Profilio: Psychometric Profiling to Boost Social Media Advertising

Fabio Celli
Profilio Company
Via Sommarive 18
Trento, Italy 38123
fabio.celli@profilio.co

Pietro Zani Massani
Profilio Company
Via Sommarive 18
Trento, Italy 38123
pietro.zani@profilio.co

Bruno Lepri
FBK and Profilio Company
Via Sommarive 18
Trento, Italy 38123
lepri@fbk.eu

ABSTRACT

Profilio Company is a startup in its early business development stage that has developed a profiling solution for the field of paid social media advertising. In particular, the solution is designed for the enrichment of Customer Relationship Managements data and for segmentation of customer audiences. Three different Proof of Concepts with different clients have showed that the solution reduces the costs of paid social media advertising in different settings and with different advertising targets, especially starting from large audiences. In this paper we report the details about Profilio's business idea, the development of Profilio's technologies and the results of the Proof of Concepts.

CCS CONCEPTS

•**Human-centered computing** →**User models;** *Collaborative and social computing;*

KEYWORDS

User Profiling, Personality Computing, Social Media, Advertising, CRMs

ACM Reference format:
Fabio Celli, Pietro Zani Massani, and Bruno Lepri. 2017. Profilio: Psychometric Profiling to Boost Social Media Advertising. In *Proceedings of MM'17, October 23–27, 2017, Mountain View, CA, USA.*, , 5 pages.
DOI: https://doi.org/10.1145/3123266.3129311

1 MOTIVATION AND BACKGROUND

Popular social media platforms such as Facebook, Twitter, LinkedIn and YouTube are more and more offering novel ways to advertise brands. For example, Facebook provides to the advertisers options such as promoted posts, sponsored stories, page post ads, etc. Moreover, Facebook, Twitter and LinkedIn have developed a targeting technology which allows advertisers to reach a specific audience (e.g. males vs. females, people with specific interests or living in specific places, different age groups, etc.). These advanced targeting options provide a level of personalization not achievable on other advertising channels (e.g. TV, newspapers, websites, etc.). For example, advertisers may reach specific audiences by looking at their self-reported interests, skills, specific pages they are engaged with, and so on. The interest targeted by this strategy, called *interest targeting*, can be as general as an industry (e.g. fashion industry) or as specific as a product (e.g. sunglasses). Currently, Facebook, Twitter, and LinkedIn are offering this targeting strategy.

Other examples of targeting strategies are *behavioral targeting*, where an advertiser can reach specific customers based on their purchase behaviors [22, 31] and *connection targeting*, where an advertiser can reach people who have a specific kind of connection to a given page, app, event or group. Both types of targeting, currently offered by Facebook, LinkedIn and Twitter, take customers' past behavior into account.

Finally, Facebook and Twitter offer *custom targeting*, that permits advertisers to reach specific audiences by uploading a list of email addresses, phone numbers, usernames or users IDs[1]; while Facebook and LinkedIn offer *lookalike targeting*, a strategy that permits to reach new people who are similar to an audience of interest. More specifically, custom targeting permits to an advertiser to upload and target directly an already known group of people. Instead, lookalike targeting helps companies extend their custom audiences to reach new, similar users. Thus, for those businesses looking to acquire new customers through social media advertising, lookalike targeting may work as a very effective acquisition tool.

Given all these targeting strategies, companies have doubled social media advertising budgets worldwide over the past 2 years, going from $16 billions in 2014 to $31 billions in 2016 [3]. Specifically, paid social media advertising is primarily used to support branding-related efforts, such as the consumer's ability to recognize or recall a brand, *brand awareness*, that is central to purchasing decision making [14, 17]. Let us take as example the fashion industry. According to the McKinsey Global Fashion Index, this industry is now worth about $2.4 trillion; however 2016 was one of the fashion industry's toughest years [4]. This is due to many factors such as the competition of the emerging markets like China and India, the stagnating economies in the Western countries, the increasing volatility and speed of the market, and the growth of athletic wear. At the same time, consumers have become more demanding and less predictable in their purchasing behavior, and this translates into higher advertising costs for brand awareness. Thus, this situation requires more powerful technological solutions for profiling and targeting consumers.

[1]Facebook calls its own audiences *custom audiences*, while Twitter calls its own ones *tailored audiences*

MM'17, October 23–27, 2017, Mountain View, CA, USA.
 ISBN 978-1-4503-4906-2/17/10…$15.00
DOI: https://doi.org/10.1145/3123266.3129311

Crucially, consumers are nowadays leaving a lot of digital footprints (e.g. Facebook profile pictures, tweets, Facebook likes and statuses, pictures on Instagram, videos on YouTube, networks of acquaintances on LinkedIn and of friends on Facebook, etc.) that can be exploited to analyze and predict their behaviors, interests, political and sexual orientations, and psychometric characteristics (e.g. personality traits, emotional dispositions, etc.). Profilio Company, launched at beginning of March 2017, has therefore developed a profiling technology, based on these different (multimodal) sources of digital footprints (e.g. images, textual contents, demographics, social media activities, etc.), that is able to predict a set of behavioral variables, including purchase motivations, job performance and subjective well-being. There are many potential applications of this technology, ranging from credit scoring to human resources, but we aim to apply it to the market of paid social media advertising, starting vertically from the field of fashion and then expanding to other fields. In Section 2 we describe our business idea, our technology and the science behind, in Section 3 we report some results obtained with our current customers and in Section 4 we draw our conclusions and our perspective for the future.

2 BUSINESS IDEA

Profilio Company is the result of years of research on psychometric profiling and automatic behavior understanding from multimodal sources of data (e.g. text, images, social media activities, ego-network characteristics, etc.). Specifically, Profilio Company offers a technological solution that responds to the growing demand of understanding and predicting consumers' behaviors. Starting from public digital footprints, like social media profile pictures and text, and a growing number of other data sources including demographics, Profilio Company's psychometric engine predicts a set of behavioral dimensions such as:

- purchase behavior (attention to advertising, impulsive or compulsive buyers, high spending customers),
- purchase motivations (probability to seek for a sense of belonging to a brand or to display a status),
- cultural attitude (attitude towards innovation or conservation),
- job performance (ability to manage stress, individual or group task proficiency, leading ability),
- subjective and physical well-being (life satisfaction, health probability),
- relationships (anxiety, avoidance and relationship quality),
- news sharing (probability to share news with negative mood or to share fake news),
- music preference (probability to like complex, dance, rebel or conventional music),
- tourism attitude (probability to seek adventure in travel and to experience satisfaction from hospitality),
- social attitude (probability to have pro-social or antisocial behavior, attractiveness and social views).

We deliver this technological solution in a scalable way by means of APIs and we provide a profiling service to segment and create custom audiences. The value of our technological solution lies in:

- the enrichment of Customer Relationship Managements' (CRM) customer data, that can be exploited for business intelligence (i.e. customer segmentation),
- the production of custom audiences that reduce the costs of paid social media advertising.

The development of this solution and the business idea is based on two pillars: a solid scientific work behind our technology and a competitive advantage.

2.1 The Science Behind our Solution

Our technology is based on well-established theories in social psychology and on a decade of research in the fields of personality computing and human behavior understanding.

Research in social psychology has proposed and validated theories to predict and explain individual behaviors and preferences with psychological models, such as the Myers Briggs model [19]. For our solution we adopted the most widely accepted personality model in the scientific community: the Five Factor Model, that defines five traits, namely Openness, Conscientiousness, Extroversion, Agreeableness and Emotional Stability (often conversely referred to as Neuroticism) [10].

Research made great progresses in recent years by exploiting big data of digital footprints from social media as an excellent ground for personality computing and human behavior understanding. For example, nowadays it is possible to train machine learning models to predict personality types from Twitter, Facebook or Linkedin, using features such as number of followers, density of subject's network, number of hashtags, Facebook Likes and other language independent features [9, 11, 12, 15, 16, 21, 23, 29].

An interesting finding is that the computational models based on the subjects' interests are significantly more correlated with personality self-ratings than average human judgments [33]. Sometimes the availability of these features could be subject to limitations due to privacy settings, but more recent work showed how it is possible to predict personality types from public profile pictures [8, 26] exploiting techniques such as Bag-of-Visual-Words, Convolutional Neural Networks or other low level features [7, 27, 32], that are able to capture information from technical variables of images such as pixel patterns, color scales or variations in brightness. Profile pictures (not necessarily faces) convey a lot of information about a user and, according to literature in psychology, are directly connected to their identity [2, 30]. For example, results in this field revealed that extroverted and emotionally stable people tend to have pictures where they are smiling and appear with other people; introverts tend to appear alone and with less bright colors; neurotics tend to have images with close-up faces or without humans and uncooperative people tend to have pictures with few colors.

In addition to these findings there is a rich scientific literature about behavioral dimensions correlated to personality, such as socioeconomic status, life satisfaction, job performance, relationship quality and cognitive ability [13, 20, 24]. The accuracy of the system, evaluated with training/test split and backtests for all the predicted dimensions, is around 80%.

Exploiting data from Facebook and Twitter collected since 2014 within a research project [8], we put together all these scientific findings in a customized system involving semi-supervised [34] and

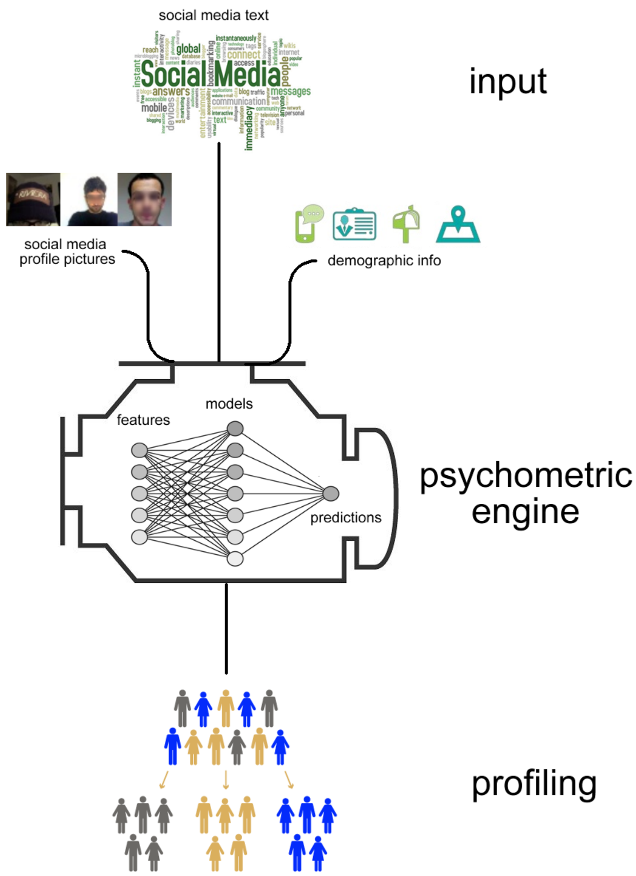

Figure 1: A schematic representation of Profiling technology

Deep Learning techniques (e.g. Convolutional Neural Networks) [5], beside other machine learning algorithms such as Support Vector Machine regressors [28] and enhanced rule-based Ripper algorithms [25]. A representation of the system is depicted in Figure 1.

The system can take as input social media profile pictures (blurred in Figure 1 for privacy reasons), text or demographic information. A first level of technical feature extraction feeds several different layers of models that output the final predictions and delivers a user profile. For business reasons we can not disclose other information about the system.

As our technological solution produces predictions of individual psychometric traits, we must be compliant on two main legal aspects:

- User consent on data treatment. For this reason, we use exclusively public data as sources for our predictions (e.g. profile pictures on Facebook, textual content on Twitter, etc.) or data owned/authorized by our clients (e.g. email addresses and demographics), thus they bear the responsibility for that data;
- Discrimination. For this reason, our technological solution avoids to produce discriminative predictions regarding political views, religious views, and sexual orientations.

2.2 Market, Competitors and Competitive Advantage

Profilio's reference market is divided into two main segments:

- sales of data enrichment services for Customer Relationship Managements (CRMs) of e-commerce portals (examples of target in the field of fashion are Zalando and Yoox-Net-A-Porter) and large consulting firms (e.g. Bain & Company, McKinsey, etc.)
- sales of custom audiences profiling for paid social media advertising to large groups in the field of fashion and luxury (Kering holding who owns Gucci, Puma, and Yves Saint Laurent among the other brands, Prada group, AEFFE group, etc.)

The business model envisages two sources of income linked to each other:

- income from sales of profiling services for data enrichment and custom audiences' generation,
- income from research projects finalized to improve the profiling technologies.

The solutions available from the major global competitors are general purpose products (with the exception of Cambridge Analytica that is focused on electoral processes) and mainly designed for the US market. Apply Magic Sauce[2] is a company that provides an API service predicting personality and other personal dimensions from Likes and text in social media, sold with two types of subscriptions: basic ($ 500 / month) and pro ($ 3000 / month). Cambridge Analytica[3] is a company that combines data mining and psychometric profiling with strategic communication for electoral processes. The company is heavily funded by the family of Robert Mercer, an American hedge-fund billionaire [1]. In 2015 it became known as the data analysis company working for Ted Cruz and Donald Trump's presidential campaigns. IBM Watson Personality Insights[4] is a personality prediction service (metrics and summary on personality and other business dimensions) delivered through APIs. Prediction is made from various textual data (email, social media, generic texts) in a limited number of languages at a price of about $ 0.01 per call. Crystal[5] is a service that predicts 4 dimensions of digital footprint personality with the DISC (Dominance, Inducement, Submission, and Compliance) model [18] and specializes in the sales sectors (predicts the most suitable loyalty to convince the customer based on the personality profile) and hiring (predicts the characteristics of candidates in the human resources sector) starting from LinkedIn, email and calendar data. Pricing and service delivery are customized.

In the reference market, knowledge acquisition from digital footprints is hampered by a number of factors:

- social media privacy settings, that prevent the collection of some private kind of data, such as text and likes. The

[2] http://applymagicsauce.com/
[3] https://cambridgeanalytica.org
[4] https://watson-pi-demo.mybluemix.net
[5] https://www.crystalknows.com

solution offered by Profilio Company includes profiling from public sources like profile pictures.
- limitations in the availability of data sources: many large companies, even in the field of luxury and fashion, have only demographic data of their customers in their CRMs, and this prevents the application of many profiling models; Profilio Company offers a solution that is domain-adaptive and is constantly concentrating the efforts towards enlarging the number of data sources that can be profiled.
- language limitations, that is especially important in a market like EU. Profilio's technology is language independent and can be applied in different language areas.

These factors, plus the fact that the application of these technologies in the field of fashion is very recent, represent a competitive advantage for Profilio Company.

3 CASE HISTORY AND RESULTS

Since March 2017, Profilio Company has ran three different Proof of Concepts (PoC) with different clients, (i) a big fashion group, (ii) a small startup promoting a discount app and (iii) a medium e-commerce of fashion, design and food & wine. In each PoC we tested the results of same advertising copyright with different audiences in a span of time of one week.

Table 1: results of three different Proof of Concepts with three different clients in a time span of one week.

PoC 1: big fashion group, target: video brand awareness					
audience	lookalike	country	CPV	CTR	Avg%ViewTime
fb fans	3%	UK	0.14	1%	16.6%
purchase	3%	UK	0.04	24.4%	46.5%
Profilio 1	3%	UK	0.04	24.6%	48.6%
PoC 2: small startup, target: app install					
audience	lookalike	country	CPA	CPuC	CPShare
fb fans	3%	IT	0.46	0.18	26.1
Profilio 2	3%	IT	0.43	0.16	19.7
PoC 3: e-commerce, target: lead generation					
audience	lookalike	country	CPL	CPuC	CPShare
fb fans	1%	IT	1.12	0.18	115.8
subscribe	1%	IT	1.02	0.26	262
purchase	1%	IT	0.89	0.23	197.7
Profilio 3	1%	IT	0.77	0.17	76.5

The results reported in Table 1 refer to the three clients in a time span of one week. The target is different for each client: in PoC 1 (big fashion group) the target is brand awareness with a 15 seconds viral videoclip. As evaluation metrics we used Cost Per 3 seconds View (CPV), Click Through Rate (CTR) and average percentage of videoclip view time (Avg%ViewTime). Starting from the fashion group's Facebook fans, we created a custom audience (Profilio 1) with users that our system predicted having high attention to advertising. We compared this audience against two control groups: (i) the Facebook fanbase of the fashion group and (ii) an audience with users that purchased items of the fashion group. In order to have audiences comparable also by potential reach, we created Facebook lookalike of all the audiences in order to reach 1.4M potential customers in UK with each audience. Our results show that Profilio Company's profiling solution helps reducing CPV and improving CTR and the average percentage of videoclip view time. It is interesting to note that profiling from a very large Facebook fanbase obtains better results than using an audience that purchased brand's items, at the same cost per view. In PoC 2 (small startup) the target is an App install action. Again, we created a custom audience (Profilio 2) starting from the startup's Facebook fans, profiling users with high attention to advertising. We compared the results against a control group with the startup Facebook fans, creating lookalike audiences in order to have a comparable potential reach (150K potential customers). Our results show that the profiled audience reduces the Cost per App Install (CPA), the Cost Per unique Clicks (CPuC) and the Cost Per Shares. Finally, in PoC 3 (medium e-commerce) the target is lead generation, defined as the registration to the e-commerce website. Hence, we used as evaluation metrics the Cost Per Lead (CPL), the Cost Per unique Clicks (CPuC) and the Cost Per Shares. In this case, we have started creating our profiled audience from a mixture of the e-commerce Facebook fans, Facebook users who reacted in the e-commerce Facebook page and subscribed users. Again, we created lookalike audiences in order to have a comparable potential reach (in this case 300K potential customers) between experimental and control audiences. Results show that our profiling helps reducing CPL, CPuC and CPShare.

4 CONCLUSION AND FUTURE

Profilio Company is a startup in its early business development stage. We have developed a Minimal Viable Product (MVP) with a psychometric engine as its core. The solution, that delivers the enrichment of CRMs' data and the profiling of custom audiences, reduces the costs of paid social media advertising in different settings and with different advertising targets. The Proof Of Concepts we ran with our first customers showed that our profiling solution is effective, especially starting from large audiences. This is one of the reasons why we target large fashion groups as our ideal clients. We have several targets for the future of Profilio Company. First of all, we are working for improving the performance of our psychometric engine by enlarging the number of social media we use for training our models, as scientific literature suggests that rich personas can be extracted by looking at different social media as different points of view on customers' profiles [6]. We are also working to improve machine learning models by leveraging state-of-the-art approaches in deep learning. Another goal is to automate also the generation of advertising strategies customized for the specific psychometric characteristics of our clients' customers. More specifically, we plan to investigate the advertising effectiveness of generating targeted multimodal (textual, visual and audio) messages. From a business point of view, we are working to find new clients, to explore new markets (e.g. from Italy to Europe and then United States), and to find investors eager in taking the company to the next level. As people are producing always larger samples of digital footprints and privacy becomes more and more a serious issue, predictive profiling will become more and more useful for business companies. Profilio Company is working in this direction, with the mission to automatically scouting the value of people behaviors, and turn it into knowledge.

REFERENCES

[1] 2015. Cruz-Connected Data Miner Aims to Get Inside U.S. Voters' Heads. https://www.bloomberg.com/politics/features/2015-11-12/is-the-republican-party-s-killer-data-app-for-real-. (2015). Retrieved: February 2016.

[2] 2016. *Analyzing Personality through Social Media Profile Picture Choice.*

[3] 2016. Social media advertising usage statistics. https://blog.hootsuite.com/social-media-advertising-stats. (2016). Retrieved: May 2017.

[4] 2017. The state of fashion. http://www.mckinsey.com/industries/retail/our-insights/the-state-of-fashion. (2017). Retrieved: May 2017.

[5] Yoshua Bengio. 2009. Learning deep architectures for AI. *Foundations and trends® in Machine Learning* 2, 1 (2009), 1–127.

[6] Kseniya Buraya, Aleksandr Farseev, Andrey Filchenkov, and Tat-Seng Chua. 2017. Towards user personality profiling from multiple social networks. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence.*

[7] Fabio Celli. 2015. *IATO: Feature Extraction for image analysis.* Technical Report. University of Trento.

[8] Fabio Celli, Elia Bruni, and Bruno Lepri. 2014. Automatic Personality and Interaction Style Recognition from Facebook Profile Pictures. In *Proceedings of the ACM International Conference on Multimedia.* ACM, 1101–1104.

[9] Fabio Celli and Luca Rossi. 2012. The role of emotional stability in Twitter conversations. In *Proceedings of the Workshop on Semantic Analysis in Social Media.* Association for Computational Linguistics, 10–17.

[10] Paul T Costa and Robert R McCrae. 2008. The Revised NEO Personality Inventory (NEO-PI-R). *In G.J. Boyle, G Matthews and D. Saklofske (Eds.). The SAGE handbook of personality theory and assessment* 2 (2008), 179–198.

[11] Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. 2011. Predicting personality from twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.* IEEE, 149–156.

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

[13] Ashley S. Holland and Glenn I. Roisman. 2008. Big Five personality traits and relationship quality: Self-reported, observational, and physiological evidence. *Journal of Social and Personal Relationships* 25, 5 (2008), 811–829.

[14] Wayne D. Hoyer and Steven P. Brown. 1990. Effects of brand awareness on choice for a common, repeat purchase product. *Journal of Consumer Research* 17 (1990), 141–148.

[15] Michal Kosinski, Yoram Bachrach, Pushmeet Kohli, David Stillwell, and Thore Graepel. 2014. Manifestations of user personality in website choice and behaviour on online social networks. *Machine learning* 95, 3 (2014), 357–380.

[16] Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110, 15 (2013), 5802–5805.

[17] Emma D. Macdonald and Byron M. Sharp. 2000. Brand awareness effects on consumer decision making for a common, repeat purchase product: A replication. *Journal of Business Research* 48 (2000), 5–15.

[18] Wilhelm Marston. 1928. *Emotions of normal people.* Kegan Paul Trench, Trubner & Co. Ltd.

[19] Isabel Briggs Myers and Peter B Myers. 2010. *Gifts differing: Understanding personality type.* Davies-Black Publishing.

[20] Andrew Neal, Gillian Yeo, Annette Koy, and Tania Xiao. 2012. Predicting the form and direction of work role performance from the Big 5 model of personality traits. *Journal of Organizational Behavior* 33, 2 (2012), 175–192. https://doi.org/10.1002/job.742

[21] Thin Nguyen, Dinh Phung, Brett Adams, and Svetha Venkatesh. 2011. Towards discovery of influence and personality traits through social link prediction. In *Proceedings of the International AAAI Conference on Weblogs and Social Media.* 566–569.

[22] Sandeep Pandey, Mohamed Aly, Abraham Bagherjeiran, Andrew Hatch, Peter Ciccolo, Adwait Ratnaparkhi, and Martin Zinkevich. 2011. Learning to target: What works for behavioral targeting. In *Proceedings of the 20th Conference on Information and Knowledge Management (CIKM).* ACM, 1805–1814.

[23] Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. 2011. Our Twitter profiles, our selves: Predicting personality with Twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.* IEEE, 180–185.

[24] Brent W. Roberts, Nathan R. Kuncel, Rebecca Shiner, Avshalom Caspi, and Lewis R. Goldberg. 2007. The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science* 2, 4 (2007), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

[25] Bakhtawar Seerat and Usman Qamar. 2015. Rule induction using enhanced RIPPER algorithm for clinical decision support system. In *Intelligent Control and Information Processing (ICICIP), 2015 Sixth International Conference on.* IEEE, 83–91.

[26] Cristina Segalin, Fabio Celli, Luca Polonio, Michal Kosinski, David Stillwell, Nicu Sebe, Marco Cristani, and Bruno Lepri. 2017. What your Facebook profile picture reveals about your personality. In *Proceedings of the International Conference on ACM Multimedia (ACM-MM).* ACM.

[27] Cristina Segalin, Dong Seon Cheng, and Marco Cristani. 2017. Social Profiling Through Image Understanding. *Comput. Vis. Image Underst.* 156, C (March 2017), 34–50. https://doi.org/10.1016/j.cviu.2016.10.013

[28] Shirish Krishnaj Shevade, S Sathiya Keerthi, Chiranjib Bhattacharyya, and Karaturi Radha Krishna Murthy. 2000. Improvements to the SMO algorithm for SVM regression. *Neural Networks, IEEE Transactions on* 11, 5 (2000), 1188–1193.

[29] Jacopo Staiano, Bruno Lepri, Nadav Aharony, Fabio Pianesi, Nicu Sebe, and Alex Pentland. 2012. Friends don't lie: inferring personality traits from social network structure. In *Proceedings of the 2012 ACM conference on ubiquitous computing.* ACM, 321–330.

[30] Simine Vazire and Samuel D Gosling. 2004. e-Perceptions: personality impressions based on personal websites. *Journal of Personality and Social Psychology* 87, 1 (2004), 123.

[31] Ju Yan, Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen. 2009. How much can behavioral targeting help online advertising?. In *Proceedings of the 18th International Conference on World Wide Web (WWW).* ACM, 261–270.

[32] Jun Yang, Yu G. Jiang, Alexander G. Hauptmann, and Chong W. Ngo. 2007. Evaluating bag-of-visual-words representations in scene classification. In *MIR '07: Proceedings of the international workshop on Workshop on multimedia information retrieval.* ACM, New York, NY, USA, 197–206. https://doi.org/10.1145/1290082.1290111

[33] Wu Youyou, Michal Kosinski, and David Stillwell. 2015. Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences* 112, 4 (2015), 1036–1040.

[34] Xiaojin Zhu. 2006. Semi-supervised learning literature survey. *Computer Science, University of Wisconsin-Madison* 2 (2006), 3.