CMPS 299

American University of Beirut

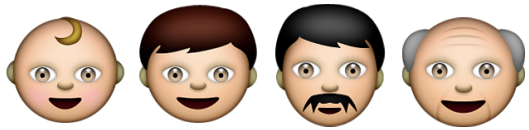
Lemira Hala Chehab
Nisreen Mansouri
Vanessa Boghos

December 18, 2017

Overview: Our Aim

To see if there is a correlation between phone activity and a persons traits:

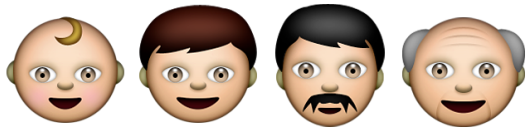
1. Age



Overview: Our Aim

To see if there is a correlation between phone activity and a persons traits:

1. Age

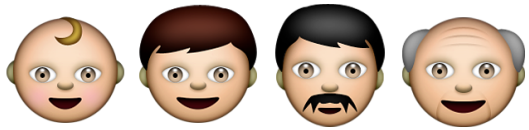


2. Gender

Overview: Our Aim

To see if there is a correlation between phone activity and a persons traits:

1. Age



2. Gender

3. Mood

Main Features

1. Monitor users activity

Main Features

1. Monitor users activity
2. Collect data from many users

Main Features

1. Monitor users activity
2. Collect data from many users
3. Apply machine learning techniques

Technologies Used

1. **Android Studio:** SDK used to develop the App

Technologies Used

1. **Android Studio:** SDK used to develop the App

2. **Google Firebase:** Real time database for the App

Technologies Used

1. **Android Studio:** SDK used to develop the App

2. **Google Firebase:** Real time database for the App

3. **Java SDK:** Used for features extraction and neural network implementation

Logging

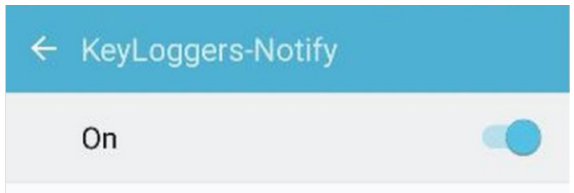
1. **Frontend**
    1. User fills in his data in app
    2. Android service monitors persons activity
    3. Activity includes apps entered, scrolls, taps, typing

Logging

1. **Frontend**
   1. User fills in his data in app
   2. Android service monitors persons activity
   3. Activity includes apps entered, scrolls, taps, typing
2. **Backend**
   1. Logs are sent to Firebase Database in real time
   2. Log contains: context, date, type

```
"context" : "[Facebook]",
"dateAndTime" : "1510588279473",
"type" : "CLICKED"
```


1% 7:31 PM
KeyLoggers-Notify
First Name:
Last Name:
Age:
Gender:
Mood:
Comment:
ADD SESSION


KeyLoggers-Notify
On

Data Collection

- Collected data from *203* users
- Users were asked to use the phone for *3 minutes*
- They would use the phone as if it were their own
- Users would go on various apps: facebook, fruit ninja, etc.

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games
16. How many times the user used the camera

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games
16. How many times the user used the camera
17. How many times accessed the Phone

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games
16. How many times the user used the camera
17. How many times accessed the Phone
18. Numbers of calls

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games
16. How many times the user used the camera
17. How many times accessed the Phone
18. Numbers of calls
19. Word count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games
16. How many times the user used the camera
17. How many times accessed the Phone
18. Numbers of calls
19. Word count
20. Search count

Features Extraction and Data Filtering

We decided to extract the 21 features:

1. Click count
2. Long click count
3. Scrolls count
4. Text count
5. Focused count
6. How many times the window was changed
7. Total log count
8. Time spent in Facebook
9. Time spent in Whatsapp
10. Time spent in Instagram
11. Time spent in Camera
12. Time spent in Gallery
13. Time spent in Email
14. Time spent in Youtube
15. Time spent in Games
16. How many times the user used the camera
17. How many times accessed the Phone
18. Numbers of calls
19. Word count
20. Search count
21. Number of videos watched on youtube

Filtered Data Example

Filtered logs

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
9.0	0.0	256.0	0.0	2.0	22.0	289.0	349232.0	0.0	68374.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0

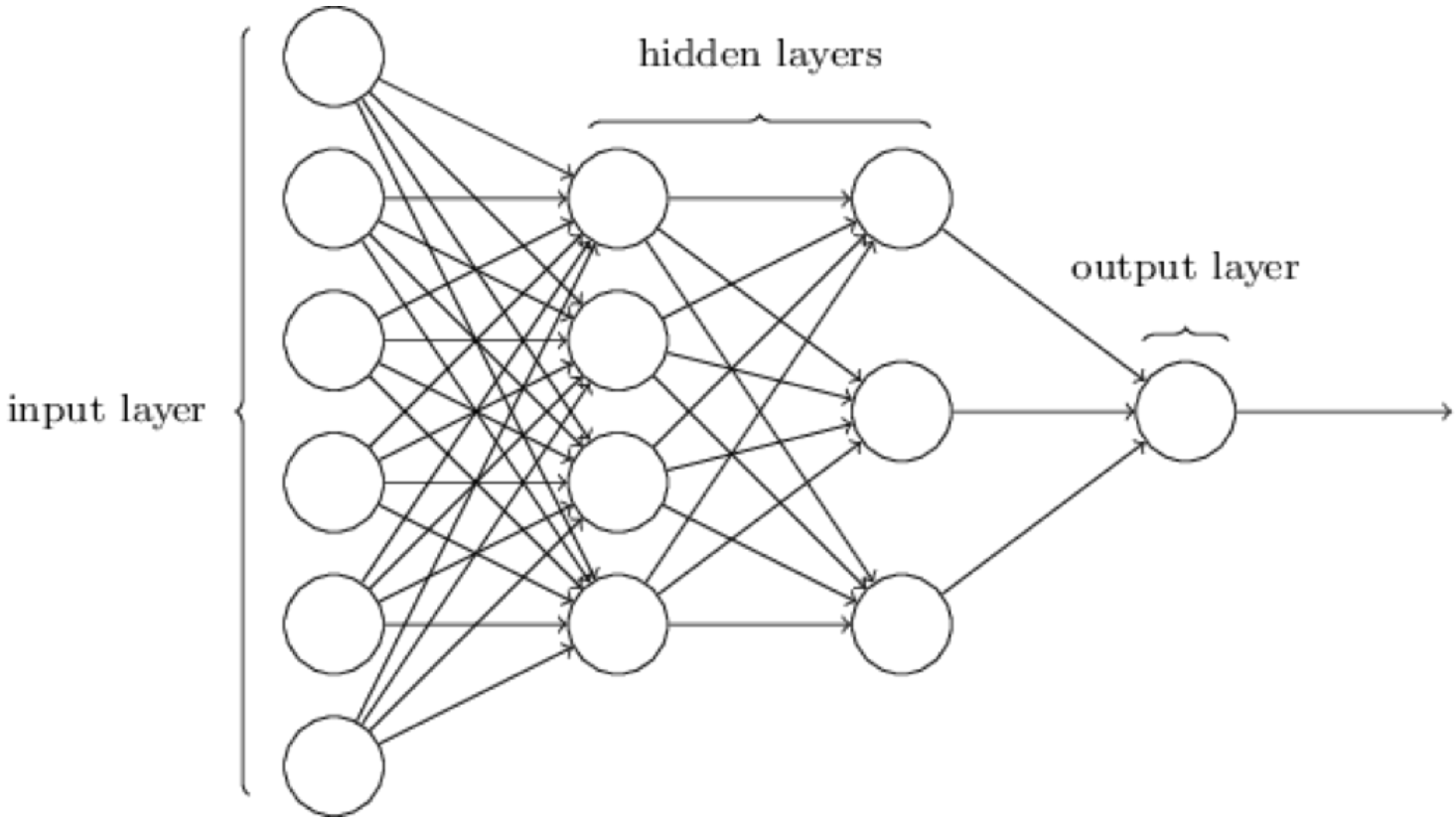
Filtered User Info

Age	*Gender*	*Mood*
20.0	1.0	0.0

Gender:
0: Female
1: Male

Mood:
0: Neutral
1: Happy
2: Sad
3: Tired
4: Angry

Figure: Implemented Neural Network



Machine Learning

1. **Input vector:** Contains all the features extracted

Machine Learning

1. **Input vector:** Contains all the features extracted
2. **Output vector:** age, mood, gender

Machine Learning

1. **Input vector:** Contains all the features extracted
2. **Output vector:** age, mood, gender

1. **Training data on different:**

Machine Learning

1. **Input vector:** Contains all the features extracted
2. **Output vector:** age, mood, gender

1. **Training data on different:**
    1. **Number of hidden units**

Machine Learning

1. **Input vector:** Contains all the features extracted
2. **Output vector:** age, mood, gender

1. **Training data on different:**
    1. **Number of hidden units**
    2. **Epoch**

Machine Learning

1. **Input vector:** Contains all the features extracted
2. **Output vector:** age, mood, gender

1. **Training data on different:**
    1. **Number of hidden units**
    2. **Epoch**
    3. **Activation Functions (Sigmoid, Linear, Softmax)**

Code for ML

Figure: Sample Java code

```
for (int units = 1; units < 50; units++) {
    for( int i = 0; i < epoch.length; i++) {
        //train
        TrainingNN nn = new TrainingNN(21, 1, units, DefaultSettings.activationFunctionOuterLinear);
        nn.train(trainingInput, trainingOutputAge, epoch[i], 0.000001);

        //test
        for( int j = 0; j < testingInput.length; j++)
            nnOutputs[j] = nn.forwardPrograpagation(testingInput[j]);

        double mse = getmse(nnOutputs, testingOutputAge);

        if(mse < min_mse) {
            min_mse = mse;
            min_nn = nn;
            nnBestOutput = nnOutputs;
        }
    }
}

double accuracy = calculateAccuracy(testingOutputAge, nnBestOutput);
```

Accuracy

1. **Training:**
    1. Calculate the mean squared error (MSE)
    2. Using minimum MSE, we found the optimal Epoch, function, and number of hidden units

Accuracy

1. **Training:**
   1. Calculate the mean squared error (MSE)
   2. Using minimum MSE, we found the optimal Epoch, function, and number of hidden units
2. **Testing:**
   1. Determine Accuracy (expected data vs actual data)

Findings: Gender

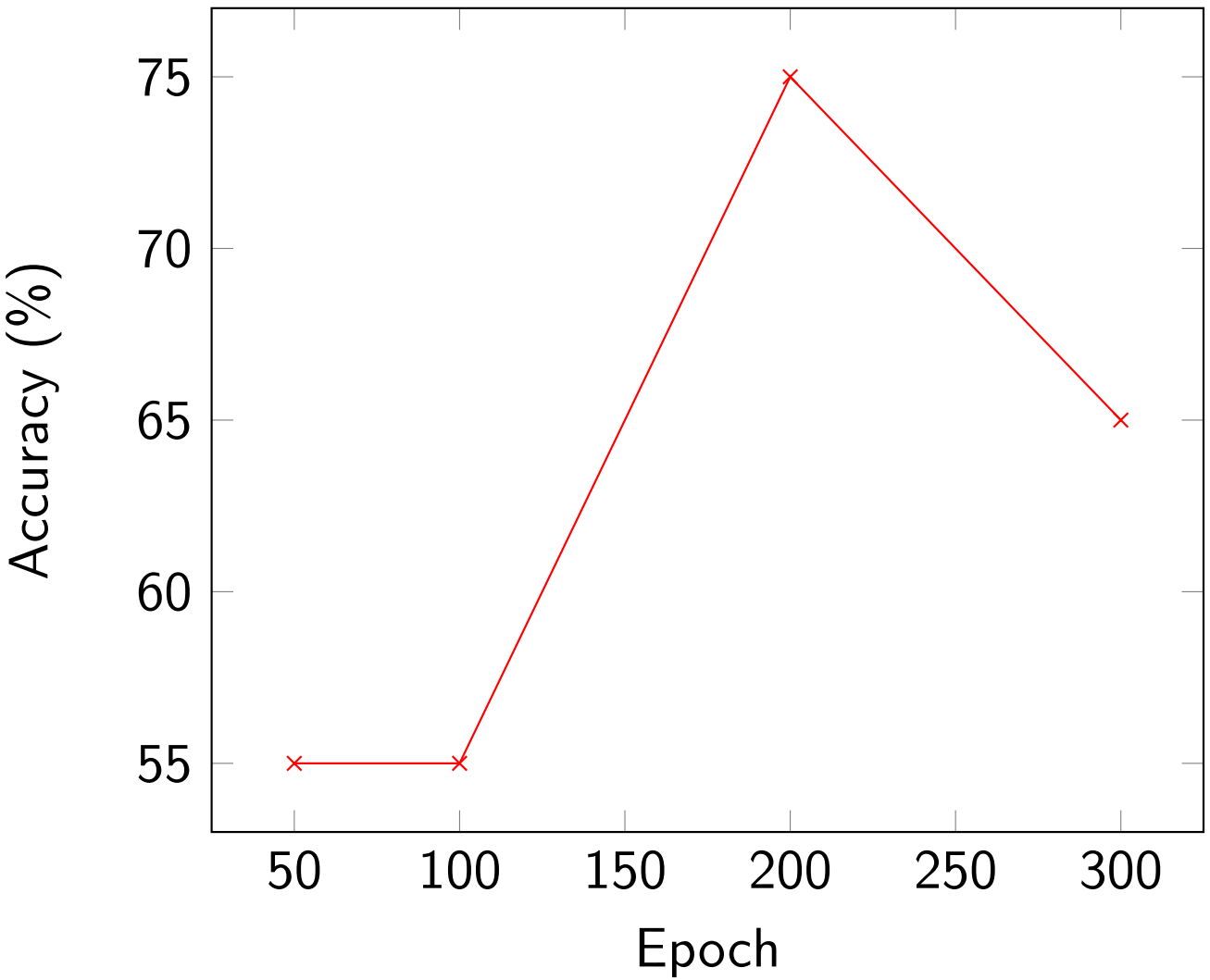
Table: Training on 90% of the data - Linear function

Epoch	Accuracy in %	Hidden Neurons
100	35	19
500	45	48
800	55	89
1000	65	48
1200	60	64
2000	55	35

Table: Training on 90% of the data - Sigmoid function

Epoch	Accuracy in %	Hidden Neurons
50	55	26
100	55	90
200	75	78
300	65	84

Findings: Gender Graph example



Findings: Age

Table: Training on 90% of the data

Epoch	Accuracy in %	Hidden Neurons
70	55	28
80	60	45
90	60	27
100	55	45

Findings: Mood

Table: Training on 90% of the data

Epoch	Accuracy in %	Hidden Neurons
10	35	26
25	25	43
50	20	11
100	25	18
200	20	49

Q&A