



USER ROUTINE PREDICTIVE ANALYSIS

BIG DATA ANALYTICS AND APPLICATIONS PROJECT



SPRING 2016

SUBMITTED BY

RAKESH REDDY BANDI (02)

NIHAR DUDAM (06)

SRICHARAN PAMURU (20)

MARK SCHULTZ (26)

I. Introduction:

Big data has become a buzzword in the field of Computer Science. As technology advances, we are finding new ways to collect and analyze data. With the invention of the smart watch, we have been able to expand this collection of data to areas that had not been possible years ago. In addition, it has never been easier for the user to retrieve data from their device when needed. Our project combines these two conveniences into one application. We will use big data to measure and map a User's routine, and use this routine to make predictions and suggestions about the future.

We plan on applying several sources for data collection. First of all, the applications collect data from movements of the watch to classify patterns in certain movements. It will use these movements to guess the activity that the User may be doing at the moment. It will also collect data from the heart rate monitor of the smartwatch to measure the rigor of the User's activity. GPS data on the watch will determine the location of the user. We will also use datetime metrics to determine the day and the time that the user typically does certain activities.

The sources of data above will be collected by the user, but we also want to combine this with several API's to make better predictions, and add suggestions to give back to the user. We will use speech to text API to query the application for information. We will use the Google Location API to develop a richer context of the user's most frequented locations, and suggest possible locations that might interest the user. We will implement the Yelp API, so that when the application predicts when the user eats, it will make suggestions on places to go.

All of the data we will collect will be organized using the Hadoop file system and implemented using Amazon Web Services. We will discuss the technologies more in this documentation. The project is broken down into four phases. For the first Increment, we began the data collection to measure several common movements. We are initially using Naïve Bayes to classify these movements. We have also began research on several of the API's that we plan on implementing.

II. Project Goal and Objectives:

The main implementation of the project will consist of the following specification based objective and subjective analysis along with a sketched perspective of goal:

1. Overall Goal:

The goal of the project is to predict a user's activity at any time during a user's schedule. A large dataset is provided by the smartwatch and the machine is trained on the basis of machine learning algorithms as to how the coordinates will affect the motion of the sensor. Machine is trained and the provided data is analyzed based on the reading from the various sensors of the smartwatch. The location information will provide a set of variable features such as restaurants, parks, etc. near to the naive user. APIs will provide the analysis of information and data regarding the review and further circumspect options are brought forward to help the smartwatch analyze the data and provide a predictive routine. A variable set of sensors will collect the proximity, heart rate, accelerometer based motion readings, etc. to analyze data set. The analysis of this large datasets will return recommended notifications to the user for a viable routine.

2. Specific objectives:

- Predict user routine by analysis of large datasets through machine learning techniques.
- Return geolocation based places in proximity such as restaurants, parks, etc.
- Sensor dataset provides specific activity based readings.
- Provide recommended services by sending notification to user depending on time and routine.

3. Specific features:

Sensors in the watch serve as primary source for providing data which is required for analysis. The accelerometer will provide a three dimensional reading with regard to the position. The heart rate sensor will provide the heart beat rate which can be used to obtain the magnitude of force exerted while performing a particular activity. GPS will provide information about the location and the APIS will return data to the user depending upon the routine's requirements and suggestive

notifications will be sent to the smartwatch while the connection to another remote device or machine such as a smartphone holds true. It will create a complicate alliance and thus the machine is to be trained using various machine learning algorithms. Data clusters are classified using the classification techniques intended to provide more information about the routine prediction. There are other sensors in the watch such as ambient light sensor, gyroscope, magnetic field sensor, step counter, gravity, linear acceleration, significant motion, etc. which are used collaboratively to obtain a dataset.

4. Significance:

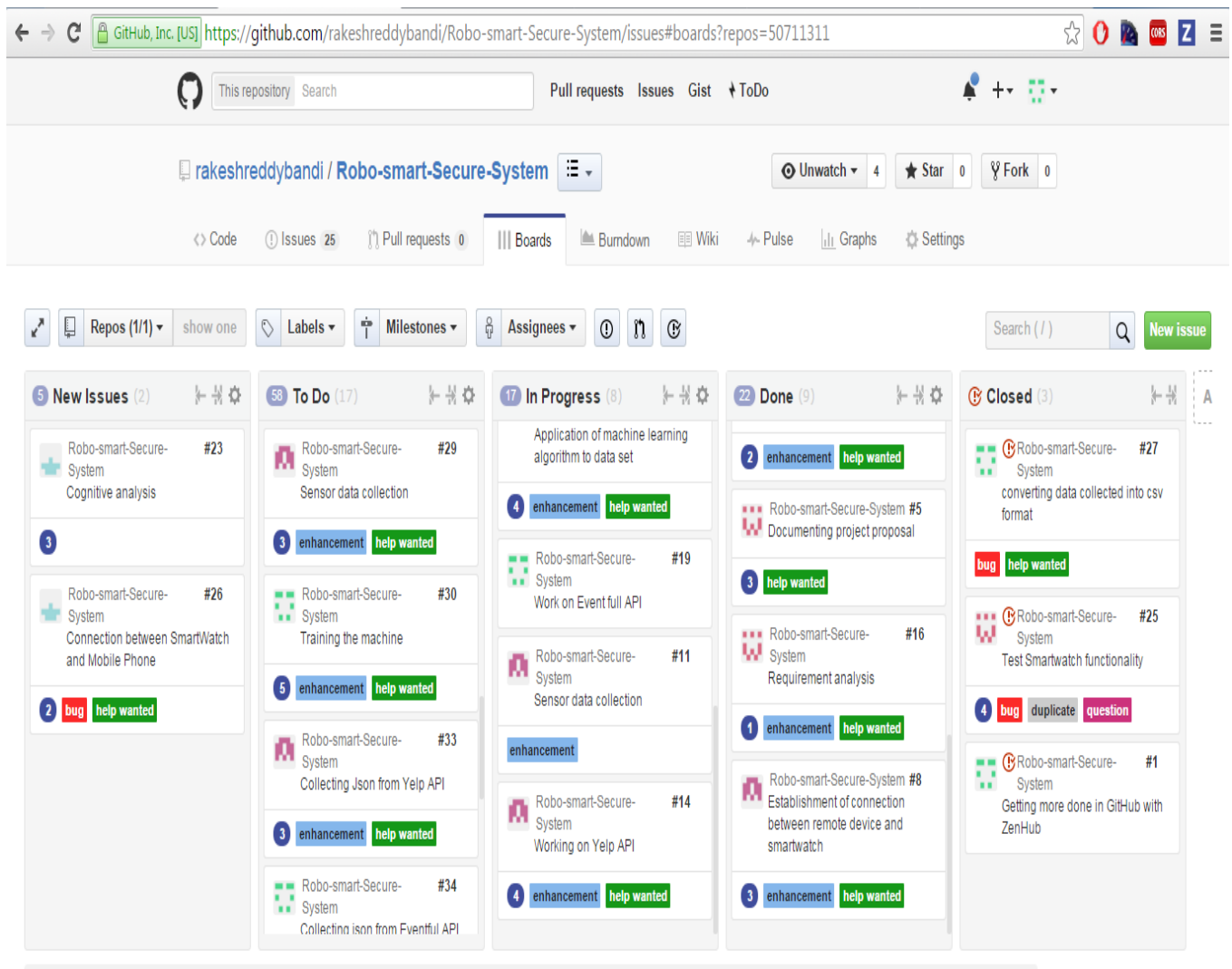
The most important aspect of the project is that an analysis of big data is done in order to predict the routine of a specific user. It will generate a pattern based on the different readings to generate a more sophisticated clustered activity set. This will help in obtaining a clear cut view of a user's routine plus a more generative dimension to a time schedule. Sensors present in the smartwatch will be crucial in collection of data, which in turn is fed into the machine through various machine learning techniques. A stable schedule would prompt the user to receive a collaborative approach to obtain the output as a predictive routine analysis.

III. Project Plan:

1. Schedule for entire project:

The entire schedule for the project is shown as follows where the respective use case and process step up with all the modules and the time frame is categorized specifically to illustrate the step by step fulfilment of the entire processes done such as feature design and feature implementation.

Screenshot 1: The following screenshot demonstrates the overall perspective of implementation in which the project is being done. The To do and In Progress have a significantly high number of modules and are continued in the following images.



Screenshot 2:

The following screenshots shows the second listing of the To Do phases and In Progress where there is an increase in the number of elements and tasks to be done. There is an incentive in populating the respective tasks and assign them to the members as per core requirement and generative design.

The screenshot displays the GitHub interface for the repository 'rakeshreddybandi / Robo-smart-Secure-System'. The URL in the browser is <https://github.com/rakeshreddybandi/Robo-smart-Secure-System/issues#boards?repos=50711311>. The repository has 4 Unwatch, 0 Stars, and 0 Forks. The navigation bar includes links for Code, Issues (25), Pull requests (0), Boards (selected), Burndown, Wiki, Pulse, Graphs, and Settings. The Kanban board is organized into five columns: New Issues (2), To Do (17), In Progress (8), Done (9), and Closed (3). Each column contains a list of issues with their titles, numbers, and labels.

Column	Issue Number	Issue Title	Labels
New Issues (2)	#23	Robo-smart-Secure-System Cognitive analysis	
	#26	Robo-smart-Secure-System Connection between SmartWatch and Mobile Phone	bug, help wanted
To Do (17)	#31	Robo-smart-Secure-System Predicting the user activities	enhancement, help wanted
	#32	Robo-smart-Secure-System Adding SmartWatch data as training data	enhancement, help wanted
	#24	Robo-smart-Secure-System Notification to Smart Watch	enhancement, help wanted
	#37	Robo-smart-Secure-System	
In Progress (8)	#17	Robo-smart-Secure-System Work on Geo Location API	enhancement
	#12	Robo-smart-Secure-System Application of machine learning algorithm to data set	enhancement, help wanted
	#19	Robo-smart-Secure-System Work on Event full API	
	#11	Robo-smart-Secure-System Sensor data collection	
Done (9)	#10	Robo-smart-Secure-System Research on APIs	enhancement, help wanted, question
	#6	Robo-smart-Secure-System Refine and rebuild project proposal	enhancement
	#22	Robo-smart-Secure-System Modular Description	enhancement, help wanted
	#5	Robo-smart-Secure-System Documentation project proposal	
Closed (3)	#27	Robo-smart-Secure-System converting data collected into csv format	bug, help wanted
	#25	Robo-smart-Secure-System Test Smartwatch functionality	bug, duplicate, question
	#1	Robo-smart-Secure-System Getting more done in GitHub with ZenHub	

Screenshot 3:

The following screenshot shows the phase where the tasks in progress have been concentrated mainly on the design specifications. It combines along with the done tasks to stabilize the development of project in due course to simultaneously bring out the working modules.

The screenshot displays the GitHub interface for the repository 'rakeshreddybandi / Robo-smart-Secure-System'. The URL in the browser is <https://github.com/rakeshreddybandi/Robo-smart-Secure-System/issues#boards?repos=50711311>. The repository has 4 Unwatch, 0 Stars, and 0 Forks. The main navigation bar includes links for Code, Issues (25), Pull requests (0), Boards, Burndown, Wiki, Pulse, Graphs, and Settings. The 'Boards' tab is active, showing a Kanban board with the following columns:

- New Issues**: 3 issues. The first issue is 'Robo-smart-Secure-System #23 Cognitive analysis' with a 'help wanted' label. The second issue is 'Robo-smart-Secure-System #26 Connection between SmartWatch and Mobile Phone' with 'bug' and 'help wanted' labels.
- To Do**: 17 issues. The first issue is 'Robo-smart-Secure-System #13 Testing' with 'enhancement' and 'help wanted' labels. The second issue is 'Robo-smart-Secure-System #15 Maintenance' with 'bug', 'enhancement', and 'help wanted' labels. The third issue is 'Robo-smart-Secure-System #18 Debugging' with 'bug', 'help wanted', and 'question' labels. The fourth issue is 'Robo-smart-Secure-System #31'.
- In Progress**: 8 issues. The first issue is 'Robo-smart-Secure-System #4 Analyzing the machine learning algorithms' with a 'help wanted' label. The second issue is 'Robo-smart-Secure-System #20 Architecture design' with 'enhancement' and 'help wanted' labels. The third issue is 'Robo-smart-Secure-System #21 Working on Weather API' with 'enhancement' and 'help wanted' labels. The fourth issue is 'Robo-smart-Secure-System #17 Work on Geo Location API'.
- Done**: 9 issues. The first issue is 'Robo-smart-Secure-System #2 Project Proposal Discussion' with an 'enhancement' label. The second issue is 'Robo-smart-Secure-System #7 Knowledge discovery of smarwatch' with 'enhancement' and 'question' labels. The third issue is 'Robo-smart-Secure-System #3 Finalizing project plan' with an 'enhancement' label. The fourth issue is 'Robo-smart-Secure-System #10 Research on APIs'.
- Closed**: 3 issues. The first issue is 'Robo-smart-Secure-System #27 converting data collected into csv format' with 'bug' and 'help wanted' labels. The second issue is 'Robo-smart-Secure-System #25 Test Smartwatch functionality' with 'bug', 'duplicate', and 'question' labels. The third issue is 'Robo-smart-Secure-System #1 Getting more done in GitHub with ZenHub'.

Screenshot 4:

The screenshot here shows the listing done where the penultimate requirement analysis and engineering is done to bring out the resulted output that is tested. The maintenance is to be done and especially debugging to negate results which are not in favour of project deployment.

The screenshot displays the GitHub interface for the repository 'rakeshreddybandi / Robo-smart-Secure-System'. The URL in the browser is <https://github.com/rakeshreddybandi/Robo-smart-Secure-System/issues#boards?repos=50711311>. The repository has 4 Unwatch, 0 Star, and 0 Fork. The navigation bar includes links for Code, Issues (25), Pull requests (0), Boards, Burndown, Wiki, Pulse, Graphs, and Settings. The main content area shows a Kanban board with five columns: New Issues (2), To Do (17), In Progress (8), Done (9), and Closed (3). Each column contains a list of issues with their titles, numbers, and labels like 'enhancement', 'bug', 'help wanted', 'duplicate', and 'question'.

Column	Count	Issue #	Title	Labels
New Issues (2)		#23	Robo-smart-Secure-System Cognitive analysis	
		#26	Robo-smart-Secure-System Connection between SmartWatch and Mobile Phone	bug, help wanted
To Do (17)	3		enhancement help wanted	
		#36	Robo-smart-Secure-System Refining all the features	
	8		enhancement help wanted	
		#38	Robo-smart-Secure-System Final Documentation	
In Progress (8)			Application of machine learning algorithm to data set	
	4		enhancement help wanted	
		#19	Robo-smart-Secure-System Work on Event full API	
Done (9)	2		enhancement help wanted	
		#5	Robo-smart-Secure-System Documenting project proposal	
	3		help wanted	
Closed (3)		#27	Robo-smart-Secure-System converting data collected into csv format	bug, help wanted
		#25	Robo-smart-Secure-System Test Smartwatch functionality	
	4		bug duplicate question	

Screenshot 5:

The entire project schedule has been roughly shown here. A detailed insight of project plan is done in such a manner that there might be few more additional development or phase build in due course of the project. It shows a correlation of the various modules included as per specification and features.

The screenshot displays the GitHub interface for the repository 'rakeshreddybandi / Robo-smart-Secure-System'. The top navigation bar includes links for Pull requests, Issues, Gist, and To Do. The repository name and statistics (Unwatch, Star, Fork) are visible. Below the repository name, there are tabs for Code, Issues (25), Pull requests (0), Boards, Burndown, Wiki, Pulse, Graphs, and Settings. The 'Boards' tab is active, showing a Kanban board with five columns: New Issues (2), To Do (17), In Progress (8), Done (9), and Closed (3). Each column contains a list of issues with their titles, numbers, and labels. The issues are categorized by labels such as 'enhancement', 'bug', 'help wanted', 'duplicate', and 'question'. The board provides a visual overview of the project's progress and upcoming tasks.

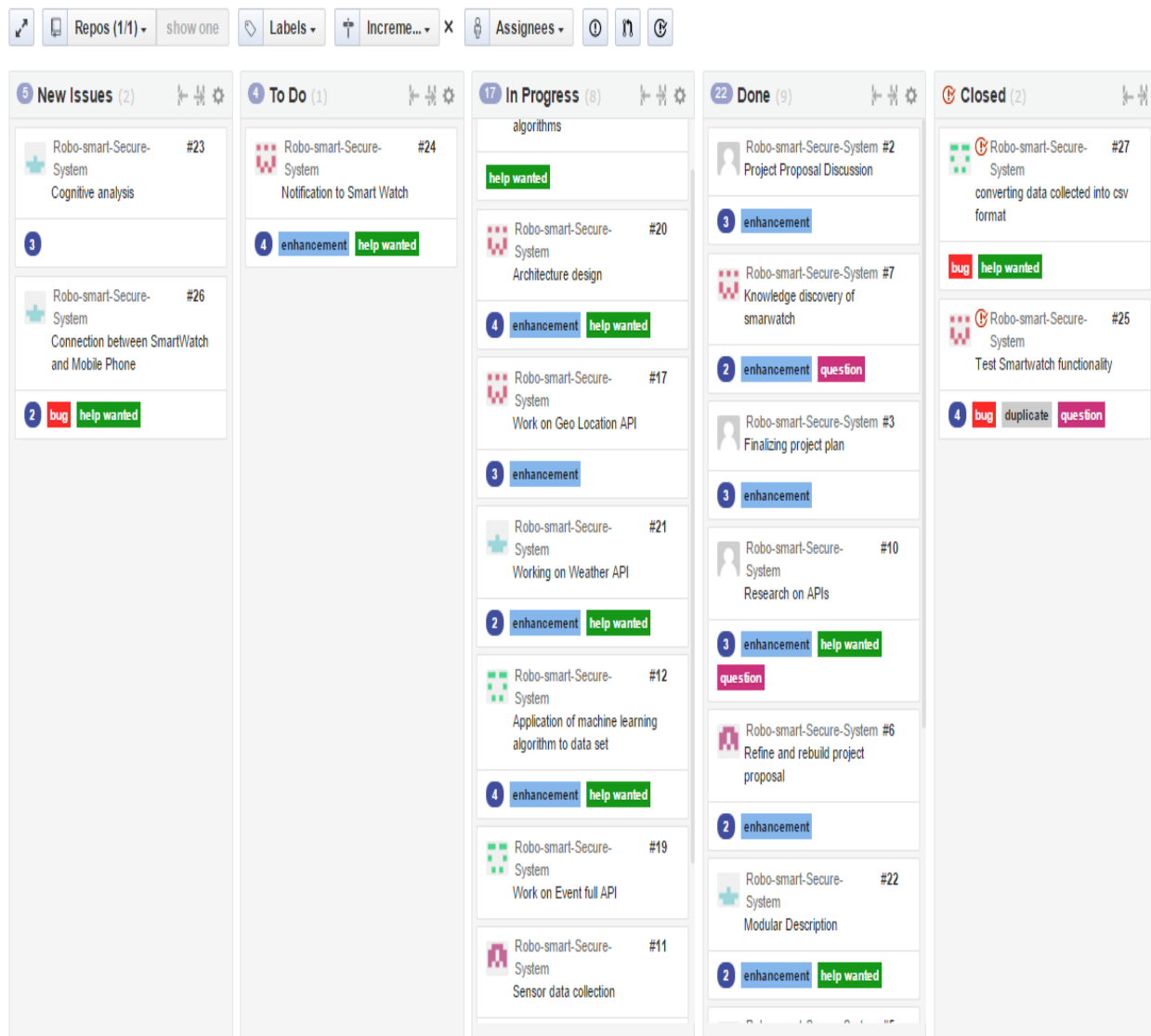
Column	Issue Number	Issue Title	Labels
New Issues (2)	#23	Robo-smart-Secure-System Cognitive analysis	
	#26	Robo-smart-Secure-System Connection between SmartWatch and Mobile Phone	bug, help wanted
To Do (17)	#34	Robo-smart-Secure-System Collecting json from Eventful API	enhancement
	#35	Robo-smart-Secure-System Analyze Location Data	enhancement, help wanted
	#36	Robo-smart-Secure-System Refining all the features	enhancement, help wanted
	#38	Robo-smart-Secure-System	
In Progress (8)		Application of machine learning algorithm to data set	enhancement, help wanted
	#19	Robo-smart-Secure-System Work on Event full API	
	#11	Robo-smart-Secure-System Sensor data collection	
	#14	Robo-smart-Secure-System Working on Yelp API	enhancement, help wanted
Done (9)	#2	Robo-smart-Secure-System	enhancement, help wanted
	#5	Robo-smart-Secure-System Documenting project proposal	help wanted
	#16	Robo-smart-Secure-System Requirement analysis	enhancement, help wanted
Closed (3)	#27	Robo-smart-Secure-System converting data collected into csv format	bug, help wanted
	#25	Robo-smart-Secure-System Test Smartwatch functionality	bug, duplicate, question

2. Project Timeline:

The implementation of the project is covered with a specific timeline and this is shown as follows to implement it using Zenhub:

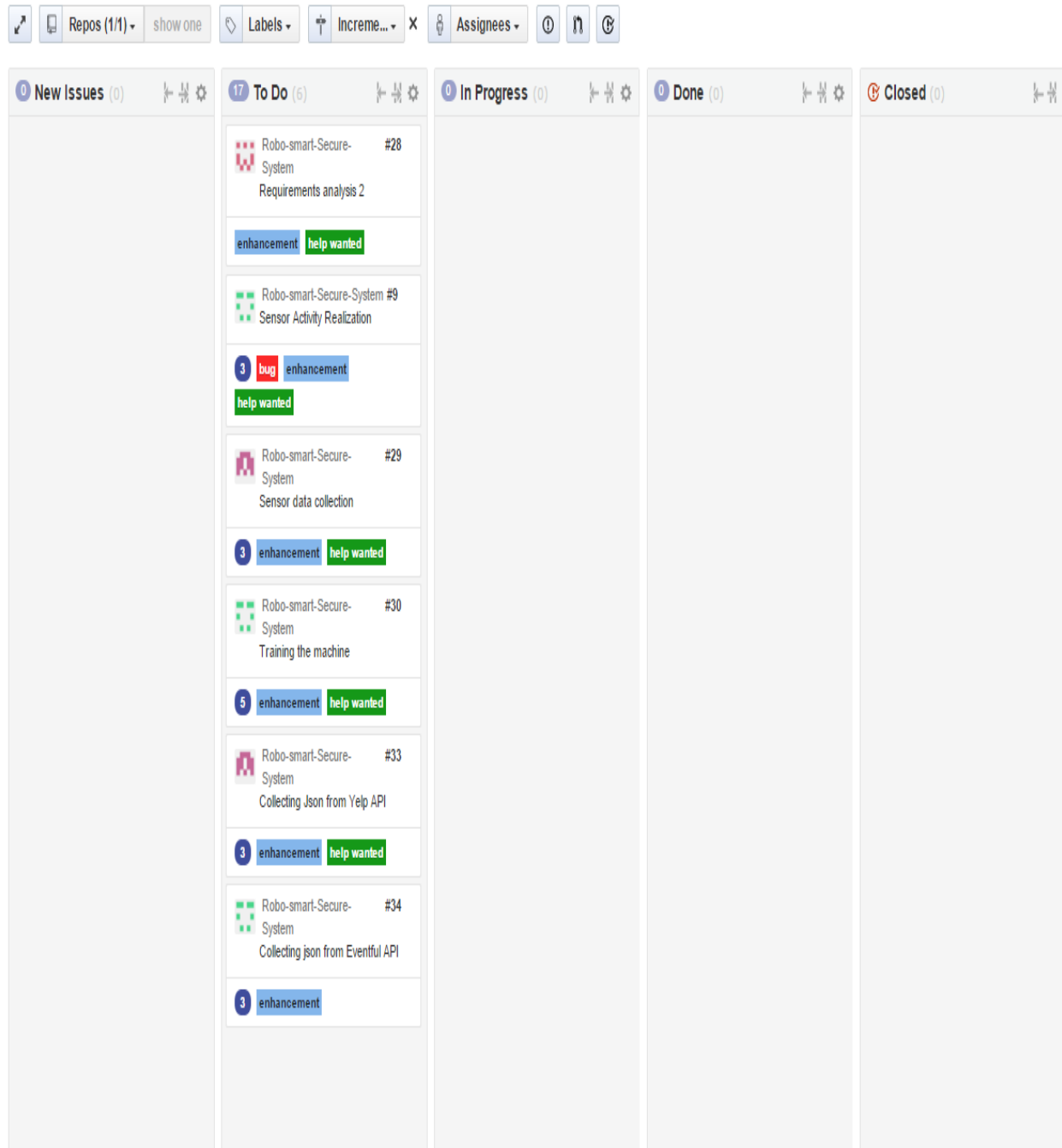
Increment 1:

The first increment has the following tasks to be completed by 02/19/2016 with tasks assigned as such:



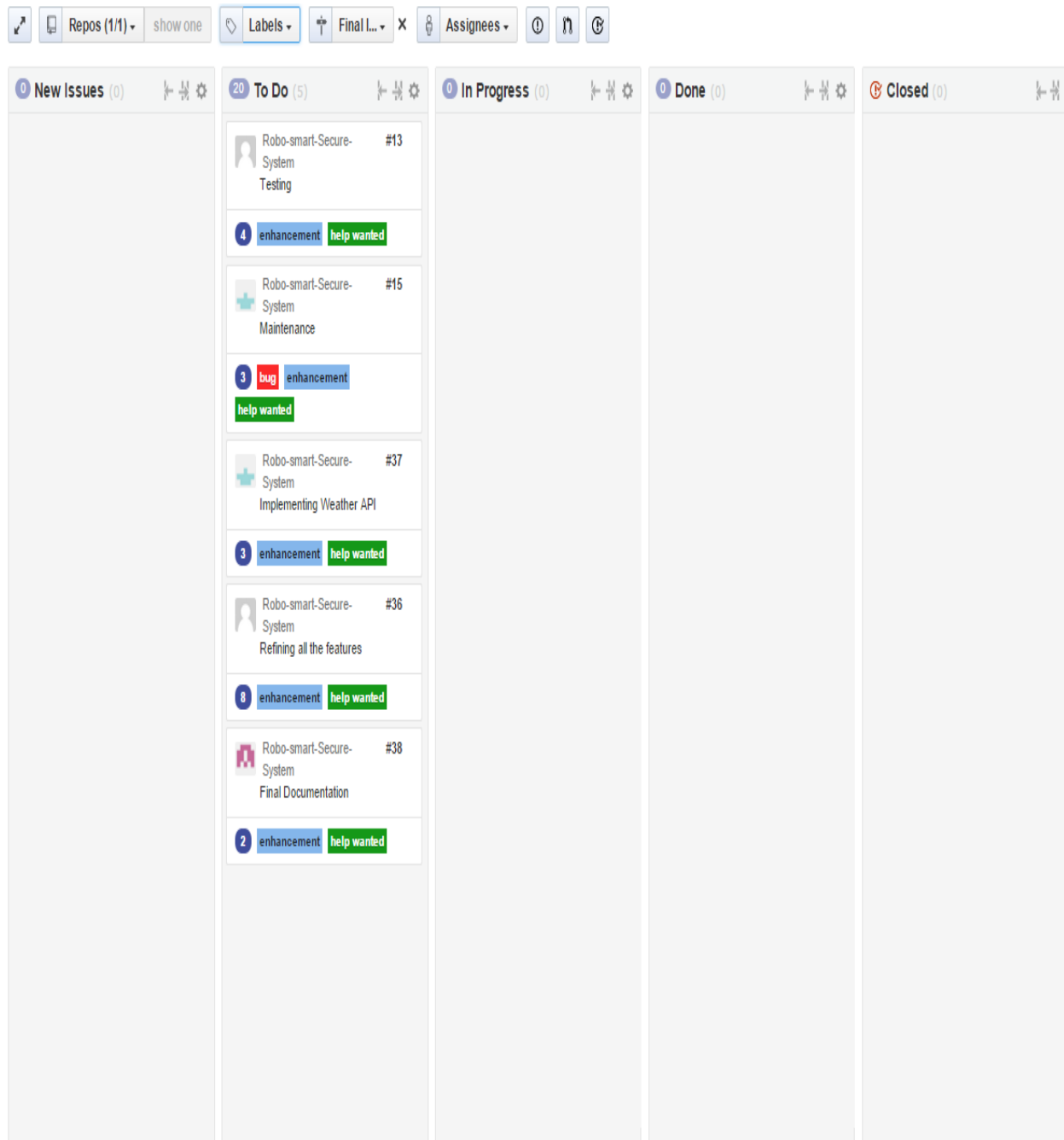
Increment 2:

The timeline for the second increment due on 03/11/2016 will have the graph as follows:



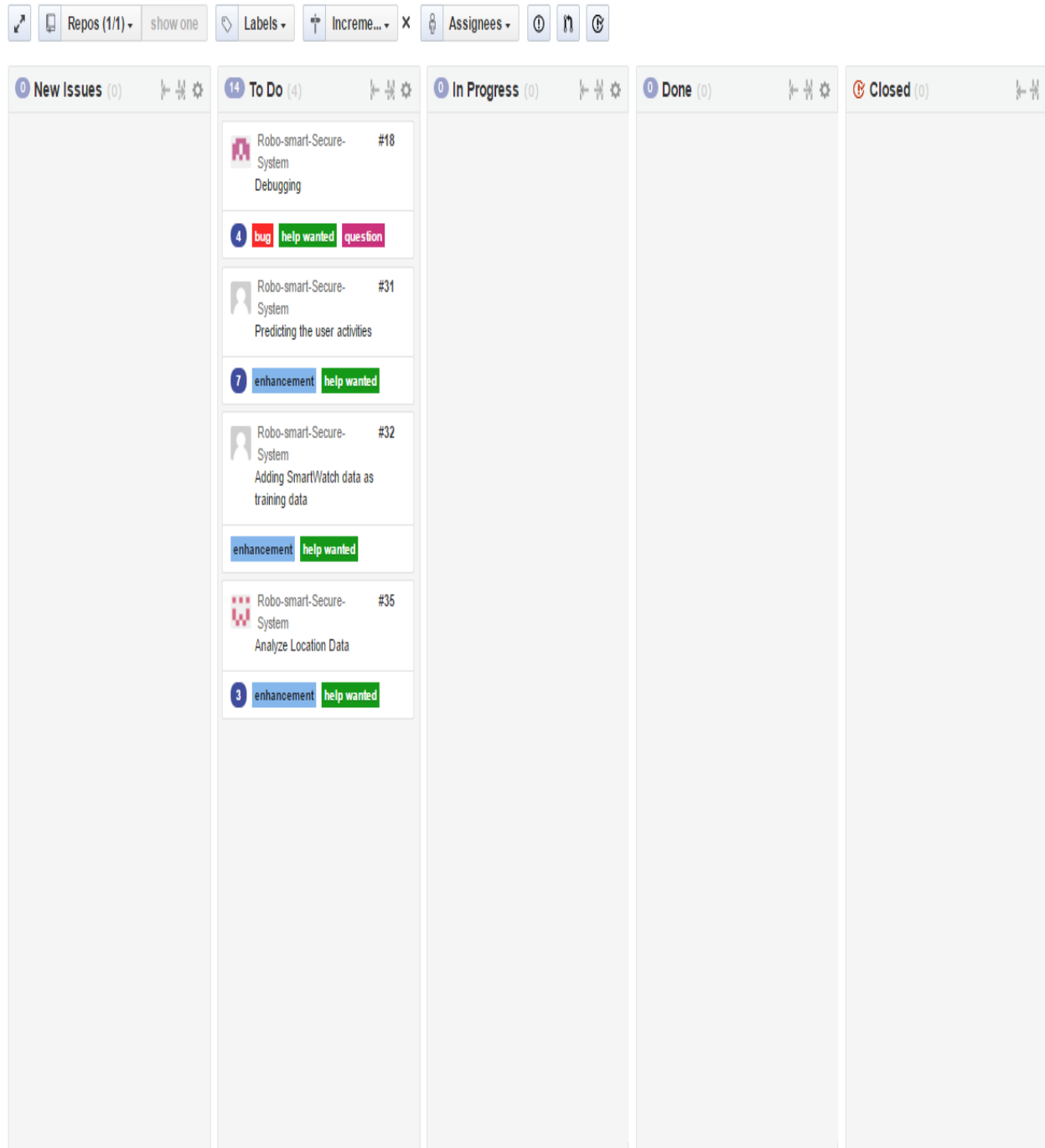
Increment 3:

The timeline for the third increment due on 04/06/2016 will have the graph as follows:



Final Increment:

The timeline for the final increment due on 04/29/2016 will have the graph as follows:

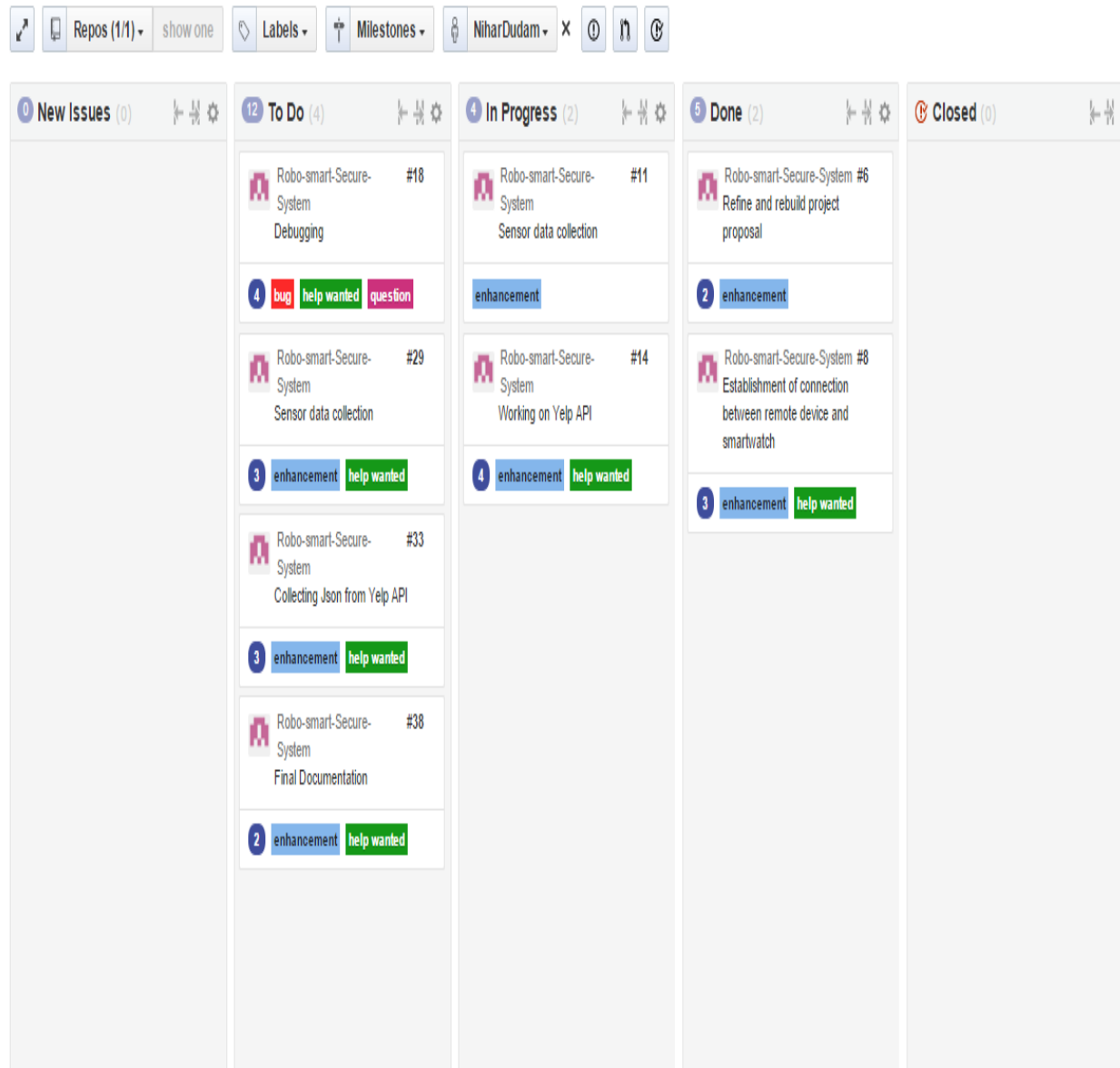


3. Project Members and Task Responsibility:

This section deals with the timeline of the project for each member along with the task responsibility:

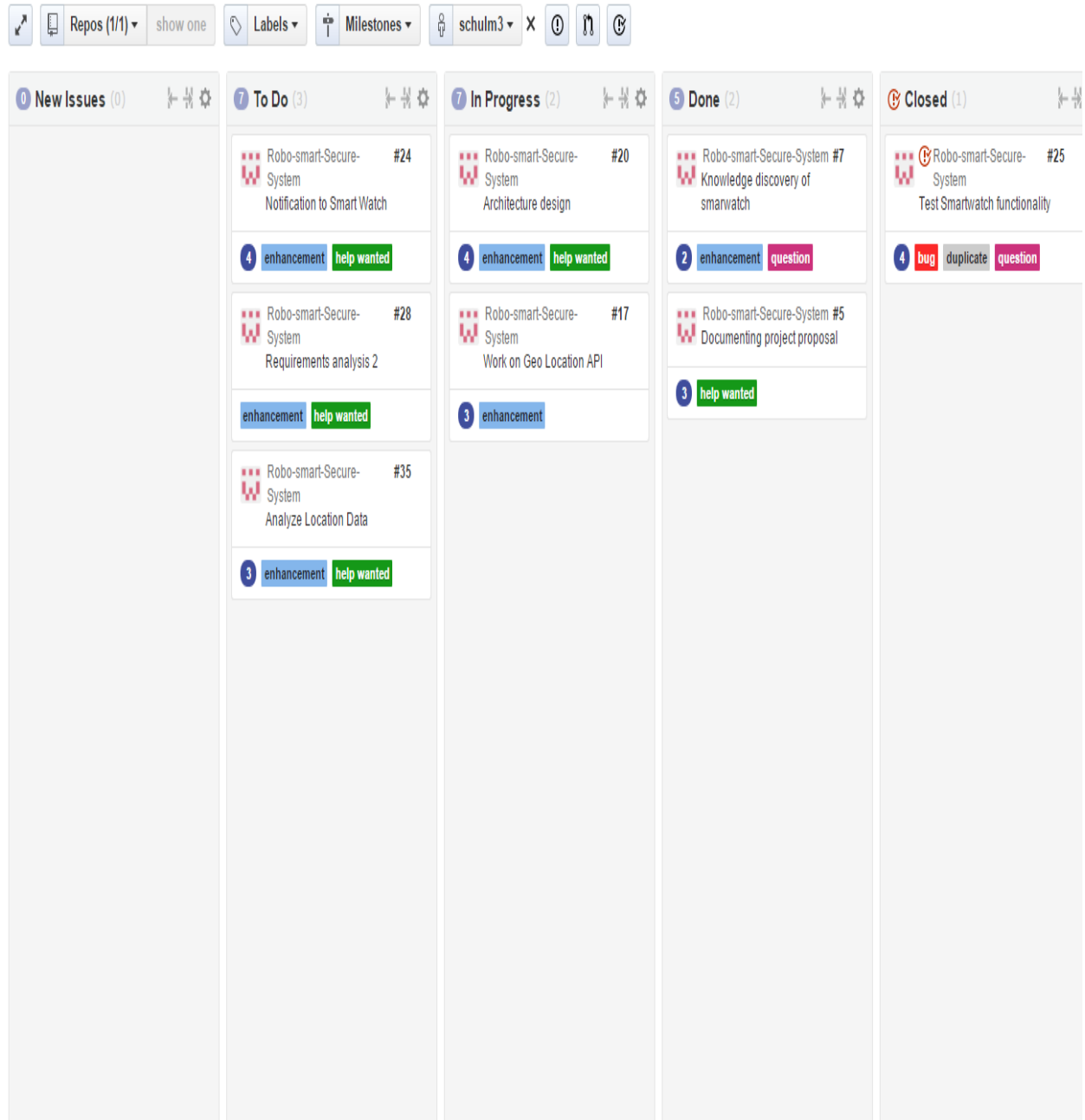
a. Nihar Dudam

The tasks assigned to Nihar in order to facilitate the completion of the project has the following diagram:



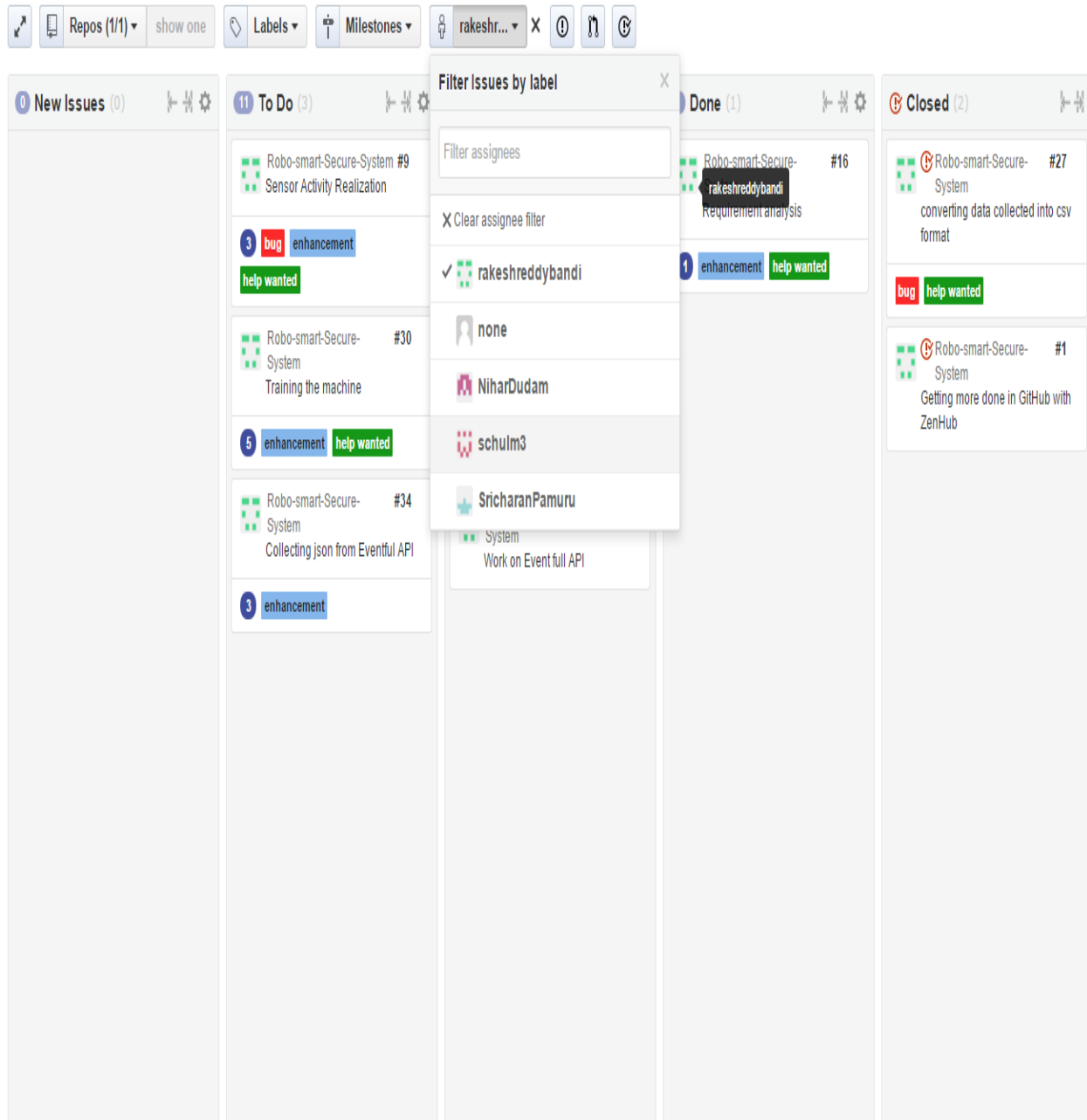
b. Mark Schultz

The tasks assigned to Mark in order to facilitate the completion of the project has the following diagram:



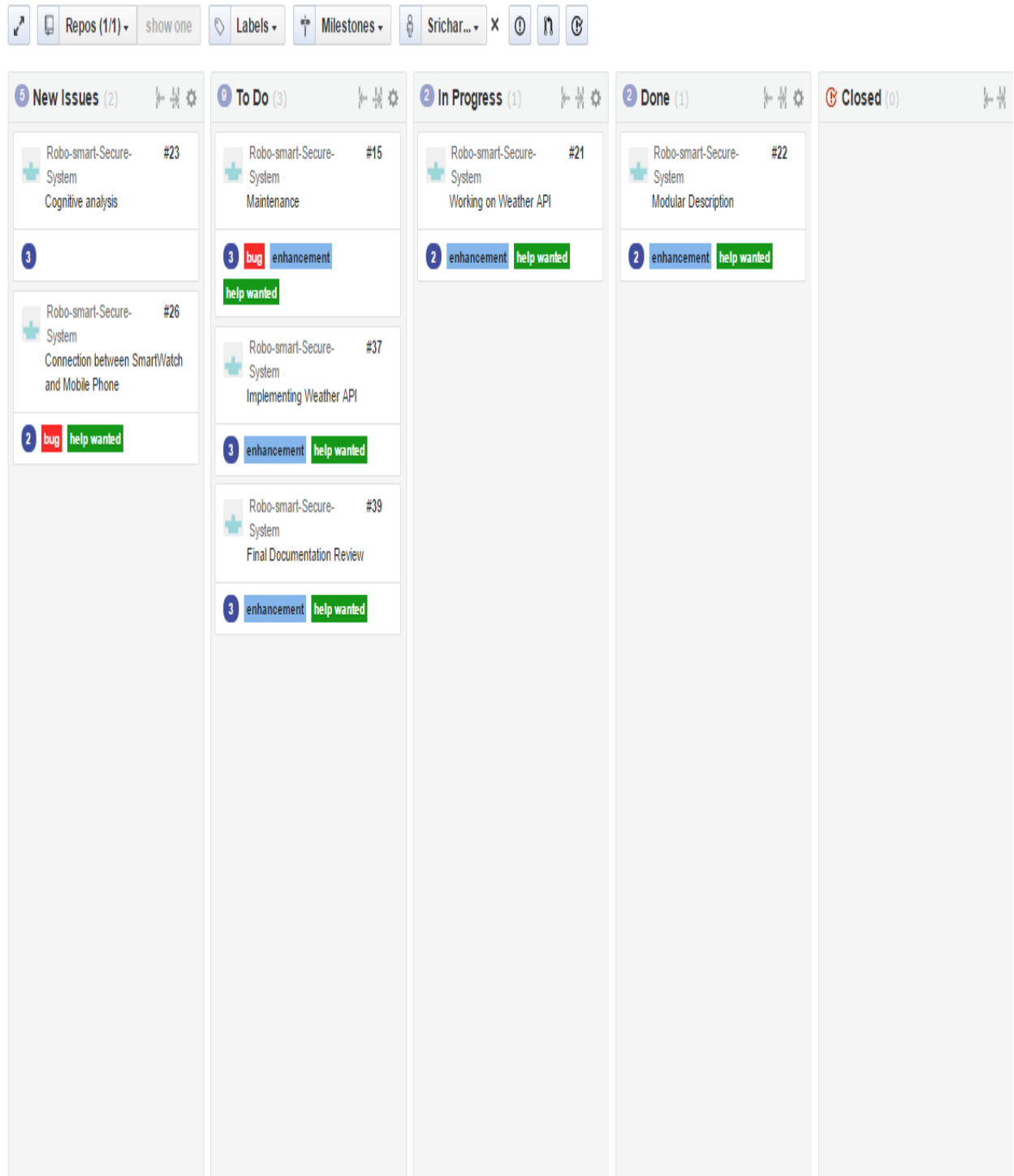
c. Rakesh Reddy Bandi

The tasks assigned to Rakesh in order to facilitate the completion of the project has the following diagram:



d. Sricharan Pamuru

The tasks assigned to Sricharan in order to facilitate the completion of the project has the following diagram:



4. Burndown charts:

The burndown charts would basically represent the time deployment of project as per the various phases. It would instantiate the modules of time and its according development over the schedule.

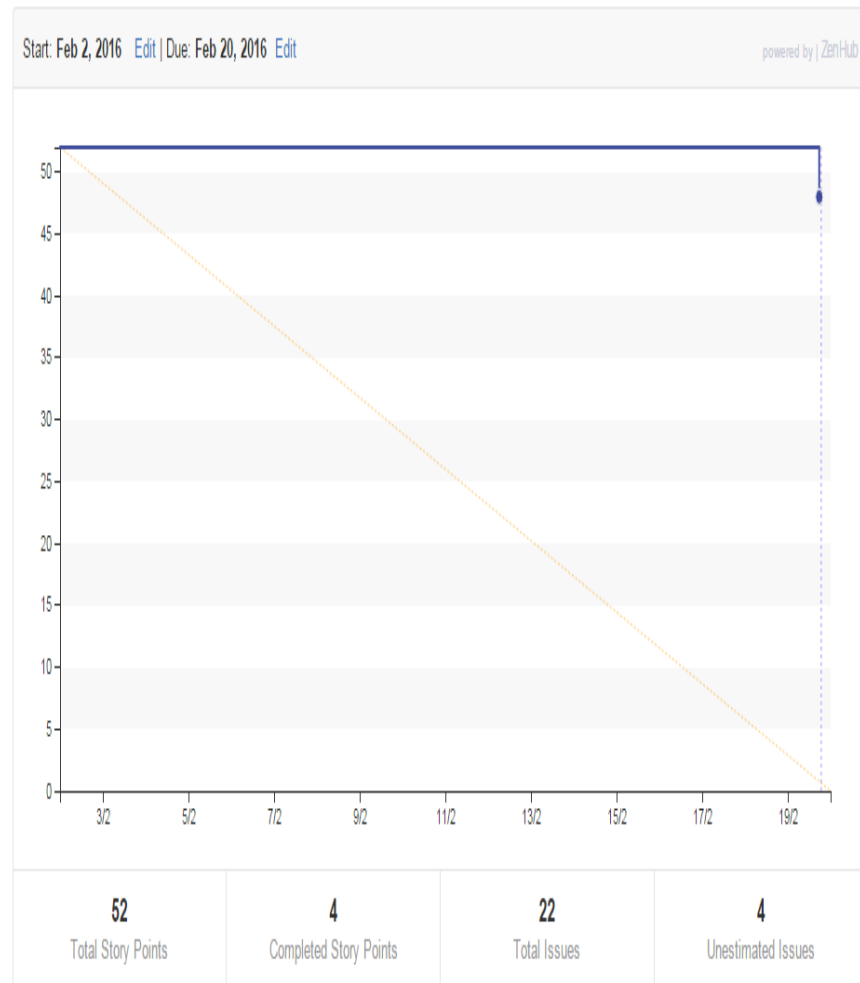
Increment 1:

The burndown chart of increment 1:

Increment 1

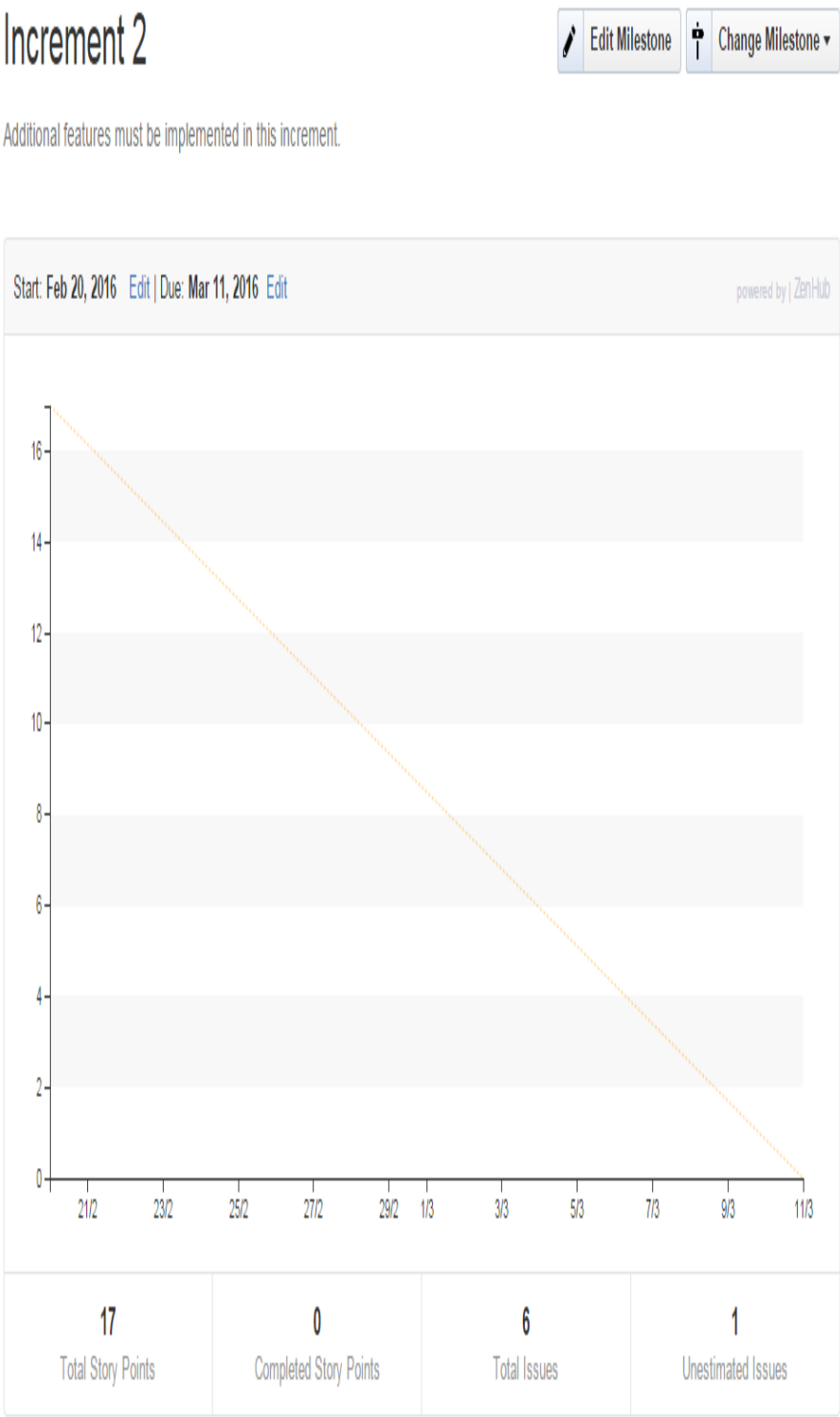
[Edit Milestone](#) [Change Milestone](#)

In this increment group need to discuss about project plan and features. How can we implement all the features discussed ?. We should also finalize about the first increment tasks to be done.



Increment 2:

The burndown chart of increment 2 is as follows:



Increment 3:

The burndown chart of increment 3 is:

Increment 3

Discussed features must be implemented in this increment.

Edit Milestone

Change Milestone ▼

Filter milestone

powered by | ZenHub X

Filter milestone

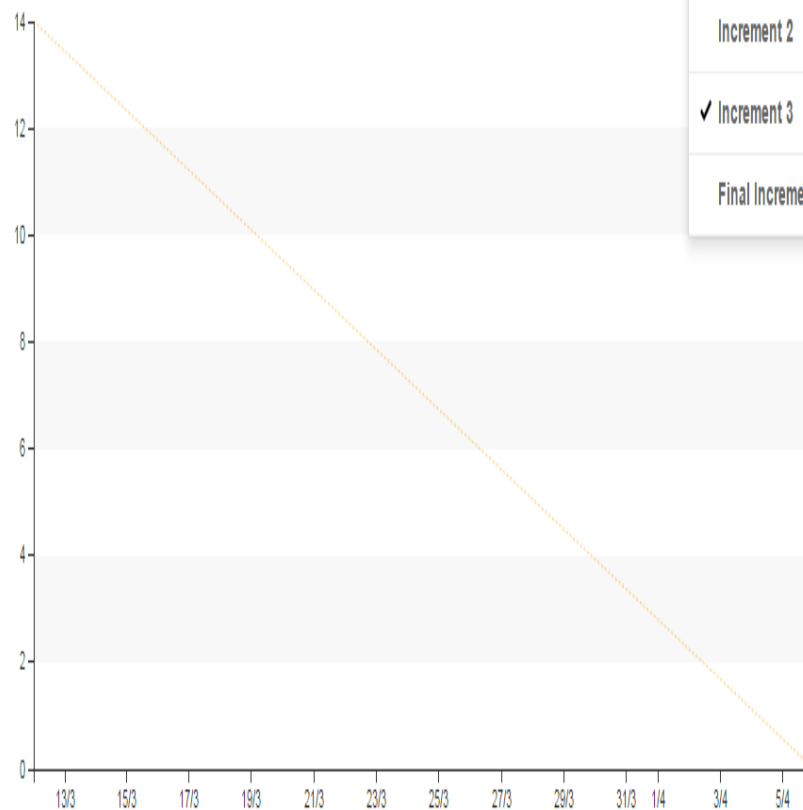
Increment 1

Increment 2

✓ Increment 3

Final Increment

Start: Mar 12, 2016 Edit | Due: Apr 6, 2016 Edit



14

Total Story Points

0

Completed Story Points

4

Total Issues

1

Unestimated Issues

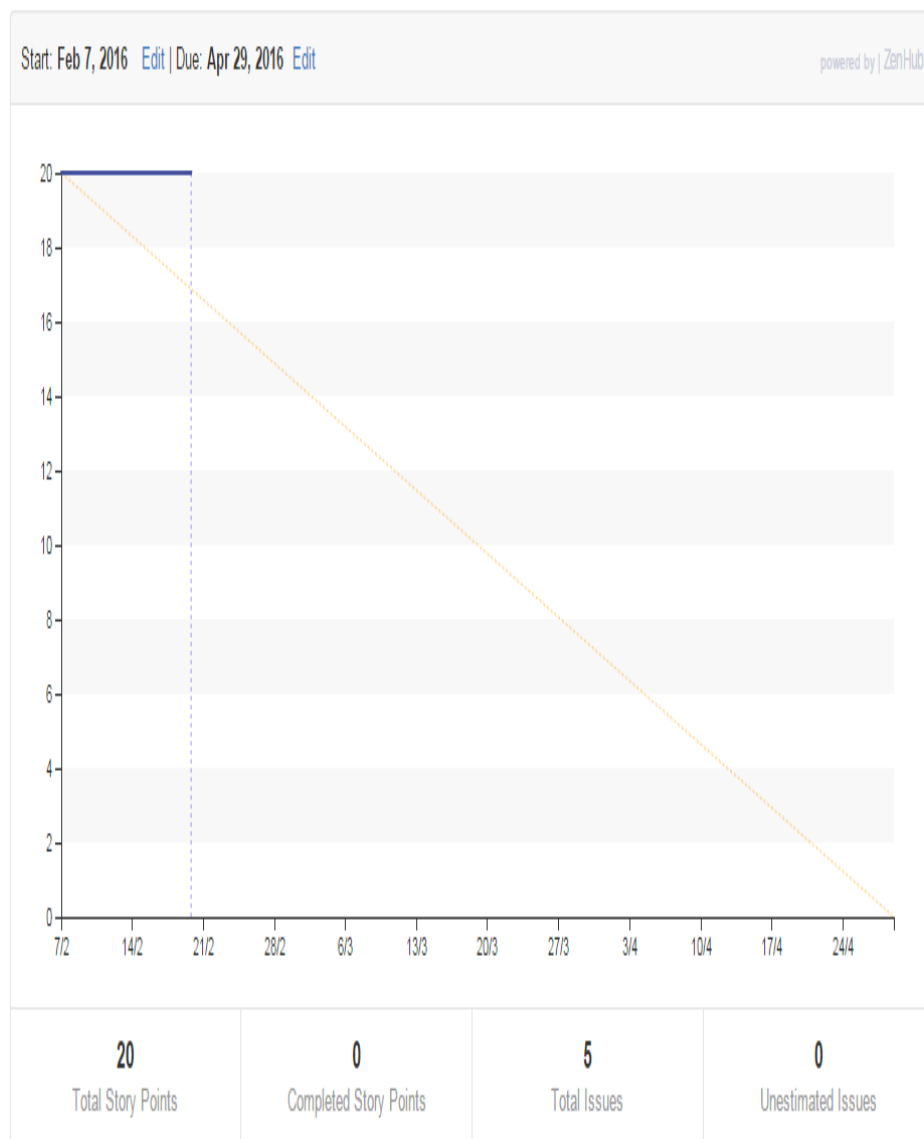
Final Increment:

The final increment has the burndown chart as follows:

Final Increment

 Edit Milestone  Change Milestone ▼

Testing the various features implemented in the project and deploying.



IV. First Increment Report

1. Existing API

Various API's used in this project are:

- a. Yelp API
- b. Eventful API
- c. Alchemy API
- d. Weather API
- e. Google Maps API

a. Yelp API:

By using Yelp API we can

- Get top 40 best results (e.g. Restaurants) as per geographical location
- Sort out the results based on ratings, distance etc.
- For every query we will be getting JSON format data which can be analyzed.

All the results are based on location. We can specify location in various formats.

1. Specify location by neighborhood, Address or city. We can even provide latitude and longitude as parameters.

Sample request:

<https://api.yelp.com/v2/search/?location=kansas city, MO&sort=2&limit=1>

2. Specify location by Geographical bounding box

Sample request:

<https://api.yelp.com/v2/search?term=food&bounds=37.925696,-122.500000|37.788022,-122.399797&limit=2>

Oath authentication must be provided so as to access the yelp data. Yelp is using Oath(1.0) authentication so as to authenticate users.

Query:

<https://api.yelp.com/v2/search/?location=kansas city, MO&sort=2&limit=1>

Response in JSON:

```
{
  "region": {
    "span": {
      "latitude_delta": 0.0,
```

```
    "longitude_delta": 0.0
  },
  "center": {
    "latitude": 39.0572982,
    "longitude": -94.6056887
  }
},
"total": 40,
"businesses": [
  {
    "is_claimed": true,
    "rating": 5.0,
    "mobile_url": "http://m.yelp.com/biz/bob-wasabi-kitchen-kansas-city?utm_campaign=yelp_api&utm_medium=api_v2_search&utm_source=oFtxuG9QYe-8GFrmKoqAqw",
    "rating_img_url": "http://s3-media1.fl.yelpcdn.com/assets/2/www/img/f1def11e4e79/ico/stars/v1/stars_5.png",
    "review_count": 81,
    "name": "Bob Wasabi Kitchen",
    "rating_img_url_small": "http://s3-media1.fl.yelpcdn.com/assets/2/www/img/c7623205d5cd/ico/stars/v1/stars_small_5.png",
    ,
    "url": "http://www.yelp.com/biz/bob-wasabi-kitchen-kansas-city?utm_campaign=yelp_api&utm_medium=api_v2_search&utm_source=oFtxuG9QYe-8GFrmKoqAqw",
    "categories": [
      [
        "Sushi Bars",
        "sushi"
      ],
      [
```

```
    "Japanese",
    "japanese"
  ]
],
"phone": "8167535797",
"snippet_text": "Phenomenal. It's really the only word I have. \n\nNever had fish
this fresh in mid-America and when you combine that with the friendliness of the...",
"image_url": "http://s3-
media2.fl.yelpcdn.com/bphoto/bDx_1_JucKzAAdfNKeezsA/ms.jpg",
"snippet_image_url": "http://s3-
media4.fl.yelpcdn.com/photo/ri2jL_uSITEPXmuwjIj6oA/ms.jpg",
"display_phone": "+1-816-753-5797",
"rating_img_url_large": "http://s3-
media3.fl.yelpcdn.com/assets/2/www/img/22affc4e6c38/ico/stars/v1/stars_large_5.png",
"id": "bob-wasabi-kitchen-kansas-city",
"is_closed": false,
"location": {
  "city": "Kansas City",
  "display_address": [
    "1726 W 39th St",
    "Volker",
    "Kansas City, MO 64111"
  ],
  "geo_accuracy": 8.0,
  "neighborhoods": [
    "Volker"
  ],
  "postal_code": "64111",
  "country_code": "US",
  "address": [
```



```

        "1726 W 39th St"
    ],
    "coordinate": {
        "latitude": 39.0573319047689,
        "longitude": -94.6053200215101
    },
    "state_code": "MO"
}
}
]
}

```

Implementation:

Hence by using Yelp API we can get top restaurants, hangout places near the user. We can access the location of the user through smart watch. We can analyse the location data and predict the location he would visit. Hence we could predict the co-ordinates or specific location of the user and by using yelp API we could get the top results and display to the user. By this we can predict daily routine of the user and help him to manage his daily activities.

b. Eventful API:

Eventful has largest data collection of events which are going on in local places of a specific location. Events can be sports, concerts, political rallies and etc. Eventful.com is built on unique , open platform that enables us to use Eventful data and its features via Eventful API.

API Tasks:

1. Event Search : We can get list of events going on based on various keywords like location, time period, category etc.
2. Venue Search : We can get list of venues based on name, type, location and etc.

Here we will be getting XML Response which need to be analyzed. Eventful API requests are accessed using Oath 1.0 authentication

Sample request :

<http://api.eventful.com/rest/events/get?...&id=E0-001-000278174-6>

XML Response:

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<event id="E0-001-000278174-6">
  <title>Martini Tasting</title>
  <url>http://eventful.com/events/west-chester/ music-martini-tasting-/E0-001-000278174-6?utm_source=apis&utm_medium=apim&utm_campaign=apic</url>
  <description>Sample drinks made with featured vodkas.</description>
  <start_time>2005-08-24 12:00:00</start_time>
  <stop_time></stop_time>
  <all_day>0</all_day>
  <tz_id>127</tz_id>
  <tz_olson_path></tz_olson_path>
  <tz_country></tz_country>
  <tz_city></tz_city>
  <venue_name>Encore Cafe - West Chester</venue_name>
  <venue_id>V0-001-000108360-1</venue_id>
  <venue_type></venue_type>
  <venue_display>1</venue_display>
  <address></address>
  <city>West Chester</city>
  <region>Ohio</region>
  <postal_code>45069</postal_code>
  <country>United States</country>
  <latitude>39.2569</latitude>
  <longitude>-84.5465</longitude>
  <geocode_type>Zip Code Based GeoCodes</geocode_type>
  <created>2005-07-28 13:09:09</created>
  <modified>2005-10-03 12:10:37</modified>
  <owner>mark</owner>
  <privacy>1</privacy>
  <free>no</free>
  <price>45</price>
  <withdrawn>0</withdrawn>
  <withdrawn_note></withdrawn_note>
  <links>
    <link id="312146">
      <url>http://frontier.cincinnati.com/calendar/proddisplay.asp?prodid=28356</url>
      <type>Official Site</type>
      <description>Details</description>
      <username>mark</username>
      <time>2005-07-28 13:09:09.372445</time>
    </link>
  </links>
</event>
```

```

    </link>
  </links>
  <comments></comments>
  <trackbacks></trackbacks>
  <images></images>
  <tags>
    <tag id="dining">
      <title>Dining</title>
      <owner>evdb</owner>
    </tag>
  </tags>
  <calendars></calendars>
  <going></going>
  <properties>
    <property>
      <id>123</id>
      <name>Rating</name>
      <value>5 stars</value>
    </property>
  </properties>
  <categories>
    <category>
      <id>business</id>
      <name>Business, Finance, Investing</name>
    </category>
  </categories>
</event>

```

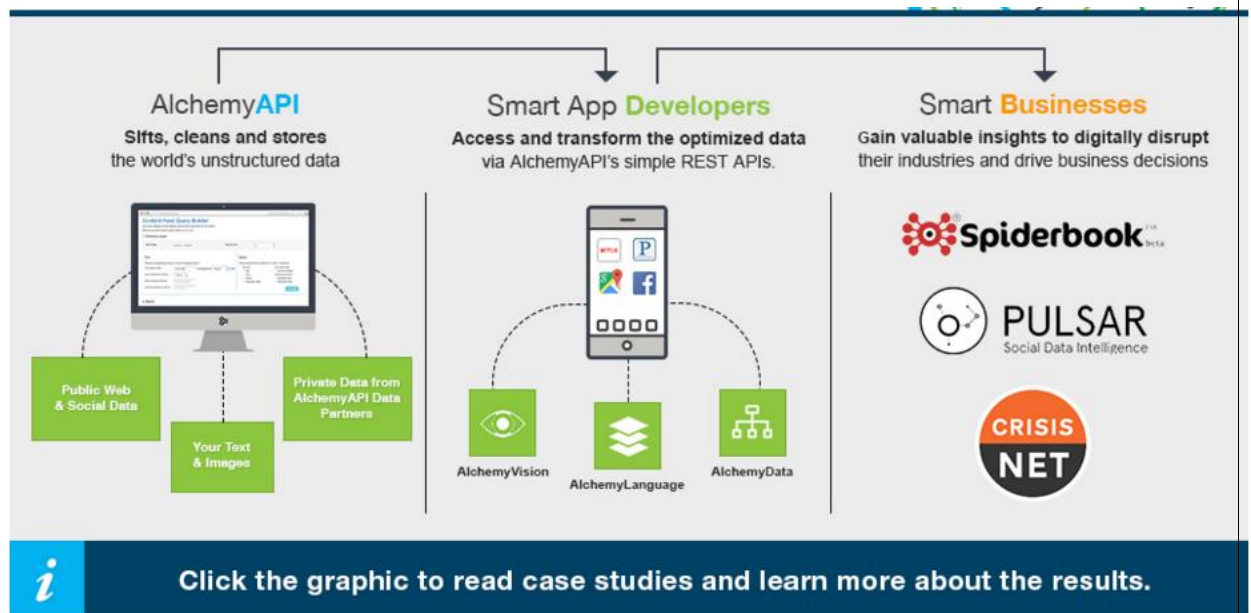
We can retrieve event details from this XML Response and send a notification to user regarding the specific event whenever user is residing near the happening place through smart watch. By sending notification user can be aware of the important events happening.

c. **Alchemy API :**

Alchemy API uses natural language processing technology and various machine learning technology to predict and extract semantic meta-data from various information available on places, people, topics, facts, relationships, languages and authors.

In this project we have implemented Sentimental Analysis, taxonomy, Concept tagging, Entity extraction, keyword extraction, Image Extraction and combined call.

Working of Alchemy API :

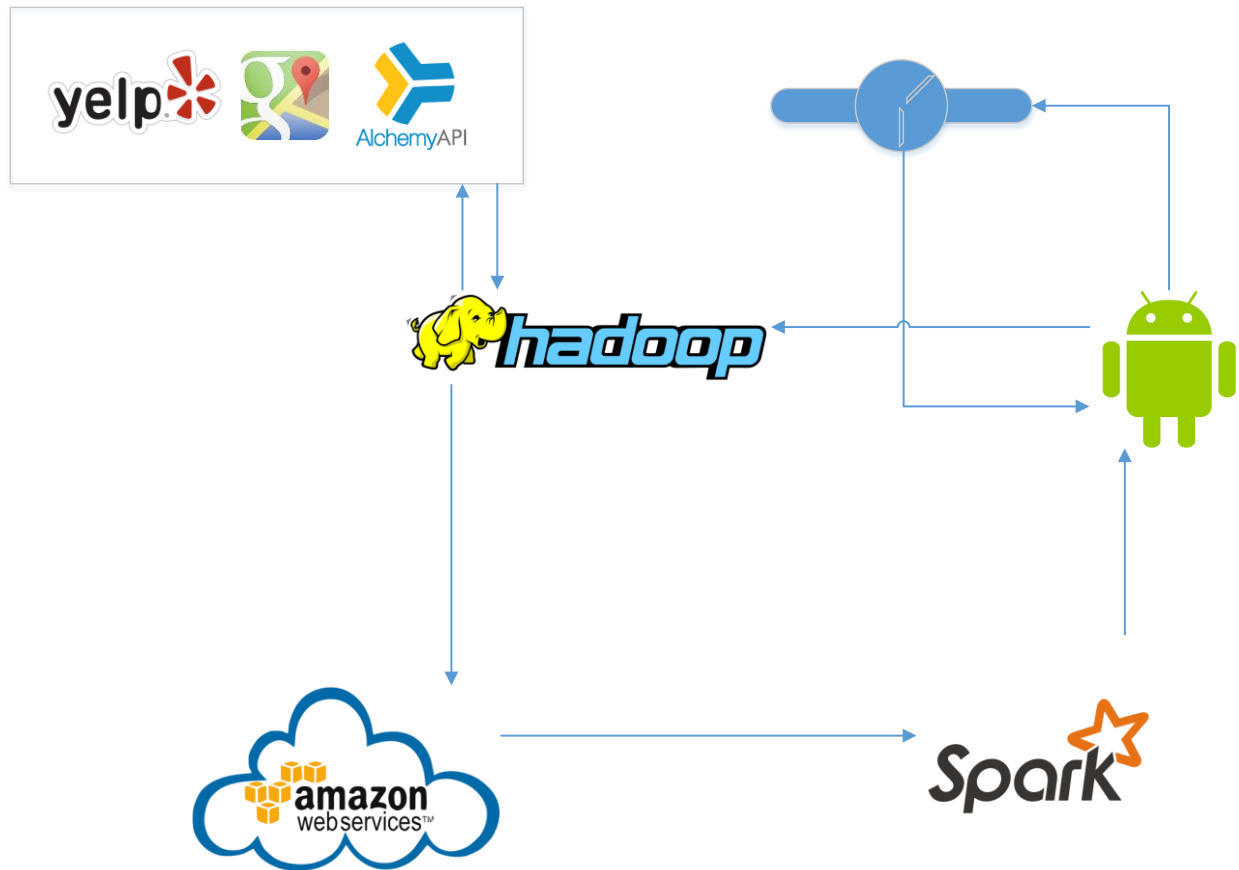


We need to get an Alchemy API key to access the API data.

With this implementation user can make use of various features like, he can get sentimental analysis of the data which can be provided using an URL. He can also judge the details of the image.

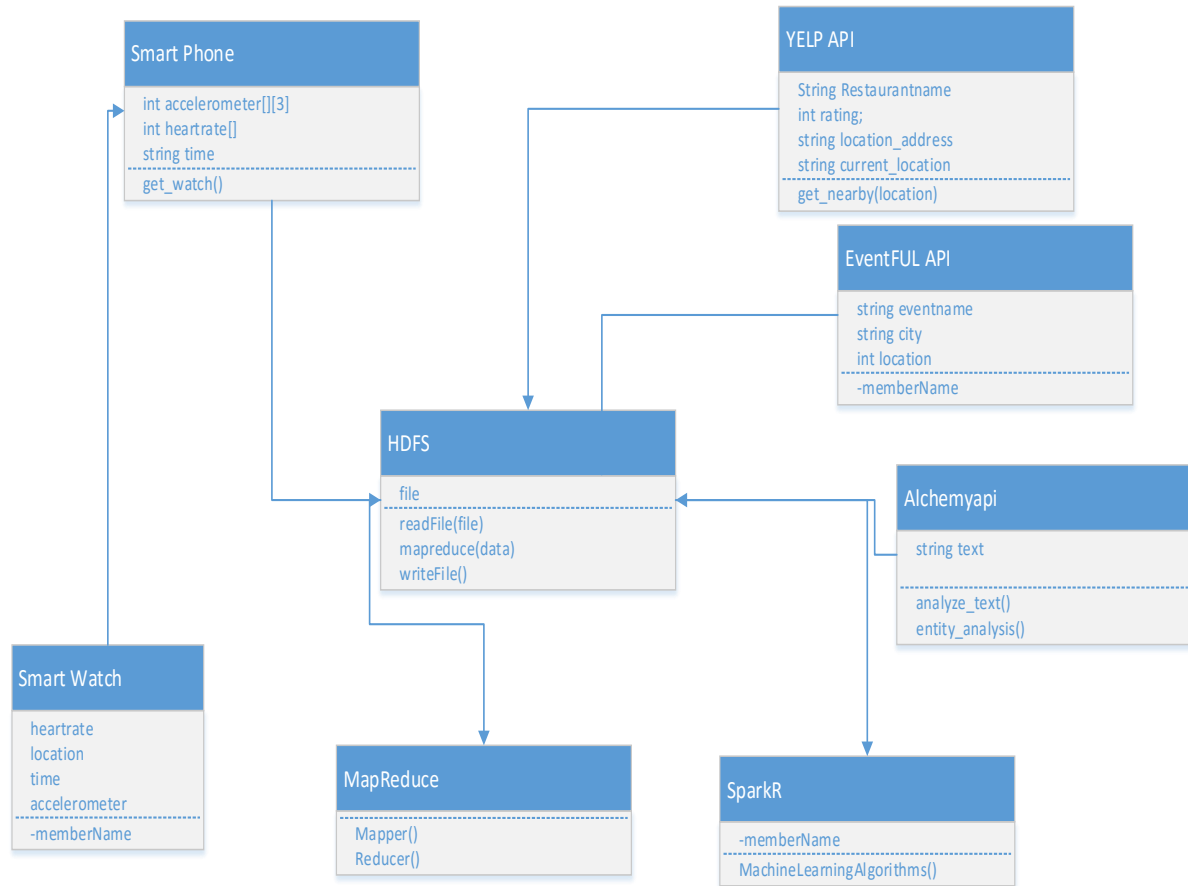
2. Design of Features:

a. Architecture:



Our application begins and ends with the smartwatch user interface. The user will input to the watch. watch sends data to the Hadoop File System through the Android application. The Hadoop System will organize data that comes both from the user and from APIs. The application will then send data to the Amazon Web Services to run analytics on the data. We will implement MapReduce programs in Spark and Spark R. We will use several machine learning algorithms for prediction. After running the Query in spark, the data is sent back to the Android application, which can then be viewed by the user on their smartwatch.

b. Class Diagram:

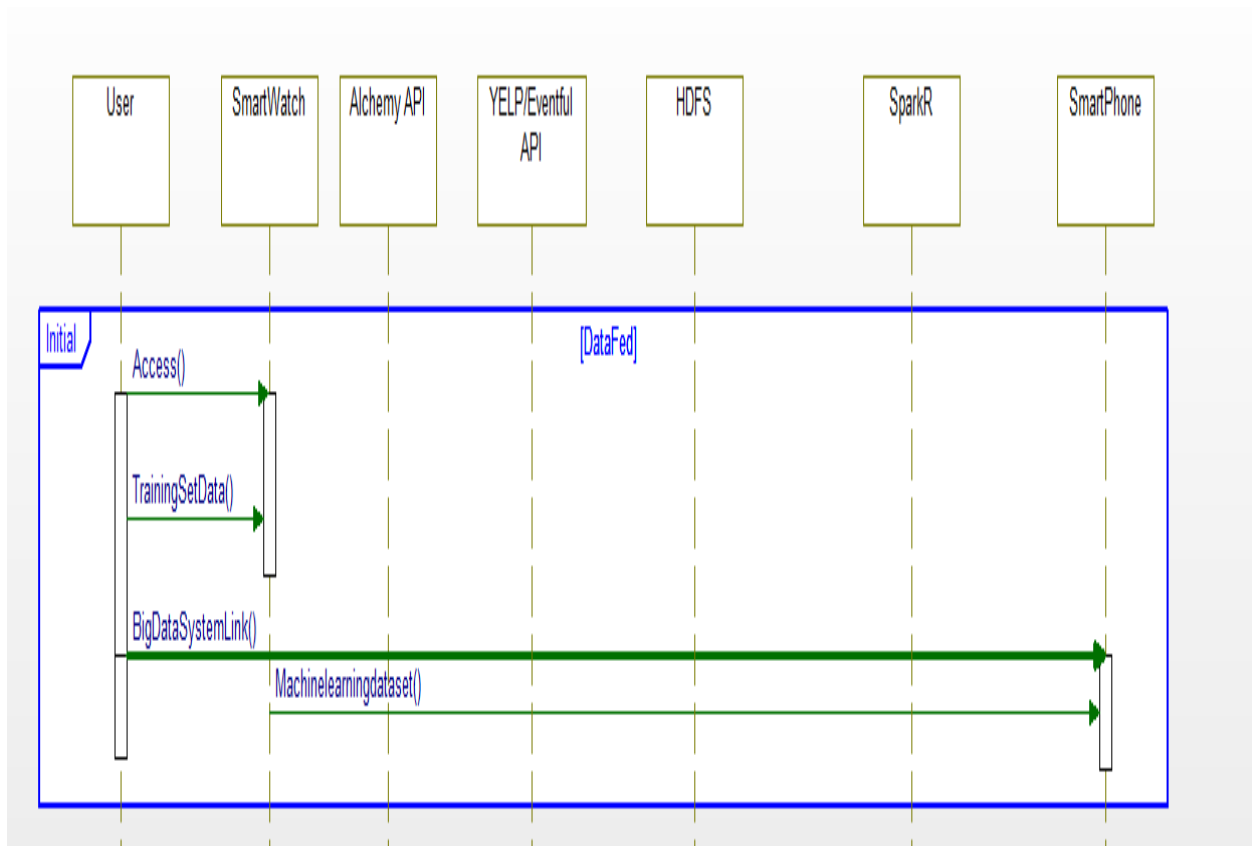


The class diagram we have drawn outlines the structure of our current project. We have three classes for three APIs, which provide data to the HDFS class. Also providing data to the HDFS is the smart phone, which calls `get_watch()` to retrieve smart watch data. The smart watch collects heartrate, location, time, and accelerometer data. HDFS will implement MapReduce and SparkR classes to analyze data it is given. SparkR will implement several of the machine learning algorithms, which will be discussed in more detail later in this paper.

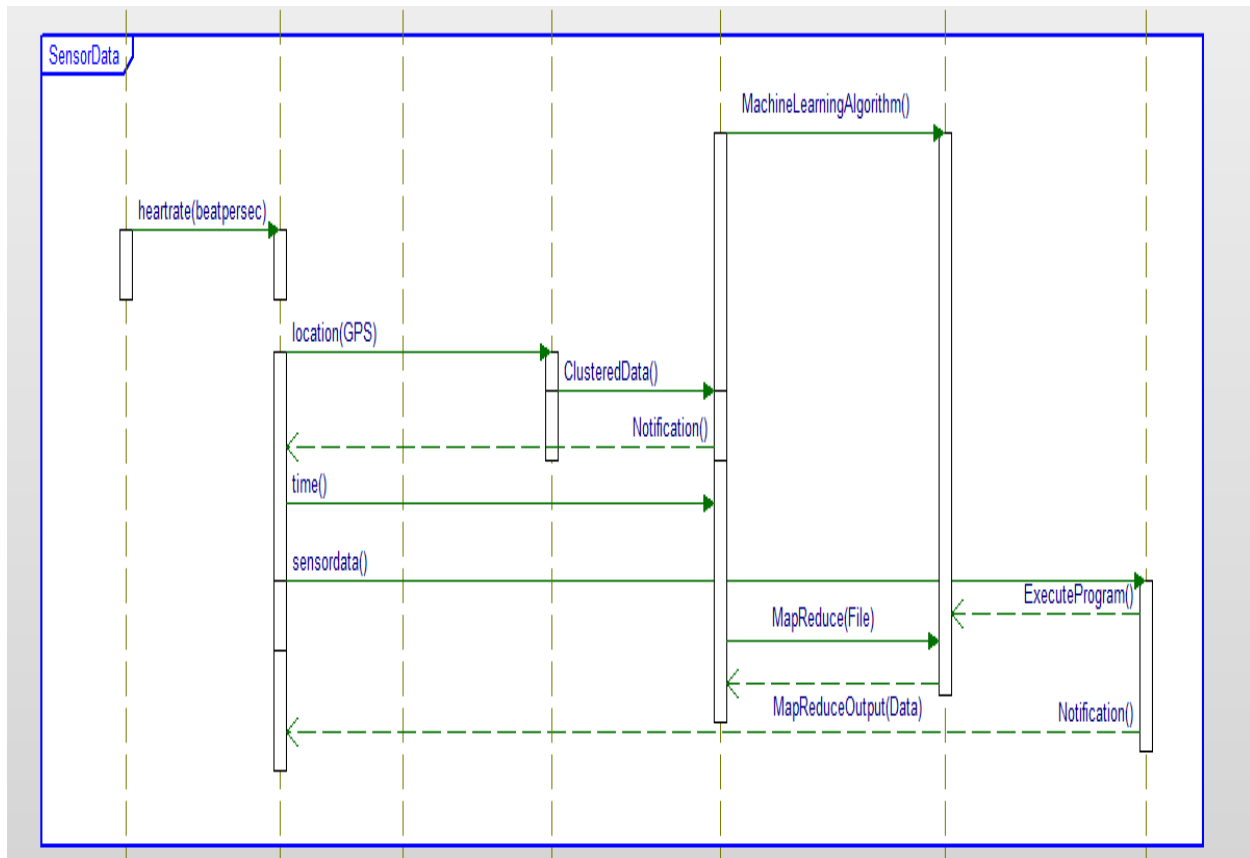
c. Sequence Diagram:

The sequence diagram is shown as below. The sequence diagram is

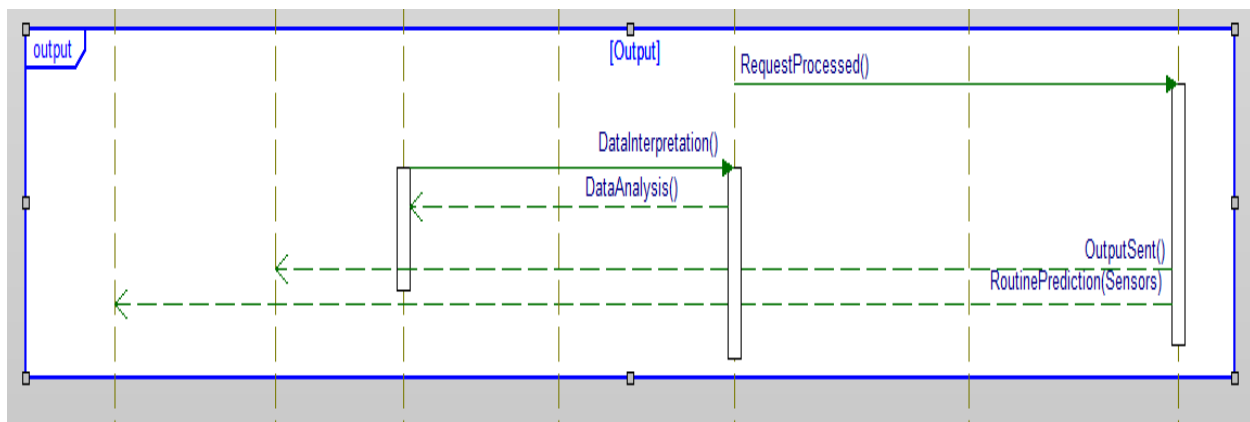
Disintegrated and attached:



The datafed process would mainly consist of all the classes to initially set up communication. The access and training set data is collected and then related with the smartwatch by the user as there is a setup of data fed into the remote machine. The machine learning dataset would be communicated through the smart phone and the device learns as to what the user pattern is.



The sensor data would basically do the operations involved mainly from the HDFS to obtain a programmed output that provides the user as to what the routine is or rather should be.



The final output would basically be a clear cut approach and the user gets a routine predicted.

d. Machine Learning Algorithms:

In Increment one, we have implemented a naïve bayes algorithm in R to classify datasets of user movements. The naïve bayes is a classifier that uses Bayes theorem to determine the most likely outcome. Bayes Theorem is written as such: $P(A|B) = \frac{P(A)*P(B|A)}{P(B)}$. In this equation, P(A) is the prior, P(B|A) is the likelihood, and P(B) is the evidence. This algorithm relies on P(A) and P(B) being independent of one another. While in reality, this is rarely the case, it provides an estimate that is extremely effective in classification.

In future increments, we plan on using a several more machine learning algorithms implemented on SparkR to scale larger sets of data.

e. Datasets:

A dataset is a collection of related sets of information iterated across an activity centric output. The dataset collected as input from the smartwatch will be fed via machine learning algorithms to the remote machine for it to be analyzed. The data/information is obtained from the different sensors which are required for providing an analytic comparison. Output is retrieved from the values fed. Supervised learning is in the initial stages since the machine is trained or rather intelligence provided regarding what to react in which case. After the training data set is provided, unsupervised learning comes to the frame. Here, the machine is able to analyze the data and compare with the training set. It will also be able to learn in case the user does an activity different from the routine. A notification is sent but nonetheless learning is also done by the machine, which in our case is the smart phone connected via USB.

3. Implementation

Mobile Client Implementation (Smart Watch, Smart Phone, Robot):

Mobile Application for Collecting the Sensor Data:

- We have developed an application which is specifically for collecting the user smart phone and smart watch sensor data.
- In this application we have collected the Sensor data and we stored in to the csv files in the Device memory.
- Later we use this data to train over model to predict the user patterns regularly.
- This application mainly gets the all sensor data values from the Sensors such as Gyroscope, Accelerometer, Proximity, Light, Magnetic and Heart Rate in the Smart watch and smart phone.
- Apart from this we also collected the Location's latitude and longitude at which the sensor data collected and we store the data in csv file.

Text Analysis and Image Analysis Application:

- We have implemented the Text and Image analysis of Data and Image respectively using the alchemy API.
- The Smart phone calls from the Alchemy API to retrieve entities, text, and image data.
- The application can analyze text that is given to it. With this text, it can determine important entities, such as a name or a place. It can also look for keywords in the text.
- The application also is able to perform sentiment analysis, and can rate the text as positive or negative. The application can also look at images and classify them based on their characteristics. In our testing, the image is successfully classified. Once these features are performed, the smart watch is given a notification with the analysis.

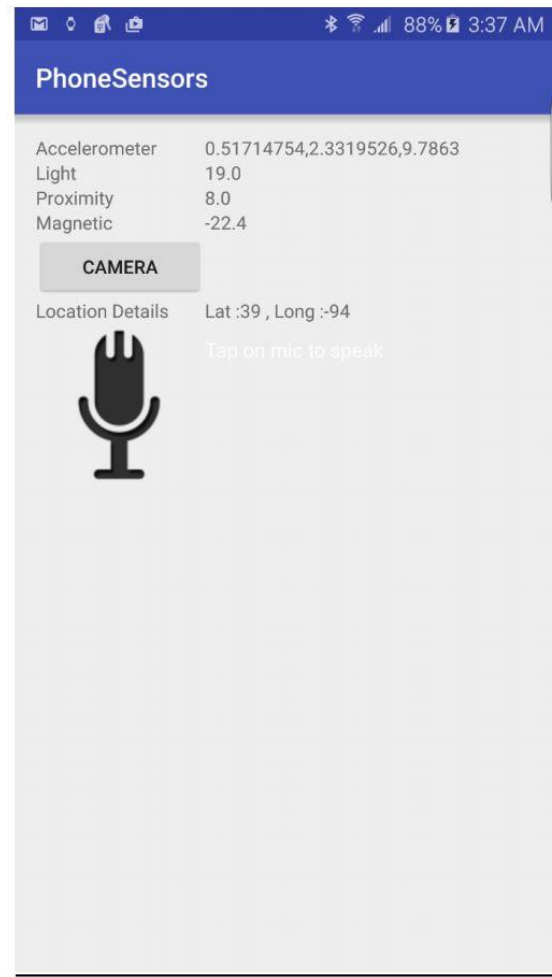
Machine Learning Application:

Machine Learning Classification Algorithm- Naïve Bayes Classification Algorithm to predict the User Activity:

- We have predicted the user activity using the Naïve Bayes Classification algorithm.
- We have trained the model based on the data obtained for different activities such as walking, getting out of bed and climbing up the stairs.
- So using the training data, we have predicted the user activity for the particular accelerometer readings using the Naïve Bayes Classification algorithm.
- Naïve Bayes Classification algorithm:
 1. Calculate the Mean and Standard deviation for each of column in the data set.
 2. The likelihood for each of the axis for activity is calculated.
 3. The likelihood of each of the axis reading was calculated using the Gaussian distribution function.
 4. Using the Naïve Bayes posterior probability formula we have computed the probability of occurrence of posterior.
 5. Posterior probability is directly proportional to product of prior probability and likelihood.
- Using the posterior probability we have predicted which activity is most likely to perform.

4. Deployment:

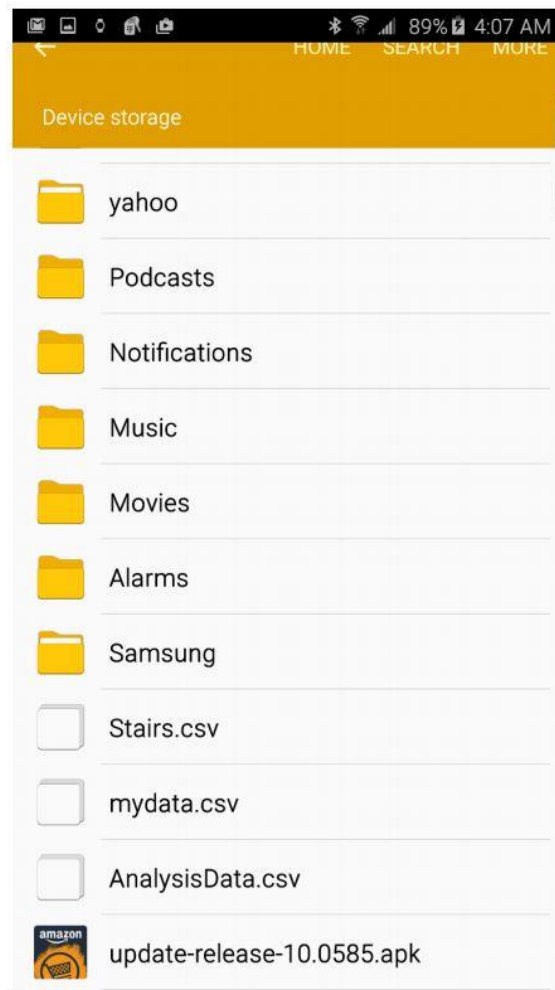
1. SmartPhone/SmartWatch Sensors Data Storage Application:



The various sensors in both mobile and smart watch gives the data values continuously for the user activities.

We store these values in the Internal storage of the Device.

2. . SmartPhone/SmartWatch Sensors Data Storage Application File Explorer View:



We can see the data sets as .CSV files in the Internal memory of Android Devices.

Machine Learning Model:

1. Model predicting the User Activity as Getting up the Bed

```
Console ~/
> probw
[1] 1.612338e-24
> probG
[1] 3.057487e-10
> probC
[1] 1.011029e-22
>
> maxprob=max(probc,probG,probw)
> if(maxprob==probc){
+   print("climbing stairs")
+ }else if(maxprob==probG){
+   print("Getting up bed")
+ }else{
+   print("walking")
+ }
[1] "Getting up bed"
> attributes(tree)
```

2. Model Predicting the User Activity as Climbing the Stairs

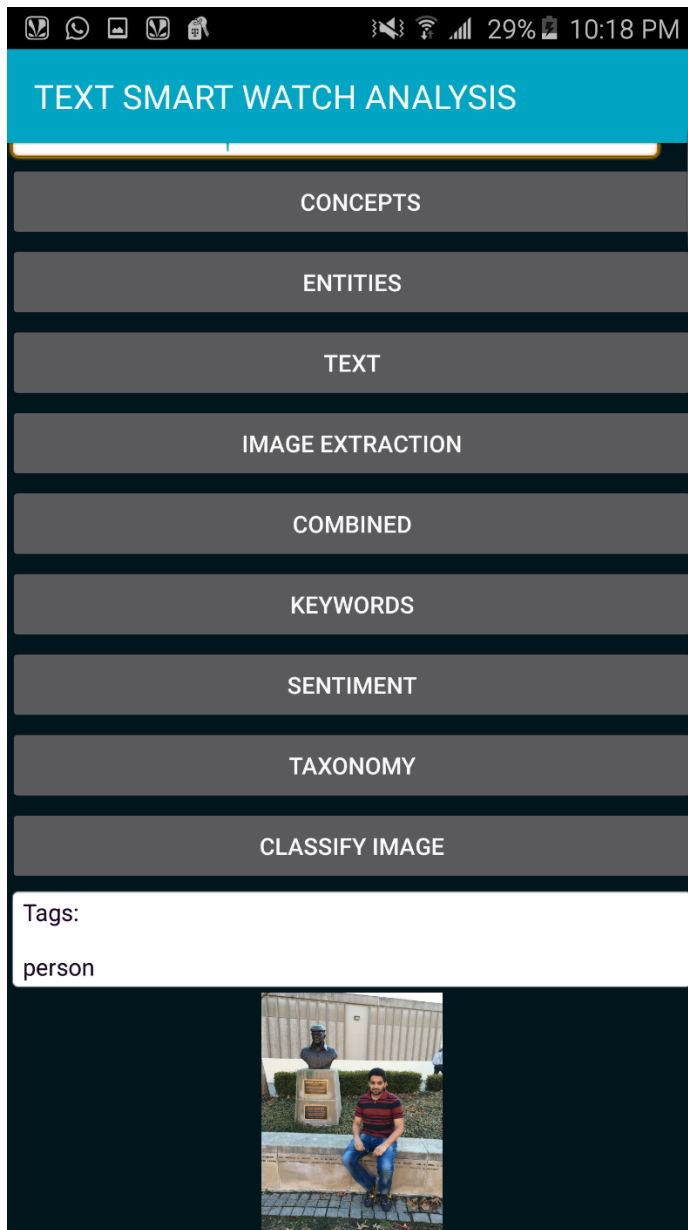
```
Console ~/
> probw
[1] 4.416484e-09
> probG
[1] 6.389902e-10
> probC
[1] 1.2815e-08
>
> maxprob=max(probc,probG,probw)
> if(maxprob==probc){
+   print("climbing stairs")
+ }else if(maxprob==probG){
+   print("Getting up bed")
+ }else{
+   print("walking")
+ }
[1] "climbing stairs"
> attributes(tree)
```

3. Model Predicting the User Activity as Walking

```
Console ~/
> probw
[1] 1.655857e-08
> probG
[1] 3.794086e-09
> probC
[1] 1.932059e-09
>
> maxprob=max(probc,probG,probw)
> if(maxprob==probc){
+   print("Climbing stairs")
+ }else if(maxprob==probG){
+   print("Getting up bed")
+ }else{
+   print("walking")
+ }
[1] "walking"
> attributes(tree)
```

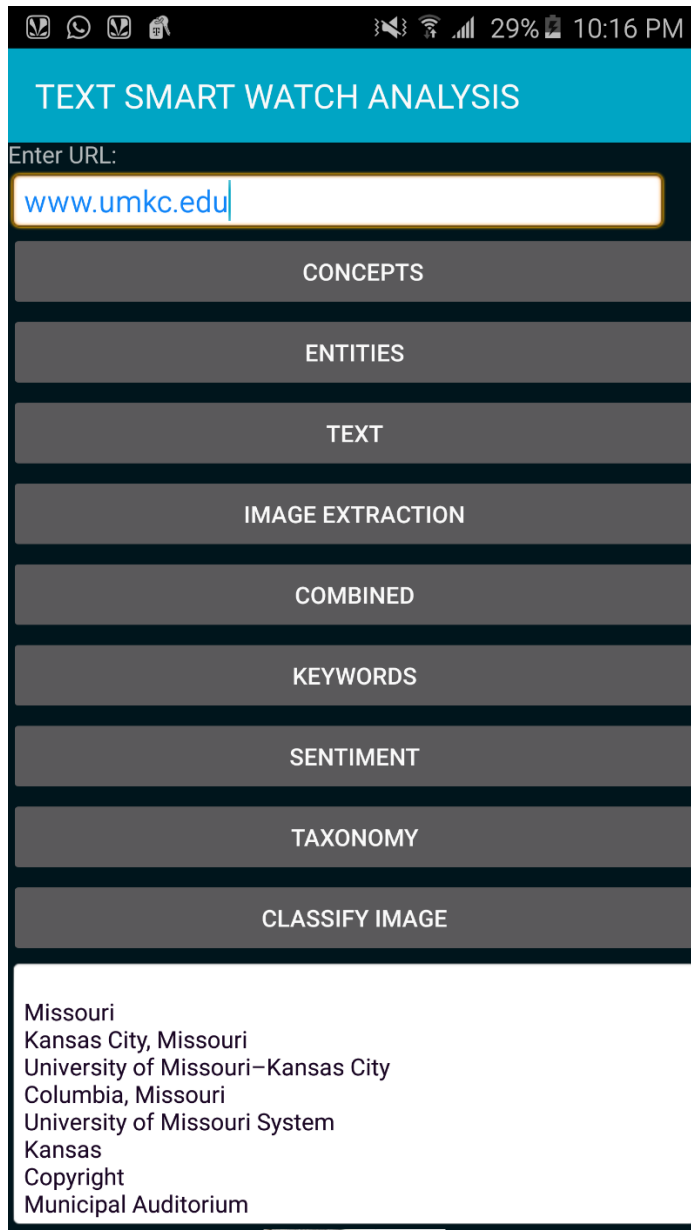
TEXT AND IMAGE ANALYSIS APPLICATION-ALCHEMY API:

1.



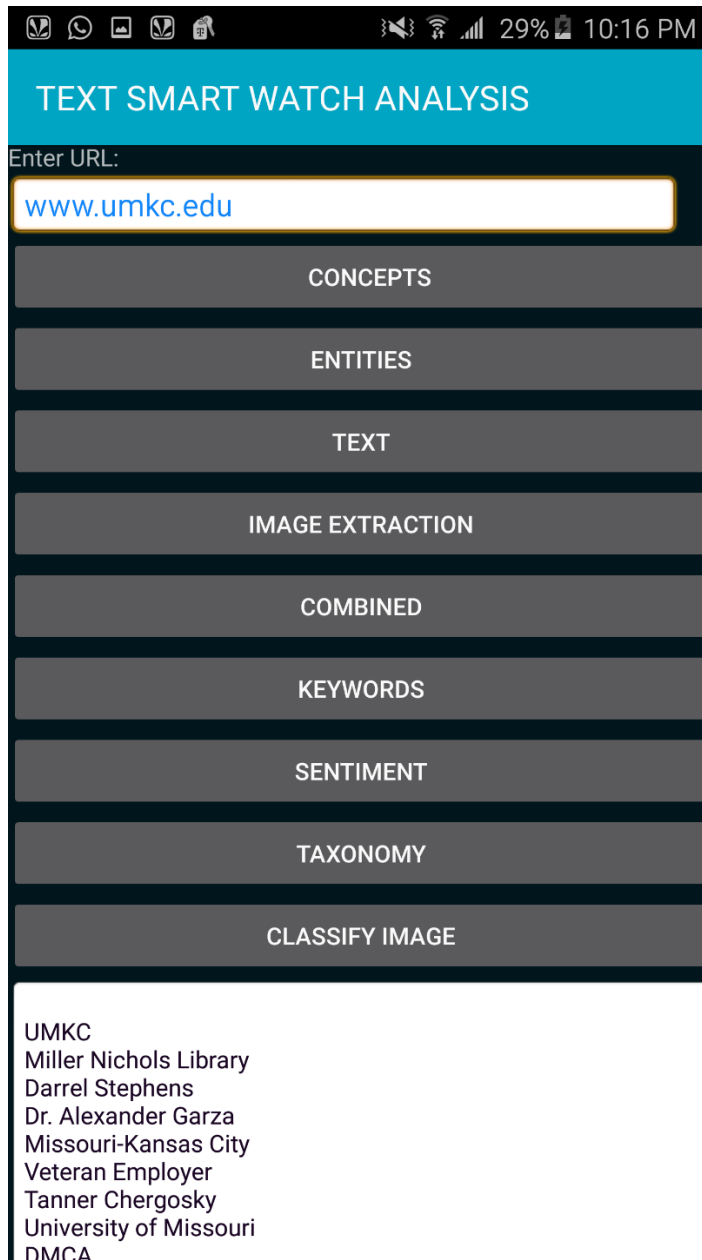
The above figure illustrates about classify image button where it does image analysis of the image provided and gives respective tags.

1.



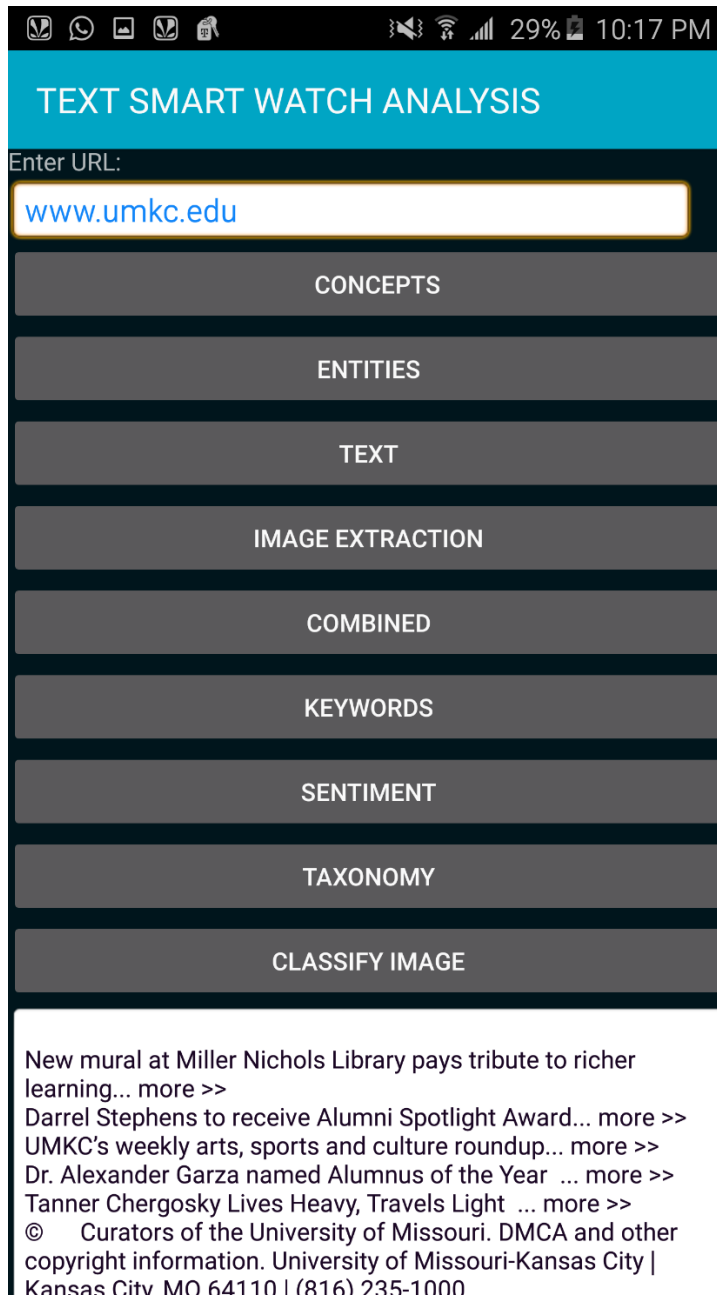
The above figure explains about the **concepts of** URL mentioned in the text box.

2.



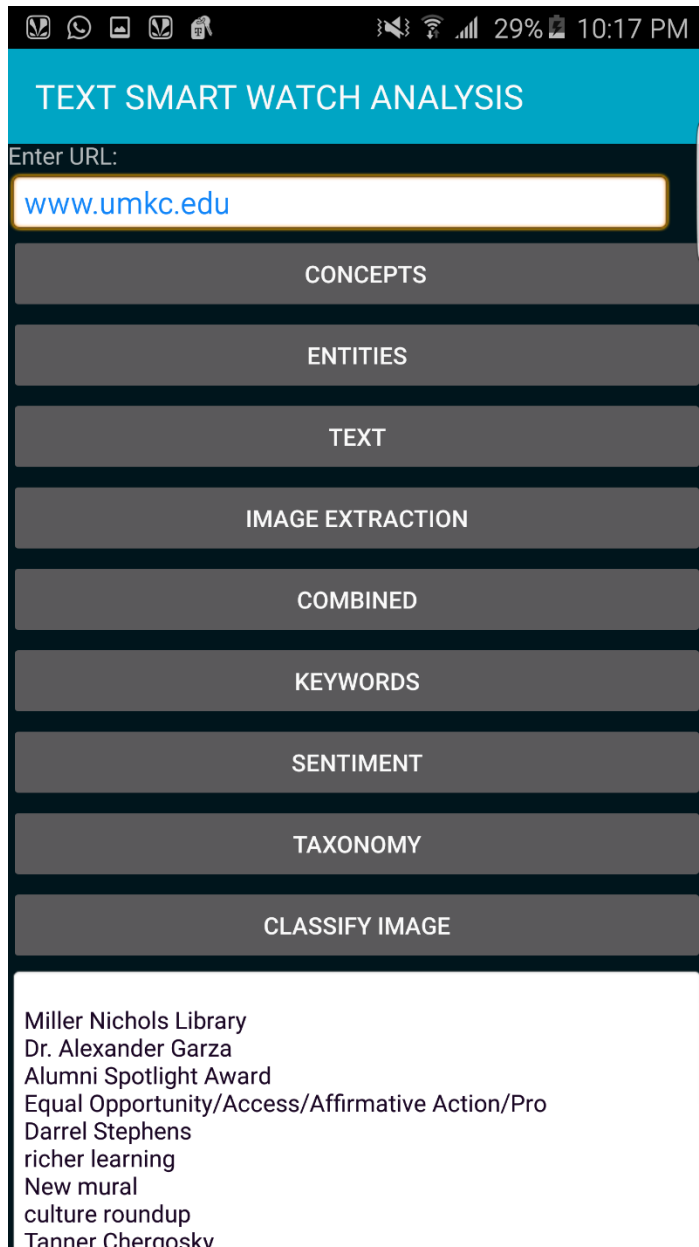
The above figure illustrates about the various **entities** in the URL.

3.



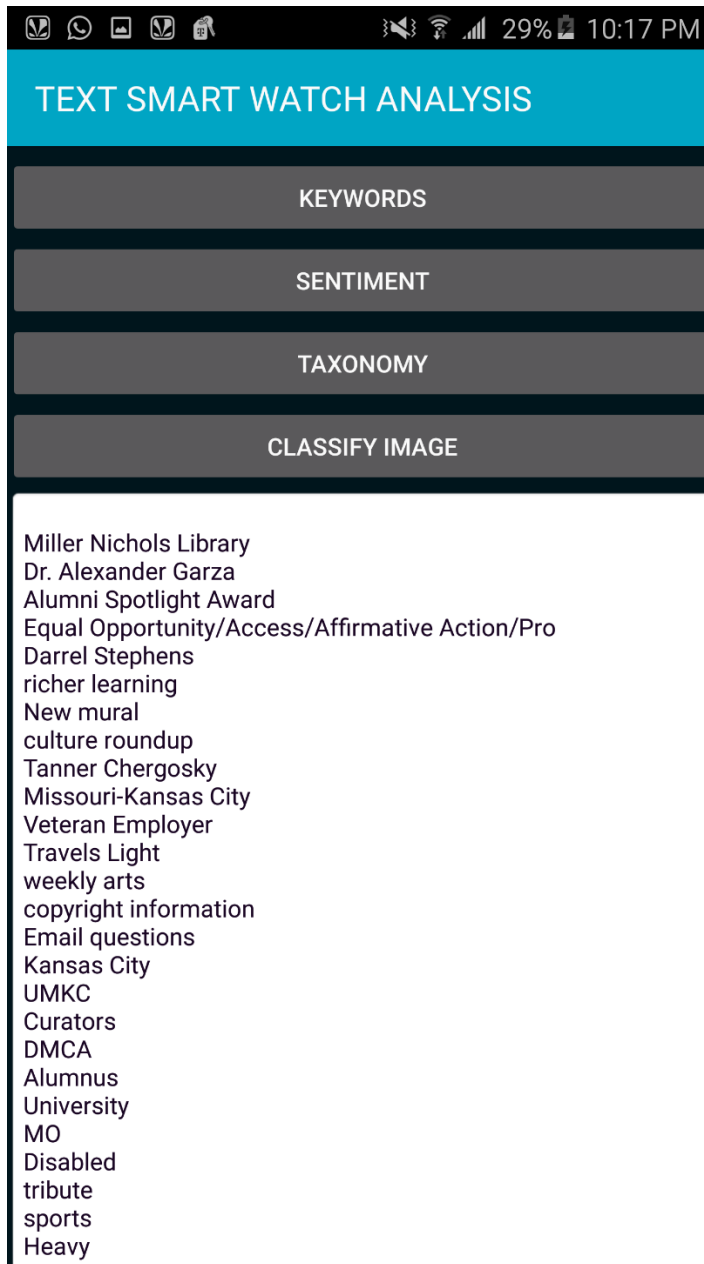
The above figure explains about the **text** present in the URL.

4.



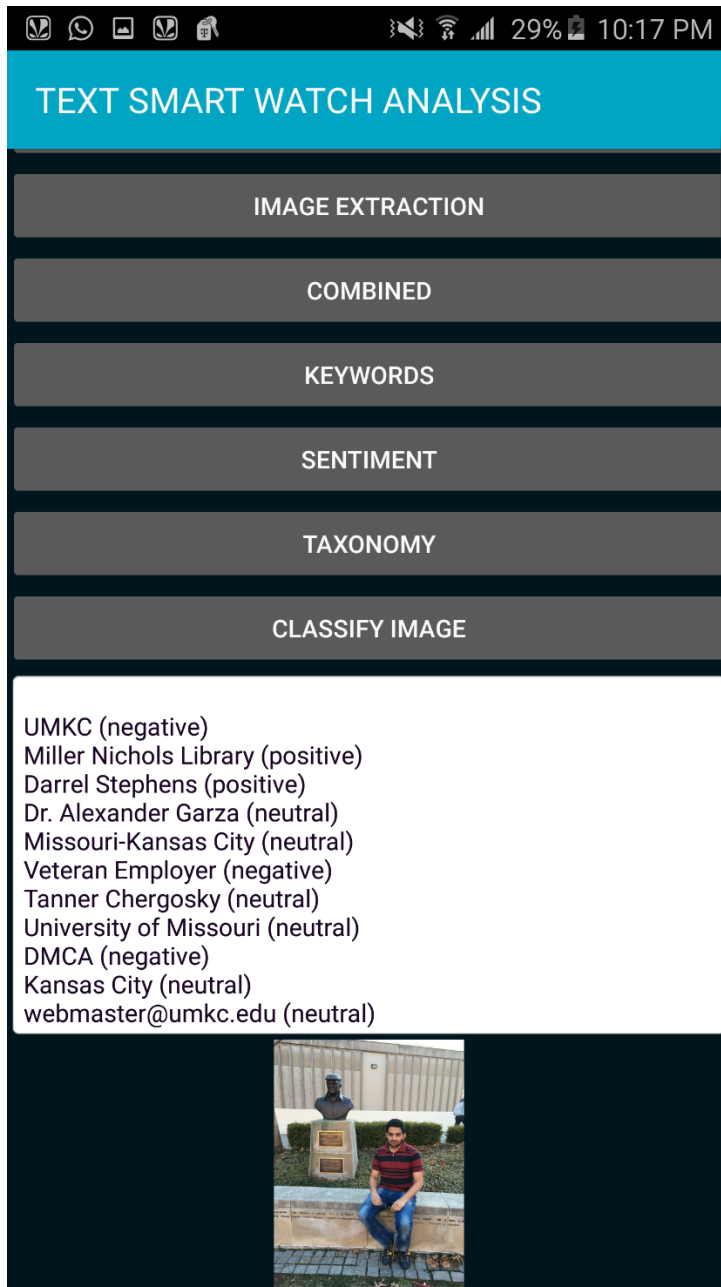
The above figure explains about the **combined** feature of the application.

5.



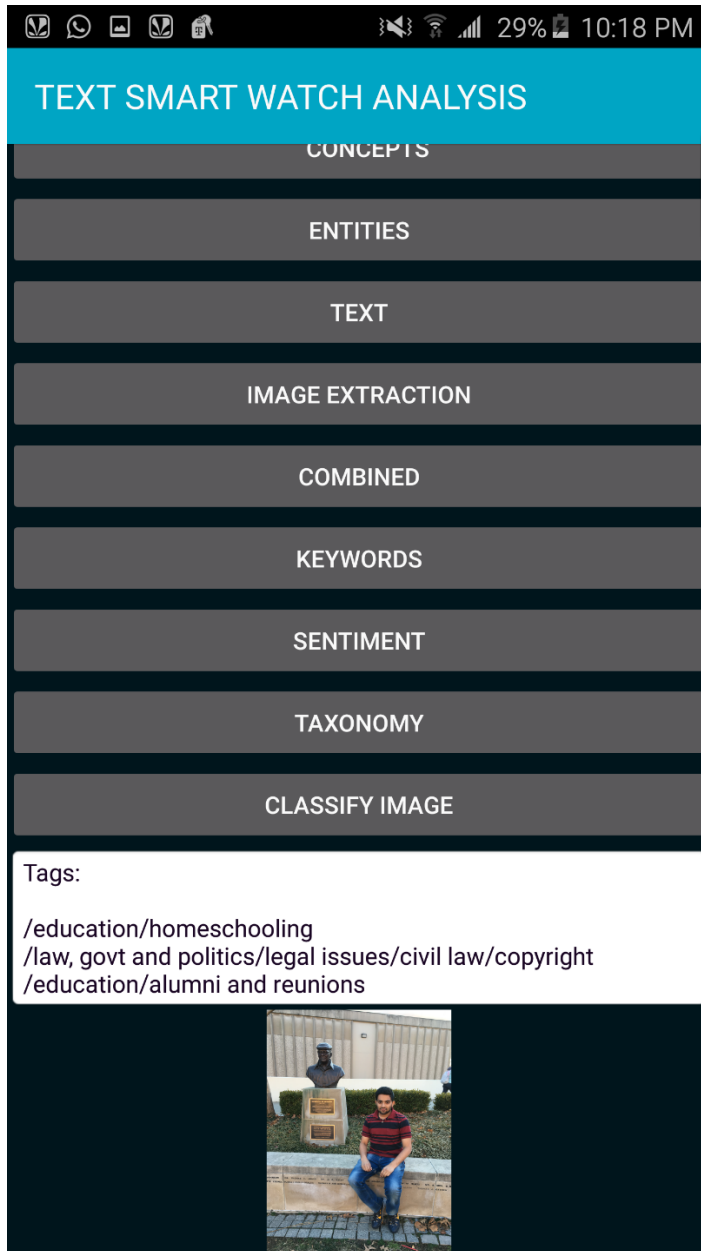
The above figure explains about the **keywords** button in the application.

6.



The above figure explains about **sentimental analysis** in application.

8.



The above figure illustrates about the **taxonomy** of the URL.

5. Project Management

Increment 1:

The following image shows how the image is categorized and the implementation using the Zenhub Tool.

The screenshot displays a GitHub repository page for 'rakeshreddybandi / Robo-smart-Secure-System'. The interface includes a search bar, navigation links for Pull requests, Issues, Gist, and To Do, and a repository overview section showing 4 Unwatch, 0 Star, and 0 Fork. The main content area features a Kanban board with five columns: New Issues (2), To Do (1), In Progress (8), Done (9), and Closed (2). Each column contains a list of issues with their respective labels and status. The issues are categorized by labels such as 'enhancement', 'help wanted', 'bug', and 'duplicate'. The board also includes a search bar and a 'New issue' button.

Column	Count	Issue #	Issue Title	Labels
New Issues (2)		#23	Robo-smart-Secure-System Cognitive analysis	
		#26	Robo-smart-Secure-System Connection between SmartWatch and Mobile Phone	bug, help wanted
To Do (1)		#24	Robo-smart-Secure-System Notification to Smart Watch	enhancement, help wanted
In Progress (8)			Application of machine learning algorithm to data set	enhancement, help wanted
		#19	Robo-smart-Secure-System Work on Event full API	
		#11	Robo-smart-Secure-System Sensor data collection	
			enhancement	
		#14	Robo-smart-Secure-System Working on Yelp API	enhancement, help wanted
Done (9)		2	enhancement help wanted	
		#5	Robo-smart-Secure-System Documenting project proposal	
		3	help wanted	
Closed (2)		#27	Robo-smart-Secure-System converting data collected into csv format	bug, help wanted
		#25	Robo-smart-Secure-System Test Smartwatch functionality	bug, duplicate, question
		3	enhancement help wanted	

The tasks in progress and the tasks to be done are basically brought together and will form a crucial part of the schedule and tasks are divided between members.

The screenshot displays the GitHub interface for the repository 'rakeshreddybandi / Robo-smart-Secure-System'. The URL in the browser is <https://github.com/rakeshreddybandi/Robo-smart-Secure-System/issues#boards?repos=50711311&milestone=Increment%201-2016-0>. The repository has 4 unwatchers, 0 stars, and 0 forks. The 'Boards' tab is selected, showing a Kanban board with the following columns:

- New Issues (2)**:
 - Issue #23: Robo-smart-Secure-System, Cognitive analysis (3 votes)
 - Issue #26: Robo-smart-Secure-System, Connection between SmartWatch and Mobile Phone (2 votes, bug, help wanted)
- To Do (1)**:
 - Issue #24: Robo-smart-Secure-System, Notification to Smart Watch (4 votes, enhancement, help wanted)
- In Progress (8)**:
 - Issue #17: Robo-smart-Secure-System, Work on Geo Location API (3 votes, enhancement)
 - Issue #12: Robo-smart-Secure-System, Application of machine learning algorithm to data set (4 votes, enhancement, help wanted)
 - Issue #19: Robo-smart-Secure-System, Work on Event full API (3 votes, help wanted)
 - Issue #11: Robo-smart-Secure-System, Sensor data collection (3 votes, help wanted)
- Done (9)**:
 - Issue #6: Robo-smart-Secure-System, Refine and rebuild project proposal (2 votes, enhancement)
 - Issue #22: Robo-smart-Secure-System, Modular Description (2 votes, enhancement, help wanted)
 - Issue #5: Robo-smart-Secure-System, Documenting project proposal (3 votes, help wanted)
 - Issue #16: Robo-smart-Secure-System, Requirement analysis (3 votes, help wanted)
- Closed (2)**:
 - Issue #27: Robo-smart-Secure-System, converting data collected into csv format (4 votes, bug, help wanted)
 - Issue #25: Robo-smart-Secure-System, Test Smartwatch functionality (4 votes, bug, duplicate, question)

The screenshot displays a GitHub repository page for 'Robo-smart-Secure-System' by user 'rakeshreddybandi'. The page features a Kanban board with five columns: 'New Issues' (2 items), 'To Do' (1 item), 'In Progress' (8 items), 'Done' (9 items), and 'Closed' (2 items). Each column contains task cards with details such as issue numbers, titles, and labels like 'enhancement', 'bug', and 'help wanted'. The interface includes a search bar, navigation links for Pull requests, Issues, Gist, and To Do, and a repository overview section with statistics like 'Unwatch', 'Star', and 'Fork'.

Column	Count	Task Details
New Issues	2	<ul style="list-style-type: none"> #23: Robo-smart-Secure-System Cognitive analysis (3 labels) #26: Robo-smart-Secure-System Connection between SmartWatch and Mobile Phone (2 labels: bug, help wanted)
To Do	1	<ul style="list-style-type: none"> #24: Robo-smart-Secure-System Notification to Smart Watch (4 labels: enhancement, help wanted)
In Progress	8	<ul style="list-style-type: none"> #4: Robo-smart-Secure-System Analyzing the machine learning algorithms (1 label: help wanted) #20: Robo-smart-Secure-System Architecture design (4 labels: enhancement, help wanted) #21: Robo-smart-Secure-System Working on Weather API (2 labels: enhancement, help wanted) #17: Robo-smart-Secure-System Work on Geo Location API
Done	9	<ul style="list-style-type: none"> #2: Robo-smart-Secure-System Project Proposal Discussion (3 labels: enhancement) #7: Robo-smart-Secure-System Knowledge discovery of smarwatch (2 labels: enhancement, question) #3: Robo-smart-Secure-System Finalizing project plan (3 labels: enhancement) #10: Robo-smart-Secure-System Research on APIs
Closed	2	<ul style="list-style-type: none"> #27: Robo-smart-Secure-System converting data collected into csv format (1 label: bug, help wanted) #25: Robo-smart-Secure-System Test Smartwatch functionality (4 labels: bug, duplicate, question)

Therefore all the tasks have been listed and they coercively form a part of the schedule and member assigned tasks are copulated together.

- Implementation status report:

a. Work completed

Member	Task Description	Member Responsibility	Contribution %	Time (hours)	Comments /Issues
Rakesh	1. Collect data Smart Watch to phone. 2. Machine Learning Naïve Based Algorithm Implementation. 3. Alchemy API –Sentimental Analysis, Taxonomy	Develop, design, build and testing the task	100	18	Running large data sets Machine Learning algorithms in Spark R have to be done
Nihar	1. Collect data Smart Watch to phone. 2. Alchemy API –Image Extraction, Concept Tagging 3. Yelp API	Develop, design, build and testing the task	100	20	
Mark	1. Notification Sending to Smart Watch 2. Machine Learning Naïve Based Algorithm Implementation. 3. Alchemy API –Entity Extraction, Keyword Extraction	Develop, design, build and testing the task	100	22	
Sricharan	1. Notification Sending to Smart Watch 2. Alchemy API –Combined call and Classify Image 3. Eventful API	Develop, design, build and testing the task	100	20	

b. Work yet to be completed

Member	Task Description	Member Responsibility	Time (hours)	Comments /Issues
Rakesh	1. Implementing Machine Learning Algorithms in Spark R 2. Sending notification of predicted data. 3. Predicting the user patterns regularly. 4. Implementing API's.	Develop, design, build and testing the task	40	
Nihar	1. Implementing Machine Learning Algorithms in Spark R. 2. Collecting more sensor data sets. 3. Predicting the user patterns regularly. 4. Implementing API's.	Develop, design, build and testing the task	40	
Mark	1. Implementing Machine Learning Algorithms in Spark R 2. Sending notification of extracted data 3. Improving GUI. 4. Predicting the user patterns regularly.	Develop, design, build and testing the task	40	
Sricharan	1. Implementing Machine Learning Algorithms in Spark R 2. Sending notification of extracted data. 3. Testing the implemented features. 4. Predicting the user patterns regularly.	Develop, design, build and testing the task	40	

6. Bibliography:

- i. Komninos, Andreas, and Mark Dunlop. "Text input on a smart watch." Pervasive Computing, IEEE 13.4 (2014): 50-58.
- ii. Lutze, Rainer, and Klemens Waldhor. "A Smartwatch Software Architecture for Health Hazard Handling for Elderly People." Healthcare Informatics (ICHI), 2015 International Conference on. IEEE, 2015.
- iii. <http://www.motorola.com/us/products/moto-360>
- iv. <https://www.thegrommet.com/robome-app-enabled-robot>
- v. <https://www.yelp.com/developers/documentation/v2/overview>
- vi. <https://api.eventful.com/>