MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190409

Towards Emotion-Sensitive Conversational User Interfaces in Healthcare Applications

Kerstin Denecke[a], Richard May[a], Yihan Deng[a]

[a] *Bern University of Applied Sciences, Institute for Medical Informatics, Bern, Switzerland*

Abstract

Perception of emotions and adequate responses are key factors of a successful conversational agent. However, determining emotions in a healthcare setting depends on multiple factors such as context and medical condition. Given the increase of interest in conversational agents integrated in mobile health applications, our objective in this work is to introduce a concept for analyzing emotions and sentiments expressed by a person in a mobile health application with a conversational user interface. The approach bases upon bot technology (Synthetic intelligence markup language) and deep learning for emotion analysis. More specifically, expressions referring to sentiments or emotions are classified along seven categories and three stages of strengths using treebank annotation and recursive neural networks. The classification result is used by the chatbot for selecting an appropriate response. In this way, the concerns of a user can be better addressed. We describe three use cases where the approach could be integrated to make the chatbot emotion-sensitive.

Keywords:

Conversational user interface, sentiment analysis, deep learning, natural language processing

Introduction

A fundamental shift in healthcare takes place, driven by an aging population and the increasing incidence of chronic conditions that are induced by behavior. Changing individual behavior is increasingly at the center of healthcare. The reactive system where a patient's acute illnesses are treated is evolving to a treatment more centered on patients, prevention, and the ongoing management of chronic conditions. Beyond, mobile health applications are increasingly used by patients offering the opportunity to collect health data, to continuously monitor the personal health and to be accompanied by a personal health coach all over the day.

To realize such applications, conversational user interfaces (CUI) gained in interest for mobile health applications in the last years [1]. Different terms have been used for a conversational user interface or agent such as: chatbot, machine conversation system, virtual agent, dialogue system, and chatterbot. A CUI-based system is a computer program that interacts with users using natural language (written or spoken). The purpose of such system is to simulate a human conversation. There are two types of chatbots: Unintelligent chatbots interact using a predefined conversation flow. Intelligent chatbots use machine learning to automatically generate responses on the fly. The chatbot architecture integrates a language model and computational algorithms to generate an informal chat communication between human and computer using natural language. Interacting with intelligent agents is not a new topic, but reliable linguistic functionality, availability as services and inclusion of intelligence through machine learning and deep learning has increased its popularity.

CUI have been used in health related applications for example to achieve a health behavior change [2] or to support disease self-management. Lokman and Zain [3] introduced a chatbot that serves as a virtual dietitian for diabetic patients. The chatbot asks questions and gives at the end a diet advice suitable for the current diabetic situation. The conversation is going along a path that is remembered by the system to consider all answers in the decision making.

Similar to this virtual dietitian, the majority of existing medical chatbots only allows for constrained user input (e.g. multiple choice of several options instead of natural language input) to avoid misunderstandings that can occur within natural language interpretation. However, a more realistic, natural interaction requires a natural language user interface without predefined answers, but with integrated information extraction and natural language processing capabilities to recognize and interpret the content correctly. In this work, we are focusing on such chatbots with unconstrained natural language input capabilities. In these applications, users communicate with the system in their own words. In this way, it is possible to express personal emotions and sentiments in certain situations, for example when a person is not feeling well. However, this requires CUI-based applications that can analyze and interpret emotions from user input which are so far rarely available.

The objective of this work is to introduce a concept for analyzing emotions and sentiments resulting from interactions with a CUI. More specifically, our ultimate goal is to equip a health chatbot with capabilities to determine emotions and sentiments expressed in a chat with a CUI-based application in order to better focus suggestions and to increase the impact of recommendations of the application on the behavior of a user (e.g. increase compliance, react on negative emotions).

The main questions addressed in this paper are:

- How emotion and sentiment analysis could be integrated in mobile health applications with CUI?
- Which use cases exist for emotion-sensitive CUI in healthcare?

Conversational agents and sentiment analysis

There are several mobile health applications using CUI available. Some of them have been studied with respect to efficacy in clinical trials; most of them not [1]. Amoto et al. [4] introduced a chatbot-based recommender systems HOLMeS (Health On-Line Medical Suggestions). The system is designed to autonomously interact with a user by understanding natural language in a chat and acting as a human physician. It provides general information on itself or the affiliated medical center, collects patient information and enables a patient to book an appointment in the affiliated medical center. It is implemented using the IBM Watson Conversation Service and trained via the Bluemix platform. VPBot, a SQL-based chatbot that simulates a patient that medical students can interview. VPBot was successfully exploited in Harvard Medical School's virtual patient program [5]. In such use cases for education or information provision, emotions and sentiment are not extremely relevant or do not even occur. Emotion and sentiment in this context comprises subjective, emotional statements e.g. description of the health status ("I am feeling well"), outcome of a treatment or experiences with it ("The therapy session was helpful"), emotions ("I feel enirely loss").

However, there are use cases where it becomes essential to identify and analyse expressions bearing emotions. Woebot [6] is a chatbot designed for supporting cognitive behavioral therapy. The chatbot allows to enter emotions by selecting terms from a list of suggestions. This limits the user to comprehensively express his or her actual emotions and feelings. In that case, natural language processing and sentiment analysis could be very useful to assess and address the user’s mood and situation carefully. When it comes to mobile applications that aim at encouraging behavior changes, behavior models and motivation strategies have to be considered.

Aberg [7] analyzes different motivation strategies realized as chatbots as a mean to motivate people to live more sustainable lives. The effect of motivational factors from behavioral psychology were tested, and as well as the impact on people’s food consumption habits. The findings of this paper were based on three chatbot prototypes; one that is built on the motivational factor of information; a second one that is implemented on the motivational factor of goal-setting, and a third one that follows the motivational factor of comparison. The result from the user interviews indicates that chatbots can affect and motivate people to consume food in a more sustainable way.

Analyzing emotions and sentiments resulting from interactions with CUI has so far only rarely been addressed. There are multiple ways to enable a chatbot to choose an emotion category for a response. On the one hand, the chatbot can be equipped with a personality and background knowledge. On the other hand, training data can be used to find the most frequent response emotion category for an emotion in a given response and use this as the response emotion.

Previous research by Skowron proposed affect listeners, i.e. conversational systems that can respond to user’s utterances on a content-, but also on an affect-level [8]. Zhou et al. [9] describe an emotional chatting machine that can generate appropriate responses fitting in content and emotion to a user’s response. The architecture consists of a recurrent neural network enabled with GRU cells with attention mechanism. It contains three different mechanisms for generating responses with a specific emotion: External knowledge serves to model emotions explicitly using an external emotion vocabulary. Internal memory captures emotion dynamics and finally, different emotion categories are represented as embedded vector. Socher et al. introduced a sentiment treebank that includes fine grained sentiment labels for parsing trees of sentences. On this treebank, they applied recursive deep models to predict sentence level sentiment. With a relatively complicated treebank annotation, the proposed method has better recognized the negated sentiment and achieved more than 80% overall accuracy [10].

Sentiment and emotion analysis in a medical context has been mainly addressed for web content. Denecke and Deng reviewed the state of the art and studied the challenges of sentiment analysis in medical settings [11]. They found out that given the varying usage and meanings of terms, sentiment analysis from medical documents requires a domain-specific sentiment source and complementary context-dependent features to be able to correctly interpret the implicit sentiment. The challenges of sentiment and emotion analysis in health chatbots have not yet been considered so far. Further, health applications equipped with emotion and sentiment analysis are still missing. In contrast to existing work, our aim is not to create a chatbot that formulates its responses with certain emotion terms, but to develop a method to analyse a user statement to select an appropriate, motivating or encouraging response given a specific user emotion.

Material and methods

Synthetic intelligence markup language (SIML)

SIML, pronounced "si mal", is used to build the chatbot’s brain in our application (https://simlbot.com). It is a derivative of Extensible Markup Language (XML) and is able to react to user input, collect and manage data, learn from it and generate new content for conversations. The official parser was developed for C# and can be used on all Windows platforms supporting .NET Framework 4.5 or higher as well as under Linux and Mac (Mono). SIML contains two specifications: SIML Classic and SIML Modern/OSCOVA. All specifications are interpreted differently and are aimed at different target groups. SIML Classic is used in our application because we resist on using internal natural language processing as SIML Modern offers.

In SIML, all data is arranged in the form of a large decision tree and can be addressed by pattern matching. SIML files consist of a large number of tags. The basic structure is created by concepts and models, which are enclosed by the SIML tag. Concepts are the basic unit of knowledge storage in SIML. Each concept is a rule for matching an input and converting to an output, and consists of a pattern, which matches against the user input, and a response, which is used in generating the chatbot answer. In SIML, it is possible to use regular expressions and loops within a model, to interpret JavaScript and to create own tags to call specific methods of the program code. Another function of SIML is the usage of mappings which we will use to select an appropriate response for a user statement containing emotions and sentiments.

Sentiment and emotion analysis

Traditional sentiment analysis approaches classify texts or sentences according to their overall sentiment which can be positive, negative, neutral or even have more fine-grained sentiment categories. They often base upon a bag of word representation. Additionally, the polarity shifting and syntactic structures are transformed into rules to regularize the composition of sentiment at sentence level. However, the performance of bag of words based methods has not reached an accuracy of 80%.

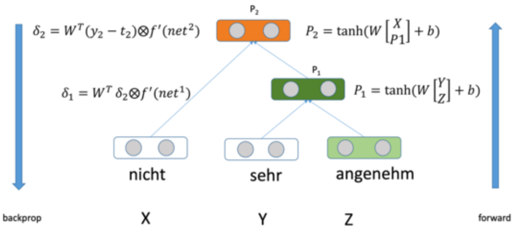

Figure 1. Compositional function of recursive neural networks and backpropagation of errors. Light green color indicates the positive polarity of joy, dark green represents intensified polarity and joy while orange shows negated polarity of joy.

A tree-based composition of a sentiment treebank has been proposed by Socher et al. [10] in order to improve the accuracy. The recursive neural network that is used in their approach requires a well-formed tree annotation corpus as training data. A recursive neural network is a type of deep neural network created by applying the same set of weights recursively over a structured input, in this case a linguistic parse tree of a sentence. The annotation of a sentiment treebank is quite laborious and sophisticated, since the polarity of the sentence (polarity at root) must be determined by a composition of all subtrees. The annotator needs to judge the polarities at different levels and provides also an overall value to the tree root. Our approach for an emotion analyzer is inspired by the work of Socher et al. [10], see Figure 1 and 3.

The classification of user statements in the context of a chatbot is an important step for automatic, appropriate answering of user statements by a chatbot. The concrete task is to categorize a user statement into predefined emotional categories. We currently employ seven axes of emotions (disgust, joy, surprise, anger, fear, sadness, contempt), see Figure 2. The polarity is transformed to a respective emotional tree. For each emotion class, a treebank classifier is trained. The emotion of one sentence is represented as the following polarity vector:

$P_i = (x_{anger}, x_{disgust}, x_{joy}, x_{surprise}, x_{fear}, x_{contempt}, x_{sadness})$

where P_i represents the i-th user statement of the chatbot conversation. The value x represents the normalized strength of the emotion where we distinguish 3 classes (low, medium, high). To obtain a vector of polarity for a user statement, seven instances of the classifier based on recursive neural networks pass through the seven polarity treebanks for initial pre-training.

Before the classification starts, the user statement is linguistically parsed into a syntactic tree. Each node of the tree is represented by word vectors learned from a corpus of German Wikipedia and medical forum entries. During training, one feedforward step is conducted to learn the structure (bottom-up) whereas a back-propagation step is performed to adjust the errors (top-down).

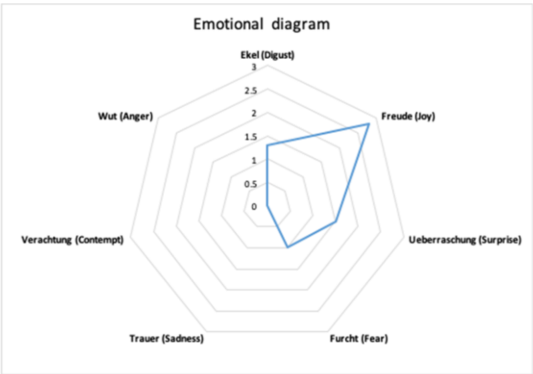

Figure 2: Emotional diagram with seven axes: disgust, joy, surprise, fear, sadness, contempt, anger

In Figure 2, we can see one example of recursive neural networks that we will apply for the emotion detection. The leaf nodes represent the emotion-related terms in the parsed sentence. In this example, the polarity of "angenehm" (pleasant) has been intensified by "sehr" (very) and finally negated by "nicht" (not). Apparently, the syntactic tree has better presented the semantic scope with negation. For the negation parsing, we have chosen FastContext[1] [12] with German negation corpus.

The objective of the optimization is maximizing the probabilities of the correct prediction while minimizing the cross-entropy error between the predicted distribution y_i and target distribution t_i. For the classification into one of the three polarity strength classes, the posterior probability over labels given the word vector via $Y^x = \mathrm{softmax}(W \cdot X)$

where $W \in \mathbb{R}^{3Xd}$ is the emotion classification matrix.

Results

Concept for an emotion-sensitive chatbot

Our concept for integrating emotion analysis into a SIML-based chatbot is shown in Figure 4. A user input resulting from an interaction with the chatbot application is analyzed by the above mentioned emotion analysis algorithm. The resulting polarity vector with strength per emotion class is used to select the corresponding answers based on predefined selecting metrics. SIML mappings assign an emotion label after receiving the emotion vector. Depending on the emotion, we use loops within the models to find matching answers to the user input in the chatbot's knowledge base. These answers can be randomly selected from a pool of matching answers to make the conversation more natural. The answer is displayed on the user interface of the application. In this way, the conversation can be continued coherently considering the emotional fluctuation of users. At the current step, we focus on recognizing the primary emotional polarity of a user input. For treebank construction, we will select 100 sentences for each of the seven classes of emotions (see Figure 3) and parse them into a tree structure. These tree structures will afterwards be annotated with polarity values ranging from 1 to 3 by human annotators.

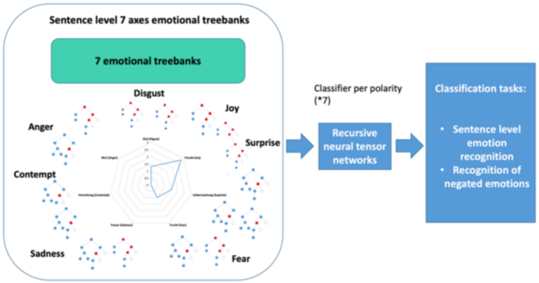

Figure 3. Training of the emotion analyzer based on annotated treebanks

Use case scenarios

Traditional models of care delivery basically base upon face-to-face interactions between clinicians and patients. However, new technologies are augmenting this interaction model and fundamentally transforming the ways in which clinicians deliver care to individuals. Mobile apps, for example, can facilitate tracking and monitoring. These remote and self-care-oriented technologies may help creating a truly interactive healthcare ecosystem for patients. The following use cases base on mainly two paradigms: Engaging individuals by interacting with a CUI and utilizing remote and self-care-oriented technologies to support and empower individuals. Emotion and sentiment analysis will help to realize these two paradigms.

[1] https://github.com/jianlins/FastContext

Addressing concerns during medication management

eMMA, the electronic medication management assistant [14], is a mobile application to support in the medication management. In its current version, it is designed to collect compliance information from a patient, to provide information on interactions between food and medications and to answer his or her questions on the medication [14]. Equipping eMMA with capabilities for automatically recognizing and classifying concerns of a user mentioned in the application could help to address appropriately the concerns of the user. In this way, the compliance could be improved. The chatbot will not convince the user, but recognizes concerns which are stored for discussion with the physician and provides information for engaging the user to take the medication as recommended by the physician. In case the system detects serious concerns and symptoms, it could suggest to contact the physician.

Addressing problems and concerns during self-anamnesis

Ana is a system for collecting the medical history within the context of music therapy using a CUI [15]. Equipping Ana with emotion and sentiment analysis capabilities would enable the system to identify situations when the user runs into problems and concerns because of the questions he or she has to answer. In a usability study with that system [15], the users confirmed, that they liked the possibility of entering free text to communicate with the chatbot. However, they asked for better interactions with the chatbot, in particular when she states that she is not feeling well. Such expressions could be detected and analyzed by our method for providing an appropriate reaction.

Determine cognitions and suggest behavior changes

Another use case is to recognize the tone of a client's response to suggestions made by a chatbot. The suggestions might be for example therapeutic suggestions within a mobile cognitive behavior therapy (CBT) or motivational suggestions within a mobile application targeting at achieving a personal health goal. Within the HABIT project, a mobile application for CBT has been developed that uses a chatbot interface for delivery of CBT [16]. CBT shares the idea that behavior change may be affected via cognitive change [17]. A chatbot system for CBT requires facilities to assess a client's mood overall, or the tone of the response to a particular suggestion in the therapeutic chat. For example a negative answer, high emotional, neglecting a suggestion of the therapist might be interpreted as low acceptance of the immediately prior recommendation. Our emotion analyzer could support in automatically analyze the responses of a user, supporting the therapist in analyzing the chatlog afterwards or enabling chatbot responses that address these moods and tones of user statements.

Discussion

In the general domain, the objective of enriching a chabot with emotions is to achieve a more human-like interaction style. In contrast, in our scenarios the objective is to determine the emotion of a patient for future interpretation by the treating physician or by addressing concerns directly by the chatbot.

Sentiment analysis in the medical domain differs from sentiment analysis in other domains [11]. We address this fact, by training the treebank classifier on a data set of health-related content. Additionally, health-related lexicons could be integrated in the parsing process. Asgar et al. [18] suggest a bootstrapping model and a dataset of health reviews to learn a health-related sentiment lexicon. Rane et al. introduce their concept for using sentiment analysis to improve emotional health of a user [19]. More specifically, they analyze the sentiment or emotion of a user and display specific media to counter the emotion. They used a Naïve Bayes classifier based on supervised learning that was trained on a twitter data set. Besides the different technology, our approach aims at considering the peculiarities of sentiment in the health care domain, which is lost when training is realized with a non-domain-specific data set.

The use of ontologies to support interpretation of several types of data (audio, video, and text) from a chatbot's environment has been analyzed in previous work. Formal concept-based rules are suggested to express affective behavior aiming at improving the empathy of bots [12]. The proposed technique relies on semantic technologies such as OWL and SWRL languages. Affective states are exploited to improve the bot's empathy in the interaction based on an emotion ontology. In a chatbot conversation, user statements might be short. This makes it even more challenging to find appropriate methods. It is still unknown to what extent the previous statements in a conversation and the user context can contribute to the interpretation of expressed emotions.

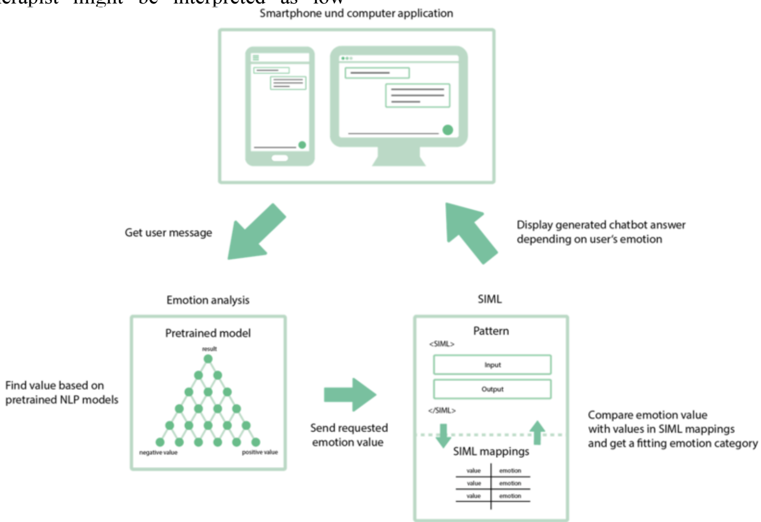

Figure 4: System architecture

Conclusions

In this paper, we introduced a concept for an emotion analyzer for natural language statements as they result from a chatbot conversation. This approach is currently implemented. In future work, we will evaluate the above mentioned method. For this purpose, we will compare the recursive neural networks with the traditional kernel based method in terms of sentence level polarity recognition and the detection of negation. We will run the sentiment analysis using SVM or convolutional neural network on the sentence level annotation to serve as benchmark [20]. Further, we will extend one of the introduced mobile applications with the emotion analyzer.

References

[1] L. Laranjo, A.G. Dunn, H. Ly Tong et al. Conversational agents in healthcare: a systematic review. Journal of the American Medical Informatics Association, 25(9), 2018, 1248–1258

[2] A. Fadhil and S. Gabrielli. Addressing challenges in healthy lifestyles: The Al-chatbot approach. In Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth '17, New York, USA, 2017, pages 261-265

[3] A.S. Lokman and J.M. Zain. An architectural design of virtual dietitian (vidi) for diabetic patients. Proceedings of the International Conference on Computer Science and Information Technology, 2009, pages 408-411

[4] F. Amato, S. Marrone, V. Moscato, G. Piantadosi, A. Picariello, C. Sansone: Chatbots Meet eHealth: Automatizing Healthcare. In: Impedeovo D, Pirlo G (eds.) Proc. of the Workshop on Artificial Intelligence with Applications in Health, CEUR Workshop Proceeedings, volume 1982, 2017, pages 40-49, http://ceur-ws.org/Vol-1982/

[5] B.P. Kerfoot, H. Baker, T.L. Jackson, W.C. Hulbert, D.D. Federman, R.D. Oates, and W.C. DeWolf. A multiinstitutional randomized controlled trial of adjuvant web-based teaching to medical students. Academic Medicine: Journal of the Association of American Medical Colleges, 81:224-30, 2006

[6] K.K. Fitzpatrick, A. Darcy, M. Vierhile. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. JMIR Ment Health. 2017 Jun 6;4(2):e19. doi: 10.2196/mental.7785

[7] J. Aberg: Chatbots as a mean to motivate behaviour change. Master's thesis in Interaction Technology and Design. Umea University, Sweden, 2017

[8] M. Skowron: Affect listeners: acquisition of affective states by means of conversational systems. COST'09 Proceedings of the Second international conference on Development of Multimodal Interfaces: active Listening and Synchrony, 2009, pages 169-181

[9] H. Zhou, H. Huang, T. Zhang, X. Zhu and B. Liu. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. AAAI Conference on Artificial Intelligence, 2018

[10] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. Manning, A. Ng and C. Potts. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. Conference on Empirical Methods in Natural Language Processing (EMNLP 2013), 2013

[11] K. Denecke and Y. Deng. Sentiment analysis in medical settings: New opportunities and challenges. Artificial Intelligence in Medicine, 2015, 64(1):17-27 10.1016/j.artmed.2015.03.006

[12] A. Justo, F.R. Jensen, R. Bonacin, I. Calado, J.C. Dos Reis. Exploring Ontologies to Improve the Empathy of Interactive Bots. In Proceedings of the 27th IEEE International Conference on Enabling Technologies (WETICE'18). Paris, France, 2018, pp. 261-266

[13] W.W. Chapman, D. Hillert, S. Velupillai S, et al. Extending the NegEx lexicon for multiple languages. Stud Health Technol Inform. 2013;192:677-81.

[14] M. Tschanz, T.L. Dorner, J. Holm, K. Denecke: Using eMMA to manage medication. IEEE Computer, vol. 51, issue 08, pages 18-25, 2018

[15] K. Denecke, S. Lutz Hochreutener, A. Pöpel, and R. May: Self anamnesis with a Conversational User Interface: Concept and Usability Study. Methods of Information in Medicine, 2018 (to appear)

[16] J. Warren, E. Tempero, I. Warren, A. Sathianathan et al.: Experience Building IT Infrastructure for Research with Online Youth Mental Health Tools, Proceedings of the ASWEC conference 2018

[17] K.S. Dobson, D.J.A. Dozois. Historical and Philosophical Bases of the Cognitive-Behavioural Therapies. In Dobson KS (Ed.) Handbook of Cognitive-Behavioural Therapies, The Guilford Press, 2001

[18] M.Z. Asghar, S. Ahmad, M. Qasim, S.R. Zahra, F.M. Kundi. SentiHealth: creating health-related sentiment lexicon using hybrid approach. Springerplus. 2016 Jul 20;5(1):1139. doi: 10.1186/s40064-016-2809-x. eCollection 2016

[19] P. Rane, K. Bhansali and S. Nair: Sentiment Analysis to Improve Emotional Health of User. International Journal of Computer Applications 120(1):21-24, June 2015

[20] J. Shi, J.F. Hurdle. Trie-based rule processing for clinical NLP: A use-case study of n-trie, making the ConText algorithm more efficient and scalable. J Biomed Inform. 2018 Sep;85:106-113. doi: 10.1016/j.jbi.2018.08.002

Address for correspondence

Kerstin Denecke

Bern University of Applied Sciences, Institute for Medical Informatics, Quellgasse 21, 2501 Biel, Switzerland,

E-Mail: kerstin.denecke@bfh.ch