

MODELAGEM LINGUÍSTICA DA PERCEÇÃO SOCIAL DE DESASTRES NATURAIS: MARCADORES EMOCIONAIS EM UM CORPUS DE TELEJORNAIS

Marllon Felipe Muniz de Farias Alves & Margarethe Born Steinberger-Elias

Universidade Federal do ABC (UFABC)

Mestrado em Engenharia da Informação

Área de concentração: Sistemas Inteligentes

Linha de Pesquisa: Inteligência Social

Pesquisa financiada pela Capes com início em: 03/2012

Previsão de defesa da Dissertação: 03/2014

Este trabalho apresenta um estudo piloto para reconhecimento automático de expressões linguísticas que denotem emoção em um conjunto selecionado de notícias sobre desastres naturais. A extração automática de informações textuais é também chamada de mineração textual. Com o avanço das técnicas de PLN (MINSKY, 1968), foram desenvolvidos sistemas capazes de extrair coleções sistemáticas de textos (corpora) e sistemas capazes de buscar informações relevantes para usuários em textos sobre domínios específicos de conhecimento (SARDINHA, 2005). Todavia, textos são estruturas complexas, e a velocidade das buscas tende a comprometer a qualidade com que a informação é identificada.

Recuperar informação com qualidade significa dinamizar o potencial de associações de conteúdos, o que depende de formas de agregação de inteligência (capacidade de aprendizado) ao processo. Ou seja, um sistema que seja capaz de aprender com as buscas já realizadas e que possa prever novas demandas será mais ágil para criar novos percursos de reconhecimento associativo e gerar conhecimento mais relevante em relação ao corpus estudado (DOWMAN, 2002).

Neste trabalho são analisados relatos jornalísticos televisivos para a construção de um corpus geral (corpus de referência) sobre desastres naturais e de um corpus específico, que permita identificar marcadores de emoção, que são palavras que possuem carga emocional sobre o indivíduo em si ou sobre o imaginário coletivo, podendo indicar alegria, tristeza, surpresa, medo, entre outras emoções. Estes vocábulos podem ser classificados como marcadores de valência emocional positiva, negativa ou neutra, porém não de forma binária, pois cada emoção pode ser expressa em diferentes intensidades e formas (implícita ou explicitamente) (MIHALCEA & STRAPPARAVA, 2007). Segue uma descrição das etapas necessárias ao desenvolvimento do estudo: .

1. Definição dos parâmetros para coleta do material
2. Coleta e pré-seleção do material
3. Definição do processo de classificação e filtragem do material
4. Identificação de marcadores de emoção em relatos transcritos do estudo-piloto

5. Avaliação entre o nível de impacto social das notícias do estudo-piloto causado entre os agentes (pessoas ou entidades envolvidas no relato) e a aplicação dos marcadores de emoção (STEINBERGER, 2010).

Identificadas e analisadas, as expressões encontradas são classificadas, um estudo preliminar aponta cinco classes de marcadores de sentimento: *advérbios de intensidade*, *adjetivos de intensidade*, *quantificadores de generalização*, *sentenças expressivas (núcleo “negativo”)* e *verbos de opressão*. O quadro abaixo apresenta um exemplo de expressões e palavras classificadas de acordo com grupo de marcadores emocionais pertinente.

Quadro 1 - Classificação de marcadores emocionais

Advérbios de intensidade	Adjetivos de intensidade	Quantificadores de generalização	Sentenças expressivas (núcleo “negativo”)	Verbos de opressão
Ainda mais	Único	Tudo	A calamidade nossa é essa	Obrigar
Bem	Forte	Toda	Não é exagero	Ter que
Muito	Gigantesca	Permanentemente	Medo é reviver uma das maiores tragédias	Ameaçar
Já	Difícil		A gente espera a morte num lugar desses	Corre pra salvar o pouquinho
			Não deixam esquecer o pesadelo	

Mesmo através de análise preliminar, já é possível notar alta concentração de marcadores emocionais nas sentenças produzidas pelas vítimas (pessoas impactadas diretamente) do evento, a maioria com caráter negativo podendo ser classificadas como expressões de medo e/ou tristeza.

Referências bibliográficas

1. DOWMAN, M. et al. Web-Assisted Annotation, Semantic Indexing and Search. International World Wide Web Conference. Chiba, Japão: [s.n.]. 2005. p. 10.
2. MIHALCEA, R., & STRAPPARAVA, C. Affective Text. 4th International Workshop on Semantic Evaluations. Praga. 2007. p. 5.
3. MINSKY, Y. M. (editor). **Semantic information processing**. Cambridge: The MIT Press, 1968.
4. SARDINHA, T.B. **Introdução à Linguística de Corpus**, São Paulo: Manole, 2005.
5. STEINBERGER, M.B. Estudo das condições de produção de relatos sobre desastres e catástrofes na América Latina. Grupo de Pesquisa do CNPq sobre Geografias da Comunicação, Anais do Colóquio Binacional Brasil-EUA de Ciências da Comunicação, Universidade de Caxias do Sul, RS, 2010.