessa forma, supõe-se que quando o aluno possui os níveis muito baixo, baixo ou médio em pelo menos dois fatores motivacionais e estado de ânimo negativo, o estudante encontra-se na categoria desmotivado. Nos clusters 1 e 8, o estado de ânimo é ambíguo, os fatores motivacionais de confiança são muito baixo, o esforço é médio ou muito alto, e a independência é baixo ou médio. Assim, constatou-se que quando o aluno possui o estado de ânimo ambíguo e a confiança muito baixo, sua categoria será de desmotivado, mesmo possuindo um esforço médio ou muito alto. Sugerindo que o aluno que é esforçado, mas não confiante e possui um estado de ânimo ambíguo, encontra-se na categoria desmotivado.

Dessa forma, podemos concluir que nos contextos investigados, o estado de ânimo aliado aos fatores motivacionais são importantes variáveis para as categorias que os alunos se encontram. Considera-se que esses resultados são um valioso ganho, pois, a confiança está ligada ao grau de segurança que o aluno possui para o desenvolvimento de suas atividades, já o esforço está relacionado ao quanto o estudante se dedica as suas tarefas e a independência está associado a autonomia dos alunos. Assim, quando esses fatores são relacionados com o estado que o aluno se sente, isso gera a motivação ou não.

Nesse sentido, o professor, de posse de tais informações, pode tomar decisões proativas e tentar diminuir o número de alunos desmotivados, o que pode gerar um menor número de retenções e evasões nos cursos. Com os dados coletados na pesquisa foi possível compreender a relação da motivação do aluno para o processo de aprendizagem, além da possibilidade de desenvolver estratégias pedagógicas condizentes com o perfil de cada estudante, principalmente ao encontrar o estado de ânimo negativo e relacionado à desmotivação.

## 6. Considerações Finais

Este estudo investigou a existência de padrões de fatores de predominância de estados de ânimo e motivacionais dos alunos em um ambiente virtual de aprendizagem. Para tanto, buscou-se identificar, através de técnicas de Mineração de Dados Educacionais, possíveis padrões relativos aos traços de personalidade e aos fatores motivacionais de confiança, esforço e independência. O método aplicado à mineração foi a clusterização, usando o



algoritmo K-means. Desse modo, foi pré-estabelecido a geração de dez clusters. Como resultados obtidos, a análise apontou a viabilidade de geração de perfis motivacionais de alunos a partir da mineração de dados educacionais.

Considera-se que a geração de perfis detalhados sobre o estado afetivo de motivação, a partir dos fatores emocionais obtidos através do comportamento observável em um AVA, pode subsidiar a criação de ferramentas de apoio ao professor.

Cabe salientar que os dados apresentados são identificações de variáveis encontradas no decorrer das análises, não sendo realizado um aprofundamento sobre a relação causal, ou não, destas inferências. É necessário novos estudos que aprofundem os valores de outras variáveis e se as mesmas influenciam, ou não, os indicadores de estado de ânimo dos alunos. Por fim, como trabalhos futuros, acredita-se que seja necessária uma maior amplitude das investigações com um nível maior de abrangência.

## Referências

BERCHT, M. Em direção a agentes pedagógicos com dimensões afetivas. (Doutorado: PGCC), Instituto de Informática, Porto Alegre: UFRGS, 2001, 152p. Tese de Doutorado.

DE AMO, S. Técnicas de mineração de dados. Jornada de Atualização em Informática, v.1, p. 1 – 43, 2004.

HAN, J.; KAMBER, M. Data mining: concepts and techniques. Elsevier, v.3, 2011, p.1 – 5.

LONGARAY, A. N. C.; RIBEIRO, A. C. R.; BEHAR, P. A. Estratégias Pedagógicas para a Educação a Distância: um olhar a partir dos estados de ânimo do aluno. In: Congresso Internacional de Informática Educativa - TISE, 9, 2013, Porto Alegre. Anais. Porto Alegre: Congresso Internacional de Informática Educativa, 2013, p. 606 – 609.

LONGHI, M. T.; BEHAR, P. A.; SIMONATO, G. Os fatores motivacionais e os estados de ânimo em ambientes virtuais de aprendizagem. In: Congresso Iberoamericano de Informática Educativa, 2010, Santiago do Chile, Chile. Anais. Chile: Congresso Iberoamericano de Informática Educativa, 2010. p. 551 - 558.

LONGHI, M. T. Mapeamento de aspectos afetivos em um ambiente virtual de aprendizagem. (Doutorado: PPGIE), Programa de Pós-Graduação em Informática na Educação, Porto Alegre: UFRGS, 2011, 273p. Tese de Doutorado.

LONGHI, M. T.; BEHAR, P. A.; BERCHT, M. Inferência dos estados de ânimo do aluno em um ambiente virtual de aprendizagem baseada em redes bayesianas. Informática na educação: teoria & prática, v. 14, n. 1, p. 93 – 105, jun. 2011.

LONGHI, M. T.; BEHAR, P. A.; BERCHT, M. Mapeamento de aspectos afetivos em um ambiente virtual de aprendizagem. Brazilian Symposium on Computers in Education, v. 23, n. 1., p. 1 – 10, 2012.

NASCIMENTO, R. L. S. do; CRUZ JUNIOR, G. G. da; ARAÚJO FAGUNDES, R. A. de. Mineração de Dados Educacionais: Um Estudo Sobre Indicadores da Educação em Bases de Dados do INEP. RENOTE, v. 16, n. 1, p. 1 - 11, 2018.

OCUMPAUGH, J.; BAKER, R.; GOWDA, S.; HEFFERNAN, N.; HEFFERNAN, C.

Population validity for Educational Data Mining models: A case study in affect detection. *British Journal of Educational Technology*, v. 45, n. 3, p. 487 - 501, 2014.

PASQUALI, L.; AZEVEDO, M. M.; GHESTI, I. Inventário fatorial de personalidade: manual técnico e de avaliação. São Paulo: Casa do Psicólogo, 1997.

PIAGET, J. O nascimento da inteligência na criança. *Mental*. Rio de Janeiro: Zahar, v. 258, 1986, p. 240 - 259.

ROMERO, C.; VENTURA, S. Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, v. 3, n. 1, p. 12 - 27, 2013.

SANTOS, F. D.; BERCHT, M.; WIVES, L. K.; CAZELLA, S. C. Análise de evidências do estado de ânimo desanimado de alunos de um AVEA: uma proposta a partir da aplicação de regras de associação. In: Congresso Brasileiro de Informática na Educação, 4, 2015, Alagoas. *Anais do CBIE*, 2015, p. 1 - 10.

SILVA, J. C. S.; RAMOS, J. L. C; RODRIGUES, R. L.; SOUZA, F. da F. de; GOMES, A. S. Análise do engajamento de estudantes com base na Distância Transacional a partir da Mineração de Dados Educacionais. *RENOTE*, v. 14, n. 1, p. 1 - 11, 2016.

SCHERER, K. R. On the nature and function of emotion: A component process approach. *Approaches to emotion*, v. 2293, n.1, p. 300 - 317, 1984.

SCHERER, K. R. What are emotions? And how can they be measured? *Social science information*, v. 44, n. 4, p. 695 - 729, 2005.