UNIVERSITÀ DEGLI STUDI DI TRENTO

Department of Information Engineering and Computer Science

Bachelor's degree in
Computer Science

FINAL DISSERTATION

TITLE TODO

Supervisor
Alberto Montresor
Co-Supervisor
Daniele Miorandi

Student
Stefano Perenzoni

Academic year 2019/2020

Contents

Abstract

Abstract

Chapter 1

Introduction

During my internship at U-Hopper, I had the opportunity to develop this Thesis as a result of my experience inside the company. **U-Hopper is a research-intensive deep-tech SME, headquartered in Trento, providing big data-enabled solutions and technologies for the government, retail and manufacturing sectors. U-Hopper has received numerous awards for its innovative solutions, including, among the others, the Lamarck prize (2013), a EC Seal of Excellence (2015), the Innov@Retail prize (2016) and a nomination for the 2017 EC Innovation Radar Awards..** The company is active in many different domains such as retail and tourism and offers a variety of competences including chatbots, analytics, and machine learning. Thanks to Tapoi[1], an innovative data intelligence solution, U-Hopper is also into the sector of user profiling. It allows businesses to deliver personalized experiences to their customers through the mining and analysis of their activities on social networks. Thus, the extraction of behavioural insights can be a valuable aspect since being aware of how an individual comes to a decision helps to provide each customer with the right tailored content.

According to neuroscientists Adelstein et al., personality describes human behavioural responses to wide classes of external stimuli. It works as an adaptive system [2]. Psychologists claim that in this system, the constructs represent interacting psychological mechanisms that cause the flow of behaviour and experience; and the adaptive function allows the whole organism to survive and fulfil its needs [6]. The parameters of the adaptive system represent the variation of the same from person to person and, therefore, characterize uniquely every individual. These parameters are also referred to as personality traits in several different personality models studied over the years. Each model includes its range of traits which combinations describe several personality types.

Psychology researchers have shown clear connections between general personality traits and many types of behaviour. First of all, some fundamental traits describe the type of relationship a person has with the outside world and the way he or she communicates [17]. In the field of Human-Computer Interaction, users prefer interfaces designed to represent personalities that most closely matched their own [21]. Some studies have also suggested connections between customer personality and marketing. Through techniques more focused on the target audience, it is possible to profile individuals, and tailor advertisement automatically displayed based on their personality.[5] Therefore, the ability to identify the personalities of people or even better, details of their personality traits is a significant competitive advantage since we would have a well-defined representation of the customer's reasoning process.

1.1 Personality models

In psychology, an individual's personality is described through a specific model comprised of different traits to provide a better understanding of human's unique characteristics. **Qualcosa sull'importanza di un modello per la personalitÃ TODO, poi spostare la riga sotto**

---

[1] `www.tapoi.me`

4

all'interno della subsection While several models exist, the *Big Five*, also known as the *five-factor model* and the *OCEAN model* is one of the most well-researched and widely accepted taxonomies among scientists.[20, 19]

1.1.1 Big-Five factor structure

It formalizes personality along 5 domains, namely Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism. Each one of these traits is continuous and usually ranges on a scale from 1 to 5. At the two extremes of each trait, two separate aspects reflect a particular behaviour. High openness marks imagination, creativity, and curiosity in learning and exploring new things. Conscientiousness represents self-discipline and attention to details. Extroversion measures preferences for interacting with other people. Agreeableness reflects the extent to which a person is generous, trustworthy and always willing to help others. Finally, a high score on neuroticism indicates a tendency to get stuck in negative emotions. For example, conscientiousness is bounded by carelessly at the lowest end and by organization and efficiency at the greatest one. Also, neuroticism, which quantifies the tendency to experience negative emotions, is characterized respectively by nervousness and confidence. "Breve descrizione di ognuno dei 5"

Since its first definition, this model rapidly became one of the standards the psychological community, largely accepted by the most share of scientist. However, concerning the application of personality insights in work environment and marketing received some critics. "Critica al modello, vedi excel del problema per il paper". "Trova paper su critica al modello per quanto riguarda l'applicazione a campi esterni alla psicologia" [13, 22].

1.1.2 Myers-Briggs type indicator

"Introduzione al MBTI, caratteristiche. Dimensioni e problema di classificazione" "Citazione al paper di overlap tra i due modelli"

1.2 Problem statement

"To facilitate communication, it is possible to use personality models to gain a better understanding of what drives the person of interest."

"The personality of an individual is defined as the set of responses to external stimuli"

1.3 Customer insights

1.4 Objectives

1.5 Outline

Chapter 2 describes the state of the art. It could change once that the system design is defined. Chapter 3 introduces the design solution. It focuses on used components and algorithms, their logic and their interfaces. Chapter 4 shows how the mentioned components are implemented and integrated. It follows the implementation of the algorithm and the evaluation of a general prototype of the proposed system. Chapter 5 concludes the thesis with some observations and future work proposals.

Chapter 2

State of the Art

This section and its subsections could change significantly.

    This chapter presents the current state of the art regarding insights extraction on social media. First of all, it is shown what research has achieved concerning Social Media Analysis. Following, it is explained how these aforementioned analyses have been used in dealing with psychometric profiling of online users. Finally, there is a focus on what has been done in terms of providing actionable personality insights.

    "Psychometric profiling is the process by which your actions are used to infer your personality."

2.1   Social media analysis

2.2   Psychometric Profiling

2.3   Personality insights

Far vedere diversi tipi di analisi fatti per la personalitÃ Poi dire che molti si basano sul big5 e far vedere quelli sull'MBTI

Chapter 3

Design and methodology

3.1 Components logic

3.2 Algorithms

Chapter 4

Implementation and evaluation

4.1   Components interactions

4.2   Algorithms implementation

4.3   Evaluation metrics

4.4   Performance evaluation of the system

Chapter 5

Conclusions

5.1 Limitations

5.2 Future work

References

[1] Muhammad M Abdul-Mageed et al. "Recognizing pathogenic empathy in social media". In: *Eleventh International AAAI Conference on Web and Social Media*. 2017.

[2] Jonathan S Adelstein et al. "Personality is reflected in the brain's instrinsic functional architecture". In: *PloS one* (2011).

[3] Yair Amichai-Hamburger and Gideon Vinitzky. "Social network use and personality". In: *Computers in human behavior* 26.6 (2010), pp. 1289–1295.

[4] Danny Azucar, Davide Marengo, and Michele Settanni. "Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis". In: *Personality and individual differences* 124 (2018), pp. 150–159.

[5] Yoram Bachrach et al. "Personality and patterns of Facebook usage". In: *Proceedings of the 4th annual ACM web science conference*. 2012, pp. 24–32.

[6] Colin G DeYoung. "Toward a theory of the Big Five". In: *Psychological Inquiry* 21.1 (2010), pp. 26–33.

[7] Golnoosh Farnadi et al. "A multivariate regression approach to personality impression recognition of vloggers". In: *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. 2014, pp. 1–6.

[8] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. "Predicting personality traits with instagram pictures". In: *Proceedings of the 3rd Workshop on Emotions and Personality in Personalized Systems 2015*. 2015, pp. 7–10.

[9] John W Foreman. *Data smart: Using data science to transform information into insight*. John Wiley & Sons, 2013.

[10] Jennifer Golbeck, Cristina Robles, and Karen Turner. "Predicting personality with social media". In: *CHI'11 extended abstracts on human factors in computing systems*. 2011, pp. 253–262.

[11] Jennifer Golbeck et al. "Predicting personality from twitter". In: *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE. 2011, pp. 149–156.

[12] Sharath Chandra Guntuku et al. "Studying personality through the content of posted and liked images on Twitter". In: *Proceedings of the 2017 ACM on web science conference*. 2017, pp. 223–227.

[13] Leaetta M Hough and Adrian Furnham. "Use of personality variables in work settings". In: *Handbook of psychology* (2003), pp. 131–169.

[14] Kokil Jaidka, Sharath Chandra Guntuku, and Lyle H Ungar. "Facebook versus Twitter: Differences in Self-Disclosure and Trait Prediction". In: *Twelfth International AAAI Conference on Web and Social Media*. 2018.

[15] Margaret L Kern et al. "Gaining insights from social media language: Methodologies and challenges." In: *Psychological methods* 21.4 (2016), p. 507.

[16] Michal Kosinski, David Stillwell, and Thore Graepel. "Private traits and attributes are predictable from digital records of human behavior". In: *Proceedings of the national academy of sciences* 110.15 (2013), pp. 5802–5805.

[17]   Ana Carolina ES Lima and Leandro N de Castro. "Predicting temperament from Twitter data". In: *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*. IEEE. 2016, pp. 599–604.

[18]   Fei Liu, Julien Perez, and Scott Nowson. "A language-independent and compositional model for personality trait recognition from short texts". In: *arXiv preprint arXiv:1610.04345* (2016).

[19]   Robert R McCrae and Paul T Costa. "Validation of the five-factor model of personality across instruments and observers." In: *Journal of personality and social psychology* 52.1 (1987), p. 81.

[20]   Robert R McCrae and Oliver P John. "An introduction to the five-factor model and its applications". In: *Journal of personality* 60.2 (1992), pp. 175–215.

[21]   Clifford Nass and Kwan Min Lee. "Does computer-generated speech manifest personality? An experimental test of similarity-attraction". In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 2000, pp. 329–336.

[22]   Wendy Patton and Mary McMahon. *Career development and systems theory: Connecting theory and practice*. Vol. 2. Springer, 2014.

[23]   James W Pennebaker and Laura A King. "Linguistic styles: Language use as an individual difference." In: *Journal of personality and social psychology* 77.6 (1999).

[24]   Barbara Plank and Dirk Hovy. "Personality Traits on Twitter—Or—How to Get 1,500 Personality Tests in a Week". In: *The 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA), EMNLP 2015*. 2015.

[25]   Daniele Quercia et al. "Our twitter profiles, our selves: Predicting personality with twitter". In: *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE. 2011, pp. 180–185.

[26]   Xianzhi Ruan, Steven Wilson, and Rada Mihalcea. "Finding optimists and pessimists on twitter". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2016, pp. 320–325.

[27]   Youngseo Son et al. "Recognizing counterfactual thinking in social media texts". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2017, pp. 654–658.

[28]   Christian Torrero, Carlo Caprini, and Daniele Miorandi. "A Wikipedia-based approach to profiling activities on social media". In: *arXiv preprint arXiv:1804.02245* (2018).

[29]   Ben Verhoeven, Walter Daelemans, and Barbara Plank. "Twisty: a multilingual twitter stylometry corpus for gender and personality profiling". In: *Proceedings of the 10th Annual Conference on Language Resources and Evaluation (LREC 2016)/Calzolari, Nicoletta [edit.]; et al.* 2016, pp. 1–6.

[30]   Wu Youyou, Michal Kosinski, and David Stillwell. "Computer-based personality judgments are more accurate than those made by humans". In: *Proceedings of the National Academy of Sciences* 112.4 (2015), pp. 1036–1040.

[31]   Mohammadzaman Zamani, Anneke Buffone, and H Andrew Schwartz. "Predicting Human Trustfulness from Facebook Language". In: *arXiv preprint arXiv:1808.05668* (2018).