

Weekly Wrap-up

Progress Highlights and Insights



Contents

Thesis Writing

Alignment

What's next

Thesis Writing

- Added SHAP and Lime in literature review chapter
- Finished methodology, implementation and results chapters

Alignment

- Main goal
- Phases
- Final output





Main Goal

- The main goal of this project was to
 - Predict entrepreneurial personality through text.
 - Explain Prediction
 - With Higher focus on the explanation output.

Phases

- The project can be seen as 3 phase process



Phase 1

- To build a classifier to predict the entrepreneurial personality we will
 - Build personality traits extraction model
 - Use the traits extracted a way to identify entrepreneurs.
 - The traits can serve as both a way to classify the entrepreneurs and non entrepreneurs and a way of explaining the decision
 - Example :
 - Ali is an Entrepreneur cause he is INTJ and INTJ is an entrepreneurial personality type.



Phase 2

- The second phase is about explaining the model build for trait extraction through LIME or SHAP
- As we observed last week LIME and SHAP were able to produce similar results



Phase 3

- Apply the Personality trait on a number of entrepreneurs
 - Extract the entrepreneurs personalities
 - Build a set of personality types that match the entrepreneurial personality
 - Return an output with a prediction and two pieces of explanation
 - The personality type
 - Why the text fits this personality types



Example Output

- Example:
 - Ali is an entrepreneur cause his personality is **INTJ**, He is I(introvert) cause his text has words such as **alone**, **lonely**, he is a f(feeling type of person) cause his text mentions words like **feel**, **compassionate**, etc.



Concept Transformer

- Although the concept transformer architecture could not be used
- It inspired the whole idea where we use the personality traits (concepts) to explain the prediction
- On top of that we added a layer of more basic explanation using LIME or SHAP

What's next

- Working on Phase 3
- Suggestions





Working on Phase 3

- We applied the trait extraction on 2 entrepreneurs only so we only have INTJ in our set of personality types.
- Should look for more Entrepreneurs
 - Maybe non-famous ones
- Collect tweets, Or use tweets collected already.



Suggestions

- We are all on the same line
- Open to any suggestions or additions!

References

- A Unified Approach to Interpreting Model Predictions
- Interpretation of multi-label classification models using shapley values
- <https://shap.readthedocs.io/en/latest/>
- <https://christophm.github.io/interpretable-ml-book/shapley.html>
- <https://lime-ml.readthedocs.io/en/latest/lime.html>



Thank You

