# Model iteration log

Group members: Celine, Deuza, Emil, Jack
Mentor: Mekselina

| nr. | Model details | | Data details | | | | | Parameters | Metrics *round to 3 decimals!* | | | | | Results & Observations | | | Student(s) responsible | Github link |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| nr. | Model type | Framework/library used | Dataset | Preprocessing steps | Train/Val/Test split | Features used (if applicable) | Data augmentation (if applicable) | Hyperparameters | Accuracy | Precision | Recall | F1 score | Other metrics (if applicable) | Strenghts | Weaknesses | Other notes | | |
| 1 | LR | Scikit-learn, Pandas, NumPy, TfidfVectorizer | preprocessed_dataset.csv | Fill missing values in Corrected_Emotion using Pipeline_Emotion<br>- Label encoding for categorical target variable<br>- TF-IDF vectorization (max_features=5000, stop_words='english')<br>- Standardization using StandardScaler | Train: 80%<br>Test: 20% | TF-IDF text features word embeddings | none | class_weight='balanced'<br>- C=0.5 | 62% | 63% | 62% | 0.62 | | - Simple and interpretable model<br>- Handles class imbalance using class_weight='balanced' | Limited ability to capture deep semantic meanings in text<br>- TF-IDF ignores word order and context | | jack | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/3080cd6890015cbaf0d7855d0ae40879c2340643/Task_5/LRTT1.ipynb |
| 2 | Transformer(Bert) | torch,transformers,sklearn,pandas,numpy | FinalDataset.csv | Load the dataset from 'C:/Users/jack/Downloads/FINAL_DATASET (2).xlsx'.<br>-Drop rows where 'Translation' or 'Corrected_Emotion' is NaN.<br>-Encode the 'Corrected_Emotion' (target variable) using LabelEncoder().<br>-Extract the 'Translation' column as input text and 'Corrected_Emotion' as target labels.<br>-Split the data into training and validation sets (80% training, 20% validation).<br>-Tokenize the text data using DistilBertTokenizer.<br>-Create a custom Dataset class (EmotionDataset) to handle tokenization, padding, and truncation.<br>-Convert text to tokenized tensors with input_ids and attention_mask.<br>-Handle potential NaN values in text by replacing them with an empty string.<br>-Prepare DataLoader objects for training and validation sets. | Train: 80%<br>Test: 20% | Word embeddings (Random) | none | Batch Size: 16 (used in DataLoader)<br>Epochs: 3<br>Learning Rate: 1e-5 (set in AdamW)<br>Optimizer: AdamW<br>Loss Function: CrossEntropyLoss | 72% | 72% | 45% | 51% | | | Constant Learning Rate (1e-5) – No scheduler for adaptive optimization.<br>Small Batch Size (16) – Could lead to unstable training; larger sizes need more memory.<br>No Early Stopping – Risk of overfitting with a fixed 10-epoch setup. | | jack | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/3080cd6890015cbaf0d7855d0ae40879c2340643/Task_5/transformerBertModelITT1.ipynb |
| 3 | RNN | TensorFlow, Keras, Pandas, NumPy | ver_2_FINAL_DATASET.XLSX | Tokenization, Padding, Label Encoding, Sentiment Score Normalization | Train: 80%<br>Test: 20% | Word embeddings, Sentiment Scores | None | Batch size: 32, Learning rate: 0.001, RNN units: 128-64 | 18% | 11% | 18% | 10% | None | Handles Sequential text and sentiment features | Struggles with long range dependencies and potential overfitting | Try a higher learning rate | Celine | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-CelineWu231265/blob/6d4a1a141d5198a5f96a2b56ed754f61bb59bc1a/Sentiment_score_it1.ipynb |
| | BERT-tiny (Transformer) | HuggingFace Transformers, PyTorch | cropped_df.csv | rop NaNs, shuffle, map labels, tokenize text with max_length=128 | 72% train / 8% val / 20% test | | None | learning_rate=5e-5, epochs=1, batch_size=16, weight_decay=0.01 | 68% | 70% | 68% | 69%% | n/a | Very fast training, good initial performance with minimal setup | Undertrained (only 1 epoch), tiny model may not capture complex patterns | Consider training for more epochs, adding test | Deuza | |
| 4 | BiLSTM | TensorFlow/Keras | augmented_dataset.csv | Tokenization, Stopword Removal, Lemmatization, POS tagging, Sentiment Scores, TF-IDF | 70% / 10% / 20% | Word embeddings (GloVe), POS tags, Sentiment scores, TF-IDF, Negation counts | Paraphrasing (Parrot), Back-Translation | Batch size: 32, Learning rate: 1e-4, LSTM units: 64, Dropout: 0.5, Optimizer: Adam | 21% | 21% | 21% | 21% | | Handles sequential data well | Struggles to distinguish emotions, low recall, potential underfitting | Need other methods to balence the dataset | Deuza | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-DeuzaVarela235065/blob/main/task_4/attempt_model4.ipynb |
| 5 | BiLSTM + Embeddings | TensorFlow/Keras | goemotions_1.csv, goemotions_2.csv, goemotions_3.csv | Tokenization, Stopword Removal, Lemmatization, POS tagging, Sentiment Scores, TF-IDF | 70% / 10% / 20% | Pretrained GloVe embeddings, Sentiment Scores, POS Features, TF-IDF, Negation Counts | Paraphrasing (Synonym Replacement), Back-Translation - Atempt | Batch Size: 16, Learning Rate: 1e-4, LSTM units: 128-64-32, Dropout: 0.5, Optimizer: Adam | 31% | 28% | 32% | 27% | | Handles embeddings | Low accuracy, Fails to separate emotions well, Possible underfitting | | Deuza | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-DeuzaVarela235065/blob/main/task_4/fail4.ipynb |
| 6 | BiLSTM + Embeddings | TensorFlow/Keras | goemotions_1.csv, goemotions_2.csv, goemotions_3.csv | Tokenization, Stopword Removal, Lemmatization, POS tagging, Sentiment Scores, TF-IDF | 70% / 10% / 20% | Pretrained GloVe embeddings (300D), Sentiment Scores, TF-IDF, Negation Counts | Paraphrasing (Synonym Replacement), Back-Translation, SMOTE | Batch Size: 16, Learning Rate: 1e-4, LSTM units: 128-64, Dropout: 0.5, Optimizer: Adam | 51% | 53% | 51% | 50% | | Handles embeddings well | Fails to separate emotions well, struggles with "anger" and "surprise", Possible underfitting | refine lr tuning | Deuza | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-DeuzaVarela235065/blob/main/task_4/attempt_model3_1.ipynb |
| 7 | RNN + GRU + GloVe | TensorFlow, Keras, Pandas, NumPy | ver_2_FINAL_DATASET.XLSX | GloVe 100D embeddings, class weighting, sentiment score normalization | Train: 80%<br>Test: 20% | Pretrained GloVe embeddings, Sentiment Scores | None | Batch size: 32, Learning rate: 0.0005, GRU units: 128-64, Class weights applied | 21% | 10% | 21% | 12% | | Uses pretrained embeddings, handles class imbalance better | still struggles with class imbalance, potential overfitting | Use TF-IDF in the next iteration | Celine | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-CelineWu231265/blob/6d4a1a141d5198a5f96a2b56ed754f61bb59bc1a/Sentiment_score_it2.ipynb |
| 8 | Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | MELD dataset | Tokenizing, lowecasing, stemming, removing stop words and punctuation, and stemming | Data split as provided in MELD_train.xlsx (training) and MELD_test.xlsx (testing); no explicit validation set mentioned. | Words extracted and processed from sentences; represented as frequency vectors (Bag of Words) | None | Laplacian Smoothing used in probability calculation. No other hyperparameters adjusted. | 50.88% | 46.38% | 50.88% | 43.83% | none | Simple and fast to implement, effective for small datasets, robust to noise the data | Assumes feature independence which is often not true in real data, can perform poorly with many features due to probability dilution, and overall performs badly. | Model performance can be increased by using NLP features such as TF_IDF. | Emil Fox | (look for iteration number in notebook) |
| 9 | Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | MELD dataset | Tokenizing, lowecasing, stemming, removing stop words and punctuation, and stemming | Data split as provided in MELD_train.xlsx (training) and MELD_test.xlsx (testing); no explicit validation set mentioned. | Words extracted and processed from sentences; represented as frequency vectors (Bag of Words), SMOTE and TF-IDF vectorizer | None | Laplacian Smoothing used in probability calculation. No other hyperparameters adjusted. | 32,45% | 44,76% | 32,45% | 35,73% | none | same ones but performance is worse | same ones but performance is worse | The performance went down after trying to balance the dataset with SMOTE | Emil Fox | (look for iteration number in notebook) |
| 10 | Multinomial Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | Emotions dataset for NLP from a kaggle comp (cropped_df) | Tokenizing, lowecasing, stemming, removing stop words and punctuation, and stemming | Data split as provided in MELD_train.xlsx (training) and MELD_test.xlsx (testing); no explicit validation set mentioned. | Words extracted and processed from sentences; represented as frequency vectors (Bag of Words), SMOTE and CountVectorizer | None | Laplacian Smoothing used in probability calculation using MultinomialNB. No other hyperparameters adjusted. | 76,79% | 83,33% | 58,14% | 61,60% | none | same ones but performance is better | same ones but performance is better | I used a different dataset and the performance scores were better | Emil Fox | (look for iteration number in notebook) |
| 11 | Multinomial Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | Peer made emotion classification dataset | tokenization, conversion to lowercase, and stemming | 70/30 split and no val | token counts of words as features extracted using countvectorizer, bag of words | None | Laplacian Smoothing used in probability calculation using MultinomialNB. No other hyperparameters adjusted. | 64,47% | 31,31% | 22,39% | 22,81% | none | same here with a different dataset | same here with a different dataset | I changed the dataset again to the main one we have to use and the performance was not great on it | Emil Fox | (look for iteration number in notebook) |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | Multinomial Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | Peer made emotion classification dataset with new data and augmented | tokenization, conversion to lowercase, and stemming | 70/30 split and no val | token counts of words as features extracted using countvectorizer, bag of words | Yes there was data augmentation done on the dataset | Laplacian Smoothing used in probability calculation using MultinomialNB. No other hyperparameters adjusted. | 34,52% | 34,02% | 34,45% | 33,91% | none | same here with a modified dataset | same here with a modified dataset | The dataset was the same but just modified, there was more data added and data augmentation was peformed, a slight increase can be seen | Emil Fox | (look for iteration number in notebook) |
| 13 | Multinomial Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | Peer made emotion classification dataset with new data and augmented | tokenization, conversion to lowercase, and stemming | 80/20 split and no val | token counts of words as features extracted using countvectorizer, bag of words | Yes there was data augmentation done on the dataset | Laplacian Smoothing used in probability calculation using MultinomialNB. No other hyperparameters adjusted. | 37,45% | 37,21% | 37,21% | 36,99% | none | same here with a modified dataset | same here with a modified dataset | This time I changed the splitting of the train and test sets and there was a very slight increase | Emil Fox | (look for iteration number in notebook) |
| 14 | Multinomial Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | combined data set of peer augmented data and kaggle emotion dataset | tokenization, conversion to lowercase, and stemming | 80/20 split and no val | token counts of words as features extracted using countvectorizer, bag of words | Yes there was data augmentation done on the dataset | Laplacian Smoothing used in probability calculation using MultinomialNB. No other hyperparameters adjusted. | 86,63% | 80,39% | 80,88% | 80,59% | none | same but with two combined datasets | same but with two combined datasets | This time I combined a new dataset to the current one, and the scores dramatically grew | Emil Fox | (look for iteration number in notebook) |
| 15 | Multinomial Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | dataset of peer augmented data | tokenization, conversion to lowercase, and stemming | 80/20 split and no val | Sentiment score, token counts of words as features extracted using countvectorizer, bag of words | Yes there was data augmentation done on the dataset | Laplacian Smoothing used in probability calculation using MultinomialNB. No other hyperparameters adjusted. | 37,09% | 36,93% | 36,85% | 36,63% | none | the same again with now again only 1 dataset | similar things | I went back to only using the previous dataset and added sentiment score | Emil Fox | (look for iteration number in notebook) |
| 16 | Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | dataset of peer augmented data and online kaggle dataset | tokenization, conversion to lowercase, and stemming | 80/20 split and no val | token counts of words as features extracted using countvectorizer, bag of words | Yes there was data augmentation done on the dataset | Laplacian Smoothing used in probability. No other hyperparameters adjusted. | 68,81% | 65,64% | 68,19% | 64,17% | none | same but with two combined datasets | same but with two combined datasets | This time I corrected the preprocessing of the dataset and how they are combined as it was not done correctly for the last iterations. | Emil Fox | (look for iteration number in notebook) |
| 17 | Naïve Bayes Classifier | Python, NLTK, Pandas, NumPy, scikit-learn | New dataset using the same two from before but its been fully combined and pre processed | tokenization, conversion to lowercase, and stemming | 80/20 split and no val | token counts of words as features extracted using countvectorizer, bag of words | Yes there was data augmentation done on the dataset | Laplacian Smoothing used in probability. No other hyperparameters adjusted. | 80,50% | 80,41% | 80,50% | 80,33% | none | same but with two combined datasets | same but with two combined datasets | This time I manually caclulated the frequencies, probabilities and predictions instead of using multinomialNB | Emil Fox | (look for iteration number in notebook) |
| 18 | RNN + GRU + GloVe | TensorFlow, Keras, Pandas, NumPy, Gensim | ver_2_FINAL_DATASET.XLSX | Added TF-IDF features, MinMaxScaler applied, class weights | Train: 80%<br>Test: 20% | Pretrained GloVe embeddings, Sentiment Scores, TF-IDF features | None | Batch size: 32, Learning rate: 0.0005, GRU units: 128-64, Binary Crossentropy | 22% | 10% | 22% | 12% | none | Used both embeddings and TF-IDF for better representation | More features can introduce redundancy | Try Bi-GRU for improved contextual understanding | Celine | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-CelineWu231265/blob/13084203adcda16c7c5362b44def285f25d995bc/Sentiment_score_TFIFD_it3.ipynb |
| 19 | RNN + Bi-GRU + GloVe | TensorFlow, Keras, Pandas, NumPy, Gensim | ver_2_FINAL_DATASET.XLSX | Changed GRU to Bi-GRU for better context capture, trained embeddings | Train: 80%<br>Test: 20% | Pretrained GloVe embeddings, Sentiment Scores, TF-IDF features | None | Batch size: 64, Learning rate: 0.0001, Bi-GRU units: 128-64, Binary Crossentropy | 34% | 34% | 34% | 33% | none | Bidirectional GRU captures long-term dependencies better | Slower training and potential overfitting | Will add POS tag features for the next iteration | Celine | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-CelineWu231265/blob/6d4a1a141d5198a5f96a2b56ed754f61bb59bc1a/Sentiment_score_TFIFD_it4.ipynb |
| 20 | RNN + Bi-GRU + GloVe | TensorFlow, Keras, Pandas, NumPy, Gensim | ver_2_FINAL_DATASET.XLSX | Added POS tag features, MinMaxScaler applied, Dropout added | Train: 80%<br>Test: 20% | Pretrained GloVe embeddings (50D), Sentiment Scores, TF-IDF, POS Tags | None | Batch size: 64, Learning rate: 0.0001, Bi-GRU units: 128-64-32, Dropout 0.3, Categorical Crossentropy | 31% | 32% | 31% | 31% | none | More structured linguistic features, Dropout improves generalization | Feature complexity can increase the risk of redundancy | Use data augmentation | Celine | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-CelineWu231265/blob/13084203adcda16c7c5362b44def285f25d995bc/rnn_pos_tfidf_sent_it5.keras |
| 21 | RNN | TensorFlow, Keras, Pandas, NumPy, NLTK, NLP-Aug | cropped_df.csv | Applied Synonym Augmentation (WordNet), Balanced dataset using oversampling, Label Encoding, Tokenization, Padding | Train: 80%<br>Test: 20% | Word embeddings (Random), Augmented Data | Synonym replacement using WordNet | Batch size: 32, Learning rate: 0.001, RNN units: 128-128, Dropout 0.5 | 75% | 79% | 74% | 74% | none | Helps address class imbalance, generates synthetic data | Potentially introduces noisy augmented data | Test different augmentation methods (back-translation, paraphrasing) | Celine | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-CelineWu231265/blob/13084203adcda16c7c5362b44def285f25d995bc/RNN_it6.ipynb |
| 22 | LR | Scikit-learn, Pandas, NumPy, TfidfVectorizer | cropped.csv | Fill missing values in Corrected_Emotion using Pipeline_Emotion<br>- Label encoding for categorical target variable<br>- TF-IDF vectorization (max_features=5000, stop_words='english') | Train: 80%<br>Test: 20% | augmented data | Synonym replacement using WordNet | class_weight='balanced'<br>- C=0.5 | 79% | 79% | 79% | 79% | | - Simple and interpretable model<br>- Handles class imbalance using class_weight='balanced' | Limited ability to capture deep semantic meanings in text<br>- TF-IDF ignores word order and | | | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/185af85134a40abf100219c |
| 23 | Transformer(Bert) | | cropped.csv | Load the dataset from 'C:/Users/jack/Downloads/FINAL_DATASET (2).xlsx'.<br>-Drop rows where 'Translation' or 'Corrected_Emotion' is NaN.<br>-Encode the 'Corrected_Emotion' (target variable) using LabelEncoder().<br>-Extract the 'Translation' column as input text and 'Corrected_Emotion' as target labels.<br>-Split the data into training and validation sets (80% training, 20% validation).<br>-Tokenize the text data using DistilBertTokenizer.<br>-Create a custom Dataset class (EmotionDataset) to handle tokenization, padding, and truncation.<br>-Convert text to tokenized tensors with input_ids and attention_mask.<br>-Handle potential NaN values in text by replacing them with an empty string.<br>-Prepare DataLoader objects for training and validation sets. | Train: 80%<br>Test: 20% | Word embeddings (Random), Augmented Data | none | Batch Size: 16 (used in DataLoader)<br>Epochs: 10<br>Learning Rate: 1e-5 (set in AdamW)<br>Optimizer: AdamW<br>Loss Function: CrossEntropyLoss | 87% | 87% | 87% | 87% | | | Constant Learning Rate (1e-5) – No scheduler for adaptive optimization.<br>Small Batch Size (16) – Could lead to unstable training; larger sizes need more memory.<br>No Early Stopping – Risk of overfitting with a fixed 10-epoch setup. | | jack | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/185af85134a40abf100219c5e6a8dbd18185ac08/Task_5/bert2.ipynb |
| 24 | Transformer(RoBERTa) | torch,transformers,sklearn,pandas,numpy<br>torch, transformers, pandas, datasets, sklearn | comb_cropped_df2 (train), group 16 test data (ours) | tokenization using robertatokenizer, truncation and padding added to max, casting labels to classlabel | comb_cropped_df2(train), output of CIA (test) | input text tokenized into input data | none | epochs: 4, batch size: 16, weight decay: 0.01, early stopping with patience 3 | 50% | I forgot to define metric | I forgot to define metric | 48% | Eval loss, train loss | uses a powerful pretrained language model roberta-base | no handling of mislabeled data, no class balancing, no cusom data augmentation | eval strat | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/Roberta.ipynb |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 25 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 16 test data (test) | tokenization using robertatokenizerfast, truncation and padding added to max, casting labels to classlabel | combined_allgroups (train), output of CIA (test) | input text tokenized into input data | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 47% | 66% | 35% | 46% | Eval loss, train loss | uses an even more powerful pretrained language model roberta-large | no handling of mislabeled data, no class balancing, no cusom data augmentation | eval strat | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 26 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | go_emotions_1 (train), group 16 test data | mapped joy to happiness, tokenized with robertatokenizerfast | go_emotions_1 (train), group 16 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 50% | not available | not available | 49% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 27 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 9 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 9 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 70% | not available | not available | 71% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 28 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 25 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 25 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 80% | not available | not available | 74% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 29 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 27 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 27 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 67% | not available | not available | 71% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 29 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 7 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 7 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 69% | not available | not available | 70% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 30 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 23 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 23 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 64% | not available | not available | 65% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 31 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 15 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 15 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 60% | not available | not available | 60% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 32 | Transformer(RoBERTa) | torch, transformers, pandas, datasets, sklearn | combined_allgroups (train), group 28 test data | mapped joy to happiness, tokenized with robertatokenizerfast | combined_allgroups (train), group 28 test data | lexicon features from NRC emotion lexicon | none | epochs: 12, batch size: 16, lr: 3e-5, weight decay: 0.01, early stopping with patience 3 | 74% | not available | not available | 75% | Eval loss, train loss | pretrained language model roberta-large, context aware classification, | no handling of mislabeled data, no class balancing, no cusom data augmentation, limited by quality of emotional data labeling | none | emil | https://github.com/BredaUniversityADSAI/2024-25c-fai2-adsai-group-group16/blob/d7eabccc98d4d21f03928eccf36eec57d222178b/Task_5/RoBERTa/RobertaV2.ipynb |
| 33 | | | | | | | | | | | | | | | | | | |
| 34 | | | | | | | | | | | | | | | | | | |
| 35 | | | | | | | | | | | | | | | | | | |
| 36 | | | | | | | | | | | | | | | | | | |
| 37 | | | | | | | | | | | | | | | | | | |
| 38 | | | | | | | | | | | | | | | | | | |
| 39 | | | | | | | | | | | | | | | | | | |
| 40 | | | | | | | | | | | | | | | | | | |

*you can add more rows to the table if needed*