Digitalization and Management Innovation II
R.J. Dwyer (Ed.)

doi:10.3233/FAIA230711

Utilization of Transformer-Based Language Models in Understanding Citizens' Interests, Sentiments and Emotions Towards Public Services Digitalization

Gloria HRISTOVA[a,1] and Nikolay NETOV[a]
[a] *Faculty of Economics and Business Administration, Sofia University "St. Kliment Ohridski", Bulgaria*

Abstract. We live in an era of digital revolution not only in the industry, but also in the public sector. User opinion is key in e-services development. Currently the most established approaches for analyzing citizens' opinions are surveys and personal interviews. However, governments should focus not only on developing public e-services but also on implementing modern solutions for data analysis based on machine learning and artificial intelligence. The main aim of the current study is to engage state-of-the-art natural language processing technologies to develop an analytical approach for public opinion analysis. We utilize transformer-based language models to derive valuable insights into citizens' interests and expressed sentiments and emotions towards digitalization of educational, administrative and health public services. Our research brings empirical evidence on the practical usefulness of such methods in the government domain.

Keywords. Digitalization, e-government, sentiment analysis, natural language processing, transfer learning, transformer-based language models, emotion mining

1. Introduction

We live in an era of digital revolution. Our lives are being transformed by the idea of carrying out daily activities in a faster, cheaper and more convenient way. Digital transformation is underway not only in the industry, but also in the public sector. Policy makers are becoming increasingly aware of the need for sustainable digital processes in the administrative, health, educational and other public sectors. The COVID-19 pandemic was yet another strong evidence in support of this since remote access allows for faster services without physical contact between citizens and administrative officers. Currently, e-government development is a top priority in the political agenda globally.

Aimed at tracking the digital transformation of governments worldwide, the United Nations organization developed the "E-Government Development Index" (EGDI). Bulgaria is ranked 52th among 193 countries with EGDI=0.7766 in 2022. For comparison,

[1] Corresponding author: Gloria Hristova, Faculty of Economics and Business Administration, Sofia University "St. Kliment Ohridski", Bulgaria; E-mail: g.hristova@feb.uni-sofia.bg.
The presentation and dissemination of these research results is supported in part by National Science Fund Project КП-06-Н45/3/30.11.2020 "Identifying citizens' attitudes and assessments about access, quality, and usage of electronic public services".

the top three countries are Denmark with EGDI=0.9717, Finland with EGDI=0.9533 and Republic of Korea with EGDI=0.9529. Digital transformation in leading countries includes the utilization of modern technologies as artificial intelligence (AI) and big data analysis [1]. One of the components of EGDI is the Online Services Index (OSI) which measures the scope and quality of online services in a given country. OSI is measured by experts in assessing public sector online services who analyze how quickly and intuitively the average user finds information and features on government sites and electronic portals. The value of OSI for Bulgaria is below the average measured in sub-region and region countries which reveals that more efforts should be put into the development of user-friendly and intuitive online public services.

User opinion is key in services digitalization. Digital revolution affects all aspects of daily life, including communication - in the last few years, social networks became a main place to share opinions about products and services. Currently, the most established approaches for analyzing citizens' opinions on public services are surveys and personal interviews. However, opinions expressed by individuals on social networks can provide valuable insights into public behavior, which may not be immediately apparent through traditional means of data collection. The utilization of such unstructured data often characterized with big volume is a serious challenge for governments. The last underlines the need to develop and assess different methodologies for knowledge extraction from unstructured data in the government domain.

Modern technologies for text analytics and natural language processing (NLP) might be the key to overcoming some of these real-world challenges in the public sector. Such approaches can enable authorities to perform big data analysis and to quickly and automatically identify trends in public behavior, opinions, and needs, which can help inform government decisions and improve satisfaction with public services. Our research is focused on exploring this topic in more depth and the main aim of the current study is to investigate the applicability and practical usefulness of state-of-the-art NLP technologies for sentiment analysis and topic modeling in gaining valuable insights on public opinion. Previous research is focused on the application of traditional machine learning algorithms for opinion mining. However, our study makes an attempt to overcome some of the challenges in the field by developing an analytical approach for public opinion analysis based on pre-trained transformer-based language models.

In the last few years, we have witnessed the emergence of many advances in the NLP field based on Transformer models and transfer learning [2]. Our research brings empirical evidence on the practical usefulness of such methods in the government domain. Our methodology is tested empirically within a case study on mining citizens' opinions on Bulgarian e-government. To accomplish our main aim, we scrape a large sample of public comments posted in popular Bulgarian social networks in an attempt extract valuable insights into citizens' interests and expressed sentiments and emotions towards digitalization of educational, administrative and health public services. To the best of the authors' knowledge, the current empirical study is the first applying a combination of transformer-based language models for sentiment analysis and topic modeling of citizens' opinion in the Bulgarian government sector.

2. Related work

In this section we provide a review of related work. First, we present an overview of the most recent research in applying NLP for public opinion analysis in the e-government domain. Next, we provide a brief review of recent advances in sentiment analysis and topic modeling since these two techniques are at the core of our methodology.

2.1. Analysis of Public Opinion on Government E-services

Verma provides the first comprehensive review of research applying sentiment analysis for smart society [3]. Findings suggest that research on social benefits of citizen opinion mining for e-governance is still in a developing phase and will gain more and more traction among the research community [3]. Furthermore, social networks are becoming "citizen-government collaboration platforms". An overview of machine learning utilization within governmental applications is provided in [4]. Authors suggest that social media sentiment analysis and application of machine learning techniques, in general, allows governments to understand better the needs of citizens, while the last get the opportunity to influence the process of service development. However, working with unstructured big data is considered a particularly challenging task in this domain [4].

N'Diaye et al. propose a system aimed at enhancing the effectiveness of e-government [5]. The system tracks public sentiments towards government decisions and policies utilizing data from Facebook pages related to the Mauritania's government. In [6] is developed a fine-grained sentiment analysis method based on deep learning techniques and aimed at understanding aspects of citizens' opinions on government apps. Al-Qudah et al. employ opinion mining techniques to explore citizens' attitudes towards an e-payment service in Jordan used for telecommunication bills, public utilities etc. [7]. Andoh et al. perform a comparison between several supervised machine learning algorithms for mining public sentiments expressed towards the Ghanaian Government in Twitter [8]. In [9] is proposed an approach based on sentiment analysis and topic modeling techniques aimed at understanding public opinion on governments implementing digital contact tracing to prevent the spread of COVID-19. In [10] are studied the advantages and limitations of using a novel topic modeling algorithm in big data analysis in the public domain. Drawn are useful insights into the general perception of distance learning in Bulgaria in light of the COVID-19 pandemic. In [11] is presented the general framework and design of a conceptual model of ML-based AI system for mining public opinion on e-government services.

To the best of the authors' knowledge, this is the first study aimed at applying a combination of transformer-based language models for mining citizens' interests, attitudes and emotions towards e-government development and e-services provision in Bulgaria. This finding is also supported by previous research in the field [12, 13].

2.2. Recent Advances in Sentiment Analysis and Topic Modeling of Textual Data

Sentiment analysis has numerous applications in various domains when people's opinion on various subjects plays a crucial role in decision-making and management. The increased awareness of policy-makers about the importance of interaction with the public in social networks recently led to applications of sentiment analysis in the government domain too [3,13]. Sentiment analysis can be divided into two general categories, namely - opinion mining and emotion mining [14]. While opinion mining mainly aims at understanding the polarity of expressed attitudes (dividing them to positive, negative or neutral), emotion mining aims at capturing finer aspects of people's opinions (for example, expressions of fear, anger, joy etc.) [14]. Regardless of their main differences, both types of sentiment analysis could be performed using analogous techniques.

For many years, the traditional approach for sentiment analysis mainly consisted of the utilization of sentiment lexicons [15]. They are fast, suitable for large volumes of data and yield results that are easily interpretable. Nevertheless, a considerable limitation is that lexicons are restricted to the words they contain and results might highly depend on data domain. In case labeled data is available, various machine learning algorithms [15] might be used in both opinion and emotion mining. In the development of the field, several challenges have emerged (more through discussion on the subject could be found in [15]). First, domain adaptation is crucial - models trained on one data domain may not perform well in a different domain since language and sentiment expressions can vary significantly. Another challenge stems from the notion of contextual understanding - models might struggle to understand the context of a given text, such as sarcasm or irony, which can lead to misinterpretation of the sentiment. Another problem is limited labeled data – development of robust models often requires large amounts of labeled data, which could be expensive and very time-consuming to obtain.

The Transformer architecture [2] and applications of transfer learning led to the development of models with state-of-the-art performance on many NLP tasks. Such models have the potential to overcome some of the mentioned challenges in sentiment analysis and are a promising research direction to improve the accuracy and applicability of current sentiment analysis methods [16]. Recently, we observe an ever-increasing volume of pre-trained models that can be used as an "off-the-shelf" solution for a wide range of problems including opinion and emotion mining. Such models could be applied in completely unsupervised settings. One novel pre-trained model for sentiment analysis is SiEBERT which potential is demonstrated in [16, 17, 18, 19].

Topic modeling described by some authors as a "revolution in text mining" [20] is another powerful NLP method that has many practical applications. In the early work of DiMaggio et al. [21] are presented prominent ideas revealing the potential of topic modeling in the government domain. Topic extraction is essentially a machine learning technique used for discovering the semantic structure and hidden knowledge in large volumes of textual data [20]. Fundamental algorithms for non-probabilistic topic modeling are Latent Semantic Analysis and Non-Negative Matrix Factorization [20]. Probabilistic approaches like Latent Dirichlet Allocation (LDA) aim to improve the algebraic methods. Until the "transformers revolution" in artificial intelligence, LDA was probably the most popular approach for topic modeling. Although not as fast as in the field of sentiment analysis, new approaches for topic modeling based on transformer language models also emerged. The potential of the novel BERTopic algorithm [22] has been demonstrated in an increasing volume of studies [10, 17, 23, 24, 25, 26].

In a nutshell, prior to the emergence of the transformer-based architecture and large-scale pre-trained models, NLP methods relied mostly on rule-based approaches, statistical models, and machine learning algorithms, which often required significant manual effort and engineering to achieve good performance [27]. However, pre-trained models have overcome some of these limitations to a given extent and introduced new capabilities in both sentiment analysis and topic modeling. Several advantages of such models compared to other NLP methods might be outlined [27]. The most important among these are better generalizability and contextual understanding of textual data since such models are trained on large volumes of data and their intrinsic architecture allows them to capture the nuances and complexities of natural language more accurately than other methods. Furthermore, when it comes to domain adaptation, transformer models have shown promise in adapting to new domains with minimal additional training data.

3. Methodology

This section describes the proposed NLP approach for analyzing citizens' interests, sentiments and emotions towards public e-services. Our methodology is tested empirically on comments about digital educational, administrative and health public services in Bulgaria. Some preliminary results based solely on keywords data filtering without the usage of topic modeling and pre-trained language models in sentiment extraction could be found in [28].

3.1. Sample Development and Initial Text Data Processing

To test empirically our NLP approach, we scrape a large sample of textual data posted in diverse online media platforms in Bulgaria according to a list of specific keywords related to educational, administrative and health electronic public services. The data scraping process we follow extracts not only public comments, but also news or other types of textual data related to the chosen types of e-services. Therefore, initially scraped data should be filtered to contain only citizen comments. Next, various text data preparation techniques are applied, including – 1. Text data deduplication; 2. Translation to English (this step is necessary since the chosen pre-trained language model for polarity categorization works only on English data); 3. Case normalization; 4. Removal of text elements considered as not important for subsequent analyses, including - HTML tags, URLs, punctuation, digits, stop words; 5. Word tokenization.

3.2. Opinion Mining with SiEBERT

Our approach for opinion mining leverages the power of pre-trained language models. As mentioned earlier, such methods are very efficient in capturing text semantics and are very suitable in unsupervised tasks when little or no training data is available. We utilize the language model SiEBERT [29] for extraction of citizens' expressed sentiments towards digital public services. Hartmann et al. [29] suggest that the language model RoBERTa is better suited for sentiment analysis tasks compared to other popular (but conceptually very different) transfer learning models (BERT, ULMFiT and XLNet). Therefore, the authors propose SiEBERT which is a fine-tuned checkpoint of RoBERTa [30] for sentiment polarity classification – the model is trained on a large volume of texts from publicly available sentiment datasets. SiEBERT classifies English texts into two sentiment categories ("positive" and "negative"). The main reason behind choosing SiEBERT in our study is that it is a "general-purpose" sentiment analysis model, which means that it can be applied on textual data from different domains. The model is trained on a large volume of various types of textual data in different domains (for example, tweets, social media, reviews on goods and services etc.). SiEBERT reaches an average accuracy of 93.2% on previously unseen data, as reported by Hartmann et al. [29].

3.3. Emotion Mining with NRC Emotion Lexicon

We employ the NRC Emotion Lexicon in the analysis of citizens' emotions. The reason behind choosing a lexicon-based approach for emotion detection is that currently available pre-trained models capture 6 basic emotions – anger, disgust, fear, joy, sadness and surprise. However, the NRC Emotion Lexicon which is probably the most popular lexicon used by researchers [31, 32, 33] for emotion mining captures all these emotions plus additional two – “anticipation” and “trust”. Based on evidence from previous research [34, 35], we believe that these two emotions, especially “trust”, are crucial in usage of digital public services and e-government development. The NRC Emotion Lexicon is general-purpose, meaning it could be applied on text data in various domains, including social networks data – more information on the development of the lexicon is provided in [36].

Each word part of the lexicon might be related to more than one of the eight emotions. Applying the lexicon, we generate a list of associated emotions represented in each comment. One considerable limitation is that the text might contain emotions expressed by citizens using words that are not part of the dictionary. Consequently, these emotions will not be detected using this approach. We calculate the frequency of appearance of each emotion for each public comment d_i, applying the following formula:

$$S_i(\text{"}trust\text{"}) = \frac{p_i}{Q_i}, \text{ where} \tag{1}$$

d_i - a given public comment part of the sample, $i \in 1 \ldots N$.
N - overall number of public comments in the sample.
p_i - number of terms in d_i related to the “trust” emotion. “Trust” is used only as an example - the calculation is the same for rest of the emotion categories.
Q_i - count of all terms in d_i related to each of the eight emotions (a term might be counted several times if it is associated to more than one emotion category).

After calculating S_i for each emotion and public comment in the sample, we categorize texts based on the dominant emotions - as such are considered those emotions with highest value of S_i in a given comment. If S_i has an equal value for several emotions, then the given document will be labeled as expressing all these emotions. If $S_i = 0$ for a given emotion, this means that d_i does not contain any terms related to this emotion.

3.4. Topic Modeling with BERTopic

Similar to our approach for opinion polarity categorization, in order to capture citizens’ main interests expressed in discussions about digital public services, we employ a topic extraction technique that utilizes pre-trained language models. BERTopic, proposed by Grootendorst in 2022, is a neural topic modeling algorithm with increasing popularity among the research community [22]. The algorithm employs the Sentence-BERT (SBERT) framework for document embeddings. We use the “all-MiniLM-L6-v2” model for turning public comments into vector representations as suggested by the author of BERTopic [37]. Topic extraction is performed by applying the HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) algorithm on document embeddings. Each of the generated clusters of public comments is considered a “topic”.

Key element in the application of BERTopic is the development of topic representations. For this purpose, the algorithm employs a class-based TF-IDF (term frequency-inverse document frequency) procedure which helps in the subjective task of “labeling” topics. BERTopic develops cluster-word representations utilizing a modification of the well-known information retrieval metric TF-IDF. Eq. (2) illustrates the calculation of c-TF-IDF used to form topic representations:

$$\text{c-TF-IDF}(t,c) = tf_{t,c} \times \log\left(1 + \frac{A}{f_t}\right), \text{ where} \tag{2}$$

$tf_{t,c}$ – frequency of term t in cluster C.
A - average number of words per cluster.
f_t - frequency of term t across all clusters.

Eq. (2) is used to calculate the importance of words with respect to clusters - c-TF-IDF(t,c) is interpreted as the weight of term t in cluster C. Applying Eq. (2) enables us to extract the most important words within each cluster and use them to infer the topic of public comments. Using HDBSCAN for topic extraction might lead to the generation of many "outlier" documents that are not assigned to any of the generated clusters. However, since our samples are small in size, we utilize a simple technique for outlier reduction as suggested by Grootendorst [22] – "outlier public comments" are assigned to the most similar to them topics based on their c-TF-IDF representations.

4. Empirical study

The proposed methodology for public opinion analysis is tested empirically within a case study in the government sector of Bulgaria (ranked 52th in EGDI ranking). As a first step, we scraped a large sample of text data related to electronic public services and posted on various online platforms in Bulgaria. However, the size of the sample decreased substantially after initial data cleaning, deduplication and removal of observations that are not citizen comments (for example, news, announcements in government and radio websites). Furthermore, to develop a more focused sample, we additionally filtered comments about education according to selected keywords related to digital educational services. Table 1 provides details on the number of observations in each text data sample (related to educational, administrative and health services, respectively). Figure 1 presents word cloud graphs revealing the general context of discussions in each category. Public comments were posted in different Bulgarian social networks between 2019 and end of 2021. This period includes the COVID-19 pandemic providing an important context to our analysis. We assume that the pandemic had a strong impact on discussions on the existence and current condition of public e-services.

To extract the main topics of citizen interest related to each type of public service, we performed several experiments by fine-tuning important hyper-parameters of BERTopic. These include: 1. Diversification of topic representations (fine-tuned by calculating Maximal Marginal Relevance); 2. Automatic topic reduction (HDBSCAN is used to reduce topics based on their similarity); 3. Text vectorization (fine-tuning the n-gram range used in developing topic representations); 4. Minimum topic size (fine-tuning the minimum size of the generated topics). The following three sections present the main results from the analysis of citizens' interests, sentiments and emotions.

Table 1. Number of citizen comments in each of the analyzed samples of textual data

Digital Educational Services	Digital Administrative Services	Digital Health Services
2748	644	202

Figure 1. Word cloud graphs revealing the general context of public comments related to electronic educational, administrative and health services.

4.1. Digital Educational Services

Most of the public comments in our sample discuss educational services. Initially, the application of BERTopic on data combined with hyperparameter tuning led to the extraction of 34 topics related to electronic educational services. Based on a careful analysis of keywords and intertopic distance map generated with the help of the UMAP algorithm, we manually merged some of the topics that are very related to each other. Thus, the final model contains 25 topics revealing different aspects of citizens' interests related to digital educational services. Next, we applied sentiment and emotion mining on public comments as described in the "Methodology" section of the current paper and analyzed the results. In Table 2, we report seven of the most interesting topics revealed in discussions on educational e-services. The first topic is mainly about distance education for children (in general). Almost 50% of all comments related to educational e-services fall under this topic - a possible explanation is the application of outlier reduction which means that many of the comments initially treated as outliers by HDBSCAN were probably merged to this "more general" topic. Figure 2 reveals the distribution of positive vs. negative comments in each of the analyzed in detail topics. We observe that 60.13% of all comments in Topic 1 are negative.

The BERTopic algorithm discovered citizen comments related to learning in an electronic environment (Topic 2), government measures to curb the COVID-19 pandemic applied in class, like wearing masks and quarantine (Topic 3), COVID-19 spread in schools (Topic 4), university e-learning (Topic 5), technical aspects of e-learning (Topic 6) and discussions regarding distance education in other countries (Topic 7). According to our analysis, citizens have expressed mainly positive sentiments on learning in an electronic environment, university e-learning and technical aspects of e-learning. Negative sentiments dominate in discussions about public policies to curb the COVID-19 pandemic applied in schools and distance education in other countries. Emotion mining reveals that citizens have expressed their opinions by mainly using words associated with trust and anticipation (Figure 3). This observation is valid not only for public comments on educational e-services, but also for administrative and health e-services. An interesting observation is that university e-learning is associated with equal presence of joy and fear in public comments. Figure 3 allows drawing additional diverse insights into the emotions contained in citizens' comments about educational e-services.

4.2. Digital Administrative Services

BERTopic initially extracted 50 topics of citizen interest regarding administrative e-services. Following the same analytical procedure described earlier, we merged some of the most similar topics to develop a final model containing 27 topics strongly focused on administrative e-services. Table 2 reveals four of the most frequently discussed topics in this sample. Topic 8 is about e-government systems and usage of electronic signature. Again, trust is the dominant emotion in comments implying that utilization of e-services depends a lot on the level of trust in them. We observe higher levels of anger and sadness in Topic 8 compared to the rest. Public comments in Topic 9 contain predominantly negative sentiments and focus on the political aspects of e-governance and important political figures and parties. Topic 10 is devoted to electronic voting and we observe a prevalence of positive attitudes - again, citizens have used word expressions mainly associated with trust. As observed in discussions regarding educational e-services, we find a topic containing comments on government measures aimed at limiting the spread of COVID-19 in general (Topic 11). Fear, trust and sadness are the dominant emotions in this topic with sentiments being almost equally divided between positive and negative.

4.3. Digital Health Services

The sample of comments related to e-health is rather small in size compared to rest of the samples. Nevertheless, BERTopic managed to extract 6 topics of citizen interest and four of them are displayed in Table 2. We found citizens' comments related to e-health services in Bulgaria in general (Topic 12) and usage of electronic health documents like e-prescriptions (Topic 13). Expressed sentiments towards these two topics are rather negative. Topic 14 is focused on the digital aspects of COVID-19 testing and vaccination (issue of digital certificates). Topic 15 contains comments about digital health platforms and mentions the name of a specific Bulgarian company providing telemedicine services. Apart from the last topic in which citizens' sentiments are mostly positive, in rest of the topics related to e-health services in Bulgaria, citizens have expressed mainly negative attitudes. Emotion mining reveals that fear and sadness are far more dominant in this sample. In Topic 12 and Topic 14 we observe a more balanced distribution of all the eight emotions under analysis (compared to rest of the topics). A possible interpretation is that citizens might have been more expressive when discussing health services.

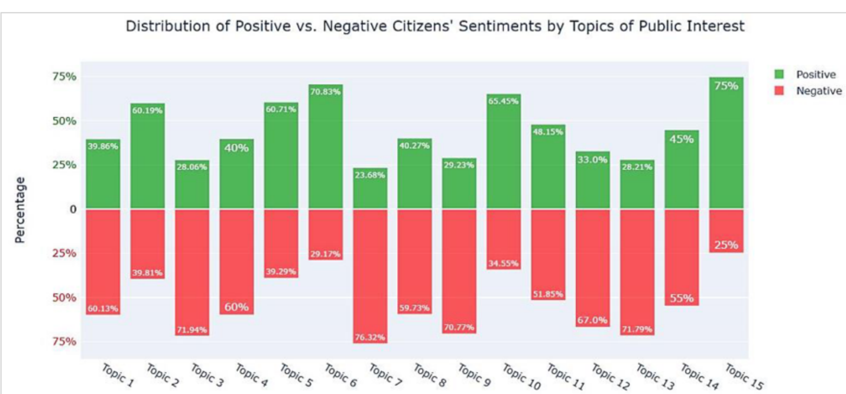

Figure 2. Visualization of the distribution of citizens' sentiments (positive vs. negative) by topics

Table 2. Main topics of citizen interest regarding digital educational, administrative and health public services

	Topics and topic representations (associated keywords and phrases)
Digital Educational Public Services	**Topic 1**: Distance education for children (in general) - *learning, children, school, distance, distance learning, online, going, teachers, time, child*
	Topic 2: Learning in an electronic environment – *school(s), students, education, training, electronic, environment, learning, electronic environment, minister*
	Topic 3: Government measures to curb the COVID-19 pandemic applied in class - *children, school(s), masks, quarantined, students, teachers, child, distance, class*
	Topic 4: COVID-19 spread in schools (*people, measures, minister, health, infected, coronavirus, number, students, cases, new*)
	Topic 5: University E-learning (*lectures, online lectures, online, exercises, attendance, lectures exercises, faculty, Sofia, book, even*)
	Topic 6: Technical aspects of E-learning (*laptop, learning, AMD, RAM, distance learning, distance, used, display, Ryzen, SSD*)
	Topic 7: Distance education in other countries (*school(s), closed, children, students, Germany, Sweden, distance, learning, kindergartens*)
Digital Administrative Public Services	**Topic 8**: E-government systems and usage of electronic signature (*electronic, system, health, service(s), signature, e-government, minister, electronic signature, order*)
	Topic 9: Political figures and parties *(Borisov, GERB, people, government, Bulgaria, would, power, party, like, one)*
	Topic 10: E-voting *(election(s), voting, electronic, vote, e-government, code, act, Bulgaria, system)*
	Topic 11: Government measures to curb the COVID-19 pandemic in general *(measures, people, crisis, situation, virus, municipalities, pandemic, one, need, time)*
Digital Health Public Services	**Topic 12**: E-health services in general *(people, health, electronic, time, system, many, years, data, even, doctors)*
	Topic 13: E-health documents *(electronic, prescription(s), pharmacy, recipes, ordinance, electronic prescriptions, paper, doctors, electronic recipes)*
	Topic 14: Digital solutions/processes for COVID-19 testing and vaccination *(test(s), vaccination, people, certificate, get, end, health, electronic, covid)*
	Topic 15: E-health systems/platforms *(health, care, change, medical, system, model, health care, establishment, doctors, Healee)*

Note: The reader should note that uniform numbering of topics is used throughout the work to avoid confusion.

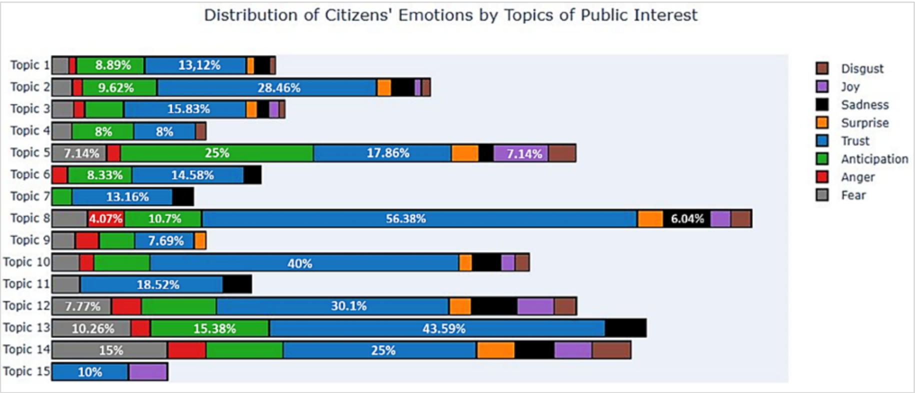

Figure 3. Visualization of the distribution of citizens' emotions by topics. Emotions are represented in different colors and their dominance is measured as a percentage. For example, "trust" is the dominant emotion in 56.38% of all comments in Topic 8 (percentages are displayed only for the more dominant emotions).

5. Discussion and Conclusions

We accomplished the main aim of the current study by proposing an analytical approach for public opinion analysis based on pre-trained transformer-based language models. We brought empirical evidence on the practical usefulness of such approaches in the government domain by presenting a case study on understanding citizens' interests, sentiments and emotions towards educational, health and administrative e-services in Bulgaria. Our approach could be implemented in a completely unsupervised way, unlike traditional machine learning methods which would require the availability of training data. The BERTopic algorithm discovered specific and meaningful topics of public interest. Nevertheless, a comparison between BERTopic and other approaches for topic extraction should be performed since this would allow a thorough exploration of the efficiency of the algorithm, and how it compares to other approaches in terms of accuracy, scalability, and computational resources required. This is a promising direction for future research, and has the potential to contribute to the development of more efficient techniques for analyzing and understanding large-scale text data.

Our study reveals that despite their limitations, emotion lexicons combined with the powerful pre-trained language models might provide very important context into the analysis of public opinion. Overall, the ability to make different intersections in the analysis is a valuable feature of the proposed approach, as it allows to gain deeper insights into citizens' opinions and the patterns that exist within them. Our methodology might also be applied in studying the public opinion on other digital services, important issues or emerging problems in the government domain. Furthermore, our approach utilizes data from different social networks used in the country, unlike other studies in the field that use mainly Twitter data for studying public opinion [13].

Our analysis revealed different topics of public interest related to digital educational, administrative and health services. Citizens have expressed mainly negative sentiments towards the discovered topics. This finding is not surprising since e-government and usage of public e-services are very controversial topics in Bulgaria. Furthermore, the political environment and government instability in the last few years are other important factors that might explain citizens' negative attitudes. Results from emotion mining indicate that citizens have expressed their opinions using words mainly related to trust and anticipation. Trust is a critical factor in the adoption of digital public services [34, 35]. If citizens do not trust the services, they might be hesitant to use them or may not use them at all, which can hinder the overall digital transformation efforts. Therefore, building trust in electronic public services is essential to ensure their successful adoption and implementation.

Finally, it is important to outline some of the limitations of the current study. First, we should note that not all Bulgarian citizens have access to or regularly use social networks, so the opinions and behaviors of this particular subset of the population studied in the current work may not accurately represent the broader population. Another limitation stems from the usage of specific keywords in data scraping since there is always a risk of missing relevant information and not capturing the full range of opinions related to the topic under study. A potential limitation of using the NRC lexicon for emotion analysis is that it might not be equally comprehensive of all the emotion categories that it includes. The last might lead to potential biases in the emotion analysis results. To mitigate this limitation, as a future direction of our study we plan to consider other emotion lexicons and combine them with approaches based on transfer learning.

Human annotation might also be very helpful in gaining a more comprehensive understanding of the emotions presented in public comments.

The combination of state-of-the-art NLP technologies for sentiment analysis and topic modeling led to the extraction of valuable insights into citizens' opinions and we believe that such an approach might successfully complement traditional approaches for public opinion analysis (e.g., surveys and interviews). Furthermore, results might be useful in designing more relevant questionnaires accurately reflecting current issues in the public domain. The fusion of knowledge obtained through diverse statistical techniques (traditional vs. NLP methods) would lead to even more informative insights into people's attitudes, expectations and interests. Our results might be found useful by social science researchers, policy-makers and experts involved in implementing modern techniques for data analysis and artificial intelligence in the government domain.

References

[1] E-Government Survey 2020: Digital Government in the Decade of Action for Sustainable Development. United Nations, Department of Economic and Social Affairs; 2020.

[2] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Advances in neural information processing systems. 2017; 30.

[3] Verma S. Sentiment analysis of public services for smart society: Literature review and future research directions. Government Information Quarterly. 2022; 39(3): 101708.

[4] Alexopoulos C, Lachana Z, Androutsopoulou A, Diamantopoulou V, Charalabidis Y, Loutsaris MA. How machine learning is changing e-government. In Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance; 2019, April. p. 354-363.

[5] N'Diaye ACM, Chrif MEMEA, El Mahmoud BM, El Beqqali O. Apply sentiment analysis technology in social Media as a tool to enhance the effectiveness of e-government: Application on Arabic and Mauritanian dialect 'HASSANIYA'. In 2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS); 2021, October: IEEE. p. 1-5.

[6] Rongxuan S, Bin Z, Jianing M. End-to-End Aspect-Level Sentiment Analysis for E-Government Applications Based on BRNN. Data Analysis and Knowledge Discovery. 2022; 6(2/3): 364-375.

[7] Al-Qudah DA, Ala'M AZ, Castillo-Valdivieso PA, Faris H. Sentiment analysis for e-payment service providers using evolutionary extreme gradient boosting. IEEE Access. 2020 Oct 19; 8: 189930-44.

[8] Andoh J, Asiedu L, Lotsi A, Chapman-Wardy C. Statistical analysis of public sentiment on the ghanaian government: a machine learning approach. Advances in Human-Computer Interaction. 2021; 2021: 1-7.

[9] SV P, Ittamalla R. General public's attitude toward governments implementing digital contact tracing to curb COVID-19–a study based on natural language processing. International Journal of Pervasive Computing and Communications. 2022; 18(5): 485-490.

[10] Hristova G, Netov N. Media Coverage and Public Perception of Distance Learning During the COVID-19 Pandemic: A Topic Modeling Approach Based on BERTopic. In 2022 IEEE International Conference on Big Data (Big Data); 2022 Dec 17: IEEE. p. 2259-2264.

[11] Hristova G, Bogdanova B, Netov N. Design of ML-based AI system for mining public opinion on e-government services in Bulgaria. In AIP Conference Proceedings; 2022, September: AIP Publishing LLC. p. 020005.

[12] Hristova G. Text Analytics in Bulgarian: An Overview and Future Directions. Cybernetics and Information Technologies. 2021; 21(3): 3-23.

[13] Hristova G, Bogdanova B, Netov N. Data mining of public opinion: An overview. In AIP Conference Proceedings; 2022, September: AIP Publishing LLC. p. 020004.

[14] Yadollahi A, Shahraki AG, Zaiane OR. Current state of text sentiment analysis from opinion to emotion mining. ACM Computing Surveys (CSUR). 2017; 50(2): 1-33.

[15] Wankhade M, Rao ACS, Kulkarni C. A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review. 2022; 55(7): 5731-5780.

[16] Vitman O, Kostiuk Y, Sidorov G, Gelbukh A. Sarcasm Detection Framework Using Emotion and Sentiment Features. arXiv:2211.13014v2 [cs.CL]. 2022.

[17] Ng QX, Lim SR, Yau CE, Liew TM. Examining the Prevailing Negative Sentiments Related to COVID-19 Vaccination: Unsupervised Deep Learning of Twitter Posts over a 16 Month Period. Vaccines. 2022; 10(9): 1457.

[18] Rozado D, Hughes R, Halberstadt J. Longitudinal analysis of sentiment and emotion in news media headlines using automated labelling with Transformer language models. Plos one. 2022; 17(10): e0276367.

[19] Beese D, Altunbaş B, Güzeler G, Eger S. Did AI get more negative recently?. Royal Society Open Science. 2023; 10(3): 221159.

[20] Kherwa P, Bansal P. Topic modeling: a comprehensive review. EAI Endorsed transactions on scalable information systems. 2019; 7(24).

[21] DiMaggio P, Nag M, Blei D. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. Poetics. 2013; 41(6): 570-606.

[22] Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv:2203.05794 [cs.CL]. 2022.

[23] Abuzayed A, Al-Khalifa H. BERT for Arabic topic modeling: An experimental study on BERTopic technique. Procedia computer science. 2021; 189: 191-194.

[24] Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Frontiers in Sociology. 2022; 7: 886498.

[25] Alqahtani E, Janbi N, Sharaf S, Mehmood R. Smart homes and families to enable sustainable societies: A data-driven approach for multi-perspective parameter discovery using bert modelling. Sustainability. 2022; 14(20): 13534.

[26] Ebeling R, Sáenz CA, Nobre JC, Becker K. Analysis of the influence of political polarization in the vaccination stance: the Brazilian COVID-19 scenario. In Proceedings of the International AAAI Conference on Web and Social Media; 2022. p. 159-170.

[27] Kalyan KS, Rajasekharan A, Sangeetha S. Ammus: A survey of transformer-based pretrained models in natural language processing. arXiv:2108.05542 [cs.CL]. 2021.

[28] Hristova G, Netov N. Analysis of Public Sentiments and Emotions in the Government Domain. Industry 4.0. 2023; 8(1): 32-5.

[29] Hartmann J, Heitmann M, Siebert C, Schamp C. More than a feeling: Accuracy and application of sentiment analysis. International Journal of Research in Marketing. 2022.

[30] Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. Roberta: A robustly optimized bert pretraining approach. arXiv:1907.11692 [cs.CL]. 2019.

[31] Tokarchuk O, Barr JC, Cozzio C. How much is too much? Estimating tourism carrying capacity in urban context using sentiment analysis. Tourism Management. 2022; 91: 104522.

[32] Agrawal SR, Mittal D. Optimizing customer engagement content strategy in retail and E-tail: Available on online product review videos. Journal of Retailing and Consumer Services. 2022; 67: 102966.

[33] Borghi M, Mariani MM. The role of emotions in the consumer meaning-making of interactions with social robots. Technological Forecasting and Social Change. 2022; 182: 121844.

[34] Alkraiji A, Ameen N. The impact of service quality, trust and satisfaction on young citizen loyalty towards government e-services. Information Technology & People. 2022; 35(4): 1239-1270.

[35] Lee T, Lee BK, Lee-Geiller S. The effects of information literacy on trust in government websites: Evidence from an online experiment. International Journal of Information Management. 2020; 52: 102098.

[36] Mohammad SM, Turney PD. Crowdsourcing a word–emotion association lexicon. Computational intelligence. 2013; 29(3): 436-465.

[37] Reimers N, Gurevych I. Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv:1908.10084 [cs.CL]. 2019.