Determining of the User Attitudes on Mobile Security Programs with Machine Learning Methods

Rıdvan Yayla
Information Technology Department, Bursa Technical University
Mimar Sinan Dist. Mimar Sinan Boulevard Eflak Str. 16310, Bursa, Turkey
E-mail: ridvan.yayla@btu.edu.tr

Turgay Tugay Bilgin
Computer Engineering Department, Bursa Technical University,
Mimar Sinan Dist. Mimar Sinan Boulevard Eflak Str. 16310, Bursa, Turkey
E-mail: turgay.bilgin@btu.edu.tr

Security plays an important role in today's virtual world. Cybersecurity software has been widely used by the development of portable virtual environments. Smartphones occur in an important part of our lives. Daily routines are performed over mobile phones, especially after the COVID-19 pandemic process. Due to its ease of use, compulsory or optional mobile phone use also brought about many security concerns. Mobile security software is used for different purposes such as virus removal and protection of personal information according to user preferences. In the field of natural language processing, user preferences can now be analyzed on the basis of machine learning methods with sentiment analysis. In this paper, the preference reasons for mobile security software have been analysed with machine learning methods based on user comments and sentiment analysis. In the study, all user comments have been classified into 10 main categories and the user preferences of mobile security programs have been analysed.

Povzetek: Prispevek opisuje ugotavljanje uporabniškega odnosa do programov mobilne varnosti z metodami strojnega učenja.

1 Introduction

Security plays an important role in the several web applications. Mobile applications are widely used by users, especially after the Covid-19 pandemic process. While mobile applications convenience people, they also bring about some security problems. Due to security gaps, users become the main target of hackers. While the number of mobile phones increases each day, security threats, which capture personal information with social engineering methods, also increase at the same rate.

Nowadays, there are many security threats for smartphones, such as advertisements, undesired information sharing with synchronization, or fake links. Smartphone users try preventing security threats with mobile security applications and firewalls. Mobile security programs carry vital importance to the users for providing protection. While some users prefer security programs for cleaning trash files and folders, many prefer them for personal information privacy. Over the past decade, phone manufacturers have developed many smartphone models and many users have bought a smartphone for their routines. The hackers have also targeted mobile applications recently due to the trending usage. This situation also brought about mobile security concerns. Due to the security gaps, threats increased in vulnerable mobile phones. Several mobile security programs have been developed against threats and they have used by users for different reasons.

On the other hand, evaluation scales based on different structures such as questionnaire, comment text or rating star are also widely used among web and mobile users recently. A comment or rating can be useful for other users in terms of buying a product or installation a program because users have an idea for product or application quality. Google Play Store (GPS), which offers various digital content to its consumers, offers an evaluation possibility of the applications to the users. While user attitudes on mobile security in the literature are handled classical methods such as generally surveys, the GPS comments on mobile security programs are analysed with a new approach based on machine learning methods in this paper.

2 Methodology

Different kinds of contents such as music, movies, books, programs, and games are used on GPS due to the smartphone's popularity. Additionally, mobile security programs have also been preferable among users due to security concerns. Android developers have developed several mobile security applications to meet these requirements for different purposes. At the same time, a

lot of research has also been made about the user attitudes on mobile security recently. Most of these researches have been made by either being handled for a group of people or a local area by using classical methods.

In 2008, a study that is conducted with a survey on 300 mobile-phone users in Oman investigated user attitudes toward e-commerce and other mobile devices [1]. On the other hand, Tambe and Kulal developed an offline Android mobile security application when the smartphone is stolen or lost by a thief [2]. In 2016, another survey that is performed with 301 attendees is made of awareness of mobile device security [3]. Özkan and Bıçakçı performed an analysis for two-factor authentication against account-hijacking attacks. They analysed eleven different Android authenticator applications and used different engineering techniques and open-source tools in their study [4]. Ophoff and Robinson made an online survey with 619 South African mobile users for exploring end-user smartphone security awareness. Moreover, the survey questions based on mobile security have been prepared by these researchers [5]. Ziqiang et al. proposed an approach combining the static and dynamic security detection methods to detect client-side. They divided the mobile application security detection into two parts server and client security detection. Additionally, they developed an automated platform for mobile threat detection [6]. Benenson et al. made an interview with 24 mobile users that are between 18 and 50 years old for satisfying the sentiment analysis requirement. He consulted security experiments and attitudes of these 24 mobile users for his study. As a result of the interviews, he suggested some hypotheses [7].

Several mobile security user attitudes have been determined by using classical methods, such as surveys, interviews as it is seen in the literature. In our study, we proposed a new approach by using sentiment analysis based on machine learning and classification methods based on word analysis. The GPS comments are investigated by confining with Turkish comments as a prototype study in study. Additionally, the study can be also extended with different languages such as English, Italian, Hungarian, or Slovenian provided that the categories and searched words should be defined in the desired language.

In the application, 249 mobile security programs on GPS are examined and a sentiment analysis based on the combination of user comments and ratings has been conducted. The received information with scraped method on GPS is processed and the user comments with ratings are analysed as positive, negative, and neutral with sentiment analysis.

Moreover, the criteria by which the users use the programs were analysed by being categorized into 10 categories. The aim of the study is to propose a new approach based on the machine learning method that contains sentiment analysis by considering user ratings instead of the classical methods. The working principle of the proposed method is shown in Figure 1.

3 Materials and methods

3.1 Baseline algorithm

Sentiment analysis has a baseline algorithm with different classification approach. Firstly, the tokenization process is applied for the words of each comment text. Phone numbers, dates, special markups such as user names, and emoticons that are required are examined and extracted in text. Emoticons of the text are also examined for sentiment tokenization because of they can express a positive such as smile face or negative sentiment such as sad face. Moreover, the stopwords are also extracted from the text. Negation is another important process for extracting features for sentiment classification [8]. The negation is determined by using adjectives or all words in sentence. The last process is classification by using a classifier in the model.

There are different classifiers for sentiment analysis.
1- Naïve Bayes,
2- Maximum Entropy
3- Support Vector Machine (SVM)

Maximum Entropy (MaxEnt) and SVM tend to do better than Naïve Bayes algorithm. In this study, SVM algorithm is preferred for sentiment classifier due to the ease of use.

3.2 Support vector machine (SVM)

SVM is a supervised machine learning algorithm that can be used for both classification or regression challenges [9]. The sentiment analysis, which is based on natural language processing (NLP) and machine learning algorithms, is made by the Support Vector Machine (SVM). SVM basically divided into two classes and it is a binary classification method. For multiclass classification, the same principle is utilized after breaking down the multi-classification problem into smaller subproblems, all of which are binary classification problems [10]. The multiclass problem is broken down to multiple binary classification cases, which is also called one-vs-one. This method is called linear SVM for multi-class classification. The comment sentiment is investigated as three main emotions that are positive, negative and neutral by using linear SVM. Additionally, the Stochastic Gradient Descent (SGD) algorithm is used for the text classification in each comment of determined sentiment.

3.3 Stochastic gradient descent (SGD)

SGD is a fundamental machine learning approach that can be applied to large-scale and diluted machine learning problems frequently encountered in text classification and natural language processing [11]. SGD is also a good optimization algorithm by productivity and ease of application. The advantage of SGD is to update each
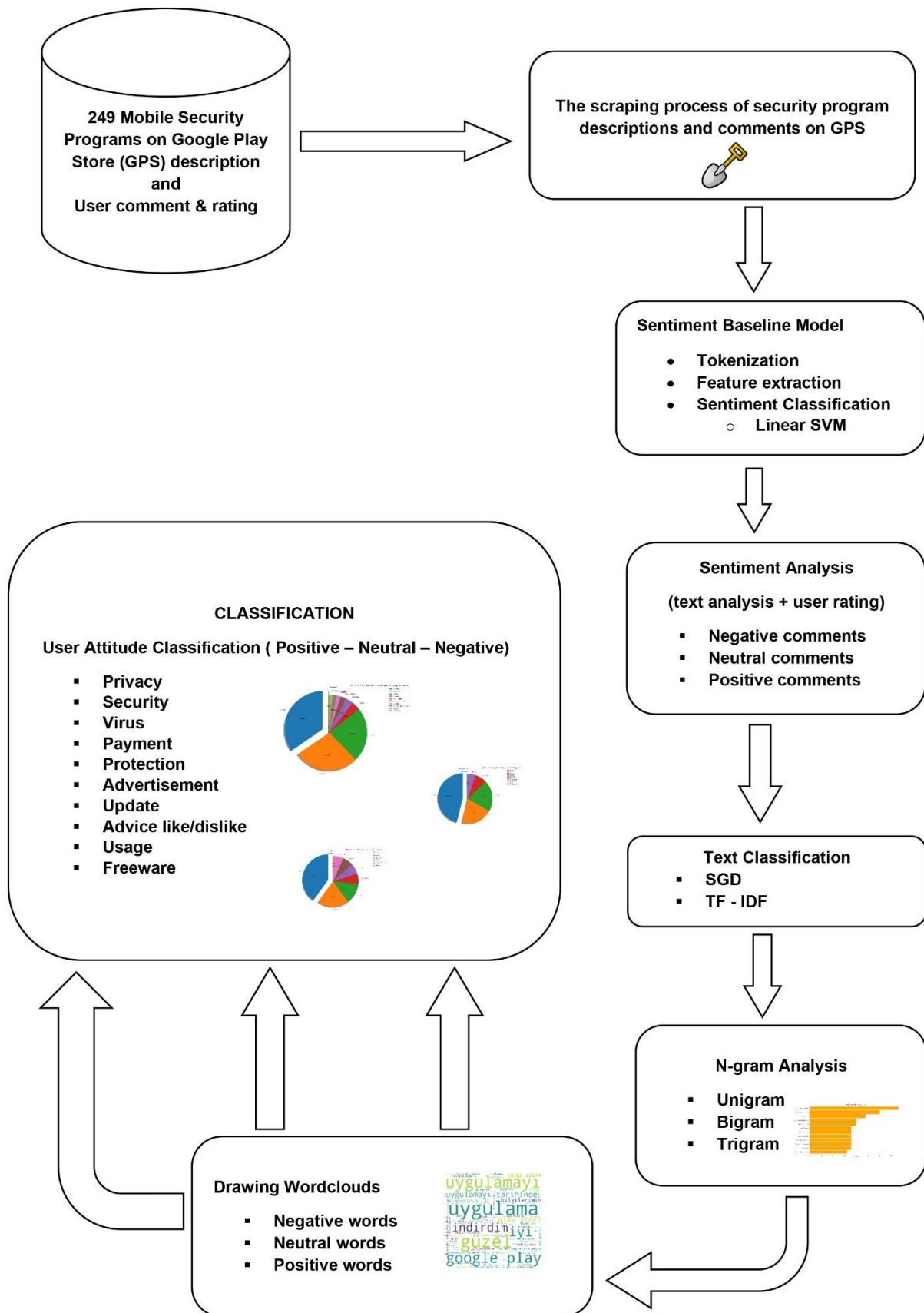
Figure 1: Working principle of the proposed method.

sample in each step by decreasing optimization calculations, especially in big datasets. SGD refers to

calculating the derivative from each training data instance and calculating the update immediately [12]. SGD

algorithm is used based on Term Frequency – Inverse Document Frequency (TF-IDF) weight factor in text classification.

TD-IDF is a calculated weight factor of a word that shows the importance in a text by using a statistical method .Thanks to TF-IDF, it is determined in which reason the GPS users write their ratings about mobile security programs. TF-IDF weights [13] are calculated as follow:

$$weight_{w,t} = \begin{cases} \log(tf_{w,T} + 1) \log \frac{n}{x_w}, & f_{w,T} \geq 1 \\ 0, & otherwise \end{cases} \quad (1)$$

3.4   N-gram model

N-gram means a sequence of N words and it helps predict the next item in a sequence. Although the n-gram model has more than one method, it consists of 3 main methods. Unigram express n-gram of size 1, bigram refers to n-gram of size 2 and trigram means n-gram of size 3 [14]. The size of n-gram can also be increased, such as four-gram, five-gram. The n-gram model can be explained with a simple sentence as following:

Example sentence:

"The mobile security application is very good."

1- Unigram: Each single word is considered for the recurrent calculation.
2- "The", "mobile", "security", "application", "is", "very", "good"
3- Bigram: Each pair of words is considered for the recurrent calculation.
4- "The mobile", "mobile security", "security application", "application is", "is very", "very good".
5- Trigram: Each three sequence of words are considered for the recurrent calculation.
6- "The mobile security", "mobile security application", "security application is", "application is very", "is very good"

In the study, the n-gram analysis is performed as unigram, bigram and trigram for all comments. Additionally, the keyword-processing algorithm is separately used at positive, negative and neutral comments. Keyword processing assigns a rate to a word and it gives a score to the word. Finally, it gives a percentage score to the text as a whole. The words are drawn by starting from the highest score word by word cloud [15]. This process is executed all positive, negative and neutral text separately and the most used words are shown in each sentiment group as a word diagram.

3.5   Model accuracy

The classification model that follows the working principle was evaluated with standard metrics called accuracy, precision, recall and F1-score where TP is true positive, TN is true negative, FP is false positive and FN is false negative [16]. Accuracy is a statistical measure which is defined as the division of the correct predictions (TP & TN) made by a classifier divided by the sum of all predictions made by the classifier, including FP and FN [17]. The accuracy is computed as follow:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Precision is defined as the ratio of the correctly identified positive cases in all predicted positive cases [18]. Precision is computed as follow:

$$precision = \frac{TP}{TP+FP} \quad (3)$$

Recall is the sensitivity of the model and it is defined as the ratio of the correctly identified positive cases to all the actual positive cases, which is the sum of FN and TP [19]. Recall is also shown as follow:

$$recall = \frac{TP}{TP+FN} \quad (4)$$

F1-Score is the harmonic meaning of the precision and recalled by taking into FP and FN cases [20]. It shows quality performance in an unbalanced data set. It is calculated as follow:

$$F_1score = \frac{2 \times (precision \times recall)}{(precision \pm recall)} \quad (5)$$

4   System design and components

4.1   Dataset

In our study, the mobile security programs on Google Play Store are examined with Turkish user comments and votes. The user comments and votes are evaluated for the sentiment analysis together. Initially, 249 mobile security programs and related comments are handled as metadata. Google play scraper based on PHP composer is used for extracting metadata and all program names, application ids and user comments and votes are written in an Excel file [21]. Thanks to Google Play scraper, 45617 comments that belong to 249 mobile security programs have been examined with a written Python script in the application. The dataset is randomly divided into train set as 80% and test set as 20%.

4.2   Pre-process

In dataset, a sequence pre-process is performed for sentiment and text classification. Initially, the all-meta data text is translated to English for sentiment analysis [22]. The normalization pre-process is applied both Turkish and English comment texts. Firstly, all html and xml markups are examined and cleaned from texts. Secondly, the usernames, web addresses, phone numbers are also cleaned. Secondly, all comment texts are examined for the feature extraction [23]. Emoticons is the important feature for sentiment analysis and it is taken into account for emotion detection. The all words are examined for negation detection. Finally, the sentiment classification is performed for all dataset. Additionally, the user ratings are also considered for an efficiency sentiment analysis in our study.

4.3   Sentiment analysis

Sentiment analysis has been realized based on three main emotions as negative, positive, and neutral. Moreover, the user votes as rating stars are also counted in sentiment analysis examination. In literature, the sentiment analysis is made in English and there are a few sentiment analysis
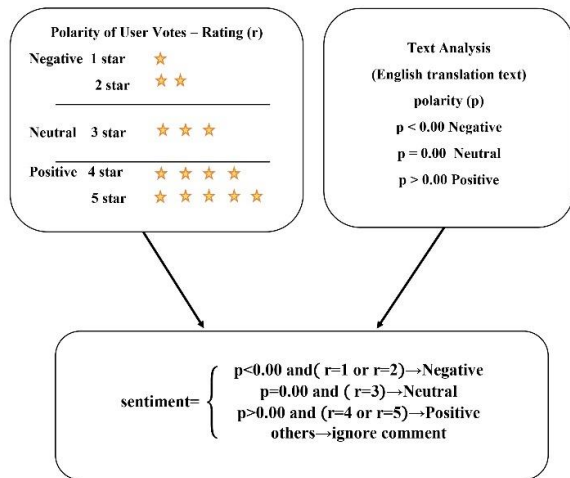
Figure 2: Sentiment analysis of the proposed method.

studies with other languages such as Spanish or Turkish [24][25]. The sentiment analysis studies based on other languages do not reflect sentiment analysis as successfully as English [26]. The diagram used in this study for the sentiment analysis is shown in Figure 2.

Instead of the usage of non-English sentiment analysis, the comments that are written as non-English are translated to English. Moreover, the sentiments are supported with the user votes related to mobile security programs for decreasing translation errors. Thanks to the translation process, the application can be executed with the other languages provided that the category names and searched words used for classification are written in the desired language.

The sentiment analysis of each comment is basically determined by sentiment polarity. Additionally, the user votes are also considered for a higher accuracy analysis in this study. As a hypothesis, the user ratings are accepted as 1 or 2 stars for negative, 3 stars for neutral, and 4 or 5 stars for positive [27]. Furthermore, when the comment text is evaluated with sentiment polarity, a polarity less than 0 is considered as negative, a polarity equal to 0 is considered as neutral, and a polarity greater than 0 is considered as positive [28]. If any user evaluation doesn't follow both rules, that comment is ignored for sentiment analysis and classification. Moreover, some users can make a negative comment and give a positive vote. These kinds of comments have been ignored by means of the proposed method.

Due to the user comments and star ratings, the user reactions are also shown in Table 1. According to the evaluation, the GPS users thought about security programs as positive 76.39 %, neutral 3.32%, and negative 1.89%. The 8395 comments are ignored according to that they don't follow the rules of the proposed method. Additionally, the numbers of positive, negative and neutral ratings according to the user ratings are also shown in Table 1.

In each original Turkish comment that follows to proposed method, the most used words (unigrams), bigrams and trigrams are determined by n-gram analysis. In addition to n-gram analysis, word clouds that contain positive, negative and neutral comment words have been also separately drawn. Finally, all comments are classified with 10 user attitude categories according to the sentiment and it has been determined whether the users like or not the security programs according to which criteria. The algorithm accuracy is evaluated with linear SVM algorithm.

4.4 Classification

When the mobile security comments are observed as a whole, 10 main Turkish classes are determined for the user attitude classification. The classification is made based on these classes. The class names (categories) based on user attitude are shown in Table 2. The class names weren't only searched, but also other words denoting the class name in each comment.

A searched word can belong to more than two classes. These searched words were taken into account for each related class. For example, let a Turkish comment is as follow:

Turkish Comment:

"Virüsü iyi temizliyor, kişisel bilgilerimi kurtardım, koruma sağlıyor, kullanışlı bir program. tavsiye ederim."

English Translation:

Comments	Rating 1	Rating 2	Percentage (%)	Total Percentage (%)	Comment Numbers
Negative	760	104	1,89 %		
	Rating 3		**Percentage (%)**		
Neutral	1513		3,32 %	81,6 %	37222
	Rating 4	**Rating 5**	**Percentage (%)**		
Positive	4293	30552	76,39 %		
Ignored Comment Numbers				18,4 %	8395
Total Comment Numbers					45617

Table 1: The numbers of rating and comments.

Class No	Class Name (Turkish Class Name)	Searched Words (Translation)
1	Virus (Virüs)	Temiz (clean), virüs, virus (virus), yok (not), bahis (bet), reklam (advertisement), sil (delete)
2	Security (Güvenlik)	Temiz (clean), güven (secure), kişisel (personal), bilgi (info), siber (cyber), saldırı (attack), koruma (protection), virüslü (infected)
3	Privacy (Gizlilik)	Kişisel (personal), bilgi (info), şahıs (person), paylaş, paylas (share)
4	Freeware (Ücretsiz)	Ücretsiz, ucretsiz (freeware), bedava (free) ,parasız (no charge), free
5	Paid (Ücretli)	Para (money), ücret (fee), kazan (win), reklam (advertisement), yıllık (annual), uyelik, üyelik (membership), abone (subscriber), deneme (trial), satış (sales), satma (sell), kart (card), sürüm (version), premium, ödeme (payment), satin (buy), fiyat (price)
6	Advertisement (Reklam)	Reklam (advertisement), afis, afiş (banner), pano (board), sürekli (permanent), izle (watch), gereksiz (unnecessary)
7	Protection (Koruma)	Virus (virus), sil (delete), tehdit (threat), koru (secure), etkili (effective), karşı (against)
8	Update (Güncel)	Performans (performance), güncel (update), otomatik (automatic), son (final), indirme (download)
9	Usage (Kullanım)	Performans (performance), kullanışlı (useful), yararlı (benefit), zararlı (harmful), işlevsel (functional), kasmıyor (twitch), ısınma (warming),şarj (charge), gereksiz (unnecessary), tüketim (consumption), bildirim (notification)
10	Advice like/dislike (Tavsiye beğeni)	öneri (offer), beğen (like), tavsiye (advice), çok iyi (very well), maşallah, harika (wonderful), kötü (bad), hiç (any)

Table 2: The categories of the study on user attitude.

"It cleans the virus well; I recovered my personal information and it provides protection. It is a useful program. I advise everybody"

In this study, all words are searched as Turkish.

1- "Virus" and "clean" words will belong to virus class,
2- "personal" and "information" words will belong to privacy and security classes
3- "protection" word will belong to protection class,
4- "useful" word will belong to usage class
5- "advise" word will belong to advice like/dislike class.

In addition to searched words, the root of the word such as "protect" word in protection class or "pay" word in paid class is also evaluated within the classes. Thereby, the frequency of the word is determined in the comment text. Even if this prototype study is prepared in the Turkish language, the word search process can be made with different languages such as Spanish or Slovenian languages. If the searched words and classes are prepared in Slovenian, the search can also be made in Slovenian.

5   Experimental results

5.1   User attitudes

When the experiment on user ratings and comments is observed, it is stated that most of the users thought positive about mobile security programs. Moreover, users have installed programs for different reasons such as privacy, security or freeware. When the comments have been separated by the sentiment analysis, the most used words have been shown by word cloud at each sentiment group. The most used positive (a), negative (b) and neutral (c) Turkish words are shown in Figure 3.
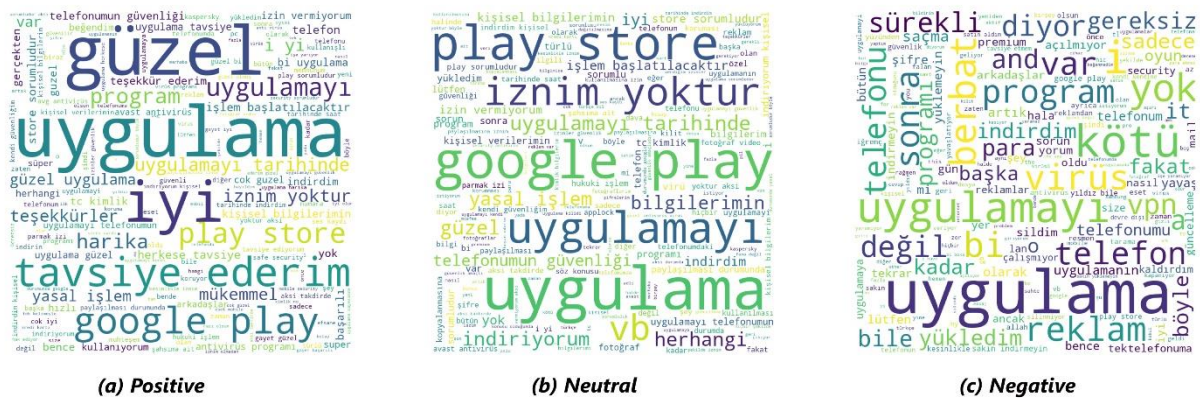
(a) Positive   *(b) Neutral*   *(c) Negative*

Figure 3: (a) Positive (b) Neutral (c) Negative Turkish words.

Categories	Positive User Attitudes		Neutral User Attitudes		Negative User Attitudes	
	Percentage	**Number**	**Percentage**	**Number**	**Percentage**	**Number**
Virus	28,1 %	9171	20,7 %	624	29,4 %	426
Security	23,7 %	7732	26,9 %	811	8,9 %	129
Privacy	22,7 %	7410	29,3 %	882	5,0 %	73
Payment	6,8 %	2220	6,0 %	180	12,4 %	179
Protection	5,0 %	1635	5,3 %	159	14,1 %	204
Advertisement	4,9 %	1590	4,9 %	149	8,7 %	126
Usage	3,3 %	1089	2,2 %	67	3,5 %	50
Update	2,5 %	827	2,3 %	69	8,8 %	128
Advice like/dislike	2,5 %	827	2,3 %	69	8,8 %	128
Freeware	0,3 %	99	0,1 %	4	0,3 %	4

Table 3: The positive (a), neutral(b) and negative (c) classified user attitudes.

When all comment texts are also investigated based on 10 categories, the user attitudes are also shown in a list for each sentiment group in Table 3. In the positive comments, it is observed that the users mostly preferred mobile security programs for privacy, security and cleaning virus. In the negative comments, it is observed that the users mostly complained about mobile security programs for privacy, security and cleaning virus. In the neutral comments, it is observed that the users mostly stayed neutral on mobile security programs about privacy and virus. The numbers and percentages of the positive, neutral and negative attitudes are also shown in Table 3.

The association probability of the words is also examined by n-gram analysis as unigram, bigram and trigram. The stop words such as "Google", "Play", "Store", "very", "for" are ignored at n-gram analysis. For all comments that follows the proposed method, the unigram, bigram and trigram analysis are shown in Table 4. The numbers of most used 10 Turkish words, pair-words, and triple-words are also shown in Table 4.

5.2 Model evaluation

In the study, the sentiment analysis based on linear SVM is performed with standard metrics which are precision, recall, and F1 score [29]. Moreover, the other classifiers that are sentiment analysis based on Naïve Bayes and Maximum Entropy are also applied for the performance test [30]. In our study, the comment text that has the positive, neutral and negative polarity are examined for the comment texts that follows the proposed method. The Linear SVM, Naïve Bayes, and Maximum Entropy classifiers are performed for three classes by using sklearn metrics library in Python. The non-normalization TP, FP, TN, and FN values that are required for the confusion matrix for each model are shown in Table 5. The total accuracy, precision, recall and F1-score values of the

UNIGRAM		BIGRAM		TRIGRAM	
Word (Translation)	Number	*Word* (Translation)	Number	*Word* (Translation)	Number
Uygulama (application)	10823	*Güzel uygulama* (beautiful, application)	2718	*Yasal işlem başlatılacaktır* (The juristicial act will be start)	520
Güzel (beautiful)	9277	*Tavsiye ederim* (I advice)	2297	*Herkese tavsiye ederim* (I advice everybody)	431
Uygulamayı (to application)	5833	*İznim yoktur* (I don't allow)	1270	*Uygulamayı telefonumun güvenliği* (application for security of my phone)	429
İyi (nice)	5729	*Yasal işlem* (juristicial act)	1016	*Telefonumun güvenliği indiriyorum* (I download for my phone security)	420
Ederim (I do)	3515	*Uygulama güzel* (application is beautiful)	960	*Uygulama tavsiye ederim* (application I advice)	357
Program (program)	3492	*Telefonumun güvenliği* (security of my phone)	914	*Uygulama güzel uygulama* (application beautiful application)	302
Kişisel (personal)	3335	*İşlem başlatılacaktır* (the act will be started)	822	*Güzel uygulama güzel* (Beautiful application beutiful)	292
Virüs (virüs)	3209	*İyi uygulama* (nice application)	746	*İznim yoktur aksi* (I don't allow, otherwise)	283
Sorumludur (responsible)	3035	*Kişisel bilgilerimin* (my personal info)	679	*Avast, antivirüs, 2021* (Avast, antivirus, 2021)	254
Tavsiye (advice)	3015	*Güvenliği indiriyorum* (I download security)	623	*Tavsiye ederim güzel* (I advice beautiful)	253

Table 4: The unigram, bigram and trigram analysis of the proposed method.

different classifiers (Naïve Bayes, Maximum Entropy and Linear SVM) are also shown in Table 6.

According to Table 6, while the accuracy of the Naïve Bayes classifier is 85 % with the proposed method, the accuracy of the Maximum Entropy is 90%. Our proposed model accuracy which is supported by linear SVM is %93. The performance results show that the proposed model that is supported by Linear SVM generates higher accuracy.

6   Conclusion

Nowadays, security plays an important role in every virtual platform. With the development of mobile phones, security concerns have become more important in mobile devices. The advantages such as usage of ease, fast accessibility of the information, and portability caused to use phones by the people at more.

After the Covid-19 pandemic, phone usage noticeably increased worldwide. As a result of this situation, the

Classifier	Sentiment	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)
Naïve Bayes	*Negative*	3	3531	3	186
	Neutral	37	3315	142	229
	Positive	3129	46	409	139
Maximum Entropy	*Negative*	11	3575	10	128
	Neutral	2	3492	5	225
	Positive	3348	18	348	10
Linear SVM	*Negative*	37	3594	42	49
	Neutral	7	3532	38	145
	Positive	3414	54	184	70

Table 5: The confusion matrix (non-normalization) values of the proposed model with Linear SVM, Naïve Bayes and Maximum Entropy.

Classifier	Metric	Negative	Neutral	Positive	Total Accuracy
Naïve Bayes	Precision	0,01	0,14	0,96	85%
	Recall	0,5	0,21	0,88	
	F1-Score	0,03	0,17	0,92	
	Support	85	198	3439	
Maximum Entropy	Precision	0,07	0,008	0,99	90%
	Recall	0,52	0,33	0,91	
	F1-Score	0,14	0,017	0,95	
	Support	87	158	3477	
Linear SVM	Precision	0,47	0,16	0,95	93%
	Recall	0,43	0,05	0,07	
	F1-Score	0,95	0,98	0,96	
	Support	86	152	3484	

Table 6: The performance results with Naïve Bayes, Maximum Entropy, and linear SVM.

mobile security requirement also increases at the same rate. Mobile phone users share their ratings about phone features and mobile application at application store markets and other platforms. The shared information has become of vital importance for the other users. By observing the user attitudes, both the developers can develop different security mobile software and the users can use the security application for the most desired request. Moreover, the user attitudes on the virtual world are not only important for security but also becomes an indispensable element for marketing, politics and society.

In our study, the usage of mobile security programs has been examined with machine learning methods and user attitudes on mobile security are also investigated by predetermined criteria. A user analysis has been made by considering the Turkish user attitudes on mobile security software as a prototype study. The study can also be expanded with user attitudes that are in different languages and topics.

Additionally, the user attitudes can be evaluated with higher accuracy by the development of machine learning techniques and a lot of prediction can be made by using the previous user attitudes. In the near future, mobile security will become an indispensable element of our lives with the minimization of communication tools and this situation will more required investigation of user attitudes.

References

[1]   N. Manochehri and M. Y. Alhinai (2006). "Mobile phone users attitude towards Mobile Commerce (m-commerce) and Mobile Services in Oman". 2nd IEEE/IFIP International Conference in Central Asia on Internet Tashkent, Uzbekistan, pp. 1-6, https://doi.org/10.1109/CANET.2006.279277.

[2]   V. Tambe, D. Chauhan, S. Kulal and S. Sherkhane (2018). "Offline Mobile Security". 2018 International Conference on Smart City and Emerging Technology (ICSCET), IEEE, Mumbai, India, pp.1-4, https://doi.org/10.1109/icscet.2018.8537303.

[3]   N. Clarke, J. Symes, H. Saevanee and S. Furnell, S. (2016). "Awareness of Mobile Device Security: A Survey of User's Attitudes". International Journal of Mobile Computing and Multimedia Communications (IJMCMC), IGI Global Publisher of Timely Knowledge7(1),pp.15-31, https://doi.org/10.4018/ijmcmc.2016010102.

[4]   C. Ozkan and K. Bicakci (2020). "Security Analysis of Mobile Authenticator Applications," 2020 International Conference on Information Security and Cryptology (ISCTURKEY), IEEE, Ankara, Turkey,pp.18-30, https://doi.org/10.1109/ISCTURKEY51113.2020.9308020.

[5]   J. Ophoff and M. Robinson (2014). "Exploring end-user smartphone security awareness within a South African context," 2014 Information Security for South Africa, IEEE, Johannesburg, South Africa, pp. 1-7, https://doi.org/10.1109/ISSA.2014.6950500.

[6]   Z. Zhou, C. Sun, J. Lu and F. Lv (2018). "Research and Implementation of Mobile Application Security Detection Combining Static and Dynamic", 2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), IEEE, Changsha, China, pp. 243-247, https://doi.org/10.1109/ICMTMA.2018.00065.

[7]   Z. Benenson, O. Kroll-Peters and M. Krupp (2012). "Attitudes to IT security when using a smartphone," 2012 Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE,Wroclaw, Poland, pp. 1179-1183.

[8]   M. Yasen and S. Tedmori (2019). "Movies Reviews Sentiment Analysis and Classification," 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), pp.860-865, https://doi.org/10.1109/jeeit.2019.8717422.

[9]   S. Ray (2017)., "Understanding Support Vector Machine (SVM) algorithm from examples", Analytic Vidhya, Retrieved from https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/

[10] K. Lavanya and C. Deisy (2017). "Twitter sentiment analysis usingmulti-class SVM," 2017 International Conference on Intelligent Computing and Control (I2C2), pp. 1-6, https://doi.org/10.1109/i2c2.2017.8321798.

[11] N. S. Huda, M. S. Mubarok and Adiwijaya (2019)."A Mlti-label Classification on Topics of Quranic Verses (English Translation) Using Backpropagation Neural Network with StochasticGradient Descent and Adam Optimizer," 2019 7th International Conference on Information and Communication Technology (ICoICT), IEEE, Kuala Lumpur, Malaysia, pp. 1-5, https://doi.org/10.1109/icoict.2019.8835362.

[12] M. S. Alsadi, R. Ghnemat and A. Awajan (2019)."Accelerating Stochastic Gradient Descent using Adaptive Mini-Batch Size," 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS), IEEE, Amman, Jordan,pp.1-7, https://doi.org/10.1109/ICTCS.2019.8923046.

[13] F. Kabir, S. Siddique, M. R. A. Kotwal and M. N. Huda (2015). "Bangla text document categorization using Stochastic Gradient Descent (SGD) classifier", 2015 International Conference on Cognitive Computing and Information Processing (CCIP), IEEE, Noida, India, pp. 1-4, https://doi.org/10.1109/CCIP.2015.7100687.

[14] A. Tripathy, A. Agrawal and S. Kumar Rath (2016). "Classification of sentiment reviews using n-gram machine learning approach", Expert Systems with Applications, pp. 117-126, https://doi.org/10.1016/j.eswa.2016.03.028.

[15] I. N. Dewi, R. Nurcahyo and Farizal (2020). "Word Cloud Result of Mobile Payment User Review in Indonesia," 2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA), Bangkok, Thailand, pp. 989-992, https://doi.org/10.1109/ICIEA49774.2020.9102048.

[16] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani and R. Budiarto (2020). "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," IEEE Access, vol. 8, pp. 90847-90861, https://doi.org/10.1109/ACCESS.2020.2994222.

[17] C. Liu, Y. Sheng, Z. Wei and Y. Yang (2018). "Research of Text Classification Based on Improved TF-IDF Algorithm," 2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE), IEEE, Lanzhou, China, pp. 218-222, https://doi.org/10.1109/IRCE.2018.8492945.

[18] V. Sundaram, S. Ahmed, S. A. Muqtadeer and R. Ravinder Reddy (2021). "Emotion Analysis in Text using TF-IDF," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) IEEE, Noida, India, pp. 292-297, https://doi.org/10.1109/Confluence51648.2021.9377159.

[19] T. Hasan, A. Matin and M. S. R. Joy (2020). "Machine Learning Based Automatic Classification of Customer Sentiment", 2020 23rd International Conference on Computer and Information Technology (ICCIT) IEEE, Dhaka, Bangladesh, pp.1-6, https://doi.org/10.1109/ICCIT51783.2020.9392652.

[20] M. Kumar Jain, D. Gopalani, Y. Kumar Meena and R. Kumar (2020). "Machine Learning based Fake

News Detection using linguistic features and word vector features," 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) IEEE, Prayagraj, India, pp. 1-6, https://doi.org/10.1109/UPCON50219.2020.9376576.

[21] Ne-Lexa (2020, March 8). google-play-scraper, https://github.com/Ne-Lexa/google-play-scraper.

[22] Abaker, A. A., & Saeed, F. A. (2021). A comparative analysis of machine learning algorithms to build a predictive model for detecting diabetes complications. Informatica, 45(1), 117-125. https://doi.org/10.31449/inf.v45i1.3111.

[23] Tiwari, P., Pandey, H. M., Khamparia, A., & Kumar, S. (2019). Twitter-based opinion mining for flight service utilizing machine learning. Informatica, 43(3),381-386.
https://doi.org/10.31449/inf.v43i3.2615.

[24] S. M. Jimenez Zafra, M. T. Martin Valdivia, E. Martinez Camara and L. A. Urena Lopez (2019). "Studying the Scope of Negation for Spanish Sentiment Analysis on Twitter," IEEE Transactions on Affective Computing, 10(1), pp. 129-141, https://doi.org/10.1109/TAFFC.2017.2693968.

[25] M. Rumelli, D. Akkuş, Ö. Kart and Z. Isik (2019). "Sentiment Analysis in Turkish Text with Machine Learning Algorithms," 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), IEEE, Izmir, Turkey, pp. 1-5, https://doi.org/10.1109/ASYU48272.2019.8946436.

[26] T. T. Zin, "Sentiment Polarity in Translation" (2020). 2020 IEEE Conference on Computer Applications (ICCA) IEEE, Yangon, Myanmar, pp. 1-6, https://doi.org/10.1109/ICCA49400.2020.9022831.

[27] Fang, X., Zhan, J. (2015). "Sentiment analysis using product review data". Journal of Big Data vol,2, 5 pp. 1-14. https://doi.org/10.1186/s40537-015-0015-2.

[28] F. Calefato, F. Lanubile, F.Maiorano, et al. (2018). "Sentiment Polarity Detection for Software Development". Empir Software Eng, Springer Link, 23, pp. 1352–1382. https://doi.org/10.1007/s10664-017-9546-9.

[29] Kaur, S., & Mohana, R. (2018). Prediction of sentiment from macaronic reviews. Informatica, 42(1),127-136.

[30] Gjoreski, H., & Kulakov, A. (2014). Machine learning approach for emotion recognition in speech. Informatica, 38(4), 377-383.