



A Philosophical Analysis of Personality Trait Detection Through Text Analysis

Mary Irene Clare O. Deleña, MS, RPsy

*Philosophy Department
De La Salle University
mary.irene.delena@dlsu.edu.ph*

Abstract: Utilizing the World Wide Web as source of the growing corpus of data, various studies have recently focused on modelling and predicting certain variables through the user's language used. Although algorithms and statistical analyses support the accuracy of techniques employed by computer scientists and computational linguists, just like any experimental studies, these methods also account for margin of error. Personality detection has been a popular topic in the field of natural language processing yet limited studies have been done in the field of philosophy of language where correlations in the rationale for such studies can be highly observed. Tracing back the philosophical influences of these strategies in understanding human behavior, (i.e. personality psychology), and also as a follow up to Wilks' article, "What Would a Wittgensteinian computational linguistics be like?" (2008), this research aims to tackle points for debate on the ability or inability of a computer to understand the nature and meaning of texts as much as human beings can. In line with this year's congress theme, "Padayon Sining: A Celebration of the Enduring Value of the Humanities", this paper argues that a research in the humanities (a philosophical analysis for instance) brings about relevant foundational issues in certain fields—in this case, on understanding personality through text analysis in the computer sciences.

Key Words: philosophy of language, text analysis, social media, natural language processing, personality

1. INTRODUCTION

Advances in computational linguistics and natural language processing traces its roots in two philosophies of language (Clark, 2014). The first is of Frege and Russell who analyzed the meaning of sentences using logic. The second is of Wittgenstein and his "meaning of use":

The context words can be obtained
empirically by automatically

analysing large amounts of text, e.g. the one billion words on Wikipedia. If we perform a similar analysis for the word "cat", we find that "dog" and "cat" tend to share context words, and hence appear close together in the information space. Words that are close together are considered to be close in meaning. (Clark, 2014)



Clark (2014) identified Frege and Russell's method as the logical approach and Wittgenstein's as the distribution vector-based approach.

In 2008, Wilks, a linguist, wrote about the influence of (the latter) Wittgenstein in the field of artificial intelligence (AI), specifically computational linguistics. Wittgenstein's "Philosophical Investigations" became the anchor of his arguments on the learning process of artificial intelligence. He wrote,

Wittgenstein's remark is situated in a work which seeks to show how language is used, and in so doing, to highlight the essentially social nature of language. To use language is to be part of a group of language users (there can be no private language), so that the meanings of words and concepts found within these shared language-games are thus perspicuous and sound. (Wilks, 2008, p. 49).

It can then be implied that with sufficient data mining, computers may be trained to use language as humans do. Moreover, what makes this analysis of language valuable is that it directly reflects the people's way of life as their expressions are found in the World Wide Web. Specifically, social media platforms such as Twitter, Facebook, and Instagram have been a popular source of data corpus.

A recent research in the field of computer science has attempted to create a model that would predict personality traits of Filipinos through their tweets (Tighe, 2017). In the research, personality traits oriented to the five-factor model (Costa and McCrae, 1985) were of consideration. Using the Big Five Inventory, a data-driven exploration of various techniques in natural language processing determined that there are certain personality traits that are easier to model than the rest. This specific example seems to be an effort towards teaching the computer the ability to detect personality through textual assessment.

2. PHILOSOPHY AND COMPUTATIONAL LINGUISTICS

This paper aims to lay out points for philosophical debate in the progressive field of computational linguistics. The next sections of this paper contain my reflections on how various philosophies of language relate to the modeling of personality through Twitter. To limit the scope, I have only focused on four philosophers of ideal language. The first section will be on Wittgenstein and latter philosophy. The second section will be on Austin and Searle and their speech act theory. The third section will be on Donellan and his descriptive descriptions. The fourth section will be on Grice and his implicatures. The fifth section will be an integration of the four.

It would be important to note that this paper does not focus on the evaluation of the methodology of training the machine to be a smarter AI. Neither does this paper focus on the critique of the efficiency of NLP as a strategy for determining and understanding personality. This paper puts into focus the philosophical implications on the generation of linguistic meanings.

3. LANGUAGE AND MEANING MAKING

Meaning making of information found in the World Wide Web can be viewed tricky to an extent. Sources of these information are people of various backgrounds and of different language games. As Wittgenstein's notion of limiting one's world through the language used is a helpful means of attaining the current baseline mode of communication, the idea of Austin and Searle, Donellan, and Grice of the nature of language being expressed should also be taken into consideration.

Above all, Wittgenstein still appears to be the most influential philosopher when it comes to machine learning. The limits of my language mean the limits of my world seems to be equivalent to the limits of my data is the limits of my world. Although machines are being trained to have



Presented at the 12th DLSU Arts Congress
De La Salle University, Manila, Philippines
February 20, 21 and 22, 2019

meaningful understanding of language, it is only limited to those that are used publicly and are available for inclusion for machine learning. There seems to be an evident significance of qualitative validation that would utilize the human capacity to appropriate contexts on specific text analyses especially when it comes to personality prediction.

4. POSSIBLE VIEWS OF PHILOSOPHERS OF LANGUAGE ON NLP

NATURAL LANGUAGE PROCESSING AND WITTGENSTEIN

Wittgenstein's influence on the field of natural language processing has already been established earlier in this paper through the distribution vector-based approach (Clark, 2014). Wittgenstein's notion that the meaning of language is generated from how it is used paves way for its creative nature (Wilks, 2008).

However, for the development of AI to be dependent on the frequency of the usage of language seems to be leading to a limitation:

On this account, replication (as a measure for success in AI development) is a limiting concept, and proves impossible for the simple reason that what creativity means is dependent on a (potentially unquantifiable) number of variables. That there are different forms of creativity already constitutes part of how we perceive creativity, and this argument might prove most fruitful for claims that aspects of non-human creativity—though they may be particularly or even substantially different from our own—should nevertheless be considered creative in some way. (Wilks, 2008, p.50)

There is a major discrepancy on how creativity is for humans and for machines. One appears to be more superior to the other as creativity of the programmed

machines are dependent on the creativity of human. For humans, thoughts are equivalent to propositions and text is the representation of reality, thus can be a medium for understanding the human psyche. Although all these can be taught to the computer, humans have the special ability to differentiate meanings through other means such as non-verbal cues, contexts, and intent. Because of subjectivity and different perspectives of understanding the world, text can also misrepresent reality.

NATURAL LANGUAGE PROCESSING AND AUSTIN AND SEARLE

Apart from the subjective meaning of texts, the validity of the linguistic value of these groups of words are also important to consider. Determining whether a text as representation of reality is successful is necessary especially in noting word usage as linked to certain traits of people. The speech act theory of Austin and Searle is about the constative and performative functions of language. Constantives can either be true or false while performative can either be happy or unhappy. (Mabaquiao, 2004). Being true or false and happy or unhappy seems to only be validated through the observation of the state of affairs. Validation of gathered tweets would be difficult, even close to impossible. This is primarily because of the remoteness of the researcher from the tweeter. Weeding out spoiled data is possible, but the sincerity of those messages gathered and the performative acts could hardly be detected, unless again, validated one by one.

NATURAL LANGUAGE PROCESSING AND DONELLAN

NLP seems to not have the ability to filter referential and attributive definite descriptions yet. There needs to be a human validation that will only be one through a qualitative analysis of the gathered data. Repetition does not always imply sameness. For instance, the word "kita" may have two different translations:

- (1) see and (2) profit.



Presented at the 12th DLSU Arts Congress
De La Salle University, Manila, Philippines
February 20, 21 and 22, 2019

It might be easier to teach the machine using the logical approach but might be difficult for the distribution vector-based approach. As exhibited in the earlier example, machine learning has its limits when it comes to repetition of words and pinpointing certain context of how they were used to represent things in the world. Computer programmed in such way might confuse sameness and equivalence with substitution. A word may be confused to mean another apart from what it was originally meant. Although the data set must possibly have included this possibility, there is still a chance that confusion might arise on which describes which. In terms of the sentence structure, in tweets, sometimes only the subject of a sentence is typed and posted. In some instances, it is only the predicate. It would then be a challenge to trace and ensure that people are referring to the same thing. In these instances, hashtags might be of utility to aid in the grouping of common tweets.

NATURAL LANGUAGE PROCESSING AND GRICE

Grice's theory of conversational implicatures consists of maxims that ensure meaningful conversations in the cooperative principle:

This cooperative principle is an umbrella term for nine components that guide how we communicate. These nine components are grouped together into four categories, called the Maxims of Conversation: the maxim of quality (truthfulness), the maxim of quantity (informativeness), the maxim of relation (relevance), and the maxim of manner (perspicuity). (p. 2)

The first maxim is quality. It pertains to the message being delivered as factual—correct and accurate. Data gathered from Twitter may or may not be always true. The rising issue of fake news is rampant and even humans take filtering what is fake and what is not as a challenging task. Familiarity with events that are being talked about is a necessity to determine its truth value. The second maxim is quantity. It

pertains to the sufficient degree of the ability of statements to relay adequate information. From its face value, a tweet is only limited to 280 characters which used to be 140. The third maxim is relation. It refers to the weight of significance of the message being given and received. In Twitter, messages are publicly relayed to specific persons. Relevance must seem questionable as it is understandable that private language used by certain groups of people cannot be easily detected by a machine that is only trained by humans themselves. All these four maxims bring about details of which aspects to look into when finding meaning in language use. In the example of using tweets in user profiling specifically determining their personality (openness, conscientiousness, extraversion, agreeableness, and neuroticism), consideration of all of Grice's maxims may raise a number of questions in meaning making. Specifically, there are factors that are not easily detected by machine learning as data gathered are randomly assigned to training and testing groups (Tighe, 2017). Some of these factors are closeness of someone to another person he/she is referring to, inside jokes, private language, and jargons. This could be a point for humans over computers as the way words are expressed in certain contexts affect the extent of understanding of the receiver of the message. Grice came up with matrices that make up meaningful conversations and meaning is dependent on the relationship of the giver and the receiver of the message. Person to person relationship is not the same as computer-person and computer-researcher relationship. When it comes to meaning making of specific statements, the mediation of a computer may be a source of confusion because of the tagging done that is based on prior training from prior data gathered.

5. CONCLUSION

Advances in AI are traced back to the influences of philosophy. Issues that arise in this science which are of interest include those that may be tackled by philosophy majors. By going back to its roots, that is philosophy, issues or limitations of the field are tackled not through focusing on the methodology by writing a critique but by

Presented at the 12th DLSU Arts Congress
De La Salle University, Manila, Philippines
February 20, 21 and 22, 2019



identifying the philosophical implications of meaning making in the World Wide Web, specifically on Twitter. Moreover, relying on the power of text should take into serious consideration other factors that are brought about by different language games that are being played by every individual. Variations in language use among Twitter users may not always be accounted for in NLP. Moreover, context and intent cannot always be determined especially that Twitter only limits each to 280 characters. Thus, it may be implied that human intervention is still necessary to validate automations performed by machines. As machines learn through the training and testing stages, it must be remembered that all these are programmed by the computer scientists from the data collected from people across the globe.

This reflection of the philosophical implications of meaning making from gathering tweets for predicting personality have looked into natural language processing from the lens of four theories from the field of the philosophy of language. These are: (1) Wittgenstein's "meaning of use", (2) Austin and Searle's "constative and performative functions of language", (3) Donnellan's "referential and attributive definite descriptions", and (4) Grice's "cooperative principle and maxims of quality, quantity, relation, and manner". All these gave rise to philosophical discussions on the ability and the inability of computers to understand and make linguistic meanings as human can. It was gathered that although the scope of reach of natural language processing is so much bigger than that of a single individual, certain limitations that pertain to attaining a deeper understanding of language that can only be observed through analyzing the context (i.e. through a longitudinal exploration, a qualitative validation, or even observation of nonverbal cues to detect sincerity and performance that comes with the text) more than the analysis of semantics and syntax.

Humans as the primary users of language are the primary sources of computers in coming up

with their own vocabulary. Personality prediction is a possibility in NLP but there is a number of important points to consider that only humans are qualified to make sense of, at least in the present time. The power of machines seems to be taking over the human ability, but it must also be kept in mind that all these forms of intelligence are called artificial for a specific reason.

6. REFERENCES

- Clark, S. (2014). Breaking new ground in Natural Language Processing. University of Cambridge. Retrieved from <https://www.languagesciences.cam.ac.uk/news/breaking-new-ground-in-natural-language-processing>
- Conversational implicatures. Retrieved from www.bu.edu/linguistics/UG/course/lx502/_docs/lx502-implicatures.pdf
- Costa, P. and McCrae, R. (1985). The NEO personality inventory. *Journal of Career Assessment*, 3(2), 123-139.
- Mabaquiao, M. (2004). Speech Act Theory: From Austin to Searle. Unpublished manuscript.
- Speaks, J. (2005). Donnellan on the referential/attribution distinction and Russell's theory of descriptions. Retrieved from <https://www3.nd.edu/~jspeaks/courses/mcgill/415/donnellan.pdf>
- Tighe, E. (2017). *Modeling Personality Traits of Filipino Twitter Users Based on*

Presented at the 12th DLSU Arts Congress
De La Salle University, Manila, Philippines
February 20, 21 and 22, 2019



Linguistic Markers (master's thesis). De La Salle University, Manila, Philippines.

Wilks, Y. (2008). What would a Wittgensteinian computational linguistics be like. In AISB Convention (pp. 1-6).