**171227**

|  |
| --- |
| this is the "original" guess the calories quiz:  <http://zephyr.sista.arizona.edu/caloriequiz/index.html>  the new one (learn2cal) is here  <http://zephyr.sista.arizona.edu/learn2cal/>  we want to improve this and base v.0.1 of the phone app on this simple quiz.  s. |

기존의 20개 질문으로 구성된 Quiz - 평균 정답률 25% (5개 정도 맞춤)

Wisdom of crowd를 이용하려는 이유? 개인의 정답률은 낮더라도 그들의 평균은 정답에 수렴한다..

실제로 적용을 해보니 정답률이 5/20 -> 7/20으로 조금 올라갔다.

**Step1.**

Guess the correlation 게임과 비슷한 방식으로 사람을 train 해보고 싶다.

개인의 bias를 바로잡아서 정확도를 올릴 수 있도록 만드는 것이 목표.

기존의 20문제 퀴즈처럼, 문제를 다 풀고 결과를 보여주는 것이 아니라, 한 문제를 풀 때마다 정답과 기타 summary를 알려주면, 유저가 스스로 자신의 bias를 확인하면서 문제를 풀 수 있으므로 정확도를 높일 수 있을 것.

결과를 Report하는 방법이 중요. 단순히 correct, wrong을 알려주는 것은 지루하고 의미가 없음.

Learning curve: 윈도우 사이즈를 정해서 시간이 지날수록 정확도의 변화가 있는 지 보고싶다. 만약 유의미한 결과가 있다면, 적정 질문 개수는 몇 개인가?

competition with other players: 가장 높은 정확도를 보인 사람이나 점수 등을 함께 보여줘서 gamify 하는 것.

Visualization: Underestimation: blue, Overestimation: red, Correct: green

**Step2.**

Crowdsourced annotation application

유저가 일정 수준의 정확도를 갖게 되었다면, 그 다음 단계로는, 정확한 칼로리를 모르는 상태에서도 (정답을 모르는 상황에서도) 칼로리를 정확하게 예측할 수 있을까?

유저가 자신이 먹을 음식을 사진으로 촬영하여 업로드하면, 사람들의 estimation을 보여주게 하면 좋겠다. 단, 다른 사람들의 도움을 받기 위해서는 자신도 다른 사람들의 사진에 대하여 estimation을 하게 만들자.

**171229**

Email로 프로젝트 자료(구글 드라이브, 드롭박스 Repository) 공유 받음.

Email 전문

|  |
| --- |
| Jinwoo,  it's probably a good idea to read a few of related papers. there's several "fields" of related papers too:  -- nutrition literature and nutrition/diet apps (e.g., myplate)  -- vision/machine-learning (e.g., Im2Cal)  -- wisdom of the crowd  here's a link to a google drive folder with several documents by Jewell Finder (the undergraduate student who's working on the project); take a look especially at "works cited" and "Jun-paper-draft":  https://drive.google.com/drive/folders/0B\_tggtWTnxQ9aDU2SktIbjB0YVU?usp=sharing  i've also shared a dropbox folder with more info. take a look at the "related work" folder there.  finally, please, keep track of what papers you're reading (especially if you find new ones).  s. |

**180102**

Health Insurance – $313.54

IFS office check in

CatCard 발급 - $25

CS department account 발급

Office key 발급 – 내일

**180103**

CS account 오류 정정

Id: jinwoocho

Office key 발급

**180104**

SI/SISTA systems, zephyr account 발급

tunnel1.sista.arizona.edu or tunnel2.sista.arizona.edu

Id: jinwoocho

**180108 (Mon)**

**180116 (Tues)**

이번주에 할 것들

Android first demo 완성하기

포함할 feature들

Demographic page -> 마지막 page로

Server to app: images, ground truth

Zephyr server에서 php 파일로 app으로부터 전달받은 filename에 해당하는 image를 DB에서 읽어오는 부분.

App to server: demographic, estimates, IP address(?), time duration per quiz

Android에서 PHP & MySQL 연동하는 방법.

* Image와 ground truth 읽기
* Filename 읽기

Image number를 사전에 할당?

* DB에 있는 unique한 image filename을 읽고, permutation으로 shuffle하여 quiz 문제 할당.

**180117 (Wed)**

Submit button 눌렀을 때 accuracy 측정하는 method 추가하기.

Submit button 눌렀을 때 결과 보여주기.

~~Image 읽어서 View에 보여주는 것 추가하기.~~

현재 몇 번째 문제인지 보여주기.

이미지를 제대로 불러오지 못했을 경우를 대비하여, Refresh 버튼 추가하기.

20문제 이상 풀었을 때, Skip 또는 End 버튼 추가하기.

SeekBar 크기 조절하기.

한 문제 풀 때 마다 시간 측정.

IP address 얻기.

최종결과 보여주기.

DB에 저장되어 있는 image와 ground truth 정보 PHP&MySQL로 가져와서 txt 파일로 저장.

Python으로 이미지 다운로드하는 과정에서 307.jpg 이미지 에러 발생.

포맷이 jpg가 아니거나 뭔가 문제가 있던 것으로 생각되어,

http://zephyr.sista.arizona.edu/learn2cal/images/307

링크 접속하여 수동으로 다운로드 함.

list.txt 파일을 만들어서, 각 라인마다 이미지 파일 이름 + ‘,’ + 칼로리 정보를 저장.

Ex)

1.jpg,467

2.jpg,907

4.jpg,97

6.jpg,396

10.jpg,348

12.jpg,536

14.jpg,360

25.jpg,160

Android project에 raw폴더 생성하여 list.txt 파일 저장하고, ImageInfo class constructor에서 txt 파일을 읽고, 파싱하여 이미지 파일 이름과 칼로리를 각각의 ArrayList에 저장.

**180125 (Thur)**

|  |
| --- |
| 1) images of food; here are some of my favorites:  <https://www2.cs.arizona.edu/~kobourov/twitter_UA.pdf> (see pictures in the slides)  <http://www.washingtonpost.com/blogs/local/wp/2014/09/29/our-favorite-foods-to-tweet-about-state-by-state/> <http://www.slate.com/blogs/future_tense/2014/10/02/university_of_arizona_researchers_analyze_food_references_in_twitter_data.html> <http://greatideas.people.com/2014/09/22/state-food-study-twitter-hashtags/>  <http://www.chron.com/life/food/article/most-tweeted-foods-by-state-5788589.php>  2) you can break the text on the first page into two parts:  -- what it is  -- rules of the game (optional)  3) summarize the two apps that you found (what they do, how they do it, etc.) with 5-6 sentences each  4) color code (red/green/blue)  5) after one game give the option to go play again or see more results  -- in both cases take the user to demographics page where they can give info and register  -- explain this is optional and data is used anonymously  -- if a user is already registered they don't see that page  -- after more results are shown, the user can have the option to play again  6) we need to keep track of registered users  7) perhaps you should copy the database and modify it so it does what you need:  -- keep track of the changes you make  -- there were other entries that we also need in the schema (e.g., window size)  8) help Jewell if you can  s. |

**180126 (Fri)**

CBP

* ~~Feature map 뽑기~~
* 기존 그래프 수정

**180127 (Sat)**

2) you can break the text on the first page into two parts:

-- what it is

-- rules of the game (optional)

3) summarize the two apps that you found (what they do, how they do it, etc.) with 5-6 sentences each.

4) color code (red/green/blue).

**180128 (Sun)**

5) after one game give the option to **go play again** or **see more results**

-- in both cases take the user to demographics page where they can give info and register

-- explain this is optional and data is used anonymously

-- if a user is already registered they don't see that page

-- after more results are shown, the user can have the option to play again

6) we need to keep track of registered users

[Play again], [See more results], [Quit] 버튼 추가

각 버튼 클릭 시, demographic 등록되어 있는 지 확인하고

등록이 안되어 있다면, Demo page 보여주기

등록이 되어 있다면 바로 [Quiz], [Result] 페이지로 넘어가기.

**[Final.class]**

onStop()에서 결과 전송하기.

**[See more results]**

Demographic page

* 왜 수집하는 것인지, 익명성 보장 된다는 것 명시하는 문구 삽입
* 데이터 저장(SharedPreferences)
* 모두 입력 했는지 확인
* UUID 추가하기

Result page

* 결과 전송
* 이전 결과 불러오기
* 지금까지의 결과 보여주기

**[Quit]**

UUID 생성

결과 전송

**[Play Again]**

게임 재시작

PHP&MySQL

MySQL Attribute 추가하기

PHP

**180129 (Mon)**

Final.class - UUID 추가

PHP 스터디

**180130 (Tue)**

Jewell이 말한 addUser.php 에서 user 정보 update가 되지 않는 문제 해결

PHP 스터디

**180131 (Wed)**

Android – MySQL 연결 작업

Insert user, Update user, Insert estimate

미팅 준비

**180201(Thur)**

|  |
| --- |
| **APP:**  **-- email to Jewell and Dane (.apk)**  **-- what happens if you're offline?**  **-- images and icon**  **-- translated related apps**  **-- immediate feedback (histogram, density)**  QUIZ:  -1) Jewell, look at the paper and the data they sent us (including spreadsheet)  -- for each image, where did it come from?  -- what databases for cals per gram?  -- what databases for unit size?  -- give an example for the 2 corn images?  -- DRAFT EMAIL ASAP  **0) get the pilot ready to go again**  **-- resolve the back-button issue?**  **-- SANITY CHECK (email any suspicious items)**  **1) database updates:**  **-- compare what Jinwoo has added**  **-- definitely record screen size to determine whether computer or phone are use (important for data analysis, e.g., errors, timing, etc.)**  **-- record where the data came from: quiz or app**  2) learning curve analysis?  3) if detecting phone we can offer them the ap  4) computer vs phone version:  -- mobile aware software design: how hard is it to do?  5) better and more complete performance analysis (like Jun's last page on the first quiz)  -- SHOULD WE CHANGE FROM ACCURACY TO PERCENT ERROR?  -- minimal first summary page: learning curve (yours and average), score (many options)  -- offer more details (record how many select this)  -- best 5, worst 5 (possibly with info when the occurred, possibly how hard they are for others) |

이제 기능적인 구현은 대부분 마무리 되었기 때문에, 완성도에 집중

**180202(Fri)**

|  |
| --- |
| **APP:**  **-- email to Jewell and Dane (.apk)**  **-- what happens if you're offline?**  **-- images and**  **-- immediate feedback (histogram, density)**  **0) get the pilot ready to go again**  **-- resolve the back-button issue?**  **-- SANITY CHECK (email any suspicious items)**  5) better and more complete performance analysis (like Jun's last page on the first quiz)  -- SHOULD WE CHANGE FROM ACCURACY TO **PERCENT ERROR**?  -- minimal first summary page: learning curve (yours and average), score (many options)  -- offer more details (record how many select this)  -- best 5, worst 5 (possibly with info when the occurred, possibly how hard they are for others) |

180209 (Fri)

|  |
| --- |
| APP:  ~~-- email to Jewell and Dane~~  ~~-- make 60 guesses by default~~  ~~-- reset guess to 0 calories each time~~  ~~-- i think it's good to see the image with the errors~~  ~~-- use 5 fixed width bars and past ones go to the left~~  -- y-axis should percent error  -- number in the bar is number of calories off  -- what happens if you're offline  -- images  ~~-- translated related apps~~  Final page:  -- overall graphic (accuracy, current-user-absolute-error-in-calories, current-user-absolute-percent-error, overall-users-absoluty-average-percent-error-histogram-with-one-bar-highlighted)  -- learning: 3 numbers: absolute percentage error for groups of 20  -- learning score: compare learning improvement to  -- learning: the curve: sliding window and average (see picture)  -- to show bias: show red/green/blue bars  -- wish-list: classify participants into groups and "profile" each new one (MDS,k-center-clustering,t-SNE)  QUIZ:  0) get the pilot ready to go again  -- resolve the back-button issue  -- Dane email's us how to "verify page from server"  -- warn at the start  1) database updates:  -- compare what Jinwoo has added  -- definitely record screen size to determine whether computer or phone are use (important for data analysis, e.g., errors, timing, etc.)  -- Jinwoo emails schema to Jewell and Dane  2) pilot again  -- Jewell will email us all when pilot test is ready  -- test is ready when it doesn't crash, questionable images are removed, basic final page, recording additional data from users (like Jinwoo)  -- we need to test the quiz on phones: action item for Jinwoo and Dane  3) better and more complete performance analysis (like Jun's last page on the first quiz)  -- SHOULD WE CHANGE FROM ACCURACY TO PERCENT ERROR?  -- minimal first summary page: learning curve (yours and average), score (many options)  -- offer more details (record how many select this)  -- best 5, worst 5 (possibly with info when the occurred, possibly how hard they are for others) |