



MINERAÇÃO DE TEXTO APLICADO A ANÁLISE DE CARTEIRA DE AÇÕES

OSMAR MACIEL REHBEIN

osmar85@gmail.com

REJANE FROZZA

frozza@unisc.br

O mercado de capitais é um ambiente em que se pode negociar o capital das empresas. Nesse mercado, pode-se negociar tanto as ações como seus derivativos, que são títulos derivados de ações, como, por exemplo, opção de compra e venda e contrato futuro. Os principais títulos negociados nesse mercado são os representativos do capital de empresas, que são as ações ou empréstimos tomados por essas empresas, via mercado, servindo para custear o desenvolvimento econômico. Assim, a empresa não fica presa a sua geração de lucros, mas também aos investimentos de seus novos investidores. Com o grande número de informações circulando na *web* e o surgimento de redes sociais que trocam mensagem, como o Twitter, criam-se oportunidades de capturar essas informações produzidas na *web* e usá-las no desenvolvimento de sistemas com soluções automatizadas que podem auxiliar os usuários ou entidades de forma a obter vantagens. Esses textos, que são produzidos por empresas e pessoas, são gravados em meios eletrônicos e muitas dessas informações são disponibilizadas na internet. Até recentemente, essas informações em texto não eram utilizadas para provocar vantagem sobre concorrentes ou auxiliar nas tomadas de decisão ou, ainda, como indicadoras de sucesso e fracasso. As mudanças começaram a aparecer com surgimento e o crescimento da Mineração de Texto, já que isso torna possível a extração de informação em meio ao imenso e crescente mundo dos textos. A mineração de texto refere-se a técnicas e processos que descobrem conhecimento inovador nos textos. Essas técnicas estão sendo empregadas em diversas áreas, como, por exemplo, para descobrir fatos da genética e na pesquisa de proteínas e, também, no mundo de negócios, com aplicações não tão pretensiosas, como a categorização automática de mensagens de correio eletrônico, em bancos de investimento, e a extração automática de resumos de documentos pesquisados, por alguns mecanismos de busca, na internet. A finalidade deste trabalho é fornecer uma ferramenta que auxilie os investidores do mercado de capitais a escolher o melhor conjunto de empresas para formar uma carteira de investimento rentável. Pretende-se demonstrar na conclusão que as notícias e os comentários que circulam na internet são capazes de prever se a ação de uma empresa vai subir ou descer. Isso será feito coletando-se, automaticamente, informações textuais diretamente de *sites* de notícias e/ou comentários do Twitter sobre as empresas que estão na bolsa de valores. Os textos coletados passarão por um processamento, abstraindo-se a orientação semântica da notícia e classificando-a como positiva, neutra ou negativa. A soma dessas classificações vai representar a tendência de subida ou descida do valor das ações da empresa analisada.

Instituição: UNISC - SANTA CRUZ DO SUL/RS