Emotions Mining Research Framework: Higher Education in the Pandemic Context

Radka Nacheva

Abstract The pandemic situation in 2020 was a challenge for the organization of the educational process in higher education. The crisis has exacerbated inequalities between universities, which funding, digital sustainability and emergency training are weaker than their national and international competitors. The poor provision of the learning process in an electronic environment has led to a number of problems. Some of them are related to the acquisition of learning material and practical skills, lack of communication between students and academic staff. The increased use of the Internet during periods of social distance has also led to an increase in participating in social media activities, which have become forums for sharing opinions and expressing emotions through text and multimedia content. In this regard, the aim of the article is to propose a research framework for evaluation of emotional attitudes in social media. The author tested the practical applicability of the proposed framework by retrieving data from the social network Twitter and applying data mining techniques for analyzing large volumes of textual content.

Keywords Emotions mining · Text mining · Social networks · Higher education · Pandemic

1 Introduction

The pandemic, which began in 2020, has faced many challenges for humanity. Many sectors have been severely affected by lockdown in countries, leading to the closure of companies or the reduction of employment in a number of sectors. The consequences are a lessening in company revenues, rising unemployment, deteriorating credit ratings of citizens and businesses and even loss of property. The International Monetary Fund (IMF) reports a deepening recession in most countries in 2020—an average of 4.4% shrinking in the global economy (International Monetary Fund 2020). According to the World Bank, that is the worst reported since World War II

R. Nacheva (✉)
University of Economics – Varna, 77, Knyaz Boris I Blvd., Varna, Bulgaria
e-mail: r.nacheva@ue-varna.bg

M. K. Terzioğlu (ed.), *Advances in Econometrics, Operational Research, Data Science and Actuarial Studies*, Contributions to Economics,
https://doi.org/10.1007/978-3-030-85254-2_18

(World Bank 2020). For the European Union, the rate of decline is even higher − 7.6% by October 2020 (International Monetary Fund 2020).

The education sector has not remained unaffected by the crisis and has undergone a number of changes to ensure the learning process in schools and universities. Despite efforts at the state, institutional and personal levels, according to UNICEF 1.6 billion scholars and students from 188 countries were affected by the crisis in April 2020 due to the closure of schools and universities (UNICEF 2020). A decline of up to 1 billion is observed for the new academic year. That is why the processes of digitalization in educational institutions are being forced. Among the tools for providing education offered by some authors are the introduction of mobile learning (Todoranova and Penchev 2020), application of artificial intelligence technologies (Petrova and Sulova 2020), gamification software (Stoyanova 2015), social networks (Malkawi et al. 2021). Some researchers also report a number of barriers to digital learning, such as the high cost of software solutions, including their acquisition and maintenance (Kuyumdzhiev 2020). This requires development of regional and national policies related to the improvement of technical, social and economic infrastructure to build conditions for the rise of educational institutions (Czaplewski and Klóska 2020).

According to the optimistic forecasts of the IMF, from the beginning of 2021 the average global growth is expected to be 5.2%, and for the EU 5% (International Monetary Fund 2020). Unfortunately, skeptical opinions are expressed regarding the recovery from the crisis. The reason is that the pandemic had a negative impact on the accumulation of human capital, and hence on active job seeking (World Bank Group 2021). Here it should be borne in mind that the quality of the workforce is an important factor in increasing the competitiveness of companies (Antonova and Ivanova 2018). This increases the requirements of the labor market, respectively to the educational institutions. The World Bank is developing scenarios for global economic growth and possible solutions to the crisis, which depend on the pace of pandemic control (World Bank Group 2021). One of the main emphases placed in them is investing funds in the development of human capital through the implementation of adequate policies in the field of education.

In this regard, **the aim of the article** is to propose a research framework for evaluation of emotional attitudes in social media. The purpose is achieved through the following objectives:

- Studying the higher education issues in a pandemic context;
- Investigating the approaches to social media mining which are applicable to emotions mining.

The author tested the practical applicability of the proposed framework by retrieving data from the social network Twitter and applying data mining techniques for analyzing large volumes of textual content.

2 Higher Education Challenges During Pandemic

In the conditions of intensified competition between universities, both nationally and internationally, one of the challenges facing them is competitive differentiation. Many of them fail to implement adequate marketing strategies to stakeholders. Successful branding must help to achieve a clear differentiation of directly competing universities through unique attributes and characteristics that are relevant to the target groups (Zhechev 2018). The rules of branding in higher education have changed with the onset of the Covid-19 crisis. It has seriously affected the attraction of new students. According to a study by the American Marketing Association from October 2020, 72% of university rectors (respectively presidents) are concerned about the decline in the value of higher education in the pandemic compared to March (48%) and April (60%) of the same year (American Marketing Association 2020). Unfortunately, there are also negative opinions among students—56% said that they can no longer afford to continue their education due to financial problems during the crisis. 36% of the parents interviewed said that they have redirected the funds for financing their children's education to cover expenses or financial losses incurred as a result of the Covid-19 crisis.

The problems that young people face during lockdown periods are far from limited to their deteriorating financial situation. A study by the University of Copenhagen shows that stress levels among people under the age of 30 increased during the pandemic period (Rohde 2021). It has been found that in participants without previous mental problems, there are those that deepen with the extension of the lockdown period. Social distance raises some physical and mental problems that lead to a decrease or lack of motivation to perform duties in a learning environment.

In an interview with the World Economic Forum, Prof. Suzanne Fortier, the Principal and Vice-Chancellor of McGill University in Montreal, Canada, a Chair of the Global University Leaders Forum too, outlined some problems in higher education that arose during periods of social distance (Fleming 2021). These include: lack of lecturer–student communication; difficulties in acquiring new knowledge and practical skills; difficulties in conducting classes, meetings, conversations due to various technical and financial reasons; problems with the social inclusion of people with special needs or people from other cultural communities. On the other hand, students in the above courses, those who come to the universities for upskilling and reskilling, typically people who are already in the workforce, have found many advantages in the flexibility of e-learning.

In addition, the challenges outlined by Bhagat and Kim can be highlighted (Bhagat and Kim 2020): lowering the quality of education as a result of distance learning; digital sustainability of universities; preparatory to and flexibility of the academic staff in providing educational services in an electronic environment; creating adequate policies for tuition, conducting exams and acquiring educational degrees that correspond correctly to the "new" reality.

Stanimirov's analysis of the educational environment states that "recognizing the challenges is the basis for generating ideas for 'closing the gaps' and synchronizing the education and labour markets, which have a cycle of 3 to 5 years" (Stanimirov 2020). As the demands of society increase, so do the demands on universities, especially in conditions of social distance, when they have to show their digital resilience and adaptability.

Given the above, as well as in view of some statistics[1] for the increased use of the Internet during the periods of social distance in 2020, it can be outlined the following main tasks of this article:

- To suggest an approach to research the emotional attitudes of Internet users, in particular in social networks;
- To study the social networks users' opinions, mining emotional attitudes about periods of social distance.

In order to accomplish the tasks, it is necessary to study approaches to social media emotions mining.

3 Social Media Mining Approaches and Technologies

Undoubtedly, social media is a big data source. One of the most popular technologies for emotions mining in social media is based on extracting knowledge from data. Data extracted from social media are unstructured and often difficult to process due to their diverse nature. The extracted content can be textual or multimedia, which implies the application of various techniques for its processing.

Sulova and Bankov suggest a four-stage approach to social media mining: Data retrieval; Text processing; Data mining and Results interpretation (Sulova and Bankov 2019). They emphasize that their proposed approach can be adapted to the specifics of using a particular social media.

Manguri et. al. apply an approach to extracting feelings from social networks, which consists of the following stages: Text data mining; Sentiment identification; Feature selection; Sentiment classification; Sentiment polarity & subjectivity (Manguri et al. 2020). As a basis for testing the approach, the authors use textual content from the social network Twitter.

Nasralah et. al. follow a comprehensive approach to retrieving textual content from social networks, which is carried out at the following main stages (Nasralah et al. 2020): Discovery and topic detection, Data collection, Data preparation and quality evaluation and Analysis and results (Fig. 1).

Another popular technology for working with large amounts of unstructured data, that we find in the literature, is artificial neural network (ANN). According to Bakaev

[1] According to Eurostat, on average in the European Union, 26% of users have used the Internet to conduct e-learning activities, 56% for social networking activities, 68% for messaging, 74% for working with e-mail. The highest use of social networks was reported in Denmark 85%, and the lowest in Germany 54% (Eurostat 2021).

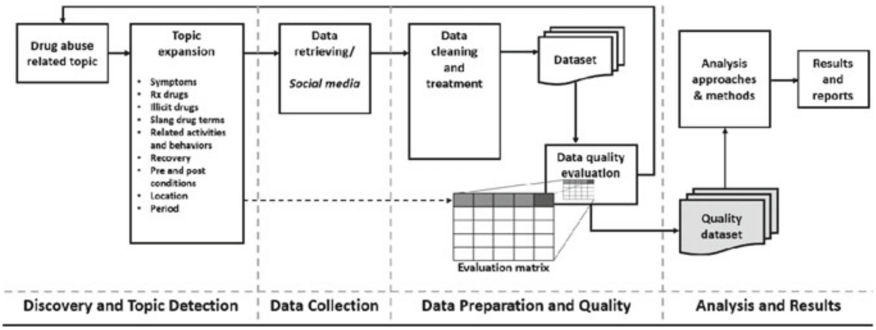

Fig. 1 Social media text mining framework by Nasralah et al. (2020)

et. al., it serves to construct a model of consumer behavior and predict the complexity of the Internet resource with which the user works (Bakaev et al. 2018).

Krebs et. al. follow a mixed approach to emotions mining in social networks based both on neural network technology and text mining (Krebs et al. 2018). Their approach, called "Pipeline for final prediction of reaction distributions", consists of several stages that take place simultaneously (Fig. 2).

Calefato et. al. propose a framework architecture of emotions mining from textual content. It consists of two main modules: Emotion Classification Module and Polarity Classification Module (Calefato et al. 2019). The authors define it as a specific solution for sentiment analysis, specifically for polarity and emotions mining from a text.

Yassine and Hajj suggest a framework architecture for emotions mining from textual content on social networks (Yassine and Hajj 2010). It consists of seven

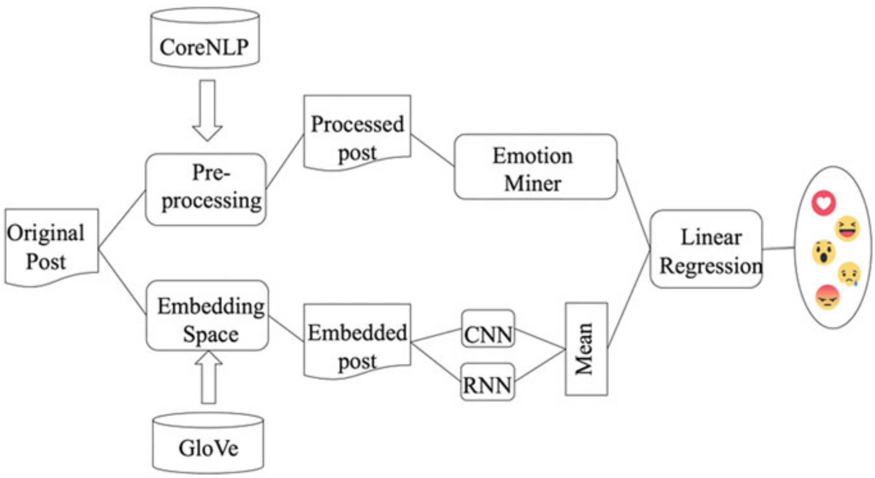

Fig. 2 Pipeline for final prediction of reaction distributions by Krebs et al. (2018)

steps: Raw data collection, Lexicon's development, Feature generation, Data pre-processing, Creating a training model for text subjectivity, Text subjectivity classification, Friendship Classification.

Luo and Yi's approach to emotions mining from online comments consist of: Dataset and Pre-processing, Setting Parameters and Determining the Number of Topics, Comparison of Sampling Time with Existing Models, Comparison of Sampling Time with Existing Models, Understandability of Results (Luo and Yi 2019).

On the basis of the cited publications, it can be concluded that the variety of approaches and technologies for emotions mining is huge. Scientists mainly focus on text mining and neural networks. That is why there are some shared stages in the approaches considered: Retrieving social media content; Content pre-processing; Classification of the content according to the identified emotions; Interpretation of results.

4 Emotions Mining Research Framework

In view of the purpose of this study and based on what has been stated so far, the author proposes a research approach to assessing emotional attitudes in social media (Fig. 3). It consists of three swim lanes: Stages, Toolkit and Artifacts (considered from bottom to top).

In this article, the author proposes an evaluation process of emotional attitudes in social media which is conducted in five stages: Topics Discovery, Social Media Connecting, Data Retrieving and Pre-processing, Modeling and Classification, Analysis and Evaluation.

For the implementation of the phases, tools for extracting, processing and evaluating content from social media are used. The output of each of the phases, resulted from a certain toolkit, are one or more artifacts: Research Plan, Data Source, Datasets, Models and Reports.

It should be noted that the proposed framework adapts to the specifics of the social media which content is being examined. For example, each social media application

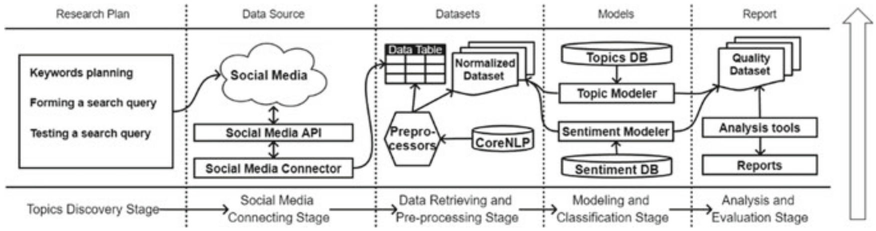

Fig. 3 Proposed Emotions Mining Research Framework

programming interface (API) has strictly individual characteristics and requires the use of different connectors for different social media mining platforms.

On Topics Discovery Stage tools for planning keywords and queries to social media are applied. This type of software evaluates the relevance of keywords and offers a set of words to form an effective query.

Based on the selected keywords set, a request is made to social media. It is necessary to use specialized data mining software. It connects to the social media API via a connector, after which the raw set of unstructured data is generated. The data mining software supports a set of pre-processors that use natural language processing methods (CoreNLP). Most often they are: Transformation, Tokenization, N-grams, Filtering, Normalization. They are applied in the order in which they are listed, and their settings depend on the objectives of the study.

As a result, a set of data is generated, which is used as a basis for the implementation of the next phase of modeling and classification. The topics are identified and the emotional attitudes are analyzed according to the classification databases. These databases contain a pre-prepared set of words related to a specific research field.

The last phase is related to the analysis of the generated models from the previous stage. Reports with the results of the study are created. Typically, they contain statistical analysis data and conclusions about the overall emotional attitudes in the specified research questions.

5 Results

It was used a software set to test the proposed framework. These are: Gephi, MeaningCloud and Orange. Gephi is an open-source and free visualization and exploration software for all kinds of graphs and networks. MeaningCloud is an Excel Add-in for text analytics (MeaningCloud 2021). Orange is an open-source data visualization, machine learning and data mining toolkit (Orange 2021).

The social network Twitter was used to test the proposed research framework. In the period 24.01.–26.01.2021 7722 tweets in English were retrieved using several keywords: university e-learning 2020; university COVID-19; financial crisis education 2020; education challenges; education crisis 2020; lockdown; higher education crisis; COVID-19 crisis 2020; education 2020; higher education COVID-19. The experiment is based on extracted text content.

Gephi was used to connect to Twitter and retrieve data. It is a powerful tool based on algorithms for visualizing and simulating graphs. It is characterized by speed. It provides a statistics and metrics framework (Gephi 2021). Downloaded tweets are exported as a csv file.

The next step is to process the raw data via Orange. The analysis of emotional attitudes is performed using the Tweet Profilier module (Fig. 4), which supports several methods of content classification. These are classes based on the classifications of Plutchnik, Ekman and Profile of Mood States (POMS). Each of them identifies a different number of basic emotions.

Fig. 4 Configuration of emotion analysis in Orange

According to Plutchik, these are: anticipation, acceptance, joy, surprise, anger, disgust, fear, sadness (Plutchik 1980). Ekman distinguishes joy, surprise, anger, disgust, fear, sadness (Ekman 1982). POMS classifies the emotions of tension, anger, vigor, fatigue, depression, confusion (Renger 1993). The common emotion for the three classifications is anger, which is considered a primary negative emotion, thanks to which individuals defend and survive, both physically and verbally.

The results generated by Orange after applying the Plutchnik classifier are shown in Figs. 5. It is noticed that the positive emotions joy and trust are the highest percentage—73.56% of all extracted tweets.

The results according to the Ekman classifier (Fig. 6) are similar to the previous one. Positive emotion joy also gives precedence—50.18% of all tweets.

In Fig. 7 it is noticed that the diagram changes when applying POMS. The reason is that this classification is oriented entirely to negative emotions.

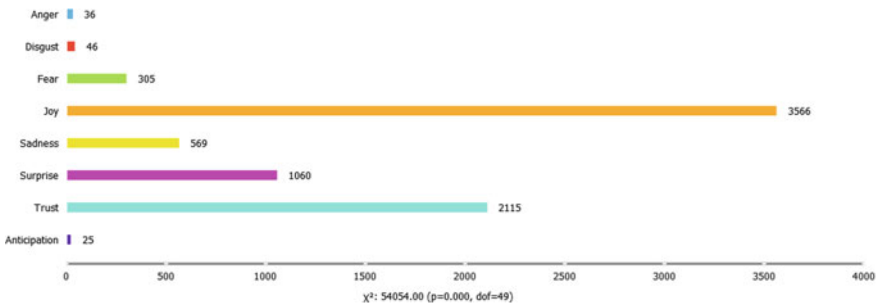

Fig. 5 Results of the emotions analysis in Orange according to the classifier of Plutchnik

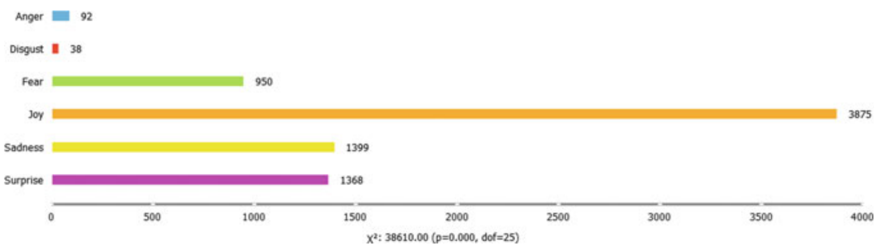

Fig. 6 Results of the emotions analysis in Orange according to the classifier of Ekman

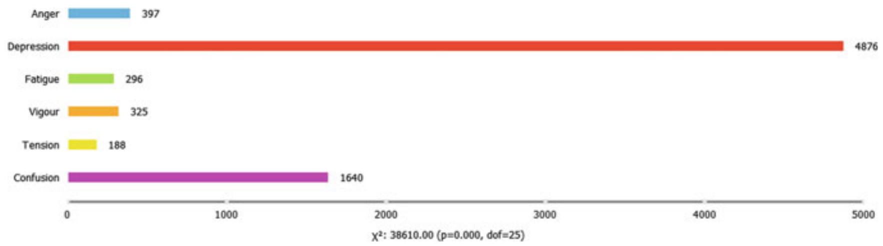

Fig. 7 Results of the emotions analysis in Orange according to the classifier of POMS

The differences in the results of the application of the different classifiers also arise from the pre-defined set of words that is applied for sentiment analysis of Orange.

For comparison, we apply MeaningCloud—an Excel plugin, through which we perform sentiment analysis. We apply the built-in basic model in the software based on WordNet. The results are summarized in Table 1. Similar to Plutchik's and Ekman's Orange classifiers, MeaningCloud recognizes that 55% of the content analyzed is positive polarity. 14% is marked with neutral polarity and 31% with negative polarity.

Therefore, it can be concluded that the emotional attitudes identified in the extracted tweets are predominantly positive. There are observed nuances of negative emotions, including anger, fear, disgust and depression**.**

According to the research framework proposed in this article, the next step is identifying the topics that excite the Twitter users. For this purpose, we use the Orange modules for Pre-processing and Topic Modeling (Fig. 8).

The applied pre-processors are:

Table 1 Relative share of polarity types

Polarity	Tweets Count	Relative share (%)
P +	1168	15
P	3077	40
NEU	1043	14
N	1469	19
N +	965	12

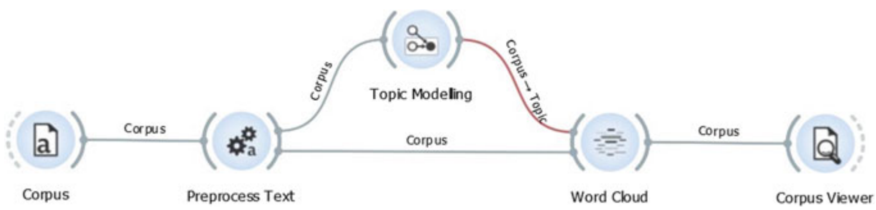

Fig. 8 Orange theme modeling configuration

Fig. 9 Twitter Topics Cloud

- Transformation (the settings are: remove URLs and accents, lowercases, parse HTML);
- Tokenization (it is applied: punctuation, white space);
- N-grams range 1–3;
- Filtering (the settings are: stopwords in English and regular expression);
- Normalization through WordNet Lemmatizer.

When modeling the topics, we use an algorithm Latent Semantic Indexing, that returns both negative and positive words and topic weights (Orange 2021). The result is shown in Fig. 9.

In conclusion of our study, we can find that users of the social network Twitter express a variety of emotional attitudes through their posts. However, positive emotions prevail, both according to two of Orange's classifiers and according to MeaningCloud's sentiment analysis. One of the possible reasons is that the publications are on topics that mainly concern the family, social contacts, politics, education, leisure activities (sports, reading, music), professional development.

6 Conclusion

Periods of social distance have an impact on people around the world in various aspects: health, professional, social, economic and financial. The first months of 2020 were the worst, as governments' decisions to close states came as a surprise, and businesses and educational institutions found it difficult to reorganize their activities, mainly in a digital environment. Over time, the "new" reality was accepted by society and people adapted their daily lives to it.

These preconditions have led the author of this article to conduct a study of the emotional attitudes of social media users, in particular the social network Twitter, based on appropriate keywords. As a result, a framework has been proposed that can be adapted depending on the needs of the research and the specific features of social media.

Acknowledgements The publication is made within project No. 8.2.2.0/18/A/021 "Perfection of the Academic Staff of Liepaja University in the Areas of Strategic Specialization—Natural Sciences, Mathematics and Information Technologies, Art, Social Sciences, Commerce and Law".

References

American Marketing Association (2020) 2020 Symposium for the marketing of higher education, November 16–19, https://www.ama.org/wp-content/uploads/2020/11/2020-Higher-Ed_Resource-Guide_v4.pdf. Accessed 27 Jan 2021

Antonova K, Ivanova P (2018) Staff leasing. Oxford Economic Papers, Oxford University Press, vol 70, 4, issue 2, pp 1416–1425

Bakaev M, Laricheva T, Heil S, Gaedke M (2018) Analysis and prediction of university websites perceptions by different user groups. In: Proceedings of XIV international scientific-technical conference on actual problems of electronics instrument engineering (APEIE), IEEE, pp 381–385

Bhagat S, Kim D (2020) Higher education amidst COVID-19: challenges and silver lining. Inf Syst Manag 37(4):366–371

Calefato F, Lanubile F, Novielli N, Quaranta L (2019) EMTk—the emotion mining toolkit. In: IEEE/ACM 4th international workshop on emotion awareness in software engineering (SEmotion). Montreal, QC, Canada, pp 34–37

Czaplewski M, Klóska R (2020) Regional policy as a factor in shaping regional development in Poland. South East European J Econ Bus 15(1):93–104

Ekman P (1982) Emotion in the human face, 2nd edn. Cambridge University Press, New York

Eurostat (2021) What did we use the internet for in 2020? https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210126-2. Accessed 27 Jan 2021

Fleming S (2021) This is how university students can emerge from the pandemic stronger, https://www.weforum.org/agenda/2021/01/online-learning-universities-covid-suzanne-fortier/. Accessed 27 Jan 2021

Gephi (2021) Gephi features, https://gephi.org/features/. Accessed 27 Jan 2021

International Monetary Fund (IMF) (2020) Real GDP growth: annual percent change, https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/EU/ADVEC/AS5/DA/WEOWORLD. Accessed 27 Jan 2021

Krebs F, Lubascher B, Moers T, Schaap P, Spanakis G (2018) Social emotion mining techniques for facebook posts reaction prediction. In: Proceedings of the 10th international conference on agents and artificial intelligence, vol 1, pp 211–220

Kuyumdzhiev I (2020) A model for timely delivery of IT solutions for Bulgarian universities. In: International multidisciplinary scientific geoconference: SGEM 20(2.1.), pp 3–10

Luo X, Yi Y (2019) Topic-specific emotion mining model for online comments. Future Internet 11(e79):18

Malkawi R, Alsmadi I, Ahmed A, Petrov P (2021) A firewall-adversarial testing approach for software defined networks. J Theoret Appl Inform Technol, Little Lion Scientific 99(1):227–241

Manguri K, Ramadhan R, Mohammed Amin P (2020) Twitter sentiment analysis on worldwide COVID-19 outbreaks. Kurdistan J Appl Res 5(3):54–65

MeaningCloud (2021) MeaningCloud official website, https://www.meaningcloud.com/. Accessed 27 Jan 2021

Nasralah T, El-Gayar O, Wang Y (2020) Social media text mining framework for drug abuse: development and validation study with an opioid crisis case analysis. J Med Internet Res 22(8):e18350

Orange (2021) Orange official website, https://orangedatamining.com/. Accessed 27 Jan 2021

Petrova R, Sulova S (2020) AI Governor for the quality and the strength of bridges. CompSysTech 20. In: Proceedings of the 21st international conference on computer systems and technologies 20. Association for Computing Machinery, pp 78–85

Plutchik R (1980) A general psychoevolutionary theory of emotion. In: Plutchik R, Kellerman H (eds) Emotion: theory, research, and experience, vol 1, Academic Press Cambridge, pp 3–31

Renger R (1993) A review of the profile of mood states (POMS) in the prediction of athletic success. J Appl Sport Psychol 5(1):78–84

Rohde A (2021) Younger people are being hit hardest by pandemic loneliness, https://www.weforum.org/agenda/2021/01/pandemic-loneliness-covid19-mental-health. Accessed 27 Jan 2021

Stanimirov E (2020) Quo Vadis, Education? In: Proceedings of Jubilee international scientific conference dedicated to the 100th anniversary of the university of economics—varna economic science, education and the real economy: development and interactions in the digital age, vol 1. Publishing house Science and Economics, pp 27–49 (in Bulgarian)

Stoyanova M (2015) Architecture of gamification system. Sci j Econ Comp Sci 2:18–33 (in Bulgarian)

Sulova S, Bankov B (2019) Approach for social media content-based analysis for vacation resorts. J Commun Softw Syst (JCOMSS) 15(3):262–271

Todoranova L, Penchev B (2020) A conceptual framework for mobile learning development in higher education. In: Proceedings of the 21st international conference on computer systems and technologies '20 (CompSysTech '20). Association for Computing Machinery, pp 251–257

UNICEF (2020) COVID-19 and children: UNICEF data hub, https://data.unicef.org/covid-19-and-children/. Accessed 27 Jan 2021

World Bank Group (2021) Global economic prospects. World Bank, Washington

World Bank (2020) COVID-19 to plunge global economy into worst recession since world war II, https://www.worldbank.org/en/news/press-release/2020/06/08/covid-19-to-plunge-global-economy-into-worst-recession-since-world-war-ii. Accessed 27 Jan 2021

Yassine M, Hajj H (2010) A framework for emotion mining from text in online social networks. In: IEEE international conference on data mining workshops. Sydney, pp 1136–1142

Zhechev V (2018) Measuring brand competitive performance—a focus on ethical brand positioning. In: Conference proceedings of 12th international days of statistics and economics. Prague, September 6–8, pp 2034–2043