Feature-level Rating System using Customer Reviews and Review Votes

Koteswar Rao Jerripothula, Member, IEEE, Ankit Rai, Kanu Garg, and Yashvardhan Singh Rautela

Abstract—This work studies how we can obtain feature-level ratings of the mobile products from the customer reviews and review votes to influence decision making, both for new customers and manufacturers. Such a rating system gives a more comprehensive picture of the product than what a product-level rating system offers. While product-level ratings are too generic, feature-level ratings are particular; we exactly know what is good or bad about the product. There has always been a need to know which features fall short or are doing well according to the customers’ perception. It keeps both the manufacturer and the customer well-informed in the decisions to make in improving the product and buying, respectively. Different customers are interested in different features. Thus, feature-level ratings can make buying decisions personalized. We analyze the customer reviews collected on an online shopping site (Amazon) about various mobile products and the review votes. Explicitly, we carry out a feature-focused sentiment analysis for this purpose. Eventually, our analysis yields ratings to 108 features for 4k+ mobiles sold online. It helps in decision making on how to improve the product (from the manufacturers perspective) and in making the personalized buying decisions (from the buyers perspective) a possibility. Our analysis has applications in recommender systems, consumer research, etc.

Index Terms—Recommender systems, natural language processing, sentiment analysis, cellular phones, reviews, decision making, text mining, web mining.

I. INTRODUCTION

With the rise of the internet and the kind of busy lifestyles people have today, online shopping has become a norm. Customers often rely on the online ratings of the previous customers to make their decisions. However, most of these ratings on the online websites are product-level ratings and lack specificity. Although products can be compared based on the product-level ratings available, there is always a class of people who prefer buying the items based on particular features. Such people have to generally go through the entire comments section to know previous customers perceptions of the products features in which they are interested. Considering the number of products present for an item (such as mobile), it becomes a tedious job for a customer to arrive at the best product for himself. Moreover, from a manufacturer’s perspective, such product level ratings hardly specify what is good or bad about the product. So, if feature level ratings are available, it gives more clarity to the seller on how to improve the product. Given all these benefits, our goal is to develop a feature-level rating system.

Although feature-level ratings can also be requested from the customers just like the product-level ratings, it is not a good proposal, for there could be too many features. Instead, it is much more practical to leverage whatever reviews and review votes that are already given by customers to provide feature-level ratings. The reviews are made up of sentences, and every sentence has some sentiment associated with it, viz. positive, neutral, or negative. Also, since they can be separated, we can always extract the sentences describing a particular feature of the product and subsequently obtain sentiment scores over such sentences. By utilizing these sentiment scores as the basis and the review votes as a support, we can build a feature-level rating system that can yield feature-level ratings, as shown in Fig. 1.

However, there are few challenges in building such a feature-level rating system. First, we need to determine which features to look for in an item. Another challenge is that there could be many words relating to the same feature; they all need to be clubbed into one feature. Second, we have to pre-process the data as some of the review comments may contain non-English languages, one word, spelling mistakes, etc. Third, we have to devise a way to transform the extracted sentiment scores into an appropriate rating for a feature of a product while incorporating the review votes.

As far as feature identification of an item (such as mobile) is concerned, we go through the word frequency table of the entire customer review data for that item, at least up to a particular frequency. Next, all the related words coming under the same feature are grouped, and the most frequent one is chosen as a representative. We call such representatives as feature keywords. Then, we perform a series of pre-processing

![Product-level Rating versus Feature-level Ratings](image-url)
steps to filter out the unnecessary data, correct the remaining, and turn it into structured data. Each review is broken into sentences, and only relevant sentences are retained. The relevant sentences are passed through sentiment analyzer to generate sentiment scores, which are then adjusted to the ratings. Scores within a particular range are given a specific rating. The ratings of the relevant sentences containing a particular feature are combined using the weighted-average \cite{10} to obtain the final ratings since all opinions are not equally valuable. We leverage review-votes to assign the required weights.

Our contributions are as follows: We develop a feature-level rating system that takes customer reviews and review votes as inputs and outputs feature-level ratings. We obtain such ratings for as many as 4k+ mobiles sold online in terms of as many as 108 features. We propose votes-aware cumulative rating and votes-aware final rating measures, a new way of accumulating and finalizing the sentiment scores. Although there are no ground truths available, we still manage to evaluate our approach by comparing the final ratings of our phone feature against overall ratings of the phone given by the customers themselves, which leads to remarkable results demonstrating the effectiveness of our method.

II. RELATED WORK

Sentiment analysis \cite{11}–\cite{18} has been an active research topic for a long period now. It has applications in health \cite{19}–\cite{21}, politics \cite{22}, \cite{23}, sports \cite{24}, \cite{25}, e-commerce \cite{26}, \cite{27}, etc. In e-commerce, customer reviews can give lots of insights about the products, as shown in \cite{28}, \cite{29}, through sentiment analysis. Specifically, \cite{11} studies trends of mobile brands on Twitter through sentiment analysis. However, the analysis is restricted to the mobile overall, not to specific features. \cite{30} tries to do so, but for limited products and limited features. While \cite{30} used SVM, a supervised learning algorithm, \cite{31} used ensembling for achieving this. In this paper, we attempt to exploit these customer reviews to provide ratings for as many as 108 features of 4k+ mobile phones sold online while incorporating review votes, which has never been done in the previous studies. Moreover, we do this in an unsupervised way, not supervised or weakly-supervised \cite{32}–\cite{35} way, thanks to the lexical approach of generating sentiment scores for a sentence. Similar to our work, \cite{36} explored digital cameras and television for the same. However, they explore only ten features. \cite{37} explored only Cannon Camera, iPhone 4s, and Mp3 player. However, we explore as many as 4k+ products and provide our recommendations. Nevertheless, to the best of our knowledge, this is the first paper to account for review votes in a feature-level rating system.

III. METHODOLOGY

The proposed method has four steps: (1) feature selection; (2) pre-processing; (3) relevant sentence extraction; (4) feature-based rating generation. Every step is described in detail in this section.

| Purpose | Characters |
|---------|------------|
| Word Formation | A-Za-z |
| Punctuation | ;:,.? (space) |
| Emotions | "<\&"|

A. Feature Selection

Let us say we collect a dataset of \( N \) feature-related words, denoted by \( W = \{w_1, \ldots, w_N\} \), by manually going through word frequency table of the entire customer review data on an item (mobile, in our case). In this way, we identify the features in which people are generally interested. Note that we neglect the words having their frequency less than 0.02% of the total number of reviews in the review data, which means they are rarely discussed feature-related words and can, therefore, be neglected. Let us say the corresponding frequencies of the feature-related words form another set denoted as \( \mathcal{Z} = \{z_1, \ldots, z_N\} \). Since the feature-related words related to a particular feature should be clubbed into one feature, we define a relationizer function denoted as \( \mathcal{R}(W, w_i) \), which returns a set of all the related words of \( w_i \) in \( W \), including itself. Note that the relationizer function discussed here as a matter of notation is manual. We now define our feature dataset, denoted by \( \mathcal{F} \), as a set of such distinct sets of related feature words, as defined below:

\[
\mathcal{F} = \{\mathcal{R}(W, w_i)\}_{i \in \{1, \ldots, N\}},
\]

where we iterate through all the words in \( W \) and form distinct sets of the related words using the relationizer function. Since different related words will form the same sets, the duplicates will be removed to make the sets left distinct. Now, let \( \mathcal{F}_k \) be the \( k^{th} \) feature words set in the \( \mathcal{F} \) feature dataset. In any \( \mathcal{F}_k \), a representative feature word is selected to identify the whole feature words set. Let us call such representatives as feature keywords. The most frequent feature word in the set is chosen as the representative (inspired by \cite{38}) or feature keyword to assign its name to the set, as shown below:

\[
\mathcal{F}_k \leftarrow w_i | i = \max \{\mathcal{Z}(i) | w_i \in \mathcal{F}_k\},
\]

where the feature set is assigned a name with the word that has the maximum frequency in the set. From now on, abusing notations a bit, \( \mathcal{F}_k \) can mean both the \( k^{th} \) feature set (a set of related feature words) and its keyword (or set’s name), according to the context.

B. Pre-processing

The review comments data is generally unstructured, for it is written by the customers online. Our goal now is to convert this unstructured data into structured data in our pre-processing steps, which means useful data is extracted, disintegrated, and corrected.

Note that the data retained after removing the characters that are useless for our purpose is what we mean as useful data. While inspecting the reviews for figuring out the features to
I have had the phone for seven days now and I am impressed. Its setup is very simple. The screen quality is great, and the processor is fast...

Fig. 2. Pre-processing: Useful Characters Retention, where unnecessary characters are removed in this illustration.

work with, we also noticed how people praise or criticize. People often use the characters required for adjective words, punctuation, or emoticons. While we retain the characters required for word formation, punctuation, and emoticons, we remove all other characters, including numbers, as shown in Fig. 2. In Table I, we give the information regarding what all characters are retained. After that, we remove any entries which are left empty because we have no use of them in the feature-level rating system. Let us consider that, for a product (not an item), we denote product review data as $D = \{C_1, C_2, \cdots, C_m\}$, comprising of $m$ useful review comments. Note that when we said customer review data of the item in the last section for feature selection, we meant review data of all the mobiles. In contrast, $D$ is the review data of just the mobile product under consideration.

By corrected data, we mean the data obtained after correcting the related feature words issue and spelling correction in the useful data just extracted. To correct the data in such a manner, we need to disintegrate the reviews into words and process them separately. We use the NLTK package of python for this purpose. It helps in disintegrating the reviews into words as tokens while neglecting the spaces. It considers even period (.) as a token, which becomes useful later while breaking the comments into sentences. Each comment can now be represented as a set of tokens, i.e., $C_j = \{t_j^1, t_j^2, \cdots, t_j^{|C_j|}\}$, where $|C_j|$ denotes the number of tokens obtained in $C_j$ and $t_j^i$ represents $i^{th}$ token of $j^{th}$ comment. We correct any token $t_j^i$ of the useful data in the following manner:

$$t_j^i = \begin{cases} \mathcal{F}_k, & \text{if } S(t_j^i) \in \mathcal{F}_k \text{ or } t_i \in \mathcal{F}_k, \forall \mathcal{F}_k \in \mathcal{F} \\ S(t_j^i), & \text{otherwise,} \end{cases}$$

where $S(\cdot)$ represents spelling correcting function (using autocorrect package of python). If a token before or after spelling correction matches with any of the members in any of our feature words sets, we replace it with the feature keyword of that set; otherwise, we replace it with the corrected token. In this way, we take care of both the related words issue (by replacing with keywords) and the spelling correction issue simultaneously. The illustrations of spelling and keywords correction are given in Fig. 3 and Fig. 4. Thus, with the useful data extracted, disintegrated, and corrected, our reviews data for a product becomes structured. Now, we can say that $t_j^i$ is $i^{th}$ token of $j^{th}$ comment of $D$.

$X_j = \{(t_j^u, \cdots, t_j^v) \mid (t_j^u, t_j^{u-1}) = \text{', }, (t_j^v, \cdots, t_j^{v-1}) \neq \text{'}, (u, v) \in \{1, \cdots, |C_j|\} \text{ and } v > u\} \tag{4}$

where we call a group of continuous tokens as a sentence if the last and previous-to-beginning token are periods (.), and if all other tokens in that group are not periods (.) However, not all the sentences are relevant for feature-based rating. We define if a sentence $X_j^i$, $i^{th}$ sentence in $X_j$, is relevant or not in the following way:

$$\rho(X_j^i) = \begin{cases} 1, & \text{if } \mathcal{F}_k \in X_j^i \text{ for any } \mathcal{F}_k \in \mathcal{F} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where $\rho(\cdot)$ is relevance function for a sentence which outputs 1 if any of our feature keywords are present in the sentence.
D. Feature-based Ratings Generation

Having extracted relevant sentences, we can go through each sentence to figure out if it mentions a particular feature, say $F_k$. If yes, we can extract the emotion of the sentence to score it. For this purpose, we extract sentiment analysis scores [39] for each of these sentences. We use [39] because it accommodates emoticons also while performing the analysis. We use their compound score as the required sentiment analysis score. It ranges between -1 and 1. We divide this range into five equal parts and assign the ratings progressively, as shown in Table II. Let the function that computes the sentiment score and assigns the appropriate rating be $\psi(\cdot)$. Then, we compute cumulative rating ($Q(\cdot)$) for each feature over the entire product review data, i.e., $D$, in the following manner:

$$Q(F_k) = \sum_{C_j \in D} \sum_{X_j^l \in X_j} \psi(X_j^l) \times \delta(F_k \in X_j^l) \times (\phi(C_j) + 1)$$

(6)

where $\delta(\cdot)$ denotes logical function to check if a sentence consists of the concerned feature or not. During our accumulation, we also consider the number of votes received to review to which the sentence belongs. These votes inform us about the strength of the opinion associated with the reviews. Let $\phi(C_j)$ denote the number of votes received for $C_j$. Here, we are assuming that any sentence equally contributes to the strength of the opinion. We adjust the votes by adding 1 to account for self-votes of the customers who originally wrote the reviews. This accumulation is illustrated in Fig. 6. Then, we compute our final rating ($A(\cdot)$) for a feature $F_k$ using the below equation:

$$A(F_k) = \frac{Q(F_k)}{\sum_{C_j \in D} \sum_{X_j^l \in X_j} \delta(F_k \in X_j^l) \times (\phi(C_j) + 1)}$$

(7)

where we divide the cumulative number of stars by the total number of votes received during the accumulation. In this way, we are essentially computing the weighted average [40], where weights are determined by the votes received. So, we now have a feature-level rating for the feature $F_k$ of a product using the customer reviews and review votes. The same proposed methodology can be applied to any number of features, any number of products, and any number of items.

IV. EXPERIMENTS

In this section, we give details of experiments conducted using the proposed methodology. First, we discuss the dataset used. Then, we discuss the features chosen from the word frequency table of the dataset. At last, we discuss our analysis of the feature-level ratings obtained using the proposed method.
### Table III

| Our features dataset: the related words are separated by || and led by a keyword |
|---|---|
| phone, product, phones, device, cell, item, smartphone, mobile, model, cellphone, products, devices, piece, cellular, smartphones, items, telephone, handset, cellphones || screen, display, glass, screens, LCD, displays, displaying || battery, batteries || camera, resolution, cameras, pixels, pixel, cam, megapixel, megapixels || price, money, buying, pay, cost, paid, sold, purchasing, spent, budget, bucks, rate, spending, pricing, pays || sim, card, dual, sim || apps, app, program, application, widgets, processes, module || android, version, operating, iOS, versions, software, windows || case, box, packaged, packing, boxes, packs || charge, charged, charging, charges, discharges, discharge, charging || charger, plug, adapter, chargers, plugs, | service, services, watch, clock, size, sizes, call, talk, voice, called, talking, dial, speak, outgoing, communications, communicating || wifi, brand, memory, data, space, || pictures, picture, photos, pics, photo, image, images, photography || touch, touchscreen || text, texts, editing, typing, qwerty, keypad, keyboards, dials || music, audio, listen, listening || internet, online, web, browsing || light, flash, flashlight, torch, warranty || bluetooth, wireless, video, videos, streaming, stream, fps, settings, setting, setup, configure, configuration, calibration, color, colors, colour || design, build, shape, compact || download, downloaded, downloads, USB, microUSB, email, emails, speed, speeds, speedy, bandwidth || headph, headph, headset, earphones, headphone, headsets || GPS, games, gaming || RAM, messages, messaging, SMS, messenger, msg, cable, cord, connector, cables || manual, instructions, instruction, booklet || processor, cpu, processors || specs, specifications, spec || hardware, | fingerprint, finger, fingers, fingerprints, switch, switched, switches || accessories, accessory || weight, bulky, lightweight || sensor, sensors, | face, faces, protection, firmware, protectors, virus, antivirus || notification, notifier, notifications, prompts || brightness, | jack, chip, charger || hotspot, graphics, GPU, graphic || icons, icon || selfie, selfies || recharge, recharging || electronics, vibrate, vibration, vibrates, | vibration, shaker, recording, recorder, zoom, lens, chat, chatting, | ringtone, ringing, | beats, overheats, temperature, overheated, temp, | voice, voice, mail, | inbox, stereo, | scroll, slider, slides, swipe || guarantee, guaranteed, languages, multimedia || compatibility, driver, drivers || pedometer, | trackpad, calculator, handsfree, autofocus, | OTG, troubleshooting, troubleshooting, | airplane, | mute, | syncs, multitask, | backlight, | permissions, reminders, | echo, | trackball, | panorama, | speech, | lockscreen, | vga, |
| Feature      | Rating | Feature    | Rating |
|-------------|--------|------------|--------|
| alarm       | 1.811  | email      | 3.107  |
| speed       | 4.977  | brightness | 3      |
| phone       | 3.524  | button     | 3.757  |
| chip        | 3.349  | weight     | 3.652  |
| music       | 3.916  | sensor     | 3      |
| camera      | 4.032  | face       | 4.538  |
| electronic  | 1      | service    | 3.123  |
| calendar    | 3.305  | files      | 3.022  |
| navigation  | 4.111  | call       | 3.785  |
| sim         | 3.255  | troubleshooting | 5 |
| multimedia  | 3      | cable      | 3.848  |
| tested      | 3.5    | zoom       | 4.788  |
| vibrate     | 2      | multitask  | 5      |
| network     | 3.464  | settings   | 3.169  |
| sound       | 3.512  | gps        | 4.879  |
| internet    | 3.985  | hardware   | 2.2    |
| charge      | 3.216  | touch      | 3.66   |
| languages   | 3      | stereo     | 3      |
| case        | 3.454  | echo       | 3.714  |
| jack        | 3      | headphondes| 3.071  |
| price       | 3.13   | memory     | 3.227  |
| calculator  | 2      | hotspot    | 3      |
| text        | 2.426  | wifi       | 4.702  |
| heats       | 4      | scroll     | 3.705  |
| light       | 4.075  | specs      | 5      |
| brand       | 3.15   | ash        | 3.286  |
| icons       | 4.667  | manual     | 3.614  |
| size        | 4.776  | warranty   | 2.737  |
| switch      | 3.957  | messages   | 3.61   |
| apps        | 3.738  | download   | 3.787  |
| protection  | 3.769  | battery    | 3.754  |
| android     | 3.514  | ringtone   | 3.333  |
| chat        | 3.917  | pictures   | 3.836  |
| watch       | 2      | bluetooth  | 3.993  |
| fingerprint | 4.206  | accessories | 2    |
| design      | 4.6    | games      | 4.016  |
| sd          | 2.739  | screen     | 3.516  |
| mms         | 5      | video      | 3.24   |
| charger     | 3.298  | color      | 3.647  |

### A. Dataset

We apply the proposed method on a dataset named Amazon Reviews: Unlocked Mobile Phones[^1] a dataset extracted by PromptCloud from the Amazon website. It consists of reviews, product-level ratings, and review votes for a total of 4418 mobile phones. There are a total of 413841 reviews present in the dataset, along with the votes obtained by them. So, there are enough reviews to carry out our sentiment analysis and obtain general insights.

### B. Features

In Table III, we list all the words which we select as feature words. As discussed earlier, they have been extracted while observing the word frequency table of the entire dataset. The feature words which are related are separated from others using |. Many of the words are just plurals of the already existing words. The related feature words are led by a feature keyword (represented in blue color), which is most prominent in the dataset amongst all the related feature words. Note that

[^1]: https://www.kaggle.com/PromptCloudHQ/amazon-reviews-unlocked-mobile-phones/data

| Phone                                      | No. of features |
|--------------------------------------------|-----------------|
| Nokia N9 - Black                           | 17              |
| ASUS ZenFone 3 ZE552KL (SHIMMER GOLD)       | 15              |
| Asus ZenFone 3 ZE552KL Moonlight White      | 14              |
| 3.5 JUNING Blue                            | 14              |
| JUNING 7-Inch - Black                      | 14              |
| 3.5 JUNING Black                           | 14              |
| 3.5 JUNING White                           | 14              |
| Asus ZenFone 3 ZE552KL Sapphire Black      | 14              |
| LG V10 H962 64GB Ocean Blue, 5.7"          | 12              |
| LG V10 H962 64GB Black, 5.7"               | 12              |
| LG V10 H962 64GB Ocean Blue, 5.7"          | 12              |
| LG V10 H962 64GB Black, 5.7"               | 12              |
| Sony Ericsson W995a Walkman (Progressive Black) | 12           |
| BLU Life View L110X 5.7-Inch (Blue)         | 12              |
| THL 5000 5" FHD IPS MTK6592T (Black)        | 12              |
| 3.5 MTK6580 JUNING G850K Black              | 12              |
| LG V10 H962 64GB 5.7-Inch (Brown Beige)     | 12              |
| SKY Devices Platinum Series 5.0W - White   | 12              |
| LG V10 H962 64GB 5.7-Inch (Opal Blue)       | 12              |
| LG V10 H962 64GB White, 5.7"               | 12              |
| Huawei P9 Lite VNS-L22 5.2-Inch (BLACK)     | 11              |
| LG Electronics G3 Stylus D690 (Black Titanium) | 11          |
| Huawei Mate 8 32GB 6-Inch (Silver)         | 11              |
| Huawei Mate 8 NXT-L29 32GB 6-Inch (Space Gray) | 11          |
| LG VX8550 Chocolate Phone (Verizon Wireless) | 11              |
| ASUS Zenfone 6 A600CG 6-inches White       | 11              |
| Futuretech A6 4.5 Inch Mik6552S (Yellow)    | 11              |
| Nokia C7 Unlocked Quadband Smartphone      | 11              |
| Huawei P9 Lite VNS-L22 5.2-Inch (WHITE)     | 11              |
| ASUS ZENFONE 6 A601CG 6"                   | 11              |
| Samsung Galaxy S2 PLUS i9150P blue-grey    | 11              |
| Futuretech A6 4.5 Inch Mik6552S (black)     | 11              |
| Nokia Asha 302                              | 10              |
| ZTE Axon Pro Phthalio Blue                  | 10              |
| Smartwatch, GEEKERA Watch Phone (Black)     | 10              |
| Yezz Andy 5E - (White)                     | 10              |
| Nokia Lumia 1520 - Black                    | 10              |
| Nokia N82 (Silver)                          | 10              |
| Huawei P9 Lite VNS-L22 (GOLD)               | 10              |
| Blackberry Torch 9800 - Black               | 10              |
| HTC One Mini 2 16GB - Silver                | 10              |
| ZTE Spro 2 Smart Projector (Silver)         | 10              |
| Nokia N79 (Silver)                          | 10              |
| Cubot X15 5.5" Inches                       | 10              |
| Samsung Evergreen A667 - Black             | 10              |
| Straight Talk Phone X2                      | 10              |
| ZTE Axon Pro                                | 10              |
| Nokia Lumia 1520 - Red                      | 10              |
| Honor 8 Dual Camera - Pearl White           | 10              |
| ZTE Axon Pro, A1P133, 32 GB, Chromium Silver | 10          |
| ZTE Axon Pro, 64 GB, Chromium Silver        | 10              |
| Unnecto Air 5.5 (Gray)                      | 10              |
| LG G3 Stylus 3G L960 (White)                | 10              |
| Nokia X5                                   | 10              |
| OnePlus White 5.5 inch                      | 10              |
as well, and this is the most discussed feature, as expected. Such consideration helps us evaluate the performance of the proposed method by comparing the feature level ratings of our phone feature with the corresponding product-level ratings already available in the dataset. We also give word cloud for actual feature keywords (i.e., except phone) in Fig. 7. Larger the keyword, most discussed it is in the reviews. It is clear from the word cloud that battery, screen, price, camera, sim, and apps are some of the most discussed features in phones.

### C. Ratings

We give sample feature-level ratings of our proposed method in Table VI, it is for Nokia C6. It can be noted how the proposed method can rate a phone on a vast number of features, and this can certainly help consumers in making their decisions on buying their mobile phones. Since we can get feature-level ratings for all the phones, to summarize our results, we report the number of phones close to different integer ratings (e.g., 2-star) for each feature in Table VI. We obtain this by rounding off the ratings calculated. It is clear from the Table that most of the phones get 3-star ratings.

| Feature       | 1-star | 2-star | 3-star | 4-star | 5-star |
|---------------|--------|--------|--------|--------|--------|
| phone         | 12     | 88     | 197    | 197    | 219    |
| screen        | 28     | 213    | 1161   | 1162   | 333    |
| battery       | 55     | 305    | 1346   | 813    | 267    |
| camera        | 38     | 219    | 706    | 1049   | 466    |
| price         | 39     | 199    | 1101   | 1575   | 378    |
| sim           | 36     | 260    | 1336   | 811    | 182    |
| apps          | 35     | 203    | 1104   | 797    | 216    |
| android       | 32     | 237    | 1097   | 803    | 219    |
| case          | 30     | 242    | 1128   | 853    | 242    |
| charge        | 52     | 319    | 1469   | 489    | 128    |
| charger       | 63     | 324    | 1260   | 458    | 119    |
| service       | 51     | 227    | 848    | 725    | 331    |
| watch         | 22     | 98     | 311    | 282    | 171    |
| size          | 10     | 52     | 384    | 802    | 526    |
| call          | 39     | 244    | 1194   | 730    | 197    |
| wifi          | 36     | 183    | 573    | 320    | 137    |
| brand         | 29     | 115    | 705    | 598    | 314    |
| memory        | 42     | 240    | 1124   | 639    | 214    |
| pictures      | 28     | 202    | 775    | 948    | 394    |
| touch         | 55     | 235    | 614    | 529    | 209    |
| text          | 39     | 204    | 842    | 468    | 192    |
| sound         | 62     | 300    | 585    | 787    | 322    |
| set           | 25     | 119    | 525    | 384    | 160    |
| network       | 40     | 185    | 845    | 484    | 141    |
| button        | 43     | 269    | 1109   | 595    | 148    |
| music         | 23     | 155    | 573    | 574    | 270    |
| internet      | 39     | 213    | 943    | 579    | 207    |
| light         | 36     | 117    | 607    | 630    | 317    |
| warranty      | 47     | 277    | 701    | 221    | 48     |
| bluetooth     | 18     | 146    | 485    | 349    | 138    |
| video         | 13     | 139    | 514    | 620    | 318    |
| settings      | 15     | 195    | 749    | 554    | 168    |
| color         | 32     | 104    | 305    | 565    | 509    |
| design        | 13     | 81     | 282    | 813    | 632    |
| download      | 22     | 155    | 635    | 457    | 165    |
| usb           | 26     | 95     | 251    | 114    | 37     |
| email         | 39     | 140    | 654    | 359    | 145    |
| speed         | 27     | 90     | 412    | 458    | 344    |
| headphones    | 62     | 255    | 629    | 393    | 144    |
| gfs           | 29     | 31     | 104    | 74     | 46     |
| games         | 22     | 89     | 330    | 434    | 256    |
| ram           | 7      | 27     | 220    | 151    | 88     |
| messages      | 27     | 198    | 688    | 299    | 107    |
| cable         | 53     | 247    | 635    | 321    | 84     |
| manual        | 38     | 267    | 656    | 299    | 122    |
| processor     | 22     | 73     | 333    | 344    | 180    |
| specs         | 17     | 69     | 344    | 342    | 231    |
| hardware      | 23     | 160    | 289    | 221    | 163    |
| fingerprint   | 35     | 134    | 437    | 320    | 138    |
| switch        | 45     | 152    | 610    | 346    | 147    |
| accessories   | 17     | 92     | 410    | 330    | 186    |
| weight        | 15     | 61     | 355    | 365    | 327    |
| sensor        | 20     | 71     | 211    | 171    | 77     |
| face          | 23     | 102    | 206    | 181    | 148    |

### Table VI

Number of phones with different integer ratings for each feature
TABLE VII
THE BEST PHONES WE RECOMMEND FOR DIFFERENT FEATURES.

| Features | Best Phones | Features | Best Phones |
|----------|-------------|----------|-------------|
| phone    | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | protection | SKY Devices Platinum Series 5.0W - Silver |
| screen   | LG V8X8500 Chocolate Phone (Verizon Wireless) | notification | Blackberry Z10 16GB - Black |
| battery  | JUNING 7-Inch - Black | brightness | ASUS Zenfone 6 A600CG White |
| camera   | Huawei P9 Lite VNS-L22 5.2-Inch (BLACK) | jack | Nokia N9 16 GB MeeGo OS - Black |
| price    | T2L 5000 5" FHD IPS MTK6592T (Black) | chip | Nokia N9 16 GB MeeGo OS - Black |
| sim      | Huawei P9 Lite VNS-L22 5.2-Inch (BLACK) | files | ASUS Zenfone 3 ZE552KL 5.5-inch (SHIMMER GOLD) |
| apps     | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | scanner | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) |
| android  | LG V10 H962 64GB Ocean Blue, Dual Sim, 5.7” | files | ASUS Zenfone 3 ZE552KL 5.5-inch (SHIMMER GOLD) |
| case     | Nokia N9 16 GB MeeGo OS - Black | microphone | Nokia N9 16 GB MeeGo OS - Black |
| charger  | Nokia N9 16 GB MeeGo OS - Black | navigation | Sony Xperia Z1 (C6902) - Black |
| charger  | JUNING 7-Inch - Black | waterproof | JUNING 7-Inch - Blue |
| service  | T2L 5000 5" FHD IPS MTK6592T (Black) | calendar | New Genuine Nokia X3-00 Unlocked GSM X3 |
| watch    | 5.5” JUNING Blue | alarm | BLU Life View L110X 5.7-Inch (Blue) |
| size     | 5.5” JUNING Blue | BLU Life View L110X 5.7-Inch (Blue) | |
| call     | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | hotspot | 5.5” JUNING Blue |
| wifi     | T2L 5000 5" FHD IPS MTK6592T (Black) | graphics | JUNING 7-Inch - Black |
| brand    | JUNING 7-Inch - Black | icons | Sony Ericsson W995a Walkman (Progressive Black) |
| memory   | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | camera | 5.5” JUNING Blue |
| pictures | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | recharge | BLU Life View L110X 5.7-Inch (Blue) |
| touch    | 5.5” JUNING Blue | electronics | Samsung S7 Galaxy (SM-G930UZKAXAA) |
| text     | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | vibrate | Sony Ericsson W995a Walkman (Progressive Black) |
| sd       | Nokia N9 16 GB MeeGo OS - Black | zoom | BLU Life View L110X 5.7-Inch (Blue) |
| network  | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| button   | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | ringtones | Samsung Evergreen A667 - Black |
| music    | Nokia N9 16 GB MeeGo OS - Black | beats | T2L 5000 5” FHD IPS MTK6592T (Black) |
| internet | Nokia N9 16 GB MeeGo OS - Black | voicemail | ZTE Axon Pro, 64 GB Phthalo Blue |
| light    | JUNING 7-Inch - Black | stereo | BLU Life View L110X 5.7-Inch (Blue) |
| warranty | Samsung Galaxy Note 5, White 32GB (AT&T) | scroll | LG V10 H962 64GB Ocean Blue, Dual Sim, 5.7” |
| bluetooth | LG V8X8500 Chocolate Phone (Verizon Wireless) | Smartwatch | GEEKERA Bluetooth Watch Phone (Black) |
| video    | JUNING 7-Inch - Black | languages | Sony Ericsson W995a Walkman (Progressive Black) |
| settings | ASUS ZenFone 3 ZE552KL 64GB Sapphire Black, 5.5-inch | multimedia | LG V10 H901 64GB T-Mobile- Space Black |
| color    | Nokia N9 16 GB MeeGo OS - Black | compatibility | LG Electronics G3 Stylus D690 (Black Titanium) |
| design   | Nokia N9 16 GB MeeGo OS - Black | driver | BLU Life View L110X 5.7-Inch (Blue) |
| download | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| download | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| email    | 5.5” JUNING Blue | BLU Life View L110X 5.7-Inch (Blue) |
| speed    | Nokia N9 16 GB MeeGo OS - Black | BLU Life View L110X 5.7-Inch (Blue) |
| headphones | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| gps      | Sony Ericsson W995a Walkman (Progressive Black) | BLU Life View L110X 5.7-Inch (Blue) |
| games    | Nokia N9 16 GB MeeGo OS - Black | BLU Life View L110X 5.7-Inch (Blue) |
| ram      | ASUS Zenfone 6 A600CG White | BLU Life View L110X 5.7-Inch (Blue) |
| messages | Blackberry Torch 9800 - Black | BLU Life View L110X 5.7-Inch (Blue) |
| table    | Nokia N9 16 GB MeeGo OS - Black | BLU Life View L110X 5.7-Inch (Blue) |
| manual   | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| processor | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| specs    | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| hardware | ASUS ZenFone 3 ZE552KL 5.5-inch (SHIMMER GOLD) | BLU Life View L110X 5.7-Inch (Blue) |
| fingerprint | LG V10 H962 64GB Ocean Blue, Dual Sim, 5.7” | BLU Life View L110X 5.7-Inch (Blue) |
| switch   | Nokia N9 16 GB MeeGo OS - Black | BLU Life View L110X 5.7-Inch (Blue) |
| accessories | ZTE Axon Pro, 64 GB Phthalo Blue | BLU Life View L110X 5.7-Inch (Blue) |
| weight   | 5.5” JUNING Blue | BLU Life View L110X 5.7-Inch (Blue) |
| sensor   | Nokia N9 16 GB MeeGo OS - Black | BLU Life View L110X 5.7-Inch (Blue) |
| face     | LG V10 H962 64GB Ocean Blue, Dual Sim, 5.7” | BLU Life View L110X 5.7-Inch (Blue) |

i.e., average, across our features. Note that not all the phones will have reviews mentioning every single feature we have selected. For example, the vga feature (at the end of the Table) is mentioned in the reviews of only 14 phones. So, only 14 phones can have ratings for the vga feature. In Table VI, we rank all our phones based on the number of features in which they are the best based on the reviews. We report only those phones which are best for at least ten features. Based on these rankings, we also report the best phone for which the required feature has been rated in Table VII so that we can recommend a phone for a given feature. For example, both Nokia N9 - Black and JUNING 7-Inch - Black have ratings of 5 for music, and we recommend Nokia N9 - Black since its best in more number of features. So, the higher the ranking, the better is the chance for the recommendation.

D. Evaluation

Although there are no ground-truth feature-level ratings available to evaluate our method, we have the purposely selected phone features that can be evaluated. We can compare our results on the phone named feature with the phones overall ratings already available in the dataset. The ratings given by
individual customers are weighted averaged (weighted by the review votes) and considered as ground-truth ratings for our phone named feature. We report different error metric values in the Table VIII while comparing with such ground truth ratings. It is impressive that, on average, our rating differs from the actual ratings by just 0.555, which is approximately just half a star, as suggested by the MAE (Mean Average Error) error metric.

Also, we report a confusion matrix for our phone feature in Table IX. Both our and ground-truth ratings are rounded to get such integer star ratings. So, we now have five classes (1-star to 5-star) into which a phone can be classified. With such discrete outputs now, we generate the confusion matrix. It suggests that our system predicts correct ratings for 2165 mobiles and the ratings within 1-star closeness for 3886 mobiles out of a total of 4141 mobiles. So, if we want the exact integer star rating, then accuracy jumps to 93.8%. Therefore, we can comfortably say that the proposed methodology does work well for the phone named feature. Note that the total number of mobiles here has changed from 4418. That is because there might be some phones which do not have any review with the feature words we have chosen as the words related to the phone named feature. Therefore, such phones do not receive the rating for their phone named feature to participate in this kind of evaluation, which requires at least one such review for consideration.

**CONCLUSION**

We have developed a system to rate mobile phones in terms of 108 features based on customer reviews and review votes. We could rate 4k+ phones; this can help make personalized buying decisions and improve the products. We accomplish this by first converting the unstructured data into structured data; then, we extract the sentences comprising our feature keywords; then, we were able to provide the feature-level ratings through sentiment analysis of these sentences. We rank the phones based on the number of features they are best at, and accordingly, we were able to recommend the best phones for a feature. We tested our methodology on the phone named feature by considering the overall customer ratings as ground-truth ratings. The performance of our method is found to be decent. We obtain MAE of only 0.555, i.e., approximately just half a star. We get 52.3% accuracy if exact integer ratings have to be predicted. However, if we can tolerate the 1-star integer rating error, the accuracy jumps to 93.8%. The proposed approach is unsupervised. As an extension, we will work on improving the performance by taking a weakly-supervised or supervised approach to this problem, for which we will have to annotate the available data in terms of all our 108 features.

**REFERENCES**

[1] V. Raghavan, G. Ver Steeg, A. Galstyan, and A. G. Tartakovsky, “Modeling temporal activity patterns in dynamic social networks,” *IEEE Transactions on Computational Social Systems*, vol. 1, no. 1, pp. 89–107, 2014.

[2] A. Farasat, G. Gross, R. Nagi, and A. G. Nikolaev, “Social network analysis with data fusion,” *IEEE Transactions on Computational Social Systems*, vol. 3, no. 2, pp. 88–99, 2016.

[3] Y. Zhu, D. Li, R. Yan, W. Wu, and Y. Bi, “Maximizing the influence and profit in social networks,” *IEEE Transactions on Computational Social Systems*, vol. 4, no. 3, pp. 54–64, 2017.

[4] X. Yang, C. Liang, M. Zhao, H. Wang, H. Ding, Y. Liu, Y. Li, and J. Zhang, “Collaborative filtering-based recommendation of online social voting,” *IEEE Transactions on Computational Social Systems*, vol. 4, no. 1, pp. 1–13, 2017.

[5] F. S. N. Karan and S. Chakraborty, “Dynamics of a repulsive voter model,” *IEEE Transactions on Computational Social Systems*, vol. 3, no. 1, pp. 13–22, 2016.

[6] F. Smarandache, M. Colhon, Ş. Vlăduţescu, and X. Negrea, “Word-level neutrosophic sentiment similarity,” *Applied Soft Computing*, vol. 80, pp. 167–176, 2019.

[7] K. Ravi, V. Ravi, and P. S. R. K. Prasad, “Fuzzy formal concept analysis based opinion mining for crm in financial services,” *Applied Soft Computing*, vol. 60, pp. 786–807, 2017.

[8] J. Xu, F. Huang, X. Zhang, S. Wang, C. Li, Z. Li, and Y. He, “Sentiment analysis of social images via hierarchical deep fusion of content and links,” *Applied Soft Computing*, vol. 80, pp. 387–399, 2019.

[9] F. H. Khan, U. Qamar, and S. Bashir, “Sentimli: Introducing point-wise mutual information with sentiwordnet to improve sentiment polarity detection,” *Applied Soft Computing*, vol. 39, pp. 140–153, 2016.

[10] K. R. Jerripothula, J. Cai, and J. Yuan, “Quality-guided fusion-based co-saliency estimation for image co-segmentation and colocalization,” *IEEE Transactions on Multimedia*, vol. 20, no. 9, pp. 2466–2477, 2018.

[11] A. Wiliam, W. K. Sasmoko, and Y. Indrianti, “Sentiment analysis of social media engagement to purchasing intention,” in *Understanding Digital Industry: Proceedings of the Conference on Managing Digital Industry, Technology and Entrepreneurship (CoMDITE 2019)*, July 10–11, 2019, Bandung, Indonesia. Routledge, 2020, p. 362.

[12] L. G. Singh, A. Anil, and S. R. Singh, “She: Sentiment hashtag embedding through multitask learning,” *IEEE Transactions on Computational Social Systems*, 2020.

[13] K.-P. Lin, Y.-W. Chang, C.-Y. Shen, and M.-C. Lin, “Leveraging online word of mouth for personalized app recommendation,” *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 1061–1070, 2018.

[14] N. Bui, J. Yen, and V. Honavar, “Temporal causality analysis of sentiment change in a cancer survivor network,” *IEEE transactions on computational social systems*, vol. 3, no. 2, pp. 75–87, 2016.

[15] R.-C. Chen et al., “User rating classification via deep belief network learning and sentiment analysis,” *IEEE Transactions on Computational Social Systems*, 2019.

[16] S. Kumar, K. De, and P. P. Roy, “Movie recommendation system using sentiment analysis from microblogging data,” *IEEE Transactions on Computational Social Systems*, 2017.

[17] M. Ling, Q. Chen, Q. Sun, and Y. Jia, “Hybrid neural network for sina weibo sentiment analysis,” *IEEE Transactions on Computational Social Systems*, 2020.

[18] K. Chakraborty, S. Bhattacharyya, and R. Bag, “A survey of sentiment analysis from social media data,” *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 450–464, 2020.
[19] F.-C. Yang, A. J. Lee, and S.-C. Kuo, “Mining health social media with sentiment analysis,” Journal of medical systems, vol. 40, no. 11, p. 236, 2016.

[20] M. Palomino, T. Taylor, A. Goker, J. Isaacs, and S. Warber, “The online dissemination of nature–health concepts: Lessons from sentiment analysis of social media relating to nature-deficit disorder,” International journal of environmental research and public health, vol. 13, no. 1, p. 142, 2016.

[21] M. T. Khan and S. Khalid, “Sentiment analysis for health care,” in Big Data: Concepts, Methodologies, Tools, and Applications. IGI Global, 2016, pp. 676–689.

[22] J. Ramteke, S. Shah, D. Godhia, and A. Shaikh, “Election result prediction using twitter sentiment analysis,” in 2016 international conference on inventive computation technologies (ICICT), vol. 1. IEEE, 2016, pp. 1–5.

[23] S.-O. Proksch, W. Lowe, J. Wäckerle, and S. Soroka, “Multilingual sentiment analysis: A new approach to measuring conflict in legislative speeches,” Legislative Studies Quarterly, vol. 44, no. 1, pp. 97–131, 2019.

[24] Y. Yu and X. Wang, “World cup 2014 in the twitter world: A big data approach to analyzing sentiments in us sports fans tweets,” Computers in Human Behavior, vol. 48, pp. 392–400, 2015.

[25] G. M. Lucas, J. Gratch, N. Malandrakis, E. Szablowski, E. Fessler, and J. Nichols, “Goaalll!: Using sentiment in the world cup to explore theories of emotion,” Image and Vision Computing, vol. 65, pp. 58–65, 2017.

[26] S. L. Addepalli, S. G. Addepalli, M. Kherajani, H. Jeshnani, and S. Khedkar, “A proposed framework for measuring customer satisfaction and product recommendation for ecommerce,” International Journal of Computer Applications, vol. 138, no. 3, pp. 30–35, 2016.

[27] J. Mehta, J. Patil, R. Patil, M. Somani, and S. Varma, “Sentiment analysis on product reviews using hadoop,” International Journal of Computer Applications, vol. 9, no. 11, pp. 0975–8887, 2016.

[28] X. Fang and J. Zhan, “Sentiment analysis using product review data,” Journal of Big Data, vol. 2, no. 1, p. 5, Jun 2015. [Online]. Available: https://doi.org/10.1186/s40537-015-0015-2.

[29] D. K. Raja and S. Pushpa, “Feature level review table generation for e-commerce websites to produce qualitative rating of the products,” Future Computing and Informatics Journal, vol. 2, no. 2, pp. 118–124, 2017.

[30] N. Nandal, R. Tanwar, and J. Pruthi, “Machine learning based aspect level sentiment analysis for amazon products,” Spatial Information Research, pp. 1–7, 2020.

[31] J. Sadhasivam and R. B. Kalivaradhan, “Sentiment analysis of amazon products using ensemble machine learning algorithm,” International Journal of Mathematical, Engineering and Management Sciences, vol. 4, no. 2, pp. 508–520, 2019.

[32] K. R. Jerripothula, J. Cai, and J. Yuan, “Efficient video object co-localization with co-saliency activated tracklets,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 3, pp. 744–755, 2019.

[33] K. R. Jerripothula, J. Cai, J. Lu, and J. Yuan, “Object co-skeletonization with co-segmentation,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3881–3889.

[34] K. R. Jerripothula, J. Cai, and J. Yuan, “Cats: Co-saliency activated tracklet selection for video co-localization,” in Computer Vision – ECCV 2016, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 187–202.

[35] K. R. Jerripothula, J. Cai, F. Meng, and J. Yuan, “Automatic image co-segmentation using geometric mean saliency,” in 2014 IEEE International Conference on Image Processing (ICIP), 2014, pp. 3277–3281.

[36] K. Zhang, R. Narayanan, and A. Choudhary, “Voice of the customers: mining online customer reviews for product feature-based ranking,” The Computer Journal, vol. 53, no. 1, pp. 11–19, 2010.

[37] K. Rufina and D. Toshniwal, “Feature based summarization of customers reviews of online products,” Procedia Computer Science, vol. 22, pp. 142–151, 2013.

[38] K. R. Jerripothula, J. Cai, and J. Yuan, “Group saliency propagation for large scale and quick image co-segmentation,” in 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 4639–4643.

[39] C. J. Hutto and E. Gilbert, “Vader: A parsimonious rule-based model for sentiment analysis of social media text,” in Eighth international AAAI conference on weblogs and social media, 2014.

[40] K. R. Jerripothula, J. Cai, and J. Yuan, “Ocec: Quality constrained co-saliency estimation for common object detection,” in 2015 Visual Communications and Image Processing (VCIP), 2015, pp. 1–4.

[41] M. Glenski, C. Pennycuff, and T. Wenninger, “Consumers and curators: Browsing and voting patterns on reddit,” IEEE Transactions on Computational Social Systems, vol. 4, no. 4, pp. 196–206, 2017.