Politecnico di Torino
Dipartimento di Automatica e Informatica
1859

DAUIN

PhD in Computer and Control Engineering
XXXV cycle

Supervisor
Maurizio Morisio

Natural Language Processing in Personality Traits and Basic Human Values Estimation of Social Media Users

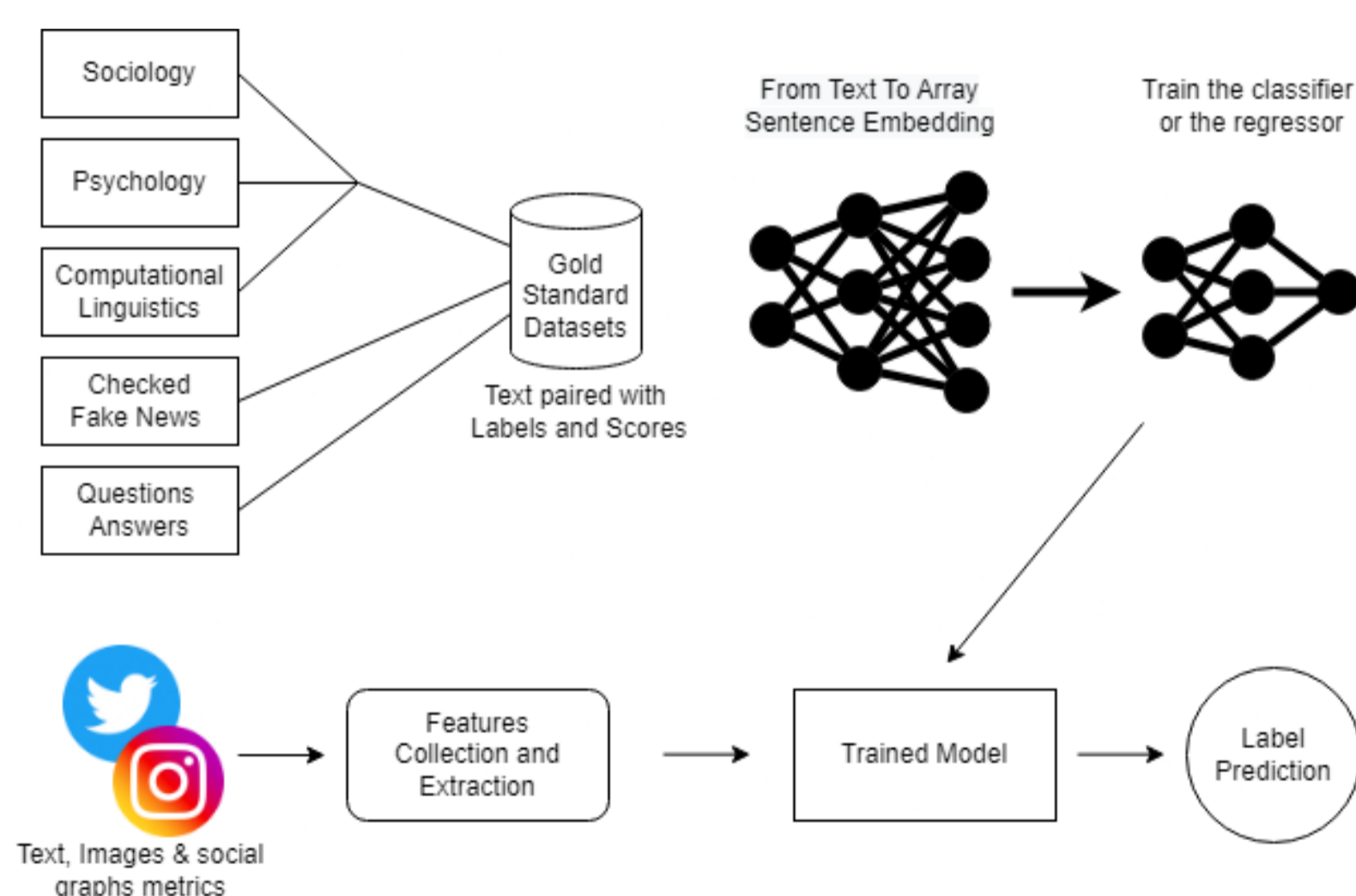PhD Candidate: *Simone Leonardi*

1. Context

Chatbots are not so smart, but even when they are able to do what you ask, their cold emotionless interaction makes you desire to stop the conversation as soon as possible. The data abundance in social media and the outstanding progress in natural language processing created the premises for more empathic A.I. dialogues.

2. Goals

Exploring some of the features of human empathy and understanding how to retrieve them from written text. Training deep neural networks to predict empathy related scores. Finally, applying these findings in real use case scenarios: influence mechanisms, fake news, education.

3. Method

Sociology, Psychology, Computational Linguistics, Checked Fake News, Questions Answers → Gold Standard Datasets — Text paired with Labels and Scores → From Text To Array Sentence Embedding → Train the classifier or the regressor

Text, Images & social graphs metrics → Features Collection and Extraction → Trained Model → Label Prediction

The field expert produced gold standard datasets. They make participants answer psychological questionnaires and collect their writings. First, I translate datasets into embedding with stacked NLP models, then I train a deep neural network to predict desired labels and scores.

The trained models are now able to predict this information receiving in input new samples. I collect data from public social media API.

4. Results

- State of the art NLP model to predict personality traits from social media posts.[1]
- Framework to detect and rank micro-influencers. It supports marketing agencies and public institutions to find effective ambassadors of their values and products.
- New NLP strategy to compute basic human values from text. There were no public alternatives at the time of my publication.
- Framework to fight fake news infodemic.[2] It reads social media posts and perceives linguistic markers of potential fake news spreaders.
- Educational chatbot to support teaching activities on Slack. It collects data during conversations to expand its knowledge base.

5. Conclusions

The results confirm that the use of NLP allows to mine useful knowledge from social media data in the context of empathy and behavior analysis.

6. References

1. Multilingual Transformer-Based Personality Traits Estimation, Leonardi et al. (2020) MDPI – Information.
2. Automated Classification of Fake News Spreaders to Break the Misinformation Chain, Leonardi et al. (2021) MDPI – Information.