



INTRODUÇÃO ÀS TÉCNICAS DE MINERAÇÃO DE DADOS UTILIZADAS NA ANÁLISE DA CESTA DE COMPRAS

Maximiliano Hugo Chaile
Geisla De Albuquerque Melo Laskoski (Orientadora)

Resumo

Em 2020, segundo estudos da Associação Brasileira de Supermercados, o setor brasileiro por intermédio da operação de todos os seus formatos e canais de distribuição atingiu um faturamento de R\$ 554 bilhões. Não obstante, à medida que o volume de dados transacionais vai aumentando, a análise de dados para obtenção de informação útil torna-se mais difícil resultando em contextos ricos em dados, mas com informações insuficientes. O presente trabalho tem assim por objetivo apresentar técnicas de mineração de dados utilizadas para entender o comportamento dos clientes em termos de consumo. À vista disso, foram selecionados três trabalhos científicos apresentados nos últimos 10 anos. No artigo de Raorane *et al* (2012), a técnica utilizada visou responder quais grupos ou conjuntos de produtos, os consumidores provavelmente comprarão em um determinado momento. Utilizando regras de associação para a análise de cesta de compras, os autores concluem que as ferramentas de mineração de dados podem ser utilizadas de forma eficaz para aprimorar os padrões associados aos comportamentos dinâmicos das transações realizadas na compra de alguns produtos específicos. No segundo artigo Raeder & Chawla (2011), os autores apresentam uma abordagem modelando os dados como uma rede de produtos. Associando ambas técnicas; tradicionais e baseadas em rede, concluem que o método é capaz de gerar entendimentos que são difíceis de obter com os métodos de análise tradicionais. Kaur & Kang (2016) referem-se ao processo de extração de dados úteis também como processo de extração do conhecimento de bases de dados. Esses processos envolvem várias etapas: como seleção, pré-processamento, transformação, mineração de dados e interpretação. Fazem menção ao artigo de Saurkar *et al* (2014), onde são descritas técnicas e algoritmos disponíveis na descoberta de padrões e regras significativas, tais como: *classificação, associação, predição, clustering e análise outliner*. Afirmam que a principal limitação dos algoritmos enumerados é que operam com dados estáticos e não capturam variações com o tempo. Para entender a dinâmica do processo de geração de dados, apresentam um algoritmo que, trabalhando com dados dinâmicos e realizando mineração periódica, examina as mudanças que ocorreram nos padrões descobertos, fornecendo assim um novo modo de ter em conta as alterações que acontecem nos dados. Os trabalhos levantados ressaltam ainda que a tomada de decisão e o entendimento do comportamento do cliente tornaram-se um assunto essencial e desafiador para manter as posições nos mercados competitivos.

Palavras-chave: mineração de dados; análise da cesta de compra; regras de associação; rede de produtos.