Identifying users communicative segments to explain digital fatigue via NLP

Roman Vit[1,†], Olexander Mazurets[1,*,†], Olexander Barmak[1,†], Maryna Molchanova[1,†] and Iurii Krak[2,3,†]

[1] *Khmelnytskyi National University, 11, Instytuts'ka str., Khmelnytskyi, 29016, Ukraine*
[2] *Taras Shevchenko National University of Kyiv, Kyiv, 64/13, Volodymyrska str., 01601, Ukraine*
[3] *Glushkov Cybernetics Institute, Kyiv, 40, Glushkov ave., 03187, Ukraine*

Abstract

The study is devoted to the interpretive detection of manifestations of digital fatigue in social media through the analysis of communicative segments. A method is proposed that shifts the focus from the classification of single messages to the detection of natural thematic groupings of texts, within which cognitive and emotional patterns of interaction are formed. The pipeline combines contextual vectorization SentenceTransformer, dimensionality reduction UMAP, density clustering HDBSCAN, lexical generalization TF-IDF and KeyBERT, and interpretive labeling of segments using Flan-T5. The experiment was performed on the Healthcare Workers Burnout Tweets corpus (1,879 messages), where the target variable reflects the presence or absence of signs of exhaustion. An explanatory map of communication segments with names and key concepts was obtained, and the proportion of messages with the label "burnout" was calculated for each segment. The results showed an uneven thematic structure of the corpus with the dominance of one large segment and higher fatigue indices in smaller, thematically narrow clusters, which confirms the contextual nature of the phenomenon. The practical suitability of the approach for building psycho-emotional stress monitoring systems in professional, educational and social environments is shown. The proposed method creates a basis for extension to multilingual corpora, temporal analysis of topic dynamics and the use of explanatory deep models to establish causal relationships between the semantics of communication and manifestations of digital fatigue.

Keywords

digital fatigue, explainable AI, communicative segments, NLP, topic modeling, cognitive overload

1. Introduction

In today's society, digital interaction has become an integral part of people's professional, educational and personal lives. According to the DataReportal Digital 2025 July Global Statshot Report [1], as of July 2025, there were approximately 5.65 billion Internet users worldwide, representing 68.7% of the world's population, and more than 5.4 billion social media accounts, representing 65.7% of the population. The average user spends approximately 6 hours and 37 minutes online each day, with more than 2 hours and 26 minutes spent on social media [2].

The Microsoft Work Trend Index 2023 [3] confirms the scale of the problem: more than 68% of workers worldwide report having less time to focus on their work due to excessive messaging and online meetings. This indicates the systemic impact of information overload on cognitive resilience, professional performance and well-being.

This level of engagement creates a constant cognitive pressure that can change the intensity of attention, emotional background and information processing strategies during digital interaction [4]. The study suggests that the state of digital fatigue is not universal, but arises within certain communicative contexts, where the number of repetitive, emotionally saturated or cognitively

* *AdvAIT-2025: 2nd International Workshop on Advanced Applied Information Technologies: AI & DSS, December 05, 2025, Khmelnytskyi, Ukraine, Zilina, Slovakia*

[1*] Corresponding author.
vit.roman.vit@gmail.com (R. Vit); exe.chong@gmail.com (O. Mazurets); alexander.barmak@gmail.com (O. Barmak); m.o.molchanova@gmail.com (M. Molchanova); yuri.krak@gmail.com (I. Krak)
0009-0009-6958-4730 (R. Vit); 0000-0002-8900-0650 (O. Mazurets); 0000-0003-0739-9678 (O. Barmak); 0000-0001-9810-936X (M. Molchanova); 0000-0002-8043-0785 (I. Krak)

overloaded topics increases. This means that different segments of digital communication can affect the mental state of the user differently: some contribute to exhaustion, others perform a compensatory or restorative function [5]. Therefore, an important task is to identify thematic and semantic characteristics of those communication segments that are associated with an increased risk of digital fatigue, as well as to explain the mechanisms of this influence using natural language processing tools.

The issue of digital fatigue is interdisciplinary and directly related to the implementation of the UN Sustainable Development Goals, in particular Goal 3 «Ensure healthy lives and well-being for all at all ages», Goal 4 «Quality education» and Goal 8 «Decent work and economic growth» [6].

The aim of the research is to develop a method for identifying communicative segments of users as potential sources that influence the manifestations of digital fatigue.

The main contributions of the article are: a conceptual model of digital fatigue is proposed, in which the communicative segment of the user is considered as a potential source of cognitive load that affects the mental state in the process of digital interaction; a method for identifying communicative segments of users is created and experimentally tested.

2. Related works

Based on a systematic analysis of publications for 2010–2025, presented in [7], the key factors, consequences and strategies for overcoming the phenomenon of digital fatigue are summarized. It is proven that excessive interaction with digital tools leads to cognitive overload, emotional exhaustion, decreased professional productivity and increased stress levels. The identified contradictions between individual research results indicate the need for a contextualized approach to organizing digital communication and developing balanced models of combining synchronous and asynchronous interaction formats. In parallel, the authors [8] note that the growing popularity of distance education formats, especially after the COVID-19 pandemic, contributed to increasing the flexibility and accessibility of learning, but at the same time caused excessive use of digital technologies. This has become one of the main factors in the emergence of digital fatigue, which manifests itself in a feeling of mental exhaustion, decreased motivation, sleep disturbance, lack of concentration and a decrease in academic performance. In this context, digital fatigue is considered as a multidimensional cognitive-emotional phenomenon that is formed under the influence of technological overload, long screen time, increased self-consciousness during video meetings, physical inactivity and lack of rest.

The study [9] analyzes the relationship between digital fatigue and marital satisfaction among adults in Turkey, taking into account gender differences. Based on Family Systems Theory and the structural equation modeling method, data from 384 respondents were processed. It was found that digital fatigue significantly reduces the level of relationship satisfaction, with the greatest impact being psychological fatigue, followed by digital addiction, physical and mental fatigue and psychosomatic manifestations. The effect was more pronounced among women. The results emphasize the systemic impact of digital fatigue on family well-being and highlight the need to strengthen the psychological resilience of families in the digital age.

The study [10] is devoted to studying the level of social media fatigue among students in the context of increasing digital activity. The aim of the work was to analyze the manifestations of social media fatigue taking into account demographic characteristics and features of technology use. The survey was conducted among 386 students studying in different years of university entry, using the social media fatigue scale and the JASP 0.19.3 software environment for descriptive data analysis. The results showed that the overall level of social media fatigue is high. The highest indicators were recorded for the cognitive component, which indicates a significant information overload of users. Along with this, a tendency was found to increase the emotional component of social media fatigue, which reflects the negative impact of social networks on the emotional well-being of students. Additionally, it has been found that the duration of social media use exacerbates the negative effects of social media fatigue, which emphasizes the importance of forming balanced digital practices in the student environment.

The impact of the COVID-19 pandemic on higher education and the associated manifestations of student burnout were examined in [11]. A qualitative narrative review of 38 peer-reviewed publications that met the requirements of SANRA examined the main factors of emotional and academic exhaustion of students, including financial instability, mental health problems, social isolation and fatigue from online learning. It was found that such conditions led to a decrease in engagement, motivation and learning performance. The analysis showed that flexible academic policies, hybrid learning models and psychological support contributed to reducing the negative consequences of burnout. At the same time, the effectiveness of artificial intelligence tools, in particular chatbots and digital academic assistants, which provided scalable emotional and academic support in a distance learning environment, was noted.

The study [12] focused on detecting signs of professional fatigue in physicians based on linguistic characteristics of clinical notes. Using data from 129,228 emergency department visits, the authors trained a model capable of recognizing notes created by physicians with high workloads who had worked at least five of the previous seven days. The model successfully identified such notes in the test sample, and also detected signs of fatigue in other conditions during night shifts and work with a large flow of patients. It was found that when the model recorded increased signs of fatigue, the quality of clinical decisions decreased, in particular, the efficiency of diagnosing myocardial infarction was 19% lower for each standard deviation increase in the estimated level of fatigue. A key linguistic feature of notes created by exhausted physicians is the predictability of the next word from the context, reflecting a decrease in cognitive flexibility. Interestingly, the model recorded a 74% higher level of predicted fatigue in texts generated by large language models than in real clinical records, indicating potential biases in generative medical texts that require further investigation.

At the same time, a systematic review cited in [13] shows that the introduction of artificial intelligence into work with electronic medical records can reduce the workload and risk of burnout of healthcare professionals. An analysis of studies (2019-2025) found positive effects of using AI scribes, clinical decision support systems, language models and NLP tools, in particular, reducing documentation time and increasing the efficiency of workflows. Methodological limitations (small samples, short follow-up periods) are noted, which require further large-scale and controlled studies to assess the long-term effectiveness and safe integration of AI into clinical practice.

The work [14] considers the use of deep learning for automatic detection of professional burnout and digital fatigue based on different types of data. The application of NLP, CNN, RNN and Transformer-models to analyze texts, voice recordings, facial reactions, activity logs and physiological indicators is considered. An end-to-end methodology for developing, training and deploying such systems is proposed, which allows to detect patterns of stress and emotional exhaustion inherent in digital fatigue and burnout. The authors emphasize the need to take into account the issues of confidentiality, ethics and bias of the models, emphasizing that the integration of psychological concepts with machine learning contributes to the transition from reactive to preventive management of the mental health of employees, increasing efficiency and well-being in the working environment. The analysis of the reviewed studies confirms that digital fatigue is a context-dependent phenomenon, and not a universal reaction to prolonged use of technologies. Its manifestations are determined by the nature of communication, the type of digital environment and the semantic load of information. In professional, educational and social contexts, different intensities of cognitive and emotional exhaustion are observed, which indicates the importance of the content and structure of communicative segments as potential sources of influence on the user's mental state. The reviewed studies have confirmed that some types of digital interaction, such as routine documentation, intensive social communications or online learning, increase the risk of fatigue, while others, on the contrary, have a compensatory effect. At the same time, a number of unsolved scientific problems remain. The mechanisms of fatigue formation in different types of digital communications require clarification, as well as long-term observations capable of tracing the dynamics of the transition from neutral to exhausting communication patterns. The impact of algorithmic content and generative models on the strengthening or modification of signs of digital fatigue has also been insufficiently studied. This determines the relevance of further development of interpretive NLP methods that would allow explaining the relationships between the subject of communication, cognitive load and manifestations of user fatigue.

This study fills the gap between descriptive studies of digital fatigue and interpretive analysis of its sources by forming the basis for building explainable, adaptive systems for monitoring the mental state of users in digital environments.

3. Method design

The conceptual model of digital fatigue is based on the assumption that the user's psycho-emotional state in the digital environment is formed under the influence of stable semantic structures of communication. Digital content is organized into communicative segments that reflect dominant topics, emotional markers, and information intensity. These segments act as sources of cognitive load, which leads to attention depletion and emotional exhaustion, forming manifestations of digital fatigue. The model assumes a causal chain: digital content, communicative segment, cognitive load, emotional exhaustion, digital fatigue. At the same time, a feedback effect is possible, when the user's current mental state influences further communicative behavior and the semantic structure of messages.

The method for identifying user's communicative segments is designed to detect and interpret thematic segments of user communication that potentially affect the manifestations of digital fatigue and cognitive exhaustion. The method provides not only automatic grouping of texts by semantic proximity, but also an interpreted explanation of each segment, which allows assessing the connections between the nature of digital interaction and the emotional state of users. Workflow of communicative segments detection and interpretation pipeline is shown in Figure 1.

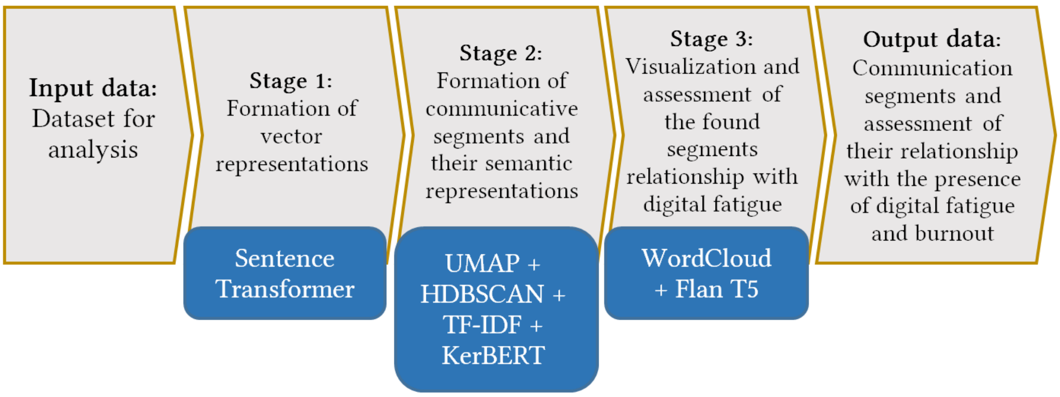

Figure 1: Workflow of communicative segments detection and interpretation pipeline.

The presented scheme reflects the method for identifying communicative segments of user texts, developed to identify the content contexts of digital interaction that potentially affect the manifestations of digital fatigue and burnout. The method is aimed at moving from superficial statistical analysis of texts to the construction of an interpreted semantic structure, in which each communicative segment is considered as a separate thematic unit [15], reflecting specific cognitive and emotional patterns of user communication.

At Stage 1, vector representations of texts are formed. This stage is responsible for transforming the unstructured text corpus into a multidimensional feature space, where each message receives its own contextual representation. For this, the Sentence Transformer model [16] is used, which provides a reflection of semantic shades of meanings and the preservation of latent connections between texts. Thus, the basis is created for further grouping of communication patterns not by formal, but by content characteristics.

Stage 2 involves the formation of communicative segments, i.e. semantically related groups of messages that form separate contexts of digital interaction [17]. At this stage, the dimensionality of embeddings is reduced using the Uniform Manifold Approximation and Projection (UMAP) method [18], which ensures the preservation of data topology, as well as clustering in the resulting space using Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) [19], which allows automatically determining the natural number of thematic groups. For each formed segment, the most characteristic lexical features are calculated using TF-IDF and KeyBERT, which allows describing its content dominants. This stage forms the basis for the semantic division of the corpus into communicative segments, within which common intentions, emotional coloring and thematic directions of user messages are revealed.

Stage 3 involves the interpretation and visualization of the obtained segments in order to explain their connection with the phenomenon of digital fatigue [20]. For this, the Flan-T5 generative model is used, which creates a generalized set of key concepts for each segment, which provides a clear interpretation of the obtained clusters (removes repetitions, adds words representing segments that are close to those found in stage 2). Visualization in the form of a WordCloud contributes to the formation of an intuitive idea of the structure of the topic and the dominance of individual concepts [21]. At the same stage, the index of association of segments with digital fatigue is calculated based on the specific share of messages marked as containing signs of exhaustion.

The result of the method is a system of communication segments, each of which has its own semantic characteristics, a set of key concepts and a level of connection with indicators of digital fatigue. This allows identifying communicative environments that contribute to an increase in cognitive load or, conversely, perform a compensatory function. The scientific novelty of the approach lies in the combination of interpreted natural language processing tools with cluster analysis to build an explanatory model of the cognitive impact of digital communications.

Its difference from existing approaches lies in the shift in focus from the classification of individual messages to the analysis of communicative segments – natural thematic groupings of texts that reflect the cognitive and emotional patterns of users' communication. This approach allows not only to assess the presence of signs of fatigue, but also to explain within which semantic contexts it occurs.

4. Experiment

4.1. Experimental datasets

The dataset «Healthcare Workers Burnout Tweets» [22] contains 1879 text messages obtained from the social network Twitter, which reflect the emotional and professional reactions of healthcare workers. Its content represents a wide range of statements related to psycho-emotional state, interaction with patients, feelings of exhaustion, support from colleagues and personal experience of working in stressful conditions of the pandemic. Each entry in the table corresponds to a separate user tweet and is accompanied by metadata describing the textual and structural characteristics of the message. In particular, for each tweet, the publication date, unique identifier, source (type of device or application from which the entry was created), number of shares and likes, as well as a number of linguistic parameters are stored, including the number of words and symbols, word length, proportion of service parts of speech, number of user mentions, hashtags and links.

Of particular importance is the variable «burnout», which acts as a target feature for training text analysis models. It indicates the presence or absence of indicators of digital fatigue or professional burnout in the message. In the markup, the value “1” corresponds to messages that show signs of emotional exhaustion, apathy, or stress, while the value “0” indicates neutral or positive messages without such manifestations. Analysis of the statistical distribution shows that the majority of tweets (64.7%, i.e. 1215 records) have markers of digital fatigue or burnout, while the rest (35.3%, or 664 records) do not contain such signs. This imbalance indicates a high representation of messages in which healthcare workers demonstrate symptoms of emotional exhaustion, which makes the corpus relevant for researching psychoemotional risks in digital communication and training models for detecting digital fatigue [23].

4.2. Experiment and setup description

The experimental software is developed in the «Google Colaboratory» cloud environment, which provides interactive execution of program code, automatic saving of results, and the ability to reproduce the experiment on different hardware platforms. The environment is based on the Linux operating system with pre-installed Python libraries and allows calculations to be performed on both the central processor and graphics accelerators. In this study, calculations were performed in CPU mode, which ensured the independence of the results from the hardware parameters.

The experimental system was developed in Python 3.10, which provides support for libraries for natural language analysis, statistical modeling, clustering, and visualization. The main stages of the experiment are implemented in the form of sequentially interconnected modules: data preparation, construction of semantic representations, clustering, identification of key tokens, generative interpretation, and graphical display of results.

Data preprocessing was performed using the pandas, «NumPy», and «re» libraries, which provide cleaning of texts from references, special characters, gaps, and duplications, as well as conversion of text fields into a standardized format for further processing.

To form vector representations of texts, the «SentenceTransformers» library was used, within which the pre-trained all-MiniLM-L6-v2 model was applied [24]. This model forms contextual embeddings that reflect the semantic similarity of texts in a multidimensional space.

To reduce the dimensionality of the vector space in the experimental setup, an implementation of the UMAP method from the umap-learn library in the Python environment was used. The use of this package provided an effective mapping of multidimensional contextual vectors into a space of lower dimension while preserving the topological relationships between the embeddings. The clustering of the obtained coordinates was implemented using the «HDBSCAN» algorithm from the «hdbscan» library, which automatically determines the number of clusters and takes into account variations in the density of the semantic space. Both libraries work in interaction with the «NumPy» and «pandas» structures, which ensures the stability of calculations and the correctness of data transfer between modules.

To determine the key tokens of each cluster, an implementation of the «TF-IDF» method from the «scikit-learn.feature_extraction.text» package [25] was used, which performs a statistical evaluation of the weights of terms within segments. Semantic refinement of key concepts was carried out using the «KeyBERT» library, which is based on the same «SentenceTransformer» architecture and allows comparing tokens by contextual similarity. The obtained descriptors form the basis for further interpretation of the thematic content of the segments.

The generative module is implemented on the basis of the «Flan-T5» model [26], which is part of the «Transformers» framework from «Hugging Face». The model is used in the «sequence-to-sequence» mode for generalized key concepts of segments [27]. To calculate the semantic centrality of selected texts, the «cosine_similarity» module from the «scikit-learn.metrics.pairwise» package was used, which ensures the selection of the most representative examples for further interpretation. The results were visualized using the «Matplotlib» and «WordCloud» libraries, which provide the construction of diagrams, word clouds and graphical dependencies between the size of segments and the proportion of messages containing signs of digital fatigue. All results, including tables and images, were stored in the cloud catalog of the «Colab» environment, which ensures full reproducibility of the experiment and the convenience of further analysis.

5. Results and discussion

Analysis of the results of clustering texts from the corpus «Healthcare Workers Burnout Tweets» made it possible to identify the natural structure of healthcare workers' communications in the digital environment and to trace the dominance of individual thematic contexts related to professional burnout and support from colleagues.

The diagram (Figure 2) displays the results of clustering the tweet corpus, in which a significant part of the messages belong to one large segment. This distribution is explained by the very nature of the original dataset: most healthcare workers' tweets have similar emotional and thematic features related to burnout, support from colleagues, or experiencing crisis situations. The high concentration of texts in Segment 1 indicates the dominance of a common communicative context that unites messages with similar semantic markers.

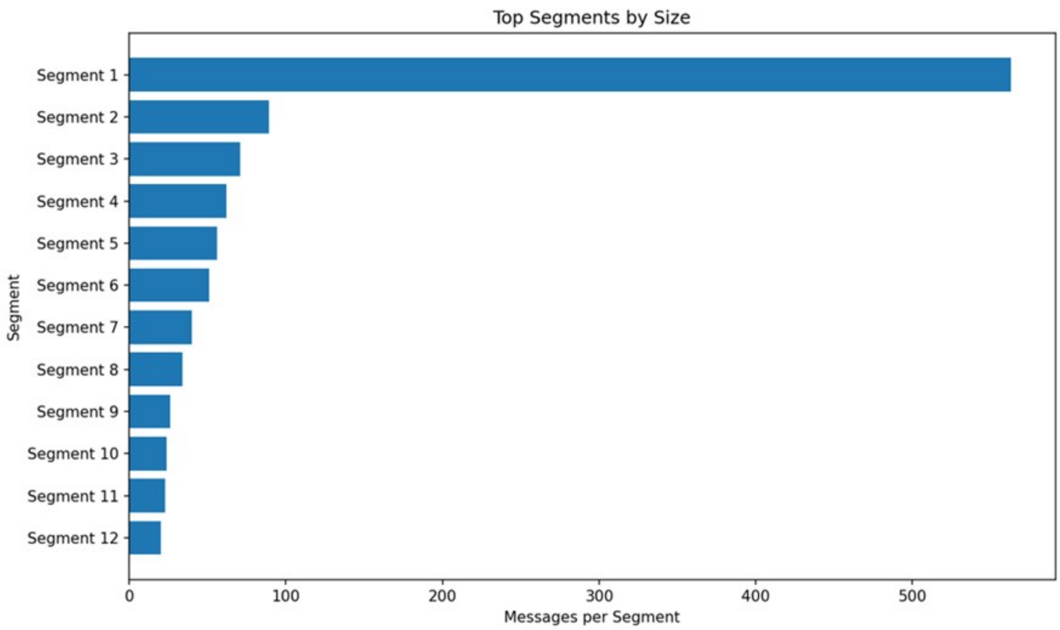

Figure 2: Distribution of tweets by communication segments after clustering.

The obtained result confirms the correctness of the algorithm, which, without having a predetermined number of clusters, isolated the natural structure of the data. The algorithmic model reflected the real thematic heterogeneity of the corpus, where the main cluster represents the central theme of the dataset, and smaller segments record local branches of communication related to individual aspects of professional experience.

The diagram (Figure 3) displays the proportion of messages with manifestations of digital fatigue in each of the clusters that were formed during the semantic segmentation of the studied dataset. The horizontal axis shows the value of the digital fatigue index, which is defined as the ratio of the number of tweets with the «burnout» = 1 label to the total number of messages in the segment.

Distribution analysis shows that most segments are characterized by a high proportion of messages that show signs of emotional exhaustion, apathy, or fatigue. The highest index values are observed in segments with a smaller amount of data, which indicates the localization of the most emotionally saturated statements in thematically narrow communicative contexts. In contrast, large segments demonstrate a moderate level of the indicator, which may be due to greater thematic diversity and the inclusion of neutral or supportive messages.

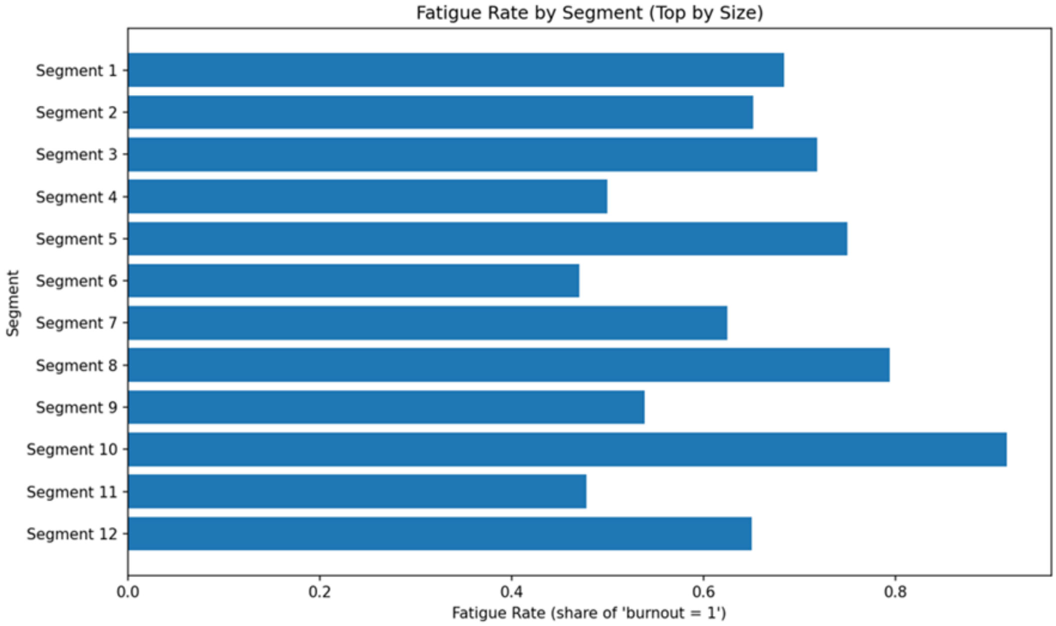

Figure 3: Messages share with digital fatigue manifestations.

The diagram (Figure 4) displays the integral dependence between the size of the segments and the level of digital fatigue in the corpus. The diagram is presented to summarize and improve the visualization of previous results, since it combines two indicators in one space: the size of the segments and the level of digital fatigue. It allows us to clearly trace the relationship between the number of messages in a cluster and the proportion of emotionally exhausted texts, which clarifies the interpretation of the previous diagrams and reinforces the conclusion about the contextual nature of the manifestations of digital fatigue.

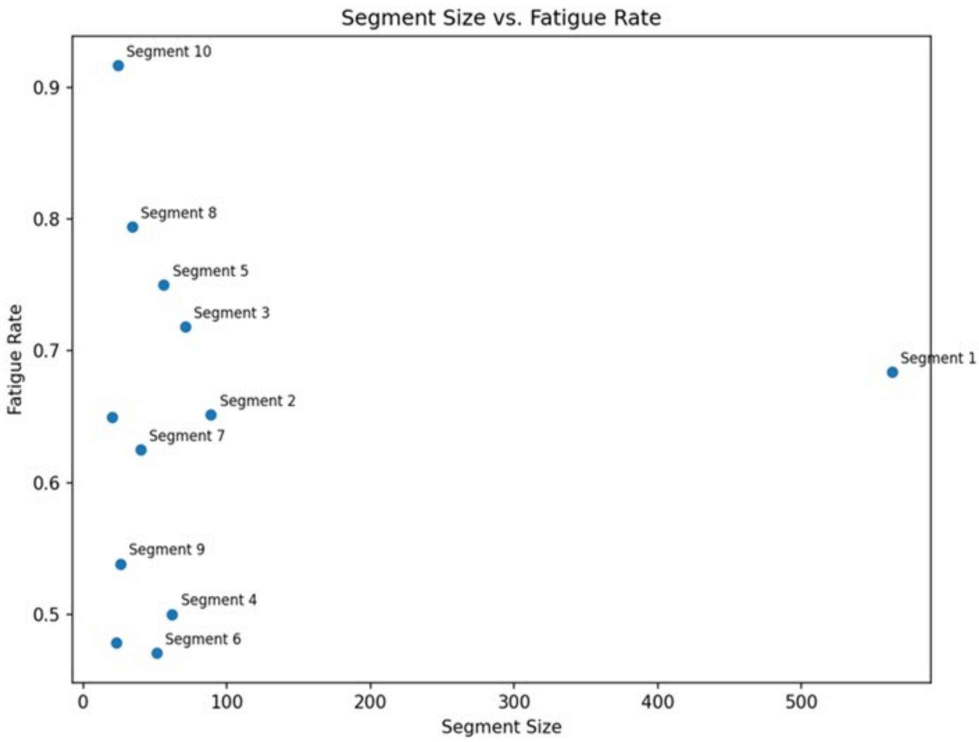

Figure 4: Integral relationship between digital fatigue level and segment size.

Figure 5 shows a cumulative coverage curve that displays the coverage of the message corpus depending on the number of segments considered, ordered by their size.

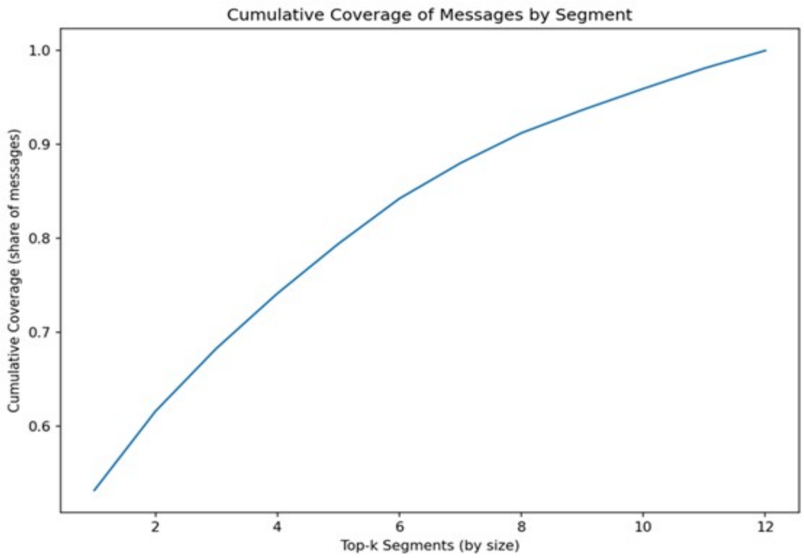

Figure 5: Cumulative coverage curve.

Cumulative coverage curve shows that larger segments cover the majority of all texts, while adding smaller clusters gradually saturates the coverage until the corpus is fully covered. This shape of the curve confirms the uneven distribution of the data, with a few dominant thematic groups containing the bulk of the messages, and the remaining segments representing narrow, less numerous, but meaningfully distinct contexts. Examples of the resulting visual semantic representations are shown in Figure 6.

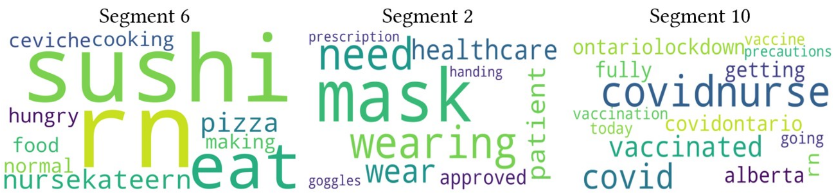

Figure 6: Examples of semantic word representations by segments.

The word clouds reflect characteristic communicative segments, in which different levels of digital fatigue are observed, which is consistent with the results of previous quantitative diagrams. Segment 6, represented mainly by everyday topics, has a low level of fatigue index, since the messages in it are neutral or compensatory in nature, not directly related to professional workload. Segment 2, formed around the lexemes «mask», «healthcare», «nurse», has an average level of digital fatigue, reflecting the working contexts of communication during the pandemic, which combine professional duties and a sense of tension. Segment 10 contains keywords related to vaccination and quarantine restrictions, and is characterized by the highest level of digital fatigue among the presented segments. This suggests that topics related to the pandemic and public discussions about COVID-19 are the most emotionally charged. Thus, the visualization confirms that the algorithm not only distinguishes thematically distinct groups of messages, but also correctly reflects their differentiation by the degree of emotional exhaustion.

The main limitation of the study is the use of a monolingual corpus of tweets, focused mainly on professional communication of medical workers during the pandemic. Such a sample limits the generalization of the results to broader types of digital interaction, where the nature of fatigue may manifest itself differently. An additional factor is the contextual variability of vocabulary inherent in social networks: not all messages with markers of emotional load directly reflect the state of digital fatigue, which creates potential distortions in the automated determination of the index. A methodological limitation also lies in the use of static parameters for dimensionality reduction and density clustering, which do not take into account the temporal dynamics of communication and the evolution of topics during the observation period. In further research, it is planned to expand the corpus using multilingual and multiplatform sources, which will allow us to investigate the phenomenon of digital fatigue in different socio-professional groups. It is planned to integrate temporal and network characteristics of communication to track the dynamics of users' emotional states over time.

6. Conclusions

The paper proposes the method that shifts the focus from the classification of individual messages to the identification of natural thematic groupings of texts, within which cognitive and emotional patterns of interaction are formed. The study confirmed the effectiveness of the proposed method for determining communicative segments of users as a means of interpreting manifestations of digital fatigue in the texts of social networks. The developed approach provides a transition from the traditional analysis of individual messages to the identification of semantically related thematic groups, which allows establishing contextual patterns of the occurrence of cognitive exhaustion. Based on a combination of the SentenceTransformer, UMAP, HDBSCAN, TF-IDF, KeyBERT and Flan-T5 methods, an explanatory model of the communicative structure of the corpus was formed, within which each segment has its own semantic profile and level of association with digital fatigue. The results obtained have both theoretical and practical significance. From a scientific point of view, they deepen the understanding of the contextual nature of digital fatigue and demonstrate the potential of interpreted NLP methods for analyzing mental states in digital environments. From an applied point of view, the proposed approach can be used to develop systems for monitoring psycho-emotional stress in professional teams, educational environments, and social networks. Further development of the work involves expanding the corpus with multilingual data, integrating temporal parameters to track the evolution of topics in dynamics, and using explanatory deep models to analyze the interaction between semantic, linguistic, and behavioral factors of digital fatigue.

Acknowledgements

The authors are grateful to Prof. Tetiana Hovorushchenko, Prof. Elena Zaitseva and other Program Committee members for organizing and conducting the 2nd International Workshop on Advanced Applied Information Technologies (AdvAIT-2025).

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] DataReportal, Digital 2025 July Global Statshot, 2025. URL: https://datareportal.com/reports/digital-2025-july-global-statshot.

[2] International Telecommunication Union (ITU), Facts and Figures 2024: Internet Use, 2024. URL: https://www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-internet-use/.

[3] Microsoft, Work Trend Index, 2025. URL: https://www.microsoft.com/en-us/worklab/work-trend-index.

[4] I. Krak, V. Didur, M. Molchanova, O. Mazurets, O. Sobko, O. Zalutska, O. Barmak, Method for political propaganda detection in internet content using recurrent neural network models ensemble, in: CEUR Workshop Proceedings, volume 3806, 2024, pp. 312–324.

[5] M.Durkot, N. Melnyk, Comparative analysis of real-time semantic segmentation algorithms, Computer Systems and Information Technologies 4 (2024) 92–97. doi:10.31891/csit-2024-4-11.

[6] V. Varriale, M. A. Camilleri, A. Cammarano, F. Michelino, J. Müller, S. Strazzullo, Unleashing digital transformation to achieve the sustainable development goals across multiple sectors, Sustainable Development 33 (1) (2025) 565–579. doi:10.1002/sd.3139.

[7] T. Supriyadi, S. Sulistiasih, K. H. Rahmi, B. Pramono, A. Fahrudin, The impact of digital fatigue on employee productivity and well-being: A scoping literature review, Environment and Social Psychology 10 (2) (2025). doi:10.59429/esp.v10i2.3420.
[8] M. L. Sharma, G. Rajan, The power of play: psychological benefits of integrating gamification in education, Digital Pedagogy 120 (2025). doi:10.25215/1300442190.14.
[9] A. Kovan, Digital fatigue within the family: psychosocial reflections of constant connectedness, Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi 27 (Aile Özel Sayısı) (2025) 333–350. doi:10.32709/akusosbil.1722337.
[10] A. Khairati, I. Ifdil, N. A. W. Zulfi, D. F. Annisa, Y. E. Putri, Digital overload: understanding social media fatigue in higher education based on demographics and technology usage, Islamic Guidance and Counseling Journal 8 (2) (2025). doi:10.25217/0020258654200.
[11] P. D. Deep, Y. Chen, Student burnout and mental health in higher education during COVID-19: online learning fatigue, institutional support, and the role of artificial intelligence, Higher Education Studies 15 (2) (2025) 381–401. doi:10.5539/hes.v15n2p381.
[12] C. C. Hsu, Z. Obermeyer, C. Tan, A machine learning model using clinical notes to identify physician fatigue, Nature Communications 16 (2025) 5791. doi:10.1038/s41467-025-60865-4.
[13] B. Sarraf, A. Ghasempour, Impact of artificial intelligence on electronic health record-related burnouts among healthcare professionals: systematic review, Frontiers in Public Health 13 (2025) 1628831. doi:10.3389/fpubh.2025.1628831.
[14] T.S, Saranya, S. Pradhan, P. Priyanshi, Application of Deep Learning Techniques to Detect Workplace Burnout (June 05, 2025). URL: http://dx.doi.org/10.2139/ssrn.5285816.
[15] O. Barmak, O. Mazurets, I. Krak, A. Kulias, A. Smolarz, L. Azarova, K. Gromaszek, S. Smailova, Information technology for creation of semantic structure of educational materials, in: Proceedings of SPIE – The International Society for Optical Engineering, volume 11176, 2019, p. 1117623. doi:10.1117/12.2537064.
[16] Hugging Face, Sentence Transformers, 2025. URL: https://huggingface.co/sentence-transformers.
[17] S.R. Pokhrel, T. Claire, On harnessing semantic communication with natural language processing, IEEE Internet of Things Journal 12(11) (2025) 15022–15031.
[18] B. Li, Y. Zheng, R. Ran, 2DUMAP: two-dimensional uniform manifold approximation and projection for fault diagnosis, IEEE Access 13 (2025) 12819–12831.
[19] T. Blachowicz, J. Wylezek, Z. Sokol, M. Bondel, Real-time analysis of industrial data using the unsupervised hierarchical density-based spatial clustering of applications with noise method in monitoring the welding process in a robotic cell, Information 16 (2) (2025) 79.
[20] I. Krak, O. Zalutska, M. Molchanova, O. Mazurets, E. Manziuk, O. Barmak, Method for neural network detecting propaganda techniques by markers with visual analytic, in: CEUR Workshop Proceedings, volume 3790, 2024, pp. 158–170.
[21] S. Wadhwa, A. Shukla, Sentiment analysis and word cloud analysis on e-waste news headlines using Python, Int. J. Environ. Waste Manag. 37 (2) (2025) 201–214.
[22] Kaggle, Healthcare Workers Burnout Dataset, 2025. URL: https://www.kaggle.com/datasets/mindyng/healthcareworkersburnout.
[23] O. Sobko, O. Mazurets, M. Molchanova, I. Krak, O. Barmak, Method for analysis and formation of representative text datasets, in: CEUR Workshop Proceedings, volume 3899, 2025, pp. 84–98. URL: https://ceur-ws.org/Vol-3899/paper9.pdf.
[24] Hugging Face, Sentence Transformers: all-MiniLM-L6-v2, 2025. URL: https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2.
[25] scikit-learn, Feature Extraction API Documentation, 2025. URL: https://scikit-learn.org/stable/api/sklearn.feature_extraction.html.
[26] Hugging Face, Transformers Model Documentation: FLAN-T5, 2025. URL: https://huggingface.co/docs/transformers/model_doc/flan-t5.
[27] M.M. Lindemann, Enhancing structural inductive biases of sequence-to-sequence models for semantic parsing and beyond, 2025. doi:10.7488/era/6221.