**A predição de medidas de personalidade utilizando processamento de linguagem natural pode ser melhorada com controle de aquiescência?**

Araê Cainã

Prof. Dr. Ricardo Primi

**INTRODUÇÃO**

O modelo do Big 5 (cinco grandes fatores) tem como objetivo descrever traços e elementos da personalidade. Para sua medição, são utilizadas diferentes estratégias, como instrumentos de auto-relato (Soto & John, 2017; Xie & Cobb, 2020), instrumentos de hetero-relato (Borkenau et al, 2014), entrevistas estruturadas (Trull et al, 1998) e, mais recentemente, pegadas digitais (Basaran & Ejimogu, 2021) uso de textos usados em redes sociais produzidos por indivíduos (Neuman & Cohen, 2014; Yamada et al, 2019).

O uso de textos para aferição da personalidade depende da transformação de palavras, frases e parágrafos em vetores numéricos. Tais indicadores podem ser produzidos a partir da contagem de termos e sua importância relativa em um texto (bag of words, tf, tf-idf) e/ou de mecanismos que capturam o uso de palavras dentro de um contexto utilizando modelos de linguagem (como o BERT). Os vetores são utilizados como variáveis independentes para predição de medidas de personalidade, utilizando de modelos clássicos - como regressão - até o uso de aprendizagem de máquina supervisionada. (Santos & Paraboni, 2022).

Embora haja sucesso na predição da personalidade utilizando textos (Siddique et al, 2019), ainda resta saber se o controle de viéses de resposta dos instrumentos cujos escores são utilizados como variável dependente podem melhorar a precisão dos modelos atuais. Mais especificamente, será que o controle por aquiêscência poderia melhorar a medida de personalidade aferida por textos? Neste estudo iremos controlar o cálculo do escore de um instrumento que mede personalidade por aquiescência, e utilizar este escore como variável predita por diferentes modelos de representação de texto, verificando se existe melhora dos modelos preditivos após a correção deste viés de resposta

**REFERÊNCIAS**

Basaran, S.; Ejimogu. O. H (2021) A Neural Network Approach for Predicting Personality From Facebook Data. *Sage Open 1(15).* https://doi.org/10.1177/21582440211032156

Borkenau, P; Mosch, A; Tandler, N; Wolf, A. (2014) Accuracy of Judgments of Personality Based on Textual Information on Major Life Domains. *Journal of Personality 84(2)* DOI: 10.1111/jopy.12153

Neuman, Y.; Cohen, Y (2014) A Vectorial Semantics Approach to Personality Assessment. *Scientific Reports* DOI: 10.1038/srep04761

Siddique, F. B., Bertero, D.; Fung, P. (2019) GlobalTrait: Personality Alignment of Multilingual Word Embeddings. *Proceedings of the aaai conference on artificial intelligence (Vol. 33, No. 01, pp. 7015-7022).* <https://doi.org/10.48550/arXiv.1811.00240>

Santos, V. G.; Paraboni, I. (2022). Myers-Briggs personality classification from social media text using pre-trained language models. *Journal of Universal Computer Science, 28(4).*

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*(1), 117–143. [https://doi.org/10.1037/pspp0000096](https://psycnet.apa.org/doi/10.1037/pspp0000096)

Trull, T. J.; Widiger, T. A., Useda, J. D.; Halcomb, J; Doan, B.; Axelrod, S. T. (1998) A Structured Interview for the Assessment of the Five-Factor Model of Personality. *Psychological Assessment 10(3)*. DOI: 1040-3590/98/J3.00

Xie, D. Cobb, C. L (2020) Revised NEO Personality Inventory. *The Wiley Encyclopedia of Personality and Individual Differences: Measurement and Assessment.* 335-350. DOI:10.1002/9781119547167

Yamada, K. Sasano, R. Takeda, K (02019) Incorporating Textual Information on User Behavior for Personality Prediction. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, pages 177–182 Florence, Italy.*