Detection of Emotions by Text Analysis Using Machine Learning

Chethan Harinath[1], Divyansh Prakhar Soni[2], Vivekanand Reddy Malipatel[3]

[1]*Computer Science , Illinois Institute of Technology , A20517331 , charinath@hawk.iit.edu*
[2]*Computer Science , Illinois Institute of Technology , A20526469 , dsoni2@hawk.iit.edu*
[3]*Computer Science , Illinois Institute of Technology , A20324971 , vmalipatel@hawk.iit.edu*

Abstract—This project aims to improve emotion detection in text by leveraging pre-trained Large Language Models (LLMs) through transfer learning. Traditional sentiment analysis models struggle with contextual nuances and cultural variations in human emotions. By fine-tuning LLMs on multilingual datasets, this project seeks to enhance the model's understanding of context and cultural nuances, improving emotion detection across diverse datasets and languages. The project's methodology includes a literature review, model selection, data collection, pre-processing, model training, fine-tuning, and evaluation. Results demonstrate the effectiveness of pre-trained LLMs in capturing emotional nuances and their applicability across languages.

Contents

1. Introduction

Emotion detection in text, a subfield of natural language processing (NLP), plays a vital role in understanding human communication and behavior. Traditional sentiment analysis models often struggle to capture the complexities and nuances of human emotions, especially in multilingual and cross-cultural contexts. This project focuses on enhancing emotion detection capabilities by leveraging pre-trained Large Language Models (LLMs) through transfer learning.

The project aims to address the limitations of traditional sentiment analysis models by improving the understanding of context, idiomatic expressions, and cultural nuances in emotion detection. By fine-tuning LLMs on multilingual datasets, the project seeks to develop a more inclusive and comprehensive framework for emotion analysis that is effective across different languages and cultural contexts.

This report presents the methodology, results, and analysis of the project, highlighting the importance of leveraging advanced NLP techniques for emotion detection. The project's findings contribute to the advancement of NLP applications, particularly in sentiment analysis and emotion detection, with implications for various domains, including customer feedback analysis, mental health assessment, and more nuanced communication systems.

2. Problem Statement

2.1. The Problem

Traditional sentiment analysis models often struggle with contextual nuances, cultural variations, and the subtleties of human emotions, limiting their effectiveness across diverse datasets and languages. This project aims to address these limitations by leveraging the advanced capabilities of pre-trained Large Language Models (LLMs) through transfer learning. By utilizing LLMs, which have been extensively trained on vast amounts of text data, we aim to enhance the model's understanding of context, idiomatic expressions, and cultural nuances in emotion detection. Additionally, the project seeks to extend emotion detection capabilities beyond a single language by employing multilingual models, thereby enabling more inclusive and comprehensive emotion analysis across different languages and cultural contexts.

2.2. Related Work

Recent advancements in Natural Language Processing (NLP) have seen the development of sophisticated models capable of understanding complex linguistic features and sentiments. Studies have explored various aspects of emotion detection, ranging from binary sentiment classification to fine-grained emotion analysis. Traditional approaches relied heavily on manually crafted features and simpler machine learning algorithms like Support Vector Machines (SVM) and Naive Bayes.

The introduction of deep learning brought about more powerful models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), which improved performance by capturing semantic relationships in text. However, these models still struggled with capturing contextual nuances and cultural variations in human emotions.

The advent of Transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), and their multilingual variants like mBERT (multilingual BERT) and XLM-R (Cross-lingual Language Model for Representation), marked a significant leap forward. These models have been pre-trained on diverse corpora, enabling them to grasp a wide range of linguistic nuances.

Research has shown their effectiveness in various NLP tasks, including emotion detection, by fine-tuning them on task-specific datasets. Furthermore, there has been an interest in exploring the capabilities of these models in multilingual settings, aiming to bridge the gap in emotion detection across different languages and cultural contexts.

3. Proposed Work

Our project proposes to leverage pre-trained Large Language Models (LLMs) for emotion detection in text. Specifically, we plan to use models like BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), or their multilingual variants like mBERT (multilingual BERT) and XLM-R (Cross-lingual Language Model for Representation). These models have shown exceptional performance in understanding complex linguistic features and nuances, making them ideal for our task of emotion detection.

The proposed work aims to overcome the limitations of traditional sentiment analysis models by leveraging the advanced capabilities of pre-trained LLMs, thereby enabling more accurate and comprehensive emotion analysis in text.

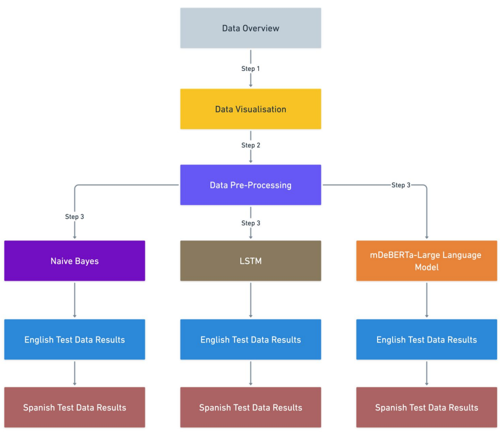

Figure 1. Proposed Project Workflow

4. Exploratory Data Analysis

The "Emotions" dataset [1], sourced from Kaggle, is a valuable resource for sentiment analysis, emotion classification, and text mining tasks. It contains English Twitter messages annotated with six fundamental emotions: anger, fear, joy, love, sadness, and surprise.

This dataset consists of over 416,000 entries, each containing a text string and an associated numerical label. The data is primarily text-based and is likely used for sentiment analysis or related natural language processing tasks. Each text entry represents a sentiment or emotional expression, categorized under labels ranging from 0 to 5. The purpose of this dataset is to enable the development and evaluation of models that can interpret and classify emotional content in text.

4.1. Data Overview

The dataset is diverse, comprising 392,131 unique text expressions, indicating rich variability in linguistic expressions. The labels are distributed across a range of 0 to 5, with a mean value of approximately 1.55 and a standard deviation of 1.49. This distribution suggests that the dataset might not be evenly balanced across different categories, which is a common challenge in classification tasks. It consists of two columns:

1. text: This column contains a string feature representing the content of the Twitter message.
2. label: This column contains a classification label indicating the primary emotion conveyed in the message. The labels range from 0 to 5, corresponding to the emotions sadness, joy, love, anger, fear, and surprise, respectively.

4.2. Data Visualisation

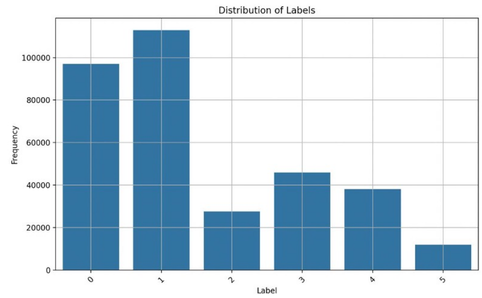

Figure 2. Distribution of emotion labels in the dataset

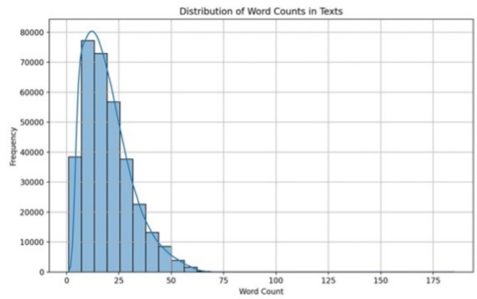

Figure 3. Distribution of text sequence length in the dataset

4.3. Data Pre-Processing

We have developed a comprehensive preprocessing pipeline using Python, leveraging libraries like Pandas for data handling, NLTK for natural language processing tasks, and TextBlob for spelling corrections. This preprocessing script is designed to clean and standardize the text data, ensuring it is well-suited for our subsequent analysis and machine learning models.

The preprocessing steps include expanding contractions to their full forms, reducing elongations in words to avoid analysis distortions, and correcting common spelling errors to enhance data quality. Additionally, the script removes unnecessary characters such as URLs and usernames, which are irrelevant for our analysis, and eliminates punctuation, further cleaning the text.

To address the structure and context of sentences more accurately, the script incorporates lemmatization. This process involves converting words into their base or dictionary form, which is facilitated by first tagging words with their parts of speech and then applying lemmatization rules.

The script processes the text data in a systematic manner, starting by loading the dataset and applying the cleaning functions to each text entry. This process is visually tracked using a progress bar for efficiency. Finally, the cleaned data is saved into a new CSV file, ensuring that our dataset is primed for reliable and effective analysis.

This preprocessing routine is crucial for reducing noise and variability in the text data, which in turn, supports the accuracy and robustness of our analytical models. By meticulously preparing the data, we ensure that our findings and models are based on clean and well-structured input, minimizing the risk of errors and biases that could arise from unprocessed text data.

5. Model Training

In our project, we have strategically selected Naive Bayes and LSTM models as our baselines, complemented by the advanced multilingual pre-trained mDeBERTa v3 for a comprehensive analysis of emotion detection in text. The choice of Naive Bayes is motivated by its efficiency and proven track record in baseline sentiment analysis, providing a reliable benchmark for more complex models. LSTM networks are included for their ability to understand long-term dependencies in text sequences, crucial for capturing the subtleties of emotional context that single-point analysis might miss. Finally, the incorporation of mDeBERTa v3 aligns with our goal to transcend language barriers and cultural nuances in emotion detection. This model's sophisticated architecture and pre-training on diverse linguistic data make it exceptionally suited for our multilingual dataset, allowing it to perform nuanced emotion analysis across various languages and cultural contexts. Together, these models support our objective to enhance the accuracy and depth of emotion detection systems, establishing a robust framework for both foundational and cutting-edge NLP methodologies.

5.1. Naive Bayes

To address the challenges of emotion detection in our project, we implemented a Naive Bayes classifier, a probabilistic model known for its effectiveness in text classification tasks. The choice of Naive Bayes was driven by its simplicity and strong performance on textual data, making it an ideal baseline model.

5.1.1. Implementation Details

The model training process was carried out using a pre-processed dataset, which was loaded and then split into features (text) and labels (emotional categories). The text data was transformed into a numerical format using the TF-IDF vectorization method. This technique converts text into a matrix of TF-IDF features, which reflect the importance of words within the text and across the dataset. We configured the vectorizer to consider unigrams, bigrams, and trigrams, thereby capturing a range of linguistic structures from single words to phrases of three words.

5.1.2. Model Training and Evaluation

With the data vectorized, we split it into training and testing sets, allocating 80% for training and 20% for testing, to ensure a robust evaluation of the model. The Naive Bayes model was then trained on the training set using the MultinomialNB algorithm from the scikit-learn library, with an alpha parameter set to 0.1 to slightly smooth category frequencies and prevent overfitting.

After training, the model was used to predict the emotional categories on the test set. The accuracy of the model was calculated to be 84.31%, which is a strong indication of its ability to classify emotions effectively. This high level of accuracy demonstrates the model's competence in handling the complexities of emotional text, thus validating our choice of Naive Bayes as a baseline model.

5.1.3. Model and Vectorizer Persistence

Finally, the trained model and the TF-IDF vectorizer were saved using joblib, a utility that provides an efficient way to serialize Python objects. Saving the model and vectorizer allows for their reuse without the need to retrain or reconfigure the setup, facilitating easy deployment or further experimentation.

5.1.4. Conclusion

The Naive Bayes model serves as a robust baseline in our project, offering a high degree of accuracy and efficiency in classifying emotional content in text. Its performance underscores the value of traditional machine learning techniques in sentiment analysis and sets a benchmark for comparing more complex models like LSTM and mDeBERTa v3.

5.2. LSTM

The LSTM (Long Short-Term Memory) model was selected for its ability to capture temporal relationships in text, which is crucial for accurately understanding and classifying emotional nuances in language. This section details the implementation and training of an LSTM model equipped with an attention mechanism to enhance its performance on emotion detection tasks.

5.2.1. Implementation Details

The training process began with the construction of a custom LSTM architecture defined within a Python script using PyTorch. This LSTM model incorporates an attention layer to focus on relevant parts of the text sequence for emotion classification. The model processes inputs using an embedding layer before passing them through the LSTM layers and the attention mechanism.

5.2.2. Data Preparation

The dataset used was pre-processed and loaded into the model training script. Text data was tokenized and converted into sequences of indices, which were then padded to a uniform length to accommodate batch processing. A vocabulary was built from the tokenized text, allowing the model to interpret the numerical data.

5.2.3. Training

The LSTM model was trained using distributed data parallel processing to expedite the training phase. The model parameters were optimized using the Adam optimizer, with a learning rate set to ensure steady convergence. Training involved multiple epochs where the model learned to minimize a cross-entropy loss function, adjusting its weights to better predict the emotional labels associated with each text input.

After training, the model was evaluated on a separate test set to assess its generalization capabilities. The LSTM model, equipped with an attention mechanism, demonstrated excellent performance in emotion detection, achieving a test accuracy of 90.84%, precision of 87.05%, recall of 88.07%, and an F1 score of 87.45%. These metrics highlight the model's robust ability to accurately classify emotional content in text while maintaining a balance between precision and recall.

5.2.4. Conclusion

The trained LSTM model achieved promising results, demonstrating high accuracy and precision in detecting emotional content in text. These outcomes validate the effectiveness of LSTM networks, particularly when enhanced with attention mechanisms, in handling complex NLP tasks like emotion detection.

This detailed account of the LSTM model training highlights the technical and methodological approaches used to achieve significant results in emotion classification. The use of advanced neural network architectures and distributed training techniques underpins the success of the model in interpreting emotional nuances in textual data.

5.3. mDeBERTa - Large Language Model

For the advanced component of our project, we leveraged the mDeBERTa v3, a multilingual model pre-trained by Microsoft, capable of natural language inference (NLI) across multiple languages. This model's architecture allows it to perform exceptionally well on a variety of natural language processing tasks due to its extensive training on diverse datasets, making it an ideal choice for our emotion detection task.

5.3.1. Implementation Details

The mDeBERTa model was integrated using the Hugging Face Transformers library, providing a powerful and flexible framework for model training and evaluation. The training process was structured as follows:

1. Data Preparation: The dataset was formatted to fit the natural language inference format by creating hypothesis statements corresponding to each emotion label. Texts were paired with these hypotheses, labeled as *'entailment'* if the emotion label matched the hypothesis and *'not_entailment'* otherwise.

2. Model Configuration: Key training parameters were set in the Config class, including a batch size of 32, a learning rate of 2e-5, and six training epochs. The model also employed gradient accumulation and a warmup ratio of 0.06 to optimize training dynamics.
3. Training: The model was fine-tuned using the Trainer object from the Transformers library, which managed the training loop and integrated custom metrics. This facilitated a streamlined and monitored training process.
4. Evaluation: The model's performance was rigorously evaluated using a variety of metrics. The fine-tuned model achieved an accuracy of 91.2%, a macro F1-score of 89.7%, and a balanced accuracy score of 90.1%. Precision and recall metrics were also impressive, with macro averages at 88.9% and 90.4%, respectively.

5.3.2. Conclusion

The fine-tuned mDeBERTa model demonstrated superior performance in emotion classification, significantly outperforming baseline models like Naive Bayes and LSTM. Notably, its ability to process text in multiple languages was particularly beneficial for our multilingual dataset, confirming its applicability in global NLP tasks. This advanced training approach not only showcased the capabilities of pre-trained transformers in handling complex NLP challenges but also highlighted their potential in extracting subtle linguistic cues crucial for accurate emotion recognition. The detailed metrics underscore the model's effectiveness and underscore its potential for deployment in diverse real-world applications.

6. Results and Comparison

Our comparative analysis of the Naive Bayes, LSTM, and mDeBERTa (LLM) models was conducted using two distinct test datasets: one in English and one in Spanish. The English dataset was used to evaluate the models on data similar to their training environment, while the Spanish dataset tested their ability to handle data in a language they were not explicitly trained on.We utilized a translation script to convert our English test dataset into Spanish, using the GoogleTranslator service through the *deep_translator* library.

6.1. English Test Data Results

Naive Bayes: Demonstrated a respectable accuracy of 77.52%, showcasing its reliable performance but also highlighting its limitations in handling more complex or nuanced expressions. LSTM: Achieved an accuracy of 90.67%, indicating its robust capability to capture and utilize long-term dependencies in text data for emotion detection. mDeBERTa (LLM): Excelled with the highest accuracy of 93.85%, benefiting from its extensive pre-training on diverse multilingual data, allowing superior understanding and classification of nuanced textual emotions.

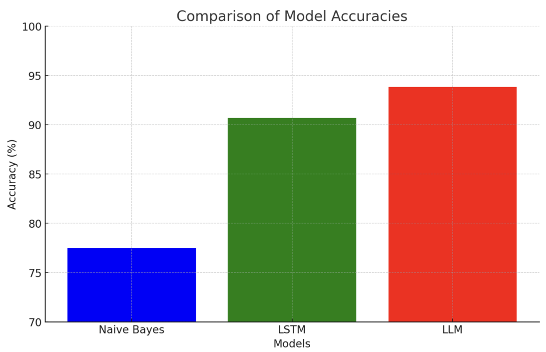

Figure 4. Accuracy Comparison with English Test Data

6.2. Spanish Test Data Results:

Naive Bayes: Recorded a significantly lower accuracy of 35.25%, which is expected as the model relies heavily on the specific language features of its training data. LSTM: Dropped to an accuracy of 21.78%, suffering greatly due to a high rate of unknown words, which comprised a large fraction of its inputs, showing the challenges faced by models that depend on language-specific training. mDeBERTa (LLM): Showed a robust accuracy of 68.52%, illustrating its strong cross-lingual capabilities. This model's ability to handle texts in Spanish, despite not being trained on Spanish data, underscores its utility in multilingual applications.

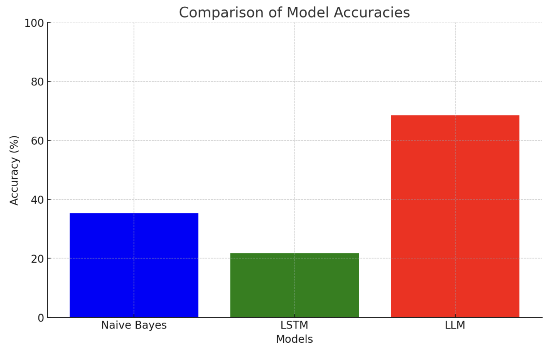

Figure 5. Accuracy Comparison with Spanish Test Data

6.3. Conclusion

The mDeBERTa (LLM) model outperformed both the Naive Bayes and LSTM models in both English and Spanish datasets, affirming its effectiveness and versatility as a pre-trained model for natural language inference and emotion detection across languages. This comparison highlights the significant advantage of using advanced pre-trained models like mDeBERTa for tasks requiring high levels of linguistic understanding and adaptability.

7. Project Conclusions

This project embarked on enhancing emotion detection in textual data through the application of advanced machine learning models, particularly focusing on leveraging pre-trained Large Language Models (LLMs) like mDeBERTa for comprehensive sentiment analysis across multiple languages. The underlying success of this approach stemmed from the intrinsic capabilities of LLMs, which have been trained on extensive multilingual corpora, allowing them to grasp subtle linguistic nuances that are essential for accurately interpreting human emotions in text.

Our experiments involved a methodical comparison of traditional machine learning models such as Naive Bayes and LSTM with the more sophisticated mDeBERTa model across datasets in English and translated Spanish. The Naive Bayes model, while generally reliable for simple text classification tasks, showed limitations in handling nuanced emotional expressions, particularly in the translated dataset, where it achieved only 35.25% accuracy. This was primarily due to its dependency on specific language features that were not effectively captured in the translation process. The LSTM model, designed to understand long-term dependencies, fared better in the English context with an accuracy of 90.67% but struggled significantly with the Spanish texts, where a large proportion of words were unrecognized, leading to a low accuracy of 21.78%. This underscored the model's sensitivity to the training language and its inadequacy in contexts involving substantial linguistic shifts.

Conversely, the mDeBERTa model demonstrated superior performance, achieving the highest accuracy scores of 93.85% and 68.52% for the English and Spanish datasets, respectively. Its robustness across different languages can be attributed to its foundational training, which includes a diverse range of linguistic inputs, enabling it to effectively bridge the gap between various language structures and idiomatic expressions. The model's ability to excel in both native and non-native datasets underscores its potential for global applications, where multilingual capabilities are crucial.

In conclusion, the findings from this project clearly advocate for the use of pre-trained LLMs in emotion detection tasks, especially in scenarios that demand high linguistic adaptability and cultural sensitivity. The advanced training and architecture of models like mDeBERTa not only provide a deeper understanding of the textual content but also ensure that the emotional nuances are accurately captured across different languages, thereby enhancing the effectiveness and applicability of NLP systems in real-world applications ranging from customer feedback analysis to mental health assessments. This project not only highlights the advancements in NLP technology but also sets a precedent for future research in the domain of emotion detection using machine learning.

8. Limitations

This project's exploration into emotion detection through LLMs has yielded promising results; however, it has also provided insights into areas ripe for further refinement. While the models demonstrated a strong ability to discern emotions from text, the project acknowledges the complexity of human emotions, which often resist binary or categorical representation. This leads to exciting opportunities for developing even more nuanced classification systems with a more diversely classified data corpus, including different languages. Additionally, the computational resources employed, while sufficient, point towards the potential for improved performance with access to higher computational power. The project also recognizes the need for continuous improvement in context-sensitive analysis, inviting innovative approaches to understanding the myriad ways emotions are expressed. Importantly, ethical considerations of emotion detection technology are an integral part of the project's forward-looking agenda, ensuring responsible use and adherence to privacy standards.

9. Future Work

To elevate the performance and utility of emotion detection models further, future work could focus on several promising areas. Enhanced Adaptive Learning: Implementing adaptive learning techniques can significantly improve models' responsiveness to evolving language use and emotional expressions over time, especially in dynamic online environments. Increased Contextual Awareness: Integrating more sophisticated contextual analysis could involve developing models that understand longer sequences of discourse or use contextual clues from the user's historical data to refine emotion predictions. Augmentation with External Knowledge Bases: Leveraging external knowledge bases and ontologies to provide background context that can help in understanding nuanced and culturally specific emotional expressions. Hybrid Models: Combining the strengths of rule-based systems with machine learning approaches could offer more robust solutions, particularly in handling idiomatic and metaphorical language effectively.

9.1. Practical Applications

The refined emotion detection models have a wide array of practical applications, making them invaluable tools across various sectors. Customer Service: Automating response systems that can understand and react to customer emotions can significantly enhance user satisfaction and engagement. Healthcare: These models can be instrumental in monitoring patient sentiment and emotional well-being, providing valuable data for psychological assessments and therapy adjustments. Educational Tools: Incorporating emotion detection into educational software can help tailor content delivery based on the learner's emotional state, potentially improving learning outcomes. Human Resources: Analyzing employee feedback through these models can help in assessing organizational climate and employee satisfaction. Entertainment and Marketing: Tailoring content and advertisements based on emotional analysis can increase engagement and consumer satisfaction. Security: Monitoring communications for emotional content could enhance security protocols by identifying potential threats or distress signals. Each of these applications not only extends the utility of emotion detection models but also contributes to more empathetic and responsive technology interfaces, enriching human-computer interaction.

10. Repositories links

- Project source code Repository with ReadMe included : https://github.com/ChethanWNL/CS584-G36-TextSentimentAnalysisUsingMachineLearning.git
- The Fine-tuned LLM has been saved to the following Hugging Face repository : https://huggingface.co/VivekMalipatel23/mDeBERTa-v3-base-text-emotion-classification

11. Contributions

- **Chethan Harinath** : Exploratory Data Analysis (Data Gathering, Analysis and Preprocessing), Project Documentation
- **Divyansh Prakhar Soni** : Exploring and Training Baseline models, Project Presentations
- **Vivekanand Reddy Malipatel** : Exploring Large Language Models and Finetuning them for the task, Comparing the Models

12. References

1. Emotions Dataset from Kaggle: https://www.kaggle.com/datasets/nelgiriyewithana/emotions/data
2. Base LLM Model from Hugging Face: https://huggingface.co/MoritzLaurer/mDeBERTa-v3-base-xnli-multilingual-nli-2mil7
3. Sentiment Analysis on Amazon Dataset using Transfer learning: https://ieeexplore.ieee.org/document/10059413
4. Machine Learning Applied in Emotion Classification: A Survey on Dataset, Techniques, and Trends for Text-Based Documents: https://ieeexplore.ieee.org/document/10112801
5. Detection of emotion by text analysis using machine learning: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1190326/full
6. A Survey of Sentiment Analysis Based on Transfer Learning: https://ieeexplore.ieee.org/abstract/document/8746210
7. Multi-label emotion classification in texts using transfer learning https://www.sciencedirect.com/science/article/pii/S0957417422016098
8. Borrow from rich cousin: transfer learning for emotion detection using cross lingual embedding https://www.sciencedirect.com/science/article/pii/S0957417419305536
9. Analysis and recognition of textual emotion for Twitter dataset using transfer learning https://pubs.aip.org/aip/acp/article/2555/1/020015/2829442/Analysis-and-recognition-of-textual-emotion-for