Heidelberg University
Institute of Computer Science
Database Systems Research Group

Project Proposal for the lecture Text Analytics

Sigmund

Depression Features in Conversational Transcripts

Team Members: Julius Daub (3536557) | Applied Informatics (M.Sc.)
julius@daubweb.com
Alexander Haas (3503540) | Applied Informatics (M.Sc.)
haas.alexanderjulian@gmail.com
Ubeydullah Ilhan (3447661) | Applied Informatics (M.Sc.)
ubeydullah.il@gmail.com
Benjamin Sparks (3664690) | Applied Informatics (M.Sc.)
benjamin.sparks@protonmail.com
GitHub: `https://github.com/gitexa/sigmund`

1 Motivation

Major depressive disorder (MDD) ranks among the leading mental health causes of the global burden of disease [1] and, with a lifetime risk of 10-20% for women and 5-12% for men, it is the most prevalent mood disorder [2]. One important impact of depression is impairment of social functioning, defined as "an individual's ability to perform and fulfill normal social roles". [3] This impairment affects mainly the social role domain, like interpersonal and marital functioning, and less the work role domain. [4] Therefore, the depression-related change of behaviour within personal relations, which can be observed for instance in conversations, can be used as a test bed to assess social functioning, both in terms of diagnostic purposes and measurement of therapeutic success. [5]

As resources in psychiatry are typically limited, depressed patients often lack professional attention due to constraints of mobility, cost and motivation. Therefore, digital solutions for mental illnesses, such as automated monitoring, promise great potential for improving diagnosis and treatment on scale. [6][7]

In this project, we examine transcripts of couple conversations from a current research project at Heidelberg University Hospital [5] by using text analytics methods. We want to identify depression-related features and quantify the differences between couples, in which one partner is suffering from MDD, and couples, in which both partners lack a history of depression. Furthermore, we use said features, to support the assessment of efficacy for a new intervention, the *Cognitively Based Compassion Training (CBCT)*.

2 Related Work

Among the many symptoms associated with major depressive disorder (MDD), one important symptom dimension is overall impairment in social functioning [5][4]. Within the *Research Domain Criteria (RDoC)* [8], a framework to structure mental illness research along different domains, our domain of interest is "systems for social processes" with its subdomain "social communication". According to Kupferberg et al. [4], further research is necessary to measure depression on the symptom level, including measures specifically related to RDoC's social domain. Therefore, our project aligns well with current psychological research needs.

Many works have examined conversation in depression [7][9][10], revealing that depression affects many linguistic properties, including phonetics, semantics, and syntax. [10] Although most of the research focused on multimodal features including acoustic features of speech, video features of facial expressions and semantic analysis of the transcript, pure analysis of the transcript has shown promising predictive capabilities on MDD diagnosis. [9][10]

Using algorithms to automatically detect depression by passive observation is a promising goal, which was also widely addressed by the research community. As data containing social behaviour of depressed persons is, on a large scale, often only available from social networks, research focused on detecting MDD using features from public social media data as Twitter and Reddit posts. [11] [12]

Apart from social media, approaches like the *Audio Visual Emotion Challenge (AVEC)* [13] provide multimodal data from psychological settings as a benchmark for algorithm testing. Although successful approaches often combine different modalities, transcript-only based analysis yield promising results. Williamson et al. showed on the 2016 AVEC dataset that semantic analysis of dialogue transcripts provided the highest performing features on depression classification compared to other modalities (mean F1 0.84 for semantic features compared 0.81 for ensemble method). [9] Morales et al. compared on the 2014 AVEC dataset, consisting of patients with labeled depression scores who answer specific questions by a human-controlled agent, the performance of speech- and text-based features for regression on the depression score, resulting in a better performance of text features in terms of RMSE.

In contrast to the strict format of QA-conversations following predefined questions and resulting answers, other works have shown, that free conversations, without prior assumptions on the topics covered and minimal information on the structure of the conversation are able to yield results comparable to those explicitly modelling topics. [7]

To summarize, there is sufficient proof in literature, that text analytics methods perform well on transcripts of conversations, even on non-guided conversations, and are able to quantify depression-related changes in social functioning, allowing predictions on depression scores.

3 Project Description

3.1 Main Goals

Our project is related to the *Social Interaction in Depression (SIDE) Study* [5] which intends to quantify the effect of a new intervention for couples in which one partner suffers from MDD.

From the perspective of the study, we follow two main goals:

1. First, we will investigate whether it is possible to distinguish between conversation transcripts of depressed and non-depressed couples. Framed differently, we want to identify depressed couples based on the conversation transcript, which can be considered a classification problem.

2. Secondly, we want to quantify the degree of depression, to be able to measure whether a (positive) effect of the novel therapy can be determined, which can be considered a regression problem on the linear Hamilton Depression Scale (HDRS) [14].

Both questions are intended to be answered solely on the basis of the transcribed conversations without having to resort to more complex methods. The underlying research question from the perspective of text analytics is: "What patterns are present in the transcript of an instructed couple's conversation that relate to MDD?"

We follow two different approaches to tackle the research question. First, the **hypothesis-driven approach** applies manual feature engineering using self-defined metrics, partly based on domain knowledge and partly based on functioning concepts from literature.

Secondly, the **explorative approach** tries to identify latent features through unsupervised learning, using mappings to a word embedding space, as can be done by word2vec [15] or lda2vec [16], identifying latent topics by latent dirichlet allocation (LDA) [17] and latent semantics by latent semantics analysis (LSA) [18].

3.2 Pipeline

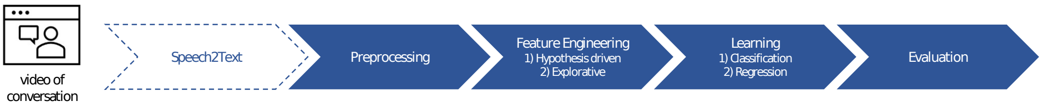

Figure 1: Overview of our text analytics pipeline

Our text analytics pipeline is visualized in figure 1. In the following sections, we describe the parts and associated challenges of the pipeline briefly.

3.2.1 Dataset and Preprocessing

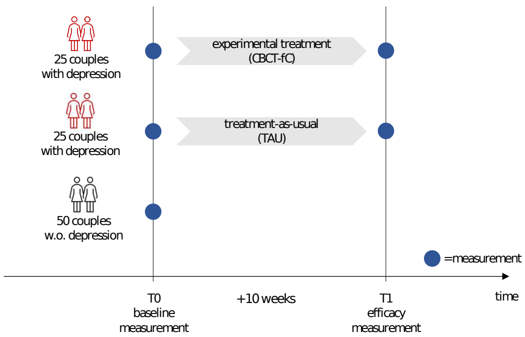

Figure 2: Overview of our dataset

In the SIDE study, many assessments were performed to determine differences between depressed and non-depressed couples, including a questionnaire to assess MDD status on the Hamilton Scale, 24h ECG, saliva and blood samples as well as an instructed conversation between the two partners of 10 minutes duration. As instruction for the conversation, the couple received a list of possible but not obligatory topics and the obligation to talk praising and positive to each other. The conversation was video-taped and gazes were tracked via eye-tracking.

For our analysis we use transcripts of the conversation, which are provided by the study team. As depicted in figure 2, a total number of 100 couples participated. At the beginning (t0), all 100 couples are tested as described above, while recording and transcribing the conversation. After 10 weeks of therapy (t1), an additional testing is performed only for the depressed couples to measure therapy outcomes. Therefore, the dataset contains a total of 150 testing samples.

Special care must be taken when preprocessing the transcripts, for example with stop-word removal, to avoid removing words whose frequent occurrence could be indicative of a depressive disorder. [9]

3.2.2 Feature Engineering and Learning

After collecting and preprocessing the data, we need to extract features. As mentioned, we use two approaches, first exploiting domain knowledge and manually crafting features, secondly retrieving latent features in an exploratory manner. We use the obtained features to distinguish between transcripts of couples affected with MDD from couples not affected with MDD, which falls into the domain of **text-classification**. Furthermore, we infer the depression score on the HDRS, which falls in the domain of **regression**. When deriving features, several other domains of text analytics are included as subtasks, for example methods for **sentiment analysis** and **word embeddings**, which have been used successfully in literature to measure the degree of depression. [9][11]

As we have domain knowledge from psychology within the group and also the research team of the SIDE study for consultation, we can split the two approaches between the group members while parallelizing our work at the same time.

3.2.3 Evaluation

To measure the quality of our predictions, we use metrics successfully proven in the context of the AVEC competition. For the classification task, we use the F1 metric, which is the harmonic mean of precision and recall, as well as the area under the receiver operating characteristic curve (AUC). For the regression task, we use the root mean squared error (RMSE). [9][10]

There is no direct baseline for evaluation, as no similar project has been carried out that we know of. As a naive baseline for the classification and regression task, classification by chance (0.5) or regression by chance (1/number of points) can be used. Another baseline can be derived from similar problems in the AVEC competition, as was done by Morales et al. [10]. As our data is different, it can serve only as orientation, and cannot be fully transferred to our results.

References

[1] A. J. Ferrari, F. J. Charlson, R. E. Norman, S. B. Patten, G. Freedman, C. J. Murray, T. Vos, and H. A. Whiteford, "Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010," *PLoS Medicine*, vol. 10, p. e1001547, nov 2013.

[2] M. Fava and K. S. Kendler, "Major depressive disorder," nov 2000.

[3] R. M. Hirschfeld, S. A. Montgomery, M. B. Keller, S. Kasper, A. F. Schatzberg, H. J. Möller, D. Healy, D. Baldwin, M. Humble, M. Versiani, R. Montenegro, and M. Bourgeois, "Social functioning in depression: A review," *Journal of Clinical Psychiatry*, vol. 61, no. 4, pp. 268–275, 2000.

[4] A. Kupferberg, L. Bicks, and G. Hasler, "Social functioning in major depressive disorder," *Neuroscience and Biobehavioral Reviews*, vol. 69, pp. 313–332, 2016.

[5] C. Aguilar-Raab, M. N. Jarczok, M. Warth, M. Stoffel, F. Winter, M. Tieck, J. Berg, L. T. Negi, T. Harrison, T. W. Pace, and B. Ditzen, "Enhancing Social Interaction in Depression (SIDE study): Protocol of a randomised controlled trial on the effects of a Cognitively Based Compassion Training (CBCT) for couples," *BMJ Open*, vol. 8, no. 9, pp. 1–15, 2018.

[6] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial," *JMIR Mental Health*, vol. 4, no. 2, p. e19, 2017.

[7] T. Alhanai, M. Ghassemi, and J. Glass, "Detecting depression with audio/text sequence modeling of interviews," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, vol. 2018-Septe, pp. 1716–1720, 2018.

[8] T. Insel, B. Cuthbert, M. Garvey, R. Heinssen, D. S. Pine, K. Quinn, C. Sanislow, and P. Wang, "Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders," jul 2010.

[9] J. R. Williamson, E. Godoy, M. Cha, A. Schwarzentruber, P. Khorrami, Y. Gwon, H. T. Kung, C. Dagli, and T. F. Quatieri, "Detecting depression using vocal, facial and semantic communication cues," *AVEC 2016 - Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, co-located with ACM Multimedia 2016*, pp. 11–18, 2016.

[10] M. R. Morales and R. Levitan, "Speech vs. text: A comparative analysis of features for depression detection systems," in *2016 IEEE Workshop on Spoken Language Technology, SLT 2016 - Proceedings*, pp. 136–143, Institute of Electrical and Electronics Engineers Inc., feb 2017.

[11] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," tech. rep., 2013.

[12] Q. Cong, Z. Feng, F. Li, Y. Xiang, G. Rao, and C. Tao, "X-A-BiLSTM: A Deep Learning Approach for Depression Detection in Imbalanced Data," in *Proceedings - 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018*, pp. 1624–1627, Institute of Electrical and Electronics Engineers Inc., jan 2019.

[13] M. Valstar, B. Schuller, K. Smith, F. Eyben, B. Jiang, S. Bilakhia, S. Schnieder, R. Cowie, and M. Pantic, "AVEC 2013 - The continuous Audio/Visual Emotion and depression recognition challenge," *AVEC 2013 - Proceedings of the 3rd ACM International Workshop on Audio/Visual Emotion Challenge*, pp. 3–10, 2013.

[14] M. HAMILTON, "A rating scale for depression," *Journal of neurology, neurosurgery, and psychiatry*, vol. 23, pp. 56–62, feb 1960.

[15] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Distributed Representations of Words and Phrases and their Compositionality," tech. rep.

[16] C. E. Moody, "Mixing Dirichlet Topic Models and Word Embeddings to Make lda2vec," may 2016.

[17] J. C. Campbell, A. Hindle, and E. Stroulia, "Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data," Tech. Rep. 4-5, 2015.

[18] S. T. Dumais, "Latent Semantic Analysis," jan 2004.