# Mineração de Dados Educacionais e Learning Analytics no contexto educacional brasileiro: um mapeamento sistemático.

**Ano:** 2020

**Relação com o trabalho a ser proposto**: Extração (Mineração) de dados e aplicação de Learning Analytics para levantar dados relevantes sobre o ensino e possíveis maneiras de aperfeiçoa-lo, ainda que nesse trabalho seja de maneira mais “geral”.

**Resumo:** A partir de um mapeamento sistemático buscou-se verificar as contribuições de Learning Analytics e Mineração de Dados Educacionais no contexto educacional brasileiro. Optou-se por fontes de busca em três revistas na área de Informática na Educação e Anais de dois eventos de relevância nacional onde foram verificados 136 artigos entre janeiro de 2008 e março de 2020. Após a aplicação dos critérios de exclusão e de qualidade foram selecionados 71 artigos. Os resultados apresentam a maior ocorrência de projetos com finalidade de analisar desempenho acadêmico e prevenção de evasão escolar, embora nos últimos anos os assuntos estejam apresentando diversificação temática. Grande parte são voltados ao Ensino Superior e na modalidade de Educação a Distância. Há variedade de tecnologias e recursos utilizados no desenvolvimento de soluções em LA ressaltando o uso de Linguagem R, MySQL e a ferramenta Weka.

**Palavras-chave:** mapeamento sistemático, mineração de dados educacionais, learning analytics.

**Problema de pesquisa/proposta**: Identificar quais ferramentas e recursos baseados em Learning Analytics e Educational Data Mining estão sendo aplicados no contexto educacional brasileiro, através de um mapeamento sistemático de literatura, a fim de oferecer subsídios a novas pesquisas, detectar lacunas e potencialidades de aprofundamento nessas áreas.

**Metodologia:** O artigo teve uma abordagem ampla e quantitativa e optou por ter como fontes de busca 3 revistas na área de Informática na Educação e Anais de 2 eventos de relevância nacional, onde foram coletados 136 artigos realizados no período entre janeiro de 2008 e março de 2020 onde após aplicados critérios de qualidade e exclusão, restaram 71 artigos, e sua metodologia dividiu-se em 5 etapas fundamentais:

O planejamento da pesquisa, onde foi criado um protocolo de pesquisa que levou em conta as questões norteadoras para condução das coletas, critérios de exclusão e qualidade, entre outros, de forma que fosse possível oferecer reprodutibilidade a esta pesquisa;

Seleção de estudos primários, que foi feita nos meios de busca selecionados seguindo o protocolo estabelecido, de forma criteriosa e sistematizada;

Avaliação de qualidade, que eliminou os artigos que não atendiam a maioria dos critérios de qualidade, ainda que passassem nos critérios de exclusão, fazendo assim com que os artigos que restassem tivessem uma profundidade maior;

Extração de dados, que foi feito através da leitura dos artigos selecionados e extraindo deles informações relevantes a pesquisa, conforme os atributos pré-definidos sistematizados em uma tabela;

Análise dos resultados, que converteu os dados presentes na tabela em gráficos e nuvens de termos, para melhor análise das questões norteadoras definidas da pesquisa.

**Fichamento:** O artigo buscou responder à questão “Quais ferramentas e metodologias de mineração de dados educacionais e learning analytics estão contribuindo com o processo de ensino-aprendizagem no contexto educacional brasileiro? ”. Em seu processo de busca, optou por utilizar uma string de busca que considerasse um termo OU outro, podendo assim filtrar artigos que falassem apenas de um dos termos, quanto de ambos. Como dito antes, esse processo de busca foi feito em revistas e anais de eventos de informática na educação, e essas fontes foram escolhidos devido a sua qualificação perante a comunidade acadêmica e o grau de confiabilidade das publicações. Da busca foram encontrados 136 trabalhos, os quais foram submetidos aos critérios de exclusão, que buscavam remover os trabalhos que não se enquadrassem no escopo do mapeamento, onde dessa etapa, restaram 73 trabalhos, onde esses foram lidos na íntegra para avaliação de qualidade, onde dessa nova etapa restaram 71 trabalhos selecionados foram dados como relevantes para o estudo. Foi utilizada uma tabela que sintetizava alguns dos dados mais relevantes extraídos de cada um desses artigos, a fim de conseguir realizar o mapeamento desejado, tais como Plataforma utilizada, Tecnologia, Recursos, entre outras informações. Na análise dos resultados foi possível visualizar devido ao aumento de trabalhos nessa área que esse tema foi se tornando mais relevante nos últimos anos e que os principais tópicos de interesse, mas não os únicos, eram o desempenho acadêmico e a evasão escolar. Outros tópicos interessantes encontrados, porém, pouco trabalhados no cenário brasileiro foram Sistemas de recomendação, predição e verificação de habilidades, ficando como sugestão para novas pesquisas. É percebido também que o foco maior desses estudos é o Ensino Superior, mesmo que tenham trabalhos para os demais e que possivelmente isso se deve ao fato de estar mais “próximo” da realidade dos pesquisadores dessas áreas. Foi levantado tamvém que essas ferramentas vêm sendo aplicadas principalmente nas modalidades de ensino Presencial e EAD. Em seguida os autores fazem uma exibição dos termos mais utilizados nessas pesquisas, dos tipos de dados, das tecnologias utilizadas para extrair e processar esses dados.

**Resultados/Observações:** Foi possível verificar que a evasão escolar e o desempenho acadêmico são os principais tópicos de interesse nas pesquisas e que geralmente são realizadas no Ensino Superior nas modalidades EAD e presencial, com público alvo basicamente de professores e alunos. Percebe-se a necessidade da ampliação da aplicação de soluções de EDM e LA no âmbito do Ensino Fundamental e Médio, uma vez que a base escolar causa forte impacto no processo de desenvolvimento acadêmico no Ensino Superior, assim como a adoção de novas metodologias mais alinhadas ao perfil de alunos da era digital.

**Áreas abertas**: Sistemas de Verificação, predição e verificação de habilidades, como sugerido pelos próprios autores. É levantado também como trabalhos futuros a possibilidade de revisões sistemáticas de artigos deste trabalho para uma análise mais qualitativa, com foco nas respostas de perguntas mais específicas.