

Effect of mental imagery parameters on the risk of experiencing loneliness in a period of isolation

Literature Review

Author: Mark Gabriel

Supervisors: Dr. Julie Ji, Dr. Debora Correa, Dr. Wei Liu

1 Abstract

The COVID-19 pandemic has forced a sudden change in society that we weren't fully prepared for. Self-isolation and physical distancing has been imposed on everyone in varying degrees and the effect of it on people's mental wellbeing is an area of interest, especially if a second or third wave of the pandemic will require people to undergo isolation again. Certain members of the community are going to be affected more than others from this rapid disruption of the day-to-day because they are not getting the usual levels of social interaction they need or are used to to maintain a healthy mental wellbeing. It is in our interest to find which members of our community are at most risk with a method that is highly accessible given the physical restrictions that such a societal condition imposes.

This study aims to be able to extract a risk level of loneliness from a person's free-text response to a scenario using natural language processing techniques. The mental imagery parameters of each response has been manually assessed and will be used to analyse what features of the corpus will be relevant in creating the model for determining a person's risk level.

2 Review of Literature

2.1 Mental Imagery

Mental imagery is the simulation of perceptual experience [4] across sensory modalities. It has been strongly linked to depression [6] which is our main area of interest. As detailed in Kosslyn et al's 2001 paper, the current techniques for measurements of a person's mental imagery involve visual tests, classification tasks, comparison tasks, questionnaires and interviews.

The usage of machine learning and natural language processing has not yet been deeply explored in extracting mental imagery parameters, though theoretically, it would be most similar to the existing method of using interviews and analyzing a person’s speech and selection of words to derive the wanted parameters, except with machine learning, the assessment would be automated rather than needing a trained professional to go through each case individually, which would be an expensive process that won’t be easily scalable in an isolation scenario where potentially a huge percentage of people are affected.

Dr. Julie Ji’s research on mental imagery, the experience of perception in the absence of external sensory input, shows that the person’s quality of mental imagery-based simulations has a significant link to maintaining and amplifying emotional states [3]. In one study on mental imagery as a motivational amplifier to promote activities as a key treatment in depression [7], the Motivational Imagery group reported higher levels of motivation, anticipated pleasure, and anticipated reward for the planned activities compared to the Activity Reminder control group and the No-Reminder control group.

2.2 Named Entity Recognition

We aim to be able to utilise the free-text responses of the survey to create a model that performs better than a regression model from the human-graded mental imagery parameters that predicts a person’s risk of loneliness. To achieve this, the collected free-text responses have to be annotated with respect to the domain of interest. In our particular research area, it will be important to tag self actions (meditate, jog, sleep in, watch TV, etc) and interpersonal actions (call friend, video call with parents, have dinner at restaurant with partner) separately, which most pre-trained models aren’t trained to do. Due to this, and also because we have a relatively small sample size, we have the luxury of using manual annotation to improve our data’s richness and context.

In the paper Multilingual Twitter Sentiment Classification: The Role of Human Annotators by Mozetic et al, [5], they find that the quality of classification models depends much more on the quality and size of training data than on the type of model trained. Human annotators can help us maximise our model’s performance by ensuring good annotation quality, which can be done by setting an overlap value (the number of times a single corpus is annotated) of at least 2.

There are a lot of tools for data annotation available. Among these options are BRAT, GATE Teamware, WebAnno, SAWT, Yedda, SANTO and UWA’s Redcoat. We can see a comparison of their features in Figure 1. With research assistants to help out with our data as annotators, an important feature we are looking for in an annotating tool is the ability to propagate documents easily and for the tool to scale the workload automatically based on the number of annotators. Redcoat [9] has been selected as the annotating tool for that reason, alongside being web-based and easy to set up and distribute.

System	Web-based project creation	Project monitoring	Curation feature	Document propagation	Class labels		
					Dynamic	Hierarchical	Multi-label
BRAT	✗	✗	✗	✗	✗	✓	✗
GATE	✓	✓	✓	✗	✗	✗	✗
SANTO	✓	✓	✗	✗	✗	✗	✗
SAWT	✓	✓	✗	✗	✗	✗	✗
YEDDA	✓	✓	✗	✗	✗	✗	✗
WebAnno	✓	✓	✓	✗	✗	✗	✗
Redcoat	✓	✓	✓	✓	✓	✓	✓

Figure 1: A comparison of existing annotation tools with Redcoat. Dynamic refers to the ability for any user to modify the class labels during annotation.

2.3 Sentiment Analysis

One of the mental imagery parameters we’re concerned with is a person’s emotional tone when they’re describing what they think or see when they’re given a scenario and a goal of starting lonely in an isolated setting and ending in a brighter, happier mental state. A person’s emotional tone can be approximated with sentiment analysis techniques. XLNet has been used to identify optimism and pessimism in twitter messages by Alshahrani et al [1], which is similar to what we’re trying to achieve. XLNet models are able to model negations and other semantic relationships by paying attention to key words, leading to a model accuracy of 96.45

A point of discussion would be if our research problem would work better with basic sentiment analysis or aspect-based sentiment analysis. Normal sentiment analysis grades the emotional tone of the overall text, while aspect-based sentiment analysis goes through the text to identify various aspects and determines the corresponding emotional tone for each one. There isn’t much available literature on when each of the techniques are relevant or when each of them should be used over the other, but the spirit of parsimony would lead us to towards using aspect-based sentiment analysis only if there are particular aspects in our research domain that we want to track individually.

For the lack of literature on comparison between the two techniques on different domains, there is merit in trying both individually, or creating a model that uses both techniques in tandem and empirically comparing which technique works best for our research problem.

2.4 Transfer Learning and XLNet

Language is complex. In order to get reasonably performing models from scratch, one would need an extremely large dataset in addition to the computing power needed to process it. This is often impossible for many researchers, and a solution to this problem is transfer learning. Transfer learning is a machine learning technique where a model is trained on a particular problem, then reused in a similar domain to carry over the knowledge of the neural network. Depending on how different the pre-trained model is to your problem, Additional training needs to be done to finetune the model to your specific domain. [8]

Model	MNLI	QNLI	QQP	RTE	SST-2	MRPC	CoLA	STS-B	WNLI
<i>Single-task single models on dev</i>									
BERT [2]	86.6/-	92.3	91.3	70.4	93.2	88.0	60.6	90.0	-
RoBERTa [21]	90.2/90.2	94.7	92.2	86.6	96.4	90.9	68.0	92.4	-
XLNet	90.8/90.8	94.9	92.3	85.9	97.0	90.8	69.0	92.5	-
<i>Multi-task ensembles on test (from leaderboard as of Oct 28, 2019)</i>									
MT-DNN* [20]	87.9/87.4	96.0	89.9	86.3	96.5	92.7	68.4	91.1	89.0
RoBERTa* [21]	90.8/90.2	98.9	90.2	88.2	96.7	92.3	67.8	92.2	89.0
XLNet*	90.9/90.9[†]	99.0[†]	90.4[†]	88.5	97.1[†]	92.9	70.2	93.0	92.5

Figure 2: Results on GLUE. * indicates using ensembles, and [†]denotes single-task results in a multi-task row. All dev results are the median of 10 runs. The upper section shows direct comparison on dev data and the lower section shows comparison with state-of-the-art results on the public leaderboard.

Popular pre-trained NLP models include Google’s BERT, Google and Carnegie Mellon University’s XLNet, and Facebook’s RoBERTa. All of these models are transformers which are deep learning models designed to handle sequential data using the concept of attention mechanisms. Attention mechanisms allow the model to access and use any previous states and utilises them based on relevance to the current node. [10]

In particular, we want a pretrained model that is strong in sentiment analysis. Looking at Figure 2 in the SST-2 column which stands for the Stanford Sentiment Treebank v2, we can see that XLNet [12] outperforms both BERT and RoBERTa in sentiment analysis on the SST-2 dataset.

While XLNet, BERT and RoBERTa all come from the same family of models, XLNet differs by introducing permutation language modelling, where all tokens are predicted in a random, permuted order rather than the traditional sequential order. This helps XLNet learn non-adjacent and bidirectional relationships between the words in the corpus. Secondly, XLNet uses Transformer-XLs as the base architecture which enables learning dependency beyond a fixed length without disrupting temporal coherence. [2] Transformer-XLs, even without the permutation language modelling, has been shown to give performance increases in NLP tasks.

For the above reasons, and for the benchmarks set out in Figure 2, XLNet will be our pretrained model of choice. In addition, HuggingFace [11] provides a transformer library that wraps all of these transformers for use with PyTorch or TensorFlow in a neat package that allows us to utilise the base XLNet model and the large XLNet model, alongside any other pretrained transformer model if we wish to make model comparisons.

References

- [1] A. Alshahrani, M. Ghaffari, K. Amirizirtol, and X. Liu. Identifying optimism and pessimism in twitter messages using xlnet and deep consensus. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2020.

- [2] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V. Le, and Ruslan Salakhutdinov. Transformer-xl: Attentive language models beyond a fixed-length context, 2019.
- [3] Julie L. Ji, David J. Kavanagh, Emily A. Holmes, Colin Macleod, and Martina Di Simplicio. Mental imagery in psychiatry: Conceptual & clinical implications. *CNS Spectrums: the international journal of neuropsychiatric medicine*, 24(1):114–126, February 2019.
- [4] Stephen M. Kosslyn, Giorgio Ganis, and William L. Thompson. Neural foundations of imagery. *Nature Reviews Neuroscience*, 2(9):635–642, Sep 2001.
- [5] Igor Mozetič, Miha Grčar, and Jasmina Smailović. Multilingual twitter sentiment classification: The role of human annotators. *PLOS ONE*, 11(5):e0155036, May 2016.
- [6] Wheatley J Wells A Fisher P Myers S Patel T, Brewin CR. Intrusive images and memories in major depression. *Behavior Research and Therapy*, 45(11):2573–2580, Nov 2007.
- [7] Fritz Renner, Fionnuala C. Murphy, Julie L. Ji, Tom Manly, and Emily A. Holmes. Mental imagery as a “motivational amplifier” to promote activities. *Behaviour Research & Therapy*, 114:51–59, March 2019.
- [8] Sebastian Ruder, Matthew E Peters, Swabha Swayamdipta, and Thomas Wolf. Transfer learning in natural language processing. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials*, pages 15–18, 2019.
- [9] Michael Stewart, Wei Liu, and Rachel Cardell-Oliver. Redcoat: A collaborative annotation tool for hierarchical entity typing. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*, pages 193–198, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.
- [11] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. Huggingface’s transformers: State-of-the-art natural language processing. *ArXiv*, abs/1910.03771, 2019.
- [12] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov, and Quoc V. Le. Xlnet: Generalized autoregressive pretraining for language understanding, 2020.