

CS 5316: Data Mining and Information Retrieval

Assignment 3

Lavanya Elango

TX State ID: A04365988

URL of web service

http://newfirebird.cs.txstate.edu/~l_e57/a3_entry.php

Summary

A web search has been implemented with Google search API that accepts a search query from user and lists 20 search results. The user can select up to five results and rerank the results based on the selection after pre-processing, using vector space model and cosine similarity.

The assignment was implemented with PHP, JavaScript using the Google custom search API.

Other tools: Postman addon in Chrome browser

Websites referred: W3 schools for JavaScript, Google documentation on API, stop word list - <http://www.ranks.nl/stopwords>

The start page where the search query is entered is

http://newfirebird.cs.txstate.edu/~l_e57/a3_entry.php

. Using the post method the search query is passed on to the next page

http://newfirebird.cs.txstate.edu/~l_e57/a3_parta.php

In this page the GET command for google API is built to include the search query from previous page. A unique API key and search engine ID created from google API/custom search was used to fetch the search results. Postman add-on in chrome browser was used to check the validity of the query.

The websites that are used by google custom search:

www.health.com
www.nih.gov
www.prevention.com
www.livestrong.com
www.eatingwell.com
www.mindbodygreen.com
www.vegetariantimes.com
www.yogajournal.com
www.mensfitness.com
www.healthywomen.org
www.Thekindlife.com
www.girlsgonestrong.com
www.well.blogs.nytimes.com

So, the queries should be based on health related topics.

As Google custom search did not give more than 10 search results, results 11 to 20 was obtained with start=11 in GET command. The results in JSON format were parsed to extract the link, title and snippet which are then displayed on the html page. The user is provided 5 options to select the results for re-ranking. The five inputs (some can be blank) and the search query are passed to the next page by POST method.

http://newfirebird.cs.txstate.edu/~l_e57/a3_partb.php

Using the search query, re-rank selections from previous page and the google search API, the documents for each search result, one document with combined results of re-ranking, and the dictionary with all words was built using JavaScript objects. The results in the order of unranked display were also stored in an array to be used later.

Preprocessing

The text in all the documents and dictionary was processed to ignore case, to look only for words with text, to remove empty or blank text and to remove stop words including the search query. The dictionary alone was processed to remove duplicate entries.

Vector space model

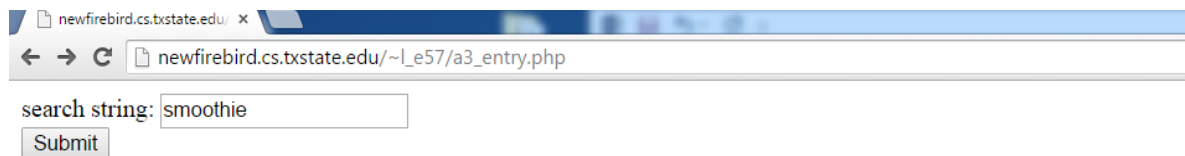
Term frequency (tf) was calculated for each word in the dictionary and normalized. The inverse document frequency (idf) was calculated for each word in the dictionary. tf.idf was calculated and normalized which created the vector space model for the documents.

Cosine similarity

To re-rank the search results, similarity between documents was calculated by cosine similarity of all documents against the re-rank document. From the cosine similarity result, the page order for highest to lowest similarity was obtained, that was used to display the results.

Experiments

Screenshot of my attempt to search for query "smoothie".



Screenshots of the fetched results in a3_parta.php

4/28/2015

newfirebird.cs.txstate.edu/~l_e57/a3_parta.php

Your query was: smoothie

[1 - 20 Super-Healthy Smoothies](#)

Dec 30, 2013 ... Blend up some refreshment with these 20 super-healthy smoothie recipes.

[2 - Quick and Healthy Smoothie Recipes | Eating Well](#)

Delicious and healthy smoothie recipes from EatingWell. Start your day off right with our great-tasting smoothie recipes.

[3 - 10 Slimming Smoothie Recipes](#)

Sip up and slim down with these delicious flat belly smoothie recipes.

[4 - smoothie - MindBodyGreen.com](#)

Healing + Warming Turmeric Smoothie With Bee Pollen. 3/20/2015 | Jamie ... 7 Ingredients To Add To Your Smoothie When You're Pregnant. 2/22/2015 | Lisa ...

[5 - Smoothie Recipes For Bloating, Fatigue, And Pain | Prevention](#)

6 Smoothies For Fatigue, Bloating, And Other Everyday Ailments. Candice Kumai October 29, 2014
Write a comment. Subscribe. X Close. State, AL, AK, AS, AZ ...

[6 - smoothies - MindBodyGreen.com](#)

Starbucks Is Now Serving Kale Smoothies (Yes, Seriously). 3/31/2015 | Emi ... 7 Ingredients To Add To Your Smoothie When You're Pregnant. 2/22/2015 | Lisa ...

[7 - Healthy and Fast Juice and Smoothie Recipes - Health.com](#)

Trying a cleanse, want to jumpstart your day, or need a post-workout treat? Try these amazing juice and smoothie recipes.

[8 - What Else Can We Add to a Smoothie? - NYTimes.com](#)

Mar 20, 2015 ... Over the years I have given you a number of Recipes for Health devoted to smoothies – fruit smoothies with nuts and seeds; smoothies with ...

[9 - Healthy Smoothie Recipes - Health.com](#)

Try a freshly blended smoothie. For a tasty, good-for-you beverage, using antioxidant-rich berries, chia seeds, and fresh herbs such as basil and rosemary ...

[10 - Banana-Maple Smoothie Recipe - Health.com](#)

Banana-Maple Smoothie - For a quick and easy breakfast option, blend this up in the morning for a healthy dose of potassium.

[11 - Thermos-Ready Smoothie Recipe | Eating Well](#)

"The great thing about smoothie recipes that you can alter them to your needs and taste. Too much sugar in these bad boys? I would cut out the apple juice and ...

[12 - Refreshing, Low-Cal Smoothies - Health.com](#)

Use these four quick and easy smoothie recipes as healthy breakfasts, snacks, and even desserts!

[13 - smoothies | Prevention](#)

Shed belly fat and satisfy your sweet tooth with these delicious fruit smoothies. 20 Super-Healthy Smoothie Recipes. 20 Super-Healthy Smoothies. The Editors of ...

[14 - Juice And Smoothie Recipes | Men's Fitness](#)

Drink Your Veggies Jamba Juice is offering small Fruit & Veggie smoothies for ... 14 Smoothie Super-Boosters Supercharge your smoothie, adding a boost of ...

[15 - Hawaiian Smoothie Recipe | Eating Well](#)

For the best taste and color, use a red-fleshed Hawaiian papaya for this exotic, lip -puckering taste of island summer. Papayas are loaded with papain,

[16 - Healthy Milk Shakes and Smoothies - Health.com](#)

These drinks combine low-fat dairy products, fruit, and even peanut butter to provide 11 recipes that contain less than 300 calories.

[17 - 5-Minute Kitchen: Blend a Green Smoothie | Men's Fitness](#)

If you're just getting your bearings around the blender, the idea of slipping spinach or kale into a smoothie might sketch you out. But we promise, if you do it right, ...

[18 - Strawberry Smoothie Recipe | Eating Well](#)

Reminiscent of an old-fashioned strawberry milkshake, this smoothie blends that cold, creamy luxury with the healthful benefits of fresh fruit.

[19 - Clean Breeze Smoothie Recipe | Eating Well](#)

This refreshing smoothie is made with cucumber and kiwi and gets a kick from ginger-flavored kombucha and fresh cilantro.

[20 - Feel-Good Smoothies - Health.com](#)

A variety of tasty smoothie recipes that will help you perk up, stay slim, and be superhealthy.

From the above 20 search results, enter the number of the corresponding result that is relevant, you may choose 1 to 5 and enter them in the box below .

Rank#1 result number:

Rank#2 result number:

Rank#3 result number:

Rank#4 result number:

Rank#5 result number:

The web results that are to be ranked are entered in the above screen shots.

4/28/2015

newfirebird.cs.txstate.edu/~l_e57/a3_partb.php

Your query was: smoothie

Your rank1# was: 16

Your rank2# was: 9

Your rank3# was: 2

Your rank4# was: 1

Your rank5# was:

[2 - Quick and Healthy Smoothie Recipes | Eating Well](#)

Delicious and healthy smoothie recipes from EatingWell. Start your day off right with our great-tasting smoothie recipes.

[3 - 10 Slimming Smoothie Recipes](#)

Sip up and slim down with these delicious flat belly smoothie recipes.

[9 - Healthy Smoothie Recipes - Health.com](#)

Try a freshly blended smoothie. For a tasty, good-for-you beverage, using antioxidant-rich berries, chia seeds, and fresh herbs such as basil and rosemary ...

[16 - Healthy Milk Shakes and Smoothies - Health.com](#)

These drinks combine low-fat dairy products, fruit, and even peanut butter to provide 11 recipes that contain less than 300 calories.

[10 - Banana-Maple Smoothie Recipe - Health.com](#)

Banana-Maple Smoothie - For a quick and easy breakfast option, blend this up in the morning for a healthy dose of potassium.

[12 - Refreshing Low-Cal Smoothies - Health.com](#)

Use these four quick and easy smoothie recipes as healthy breakfasts, snacks, and even desserts!

[13 - smoothies | Prevention](#)

Shed belly fat and satisfy your sweet tooth with these delicious fruit smoothies. 20 Super-Healthy Smoothie Recipes. 20 Super-Healthy Smoothies. The Editors of ...

[17 - 5-Minute Kitchen: Blend a Green Smoothie | Men's Fitness](#)

If you're just getting your bearings around the blender, the idea of slipping spinach or kale into a smoothie might sketch you out. But we promise, if you do it right, ...

[11 - Thermos-Ready Smoothie Recipe | Eating Well](#)

"The great thing about smoothie recipes that you can alter them to your needs and taste. Too much sugar in these bad boys? I would cut out the apple juice and ...

[4 - smoothie - MindBodyGreen.com](#)

Healing + Warming Turmeric Smoothie With Bee Pollen. 3/20/2015 | Jamie ... 7 Ingredients To Add To Your Smoothie When You're Pregnant. 2/22/2015 | Lisa ...

[14 - Juice And Smoothie Recipes | Men's Fitness](#)

Drink Your Veggies Jamba Juice is offering small Fruit & Veggie smoothies for ... 14 Smoothie Super-Boosters Supercharge your smoothie, adding a boost of ...

[18 - Strawberry Smoothie Recipe | Eating Well](#)

Reminiscent of an old-fashioned strawberry milkshake, this smoothie blends that cold, creamy luxury

with the healthful benefits of fresh fruit.

[20 - Feel-Good Smoothies - Health.com](#)

A variety of tasty smoothie recipes that will help you perk up, stay slim, and be superhealthy.

[19 - Clean Breeze Smoothie Recipe | Eating Well](#)

This refreshing smoothie is made with cucumber and kiwi and gets a kick from ginger-flavored kombucha and fresh cilantro.

[15 - Hawaiian Smoothie Recipe | Eating Well](#)

For the best taste and color, use a red-fleshed Hawaiian papaya for this exotic, lip -puckering taste of island summer. Papayas are loaded with papain,

[5 - Smoothie Recipes For Bloating, Fatigue, And Pain | Prevention](#)

6 Smoothies For Fatigue, Bloating, And Other Everyday Ailments. Candice Kumai October 29, 2014
Write a comment. Subscribe. X Close. State, AL, AK, AS, AZ ...

[8 - What Else Can We Add to a Smoothie? - NYTimes.com](#)

Mar 20, 2015 ... Over the years I have given you a number of Recipes for Health devoted to smoothies – fruit smoothies with nuts and seeds; smoothies with ...

[7 - Healthy and Fast Juice and Smoothie Recipes - Health.com](#)

Trying a cleanse, want to jumpstart your day, or need a post-workout treat? Try these amazing juice and smoothie recipes.

[6 - smoothies - MindBodyGreen.com](#)

Starbucks Is Now Serving Kale Smoothies (Yes, Seriously). 3/31/2015 | Emi ... 7 Ingredients To Add To Your Smoothie When You're Pregnant. 2/22/2015 | Lisa ...

[1 - 20 Super-Healthy Smoothies](#)

Dec 30, 2013 ... Blend up some refreshment with these 20 super-healthy smoothie recipes.

Observation

Entering only one result for re-rank (out of 5), resulted in same result being displayed first after re-ranking (this was what was intended too).

Comments

1. Part A: Only 10 search results were displayed. So to display 20 search results, GET command was created twice.
2. Part B: The system works fine for one word search query. For multiple word search query, after clicking the rerank button (in a3_parta.php), the POST method only passes the first word of the query (to a3_partb.php), rest of words are dropped while passing. So the results are obtained only from the first word of the query (is different from those obtained in previous page). If the full string is passed the results will match. Found this at the last moment - did not have the time to fix it.