Abhishek Jaiswal

Iit2013129@iiita.ac.in

Algorithm

- I have taken five attributes of a student namely : **Name(I divided it into Front Name, Last Name), College Name, Email, Phone, DOB.**
- Now, as we can observe these all **attributes have different weight value of assuring a successful map** if they are matched in the page. (By successful map I mean that we are sure that mentioned information in that page is about the particular student we are searching for).
- Example: If name of student is matched then it is of smaller importance compared to DOB, phone and Email matching.
- Based on this fact, I assigned different weight values to different attributes :

ATTRIBUTE	SCORE
Front Name	2
Full Name	4
College	5
Phone	10
DOB	9
Email	10

- Then I divided my prediction result into five possible cases : Very Sure, Sure, May be, Can't Say, No.
- I fixed a range for each of the cases to be true.

RESULT	SUM OF SCORES
Very Sure	>= 19
Sure	>= 10
May be	>= 6
Can't Say	>= 1
No	0

- I have chosen these ranges based on following observations:
  - ➢ **Very Sure** is only possible when **any two of (Phone, DOB, Email)** is matched.
  - ➢ **Sure** is possible when: **(F name + L name + College)** OR **either of (Phone, Email)** OR **(DOB with any other attribute) is matched.**
  - ➢ **May be** is possible when **any two of (F name, L name, College)** are matched.

- ➢ **Can't Say** when **any one of attribute** is matched.
- ➢ **No** when **nothing matched**.
- • To match our attributes with the Web Pages data, I **converted both of them into Lower Cases** so that cases do not cause an error in matching.
- • There are **many formats of DOB ex: (8-2-1995 OR 8/2/1995 OR 2/8/1995 OR 8 feb 1995**).
  To overcome this problem of matching, I used following steps:
  - ➢ As we can see they **all have value of day in the same format. i.e.(8)**
  - ➢ So I **first searched for the date** i.e. 8 in this example.
  - ➢ Once I found an index of matching, I **searched for matching of Month and Year in The windows of size 15 in both the directions from the Index where date was matched**.
  - ➢ For the month matching I considered both the cases for each month i.e.(2 and feb).
  - ➢ We repeat this step for each of index where we found matching of day.

- • **Program stores the result of each link that it grabbed in the Result.txt file**.
- • Program **creates a .txt file automatically for each line it grabbed, which is could be then converted to .htm to open the webpage when offline**.
- • Program gives very good results for the inputs I have tried.

SAMPLE INPUT OUTPUT RESULTS:

Enter Student Name

Abhishek Jaiswal // user entered

Enter Students College Name

IIIT Allahabad   // user entered

Enter Students email

abhsk246@gmail.com   // user entered

Enter Students contact no..

8737862286   // user entered

Enter Students DOB ex(24/2/1995) don't add a 0 before 2

8/2/1995 // user entered

Please enter the search term.

abhishek jaiswal iiit Allahabad // text you want to search on google

Please enter the number of results. Example: 5 10 20

10   // it is the number of links that you want to check for matching the student

FOLLOWING ARE THE LIST OF TOP 10 Links that came when searched for given text.

Text::Abhishek Jaiswal Profiles | Facebook, URL::=https://www.facebook.com/public/Abhishek-Jaiswal

Text::Profiles @ IIIT Allahabad, URL::=http://profile.iiita.ac.in/mmishra/

Text::Abhishek Jaiswal - Quora, URL::=https://www.quora.com/profile/Abhishek-Jaiswal-21

Text::CA Abhishek Jaiswal - LinkedIn, URL::=https://in.linkedin.com/in/abhishek-jaiswal-b4bb9a80

Text::Abhishek Jaiswal | LinkedIn, URL::=https://in.linkedin.com/in/platinumabhi

Text::abhishek_iiita - CodeChef User | CodeChef, URL::=https://www.codechef.com/users/abhishek_iiita

Text::Abhishek Jaiswal (abhishek.jaiswal.7106670)'s Developer Profile ...,
URL::=https://www.hackerearth.com/users/abhishek.jaiswal.7106670/

Text::Abhishek jaiswal - SPOJ, URL::=http://www.spoj.com/users/abhigeek

Text::Abhishek Jaiswal (@Abshk111) | Twitter, URL::=https://twitter.com/abshk111

Text::Abhishek Jaiswal @abhi._jaiswal Instagram photos | Websta ...,
URL::=http://websta.me/n/abhi._jaiswal

link grabbed :https://www.facebook.com/public/Abhishek-Jaiswal

text begins -------------------------------

first name matched -> abhishek

last name matched -> abhishek jaiswal

----

https://www.facebook.com/public/Abhishek-Jaiswal score-> 6

https://www.facebook.com/public/Abhishek-Jaiswal may be

link grabbed :http://profile.iiita.ac.in/mmishra/

text begins -------------------------------

first name matched -> abhishek

college matched -> iiit allahabad

last name matched -> abhishek jaiswal

----

http://profile.iiita.ac.in/mmishra/ score-> 11

http://profile.iiita.ac.in/mmishra/ sure

link grabbed :https://www.quora.com/profile/Abhishek-Jaiswal-21

text begins -------------------------------

first name matched -> abhishek

college matched -> iiit allahabad

last name matched -> abhishek jaiswal

----

https://www.quora.com/profile/Abhishek-Jaiswal-21 score-> 11

https://www.quora.com/profile/Abhishek-Jaiswal-21 sure

link grabbed :https://in.linkedin.com/in/abhishek-jaiswal-b4bb9a80

text begins -------------------------------

link grabbed :https://in.linkedin.com/in/platinumabhi

text begins -------------------------------

first name matched -> abhishek

last name matched -> abhishek jaiswal

college matched -> iiit allahabad

----

https://in.linkedin.com/in/platinumabhi score-> 11

https://in.linkedin.com/in/platinumabhi sure

link grabbed :https://www.codechef.com/users/abhishek_iiita

text begins --------------------------------

first name matched -> abhishek

college matched -> iiit allahabad

----

https://www.codechef.com/users/abhishek_iiita score-> 7

https://www.codechef.com/users/abhishek_iiita may be

link grabbed :https://www.hackerearth.com/users/abhishek.jaiswal.7106670/

text begins --------------------------------

first name matched -> abhishek

last name matched -> abhishek jaiswal

----

https://www.hackerearth.com/users/abhishek.jaiswal.7106670/ score-> 6

https://www.hackerearth.com/users/abhishek.jaiswal.7106670/ may be

link grabbed :http://www.spoj.com/users/abhigeek

text begins ------------------------------

first name matched -> abhishek

college matched -> iiit allahabad

last name matched -> abhishek jaiswal

----

http://www.spoj.com/users/abhigeek score-> 11

http://www.spoj.com/users/abhigeek sure

link grabbed :https://twitter.com/abshk111

text begins ------------------------------

first name matched -> abhishek

last name matched -> abhishek jaiswal

college matched -> iiit allahabad

----

https://twitter.com/abshk111 score-> 11

https://twitter.com/abshk111 sure

link grabbed :http://websta.me/n/abhi._jaiswal

text begins ------------------------------

RESULT

- All the links where score is 11 and the output is sure are actually referring to me only.
- I have tried other test cases as well , it worked correctly for all of them.