

MAIA, Luiz Cláudio Gomes. Uso de sintagmas nominais na classificação automática de documentos eletrônicos. 2008. 158f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

*Esta pesquisa verificou se ocorre aprimoramento na classificação de documentos eletrônicos com o uso de técnicas e algoritmos de mineração de texto (análise de texto) utilizando além das palavras, sintagmas nominais como indexadores. Utilizaram-se duas ferramentas nos experimentos propostos desta pesquisa o OGMA e a WEKA. O OGMA foi desenvolvido pelo autor para automatizar a extração dos sintagmas nominais e o cálculo do peso de cada termo na indexação dos documentos para cada um dos seis métodos propostos. A WEKA foi utilizada analisar os resultados encontrados pelo OGMA utilizando aos algoritmos de agrupamento e classificação, simplekmeans e NaiveBayes, respectivamente, obtendo um valor percentual indicando quantos documentos foram classificados corretamente. Os métodos com melhores resultados foram o de termos sem stopwords e o de sintagmas nominais classificados e pontuados como descritores.*