Using Information Retrieval Models to Find a Target Audience in Social Media

Jake Lasley  
 Computer Science  
 University of Texas at El Paso  
 El Paso, Texas  
 amotta3@miners.utep.edu

Valentin Becerra  
 Computer Science  
 University of Texas at El Paso  
 El Paso, Texas  
valentin.oneiro@gmail.com

Alan Motta  
 Computer Science  
 University of Texas at El Paso  
 El Paso, Texas  
 alanmotta\_22@gmail.com

ABSTRACT

Social media and social media applications are hotspots for collecting information about individuals. These websites and their component applications are an outlet for people to share their interests, details about themselves, and react to other individual’s posts. The nature of these websites brings us closer to the people and business that we know and love. We add our friends online, we follow the artists we like to listen to, and we check in to the places we love to go to. What about the places that we don’t know about, but would really enjoy being at? How about, the artist we don’t yet listen to, but would have their songs on repeat? The people we haven’t met, but would eventually be part of our close circle of friends? This proposal includes methods of information retrieval used on Instagram to help connect businesses, artists, and people closer to their areas of interest.

STATEMENT OF PROBLEM

Currently websites help businesses find their target audience, and artists discover their fanbase. However, these are services that you must pay for that make use of people’s private information. There are applications that help you discover new people, but these applications are specialized solely for that one purpose and rely on private information too. There are currently not many methods that help online profiles make these connections using publicly available data, much less any that give the people the power to do this themselves. There is a need for such methods and that also protect user’s online privacy.

OBJECTIVES

Objective 1: Use web crawlers to extract user public information.

Objective 2: Processing website text.

Objective 3: Ranking results by relevance.

Objective 4: Querying user public information

Objective 5: Classifying results

PLAN OF ACTION

1. Use web crawlers to extract user public information

We will use web crawlers to crawl through online profiles while abiding politeness policies. Also, we will only access resources that the robots.txt dictates we are allowed to access. Next, we will build a vertical search that provides higher accuracy of information.

1. Processing website text

After gathering information, we will have to decide whether it should be modified or restructured to simplify searching. Next, we will use methods to parse and recognize the content and structure of the text acquired. Then, we will tokenize, stem, and identify stop words in the text.

1. Ranking results by relevance

To process the text, we need to rank results so that they are relevant. For this we will make use of the algorithms discussed in lecture.

1. Querying user public information

To refine results we will add spell checking abilities to the user’s search. We will also use methods of query expansion by using synonyms and related words.

REFERENCES

[1] How to Find and Reach Your Target Audience on Instagram, https://sproutsocial.com/insights/instagram-target-audience/

[2] 6 WAYS TO CREATE A MORE ENGAGED INSTAGRAM AUDIENCE, https://www.businessinsider.com/instagram-demographics-2013-12

[3] Top Instagram Demographics That Matter to Social Media Marketers, https://blog.hootsuite.com/instagram-demographics/

Conference Name:ACM Woodstock conference

Conference Short Name:WOODSTOCK’18

Conference Location:El Paso, Texas USA

ISBN:978-1-4503-0000-0/18/06

Year:2018

Date:June

Copyright Year:2018

Copyright Statement:rightsretained

DOI:10.1145/1234567890

RRH: F. Surname et al.

Price:$15.00