



Silvano Nogueira Buback

**Utilizando aprendizado de máquina
para construção de uma ferramenta de
apoio a moderação de comentários**

Dissertação de mestrado

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico e Científico da PUC-Rio.

Orientador: Prof. Marco Antonio Casanova

Co-Orientador: Prof. Ruy L. Milidiú

Rio de Janeiro
Setembro de 2011



Silvano Nogueira Buback

**Utilizando aprendizado de máquina
para construção de uma ferramenta de
apoio a moderação de comentários**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico e Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marco Antonio Casanova

Orientador

Departamento de Informática - PUC-Rio

Prof. Ruy L. Milidiú

Co-Orientador

Departamento de Informática - PUC-Rio

Prof. Antonio L. Furtado

Departamento de Informática - PUC-Rio

Prof. Karin K. Breitman

Departamento de Informática - PUC-Rio

Prof. José Eugenio Leal

Coordenador Setorial do Centro - Técnico Científico – PUC-Rio

Rio de Janeiro, 01 de setembro de 2011.

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Silvano Nogueira Buback

Graduou-se em Ciência da Computação com ênfase em Sistemas de Informação na Universidade Federal do Espírito Santo onde ajudou a construir um ambiente de desenvolvimento de software baseado em ontologias. Trabalhou na Americanas.com e na Globo.com onde é responsável pela engenharia dos sistemas de autenticação e cadastro de usuários e de comentários. Desenvolveu como dissertação de mestrado uma ferramenta para auxiliar a moderação de comentários de notícias.

Ficha Catalográfica

Silvano Nogueira Buback

Utilizando aprendizado de máquina para construção de uma ferramenta de apoio a moderação de comentários / Silvano Nogueira Buback; orientador: Marco Antonio Casanova; co-orientador: Ruy Luiz Milidiú. - Rio de Janeiro: PUC-Rio, Departamento de Informática, 2011.

v., 65 f.: il. ; 29,7 cm

Dissertação de Mestrado - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática

Referencias bibliográficas incluídas.

Moderação de comentários; Aprendizado de Máquina; SVM; Classificadores Bayesianos; Boosting; Processamento de Linguagem Natural; Classificação de texto

CDD: 004

Agradecimentos

Agradeço primeiramente à Deus, que me conduziu e me protegeu durante toda essa caminhada.

À minha esposa Karol, pela imensa compreensão e paciência durante esses anos de mestrado, inclusive o apoio na revisão da dissertação.

À Globo.com, por ter acreditado em mim e oferecido a oportunidade de cursar este mestrado.

Ao meus colegas da Globo.com, pelas idéias e pela força que me deram para a conclusão desta dissertação.

Ao meu orientador Marco Antonio Casanova pela paciência, compreensão e pelos ensinamentos.

Ao meu co-orientador Ruy Milidiú pelas aulas e pelo conhecimento compartilhados.

Aos meus pais pela compreensão de eu não estar tão presente, e mesmo assim eles estão sempre torcendo por mim.

Resumo

Buback, Silvano Nogueira; Casanova, Marco Antonio; Milidiú, Ruy Luiz. **Utilizando aprendizado de máquina para construção de uma ferramenta de apoio a moderação de comentários.** Rio de Janeiro, 2011. 65p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Uma das mudanças trazidas pela Web 2.0 é a maior participação dos usuários na produção do conteúdo, através de opiniões em redes sociais ou comentários nos próprios sites de produtos e serviços. Estes comentários são muito valiosos para seus sites pois fornecem feedback e incentivam a participação e divulgação do conteúdo. Porém excessos podem ocorrer através de comentários com palavras indesejadas ou spam. Enquanto para alguns sites a própria moderação da comunidade é suficiente, para outros as mensagens indesejadas podem comprometer o serviço. Para auxiliar na moderação dos comentários foi construída uma ferramenta que utiliza técnicas de aprendizado de máquina para auxiliar o moderador. Para testar os resultados, dois corpora de comentários produzidos na Globo.com foram utilizados, o primeiro com 657.405 comentários postados diretamente no site, e outro com 451.209 mensagens capturadas do Twitter. Nossos experimentos mostraram que o melhor resultado é obtido quando se separa o aprendizado dos comentários de acordo com o tema sobre o qual está sendo comentado.

Palavras-chave

Moderação de comentários; Aprendizado de Máquina; SVM; Classificadores Bayesianos; Boosting; Processamento de Linguagem Natural; Classificação de texto.

Abstract

Buback, Silvano Nogueira; Casanova, Marco Antonio (Advisor); Milidiú, Ruy Luiz (Co-Advisor). **Using Machine Learning to build a tool that helps comments moderation.** Rio de Janeiro, 2011. 65p. MSc. Dissertation - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

One of the main changes brought by Web 2.0 is the increase of user participation in content generation mainly in social networks and comments in news and service sites. These comments are valuable to the sites because they bring feedback and motivate other people to participate and to spread the content. On the other hand these comments also bring some kind of abuse as bad words and spam. While for some sites their own community moderation is enough, for others this inappropriate content may compromise its content. In order to help these sites, a tool that uses machine learning techniques was built to mediate comments. As a test to compare results, two datasets captured from Globo.com were used: the first one with 657.405 comments posted through its site and the second with 451.209 messages captured from Twitter. Our experiments show that best result is achieved when comment learning is done according to the subject that is being commented.

Keywords

Comments Moderation; Machine Learning; SVM; Bayesian Classifiers; Boosting; Natural Language Processing; Text Classification.

Sumário

1	Introdução	11
1.1.	Organização da tese	12
2	Trabalhos relacionados	13
2.1.	Filtro anti-spam	13
2.2.	Sentiment Analysis	14
3	Aprendizado supervisionado	17
3.1.	Técnicas para aprendizado supervisionado	17
3.1.1.	Naïve Bayes	17
3.1.2.	Boostexter	18
3.1.3.	SVM	19
3.1.4.	Maximum Entropy	20
3.2.	Métodos de extração de atributos	21
3.2.1.	Bag of Words	21
3.2.2.	Corretor	21
3.2.3.	POS Tagging	23
3.2.4.	N-grams	24
4	Arquitetura	25
4.1.	Visão geral da arquitetura	26
4.2.	Camada de Aprendizagem	27
5	Experimentos	31
5.1.	Critérios para avaliação do sistema	31
5.1.1.	Precisão	31
5.1.2.	Recall	32
5.1.3.	Acurácia	32
5.1.4.	F-measure	32
5.1.5.	Coeficiente de correlação	33

5.2. Corpora	33
5.2.1. Corpus globo-comments	34
5.2.1. Corpus globo-twitter	34
5.2.2. Diferenças entre os corpus	34
5.3. Metodologia do teste	35
5.4. Resultados obtidos	36
5.4.1. Primeira etapa	37
5.4.2. Segunda etapa	38
5.4.3. Terceira etapa	40
5.5. Discussão	43
6 Conclusões e trabalhos futuros	46
7 Bibliografia	48
Apêndice A – Resultados da etapa I dos experimentos	51
Apêndice B – Distribuição dos comentários por grupo	62
Apêndice C – Parâmetros de uso da API de linha de comando	64

Lista de imagens

Figura 1 – Relacionamento entre as entidades do sistema de comentários	25
Figura 2 - Arquitetura do sistema	26
Figura 3 - Diagrama de classes da camada aprendizagem	28
Figura 4 - Validação cruzada para 4 partições	36
Figura 5 - Gráfico da probabilidade de rejeição por percentagem de erro	41
Figura 6 - Gráfico da quantidade de comentários por probabilidade de rejeição	42
Figura 7 - Gráfico da contribuição do erro geral por probabilidade	42

Lista de tabelas

Tabela 1 - Resultados obtidos em (ANDROUTSOPOULOS, KOUTSIAS, <i>et al.</i> , 2000) e (ANDROUTSOPOULOS, KOUTSIAS, <i>et al.</i> , 2000).	13
Tabela 2 - Configuração das máquinas utilizadas nos testes	37
Tabela 3 - Resultados obtidos utilizando classificadores de contexto	40
Tabela 4 - Desvio padrão variando-se o parâmetro C (SVM)	43
Tabela 5 - Desvio padrão variando-se o método de seleção de atributos	44