Exploring effective translation between affect representation schemes

Author: Mira Ilieva Email: m.h.Ilieva@student.tudelft.nl
Supervisors: Bernd Dudzik, Chirag Raman
Affiliates: EEMCS, Delft University of Technology

Text Content Analysis

01 Background

- **Affect**: the outward and inward experience of feeling, emotion, attachment, or mood.
- **Representation Scheme:** objective ways to describe emotions in a systematic manner[1].

02 Research Questions

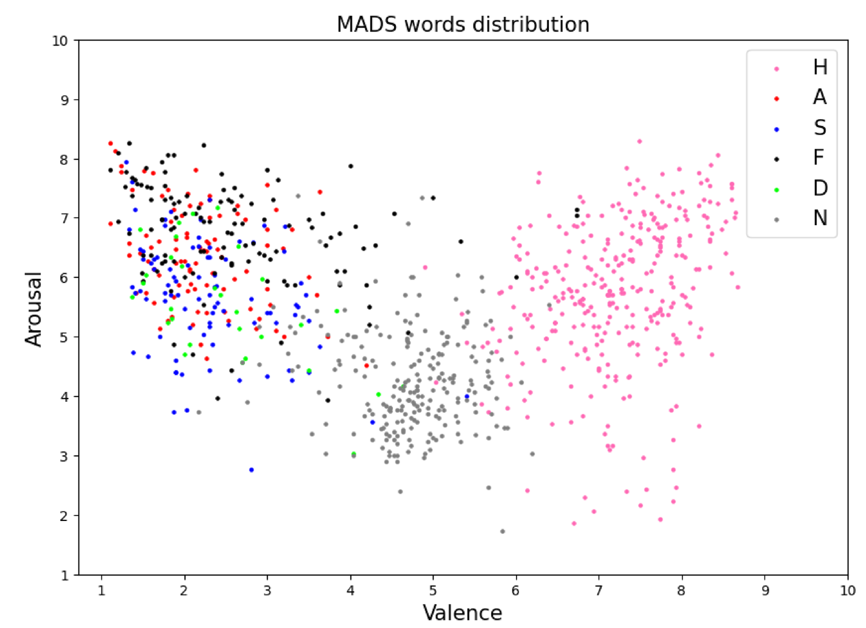
- Is it feasible to translate between representation schemes?
- Does the translation model generalise between datasets (different languages and cultures)?
- Is cross-gender translation feasible?

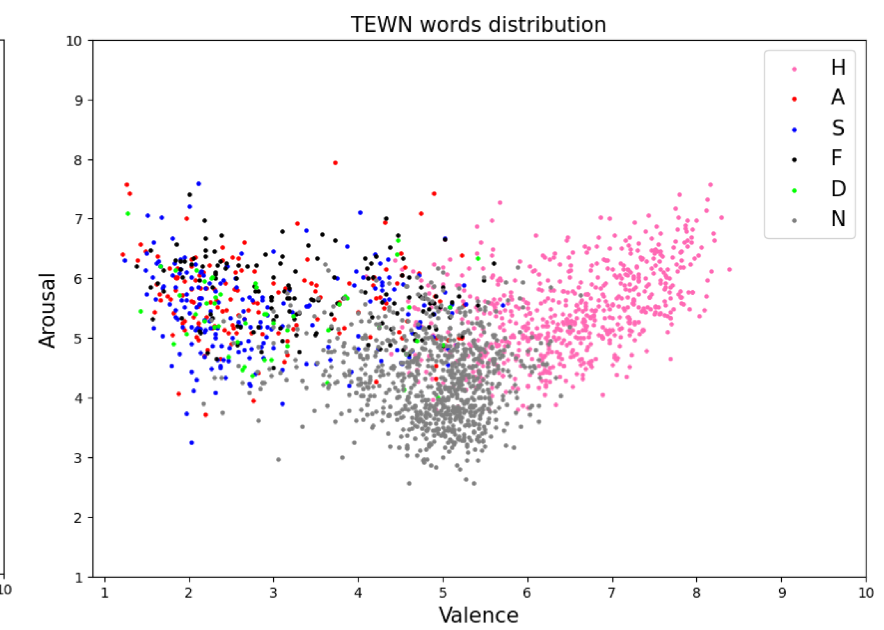
03 Datasets and Labels

- Categorical emotions representation:
  - Happiness
  - Anger
  - Sadness
  - Fear
  - Disgust
  - Neutral (added in the pre-processing)
- Dimensional emotions representation:
  - Valence: (negative - positive)
  - Arousal: (calmness - excitement)
- The two datasets we use:
  - MADS [2]
  - TEWN [3]

	Categorical	Dimensional	Per gender info
MADS	✓	✓	✓
TEWN	✓	✓	-

	# words	# raters	# of women	# of men
MADS	875	660	507	153
TEWN	2031	1527	952	757



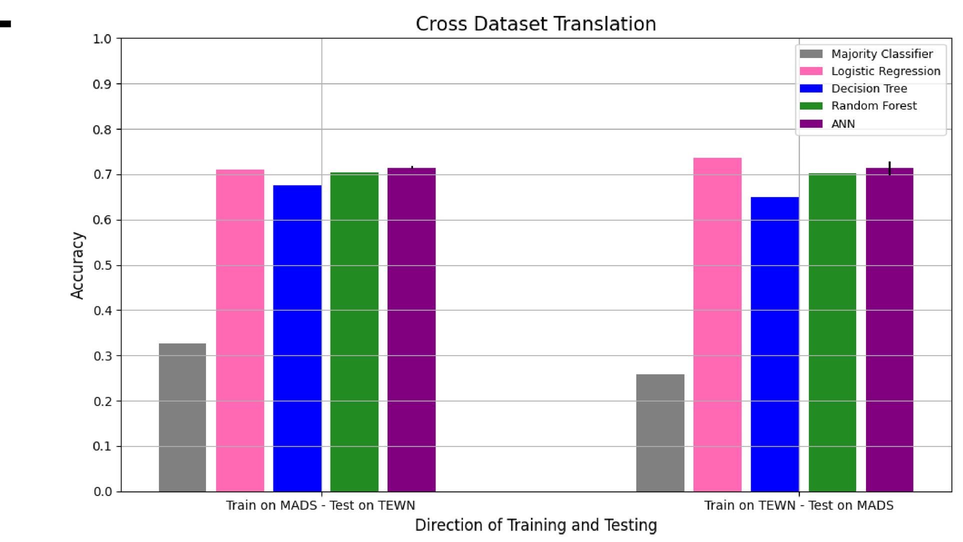


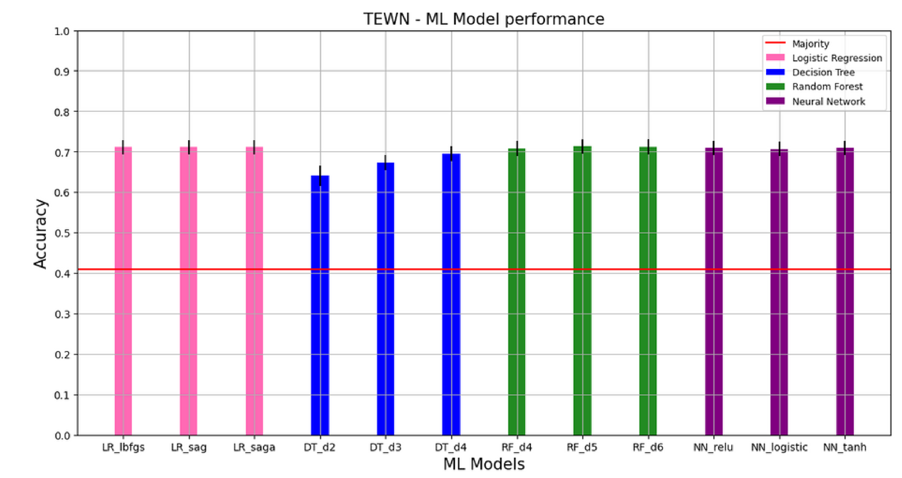
04 Methodology

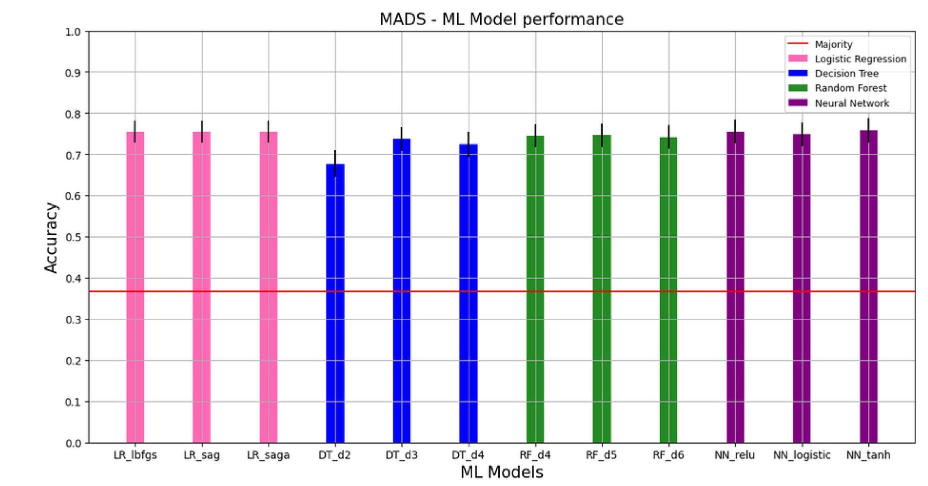
Singular dataset translation: All classifiers were executed concurrently, along with the Majority classifier, thereby ensuring consistency in data partitioning and reducing any potential biased influence on classifier performance.
Cross-dataset translation: In this study, the model was trained on the MADS dataset and subsequently tested on the TEWN dataset (as well as in the other direction).
Cross-gender translation: Update the labels (make new ones for men and women). We then proceed with training all the models on the male labels and testing on the female, while splitting the data in 80:20 and repeating 100 times in order to increase reliability. Similarly, the approach is mirrored for training on women ratings and testing on men.

05 Results/Findings

- In the case of the single dataset translations: it is possible to achieve around 75% accuracy, way over the Majority classifier.
- In the case of the cross-data we see a slight decrease to around 70% accuracy, however still .
- In the cross-gender we see that initial label overlap is around 75%, however, the accuracy of the ML model is lowered to 60%.







Word	Labelled by women	Labelled by men
to admire	D	H
to harass	F	A
to threaten	F	A
envious	D	A
to infect	D	F
embarrassment	A	F

06 Conclusion & Limitations

- The study found that the emotion labels 'H' and 'N' consistently achieved high accuracy scores due to their frequent occurrences and distinct mean Valence and Arousal values.
- Differences in Valence and Arousal values between datasets may be due to cultural disparities, but the machine learning models generalized well.
- Cross-gender translation achieved accuracies above 0.6 by aligning variations with the Dimensional Representation scheme.
- Translation between Affect Representation Schemes is feasible for distinguishing between negative, neutral, and positive words, but more work is needed for a robust framework.
- Some of the limitations are the lack of individual ratings and the limited overlap words rated in both datasets.

07 References

1. Alswaidan, N., Menai, M.E.B. A survey of state-of-the-art approaches for emotion recognition in text. Knowl Inf Syst 62, 2937–2987 (2020). https://doi.org/10.1007/s10115-020-01449-0
2. Hinojosa, J. A., Martínez-García, N., Villalba-García, C., Fernández-Folgueiras, U., Sánchez-Carmona, A., Pozo, M. A., & Montoro, P. R. (2015). Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. Behavior Research Methods, 48(1), 272–284. https://doi.org/10.3758/s13428-015-0572-5
3. Kapucu, A., Kılıç, A., Özkılıç, Y., & Sarıbaz, B. (2018). Turkish emotional word norms for arousal, Valence, and discrete emotion categories. Psychological Reports, 124(1), 188–209. https://doi.org/10.1177/0033294118814722