Analysis of Google Play store review

Sanket Prabhu: srp140430
Chirag Chudasama: cbc140130
Kunal Kapoor: kxk140330

Abstract:

Developing an app in today's smartphone ecosystems is easier than ever - the barrier of entry is low enough that a single developer can write a commercially successful app. The iOS App Store is reported to have reached the one million apps mark. As of October 2013, 50B apps have been downloaded from the app store. Furthermore, revenue earned by developers from the iOS App Store have been over $1.5B in 2012. Customer interest in apps is evident: top apps in the Google Play app store have over 1m reviews.
As a developer, it is essential to "stay on top of your game", i.e., keep your app updated with the most requested features and bug-fixes. However, most app stores provide only an average rating (out of 5) for each app. Consequently, it is difficult to identify why people like or dislike a particular. We aim to solve this problem.

Problem Statement:

The aim of this project is to collect reviews from Google Play Store and preprocess it and analyzing it to give top positive as well as top negative reviews about particular app, and to provide more detailed analyses of the reviews of an app to a developer, beyond an average rating.

System Overview:

1. To get review about particular app.

We got dataset of reviews. The datasets which we are using (2014):

a. Audible
b. Flixster
c. Jewels
d. Operamini
e. Pandora
f. Soundhound
g. Yelp

2. Processing reviews.

A review consists of the following relevant data:

a. A unique author ID,
b. Review Creation Time
c. Rating (ranging from 1 to 5)
d. Review Text

We normalize the ratings to a scale of [0, 1].
Expectedly, not all reviews possess perfect grammar or punctuation, rendering preprocessing a must. To begin, text for each review is transformed to lowercase. Subsequently, emoticons are converted into symbols. Emoticons can potentially prove to be strong indicators of opinion, and thus, are valuable. Now, exclamation marks are transformed into symbols, since we believe that they, too, can prove to be strong indicators of opinion. Finally, the review text is stripped of punctuation, and replaced by whitespace.
Image which is shown below states that we have replaced smileys with words

```
':)':    'smile',
':-)':   'smile',
':]':    'smile',
':-]':   'smile',
'(:':    'smile',
'(-:':   'smile',
';)':    'wink',
';-)':   'wink',
':P':    'cheeky',
':-P':   'cheeky',
';P':    'cheeky',
';-P':   'cheeky',
';D':    'laugh',
';-D':   'laugh',
':D':    'laugh',
':-D':   'laugh',
':(':    'sad',
':-(':   'sad',
';(':    'sad',
';-(':   'sad',
'):':    'sad',
')-:':   'sad',
':[':    'sad',
':-[':   'sad',
':/':    'wry',
':-/':   'wry',
';/':    'wry',
';-/':   'wry',
':\\':   'wry',
':-\\':  'wry',
';\\':   'wry',
';-\\':  'wry',
':\'(':  'cry',
':\'-(': 'cry',
'<3':    'love',
'>:(':   'angry',
'>:-(':  'angry',
'>;(':   'angry',
'>;-(':  'angry',
'O.O':   'bigeyes',
'o.O':   'bigeyes',
'O.o':   'bigeyes',
'-.-':   'squint'
```

List of stop words we are using are shown below:

Source: http://xpo6.com/list-of-english-stop-words/

```
a
about
above
across
after
afterwards
again
against
all
almost
alone
along
already
also
although
always
am
among
amongst
amoungst
amount
an
and
another
any
anyhow
anyone
anything
anyway
anywhere
```

3. Creating a Corpus of Subjective Words

We hypothesize that the opinion of a review is indicated by the presence of "subjective words". For example, a review is more likely to complain about crashes, and have a low rating, if it contains the word "crash". Thus, we intend to create a corpus of such "subjective" words.
We also hypothesize that a subjective word is indicative of a high or low rating.
For example, a review with the word "crash" in it is likely to have a low rating, whereas one with "stable" in it is likely to have a high rating.
Consequently, we will extract subjective words from reviews by analyzing their power to predict the rating of a review. Specifically, we are planning to apply logistic regression to review ratings.

4. Identifying Topics in a Review

Given a review text, a list of subjective words, and a list of topics that each subjective word corresponds to, the topics addressed by the review are the union of topics for each word in the review.

5. Ranking Topics in Order of Relevance

Given a list of topics, and a list of reviews associated with each topic, one can compute various metrics for each topic, such as average rating, number of reviews, etc. Consequently, one can rank the topics in order of relevance to a single or hybrid metric. For example, a hybrid metric that ranks topics in increasing order of average rating and decreasing order of number of reviews provides a list of topics that customers are most unhappy about. Similarly, hybrid metrics can be designed to provide topics that customers are most happy about, or write the longest reviews about, etc.

6. Identifying Representative Reviews for a Topic.

For each topic, we are planning to identify a list of representative reviews that summarize public sentiment on that topic. This is achieved by ranking all reviews relevant to a topic according to some metric and picking the top 3. We will chose a metric that ranks reviews in proportion to the number of times a word from the topic set appears in the review, and inversely proportional to the squared distance of the review's rating from the average rating for that topic.

Output:

We started out with the goal of providing a more detailed analyses of the reviews of an app to a developer, beyond an average rating.
We will be able to provide a list of topics that are relevant to the manner in which the developer wishes to interpret the reviews, an average rating that conveys general sentiment for that topic, and representative reviews that give insightful criticism regarding the topic.

ProcessedReviews.txt

ProcessedReviews.txt - Notepad

File Edit Format View Help

{'soundhound': [{'text': 'nice sometimes help me to find the song must have this app', 'rating': 1.0}, {'text': 'dont want this never works and i cant get rid of this useless program i have never once got a single song name ever i wish i was not forced to have this on my phone and to have it default set to start up when my search button is held down is unnecessary ', 'rating': 0.0}, {'text': 'stopped working dosent work anymore', 'rating': 0.5}, {'text': 'amazing EXCLAMATION this acc works even with humming EXCLAMATION ', 'rating': 1.0}, {'text': 'simply amazing EXCLAMATION my only problem is having used it so much for a couple years now i have too many sound searches and bookmarks i wish i could export them to the sd card and reclaim some space 16 mb on my phone that is now into the low space warning level 42 mb and no longer syncs gmail etc ', 'rating': 1.0}, {'text': 'doesnt work for telugu songs', 'rating':

FeatureWeights.txt

FeatureWeights.txt - Notepad

File Edit Format View Help

{'megaplex': 0.00037336324806314114, 'awesome': 0.7355660196008285, 'supet': 0.0012683084297670615, 'upto': 0.002254126097661932, 'tkikmoto': 0.0008306854099173308, 'sorta': -0.0005027780124755834, 'news': -0.002206490727727751, 'guy': -0.0009316047898673874, 'iit': 0.00038701751602249055, 'controlling': -0.0013731038896422569, 'grail': -0.0009531905926583916, 'quirk': 0.0004201920858945851, 'gs4': -0.0005050039459732899, 'nmore': 0.0008008852939371091, 'fgfhfg': 5.656093856286184e-07, 'lovein': 0.0007995501388486683, 'enormous': -0.00011795472287954915, 'shiny': -0.0005549228625156984, 'adictive': 0.0010987799188424238, 'maryj': 0.00031963122834456555, 'med': -0.0003805906814002371, 'chances': -0.0007111483873251186, 'fianlly': 0.00021430711873138495, 'managing': 0.0004534330492830116, 'expertly': 0.00023071369868512642, 'snaps': -0.0005190458638995153, 'vish':

Creating a Corpus of Subjective Words

The first step in our pipeline – linear regression – gives us the basic ability to predict the sentiment of a review. This is useful to identify features in a review that are good indicators of sentiment. Our values for the parameters of linear regression were: $\alpha\alpha$=0.0001, $\epsilon\epsilon$=0.001, $\boldsymbol{\phi\phi}$ is the presence-indicating feature vector of all unigrams with stop words filtered out.

The top positive words were:
love, great, good, awesome, best, !, excellent, app, nice, game, browser, cool, fast, easy, works, fun, amazing, movie, opera, addictive, perfect, music, movies, awesome, super, helpful, fantastic, better, jewels.
The top negative words were:
update, work, open, sucks, will, phone, uninstall, ads, play, bad, songs, books, poor, crap, book, crashes, useless, screen, uninstalled, force, terrible, download, 3, load, horrible, uninstalling, start, takes, waste, annoying.

Identifying Topics in a Review

Topic inference in a document is a well-studied task. However, most approaches assume a structured document with well-formed grammar and only the occasional error. However, most reviews we encountered showed poor structure, spelling and syntax. This rendered NLP techniques impossible. Consequently, we fell back to a simple technique – finding a group of subjective words with similar meanings in a review, and interpreting this set of words as a topic of discussion.
Using our technique, we obtained an average of 2000 topics per app. Note that this list contains a large spectrum of all possible topics, ranging from "ads", "bug" or "crash" all the way to "fine", "okay", etc. In the next step, we rank topics according to their relevance.

In our results, we noticed that not all topic sets are obviously a set of synonyms. For example, the set ('bring', 'bringing', 'brings', 'brought', 'play', 'played', 'playing', 'plays', 'work', 'worked', 'working', 'works') does not obviously contain pairwise synonymous words. However, one must realize that a word can be used in many senses, and can imply different meanings. In the absence of contextual meaning of a word, one must assume every meaning to be possible, leading to a certain level of ambiguity. This might also lead to two or more independent topics fused together under the same topic. However, this issue is not prevalent or alarming: the example above is the worst-case scenario.

Another issue we noticed was some topics were similar for the most part, except for one or two words, and hence appeared together in rankings everywhere. However, this issue was not too prevalent either, and hence, was deferred.

Synsets.txt

Synsets.txt - Notepad

File Edit Format View Help

```
[('motel', 'motels'), ('alternate', 'alternative', 'substitute'), ('asked', 'asks', 'demand', 'demands', 'involve', 'involved', 'involves', 'needed', 'neededs', 'needing', 'require',
'required', 'requires', 'requiring', 'takes'), ('drop', 'dropped', 'dropping', 'drops', 'spend', 'spending', 'spent'), ('subject', 'theme', 'themes', 'topic', 'topics'), ('reach',
'reached', 'reaching', 'touch', 'touched', 'touches', 'touching'), ('lines', 'pipeline'), ('mobility',), ('phrase', 'word', 'wordings'), ('brand', 'trademark'), ('fancy', 'fantasies',
'fantasy'), ('follow', 'watch', 'watched', 'watches', 'watching'), ('paint', 'rouge'), ('crossing', 'ford'), ('confused', 'lost'), ('irreplaceable',), ('mowing',), ('branched', 'fork',
'separate', 'separated', 'separating'), ('identified', 'identifies', 'identify', 'identifying', 'named', 'names', 'naming'), ('literature',), ('abusive',), ('overdue',), ('annoyance',
'annoyances', 'annoying'), ('tomahawk',), ('debacle',), ('pound', 'pounds', 'quid'), ('99',), ('hydraulic',), ('numerous',), ('audibility',), ('adobe',), ('steal', 'steals', 'stolen'),
```

Ranking Topics in Order of Relevance

The relevance of a topic depends on the desideratum of the developer. A developer may desire to look at topics with negative sentiments, positive sentiments, or topics that are popular, or any other metric.

We defined metrics for some such use-cases. For example, topics with negative sentiments can be obtained by ranking them proportional to the number of relevant reviews, and inversely proportional to the average rating. This metric is not entirely statistically sound – any topic that has enough reviews will have a mean close to the average rating – however, it works well in practice.
For Pandora, the top negative topics we found were: ('garbage', 'refuses'), ('pathetic', 'poor'), ('email', 'emailed', 'emails'), etc.

In practice, we found another heuristic extremely effective: topics ranked such that they are far from the extremes in terms of ratings, and are far from the app average, and have a sizeable number of reviews, provide an alternate view of the public sentiment that is fairly critical and insightful. For example, for Pandora, the top topics according to this metric were: ('play', 'played', 'playing', 'plays',…), ('ad', 'ads', 'advertisement', 'advertisements', 'advertising', 'adverts'), ('call', 'song', 'songs'), etc. We believe that this metric provides good results because:

(a) Most reviews that provide a 1 star / 5 star rating do so without providing any insightful criticism. On the other hand, reviews with 2-4 star ratings provide more detailed criticism. One can hypothesize that a 2-4 rating indicates the author put thought into the rating before choosing an outright good/bad (leading to a 5/1 star rating respectively) distinction, and this same thought is usually reflected in the review itself.

(b) If the average rating is high (as is the case with all the apps we happened to choose), then most reviews close to the app average, again, fail to criticize the app properly.

Other means of ranking topics could be conceived of, such as measuring the "insightfulness" of a review, and computing the average for a topic, however, such avenues were deferred. Additionally, we would like replace the use of a heuristic with that of a learning technique. However, in the face of lack of labeled data, we are unable to do so.

Output for Yelp dataset:

Positive Reviews

```
Positive metrics for yelp:
Average rating: 0.819166666667
('EXCLAMATION',)
        'you EXCLAMATION yes you EXCLAMATION you need this app EXCLAMATION i absolutely love this app EXCLAMATION it has helped me find new places to go new things to try numbers to ca
        'i love yelp EXCLAMATION it is the best way to find anything youre looking for EXCLAMATION if youre a local youll still find things and share with people who are new all the gr
        'yay EXCLAMATION mobile reviews EXCLAMATION yay EXCLAMATION finally mobile reviews EXCLAMATION increasing from 1 star to 5 so excited EXCLAMATION '
        'best app EXCLAMATION we never try anything new but with honest reviews we have branched out EXCLAMATION plus i review like crazy EXCLAMATION not just food but everything EXCLA
    1328 reviews, 0.877635542169 avg, 0.466599070328 min, 0.466599070328 max
('helpful',)
        'so helpful so helpful when im traveling and when im looking for new local places to try '
        'really helpful ive recently moved in a new area and its really helpful for me love it EXCLAMATION '
        'easy to use and very helpful EXCLAMATION very helpful app with good information its also easy to use EXCLAMATION '
        'solid i like this app because its helpful and people who make the reviews are generally honest and helpful '
    324 reviews, 0.908950617284 avg, 0.149598705906 min, 0.149598705906 max
('great', 'greater', 'greatest', 'outstanding')
        'i love yelp EXCLAMATION it is the best way to find anything youre looking for EXCLAMATION if youre a local youll still find things and share with people who are new all the gr
        'great job good job on this app one love had this app for a long time running droidx keep up the great work now its on my gs4 great job'
        'great app thanks its great has everything i need to find a place couldnt ask for better EMOTICONsmile great app keep the good work \xe2\x99\xa5\xe2\x99\xa1\xe2\x99\xa5'
        'great app for various things not only for restaurants highly recommend great app for various things not only for restaurants highly recommend its opened up new places top east
    1332 reviews, 0.913475975976 avg, 0.000833841802898 min, 0.542639719994 max
('capital', 'great', 'greater', 'greatest')
        'i love yelp EXCLAMATION it is the best way to find anything youre looking for EXCLAMATION if youre a local youll still find things and share with people who are new all the gr
        'great job good job on this app one love had this app for a long time running droidx keep up the great work now its on my gs4 great job'
        'great app thanks its great has everything i need to find a place couldnt ask for better EMOTICONsmile great app keep the good work \xe2\x99\xa5\xe2\x99\xa1\xe2\x99\xa5'
        'great app for various things not only for restaurants highly recommend great app for various things not only for restaurants highly recommend its opened up new places top east
    1328 reviews, 0.913780120482 avg, 9.76889382194e-05 min, 0.542639719994 max
('excellent',)
        'excellent application excellent application it is a boon to find out almost anything you are looking for '
        'excellent compare reviews and find out how that place you wanted to go to works even before going there excellent concept '
        'excellent excellent'
        'word of mouth from around your area in an app EXCLAMATION yelp has been nothing but great since i first started using it excellent way to share my current location with friend
    150 reviews, 0.953333333333 avg, 0.0829713412532 min, 0.0829713412532 max
('best', 'scoop')
        'looooove best app good for finding a yummy place to eat on a trip or finding the best of the best in your city '
        'the best of the best EXCLAMATION just download it dammit yelp is the best'
        'you have to get this app EXCLAMATION if youre a foodie like me and you need to find a restaurant that satisfies your tastebuds or if youre just looking for the best business f
        'best app for finding the best of '
    271 reviews, 0.960332103321 avg, 0.00169554253555 min, 0.147047059696 max
('dandy', 'great', 'greater', 'greatest', 'neat')
        'i love yelp EXCLAMATION it is the best way to find anything youre looking for EXCLAMATION if youre a local youll still find things and share with people who are new all the gr
        'great job good job on this app one love had this app for a long time running droidx keep up the great work now its on my gs4 great job'
        'great app thanks its great has everything i need to find a place couldnt ask for better EMOTICONsmile great app keep the good work \xe2\x99\xa5\xe2\x99\xa1\xe2\x99\xa5'
        'great app for various things not only for restaurants highly recommend great app for various things not only for restaurants highly recommend its opened up new places top east
```

Negative Reviews

```
Negative metrics for yelp:
Average rating: 0.819166666667
('backing', 'support', 'supporting', 'supports')
        'yelp EXCLAMATION has joined alec the promoter of stand your ground aka shoot first laws that legalize murder on the basis of racial fear the quality of the product is immateri
        'yelp inc supports lots of offensive political groups like alec so i uninstalled rather than support them '
        'yelp supports alec im uninstalling the app yelp has joined the american legislative exchange council a freedom hating organization which writes and pushes legislation on behal
        'yelp supports alec adios yelp your app was uninstalled this morning because of your support of alec ill use google and facebook from now on if i want relevant reviews '
    24 reviews, 0.114583333333 avg, -0.0155933424167 min, -0.000943560430294 max
('useless',)
        'useless in europe yelp is useless in europe the information is alway wrong and many venues are simply missing on many ocasions i went to some place that suposedly closed at 2a
        'a nice app no longer connects a mobile app that doesnt connect is well it can be useless a recent update has caused this app to fail to connect either using wifi nor my phones
        'slow to adopt this is probably applicable only in newly added countries i ve been trying to change address of my business for about 3 months upon giving up i checked my yelp m
        'useless app try tripadvisor or google'
    21 reviews, 0.142857142857 avg, -0.0151877045277 min, -0.0151877045277 max
('stupid',)
        'posting reviews as drafts is stupid write up your review and then its only a draft until you log in on the website to submit stupid '
        'so stupid EXCLAMATION whats the point of having the app if after i write a review i have to turn around and log on to yelp from my computer to post it anyway incredibly stupid
        'used to love it geez whats with everything forcing you to link upload contacts etc with social networking leave me alone EXCLAMATION if i want to then let me choose EXCLAMATIC
        'app wont let you review dumb app wont let you review places on the go stupid '
    12 reviews, 0.145833333333 avg, -0.00880979136521 min, -0.00880979136521 max
('poor', 'short')
        'search engine is poor several times ive searched and it ranks the same restaurant in another town before the town that im in also filters could use some beefing up to be pract
        'unlike most reviews on yelp i didnt get paid to give it a poor rating this app burns through my battery and locks to gps signals when it shouldnt also my results are always ex
        'good app but app is awesome but yelp as a review site is very poor and unfair to businesses and consumers'
        'poor gps implementation makes the app irrelevant the gps implementation is poor and offers me results 10 miles away what is unusual is that the app knows the locations are not
    23 reviews, 0.184782608696 avg, -0.015074600033 min, -0.0013143167642 max
('despicable', 'ugly', 'worthless')
        'no tablet interface in bed with apple yelp should be ashamed at having such an ugly and unoptimized app out this long does yelp really want apple to have a monopoly it would s
        'totally worthless cant write reviews at all it saves them to draft and tells me i have to log in to the website instead but there is no draft there so why would i use the app
        'wont even let me sign up makes this app worthless'
        'still worthless cant search more than 5 miles away i live 10 miles from the nearest town so i cant search for anything without driving closer to town filters have to be reset
    12 reviews, 0.208333333333 avg, -0.00804270587181 min, 1.76806875292e-05 max
('held', 'hold', 'holding', 'support', 'supporting', 'supports')
        'no landscape support for tablets and no way to turn off gps polling its pretty ridiculous that there still isnt landscape support for tablets after all this time also there sh
        'please provide support for more countries id so live this app if i could use it in my current location back in the usa it turned out to be one of my most used apps but now tha
        'yelp EXCLAMATION has joined alec the promoter of stand your ground aka shoot first laws that legalize murder on the basis of racial fear the quality of the product is immateri
        'yelp inc supports lots of offensive political groups like alec so i uninstalled rather than support them '
    30 reviews, 0.216666666667 avg, -0.0155933424167 min, 0.000323706030168 max
('confirm', 'confirmed', 'support', 'supporting', 'supports')
        'no landscape support for tablets and no way to turn off gps polling its pretty ridiculous that there still isnt landscape support for tablets after all this time also there sh
        'please provide support for more countries id so live this app if i could use it in my current location back in the usa it turned out to be one of my most used apps but now tha
        'yelp EXCLAMATION has joined alec the promoter of stand your ground aka shoot first laws that legalize murder on the basis of racial fear the quality of the product is immateri
        'yelp inc supports lots of offensive political groups like alec so i uninstalled rather than support them '
```

Conclusion

To summarize our work, we started out with the goal of providing a more detailed analyses of the reviews of an app to a developer, beyond an average rating. We are able to provide a list of topics that are relevant to the manner in which the developer wishes to interpret the reviews, an average rating that conveys general sentiment for that topic, and representative reviews that give insightful criticism regarding the topic.

We can improve the manner in which we infer topics from a review by applying NLP techniques on structured data, and the manner in which we signify relevance of topic or representativeness of a review using labeled data.

Software/Tools?

- Python IDE
- Visual Studio

References:

[1] C. Jones, "Apple's App Store About To Hit 1 Million Apps," 11 12 2013. [Online]. Available: http://www.forbes.com/sites/chuckjones/2013/12/11/apples-app-store-about-to-hit-1-million-apps/. [Accessed 12 12 2013].

[2] R. Baldwin, "Apple Hits 50 Billion Apps Served," 15 05 2013. [Online]. Available: http://www.wired.com/gadgetlab/2013/05/apple-hits-50-billion-served/. [Accessed 12 12 2013].

[3] Canalys, "11% quarterly growth in downloads for leading app stores," 08 04 2013. [Online]. Available: http://canalys.com/newsroom/11-quarterly-growth-downloads-leading-app-stores. [Accessed 12 12 2013].

[4] L. Zhuang, F. Jing and X.-Y. Zhu, "Movie review mining and summarization," in Proceedings of the 15th ACM international conference on Information and knowledge management (CIKM '06), ACM, New York, NY, USA, 2006.

[5] H. Tang, S. Tan and X. Cheng, "A survey on sentiment detection of reviews," Expert Systems with Applications, vol. 36, no. 7, pp. 10760-10773, 2009.

[6] "android-market-api," 04 11 2013. [Online]. Available: http://code.google.com/p/android-market-api/.