

RorschIA

Institut des Sciences du Digital, Management et Cognition - Université de Lorraine

May 30, 2024

Rachel Atherly, Alberto Lorente Galé, Aubin Medjaed and Mina Oulhen



Table of Contents

- 1 Introduction
- 2 Data
- 3 Methodology
- 4 Results
- 5 Practical Example
- 6 Emissions and Energy Consumption
- 7 Future Work

Introduction

Introduction

Supervised by Renaud EVRARD.

Interdisciplinary project (Psychology and NLP).

Goal : Automatize the encoding part of the Rorschach inklot test procedure.



Other participants : Mélanie LAURENT (MCF), Antoine FRIGAUX (Psychologist, Associate Researcher), Côme PIERRE (Psychologist), Eva RAMORINO (student M2 psychology). +Yassine BENAID (computer science)

Rorschach test

Rorschach inkblot test is 100 years old !



Rorschach test

The test :

- Gather the answers
- **Scoring of the answers**
- Calculate the psychogram



Figure: all 10 cards in order

Rorschach test scoring

Scoring use 4 labels :

- **Content**
- **Determinants**
- Location
- Quality and other notes

Objectives

Project objectives :

- Verify lack of literature connecting Rorschach and NLP
- Collect info on State of the Art
- Conceptualize the Test through NLP lens
- Define challenges and strategies to overcome them
- Apply trends to develop ML models to classify content and determinants
- Conduct the model training in an environmentaly friendly way
- Assess applicability of models

Literature and related works

No major research conducted between Machine Learning and Rorschach test.

Camati, Scaduto, and Enembreck, 2021. Machine learning models (TF-IDF) on Thematic Apperception Test, rated by Revised Morvalian System (RMS).

Interest : NLP team using the test to evaluate their models (Deep Dream).

Data

Data - 1st dataset

COVID-19 Rorschach test dataset (2017-2020)

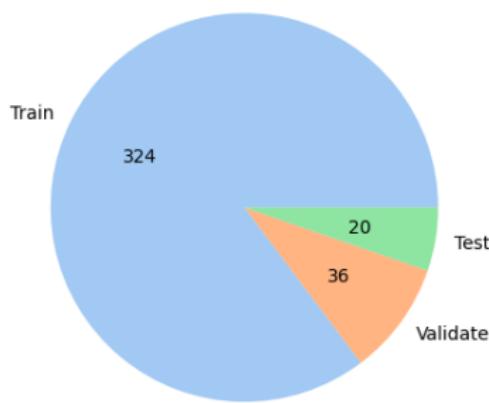
747 responses (after selection of the most qualitative)

- Inconsistent annotations
- Poor results when applied to the models
- Eventually abandoned

Data - 2nd dataset

10 protocols from the psychology department = 380 responses
More qualitative but rather low amount

Data splits for Model development



Data - 2nd dataset

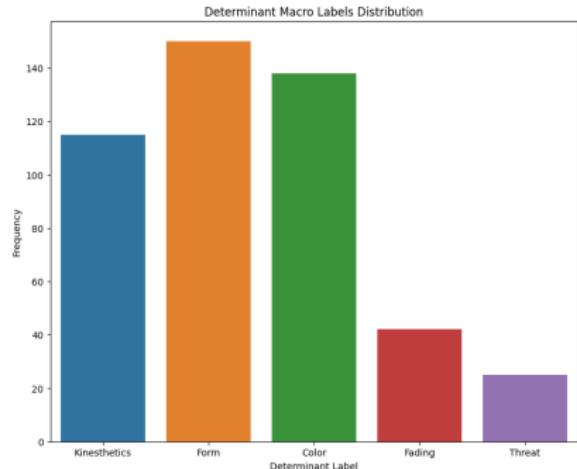
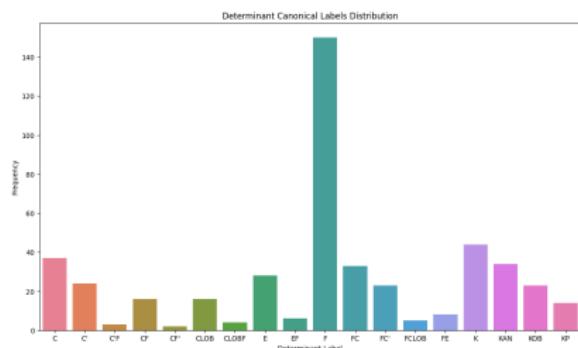
Data cleaning:

- Translation of the protocols into English (DeepL translator)
- Standardisation of content and determinant labels
- One hot encoding

Data - 2nd dataset

Uneven representativeness of the original classes

Second approach: regrouping of the labels into macro-labels



Methodology

BERT - well known transformer model

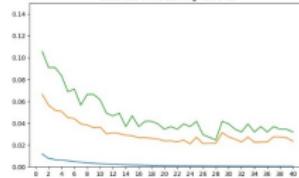
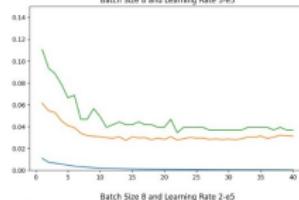
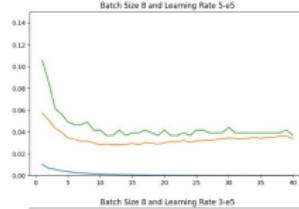
Grid search for hyperparameter tuning (batch size, learning rate, training epochs)

Evaluation Metrics: f1 scores, accuracy, training loss, validation loss, and the hamming score

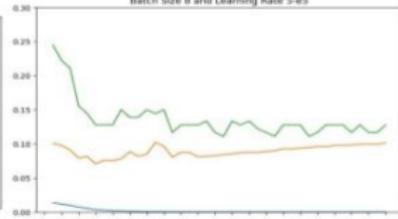
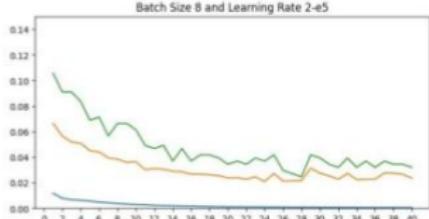
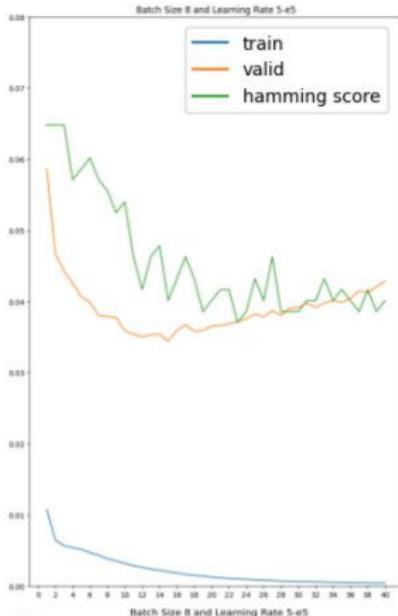
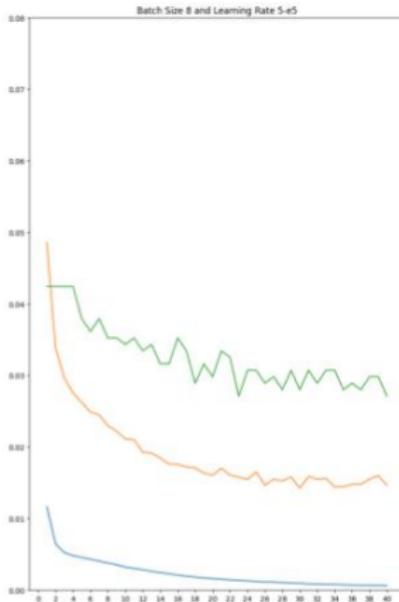
Single data representation vs different parameters per model

Evaluation

- Selection of final candidate models
- Final epoch selection
- Final model graphs
- Hamming score



Final Model Candidates



Final Model

Data Model	Learning rate	Batch size	N. epochs
Canon. Determinants	3e-5	8	15
Canon. Contents	3e-5	8	32
Macro Contents	2e-5	8	28
Macro Determinants	5e-5	8	8

Table: Final Models

Results



Results

Table: Final Results per Model

Model	F1	Accuracy	Precision	Recall	Hamming Loss*
C. Content	0.82	0.74	0.80	0.83	0.01
M. Content	0.89	0.86	0.91	0.87	0.02
C. Determinant	0.47	0.40	0.50	0.43	0.06
M. Determinant	0.59	0.44	0.62	0.57	0.20

* Hamming Loss considerations: effect of number of predictions on the total score.

Trade-offs between the two approaches

- Content representations: from 27 to 11
- Determinant representations: from 18 to 5

Trade-offs between the two approaches

- Content representations: from 27 to 11
- Determinant representations: from 18 to 5

Lower number of labels led to better results. BUT the outcome of the models is less interpretable by the psychologist.

Trade-offs between the two approaches

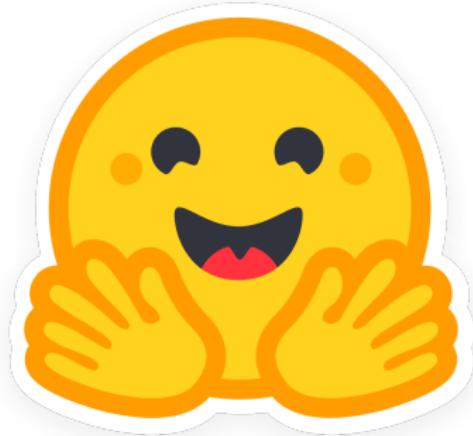
Practical solution? Use the models in tandem!



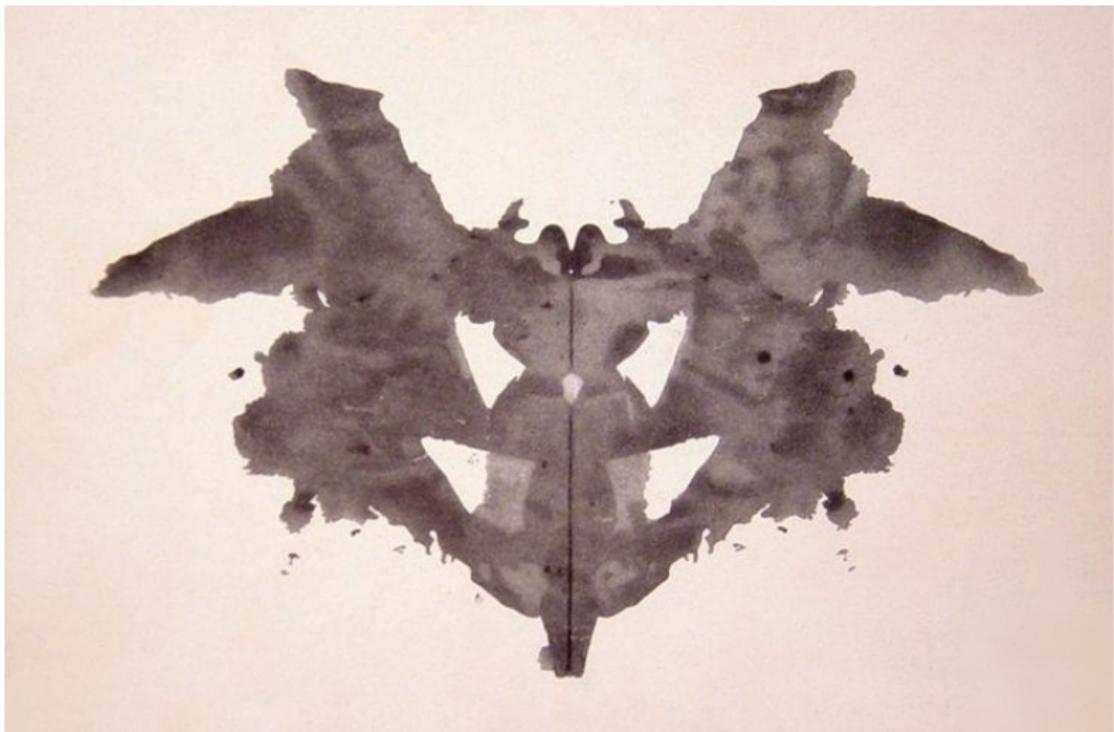
Practical Example

RorschIA App

RorschIA App hosted in



Example Blot 1



Example

Maybe a biblically accurate angel?

Example

Ezekiel 1:15-17:

"For wherever the head faced, the cherubim would go in that direction, without turning as they moved. Their entire bodies, including their backs, hands, and wings, were full of eyes all around, as were their four wheels."

Example

With the help of the Rorschach scoring cheatsheet provided by the InterPsy team we would probably code the answer as:

- Determinants: kan or k and F.
- Contents: (H)/(Ad) and Anat.

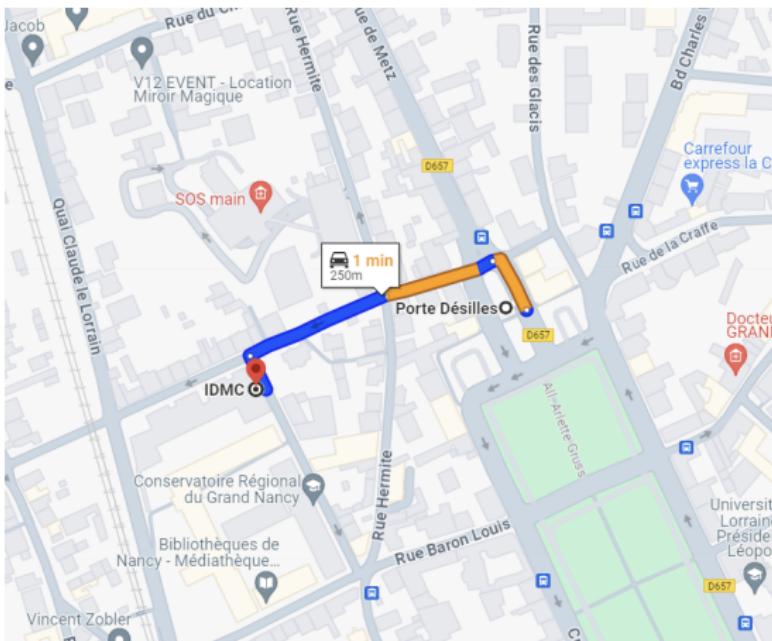
What would our models do?

Emissions and Energy Consumption

Emissions and Energy Consumption

According to Green Algorithms Calculator (Lannelongue, Grealey, and Inouye, 2021):

- Energy consumption: 1.01kWh
- Carbon footprint: 51.88gCO₂



Future Work

Future Work

- Combining both approaches
- Corpus creation
- Alternative approaches to the problem

Conclusion

Thank you!

Bibliography I



Camati, Ricardo Stegh, Alessandro Antonio Scaduto, and Fabrício Enembreck (2021). "Using the Projective Thematic Apperception Test for Automatic Personality Recognition in Texts". In: *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 78–85. URL:

<https://api.semanticscholar.org/CorpusID:252590374>.



Lannelongue, Loïc, Jason Grealey, and Michael Inouye (2021). "Green Algorithms: Quantifying the Carbon Footprint of Computation". In: *Advanced Science* 8.12, p. 2100707. DOI:

<https://doi.org/10.1002/advs.202100707>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202100707>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202100707>.