

Progress to date:

- Decided two-pronged approach of human survey and corpus-based co-occurrence analysis
- Manually Cross Referenced EmojiTracker website and EmojiNet to find popular emojis in the relevant semantic categories to determine emojis in the four groups to measure intensity
- Designed survey with fill-in-the-blank style questions to determine semantic strength within the group
- Evaluated twitter corpus based studies in classification, applied for access to developer account

Related Research evaluated:

Sheinman, Vera, Christiane Fellbaum, Isaac Julien, Peter Schulam, Takenobu Tokunaga. **"Large, Huge or Gigantic? Identifying and encoding intensity relations among adjectives in WordNet"**. Springer Science+Business (2013).

Fellbaum, Christiane and Yvette Mathieu. **"A corpus-based construction of emotion verb scales"** (2014)

Wijeratne, Sanjaya, Lakshika Balasuriya, Amit Sheth, and Derek Doran. **"EmojiNet: Building a Machine Readable Sense Inventory for Emoji."** ArXiv. October 25, 2016. Accessed September 25, 2018. <https://arxiv.org/abs/1610.07710>

Barbieri, Francesco, Miguel Ballesteros, and Horacio Saggion. **"Are Emojis Predictable?"** ACL Anthology. April 2017. Accessed September 24, 2018. <https://aclanthology.coli.unisaarland.de/papers/E17-2017/e17-2017>

Novak, Petra, Jasmina Smailovic, Borut Sluban, and Igor Mozetic. **"Sentiment of Emojis."** CoRR1509 (2015).

Wahyuni, Rinda, and Indra Budi. **"Combining Linguistic, Semantic and Lexicon Feature for Emoji Classification in Twitter Dataset."** Procedia Computer Science135 (2018): 194-201. ScienceDirect.

Current difficulties:

My biggest difficulty at the moment is getting access to a Twitter developer account. I have filled out two applications and am still waiting to hear back. I am planning to send an email this Monday to their help desk in order to try and expedite the process. Once I have this access, I am sure to run into more productive difficulties ☺

Next steps:

I want to deploy the survey and start collecting data by the end of this week, and I hope to finish collecting by about the 15th of November. I can then work on analyzing the data that I get.

I hope to have Twitter access in the next couple of days, and will then spend time writing queries to curate my dataset. I hope to analyze co-occurrence and tabulate the results about two weeks into November, so that I can then compare and synthesize them in conjunction with the survey data.