

and determine four levels of consistency. We discover companies which post consistently and find that more wealthy companies are better at maintaining consistency.

Limitations: We only consider textual content of web articles posted by the companies themselves and do not cover any form of user-generated content regarding the companies. Another important limitation is that we do not cover the visual or content-independent aspects of a brand style guide like color, typography and positioning of different sections and headers of a brand website.

Future Work: In the current work, we develop independent classifiers for each trait. However, it might be possible that one trait (weakly) implies one or more of the others; thus jointly learning all the traits together would be an important future work. We will use the insights derived from this work dealing with document-level text classification and move onto finer granularity like sentence level and identify the most contributing sentences towards the expression of a brand. We will further extend our work to develop a helper tool for the content writers and brand managers, which sentences should be modified for making the text articles more consistent with the targeted brand personality. These would be our next future endeavors.

ACKNOWLEDGMENTS

The first author is currently supported by the Junior Research fellowship from the project titled "IoTDiff: Service Differentiation for Personal Hubs", which is sponsored by Science & Engineering Research Board, New Delhi. A part of the work was carried out during the summer internship at Big Data Experience Lab, Adobe Research, Bangalore, India.

REFERENCES

- [1] Jennifer L. Aaker. 1997. Dimensions of brand personality. *Journal of marketing research* (1997), 347–356.
- [2] Pankaj Aggarwal. 2004. The effects of brand relationship norms on consumer attitudes and behavior. *Journal of consumer research* 31, 1 (2004), 87–101.
- [3] Enrique Amigó, Jorge Carrillo-de Albornoz, Irina Chugur, Adolfo Corujo, Julio Gonzalo, Edgar Meij, Maarten de Rijke, and Damiano Spina. 2014. Overview of RepLab 2014: Author Profiling and Reputation Dimensions for Online Reputation Management. In *Information Access Evaluation. Multilinguality, Multimodality, and Interaction*, Evangelos Kanoulas, Mihai Lupu, Paul Clough, Mark Sanderson, Mark Hall, Allan Hanbury, and Elaine Toms (Eds.). Springer International Publishing, Cham, 307–322.
- [4] Enrique Amigó, Jorge Carrillo-de Albornoz, Irina Chugur, Adolfo Corujo, Julio Gonzalo, Edgar Meij, Maarten de Rijke, and Damiano Spina. 2014. Overview of RepLab 2014: Author Profiling and Reputation Dimensions for Online Reputation Management. In *Information Access Evaluation. Multilinguality, Multimodality, and Interaction*, Evangelos Kanoulas, Mihai Lupu, Paul Clough, Mark Sanderson, Mark Hall, Allan Hanbury, and Elaine Toms (Eds.). Springer International Publishing, Cham, 307–322.
- [5] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Int. Res.* 16, 1 (June 2002), 321–357. <http://dl.acm.org/citation.cfm?id=1622407.1622416>
- [6] Judy Delin. 2007. Brand Tone of Voice. *Journal of Applied Linguistics* 2, 1 (2007). <https://journals.equinoxpub.com/index.php/JAL/article/view/1471>
- [7] Alecia Douglas, Juline Mills, and Raphael Kavanaugh. 2007. Exploring the use of emotional features at romantic destination websites. *Information and Communication Technologies in Tourism 2007* (2007), 331–340.
- [8] J Peter Kincaid, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. 1975. *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. Technical Report. DTIC Document.
- [9] Shibamouli Lahiri. 2015. SQUINKY! A Corpus of Sentence-level Formality, Informativeness, and Implicature. *arXiv preprint arXiv:1506.02306* (2015).
- [10] Wang Ling, Tiago Luis, Luis Marujo, Ramón Fernández Astudillo, Silvio Amir, Chris Dyer, Alan W Black, and Isabel Trancoso. 2015. Finding function in form: Compositional character models for open vocabulary word representation. *arXiv preprint arXiv:1508.02096* (2015).
- [11] Fei Liu, Julien Perez, and Scott Nowson. 2017. A Language-independent and Compositional Model for Personality Trait Recognition from Short Texts. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, Vol. 1, 754–764.
- [12] Liu Liu, Daria Dzyabura, and Natalie Mizik. 2018. Visual listening in: Extracting brand image portrayed on social media. In *Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence*.
- [13] X. Liu, A. Xu, L. Gou, H. Liu, R. Akkiraju, and H. Shen. 2016. SocialBrands: Visual analysis of public perceptions of brands on social media. In *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 71–80. <https://doi.org/10.1109/VAST.2016.7883513>
- [14] Zhe Liu, Anbang Xu, Yi Wang, Jerald Schoudt, Jalal Mahmud, and Rama Akkiraju. 2017. Does Personality Matter?: A Study of Personality and Situational Effects on Consumer Behavior. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media (HT '17)*. ACM, New York, NY, USA, 185–193. <https://doi.org/10.1145/3078714.3078733>
- [15] Natalia Maehle, Cele Otnes, and Magne Supphellen. 2011. Consumers' perceptions of the dimensions of brand personality. *Journal of Consumer Behaviour* 10, 5 (2011), 290–303.
- [16] François Mairesse, Marilyn A Walker, Matthias R Mehl, and Roger K Moore. 2007. Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of artificial intelligence research* 30 (2007), 457–500.
- [17] Brigitte Müller and Jean-Louis Chandon. 2003. The impact of visiting a brand website on brand personality. *Electronic Markets* 13, 3 (2003), 210–221.
- [18] Ellie Pavlick and Joel Tetreault. 2016. An empirical analysis of formality in online communication. *Transactions of the Association of Computational Linguistics* 4, 1 (2016), 61–74.
- [19] Joseph T. Plummer. 2000. How Personality Makes a Difference. *Journal of Advertising Research* 40, 6 (2000), 79–83. <https://doi.org/10.2501/JAR-40-6-79-83> arXiv:<http://www.journalofadvertisingresearch.com/content/40/6/79.full.pdf>
- [20] Damiano Spina, Julio Gonzalo, and Enrique Amigó. 2014. Learning Similarity Functions for Topic Detection in Online Reputation Monitoring. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '14)*. ACM, New York, NY, USA, 527–536. <https://doi.org/10.1145/2600428.2609621>
- [21] Ming-Hsiang Su, Chung-Hsien Wu, and Yu-Ting Zheng. 2016. Exploiting turn-taking temporal evolution for personality trait perception in dyadic conversations. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24, 4 (2016), 733–744.
- [22] X. Sun, B. Liu, J. Cao, J. Luo, and X. Shen. 2018. Who Am I? Personality Detection Based on Deep Learning for Texts. In *2018 IEEE International Conference on Communications (ICC)*, 1–6. <https://doi.org/10.1109/ICC.2018.8422105>
- [23] Yla R. Tausczik and James W. Pennebaker. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. (2010). <http://homepage.psy.utexas.edu/homepage/students/Tausczik/Yla/index.html>
- [24] Jing Wang and Min-Ling Zhang. 2018. Towards Mitigating the Class-Imbalance Problem for Partial Label Learning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*. 2427–2436. <https://doi.org/10.1145/3219819.3220008>
- [25] Honghao Wei, Fuzheng Zhang, Nicholas Jing Yuan, Chuan Cao, Hao Fu, Xing Xie, Yong Rui, and Wei-Ying Ma. 2017. Beyond the Words: Predicting User Personality from Heterogeneous Information. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (WSDM '17)*. ACM, New York, NY, USA, 305–314. <https://doi.org/10.1145/3018661.3018717>
- [26] Anbang Xu and Brian Bailey. 2012. What do you think?: a case study of benefit, expectation, and interaction in a large online critique community. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 295–304.
- [27] Anbang Xu, Haibin Liu, Liang Gou, Rama Akkiraju, Jalal Mahmud, Vibha Sinha, Yuheng Hu, and Mu Qiao. 2016. Predicting Perceived Brand Personality with Social Media. In *Proceedings of the Tenth International Conference on Web and Social Media, Cologne, Germany, May 17-20, 2016*. 436–445. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13078>
- [28] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1480–1489.