RorschIA : NLP and help with coding

Rachel ATHERLY, Alberto LORENTE GALE, Aubin MEDJAED, Mina OULHEN - Supervision : **Renaud EVRARD** (MCF-HDR)

other project members : Mélanie LAURENT (MCF), Antoine FRIGAUX (Psychologist, Associate Researcher), Côme PIERRE (Psychologist), Eva RAMORINO (student M2 psychology), Yassine BENAÏD (computer science)

1 Introduction

- The Rorschach inkblot test is a century old well known procedure to assess the personnality of a subject.
- Yet, the whole procedure for the test can take up to 4 hours because of a lengthy scoring process.
- Our project aim to apply the BERT transformer model to automatize the scoring process and significantly reduce the time required for it's completion.

1 Objective

Our mains objectives are :
- Conceptualizing the test through a NLP and Machine Learning lense.
- Applying classic NLP trends to develop machine learning models classifying the determinants and contents for each response to the test in a environmentally friendly manner.
- Assessing the applicability of those models and evaluate them.

2 Data

- Data from Rorschach related sources is scarce, often very inbalanced.
- Data used for this project was given anonymously by the psychology part of the RorschIA team. Responces and labels were uniformized.
- Labels were eventually grouped for the Macro models.
- Number of rows : 380. Train : 325, Validation : 36, Test : 19

3 Methodology

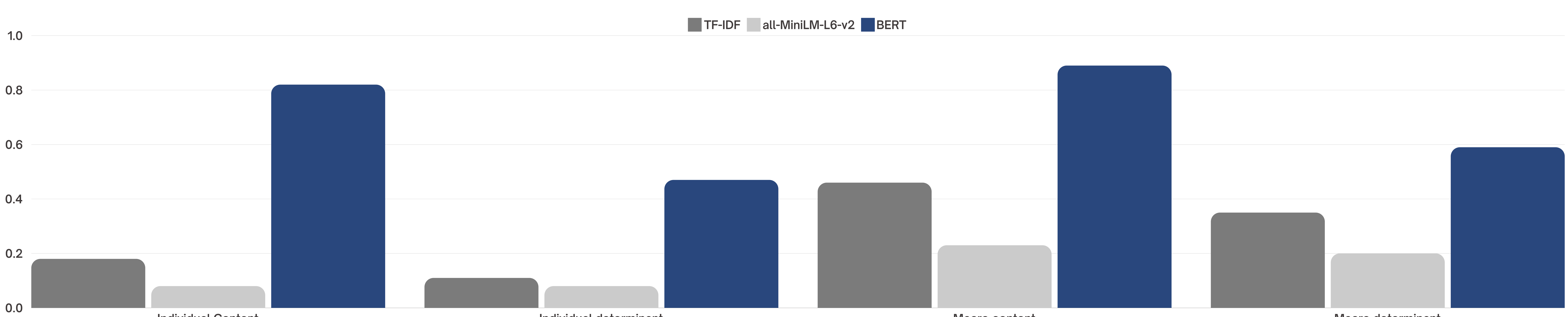
- Because of the issues found in the data, we created 4 bert-base-uncase models, two using the canonical determinants and contents, the other two using macro-grouped determinants and contents.
- Hyperparameters where fine-tuned by grid-search.

4 Results

- The performance of the models is dependant of the label classified.
- Models with grouped labels are consistently better for most metrics, showing a 7 to 12 % increase on f1 score.
- Results on macro and individual content show inpressive performances.
- Results related to determinents are mild.

Model	F1	Accuracy	Precision	Recall	Hamming loss
Individual content	0.82	0.74	0.80	0.83	0.01
Individual determinant	0.47	0.40	0.50	0.43	0.06
Macro content	0.89	0.86	0.91	0.87	0.021
Macro determinants	0.59	0.44	0.62	0.57	0.20

Performance of BERT models in the Test Split



Performance of alll models (with intial machine learning models) in the Test Split

6 Conclusion

- Approching the coding of the Rorscharch test as a multi-label classification problem, with grouping of the labels seems to fetch the best results, but at the lost of the more-precised, already in use labels.
- Traditionnal machine learning algorithms were used to assess the potential of the method. Then, we developed BERT transformer-based models for better results.
- Macro labels averaged performance is 9% higher for f1 score.
- Content models (both Individual and Macro) performances are very promising.

- Usage recommandation for potential practical testing and usage would be in assistance to a professional and not fully relying on the raw predictions output by the models.

6 Future Work

- Building a corpus project to collect anonymized Rorschach protocols and consistent annotation guidelines to tackle ambiguity and subjectivity.
- Potential merging of the two approaches by assigning first to a number of macro-categories, then depending on the output, return a canonical label for each macro-categories detected.
- Using better performing models like the newly presented GPT-4o, to notably reduce hallucinations.

Related literature

Ricardo Stegh Camati, Alessandro Antonio Scaduto, and Fabr´ıcio Enembreck. 2021.
Using the projective thematic apperception test for automatic personality recognition in texts. 2021 IEEE International Con- ference on Systems, Man, and Cybernetics (SMC), pages 78–85

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019.
Bert: Pre-training of deep bidirectional transformers for language understanding.

Michael Tanana, Kevin Hallgren, Zac Imel, David Atkins, Padhraic Smyth, and Vivek Srikumar. 2015.
Recursive neural networks for coding therapist and patient behavior in motivational interviewing, pages 71–79.

UNIVERSITÉ DE LORRAINE
Institut des sciences du Digital Management & Cognition
COMPOSANTE DE L'UNIVERSITÉ DE LORRAINE