

A Contribution to Rating and Recommendation Systems: Concepts, Development and Evaluation

Oliver Diestel

February 27, 2013

Current Situation (Just an assumption)

Google Science 2.0

Web Bilder Maps Shopping News Mehr + Suchoptionen

Umgreifend 266.000.000 Ergebnisse (0,36 Sekunden)

Science 2.0 - 8 The world's best scientists, the Internet's smartest...
[www.science20.com](#) - Diese Seite übersetzen
The world's best scientists. The Internet's smartest.
Physics - About Science 2.0 & ... - Who is Basking You On ... - Psychology

Forschungsverbund Science 2.0 - Science 2.0 - Wiki der ZfW - extem
[wiki.zfwi.uni... / Übersicht](#) - Science 2.0
19.02.2013 - **Science 2.0** befindet sich im Grundstadium der Fragen, wie das Internet mit seiner jährlichen Map 2.0 Anwendungs-Forschungs- und ...

Science 2.0 - Wikipedia, the free encyclopedia
[en.wikipedia.org/wiki/Science_2.0](#) - Diese Seite übersetzen
Science 2.0 is a somewhat controversial umbrella term, not precisely defined, which describes a range of activities, described by proponents of the term as ...
Definitions - History and background - Proliferation of Science 2.0 ... - See also

Science 2.0 Leibniz-Forschungsverbund
[www.leibniz-science20.de](#)
Willkommen in der Zukunft! 30 Einrichtungen erforschen neue Forschungs- und Publikationsprozesse. Alle Informationen über Forschungsschwerpunkte und ...

Science 2.0. Social Network für Wissenschaftler | heise online
[www.heise.de / News / 2009 + KIN 10](#)
02.03.2009 - Auf ResearchGate, einem Social Network speziell für Wissenschaftler, lassen sich jetzt Öffentliche Online-Volltexte anfragen, die alle ...

Science 2.0 - Leibniz Gemeinschaft
[www.leibniz-gemeinschaft.de/forschungsverbund... /science20/](#)
In dem Kart-Forschungsverbinder schreiben sich mehrere Leibniz-Einrichtungen zusammen.

Science 2.0 - Is an Open Access Science the Future? - Scientific...
[www.scientificamerican.com/article... /](#) - Diese Seite übersetzen
04.04.2008 - ...to transform how we do online. Here all it takes is a mouse click or a mouse click?

[illegible]

User visits the zbw.eu website searches for the information that he would like to have. Leaves the website.

Possible Situation

Google Science 2.0

Web Bilder Maps Shopping News Mehr Suchoptionen

Ungeliefert 266.000.000 Ergebnisse (0,38 Sekunden)

Science 2.0 - 8 The world's best scientists, the Internet's smartest ...
www.sciencet20.com/ - Diese Seite überbrücken
The world's best scientists. The Internet's smartest readers.
Physics - About Science 2.0 And ... - Rho is Basking Yaw On ... - Psychologe

Forschungsverbund Science 2.0 - Science 2.0 - Wiki der ZBW - extem
wiki.zbw.eu / Übersicht / Science 2.0
19.02.2013 - **Science 2.0** liefert sich im Grunde mit der Frage, wie das Internet mit seiner kollektiven Map 2.0-Anwendungs-Forschung- und ...

Science 2.0 - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Science_2.0 - Diese Seite überbrücken
Science 2.0 is a somewhat controversial umbrella term, not precisely defined, which describes a range of activities, described by proponents of the term as:
Definitions - History and background - Promotion of Science 2.0 ... - See also

Science 2.0 Leibniz-Forschungsverbund
www.leibniz-science20.de/
Willkommen in der Zukunft! 30 Einrichtungen erforschen neue Forschungs- und Publikationsprozesse. Alle Informationen über Forschungsschwerpunkte und ...

Science 2.0: Social Network für Wissenschaftler | heise online
www.heise.de / News / 2009 / KIM 10
02.03.2009 - Auf ResearchGate, einem Social Network speziell für Wissenschaftler, lassen sich jetzt öffentliche Online-Vorträge anzeigen, die alle ...

Science 2.0 - Leibniz Gemeinschaft
www.leibniz-gemeinschaft.de/forschung/leibniz - Science 20/
In der LeitF Forschungsverbünde schließen sich mehrere Leibniz Einrichtungen zusammen.

Science 2.0 - Is Open Access Science the Future? Scientific ...
www.sciencemanager.com/sci20/ - Diese Seite überbrücken
On 4/10/2008, I received new results online. And all I saw, a result but in a real day?

How to mention a user

I recently noticed this sentence in the help section:

You can @mention users anywhere in the text to point their attention, follow users and conversations and report inappropriate content by flagging it.

However, I seem to be too stupid to use it ... I tried:

- * <username>
- @<username>

What is the correct way to use this neat trick?

Well, removed the space in the first syntax

Comments

@@update tries to mention himself
5d ago (37 hours ago) · edit

1 Answer

Sort by · oldest · newest · most voted

@@update works fine, I think the problem is you are giving a space between @ and username in your first syntax.

Comments

You are correct, I wrote - at least here. Somehow the same thing doesn't seem to work over at ROS. I answered, e.g. http://answers.ros.org/question/109908/ros-for-no-new-users/?answer=109910#post-id-109909 there. Any idea why?

User visits the website, sees an interesting question from another user, reads the question and the answer, asks hopefully more questions. Goal increase the user interaction on a website.

q/A System

The screenshot shows the AskWing website interface. At the top, there are navigation tabs: "questions", "tags", "people", "badges", and "ask a question". Below these is a search bar. The main content area is divided into two columns. The left column lists questions with their titles, tags, and statistics (votes, answers, views). The right column displays a grid of contributor avatars and names. The questions listed include "My Gameduino page", "Can You plug in a Wii Nunchuck?", "Gameduino Reference Manual v01", "Due compatibility?", "Can I use the Gameduino on a 60Hz TV?", "Char or sprites with random maze?", "Gameduino with MicroSD card breakout", and "gameduino reads graphics from sd card".

The screenshot shows a specific question on the AskWing website. The question is titled "Gameduino with MicroSD card breakout" and is tagged with "gameduino" and "MicroSD". The question text asks for help with a MicroSD breakout board and pin connections. Below the question, there are two answers. The first answer is from a user named "gameduino" and provides a link to a tutorial. The second answer is from a user named "gameduino" and provides a link to a tutorial. The interface includes navigation tabs, a search bar, and a list of related questions on the right side.

General Idea: Use a Question/Answer System(Picture) to make the information retrieval process public. Use this public questions and answers to increase the user interaction on the website.

Beginning

What do we have at the beginning? Website, User and Information(Question)

After every new concept answer the question what do we have now

Rating

- ▶ Click
- ▶ active time on page

If a user stays shorter or longer on a page than the average user - it is a high or low rating.

We rate QA pages and webpages, however we only recommend qa pages.

Now we have Qustion Answer System, User and the information which user likes which question as well as our own website.

Tagging

- ▶ Standard-Thesaurus Wirtschaft, tripple store
- ▶ Stemming: Compute the root of a word.
- ▶ Computing Levenshtein distance: Calculate the distance between two words.

Principle

Every question will be tagged with a preferred label from the ZBW Thesaurus. For every word in the question find the root of the word and match it against the thesaurus. If we find a matching preferred label, Done. If we find a matching synonym take the preferred label as tag, Done.

What do we have now? Information about which user likes which question and one or more descriptor that describes the subject of the question

Input/Output

Input: $\langle \text{user_id}, \text{item_id}, \text{rating} \rangle$, [List of tags], number of recommendations

Output: Sorted list of Items with avg. rating that correspond with the pref. input label

- ▶ Item: id, name, tag
- ▶ User: id, name

Question: How does this black box work? Three part recommendation. Pipeline architecture

- ▶ Calculate similarity between items
- ▶ Add similar ratings to a user/rating vector.
- ▶ Calculate singular value decomposition
- ▶ Find similar users

Item-Based Algorithm

Display a matrix with five users and five objects not every user has a rating for every object To find similar users we need as much user/item ratings as possible so we would like to guess as much ratings as possible.

Find items that have similar ratings. Calculate the similarity values for every item.

Cosinus Similarity

$$\text{sim}(\vec{a}, \vec{b}) = \frac{\vec{a} * \vec{b}}{|\vec{a}| * |\vec{b}|}$$

Take the differences of the average rating behaviour of the user into account.

Advanced Cosinus Similarity

$$\text{sim}(a, b) = \frac{\sum_{u \in U} (r_{u,a} - \bar{r}_u)(r_{u,b} - \bar{r}_u)}{\sqrt{\sum_{u \in U} (r_{u,a} - \bar{r}_u)^2} \sqrt{\sum_{u \in U} (r_{u,b} - \bar{r}_u)^2}}$$

Calculate predictions for similar items.

$$\text{User } u, \text{ Product } p \quad \text{pred}(u, p) = \frac{\sum_{i \in \text{ratedItems}(u)} \text{sim}(i, p) * r_{u,i}}{\sum_{i \in \text{ratedItems}(u)} \text{sim}(i, p)}$$

Fill recommendation vector of each user with similar item ratings
What do we have now? Matrix with less empty fields than before

Singular Value Decomposition

Now we would like to find similar users with our matrix

Create a SVD with the matrix $M = U * \Sigma * V^t$

- ▶ U dimension $m \times m$, orthogonal matrix, spans the column space of matrix m
- ▶ Σ dimension $m \times n$, diagonal matrix having only r nonzero entries
- ▶ V dimension $n \times n$, orthogonal matrix, spans the row space of matrix m

Σ Properties

Diagonal entries of Σ have the property, that $\sigma > 0$ and $\sigma_i \geq \sigma_{i+1}$

Calculate cosinus similarity between users

Now we have a two dimensional matrix, so we can calculate the cosinus similarity between our users. Display a graph with the users, show which users are similar display the original matrix next to the graph to validate the result.

Is this a good approach?

We could have calculated the SVD directly out of the original matrix. Would we get a similar result? Is it better if we use the v matrix from the SVD to calculate the item similarity and update the SVD afterwards? Is it better if we just recommend similar items from the items the user already likes?

Evaluation

The test data: I use real data from the zbw econdesk. However we do not have real user ratings for this data. Every data gets a random quality value. 1-5 if a user would rather rate it positive or negative. Furthermore I generate 1000 test users these users will have a rating preferation, so a user might be a person that rates an item more positive or more negative. I will try to evaluate the algorithms with this test data. I might use a movie db as well for this.

Find Questions

Find questions that are highly rated by similar users that match the current topic

Display Questions

Add a tag and a div-placeholder on every webpage. Display top questions for the current user on the webpage.

Software Architecture

Scala: Finagle twitter framework

Every part of the software is a service.

The tagger, the rating algorithm and the recommendation algorithm can be used as an individual software.

Pros:

Cons:

Timetable

Start: 18.12.2012

End: 18.06.2012

1st month(18.01): Theory

2nd month(18.02): Theory + Technology

3rd month(18.03: Implementation

4th month(18.04: Implementation + First writings

5th month(18.05): Final thesis

Personal Goal

Technology implemented: 31.03.2012

Diplom Thesis ready: 01.05.2012