Intoduction – practical Motivation (Use Case) and theoretical Motivation (Research Gap), Research Question, Structural overview of thesis

Theoretical Foundation –

-> Create search funnel (either for thesis or for defense) Ein Bild, das Tisch enthält.

Automatisch generierte Beschreibung

-> Create table to compare papers

* Concepts that I need later on (some concepts can be assumed to be known (bachelor level knowledge))
  + Relation Extraction, Named Entity Recognition
  + Knowledge graphs
  + Transformers -> what is a decoder?
  + Transfer Learning
  + REBEL & Seq2Seq

FROM:

* HoWeR
* Scholar
* Web of Science
* IEEE
* ACM

Related Work

* Points of agreement / disagreement, unused problems, under-researched problems (e.g., most use same datasets for RE), conflicting results, unanswered questions
* What have others done in the direction of the research question
  + Relation Extraction on political data
  + Other Research that finetuned rebel?

1. Introduction

a. Motivation from a business perspective

b. Research gap (short)

c. Research question(s) to be answered

d. Structure of the thesis

2. Theoretical background

* Relation Extraction
* Definition
* Classical Approaches
  + Structure of the approaches
    - Named entity recognition (& Part -of speech recognition)
    - Coreference Resolution
    - Pipelines vs end-2-end
  + Knowledge Graphs
  + CNN / RNN
* Transformers
  + Definition (maybe include graph of performance increase on RE and NER due to transformers)
  + Architecture (Encoder / Decoder)
  + Seq2Seq Transformers and other relevant type of transformers
  + Types of transformers for RE

3. Related Work

* Related work on specific use case (Political use cases for transformers & PETRARCH)
* REBEL (also rough intro BART (vs BERT))
* Graph elements
* Entity hinting, entity masking
* Transformer & seq2seq very data hungry -> often additional data by using soft-labeled data (e.g., Sutskever, Vinyals, Le; Rebel)
  + REBEL
  + Related work on “soft labelled” datasets (“Distantly supervision”)

a. Introduce relevant theoretical concepts that you need for your methodology and analysis (e.g. performance measure, BERT)

b. Discuss related work in detail (i.e. others that did something similar) and what they didn’t do (i.e. the research gap)

3. Related Work

3. Methodology/Approach

a. Data (optionally as a separate chapter): describe your raw data and its source, followed by EDA and data preparation steps -> also explain annotation, relation labels (problem with some petrarch etc; problem with some relation labels being close to another)

b. Model with name: Describe your approach/model in detail, use scientific sources, explain your modelling decisions (e.g. why 2 epochs, why 10,000 data points)

4. Implementation (optional, can also be discussed in 3 b.: Provide an overview of the used software and hardware and the structure of the submitted code

5. Evaluation/Results

a. Discuss the application of your model to the data

b. Add a summary of the results and their implications

6. Conclusion