

## Improving Sentiment Analysis with Biofeedback Data

Daniel Schlör\*, Albin Zehe\*, Konstantin Kobs\*, Blerta Veseli, Franziska Westermeier,  
Larissa Brübach, Daniel Roth, Marc Erich Latoschik, Andreas Hotho

Julius-Maximilians University Würzburg  
Am Hubland, 97074 Würzburg, Germany  
{schloer, zehe, kobs, veseli, hotho}@informatik.uni-wuerzburg.de  
{franziska.westermeier, larissa.bruebach}@stud-mail.uni-wuerzburg.de  
{daniel.roth, marc.latoschik}@uni-wuerzburg.de

### Abstract

Humans frequently are able to read and interpret emotions of others by directly taking verbal and non-verbal signals in human-to-human communication into account or to infer or even experience emotions from mediated stories. For computers, however, emotion recognition is a complex problem: Thoughts and feelings are the roots of many behavioural responses and they are deeply entangled with neurophysiological changes within humans. As such, emotions are very subjective, often are expressed in a subtle manner, and are highly depending on context. For example, machine learning approaches for text-based sentiment analysis often rely on incorporating sentiment lexicons or language models to capture the contextual meaning. This paper explores if and how we further can enhance sentiment analysis using biofeedback of humans which are experiencing emotions while reading texts. Specifically, we record the heart rate and brain waves of readers that are presented with short texts which have been annotated with the emotions they induce. We use these physiological signals to improve the performance of a lexicon-based sentiment classifier. We find that the combination of several biosignals can improve the ability of a text-based classifier to detect the presence of a sentiment in a text on a per-sentence level.

**Keywords:** sentiment detection, brain-computer-interface, bio-sensing, affective computing

### 1. Introduction

Sentiment analysis has long been an active field of research in the natural language processing (NLP) community due to its widespread applicability and its potential to guide people in important decisions (Wang et al., 2012; Rill et al., 2014; Kobs et al., 2020). However, sentiment analysis for texts except tweets and product reviews, especially in languages other than English, has proven to be a challenging task, mostly due to the difficulty of getting sufficient training data (Zehe et al., 2017; Gangula and Mamidi, 2018; Schmidt and Burghardt, 2018).

According to Caicedo and Van Beuzekom (2006), emotional response typically has three components: subjective feeling (e.g., self-report), motor expression (e.g., facial expression), and physiological arousal (e.g., heart rate and brain waves). A labelling process typical for sentiment analysis is based purely on self-reports. Such reports are very time-consuming and tedious tasks, and they are highly prone to the individual’s subjective rating.

In contrast, emotion recognition, or emotion identification based on objective measurements of neurophysiological signals is common in the field of affective computing, meaning “computing that relates to, arises from, or influences emotion” (Picard, 2000, p. 1). In studies about measuring emotions using neurophysiological data, emotions are often triggered by perceptual stimuli, e.g. visual (Bhardwaj et al., 2015), auditory (Lin et al., 2010) or audiovisual stimuli (Kimmatkar and Babu, 2018). However, there still is no clear consensus about the appropriate approach to model and hence to classify emotions, i.e., if emotions are discrete constructs or if they are on continuous scales separated in groups. Various approaches exist, for example, to classify emotions in terms of valence (neutral, positive, negative), in terms of the quadrants of the

valence-arousal model (Lin et al., 2010), or even in terms of different levels of extent of valence and arousal (Horlings et al., 2008).

So far, measurements of neurophysiological signals are not common in NLP research. In this paper, we propose to merge both approaches, sentiment analysis of annotated texts and objective measurements of neurophysiological signals. Our approach uses affordable and convenient devices, i.e., a smart watch and a consumer-grade electroencephalography (EEG) headband. To this end, we

- i) make a dataset available that includes sentiment annotations, as well as two types of biofeedback data, namely heart rate and EEG data<sup>1</sup>,
- ii) perform an initial study showing that the biofeedback contains signals useful for sentiment analysis, and
- iii) discuss possible extensions and directions for future work, where we believe that incorporating information from biofeedback into sentiment classifiers will be helpful.

In our initial study using German texts, we find that either heart rate or EEG data can not be used by itself to predict sentiment as accurately as a text sentiment classifier. However, by combining a simple text sentiment classifier with heart rate and EEG data, we can improve the detection of presence or absence of sentiment in the text.

In the following Section 2 we provide an overview of related work. In Section 3 our task and approach are then described. After giving details for our dataset in Section 4, in Sections 5 and 6 we describe and discuss our results. We conclude the paper in Section 7 with a summary of our findings and an outlook on future work.

<sup>1</sup>[https://professor-x.de/datasets/dataset\\_onion\\_biofeedback.zip](https://professor-x.de/datasets/dataset_onion_biofeedback.zip)

\* equal contribution

## 2. Related Work

There is a large body of work on detecting sentiment from text. A full overview is out of scope for this paper, so we refer to the recent survey in (Zhang et al., 2018). Most recent sentiment analysis methods are based on pre-trained transformer architectures such as BERT (Devlin et al., 2018; Munikar et al., 2019). However, these models still require a rather large amount of data to fine-tune, which is not available for every language and domain.

Similarly, there exists some work investigating the detection of emotions from biofeedback data. The study by Choi et al. (2017) indicates that it is possible to detect unhappy emotions that were induced by visual stimuli from heart rate variability.

In an EEG setting, visual stimuli achieved high accuracy in emotion classification (Petranonakis and Hadjileontiadis, 2009). For other stimuli such as audio, a link from the recorded EEG data to the perceived emotion was also reported (Lin et al., 2010). Further, affect detection using an EEG was proposed to visualize emotional states of users augmenting avatar-mediated communications (Roth et al., 2019c; Roth et al., 2019b).

Using EEG data for sentiment analysis was previously proposed in (Gu et al., 2014). In their work, subjects were instructed to visualize single words in their thoughts. Their EEG response was then used as input to machine learning models to predict the valence of these words. One subject achieved better scores for concrete words, while abstract words were better estimated by lexicons.

Multimodal emotion recognition using EEG, pulse, and skin conductance with audio-visual stimuli was also performed (Takahashi, 2004).

To the best of our knowledge, our study is the first to combine lexical sentiment analysis approaches with heart rate and EEG signals collected in a natural text reading task.

## 3. Methodology

We define two separate sentence-level tasks for our study: *sentiment detection* and *sentiment classification*. The first task aims to determine whether or not a sentence conveys any emotion (regardless of its polarity), while the second provides a more fine-grained classification of sentences into the three classes *negative*, *neutral*, and *positive*. We hypothesize that biofeedback is a good indicator for at least the first task, as physiological activity can change when feeling both positive and negative emotions.

For both of these tasks, we evaluate classifiers based on a) the text of the sentence, b) the readers’ biofeedback data collected while reading the sentence, and c) a combination of both.

### 3.1. Text Based Sentiment Classifiers

Due to the small amount of available data, we use the lexicon based classifier provided by the German version of TextBlob<sup>2</sup>, which assigns each word a sentiment score from the range  $[-1, 1]$  and then calculates the overall sentiment score for a sentence. It also features a negation detection

that multiplies sentiments of negated words by  $-0.5$ . Using the resulting polarity score  $v(s)$  for one sentence  $s$ , we can define thresholds for the classification of a sentence into one of the desired classes. We classify a sentence as *positive* if  $v(s) > 0.25$ , *negative* if  $v(s) < -0.25$ , and *neutral* otherwise. In the sentiment detection setting, we classify a sentence to contain sentiment if and only if  $|v(s)| > 0.25$ .

### 3.2. Biofeedback Based Sentiment Classifiers

In this study, we compare Random Forests (RF) and linear Support Vector Machines (SVMs) for the detection and classification of sentiment from biofeedback. For both machine learning models, we use the implementation in scikit-learn (Pedregosa et al., 2011) with default parameters. We modify the number of decision trees in the Random Forest to be ten due to the faster training time and better generalization for this low data setting.

Both classifiers receive input based on the readers’ biofeedback while reading the sentence that is to be classified. Let  $B_u^c(t)$  be the value of channel  $c \in C = \{\text{heart rate, EEG}_1, \dots, \text{EEG}_n\}$  for the biofeedback data from user  $u$  at timestamp  $t$ . For each sentence  $s$ ,  $\text{begin}_u(s)$  and  $\text{end}_u(s)$  give the timestamp when reader  $u$  starts and finishes reading the sentence, respectively. All timestamps recorded for user  $u$  and channel  $c$  are given in  $T_u^c$ . Then,  $T_u^c(s) = [t_b, \dots, t_e]$  with  $\text{begin}_u(s) \leq t_i < \text{end}_u(s)$  describes all timestamps for user  $u$  and channel  $c$  which were recorded while reading the sentence  $s$ . The sample-rate  $sr_c$  describes how many timestamps and thus sensor values are recorded per second. From these time series, we derive the features for our classifiers.

#### 3.2.1. Heart Rate Features

For the heart rate data, we define  $ba_u^{hr}(s)$  as the absolute average heart rate of user  $u$  while reading sentence  $s$ :

$$ba_u^{hr}(s) = \frac{\sum_{t \in T_u^{hr}(s)} B_u^{hr}(t)}{|T_u^{hr}(s)|}. \quad (1)$$

The relative average heart rate of user  $u$  is normalized per user, given as

$$b_u^{hr}(s) = \frac{ba_u^{hr}(s) - \min(B_u^{hr})}{\max(B_u^{hr}) - \min(B_u^{hr})}. \quad (2)$$

We represent a sentence  $s$  using the values  $b_u^{hr}(s)$  for all users as well as their deltas, that is

$$\hat{b}_u^{hr}(s) = b_u^{hr}(s) - b_u^{hr}(s-1). \quad (3)$$

#### 3.2.2. EEG Features

For the EEG data, we use Fourier transformed and filtered values to better represent the common spectral bands present in brain activity (Murugappan and Murugappan, 2013). We select the time window where the reader  $u$  reads the sentence  $s$ , and select all sensor values with timestamps within this window.

$$b_u^{eegi}(s) = [B_u^{eegi}(t_b), \dots, B_u^{eegi}(t_e)] \quad (4) \\ \text{with } [t_b, \dots, t_e] = T_u^{eegi}(s)$$

<sup>2</sup><https://pypi.python.org/pypi/textblob-de/>.

For each EEG channel  $i \in \{1, \dots, 8\}$  and sentence  $s$  Fourier transformation is applied to this window, producing  $\hat{b}_u^{eegi}(s)$ .

We use  $\hat{b}_u^{eegi}(s)$  for all EEG channels and all users to represent sentence  $s$ . Note, that  $\hat{b}_u^{eegi}(s)$  contains all frequencies between 0 and  $\frac{sr_{eeg}}{2}$  in a fine-grained resolution. We reduce the number of features by a) applying a band-pass filter between 13 and 30 Hz to remove unwanted frequencies and b) applying a principal component analysis (PCA). We found 3 principal components to work best.

#### 4. Dataset

This section describes the dataset of texts annotated with heart rates, which we enrich with sentiment annotations as well as EEG data for one additional reader.

For our study, we use the BioReaderData dataset presented by Schlör et al. (2019) consisting of 4 medium-length texts in German language with different topics that should trigger different emotional reactions. The texts contained in the dataset have a length between 502 and 633 words and are described in the following:

- a) *Kangaroo*<sup>3</sup>: an excerpt from a humorous narrative book,
- b) *Dogs*<sup>4</sup>: a neutrally written factual text from National Geographic,
- c) *Genie*<sup>5</sup>: a short report about the tragic story of a feral child with many negatively connoted words, and
- d) *James*<sup>6</sup>: a neutrally written chronological description of a child’s murder.

The existing dataset contains heart rate measurements of 15 German native speakers that were reading the given texts using the BioReader app. Subjects were equipped with a Polar M600 smartwatch that measures heart rate with a sampling frequency  $sr_{hr} = 2\text{Hz}$ . The app captures the reading progress, such that heart rate data can be aligned to the text.

**Extending the Dataset with Sentiment Information** In order to perform sentiment analysis on the dataset, we let three subjects annotate each sentence in the dataset on a three-part polarity scale as either *negative*, *neutral*, or *positive*. A majority voting then determined the gold standard label, discarding all sentences where a majority vote was not possible. This resulted in a dataset with 164 sentences. A description of the texts in terms of sentence counts as well as label distribution is shown in Table 1.

<sup>3</sup>Marc-Uwe Kling, Die Känguru-Chroniken: Ansichten eines vorlauten Beuteltiers “Theorie und Praxis”, Ullstein eBooks, 2010

<sup>4</sup><https://www.nationalgeographic.de/wissenschaft/2018/07/wohin-verschwanden-die-ersten-hunde-amerikas>

<sup>5</sup><https://www1.wdr.de/stichtag/stichtag-554.html>

<sup>6</sup>[https://de.wikipedia.org/wiki/Mord\\_an\\_James\\_Bulger&stableid=176294324](https://de.wikipedia.org/wiki/Mord_an_James_Bulger&stableid=176294324)

Text	# Sentences	# Neg.	# Neu.	# Pos.
<i>Kangaroo</i>	56 (50)	20	21	9
<i>Dogs</i>	31 (31)	5	17	9
<i>Genie</i>	45 (43)	29	12	2
<i>James</i>	42 (40)	28	8	4
<b>Total</b>	174 (164)	82	58	24

Table 1: The number of sentences per text in the dataset as well as the number of sentences that are labeled as negative, neutral, and positive by a majority vote of three annotators. The number of sentences per text that received a label in the majority vote is given in parentheses.



Figure 1: OpenBCI headband as worn for EEG data collection during our study.

**Extending the Dataset with EEG Data** To extend the dataset with EEG measurements, we used a headband with an OpenBCI<sup>7</sup> Cyton board (PIC32MX250F128B micro-controller) and 8 electrodes. Electrode placements were made near the frontal and the parietal lobes at the positions Fp1, Fp2, F7, F8, T3, T4, F3 and F4 according to the 10-20 system, as these were shown to yield good features to capture the emotional state (Lin et al., 2010; Bos and others, 2006). Previous work has shown that emotion classification can be achieved with a limited number of electrodes (Bhardwaj et al., 2015). The setup is depicted in Figure 1. We presented the sentences from BioReaderData dataset to the reader while capturing their EEG data. The EEG data was obtained with a sampling rate of  $f_{eeg} = 250\text{ Hz}$ , resulting in 378704 data points.

After obtaining the EEG data, the reader was asked to review the annotated gold standard sentiment labels with respect to the perceived sentiment. The reader agreed with the gold standard label for 95% of the samples. All 8 cases of disagreement involved a sentiment change from or to neutral, indicating that these sentences can be considered borderline cases where the presence of sentiment is arguable. We use the EEG data for all sentences as biofeedback, including the sentences with disagreement since this setup is the more difficult task and also more realistic, since for

<sup>7</sup><https://openbci.com>

Classifier	Detection (RF/SVM)	Classification (RF/SVM)
Majority Vote	39.3	22.2
Stratified Random	51.2	31.0
Text	55.1	<b>46.4</b>
Heart Rate	<u>55.0/43.3</u>	<u>33.8/26.2</u>
EEG	<u>46.5/49.2</u>	<u>31.1/31.7</u>
Text, Heart Rate	<u>55.7/43.5</u>	<u>39.9/27.9</u>
Text, EEG	<u>51.2/48.6</u>	<u>36.1/34.0</u>
Heart Rate, EEG	<u>52.9/49.4</u>	<u>37.7/31.7</u>
Text, Heart Rate, EEG	<b><u>58.5/51.3</u></b>	<u>38.5/35.4</u>

Table 2: Results for sentiment detection and classification. All numbers in percent macro-averaged F1-scores. Where applicable, the first number is the performance of a Random Forest, the second number the performance of a linear SVM. The best performance for each task is given in bold, the better model for each feature set is underlined.

larger scale study the assessment of individual sentiment perception per sample will not be feasible.

## 5. Experiments

We perform experiments on the BioReaderData dataset with both classifiers, Random Forest and linear SVM, for the tasks of sentiment detection and sentiment classification. We evaluate all feature set combinations to better understand the influence a certain feature set has on the overall performance. Additionally, we employ two baselines: i) Majority vote, that always predicts the most frequent class: *non-neutral / emotional* in sentiment detection and *negative* in sentiment classification. ii) Stratified random, that takes the class distribution of the training set into account and samples the prediction from this distribution. All baselines and classifiers are evaluated using a stratified 5-fold cross-validation that is repeated 10 times. We report macro-averaged F1-scores for all methods.

### 5.1. Sentiment Detection

For the sentiment detection, we merge the positive and negative labels in the BioReaderData data. Applying all classifiers to the data results in the macro F1-scores reported in the second column of Table 2. Training a Random Forest on heart rate data of 15 subjects results in a comparable sentiment detection performance as the text based method. While text and heart rate achieve better performance than the baseline methods, using EEG data alone did not perform better than random sampling from the training data’s class distribution. Combining all three feature sets and training a Random Forest yields the best F1-score. In most cases, Random Forest performs better than SVM, which in turn works better on standalone EEG data.

### 5.2. Sentiment Classification

The third column of Table 2 describes the results for the sentiment classification task, where we have three possible classes. No model or feature combination provides a better performance than the text-based classifier in this setting. As

in the sentiment detection task, Random Forest performs better in almost all cases. Only EEG data is again better processed using a linear SVM.

## 6. Discussion

Our experiments show that the biofeedback data we have collected contains information about the sentiment that the readers experience when reading the provided texts. Using only the readers’ heart rates, we can achieve almost the same performance as a text-based classifier for the detection of sentiment in a text. Furthermore, we have shown that combining biofeedback features and lexicon-based text features can improve the overall performance over that of any of the components. Especially introducing EEG features yields a notable performance boost in comparison to heart rate plus text features. This suggests that, even though EEG features by themselves couldn’t reach competitive performance levels, signals within this data help to enrich other feature sets.

We suggest that this finding can be used to facilitate the collection of annotations for long texts: In a first step, multiple users could be asked to read, for example, a full novel while collecting their biofeedback data. After that, a classifier based on the text and biofeedback can be used to detect emotional passages in the text, which can then be manually annotated for polarity or emotions. This would filter out sentences that do not contain emotions at all and therefore do not need to be labelled, saving a large amount of time for annotation. Since our biofeedback data was obtained using a consumer grade fitness watch and an affordable EEG headband, this approach scales well to a large number of annotators. It is important to note that higher quality electrodes, as well as semi-wet and wet EEG systems may lead to better results. However, despite higher-grade EEG systems may produce better data quality, we believe that enhancing the classification through our method is possible, and further, specifically applicable to consumer applications.

For the sentiment classification, our biofeedback based approach did not yield comparable results to the text based classification. The measured physiological arousal as well as the derived features and models did not capture *what* kind of emotion was felt but just *that* an emotion was felt. For the heart rate, this result is unsurprising, since a faster heart beat can come from a negative or positive excitement, such as being scared or falling in love. For EEG data, we would have expected different results, since EEG data has already been successfully incorporated in sentiment classification contexts (Kimmatkar and Babu, 2018). However, in contrast to our experimental setup, Kimmatkar and Babu (2018) used video-clips presented to a subject instead of text and recorded the EEG data using a 62 channel system instead of the 8 channel consumer grade OpenBCI system in our experiment. In addition, our EEG-based results only rely on one subject and one repetition whereas the aforementioned study had 15 participants repeat the experiments three times. Since Lakhan et al. (2019) suggest that in general consumer grade EEG systems such as OpenBCI can be used to detect emotions successfully, we hope to improve the performance by introducing more participants in

the future, similar to the success of our human heart rate ensemble for sentiment detection.

As an additional point, we believe that biofeedback data presents a way of implicitly labelling sentences in relation to their context: medium-length texts, which are used in this study, consist of multiple sentences. While a sentence may seem neutral when judged in an isolated manner, the context of the text is very important to the person that is reading it. Biofeedback, such as heart rate or brain waves, does not just reflect the emotional state of the reader given the current sentence, but for the overall story up to that point. While many studies induced only one stimulus at a time (Choi et al., 2017; Lin et al., 2010; Gu et al., 2014), our study involved continuously reading sentences that build upon a given theme, for example humor or drama. Therefore, future labeling of sentences in texts should also consider the text before, such that the emotion that is currently induced by the text is better reflected.

This paper demonstrates a first approach, showing that biofeedback data can be used to improve text-based sentiment classifiers. Further studies will improve the data acquisition as well as processing. We are confident that the collection of a larger dataset and the inclusion of additional kinds of biofeedback will bring further improvements to the results in this first study.

## 7. Conclusion and Future Work

In this paper, we have presented an initial study about improving sentiment analysis tasks by incorporating biofeedback from subjects reading texts. We found that, while heart rate and EEG information was able to support machine learning models when detecting the presence of emotion in texts, it did not improve differentiation of said emotion as positive or negative.

In this work, we only measured physiological arousal using heart rate and EEG. In the future, we also plan to incorporate motor expression into the classification, which was, for example, proposed as classification input to analyze social interaction in virtual realities (Roth et al., 2019a). As reading usually does not induce sudden body movements, but possibly facial expressions reflecting the reader’s emotions, additionally capturing and estimating them using the front camera of a smartphone is a promising option (Tarnowski et al., 2017), which will be implemented within the BioReader app. Introducing more complex text-based sentiment and emotion classifiers can also contribute to a better classification. Especially when facial expressions recorded by the front camera are introduced, multimodal systems such as MixedEmotions (Buitelaar et al., 2018) will be an interesting tool to study.

We also want to refine our evaluation scenario by collecting a larger dataset and labeling sentences such that the story context is captured. We believe that a larger scale EEG study can further reveal insights into the emotional thought process while reading texts. We plan to include more participants as well as complex features such as differential asymmetry (DASM) and rational asymmetry (RASM) (Duan et al., 2013) and we want to incorporate artificial neural networks using EEG data in the time domain, which are able to reflect features besides the frequency space.

## 8. Bibliographical References

- Bhardwaj, A., Gupta, A., Jain, P., Rani, A., and Yadav, J. (2015). Classification of human emotions from eeg signals using svm and lda classifiers. In *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, pages 180–185. IEEE.
- Bos, D. O. et al. (2006). Eeg-based emotion recognition. *The Influence of Visual and Auditory Stimuli*, 56(3):1–17.
- Buitelaar, P., Wood, I., Negi, S., Arčan, M., McCrae, J., Abele, A., Robin, C., Andryushechkin, V., Sagha, H., Schmitt, M., Schuller, B., Sánchez-Rada, J. F., Iglesias, C., Navarro, C., Giefer, A., Heise, N., Masucci, V., Danza, F., Caterino, C., and Ziad, H. (2018). Mixed-emotions: An open-source toolbox for multi-modal emotion analysis. *IEEE Transactions on Multimedia*, PP:1–1, 01.
- Caicedo, D. G. and Van Beuzekom, M. (2006). How do you feel? *An assessment of existing tools for the measurement of emotions and their application in consumer products research*.
- Choi, K.-H., Kim, J., Kwon, O. S., Kim, M. J., Ryu, Y. H., and Park, J.-E. (2017). Is heart rate variability (hrv) an adequate tool for evaluating human emotions?—a focus on the use of the international affective picture system (iaps). *Psychiatry research*, 251:192–196.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Duan, R.-N., Zhu, J.-Y., and Lu, B.-L. (2013). Differential entropy feature for eeg-based emotion classification. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 81–84. IEEE.
- Gangula, R. R. R. and Mamidi, R. (2018). Resource creation towards automated sentiment analysis in telugu (a low resource language) and integrating multiple domain sources to enhance sentiment prediction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.
- Gu, Y., Celli, F., Steinberger, J., Anderson, A. J., Poesio, M., Strapparava, C., and Murphy, B. (2014). Using brain data for sentiment analysis. *JLCL*, 29:79–94.
- Horlings, R., Datcu, D., and Rothkranz, L. J. M. (2008). Emotion recognition using brain activity. ACM.
- Kimmatkar, N. V. and Babu, V. B. (2018). Human emotion classification from brain eeg signal using multimodal approach of classifier. In *Proceedings of the 2018 International Conference on Intelligent Information Technology*, pages 9–13.
- Kobs, K., Zehe, A., Bernstetter, A., Chibane, J., Pfister, J., Tritscher, J., and Hotho, A. (2020). Emote-controlled: Obtaining implicit viewer feedback through emote based sentiment analysis on comments of popular twitch.tv channels. *ACM Transactions on Social Computing*.
- Lakhan, P., Banluesombatkul, N., Changniam, V., Dhithijayratn, R., Leelaarporn, P., Boonchieng, E., Hompoon-sup, S., and Wilaiprasitporn, T. (2019). Consumer grade

- brain sensing for emotion recognition. *IEEE Sensors Journal*, 19(21):9896–9907.
- Lin, Y.-P., Wang, C.-H., Jung, T.-P., Wu, T.-L., Jeng, S.-K., Duann, J.-R., and Chen, J.-H. (2010). Eeg-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, 57(7):1798–1806.
- Munikaar, M., Shakyaa, S., and Shrestha, A. (2019). Fine-grained sentiment classification using bert. In *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, volume 1, pages 1–5. IEEE.
- Murugappan, M. and Murugappan, S. (2013). Human emotion recognition through short time electroencephalogram (eeg) signals using fast fourier transform (fft). In *2013 IEEE 9th International Colloquium on Signal Processing and its Applications*, pages 289–294. IEEE.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Petrantonakis, P. C. and Hadjileontiadis, L. J. (2009). Emotion recognition from eeg using higher order crossings. *IEEE Transactions on information Technology in Biomedicine*, 14(2):186–197.
- Picard, R. W. (2000). *Affective computing*. MIT press.
- Rill, S., Reinel, D., Scheidt, J., and Zicari, R. V. (2014). Politwi: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. *Knowledge-Based Systems*, 69:24–33.
- Roth, D., Bente, G., Kullmann, P., Mal, D., Purps, C. F., Vogeley, K., and Latoschik, M. E. (2019a). Technologies for social augmentations in user-embodied virtual reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, pages 1–12.
- Roth, D., Brübach, L., Westermeier, F., Schell, C., Feigl, T., and Latoschik, M. E. (2019b). A social interaction interface supporting affective augmentation based on neuronal data. In *Symposium on Spatial User Interaction*, pages 1–2.
- Roth, D., Westermeier, F., Brübach, L., Feigl, T., Schell, C., and Latoschik, M. E. (2019c). Brain 2 communicate: Eeg-based affect recognition to augment virtual social interactions. In *Mensch und Computer 2019 - Workshopband*, Bonn. Gesellschaft für Informatik e.V.
- Schlör, D., Veseli, B., and Hotho, A. (2019). Multimedia aus rezipientenperspektive: Wirkungsmessung anhand von biofeedback. *6. Tagung des Verbands Digital Humanities im deutschsprachigen Raum eV— DHd 2019: Digital Humanities multimedial & multimodal*.
- Schmidt, T. and Burghardt, M. (2018). An evaluation of lexicon-based sentiment analysis techniques for the plays of gotthold ephraim lessing. In *Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 139–149.
- Takahashi, K. (2004). Remarks on emotion recognition from bio-potential signals. In *2nd International conference on Autonomous Robots and Agents*, pages 186–191.
- Tarnowski, P., Kolodziej, M., Majkowski, A., and Rak, R. J. (2017). Emotion recognition using facial expressions. In *ICCS*, pages 1175–1184.
- Wang, H., Can, D., Kazemzadeh, A., Bar, F., and Narayanan, S. (2012). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In *Proceedings of the ACL 2012 system demonstrations*, pages 115–120. Association for Computational Linguistics.
- Zehe, A., Becker, M., Jannidis, F., and Hotho, A. (2017). Towards sentiment analysis on german literature. In *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, pages 387–394. Springer.
- Zhang, L., Wang, S., and Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4):e1253.