CareerVillage.org

Stanford University
CS+Social Good Workshop

Feb 2016

www.careervillage.org

Today's workshop

- Tackle a real-world example of how collective intelligence can be used to solve a social good challenge
- Use Python and real data to practice entity matching
- Introduce more advanced concepts for future projects

Everything you need today

Go here:

https://github.com/cs4good/social-good-106-section/tree/master/collective-intelligence

Contents:

1. PDF of this presentation
2. Python file

Agenda

- Intro to CareerVillage.org
- The problem we're focusing on today
- Instructions
- Advanced concepts

Agenda

- **Intro to CareerVillage.org**
- The problem we're focusing on today
- Instructions
- Advanced concepts

America's youth are finishing their education unready for employment

> *"For this generation of young people, the future looks bleak....No, they are not the idle youth of Greece or Spain or Egypt. They are the youth of America, the world's richest country….For this group, finding work that pays a living wage and offers some sense of security has been elusive."* - NYTimes, June 2012

500 students for every guidance counselor

HS students unaware of their career options

Students lack career goals entering college

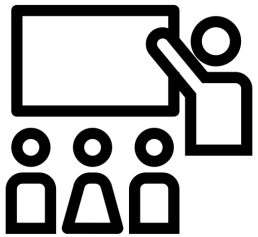
83% of college students graduating jobless

Highest youth unemployment in decades

14 Million jobs unfilled by 2020

Students need better access to career advice

Research shows having career goals and plans improves grades, degrees, and employment



Career goals boost academic performance

- ✓ Setting long-term career goals predicted positive academic performance. Source: Ting (1997)
- ✓ Uncertainty about career goals contributed to poor academic performance. Source: Altmaier, Rapaport, and Seeman (1983)

Career plans and goals boost college persistence and degree attainment

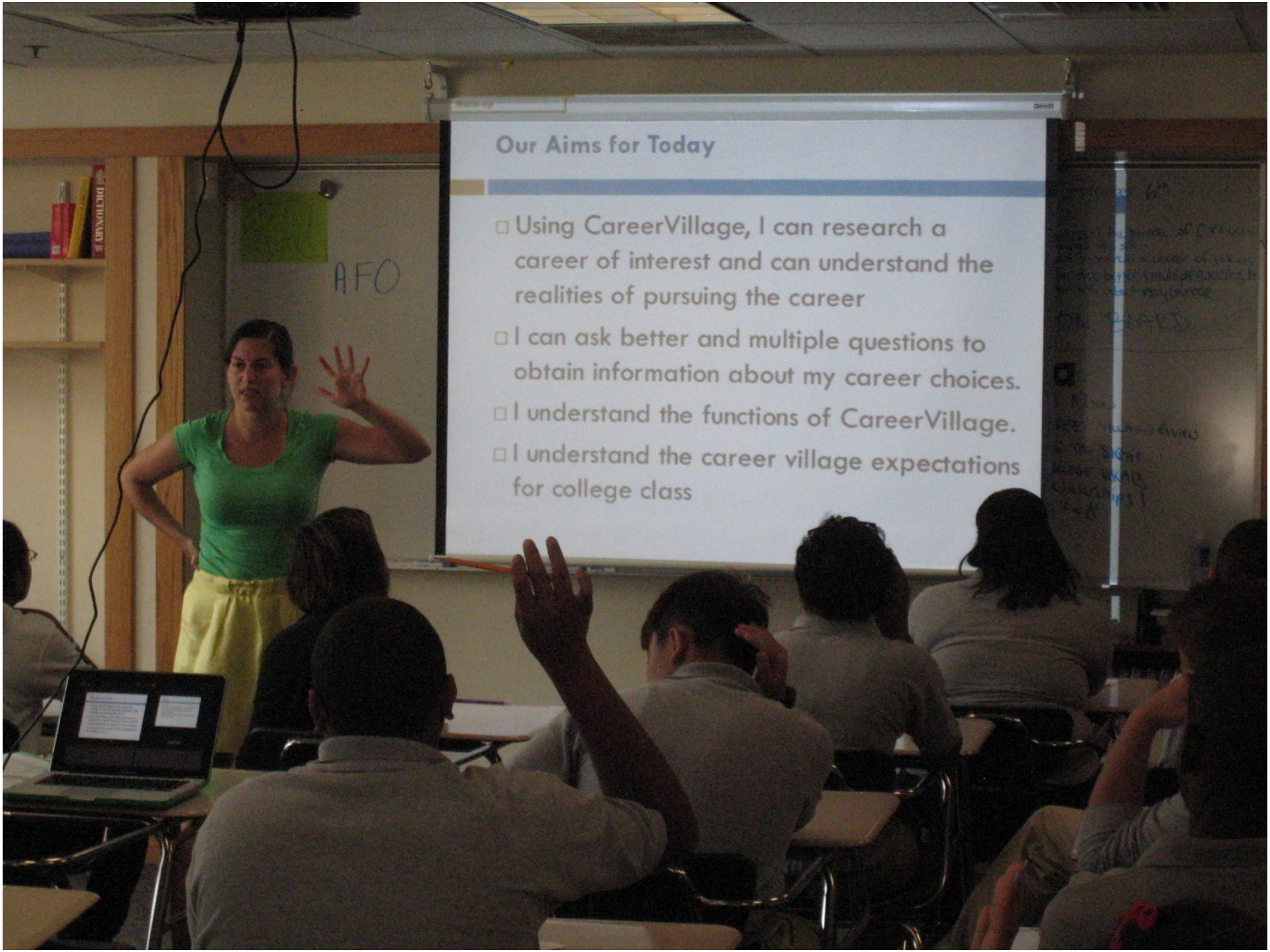
- ✓ Well-defined career plans or goals positively influence decisions to remain in college. Source: Tinto, Vincent. "Leaving college and rethinking the causes and cures of student attrition." University of Chicago Press (1993).
- ✓ Students with a defined job-related career goal were more likely to persistent in college. Source: Hull-Blanks, et al. "Career goals and retention-related factors among college freshmen." Journal of Career Development 32.1 (2005)
- ✓ Among international students career certainty leads to college persistence. Source: Singaravelu, Hemla D., Lyle J. White, and Tammy B. Bringaze. "Factors Influencing International Students' Career Choice A Comparative Study." Journal of Career Development 32.1 (2005)

Career goals boost job attainment

- ✓ Expressed career aspirations or goals have also been shown to predict job attainment (Gottfredson & Becker, 1981; Schoon & Parsons, 2002; Mau & Bikos, 2000).

Other related research reports: Schoon, Ingrid, and Samantha Parsons. "Teenage aspirations for future careers and occupational outcomes." Journal of Vocational Behavior 60.2 (2002): 262-288. Gottfredson, Linda S., and Henry J. Becker. "A challenge to vocational psychology: How important are aspirations in determining male career development?." Journal of Vocational Behavior 18.2 (1981): 121-137. McWhirter, Ellen Hawley, et al. "The effects of high school career education on social–cognitive variables." Journal of Counseling Psychology 47.3 (2000): 330. "The Path to Career Success: High School Achievement, Certainty of Career Choice, and College Readiness Make a Difference." Iowa City, IA: ACT (2009). Symonds, William C., et al. "Pathways to prosperity" Harvard GSE. (2011)


CareerVillage.org


Our Aims for Today
Using CareerVillage, I can research a career of interest and can understand the realities of pursuing the career
I can ask better and multiple questions to obtain information about my career choices.
I understand the functions of CareerVillage.
I understand the career village expectations for college class
AFO
DICTIONARY

"What are after school activities that I should do if I plan to study computer engineering in college?" – Manuel B.

"I'm looking into scholarships, but I'm scared that I'm not going to be able to find one. Is there anything that could help me sort of stand out or will i have to just hope my experiences meet and surpass the requirements?" - Ignacio F.

"How do I know if I have what it takes to be a programmer?" – Maeve K.

"How do you keep yourself motivated at work? Sometimes work can also be repetitive so how do you keep yourself motivated?" – Marissa P.

7
Likes

What are the pros and cons in a day of a pediatrician? What is the best thing about being a pediatrician?

I am a 9th grader who is interested in medicine and helping others. I have been interested in becoming this since I was younger and still have an interest in it.

Question Topics

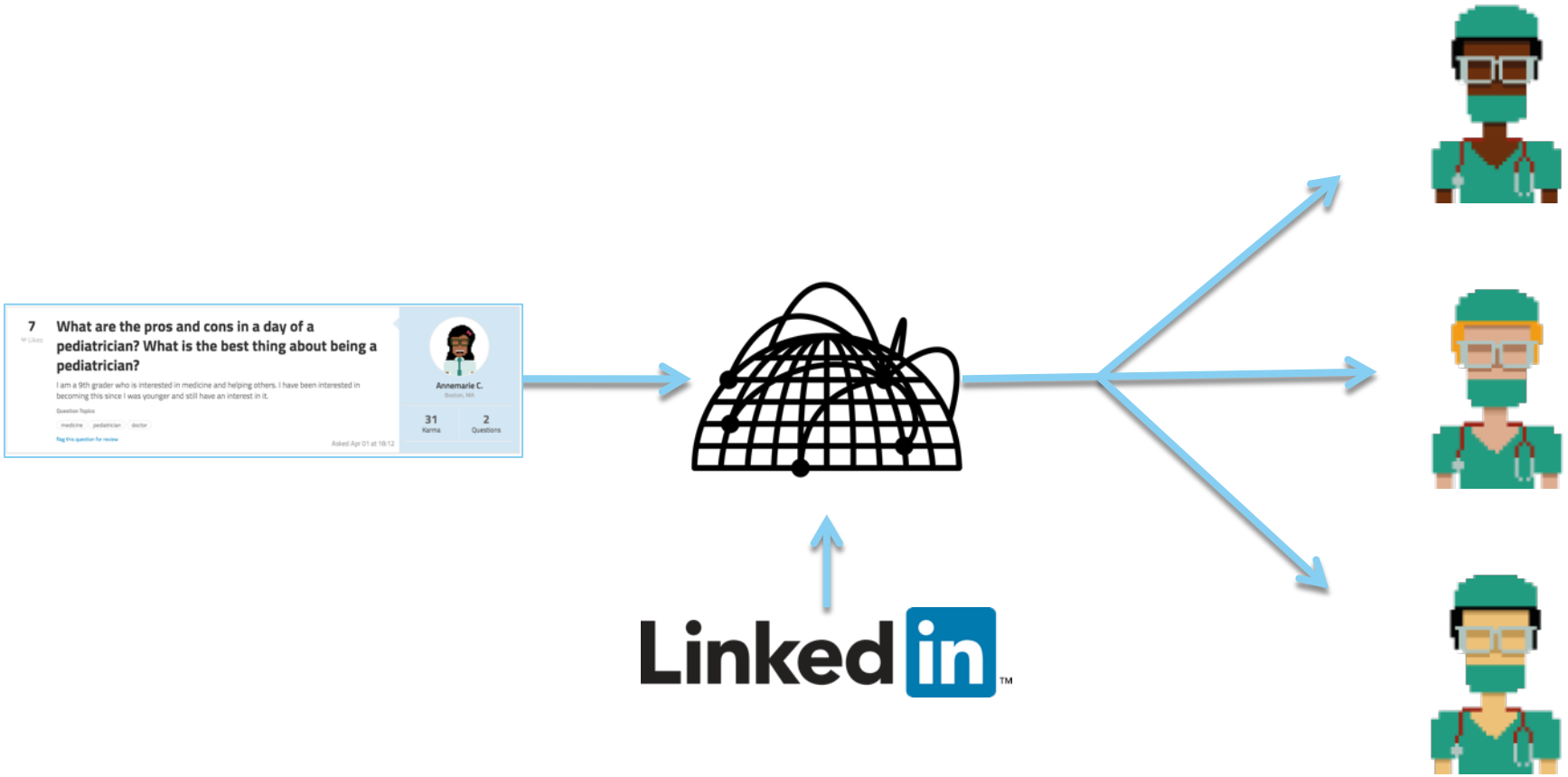
medicine | pediatrician | doctor

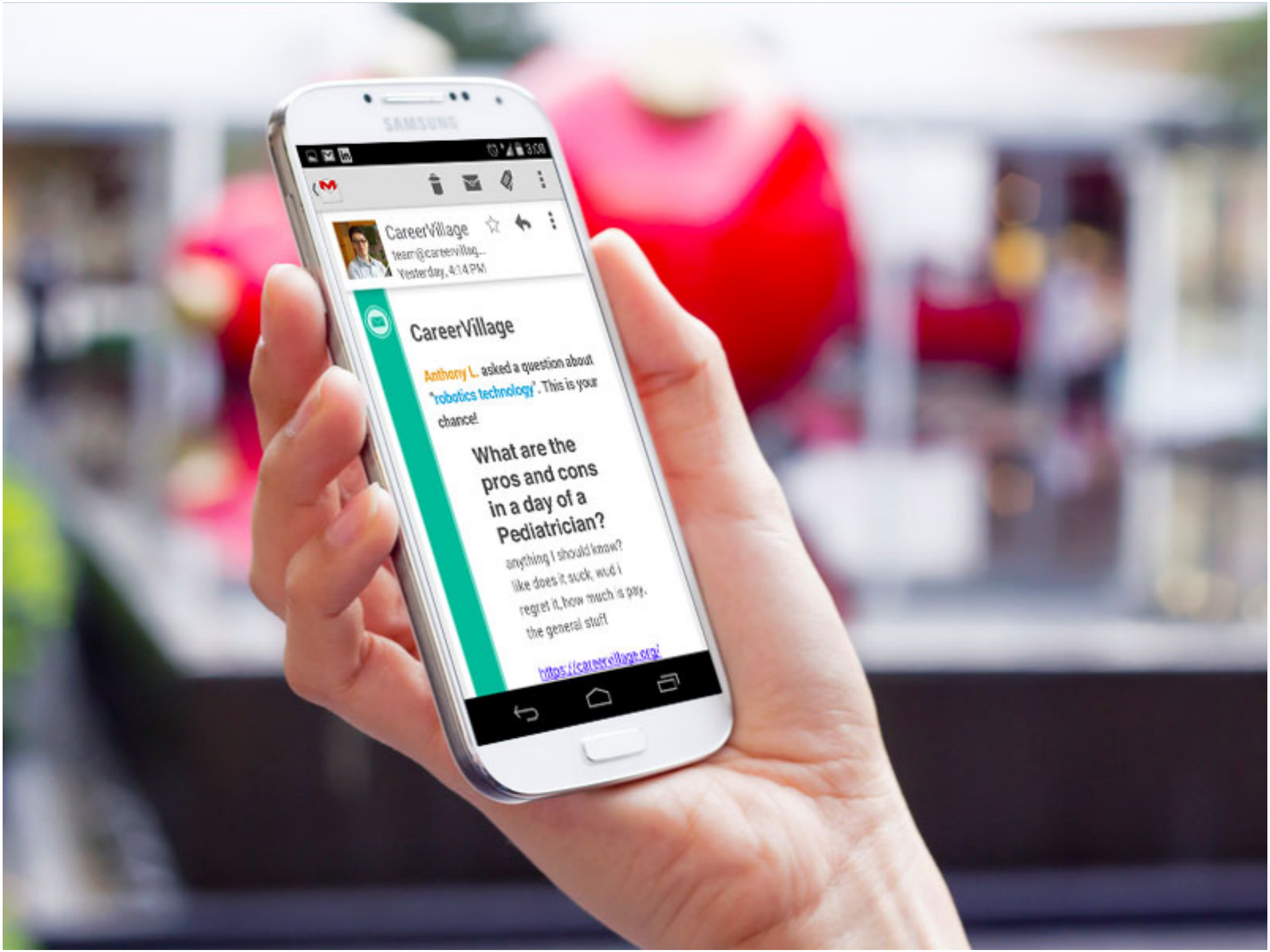
flag this question for review

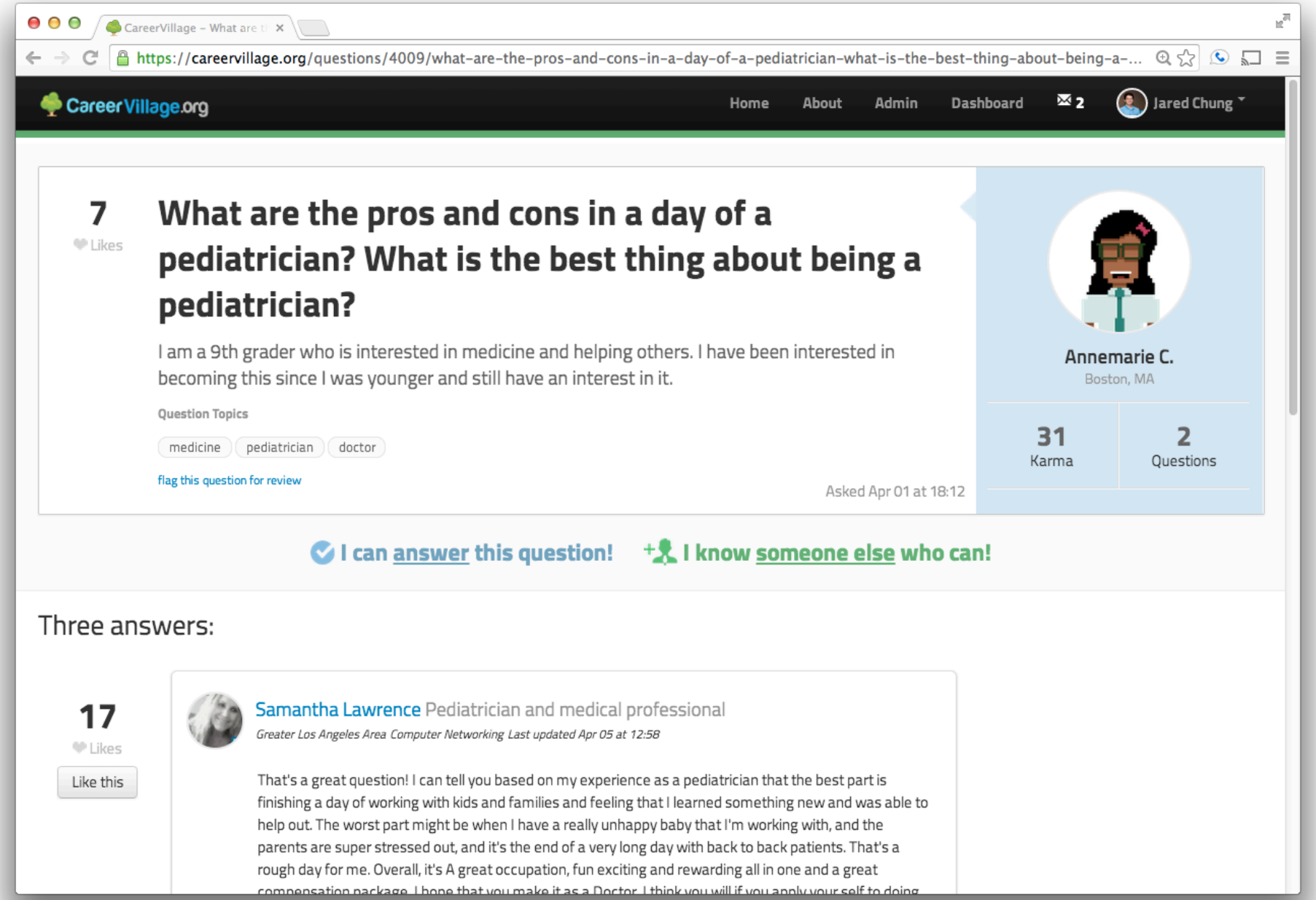
Asked Apr 01 at 18:12

Annemarie C.
Boston, MA

31	2
Karma	Questions


7
What are the pros and cons in a day of a pediatrician? What is the best thing about being a pediatrician?
I am a 9th grader who is interested in medicine and helping others. I have been interested in becoming this since I was younger and still have an interest in it.
Annemarie C.
31
Karma
2
Questions
LinkedIn


CareerVillage
team@careervillag...
Yesterday, 4:14 PM
CareerVillage
Anthony L. asked a question about "robotics technology". This is your chance!
What are the pros and cons in a day of a Pediatrician?
anything I should know? like does it suck, would I regret it, how much is pay, the general stuff
https://careervillage.org/


CareerVillage – What are th
https://careervillage.org/questions/4009/what-are-the-pros-and-cons-in-a-day-of-a-pediatrician-what-is-the-best-thing-about-being-a-...
CareerVillage.org
Home
About
Admin
Dashboard
2
Jared Chung
7
Likes
What are the pros and cons in a day of a pediatrician? What is the best thing about being a pediatrician?
I am a 9th grader who is interested in medicine and helping others. I have been interested in becoming this since I was younger and still have an interest in it.
Question Topics
medicine
pediatrician
doctor
flag this question for review
Asked Apr 01 at 18:12
Annemarie C.
Boston, MA
31
Karma
2
Questions
I can answer this question!
I know someone else who can!
Three answers:
17
Likes
Like this
Samantha Lawrence Pediatrician and medical professional
Greater Los Angeles Area Computer Networking Last updated Apr 05 at 12:58
That's a great question! I can tell you based on my experience as a pediatrician that the best part is finishing a day of working with kids and families and feeling that I learned something new and was able to help out. The worst part might be when I have a really unhappy baby that I'm working with, and the parents are super stressed out, and it's the end of a very long day with back to back patients. That's a rough day for me. Overall, it's A great occupation, fun exciting and rewarding all in one and a great

To save and rescue animals that requires only classes in business start up development, maketing, classes on how to handle animals, animal trainer, etc.

Four comments

Andrea H. Delete this May 01 at 12:18

Thank you for all the information and career options! Humanitarian efforts are truly important.

Ayan A. Delete this May 01 at 12:19

Thank you Holly. Very thoughtful.

Maeve K. Delete this May 01 at 14:56

Thank you for all this information and taking the time to write this very detailed response

Teal K. Delete this May 01 at 15:06

This is very informative. Thank you, Holly

Our volunteers work for many of America's top companies

Accenture, Adobe Systems, Aereo, Akamai, Akin Gump Strauss Hauer & Feld LLP, Allstate Insurance Agency, Amazon, American Express, AXA Group, Bank of America, Barclays, Bausch + Lomb, BEA Systems, Inc., Berklee College of Music, Best Buy, Bloomberg, Boeing, Booz Allen Hamilton, Calvin Klein, CIGNA Healthcare, Cisco Systems, Citibank, Credit Suisse, Deloitte, Delta Air Lines, Digitas, Inc., Dolby, Donna Karan International, Inc., DuPont, Electronic Arts, Ernst & Young, Facebook, Fitbit, Forbes Magazine, Ford Motor Company, GE, General Motors, Goldman Sachs, Google, Harvard University, Hertz, Honeywell, IBM, Intuit, iRobot, Johnson & Johnson, LinkedIn, Los Alamos National Laboratory, LucasArts, MAC Cosmetics, Macy's, Major League Soccer, McKinsey & Company, Merck, Merrill Lynch, Microsoft, Morgan Stanley, MTV Networks, NASA Ames Research Center, NBC-Universal, Netflix, Novartis, Partners HealthCare, PepsiCo, Pfizer Pharmaceuticals, Prudential Healthcare, PwC, Qualcomm, Raytheon, Siemens, Sony Electronics, Sony Pictures Entertainment, State Street, T-Mobile, Target, Teach for America, TechStars, Tesla Motors, The Boston Consulting Group, The College Board, The Huffington Post, The Museum of Modern Art, The New York Times, The Walt Disney Company, Thomson Reuters, TripAdvisor, U.S. House of Representatives, UBS, UNHCR, UnitedHealthcare, Universal Music Group, UPS, US Airways, US Air Force, US Army, US Department of Defense: National Security Agency (NSA), US Navy, USA Today, Verizon Wireless, Visa, Warner Brothers Studios, Williams-Sonoma, Inc., World Economic Forum, Yahoo!, Zynga, and many more…

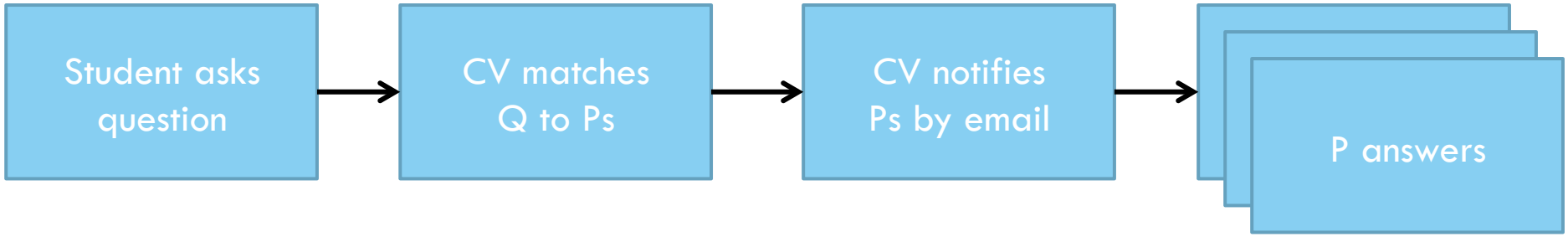
Other stuff

- Points, badges, etc.
- Social
- Career plan reviews
- Company impact
- Etc.

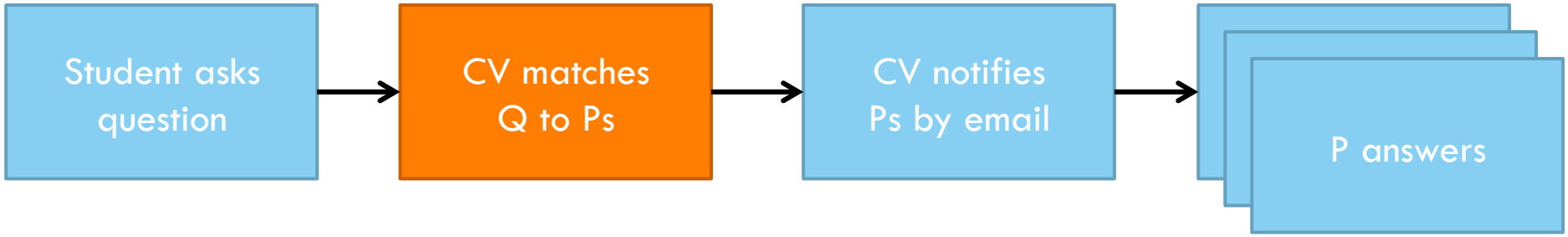
Agenda

- Intro to CareerVillage.org
- **The problem we're focusing on today**
- Instructions
- Advanced concepts

Today we get to the heart of the issue



Today we get to the heart of the issue



How do we do matching like this?

- **Machine learning**

How do we do matching like this?

- **Machine learning**

  "subfield of artificial intelligence (AI) concerned with algorithms that allow computers to learn. What this means, in most cases, is that an algorithm is given a set of data and infers information about the properties of the data—and that information allows it to make predictions about other data that it might see in the future. This is possible because almost all nonrandom data contains patterns, and these patterns allow the machine to generalize. In order to generalize, it trains a model with what it determines are the important aspects of the data."

How do we do matching like this?

- **Machine learning**
- **Collective Intelligence**

How do we do matching like this?

- **Machine learning**
- **Collective Intelligence**

  “combining of behavior, preferences, or ideas of a group of people to create novel insights.”

How do we do matching like this?

- Machine learning
- Collective Intelligence

How do we do matching like this?

- **Machine learning**
- **Collective Intelligence**

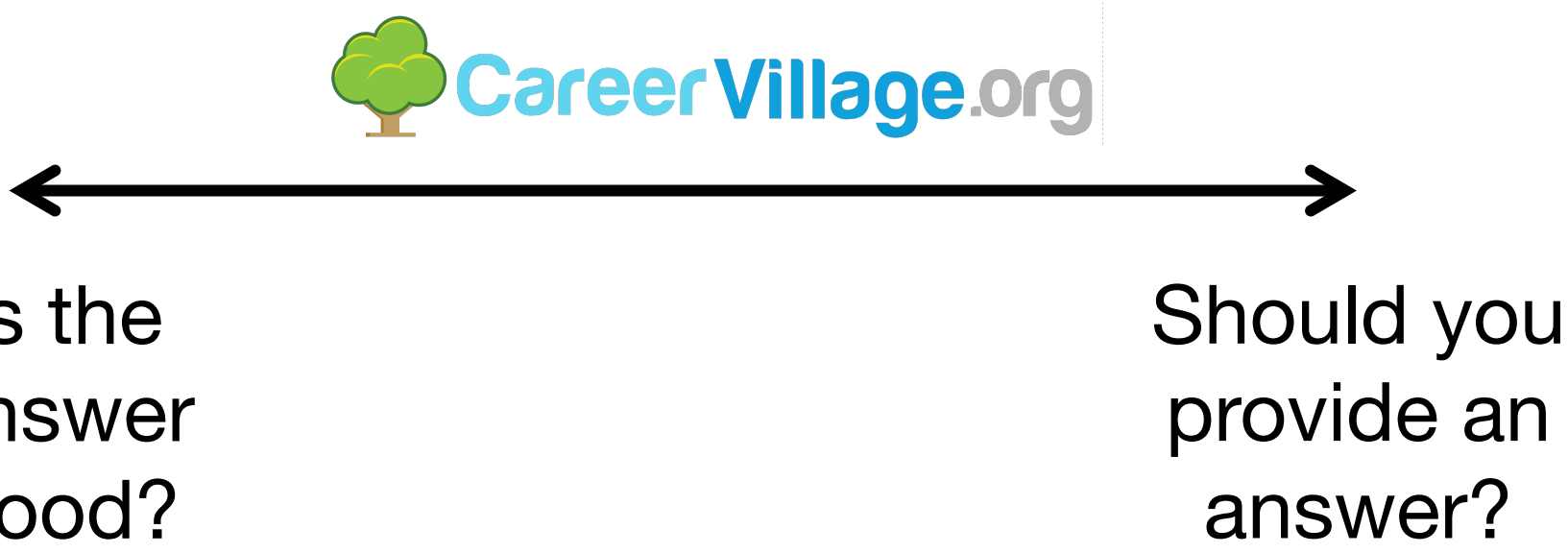
amazon

Is this product good

Which product is best for you?

How do we do matching like this?

- **Machine learning**
- **Collective Intelligence**



How do we do matching like this?

You like Apples.

- Pears are similar to Apples so you will also like Pears.
- Other people who like Apples like Pears so you will also like Pears.
- You like Apples more than Bananas and people who like Apples more than Bananas also usually like Pears more than Bananas so you will also like Pears.

Agenda

- Intro to CareerVillage.org
- The problem we're focusing on today
- **Instructions**
- Advanced concepts

How to get started

1. Get the Python file out of https://github.com/cs4good/social-good-106-section/tree/master/collective-intelligence
2. Load the CSV files from the dropbox (links in the python file)
3. Build your matching algorithm
4. Calculate your success rate

(Iterate if you can!)

What we're giving you

Data:

- **Question data:** question_id, title, body, tags, author_id, added_at
- **Answer data:** answer_id, id_of_question_being_answered, body, author_id, added_at, score
- **User data:** id, type, date_joined, headline, industry, location, answer_count, reputation, notification_setting, tags_followed_list

What we're giving you

Data:

- **Question data:** question_id, title, body, tags, author_id, added_at
- **Answer data:** answer_id, **id_of_question_being_answered**, body, author_id, added_at, score
- **User data:** id, type, date_joined, headline, industry, location, answer_count, reputation, notification_setting, tags_followed_list

How good is your matching algorithm?

Precision = % of predictions which are correct

Recall = % of correct matches predicted

F-Measure aka "F1" = (2 * P * R) / (P + R)

How good is your matching algorithm?

Precision = % of predictions which are correct

Recall = % of correct matches predicted

F-Measure aka "F1" = (2 * P * R) / (P + R)

```
Example:
Total Answers: 1,000          P = 66.67%
Correctly predicted: 800      R = 80.00%
Total predictions: 1,200      F = 0.73
```

Start with one of these...

1. If the first letter of the question's title matches the first letter of the user's headline, match them. (F1 0.00092)
2. If the question has at least 5 tags, match it with every user who follows at least 10 tags (F1 0.00253)
3. If the user is a "P" and follows at least one term the question is tagged with, then match them (F1 0.01638)

Agenda

- Intro to CareerVillage.org
- The problem we're focusing on today
- Instructions
- **Advanced concepts**

This can get SUPER sophisticated

- Group clustering
- Content-based ranking
- Semantic analysis
- Optimizations
- Genetic algorithms

Recommended resources