



UNIVERSIDADE FEDERAL DE
PERNAMBUCO

GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA
2016.1

Avaliação de técnicas de aprendizagem de
máquina para predição de engajamento com
user-generated content

PROPOSTA DE TRABALHO DE GRADUAÇÃO

Aluno: David Wilson de Farias Santos (dwfs@cin.ufpe.br)

Orientador: Tsang Ing Ren (tir@cin.ufpe.br)

Recife, 18 de Abril de 2016

Sumário

1. Resumo	3
2. Contextualização.....	3
3. Objetivos.....	4
4. Cronograma.....	5
5. Possíveis Avaliadores	5
6. Referências.....	5
7. Assinaturas.....	6

1. Resumo

Antes do advento da Web 2.0, os consumidores tinham recursos limitados para ajudar na tomada de decisão voltada para o consumo, fazendo com que eles se baseassem fortemente na informação enviesada fornecida pela publicidade tradicional. No entanto, com a chegada da Web 2.0, os usuários começaram a compartilhar suas experiências, produzindo um imenso volume de informações geradas por eles próprios, também conhecidas como UGC (*User-Generated Content* – Conteúdo Gerado pelo Usuário). O uso de imagens geradas por usuários na forma de anúncios pode aumentar a conversão de vendas, pois este tipo de conteúdo tem o poder de influenciar a decisão de potenciais consumidores, tornando-os propensos à compra. Diante desse panorama, este trabalho propõe uma avaliação de técnicas de aprendizagem de máquina que irão extrair padrões em imagens postadas nas redes sociais, com o objetivo de prever o engajamento de consumidores com tais imagens quando publicadas como anúncios.

2. Contextualização

Antes do advento da Internet, os consumidores tinham recursos limitados para ajudar a sua tomada de decisão. Alguns confiavam no que pessoas próximas diziam, e já outros se baseavam unicamente na informação que vem diretamente de propagandas. A informação oferecida por conhecidos limita os consumidores a um pequeno número de opiniões, enquanto as publicidades lhes davam apenas uma informação parcial. O que eles precisavam era de um volume de dados maior de comentários para ajudar na sua tomada de decisão.

User-Generated Content Consumers 2008-2013 (% of Internet users)						
	2008	2009	2010	2011	2012	2013
User-Generated Video	36.0 %	39.8%	42.5%	44.8%	47.2%	49.2%
Social Networking	41.2%	44.2%	46.9%	49.1%	50.5%	51.8%
Blogs	54.0%	58.0%	61.0%	64.0%	67.0%	69.0%
Wikis	33.9%	36.6%	39.0%	41.0%	42.6%	43.9%
User-Generated Content Consumers	60.0%	62.0%	64.0%	66.0%	68.0%	70.0%

Figura 1 - User-Generated Content Consumers [2]

Com a criação da Internet, veio uma nova era: a Web 2.0. De repente, os usuários da Internet tinham a capacidade de compartilhar informações uns com os outros, produzindo uma grande quantidade de conteúdo, também conhecido como UGC. A

confiança do consumidor em conteúdo gerado pelo usuário e pela publicidade foi verificada em [1], onde 66,3% dos consumidores avaliados se baseiam fortemente em conteúdo gerado por usuário ao tentar fazer decisões de compra. A Figura 1 demonstra que a tendência da UGC é avançar para blogs e redes sociais. É também interessante notar que 70% dos usuários de internet são consumidores reais de UGC, e este valor será crescente nos próximos anos.

Diante desse contexto, é de suma importância às empresas terem o conhecimento sobre o que os usuários vêm postando sobre os seus produtos nas redes sociais. Além disso, há o interesse em utilizar conteúdos UGC em publicidade por eles serem comprovadamente responsáveis por aumentar a conversão de vendas, pois elas funcionam como aprovações visuais para potenciais compradores que, quando veem uma foto que gostam ou que se identificam, se tornam mais propensos à compra [3], ou seja, a extração de informação pode ajudar a prever o engajamento e potencializar a venda de um determinado produto.

Várias informações podem ser extraídas de imagens publicadas em redes sociais, como número de curtidas, qualidade da imagem, comentários, entre outros. Esses metadados podem ser utilizados como base para um sistema que prevê os resultados de engajamento que cada imagem possui ao ser usada como anúncio. Já existem trabalhos que utilizam aprendizagem de máquina para avaliar o desempenho de seus anúncios através da CTR (taxa de cliques) [4], métrica que indica a frequência com que as pessoas que veem o seu anúncio acabam por clicar nele.

Considerando esse panorama, este trabalho propõe uma avaliação de técnicas de aprendizagem de máquina que examinará grandes quantidades de dados, metadados extraídos de imagens de rede sociais, em busca de padrões para prever quais as imagens geradas pelos usuários que apresentam maior potencial de engajamento em anúncios.

3. Objetivos

Este trabalho de graduação tem como principal objetivo a avaliação de técnicas de aprendizagem de máquina com a finalidade de prever o engajamento de uma determinada imagem em anúncios, utilizando seus metadados.

Serão selecionados os dados mais adequados obtidos a partir de imagens coletadas da rede social Instagram¹, tais como cor de fundo, presença de pessoas, qualidade da imagem, entre outros, para criar um vetor de características ótimo para nossa aplicação. A partir dessa definição, iremos preparar os dados coletados dessas imagens que servirão como base de treinamento para algoritmos de aprendizagem de máquina, onde, através de vários experimentos, será selecionado como modelo candidato, para realizar previsões de engajamento de UGC, aquele que apresentar melhor desempenho.

¹ <https://www.instagram.com/>

4. Cronograma

Nesta seção, é apresentado o cronograma previsto, na Tabela 1, para o desenvolvimento desse Trabalho de Graduação.

Atividades	Março	Abril	Maiο	Junho	Julho
Formulação da proposta e revisão literária.					
Estudo de Ferramentas					
Preparação dos dados de treinamento					
Experimentos e obtenção do modelo candidato					
Análise dos resultados.					
Preparação da apresentação					

Tabela 1: Cronograma do Trabalho de Graduação

5. Possíveis Avaliadores

Os possíveis avaliadores para o resultado a ser obtido ao final de todas as etapas descritas neste documento são:

- Tsang Ing Ren
- George Darmiton da Cunha Cavalcanti

6. Referências

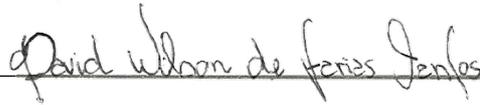
[1] Katherine A. MacKinnon, **User Generated Content vs. Advertising: Do Consumers Trust the Word of Others Over Advertisers?**, The Elon Journal of Undergraduate Research in Communications, Vol. 3, No. 1, p. 14- 22, 2012.

[2] **User-Generated Content Draws Fans.** February, 2009. <http://www.emarketer.com/Article.aspx?R=1006895>.

[3] **Social Influence: Marketing's New Frontier**, <http://corp.crowdtap.com/socialinfluence>.

[4] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, Joaquin Quiñero Candela, **Practical Lessons from Predicting Clicks on Ads at Facebook**.

7. Assinaturas



David Wilson de Farias Santos
Orientando



Tsang Ing Ren
Orientador