# Emotion Analysis
# Assignment 2 – ML-based Emotion Classification

Isabelle Mohr    Nishan Chatterjee

Institut für Maschinelle Sprachverarbeitung
Universität Stuttgart

January 10, 2021

# Selecting Datasets

**Comparison of Datasets**

| | **ISEAR** | **Tales** |
|---|---|---|
| Size | 7 665 | 15 302 |
| Granularity | descriptions | sentences |
| Topic | events | fairytales |
| Classes | Plutchik's 8 | noemo + Ekman's 6 |
| Labels | 1 per inst. | 1 per inst. |

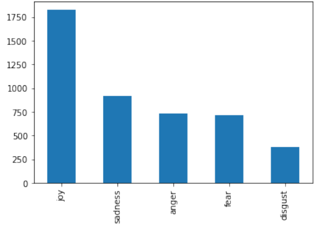
## Motivation

- Datasets of somewhat comparable size
- Granularity at similar level
- Topics unrelated- makes for interesting challenge during cross prediction
- Emotion classes overlap, can take intersection
- Both datasets only have 1 label per instance

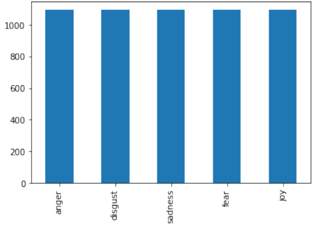
# Data Preprocessing

- The entries labelled "noemo" were dropped
- Dataset after refactoring the classes for interpretability (Example Instance)

| | text | target_emotions | label |
|---|---|---|---|
| 146188 | During the period of falling in love, each time that we met and especially when we had not met for a long time. | joy | 0 |




| target_emotions | label | data_type | text |
|---|---|---|---|
| anger | 0 | train | 622 |
| | | test | 110 |
| disgust | 4 | train | 321 |
| | | test | 57 |
| fear | 2 | train | 605 |
| | | test | 107 |
| joy | 1 | train | 1553 |
| | | test | 274 |
| sadness | 3 | train | 783 |
| | | test | 138 |




| target_emotions | label | data_type | text |
|---|---|---|---|
| anger | 2 | train | 932 |
| | | test | 164 |
| disgust | 4 | train | 931 |
| | | test | 165 |
| fear | 1 | train | 931 |
| | | test | 164 |
| joy | 0 | train | 930 |
| | | test | 164 |
| sadness | 3 | train | 931 |
| | | test | 165 |

Figure: Class Distribution after making both Tales-Emotions (Left two) and ISEAR(Right two) comparable

# Model: Feature-based

## Model Description

- kNN with $k = 5$
- Designed set of features
- All words in instance looked up in NRC Emotion Lexicon to add count to feature vector
- Also considered tense, negation, punctuation, capitals and rate of repetition (see memo for motivation)
- Scaled feature vectors to minimize distribution of 0's
- Example instance shows feature vector for instance: *"I LOVE the sun on my smooth face."*

## Example instance

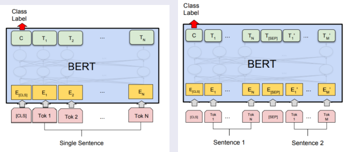
| | | |
|---|---|---|
| valence | pos | 2 |
| | neg | 0 |
| emotion | joy | 2 |
| | ... | ... |
| | anticipation | 0 |
| | disgust | 0 |
| punctuation count | . | 1 |
| | , | 0 |
| | ... | ... |
| | ! | 0 |
| | ? | 0 |
| tense | future | 0 |
| | present | 1 |
| | past | 0 |
| negation | negation count | 0 |
| capitals | capital count | 5 |
| repitition rate | repetition rate | 1 |

# Model: Base-Bert-Uncased

## Model Description

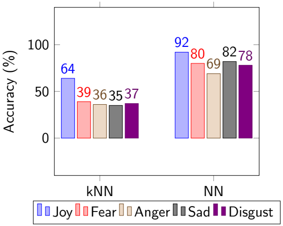
- Bidirectional encoder representations are better than embeddings using a Mask Language Model (MLM) and Next Sentence Prediction (NSP)
- Fine-tuning done with the model on our pre-processed data
- Implemented the Adam optimization algorithm
- Used a scheduled learning rate which decreases linearly during warmup and increases linearly after the warmup period

## Transformer Architecture

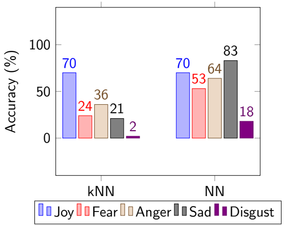



# Results: Accuracy Scores on the 5 Emotion Classes

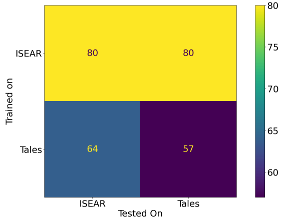
## ISEAR-ISEAR



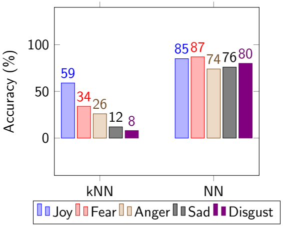

## Tales-Tales



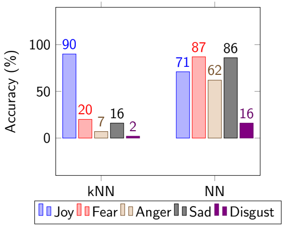

## Base-Bert Results



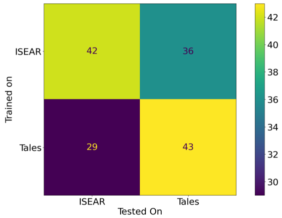

## ISEAR-Tales




## Tales-ISEAR




## kNN Results



# Discussion

- The choice for the datasets were based on easy class comparability
- Tales-Blogs were the one we wanted to compare more given its single peak across the 4 data-points from the Confusion Matrix
- The results are computed from the model state after the first epoch for better comparability across the two different approaches
- In future: design other features (tf.idf) for feature vector
- The Bert Implementation was also not run for long due to lack of computation resources and more training is expected to yield better results
- Tales-Tales lower performance in comparison could be because of:
    - Class imbalance as ISEAR is much more normally distributed
    - An odd random starting point which may average out over many models (see memo)

# ?

## Questions? Suggestions?

Look at our work: Emotion Analysis Project