

Improving Emotion Detection through Real-Time Translation of Text to ML Models Trained in Different Languages

Richard Hoehn*

Middle Tennessee State University

Dr. Jaishree Ranganathan†

Middle Tennessee State University

August 8, 2023

Abstract

This Qualifying Exam¹ research paper investigates the potential of improving Emotion Detection (ED) through real-time translation of text to multiple Machine Learning (ML) models trained in different languages specifically English and German. The research aims to address the challenges posed by the lack of comprehensive labeled datasets and language fragmentation in ED analysis research and projects.

By extending an original English language dataset with a translated dataset from German to English the training data (extended dataset) the hope is that better prediction rates can be achieved in ED analysis applications. Furthermore, translating English to German to extend the German dataset and training another ML model in parallel using PySpark and accessing both models in real-time by a simple RESTful API may improve the prediction rates by dual processing a sentence at the same time.

In order to present the results presented both orally to the MTSU's Computational Data & Science Committee and within this paper labeled datasets both in English and German were collected, translated, and two ML models built that can be accessed via an API to get a prediction on an input (POST or GET) sentence both in German or English.

*Electronic address: rhoehn@mtmail.mtsu.edu; corresponding author

†Electronic address: jaishree.ranganathan@mtsu.edu

¹PhD students of MTSU Computational and Data Science program are required to complete a Qualifying Exam in their first year in the program.

Contents

1	Introduction	4
1.1	Human Emotion as a Theoretical Framework	4
1.2	Significance of Emotion Detection (ED) in Machine Learning	5
1.3	ED challenges: Lack of Labeled Datasets and Language Fragmentation	6
1.4	Motivation for the research project focusing on extending ED datasets through translation and evaluating the impact on prediction rates	6
1.5	Objectives & Scope of the Research	6
2	Literature Review	7
2.1	Emotion Detection Importance and Diverse Populations	7
2.2	Lack of Labeled Datasets and Fragmentation caused by Different Languages	7
3	Methodology	8
3.1	Data procurement process of ED datasets in English and German	8
3.2	Development of a translation application for converting different languages into English and vice versa	9
3.3	Training Multiple ML models with PySpark	10
3.4	Evaluation Metrics and methodologies for measuring the performance of the ML models	11
3.5	Written & Oral Presentation of Research Finding	12
4	Results and Analysis	13
4.1	Prediction Results of Original & Extended Datasets	13
4.1.1	English Dataset Results	13
4.1.2	German Dataset Results	13
4.2	Analysis of the impact of Extending Datasets by Translation	14
4.3	Findings and Implications for Improving Emotion Detection	14
4.4	Conclusion and Future Work	14
5	Demonstration and Application	15
5.1	Description of the developed translation application for real-time text translation	15
5.2	Showcase of the application's functionality and its integration with the Emotion Detection process	15
5.3	Demonstration of the real-time processing of translated text and emotion detection using the ML models	15
5.4	GIT Repository Details	15
	References	16
	List of Figures	18
A	Appendix – Dataset Details	19
A.1	Translation Setup	19
A.2	Details on English Dataset	21
A.3	Details on German Dataset	23

B	Appendix – Application & Code Details	25
B.1	Application	25
B.2	Code Details	26
B.3	RESTful API Server	26

1 Introduction

Emotion detection has gained significant interest in recent years as a means to understand human behavior and improve communication effectiveness. By analyzing text or speech, these detection processes can accurately identify and classify emotions expressed by individuals or groups, often times in real-time.

By employing a variety of ED processes, including natural language processing and machine learning, applications can detect emotions with high precision. These ED systems find applications in diverse fields, such as sentiment analysis, customer feedback analysis, voting and news stances, and mental health monitoring, which all aim at contributing to a deeper understanding of human emotions and their impact on various aspects of life.

1.1 Human Emotion as a Theoretical Framework

The concept and number of emotions can vary depending on the theoretical framework or model being considered. One well-known model is the Plutchik’s Wheel of Emotions Figure 1, which proposes eight (8) primary emotions and their contrasting pairs. In this paper we will be focusing on the basic eight emotions and not on their sub-emotions.

Below are the eight primary emotions according to the Plutchik’s Wheel of Emotions [19], a detailed reference of emotions used in this research can be found in Table 1 on page 8.

- Happiness: A feeling of Joy (Fun), contentment, or delight.
- Sadness: A state of unhappiness, sorrow, or grief.
- Anger: An intense emotional response associated with feelings of displeasure, frustration, or hostility.
- Worry: An emotional response triggered by perceived threats or danger, often associated with anxiety or terror.
- Surprise: A sudden and unexpected reaction to something unusual or startling.
- Disgust and/or Hate: A strong aversion or revulsion towards something unpleasant, offensive, or repulsive.
- Trust and/or Love: A sense of reliance, confidence, or faith in someone or something.
- Enthusiasm: A feeling of excited expectation or eagerness towards future events.

It’s important to note that emotions are complex and multifaceted, and different models may propose alternative categorizations or variations in the number of primary emotions. However for the scope of this research project the most common eight (8) emotions are used as described above.

By aligning the emotional labels across the English [6] and German [21] datasets being used in this research all eight listed above were used. Details on the linkage of the two datasets is described in

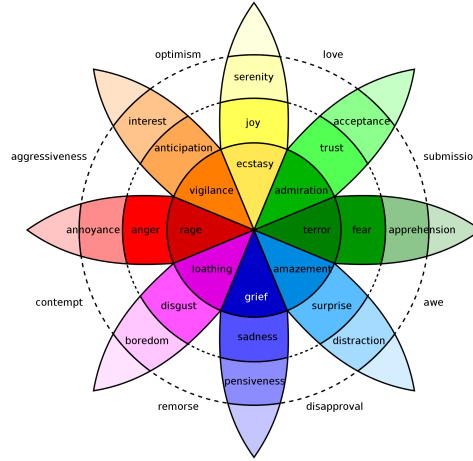


Figure 1: Plutchik Wheel of Emotions

Section 3.1 with an emphasis on the Table 1’s cross reference on page 8.

1.2 Significance of Emotion Detection (ED) in Machine Learning

Emotion Detection (ED) analysis in Machine Learning (ML) models has gained significant attention due to its potential applications in various fields, including customer sentiment analysis, stance detection in regard to a specific targets such as news and voting[10], mental health monitoring[4], and human-computer interactions (chatbots[1]) to just name a few.

By accurately identifying and understanding emotions from text data, ML applications can assist in improving user experiences, decision-making processes, and overall human-machine interactions in a positive manner[4, 10]. As an example, the primary goals of chatbots are entertainment, social contact, and novelty interaction, with a strong emphasis on productivity, scale-ability and cost reductions. In chatbots, there has been a immense requirement to simulate human-like characteristics and behavior during human-machine interactions[8]. Per Kusal et al. [1] these chatbots also improve customer engagement by offering friendliness, comfortness, flexibility, and efficient assistance. Chatbots should provide customers with more engaging responses, directly addressing their issues. The goal of deploying ED applications is for users perceive chatbots as companions rather than simple assistants. Per Adikari et al. the majority of user requests are emotional rather than informative. Chatbots gained the capacity to respond emotionally to customers due to machine learning and sentiment analysis evolution[1].

I usually know almost exactly how I feel. The problem is, I just can’t tell anyone.

– Meg Cabot

Due to the before mentioned wide scope of uses for ED this field of research and the work to advance it provides significance and importance in enabling machines to comprehend and respond appropriately to human interactions. This in turn contributes to the advancement of artificial intelligence, natural

language processing and in many cases the utilization of text-to-speech applications most all of us in today's society use and rely on [13].

Lastly, in Chen et al.[2] paper based on emotion detection for online learning, it states that with emotion recognition research, we need to focus on accuracy and real-time performance in order to apply emotional recognition based on physiological signals to solve practical problems; they key part of their claim is that in order for ED to be practical real-time methods needs to be used.

1.3 ED challenges: Lack of Labeled Datasets and Language Fragmentation

ED is an immensely growing field, however unlike Sentiment Analysis (SA) the availability of large datasets for training purposes of ML models is much smaller [17, 13]. The primary reason of the lack of data stems from the emotion detection data requiring primarily supervised learning dataset, which require time and effort to procure, and mostly created and labeled by humans. To compound on the issue of the lack of data; the many datasets that are available are in multiple languages. A significant amount of time took place to find the German labeled

1.4 Motivation for the research project focusing on extending ED datasets through translation and evaluating the impact on prediction rates

This paper is significant because an entire industry of automated emotion reading (Text, Image, Video, and Sound) technologies is quickly emerging [17, 18]. The market for emotion recognition software is forecast to reach at least \$3.8 billion by 2025 [17, 18]. And since ED is already being incorporated into products for purposes such as marketing, robotics, driver safety, customer service and a multiple of others, it is fitting to work on research to solve issues with prediction rates and lack of teaching / learning data.

With the before mentioned issues regarding the quantity (lack of) and fragmentation of the ED datasets this research was motivated by finding a way to combine different language datasets into a single set for training processes to improve the predictability of ED. The hope is that by use of translation that the predictability of ED can be improved by use of real-time language conversion and processing.

1.5 Objectives & Scope of the Research

The research project aims at answering the following research questions:

- Can by translating English data to German and extending an original German dataset increase the predictability of the ED model?
- Can by translating German data to English and extending an original English dataset increase the predictability of the ED model?
- Can by translating in real-time an input to multiple languages improve the predictability based on the combined output of two model's results.

2 Literature Review

To ensure comprehensive coverage of the latest developments and research in emotion detection and analysis for this research paper, consideration and review was only given to incorporating research papers and online news articles that were published after the year 2015, thereby providing an up-to-date perspective on the subject matter. By implementing the 2015 boundary, our methodology aims to provide a more up-to-date outlook on the subject matter at hand and what areas the ED field in moving in in respect to ML applications and processes.

2.1 Emotion Detection Importance and Diverse Populations

Emotion Detection (ED) holds significant importance in various domains, including psychology, social sciences, and human-computer interaction. It enables the analysis of human emotions at scale, providing valuable insights into individual and collective emotional states. In Chowanda et. al. paper[3] they believe that "Emotions hold a paramount role in the conversation, as it expresses context to the conversation."

Furthermore, emotions can differ across age groups, genders, cultures, and languages. Including data from diverse populations helps in developing inclusive and culturally sensitive ED models. It ensures that the models are not biased towards specific groups and can accurately detect emotions in a wide range of individuals.

2.2 Lack of Labeled Datasets and Fragmentation caused by Different Languages

The lack of diverse and large-scale data hinders the development and training of accurate ED models, limiting their effectiveness and general ability across different contexts and populations.

In recent research, efforts have been made to address these limitations by exploring new techniques for emotion detection. For instance, a study by Kapoor[7] et al. proposes analyzing the entire spectrum of the language source to predict emotional changes. This provided some help, but quickly established itself as time consuming and need of expertise in the subject matters. Overcoming the scarcity of labeled datasets and embracing innovative approaches like the before mentioned study by Kapoor[7] et al. and their analysis can contribute to the advancement of ED research. Such advancements hold the potential to enhance the accuracy and applicability of emotion detection models, enabling a deeper understanding of human emotions and more effective responses in various domains.

The paper by Mascarell[10] et al. states that unlike images language operate very differently, therefore the lack of labeled datasets poses a significant challenge in text-based Emotion Detection (ED) analysis as it limits the availability of reliable training data for ML models. This scarcity hampers the model's ability to learn and generalize emotions effectively.

Additionally, Kusal et al. further stipulates [9] that the fragmentation caused by different languages further exacerbates the issue, as it reduces the size and diversity of data available for training, resulting in limited cross-lingual generalization and potentially biased models. Overcoming these challenges is crucial for advancing ED research and improving the accuracy and applicability of emotion detection models across various languages and cultures.

3 Methodology

3.1 Data procurement process of ED datasets in English and German

The data procurement was relatively straight forward and in essence entailed using Google search terms for the Emotion Detection in English and German. Some of the main German data was obtained from the dataset built by ETH’s Emotion and Stance Detection for German Text [10] which also hosts a good website with their findings ². The English dataset was downloaded from Kaggle ³ based on Tweets collected in 2021. The English dataset contains over 38,000 tweets and their corresponding emotion labeled. Similarly the German contains over 12,000 sentences and their respective emotions associated.

An important part of the process is to make sure we match the English emotions to those of the German language emotions. Since the two languages are dissimilar we opted to match exact emotions to each other that can be seen in Table 1 below.

Emotion Datasets Labels				
English		German		Used
Name	Count	Name	Count	
Boredom	179	—	—	NO
Love	3842	Vertrauen	316	YES
Relief	1526	—	—	NO
Fun	1776	—	—	NO
Hate	1323	Ekel	29	YES
Neutral	8638	Unklar	314	YES
Anger	110	Ärger	226	YES
Happiness	5209	Freude	140	YES
Surprise	2187	Überraschung	369	YES
Sadness	5165	Traurigkeit	184	YES
Worry	8459	Angst	154	YES
Enthusiasm	759	Antizipation	774	YES
Empty	827	—	—	NO

Table 1: Dataset of English & German and their respective label counts.

Details on how we formatted the the two (English and German) datasets to each other can be viewed in the Appendix A.2 and A.3

The procurement of two ED datasets is needed. The first will be in English and the second in German. Each dataset will be translated into the other’s language for extending purposes. As an example Figure 2 depicts the original English dataset will be extended by the translated German dataset.

The reason for selecting German was based on the author⁴ being fluent in English and German

²<https://mtc.ethz.ch/research/natural-language-processing/emotion-stance.html>

³<https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text>

⁴Richard Hoehn

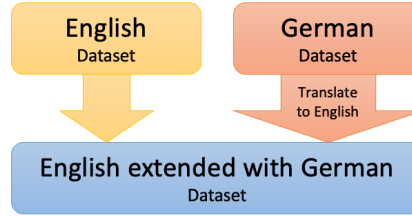


Figure 2: Example of English dataset extension by translating from German.

and hence being able to easily traverse the two languages while working on the code for dataset translations and formatting.

Once the two datasets have been translated, we now extend the original language dataset with the translated one. By this approach we are in essence extending the original datasets to create a larger one with the foreign language data. With both extended datasets we can now train our ML models for English and German.

3.2 Development of a translation application for converting different languages into English and vice versa

Translation of the emotion datasets was performed with Google Translator based on the Medium article[12] in which the author describes option for simple API translations. As an example the following is a brief introduction to Google Translator’s API:

```

1 # Using Deep Translator to leverage Google Translate
2 # Link: https://cloud.google.com/translate/docs/reference/libraries/v2/python
3 from deep_translator import GoogleTranslator
4
5 sentence = 'Chocolate milk is so much better through a straw.'
6
7 translated = GoogleTranslator(source='auto', target='de').translate(sentence)
8 print(translated) # Schokoladenmilch schmeckt durch einen Strohhalm viel besser.
9
10 translated_back = GoogleTranslator(source='auto', target='en').translate(translated)
11 print(translated_back) # Chocolate milk tastes much better through a straw.
  
```

Based on the above example the CSV and JSON datasets were individually loaded, parsed, and translated to their respective counter-part languages. Details of the application can be found in the Appendix A.2 for English and Appendix A.3 for the German. We decided to only use 1,500 row for each language so since the German (DE⁵) dataset was not much larger than that. This meant that the English (EN⁶) datasets was randomly paired down to 1,500 to match that of the German.

Each translated file now contains four (4) columns that are labeled in Tables 2 and 3 with the exact same column naming.

⁵DE = German

⁶EN = English

File: <code>pd_de_translated.csv</code>			
sentence_de	emotion_de	sentence_en	emotion_en
Original Sentence (feature)	Original Emotion (label)	Sentence Translated to English	Emotion set by Cross-Reference from German

Table 2: German CSV file structure after translation to English

File: <code>pd_en_translated.csv</code>			
sentence_en	emotion_en	sentence_de	emotion_de
Original Sentence (feature)	Original Emotion (label)	Sentence Translated to German	Emotion set by Cross-Reference from English

Table 3: English CSV File structure after translation to German

3.3 Training Multiple ML models with PySpark

Using the datasets from Tables 2, 3, and 4 that list an approximate $\sim 1,275$ and $\sim 2,775$ features for training. The algorithms tested in this research project ranged from the traditional machine learning to deep learning techniques, they are:

- **Naïve Bayes:** Per Pujan’s paper[20] on the Naïve Bayes model, it classifies and/or assigns the most likely class to a given example based on its feature vector, simplifying the task greatly by assuming that the features are independent given a class.
- **Generalised Linear Model (GLM):** Per Mueller’s paper on GLM[11] the Generalized linear models extend the concept of the well understood linear regression model.
- **Random Forest Classifier:** Per Schonlau’s paper[16] on the Random Forest Classifier, they state that random decision forests easily adapt to non-linearity found in the data and therefore tend to predict better than linear regression. More specifically, ensemble learning algorithms like random forests are well suited for medium to large datasets.

Each of these algorithms used and their results are listed in the Results and Analysis Section 4 below. In Chowanda et al. paper[3] on Text-based Emotion Techniques multiple algorithms were proposed, this research focused on just three out of their six since the others had lower prediction rates based on Chowanda et al. paper[3] paper.

In order for each of the three ML models to be processed at the same time we crated a Jupyter Notebook that starts an Apache PySpark⁷ session that is accessed via the local server running that accepts RESTful API calls with details of the server in Appendix B.3.

⁷PySpark: <https://spark.apache.org/docs/latest/api/python/index.html>

ML Models			
Model	Data	Rows for Training (85%)	Rows for Testing (15%)
A	English Original	1,275	225
B	English Extended By German	2,775	225 same as Model "A"
C	German Original	1,275	225
D	German Extended By English	2,775	225 same as Model "C"

Table 4: ML Models and their respective settings

3.4 Evaluation Metrics and methodologies for measuring the performance of the ML models

The evaluation of ML models trained on the English, German and their respective extension datasets is a critical step in assessing their performance and reliability. We chose to use the MulticlassClassificationEvaluator⁸ because is a well known tool in the field of natural language processing (NLP) research, specifically in multi class classification tasks such as emotion detection that has multiple labels associated with it; in our case eight (8) see Table 1 on Page 8. Generally this evaluator is used to evaluate the performance of machine learning (ML) models, using transformer-based architectures like BERT, RoBERTa, or DistilBERT[14], with DistilBERT[14] being the one we used in this research project.

The MulticlassClassificationEvaluator provides the essential evaluation metrics, including accuracy, precision, recall, and F1 score, for ED classifications. These metrics provide valuable insights into a model’s ability to correctly classify emotions across multiple classes and languages. Our research project is no different and the following measurements were used for our analysis:

- **Accuracy:** Accuracy measures the overall correctness of the emotion predictions made by the model. It is calculated as the ratio of correctly classified instances to the total number of instances in the dataset.
- **F1 Score:** Precision represents the proportion of correctly predicted emotions among all instances predicted as a specific emotion with respect to the labels. Recall measures the proportion of correctly predicted emotions among all instances that truly belong to a specific emotion. With these two the F1 Score ($F_1 = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$) combines precision and recall into a single metric, providing a balanced evaluation.

While evaluating the models we focused on the Accuracy and the F1 Score to illustrate the prediction results of the original versus extended datasets.

⁸URL: <https://spark.apache.org/docs/latest/api/python/reference/api/pyspark.ml.evaluation.MulticlassClassificationEvaluator.html>

3.5 Written & Oral Presentation of Research Finding

In conclusion of the project a research paper and oral presentation is required to pass MTSU Computation Science PhD Qualifying Exam. The listing the datasets used, code for the application, and details of the research project process in this research paper is openly available and hosted on a GitHub repository, allowing for easy access, reproducible, and collaborative engagement with the research findings. The repository is open and available under: <https://github.com/richardhoe/mtsu.coms.qualifying-exam>[5]

4 Results and Analysis

The following section lists the prediction rates of training and testing of the original and extended datasets both in English and German. Overall results seem to indicate that there was no measurable benefit and even negative results based on adding the extended data for Emotion Detection. In the below sub sections details on the results are listed in prose, tables, and formulas.

4.1 Prediction Results of Original & Extended Datasets

Overall the prediction results from the trained models is disappointing. Using the Random Forrest Classifier described in Section 3.4 provided the best prediction results based. Oddly, the prediction results with the extended data where lower than using the original (non-extended) data for training.

4.1.1 English Dataset Results

The Table 5 lists the results of all three ML models, their accuracy and F1 Score based on the English original and extended datasets. The improvements from the original and extended datasets are listed in the last column. In all cases the improvement to using the extended datasets was negative.

Results of English Dataset					
	Original Dataset		Extended Dataset		
	Accuracy	F1 Score	Accuracy	F1 Score	Improvement
Random Forrest	29.72%	20.32%	28.77%	19.47%	-3.20%
Naïve Bayes	5.66%	6.92%	3.77%	4.11%	-33.39%
GLM	25.94%	25.92%	22.64%	23.25%	-10.37%

Table 5: Model Testing Results on English Dataset

4.1.2 German Dataset Results

The Table 6 lists the results of all three ML models, their accuracy and F1 Score based on the German original and translated from English to German extended datasets. The improvements from the original and extended datasets are listed in the last column. Interestingly the Random Forrest Classifier was the only negative improvement whereas the GLM and Naive Bayes models had improvements. However it should be noted that the prediction rates are so low in the GLM and Naive Bayes that they should not really be considered improvements.

Results of German Dataset					
	Original Dataset		Extended Dataset		
	Accuracy	F1 Score	Accuracy	F1 Score	Improvement
Random Forrest	34.29%	17.51%	18.57%	13.44%	-45.84%
Naïve Bayes	4.29%	4.31%	6.67%	4.21%	+55.48%
GLM	11.90%	12.57%	14.76%	15.88%	+20.03%

Table 6: Model Testing Results on German Dataset

The improvements on the accuracy are based on Equation (1) listed below. It is a simple and straightforward way of comparing the Accuracy from the original trained dataset to the extended trained dataset for purposes of comparing the potential benefits of the translation.

$$Improvement(\%) = \frac{(Accuracy_{Ext}(\%) - Accuracy_{Org}(\%))}{Accuracy_{Org}(\%)} \quad (1)$$

Further

4.2 Analysis of the impact of Extending Datasets by Translation

While extending datasets for machine learning training can often improve[15] model performance, it's important to recognize that this approach has its limitations, especially in ML models that can reach a limit on training data. In our case, relying on translation services for dataset expansion might not consistently yield better results when extending the data from translated text. These services, in our case Google Translate⁹ can introduce errors or inaccuracies in translations especially in the context of emotions of a particular text, leading to noisy or incorrect training data. Consequently, the models may learn from erroneous information, which did negatively impact our models performance on the ED tasks.

4.3 Findings and Implications for Improving Emotion Detection

...

4.4 Conclusion and Future Work

It's crucial to ensure the quality and accuracy of extended data, especially when language translation is involved, to prevent introducing more noise than meaningful information into the training process. Although the research did not garner any significant improvements by translating the German to English as a means on extending the English dataset, it actually made it the prediction results worse.

Future work includes testing and translating different models by using different translation services. Using Google Cloud's translation service was adequate for this research, however there are certainly better translation services that are available but are based on paid subscriptions.

⁹Google Cloud Translate URL: <https://cloud.google.com/translate>

5 Demonstration and Application

5.1 Description of the developed translation application for real-time text translation

Details of the entire application can be found on the github repository [5].

5.2 Showcase of the application's functionality and its integration with the Emotion Detection process

...

5.3 Demonstration of the real-time processing of translated text and emotion detection using the ML models

...

5.4 GIT Repository Details

The listing the datasets used, code for the application, and details of the research project process in this research paper is openly available and hosted on a GitHub repository, allowing for easy access, reproducible, and collaborative engagement with the research findings. The repository is open and available under: <https://github.com/richardhoeHN/mtsu.coms.qualifying-exam>[5]

References

- [1] Achini Adikari et al. “A Cognitive Model for Emotion Awareness in Industrial Chatbots”. In: *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*. Vol. 1. 2019, pp. 183–186. DOI: 10.1109/INDIN41052.2019.8972196.
- [2] Mengnan Chen et al. “Research on Emotion Recognition for Online Learning in a Novel Computing Model”. In: *Applied Sciences* 12.9 (2022). ISSN: 2076-3417. DOI: 10.3390/app12094236. URL: <https://www.mdpi.com/2076-3417/12/9/4236>.
- [3] Andry Chowanda et al. “Exploring Text-based Emotions Recognition Machine Learning Techniques on Social Media Conversation”. In: *Procedia Computer Science* 179 (2021). 5th International Conference on Computer Science and Computational Intelligence 2020, pp. 821–828. ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2021.01.099>. URL: <https://www.sciencedirect.com/science/article/pii/S1877050921001320>.
- [4] Valentina Colonnello, Katia Mattarozzi, and Paolo M Russo. “Emotion recognition in medical students: effects of facial appearance and care schema activation”. In: *Medical Education* 53.2 (2019), pp. 195–205. DOI: <https://doi.org/10.1111/medu.13760>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/medu.13760>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/medu.13760>.
- [5] Richard Hoehn and Jaishree Ranganathan. *Improving Emotion Detection through Real-Time Translation of Text to ML Models Trained in Different Languages*. Version 1.0.1. 2023. URL: <https://github.com/richardhoehn/mtsu.coms.qualifying-exam>.
- [6] Kaggle and data.world. *Emotion Detection from Text*. 2021. URL: <https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text>.
- [7] Shalini Kapoor and Tarun Kumar. “Detecting emotion change instant in speech signal using spectral patterns in pitch coherent single frequency filtering spectrogram”. In: *Expert Systems with Applications* 232 (2023), p. 120882. ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2023.120882>. URL: <https://www.sciencedirect.com/science/article/pii/S0957417423013842>.
- [8] Sheetal Kusal et al. *A Review on Text-Based Emotion Detection – Techniques, Applications, Datasets, and Future Directions*. Apr. 2022.
- [9] Sheetal Kusal et al. “AI Based Emotion Detection for Textual Big Data: Techniques and Contribution”. In: *Big Data and Cognitive Computing* 5.3 (2021). ISSN: 2504-2289. DOI: 10.3390/bdcc5030043. URL: <https://www.mdpi.com/2504-2289/5/3/43>.
- [10] Laura Mascarell et al. “Stance Detection in German News Articles”. In: *Proceedings of the Fourth Workshop on Fact Extraction and VERification (FEVER)*. Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 66–77. DOI: 10.18653/v1/2021.fever-1.8. URL: <https://aclanthology.org/2021.fever-1.8>.
- [11] Marlene Müller. “Generalized Linear Models”. In: (Feb. 2004). DOI: 10.1007/978-3-642-21551-3_24.
- [12] Nidhaloff Nidhaloff. *How to translate text with python*. July 2023. URL: <https://medium.com/analytics-vidhya/how-to-translate-text-with-python-9d203139dcf5>.
- [13] Sajani Ranasinghe et al. “An Artificial Intelligence Framework for the Detection of Emotion Transitions in Telehealth Services”. In: July 2022, pp. 1–5. DOI: 10.1109/HSI55341.2022.9869503.
- [14] Victor Sanh et al. “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter”. In: *ArXiv abs/1910.01108* (2019).

- [15] Iqbal H. Sarker. “Machine Learning: Algorithms, Real-World Applications and Research Directions”. In: *SN Computer Science* 2.3 (Mar. 2021), p. 160. ISSN: 2661-8907. DOI: 10.1007/s42979-021-00592-x. URL: <https://doi.org/10.1007/s42979-021-00592-x>.
- [16] Matthias Schonlau and Rosie Yuyan Zou. “The random forest algorithm for statistical learning”. In: *The Stata Journal* 20.1 (2020), pp. 3–29. DOI: 10.1177/1536867X20909688. eprint: <https://doi.org/10.1177/1536867X20909688>. URL: <https://doi.org/10.1177/1536867X20909688>.
- [17] Jay Stanley. “*Experts Say ‘Emotion Recognition’ Lacks Scientific Foundation*”. 2019. URL: <https://www.aclu.org/news/privacy-technology/experts-say-emotion-recognition-lacks-scientific>.
- [18] Jay Stanley. “*THE DAWN OF ROBOT SURVEILLANCE*”. 2019. URL: <https://www.aclu.org/report/dawn-robot-surveillance>.
- [19] Erik Tromp and Mykola Pechenizkiy. “Rule-based Emotion Detection on Social Media: Putting Tweets on Plutchik’s Wheel”. In: (Dec. 2014).
- [20] Pujan Ziaie et al. “A Naïve Bayes Classifier with Distance Weighting for Hand-Gesture Recognition”. In: vol. 6. Jan. 2009, pp. 308–315. ISBN: 978-3-540-89984-6. DOI: 10.1007/978-3-540-89985-3_38.
- [21] ETH Zurich. *CHeeSE Dataset*. 2019. URL: <https://mtc.ethz.ch/publications/open-source/cheese-dataset.html>.

List of Figures

1	Plutchik Wheel of Emotions	5
2	Example of English dataset extension by translating from German.	9

A Appendix – Dataset Details

The following details in this appendix are for clarification purposes. All examples in this appendix are created via Panda and PySpark using data frames.

A.1 Translation Setup

All example code was created using Jupyter Notebooks.

```
1 # *****
2 # *** Setup Section ***
3 # *****
4
5 import pandas as pd
6
7 # Emotion DataFrame
8 # This was predetermined by review of the emotion from English to German
9 emotion_key = {
10     "boredom": "---",
11     "love": "Vertrauen",
12     "relief": "---",
13     "fun": "---",
14     "hate": "Ekel",
15     "neutral": "Unklar",
16     "anger": "Ärger",
17     "happiness": "Freude",
18     "surprise": "Überraschung",
19     "sadness": "Traurigkeit",
20     "worry": "Angst",
21     "enthusiasm": "Antizipation",
22     "empty ": "---",
23     "---": "Keine"
24 }
25
26 # Create a DataFrame from the emotion_key dictionary
27 df_emotions = pd.DataFrame(emotion_key.items(), columns=['emotion_en', 'emotion_de'])
28
29
30 # *****
31 # *** Translation ***
32 # *****
33 from googletrans import Translator
34 import time
35
36 # This google API take a Sentence and converts to German
37 def translate(sentence, dest_lang):
38     try:
39         translator = Translator()
40         translator.raise_Exception = True
41         translation = translator.translate(sentence, dest=dest_lang)
```

```
42     time.sleep(0.5)  # Add a 1-second delay
43     return translation.text
44 except Exception as e:
45     print(f"Translation Error: {e}")
46     return None
```

A.2 Details on English Dataset

The English dataset is based on tweets¹⁰ that are saved as a csv¹¹ file. The following code was used to filter and process the file for use in this project.

```
1 # *****
2 # *** English Data Preparations ***
3 # *****
4
5 import pandas as pd
6 from tqdm import tqdm
7
8
9 # Load English CSV File into a Dataframe
10 df_en = pd.read_csv("data/english.csv")
11 print(f'Count (row): {len(df_en)}')
12
13 # Print Columns
14 print(f'English Source Columns: {df_en.columns}\n')
15
16 # Filter only labels that we matched with the German counter part JSON File
17 filter_values = ["love", "hate", "neutral", "anger", "happiness", "surprise", "sadness", "worry", "enth
18 df_en = df_en[df_en["sentiment"].isin(filter_values)]
19
20 # Find Empty Sentence Values
21 df_en = df_en[df_en["content"].notnull() & (df_en["content"].str.strip() != '')]
22
23 # Remove unnecessary column
24 df_en = df_en.drop("tweet_id", axis=1)
25
26 # Rename Columns
27 df_en = df_en.rename(columns = {'sentiment': 'emotion_en', 'content': 'sentence_en'})
28
29 # Merge German Emotions onto English
30 df_en = pd.merge(df_en, df_emotions, on='emotion_en', how='left')
31
32 # Add German Sentence Column
33 df_en["sentence_de"] = ""
34
35 # Randomly select 1500 rows
36 df_en = df_en.sample(n=1500, random_state=2023)
37
38 df_en.to_csv('./data/pd_en.csv', index=False)
39
40
41 # Iterate over the rows with tqdm to show the progress
42 for index, row in tqdm(df_en.iterrows(), total=df_en.shape[0]):
43     # Call Translation
```

¹⁰Tweets from Twitter

¹¹csv = Comma Separated Values

```
44     sentence = translate(row["sentence_en"], 'de')
45     # Save Sentence on Column
46     df_en.at[index, 'sentence_de'] = sentence
47
48
49 df_en.to_csv('./data/pd_en_translated.csv', index=False)
```

A.3 Details on German Dataset

The German dataset was downloaded from the ETH's¹² servers via "link" as a JSON file. Since the file was JSON we will need to filter and explode the file based on the emotions that are attached to it.

```
1 # *****
2 # *** German Data Preparations ***
3 # *****
4
5 import pandas as pd
6 from tqdm import tqdm
7
8
9 # Load German JSON File into a DataFrame
10 df_de = pd.read_json("data/german.json", lines=True)
11 print(f'Count (row): {len(df_de)}')
12
13 # Print Columns
14 print(f'German Source Columns: {df_de.columns}\n')
15
16 # Explode the "article_emotion" column to distinct rows
17 df_de = df_de.explode('article_emotion')
18
19 # Filter only labels that we matched with the German counter part JSON File
20 filter_values = ["Vertrauen", "Ekel", "Unklar", "Ärger", "Freude", "Überraschung", "Traurigkeit", "Angst"]
21 df_de = df_de[df_de["article_emotion"].isin(filter_values)]
22
23 # Find Empty Sentence Values
24 df_de = df_de[df_de["title"].notnull() & (df_de["title"].str.strip() != '')]
25
26 # Remove unnecessary columns
27 df_de = df_de.drop("article_id", axis=1)
28 df_de = df_de.drop("article_stance", axis=1)
29 df_de = df_de.drop("paragraphs", axis=1)
30 df_de = df_de.drop("source", axis=1)
31 df_de = df_de.drop("snippet", axis=1)
32
33 # Rename Columns
34 df_de = df_de.rename(columns={'title': 'sentence_de'})
35 df_de = df_de.rename(columns={'article_emotion': 'emotion_de'})
36
37 # Merge English Emotions onto German
38 df_de = pd.merge(df_de, df_emotions, on='emotion_de', how='left')
39
40 # Add English Sentence Column
41 df_de["sentence_en"] = ""
42
43 # Randomly select 1500 rows
```

¹²ETH (German: Eidgenössische Technische Hochschule) or Swiss Federal Institute of Technology

```
44 df_de = df_de.sample(n=1500, random_state=2023)
45
46 df_de.to_csv('./data/pd_de.csv', index=False)
47
48
49 # Iterate over the rows with tqdm to show the progress
50 for index, row in tqdm(df_de.iterrows(), total=df_de.shape[0]):
51     # Call Translation
52     sentence = translate(row["sentence_de"], 'en')
53     # Save Sentence on Column
54     df_de.at[index, 'sentence_en'] = sentence
55
56
57 df_de.to_csv('./data/pd_de_translated.csv', index=False)
```

B Appendix – Application & Code Details

Hello World - Application & Code Details

B.1 Application

Hello World - Application

B.2 Code Details

Hello World - Code Details

B.3 RESTful API Server

The RESTful¹³ API Server will be a simple Python Flask server listening on port 8080 on the local machine.

¹³RESTful API is an interface that two computer systems use to exchange information securely over the internet.