**Plan of work (appended to thesis contract)**

Name of student Hisham Alkaed

Degree programme Text Mining M.Sc.

Student number 2671370

Name of supervisor(s)/ Dr. Ilia Markov

Chair VU Amsterdam, Faculty of Humanities, CLTL lab, Assistant Professor

Second reader/ ......................................................................................

Chair ….. ………….………………………………………………………………….

Working title

Topic Hateful meme detection. The Hateful Meme Detection challenge, created by Facebook, presents a unique problem in the field of text mining and image recognition. This problem requires advanced algorithms and techniques to analyze and extract meaningful information from multimodal features, effectively combining image and text data to identify and classify hateful content in memes. Within this context how could linguistic cues affect the performance of the ML models.

Aim and relevance ......................................................................................

…………...........................................................................

Problem definition Research question: How do linguistic cues, such as the use of specific pronouns or adjectives, contribute to the identification of hate speech in memes, and what are the limitations of using these features for automated detection?

Sub-questions:

1) late fusion v early fusion

2) Can linguistic analysis of hate speech in memes provide insights into the underlying social and cultural factors driving the spread of hate speech online?

3) What are the limitations of using linguistic cues for automated detection of hate speech in memes? How accurate are these automated systems in identifying hate speech, and what factors influence their effectiveness?

4) How much can linguistic features of hate speech in memes provide insights into the underlying social and cultural factors driving the spread of hate speech online?

Data-collection/ The dataset is already available at <https://www.kaggle.com/datasets/williamberrios/hateful-memes>. The dataset has been constructed by a professional team and the quality of it has been checked and refined to make it very interesting to the task at hand.

research method During the project the following methods will be compared and explored:

1. Late fusion of linguistic features against early fusion.
2. Linguistic features such as:
   1. gender detection
   2. hateful word per sentence rate
   3. adjective usage
   4. laughter
   5. sentiment analysis
   6. intent detection
   7. (maybe) named entity recognition: Person, Land, and ethnicity since those could possibily be referring to some kind of hateful detecetion
   8. (maybe) Content-to-function-word ratio
   9. (maybe) word embeddings

First two baseline models will be implemented or taken from literature using a straight forward algorithm. One for late fusion and one for early fusion.

Then the same algorithms will be adapted to include the presented features.

The models will be evaluated to check which features helped the most in the scenario of late fusion vs early fusion.

|  |  |
| --- | --- |
| Provisional organization of chapters | * Abstract * Introduction * Related work * Methodology   + Data   + Experiment Setup * Results * Discussion   + Error analysis * Conclusion & Future Remarks |

Provisional book list

(appendix)

Timetable

|  |  |  |
| --- | --- | --- |
| Phase | Dates | Tasks |
| orientation phase | 6/2 – 17/3 | - Explore dataset and literature  - Determine the topic and setup the thesis planning  - look up features that could be used |
| execution phase | 17/3 – 31/3 | - writing related work  - finalize features to be extracted (theoretically) |
| 31/3 – 14/4 | - implementation of features |
| 14/4 – 28/4 | - early fusion models  - late fusion models |
| 28/4 – 12/5 | - word embedding (optional)  - writing methodology |
| 12/5 – 26/5 | - writing results  - writing discussion  - start error analysis |
| 26/5 – 9/6 | - finalize error analysis  - writing error analysis  - writing conclusion & introduction & abstract |
| completion phase | 9/6 – 16/6 | - Finalizing report  - Cleaning up the code |
| 16/6 – 23/6 | Preparing presentation |

start date 01/04/2023

planned thesis presentation date /06/2023

planned graduation date 30/07/2023

Remarks .............................................….......................................