



## ANÁLISE DE SENTIMENTO DE *TWEETS*: MANIFESTAÇÕES ACERCA DO UFC 202

Luiz Antonio Lazzarin Trentinalha<sup>1</sup>, Gabriela Cristófoli Pereira<sup>2</sup>, Leonardo Pestillo de Oliveira<sup>3</sup>

<sup>1</sup> Acadêmico do Curso de Psicologia, Centro Universitário de Maringá – UNICESUMAR, Maringá-PR.  
Bolsista PIBIC-Fundação Araucária-UniCesumar. tonitrentinalha@live.com

<sup>2</sup> Acadêmica do Curso de Psicologia, UNICESUMAR

<sup>3</sup> Orientador, Doutor, Docente do Programa de Mestrado em Promoção da Saúde, UNICESUMAR

### RESUMO

O esporte apresenta uma estrutura convencionalizada, em que as normas restritas da vida diária são colocadas de lado em favor de normas morais típicas. Com isso, o objetivo do estudo é avaliar como os esportes de luta, que possibilitam o atleta chegar ao *Mixed Martial Arts* (MMA), influenciam no desenvolvimento de comportamentos morais de atletas e também espectadores do evento, e com isso descrever a política de identidade presente na modalidade. Trata-se de uma pesquisa descritiva sobre os aspectos da moralidade envolvidos no desenvolvimento de atletas que praticam artes marciais e seus espectadores. A nossa busca por discursos sobre o evento UFC (edição 202) foi conduzida no website twitter.com através da API de busca do Twitter. A coleta de dados não foi realizada diretamente com os sujeitos, mas sim por meio das mensagens divulgadas na rede social *Twitter*, ou seja, para a realização da análise de sentimento dos *tweets* foi utilizada linguagem R, um ambiente de desenvolvimento integrado para cálculos estatísticos e gráficos. Os resultados demonstraram que as mensagens enviadas pelos usuários do Twitter no dia do evento UFC 202 apresentaram uma classificação mais positiva do que negativa, demonstrando que a manifestação popular, mesmo sendo acerca de um dos eventos mais polêmicos do UFC, podem ser positivas direcionadas a eventos de artes marciais mixtas. Conclui-se que os usuários do Twitter, mesmo com a possibilidade de manifestação livre, podem emitir mensagens positivas sobre um esporte que socialmente ainda apresenta algumas dificuldades de inserção, principalmente quando se refere a crianças e adolescentes.

**PALAVRAS-CHAVE:** sentimento; twitter; psicologia; esporte.

## 1 INTRODUÇÃO

O ambiente, ou cultura em que o sujeito é educado, exerce influência na sua disposição para assumir atitudes específicas. Na tentativa de entender essas influências é necessário que se descreva quais as estratégias de socialização às quais este sujeito foi exposto. Todo contexto social é composto por regras, e o sujeito além de conhecer, aceitar e conseguir emitir juízos sobre tais regras, precisamos saber qual a real disposição deste sujeito para agir de acordo com elas.

Alguns autores utilizam o termo comportamento prosocial para discutir a questão da capacidade de agir de acordo com os preceitos morais. Este comportamento se caracteriza por ser aquele que representa uma influência positiva em outras pessoas, não significando benefício para si próprio.

A estes processos de socialização dá-se o nome de políticas de identidade, que visam dar conta da formação de identidades no seio de uma sociedade. O principal propósito disso seria garantir sua reprodução ou seja, os chamados "processos através dos quais as identificações se realizam dentro de horizontes que implicam um



determinado mundo social" (ALMEIDA 2005, p. 130). Assim como colocam Berger e Luckmann (2004) o indivíduo nasce e torna-se membro da sociedade. Viver em grupo não é uma tarefa simples, no entanto, conviver com grupos que tem diferentes regras e aspectos morais é mais problemático. Além de diferentes grupos há que se considerar o momento em que isso ocorre, no caso do esporte, em qual fase do desenvolvimento atlético ele está.

No âmbito esportivo, muito se discute a respeito de como a prática esportiva leva à atitudes de crianças e adolescentes que são condizentes com a moral social, pois a partir disso desenvolvem por exemplo a capacidade de respeitar o próximo, respeitar as regras, e lidar melhor frente à figura de autoridade. Mas os esportes de luta acabam sendo alvo de críticas a respeito da forma como são conduzidos, pois ao serem considerados esportes violentos, em que a agressividade é sinônimo de sucesso, pais e responsáveis acreditariam que este mesmo comportamento fosse levado para fora do ambiente esportivo.

Nos dias atuais, essa manifestação para fora do ambiente esportivo é ao mesmo tempo real e virtual, visto que, com o avanço da internet as pessoas passam a se manifestar de diversas formas acerca de seus sentimentos e comportamentos. O uso dos sites chamados de redes sociais é exemplo claro de como essa manifestação se torna livre e ao mesmo tempo causa espanto, dependendo da forma como é realizada.

Uma dessas redes sociais, o Twitter, oferece uma série de benefícios importantes como fonte de dados para pesquisa e investigação das mais variadas formas de expressão da população. Primeiramente, a fonte de dados é vasta e plenamente acessível. Por exemplo, em 2012 mais de 340 milhões de *tweets* foram publicados diariamente, e este conteúdo está disponível gratuitamente. Em seguida, os dados podem ser recolhidos e analisados automaticamente e em tempo real, o uso do Twitter para expressão dos sentimentos é validada pela informalidade com o que a ferramenta tem sido utilizada, em todo e qualquer contexto a rede social reflete tendências que não são discutidas em campos formais do conhecimento.

A quantidade de dados presentes na web cresce a cada dia, principalmente após a popularização das redes sociais, sendo o Twitter e o Facebook as mais conhecidas e utilizadas atualmente. Com o avanço tecnológico a quantidade de dados já existentes na web cresce historicamente, sendo que alguns cálculos realizados indicam que até 2020 essa quantidade irá quadruplicar (OLHAR DIGITAL, 2012).

Com o crescimento da quantidade de dados, surge também a possibilidade de utilização destas informações públicas como fonte para extração de características e opiniões sobre produtos, empresas e pessoas, principalmente pessoas que estão em evidência midiática. Isso porque tanto as grandes empresas, quanto as chamadas personalidades da mídia, aproveitam esta situação para monitorar sua popularidade e relevância frente à opinião dos espectadores. Mas o outro lado também deve ser observado, ou seja, como essas marcas e estas personalidades da mídia se manifestam por meio das redes sociais, e também qual a real influência dessas mensagens ao público.

Por esse motivo, surge o que se chama atualmente de Análise de Sentimento, que tem como objetivo identificar características de um texto escrito de acordo com um determinado assunto. A partir desta análise classifica-se em subgrupos pré-definidos, tais como sentimentos e emoções (que podem ser descritos como bem conhecido, triste, feliz), popularidade (com determinações de positivo e negativo) ou em outro contexto que tenha sido previamente definido (AGARWAL et al., 2011).



## 2 DESENVOLVIMENTO

A nossa busca por discursos sobre o evento UFC (edição 202) foi conduzida no website *twitter.com* através da API de busca do *Twitter*. A busca foi realizada no dia 22 de agosto de 2016, compreendendo o período prévio e posterior ao evento. O acesso ao conteúdo produzido pelo *Twitter* através da API permite um acesso de busca direto ao banco de dados da rede social.

A confecção do nosso banco de dados foi estruturado para coletar *tweets* em inglês após o término do evento, utilizando o termo #UFC202 como epicentro da extração de *tweets*. Informações gerais de funcionamento da rede social também foram extraídos, como identificar dos usuário, horário de postagem, frequência de *retweets* e conteúdo de cada *tweet*. Para a análise de sentimento e de redes, nós excluímos os *retweets* para identificar apenas o conteúdo original publicado pelos usuários do *Twitter*. A amostra final foi composta por 38,238 *tweets* e *retweets*, dos quais 26,159 foram *tweet* originais dos usuários. Técnicas de mineração de texto foram aplicadas para análise de sentimento e dos padrões semânticos dos discursos dos usuários do *Twitter*. A análise do conteúdo dos *tweets* foi conduzida inicialmente com a criação de um corpus com base nas postagens de cada usuário.

A rede foi construída tomando cada palavra da matriz de correlação como um nodo. Esses nodos foram conectados pelas hastes representando a intensidade da sua correlação (aqui expressando a co-ocorrência em parágrafos análises dos *tweets*). Esses passos representam uma configuração de rede bipartite, na qual as palavras que fossem expressas nos mesmos parágrafos teriam uma relação maior (direta), com relações indiretas estabelecidas pela presença de palavras em comum. A estrutura da rede foi definida de forma que quanto mais próximos os nodos, maior é a associação entre os palavras. Como indicadores descritivos da rede, utilizamos o conceito de *betweness* (aqui traduzido como conectores) para identificar as palavras que eram mais importantes para fazer as conexões entre os clusters da rede de conexões semânticas. Todas as análises foram conduzidas com o programa Linguagem R (*R-Project*, 2014).

## 3 RESULTADOS E CONSIDERAÇÕES FINAIS

Tabela 1. Exemplos de *tweets* positivos, negativos e neutros

	Palavras	Tweet	N (%)
Positivo	won, win, like, great	As usual @UFC #ufc202 was a great show! Another great card top to bottom.	1172 (28.8)
Negativo	fuck, lost, notorious, kill	you [curse] bum you got you're ass whooped you [curse] retarded crack head [curse] Stockton and the Diaz brothers #ufc202 #bum	477 (11.7)
Neutro	-	What do you think about the outcome of the #UFC202 yesterday?	2420 (59.5)

A Tabela 1 ilustra alguns exemplos de palavras que descrevem comportamento positivos ou pró-sociais. As palavras mais utilizadas pelos usuários e que denotam sentimento considerado positivo foram: venceu, vencer, gostar e ótimo. Em contrapartida, os termos que aparecem mais comumente para denotar posturas negativa e anti-sociais



foram: perdeu, matar, notório e “fuck”<sup>1</sup>. E os tweets classificados como neutros vinham normalmente acompanhados de uma pergunta, tal qual: “o que você achou dos resultados do UFC?”.

Uma expansão dos resultados acerca das palavras classificadas como positivas e negativas pode ser observada na Figura 1, que representa graficamente a frequência de uso das mesmas.

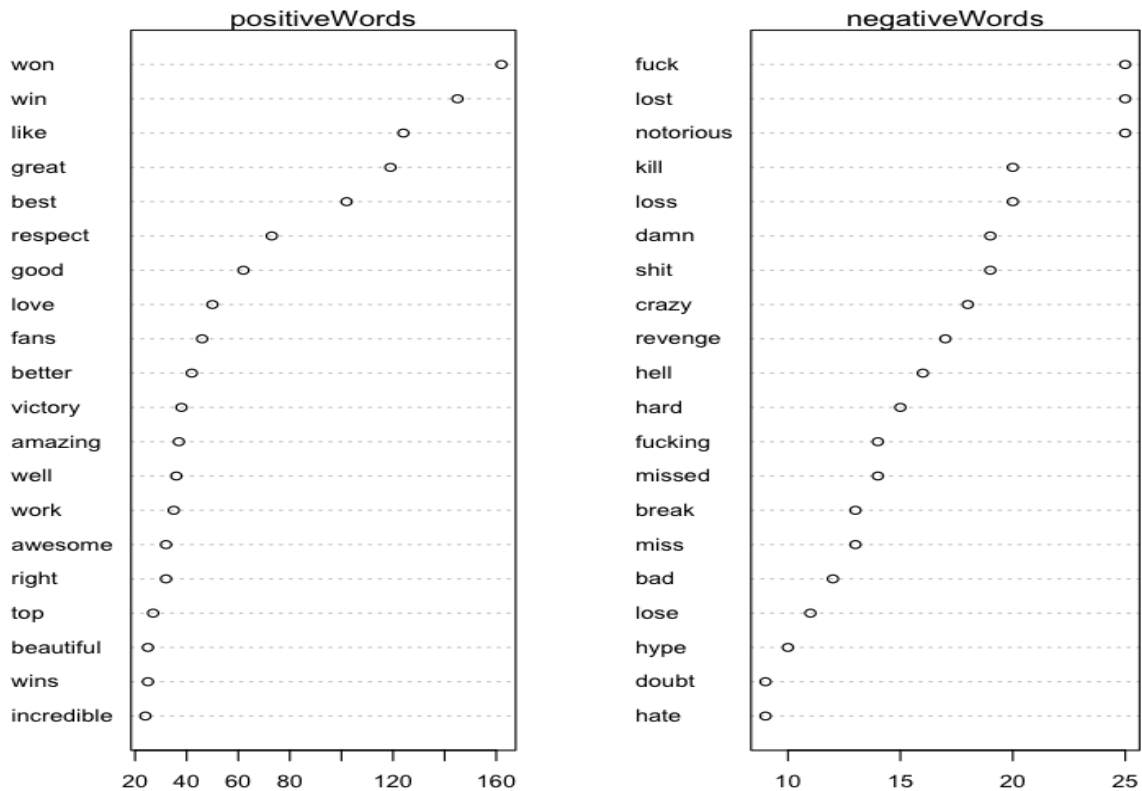


Figura 1. Palavras positivas e negativas utilizadas por usuários do *Twitter*

Pelos resultados obtidos, a Figura 1 apresenta o conjunto de palavras de atitude positiva e negativa e a quantidade de vezes em que cada uma aparece. Levando em consideração as palavras positivas, as mais citadas foram: vencer, gostar e ótimo. As menos mencionadas, apesar de aparecerem significativamente: incrível, lindo e vitórias. Dentre as palavras de cunho negativo as mais encontradas foram: matar, notório e perdeu. E as menos citadas: odiar, dúvida e extravagante.

<sup>1</sup> Palavra utilizada em inglês para expressar uma emoção negativa sobre algo.

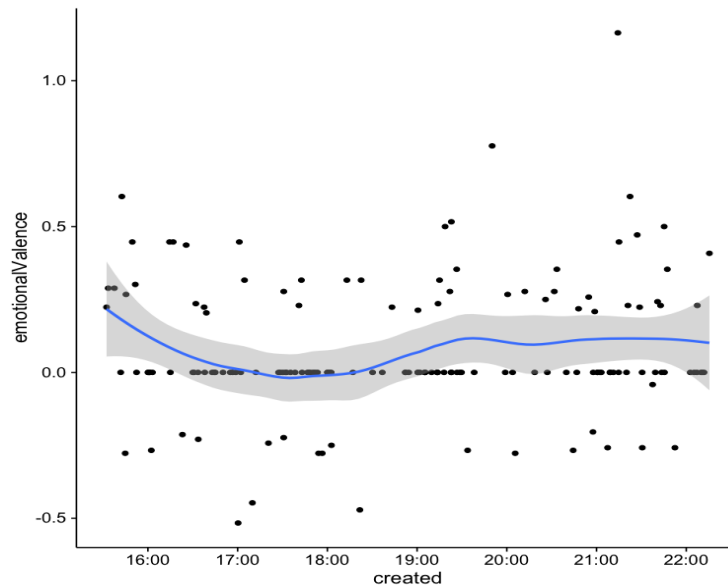


Figura 2. Flutuação do conteúdo emocional de tweets ao longo do dia do evento.

A Figura 2 descreve a flutuação do conteúdo emocional contido em cada tweet ao longo do dia do Evento. O tônico emocional de cada tweet é calculado pela presença de palavras positivas e negativas dentro de um mesmo tweet. Variando entre positivos e negativos se pode notar que em geral há um tônico positivo. Dentro do que se pode observar, essa flutuação se torna interessante pelo fato de que o conteúdo das mensagens postadas momentos antes do início do evento torna-se negativo, voltando a ser positivo no decorrer do evento e após o mesmo.

Os resultados demonstraram que as mensagens enviadas pelos usuários do Twitter no dia do evento UFC 202 apresentaram uma classificação mais positiva do que negativa, demonstrando que a manifestação popular, mesmo sendo acerca de um dos eventos mais polêmicos do UFC, podem ser positivas direcionadas a eventos de artes marciais mistas. Conclui-se que os usuários do Twitter, mesmo com a possibilidade de manifestação livre, podem emitir mensagens positivas sobre um esporte que socialmente ainda apresenta algumas dificuldades de inserção, principalmente quando se refere a crianças e adolescentes.

## REFERÊNCIAS

AGARWAL, A.; XIE, B.; VOVSHA, I.; RAMBOW, O; PASSONNEAU, R. **Sentimental Analysis of Twitter Data**. Department of Computer Science, Columbia University: New York, NY, USA, 2011. Disponível em [http://www1.ccls.columbia.edu/~beck/pubs/lsm2011\\_full.pdf](http://www1.ccls.columbia.edu/~beck/pubs/lsm2011_full.pdf). Acesso em 01 de abril de 2013.

BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2008.

BERGER, P. L.; LUCKMANN, T. **A construção social da realidade**: tratado de sociologia do conhecimento. 24. ed. Petrópolis: Vozes, 2004.

BIRD, S. NLTK: **The natural language toolkit**. In Paper presented at the proceedings of the COLING/ACL on Interactive presentation sessions. 2006.



- BLEDSOE, G. H. et al. Incidence Of Injury In Professional Mixed Martial Arts Competitions. **Journal of Sports Science and Medicine**, 2006, CSSI, 136-142.
- BREDEMEIER, B.; WEISS, M. Moral development in sport. **Exercises and Sport Science Reviews**, 1990, 18, 331-378.
- CÔTÉ, J. The influence of the family in the development of talent in sport. **The Sport Psychologist**, 1999, 13, 395-417.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- KOLLER, S. H.; BERNARDES, N. M. G. Desenvolvimento moral pró-social: Semelhanças e diferenças entre os modelos teóricos de Eisenberg e Kohlberg. **Estudos de psicologia**, 1997, 2(2), 223-262.
- LIU, B. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5, 1–167, 2012.
- MOHAMMAD, S. M., & TURNEY, P. D. **NRC Emotion Lexicon**, 2013.
- OLHAR DIGITAL. **966 exabytes**: esta é a quantidade de dados que tráfegarão na internet em 2015. Web. 05 de abril de 2012. Disponível em [http://olhardigital.uol.com.br/produtos/digital\\_news/noticias/966\\_exabytes\\_esta\\_e\\_a\\_quantidade\\_de\\_dados\\_que\\_trafegarao\\_na\\_internet\\_em\\_2015](http://olhardigital.uol.com.br/produtos/digital_news/noticias/966_exabytes_esta_e_a_quantidade_de_dados_que_trafegarao_na_internet_em_2015). Acesso em 03 de abril de 2016.
- PANG, B. & LEE, L. Opinion mining and sentiment analysis. Vol. 2, 1–135 (2008).
- PLUTCHIK, R. **EMOTION: A psychoevolutionary synthesis**. New York: Harper & Row, 1980.
- R-PROJECT. **The R Project for Statistical Computing**. Disponível em <http://www.r-project.org/>. Acesso em 01 de abril de 2016.
- VIEIRA, J. L. L. Avaliação do desenvolvimento moral de adolescentes em dilemas morais da vida diária e da prática esportiva. **Revista da Educação Física, Maringá**, 1993, v. 4, n. 01, p. 34-39.
- WEINBERG, R. S. e GOULD, D. **Fundamentos da Psicologia do Esporte e do Exercício**. 2. ed. Porto Alegre: Artmed Editora, 2001.