

Bibliography

- [1] Gang Wang, Santa Barbara, Tianyi Wang, Haitao Zheng, and Ben Y Zhao. Man vs . Machine : Practical Adversarial Detection of Malicious Crowdsourcing Workers. *the 23rd USENIX Security Symposium*, pages 239–254, 2014. URL <https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/wang>.
- [2] Michael Dewing. Social Media: An Introduction. Publicatio(Report):1–5, 2010. URL <https://goo.gl/AJNc09>.
- [3] Jan H. Kietzmann, Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3):241–251, 2011. ISSN 00076813. doi: 10.1016/j.bushor.2011.01.005.
- [4] Maeve Duggan. Mobile Messaging and Social Media 2015. (August):20, 2015. URL <http://www.pewinternet.org/files/2015/08/Social-Media-Update-2015-FINAL2.pdf>.
- [5] Dave Chaffey and D Chaffey. Global social media research summary 2016. *Smart Insights: Actionable Marketing Advice*, 2015. URL <http://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/>.
- [6] Thomas W Malone, Robert Laubacher, and Chrysanthos Dellarocas. Harnessing crowds : Mapping the genome of collective intelligence. *MIT Sloan School of Management*, 1:1–20, 2009. ISSN 15565068. doi: 10.2139/ssrn.1381502.
- [7] Thomas W. Malone. What Makes Things Fun to Learn? A Study of Intrinsically Motivating Computer Games, 1981. URL <http://eric.ed.gov/?id=EJ255512>.

- [8] P.M. Clifford. The international vocabulary of basic and general terms in metrology. *Measurement*, 3(2):72–76, 1985. ISSN 02632241. doi: 10.1016/0263-2241(85)90006-5. URL http://www.bipm.org/utils/common/documents/jcgm/JCGM{}_200{}_2008.pdf.
- [9] Alvin Enis House, Betty J. House, and Martha B. Campbell. Measures of interobserver agreement: Calculation formulas and distribution effects, 1981. ISSN 01640305.
- [10] Sotiris B Kotsiantis. Supervised machine learning: A review of classification techniques. *Informatica*, 31:249–268, 2007. ISSN 09226389. doi: 10.1115/1.1559160. URL <https://goo.gl/02Nd0i>.
- [11] Kuldeep Yadav, Ponnuram Kumaraguru, Atul Goyal, Ashish Gupta, and Vinayak Naik. SMSAssassin : Crowdsourcing Driven Mobile-based System for SMS Spam Filtering. *HOTMOBILE 2011 Phoenix*, pages 1–6, 2011. doi: 10.1145/2184489.2184491.
- [12] Twitter. Reporting Spam on Twitter., 2016. URL <https://support.twitter.com/articles/64986?lang=en{#}>.
- [13] Nexgate. State of Social Media Spam. 2013. ISSN 09645993. doi: 10.1016/S0022-3913(12)00047-9.
- [14] Networked Insights. How Dirty is Big Data ? pages 1–17, 2015. URL <http://info.networkedinsights.com/Dirty-Data-LP.html>.
- [15] Z Chu, I Widjaja, and H Wang. Detecting social spam campaigns on twitter. *Applied Cryptography and Network Security*, 2012. URL http://link.springer.com/10.1007/978-3-642-31284-7{\protect\T1\textbraceleft}{}_{{\protect\T1\textbraceright}}27http://link.springer.com/10.1007/978-3-642-31284-7{}_27.
- [16] Amazon. Amazon Mechanical Turk, 2012. URL <https://www.mturk.com/mturk/welcome>.

- [17] M. Buhrmester, T. Kwang, and S. D. Gosling. Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *on psychological science*, 6(1):3-5, jan 2011. ISSN 1745-6916. doi: 10.1177/1745691610393980. URL <https://www.mturk.com/mturk/welcome><http://pps.sagepub.com/lookup/doi/10.1177/1745691610393980><http://pps.sagepub.com/content/6/1/3.short>.
- [18] L. Von Ahn. Human computation. In *2009 46th ACM/IEEE Design Automation Conference*, pages 418-419, New York, New York, USA, 2009. ACM Press. ISBN 978-1-6055-8497-3. doi: 10.1145/1629911.1630023. URL <http://portal.acm.org/citation.cfm?id=1629911.1630023>.
- [19] Luis von Ahn, Ruoran Liu, and Manuel Blum. Peekaboom. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, page 55, 2006. doi: 10.1145/1124772.1124782. URL <http://dl.acm.org/citation.cfm?id=1124782><http://dl.acm.org/citation.cfm?id=1124772.1124782><http://portal.acm.org/citation.cfm?doid=1124772.1124782>.
- [20] Luis von Ahn. Duolingo. *Proceedings of the 2013 international conference on Intelligent user interfaces - IUI '13*, page 1, 2013. doi: 10.1145/2449396.2449398. URL <http://dl.acm.org/citation.cfm?doid=2449396.2449398><http://dl.acm.org/citation.cfm?id=2449396.2449398>.
- [21] Luis Von Ahn, Mihir Kedia, and Manuel Blum. Verbosity: a game for collecting common-sense facts. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 75-78, 2006. ISSN 978-1-60558-193-4. doi: 10.1145/1124772.1124784. URL <http://dl.acm.org/citation.cfm?id=1124784>.
- [22] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*, 6(1):319-326, 2004. ISSN 00201669. doi: 10.1145/985692.985733. URL <http://portal.acm.org/citation.cfm?doid=985692.985733>.

- [23] Severin Hacker and Luis Von Ahn. Matchin: eliciting user preferences with an online game. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1207–1216, 2009. ISSN 1605582468. doi: 10.1145/1518701.1518882. URL https://www.cs.cmu.edu/~biglou/matchin_chi.pdfhttps://www.cs.cmu.edu/~biglou/matchin_chi.pdf.
- [24] Luis Von Ahn, Shiry Ginosar, Mihir Kedia, Ruoran Liu, and Manuel Blum. Improving accessibility of the web with a computer game. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, 22:79, 2006. doi: 10.1145/1124772.1124785. URL <http://portal.acm.org/citation.cfm?doid=1124772.1124785>.
- [25] Edith Law, Luis Von Ahn, and Tom Mitchell. Search War : A Game for Improving Web Search. *Human Factors*, 25(April 2007):15217, 2009. doi: 10.1145/1600150.1600160.
- [26] Jeffrey P. Bigham, Samual White, Tom Yeh, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, and Brandyn White. VizWiz. In *Proceedings of the 23rd annual ACM symposium on User interface software and technology - UIST '10*, page 333, New York, New York, USA, 2010. ACM Press. ISBN 9781450302715. doi: 10.1145/1866029.1866080. URL <http://portal.acm.org/citation.cfm?doid=1866029.1866080>.
- [27] Alexander J Quinn and Benjamin B Bederson. Human Computation: A Survey and Taxonomy of a Growing Field. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1403–1412, 2011. ISSN 1450302289. doi: 10.1145/1978942.1979148.
- [28] Roman V. Yampolskiy. Turing test as a defining feature of AI-completeness. *Studies in Computational Intelligence*, 427:3–17, 2013. ISSN 1860949X. doi: 10.1007/978-3-642-29694-9-1.

- [29] De Wang, Danesh Irani, and Calton Pu. A social-spam detection framework. *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference on - CEAS '11*, (January):46–54, 2011. doi: 10.1145/2030376.2030382. URL <http://doi.acm.org/10.1145/2030376.2030382>.
- [30] George Forman and Ira Cohen. Learning from Little: Comparison of Classifiers Given Little Training. *Knowledge Discovery in Databases: PKDD 2004*, 19(September):161–172, 2004. ISSN 03029743. doi: 10.1007/978-3-540-30116-5_17. URL [http://www.springerlink.com/index/vlwc9y9quq3895hj.pdf\\$\\delimiter\"026E30F\\$\\nhttp://link.springer.com/10.1007/978-3-540-30116-5_{_}17](http://www.springerlink.com/index/vlwc9y9quq3895hj.pdf$\\delimiter\).
- [31] Paul Graham. A Plan for Spam. *Hackers and Painters: Big Ideas from the Computer Age*, page 271, 2004. URL <http://www.paulgraham.com/spam.html>.
- [32] Xianghan Zheng, Zhipeng Zeng, Zheyi Chen, Yuanlong Yu, and Chunming Rong. Detecting spammers on social networks. *Neurocomputing*, 159:27–34, 2015. ISSN 09252312. doi: 10.1016/j.neucom.2015.02.047.
- [33] Fangzhao Wu, Jinyun Shu, Yongfeng Huang, and Zhigang Yuan. Social Spammer and Spam Message Co-Detection in Microblogging with Social Context Regularization Categories and Subject Descriptors. *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*, pages 1601–1610, 2015. doi: 10.1145/2806416.2806560.
- [34] Kamalanathan Kandasamy and Preethi Korothe. An integrated approach to spam classification on Twitter using URL analysis, natural language processing and machine learning techniques. In *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*, pages 1–5. IEEE, mar 2014. ISBN 978-1-4799-2525-4. doi: 10.1109/SCEECS.2014.6804508. URL [http://ieeexplore.ieee.org/xpls/abs\\$\\protect\\T1\\textbraceleft\\}_{_}\\{\\protect\\T1\\textbraceright\\}all.jsp?arnumber=6804508http://ieeexplore.ieee.org/xpls/abs\\$\\}_{_}all.jsp?arnumber=6804508http://ieeexplore.ieee.org/document/6804508/](http://ieeexplore.ieee.org/xpls/abs$\\protect\\T1\\textbraceleft\\}_{_}\\{\\protect\\T1\\textbraceright\\}all.jsp?arnumber=6804508http://ieeexplore.ieee.org/xpls/abs$\\}_{_}all.jsp?arnumber=6804508http://ieeexplore.ieee.org/document/6804508/).

- [35] Ignacio Garcia. Learning a language for free while translating the web. Does Duolingo work? *International Journal of English Linguistics*, 3(1):19–25, 2013. ISSN 1923869X. doi: 10.5539/ijel.v3n1p19. URL [http://search.proquest.com.myaccess.library.utoronto.ca/docview/1415387618?OpenUrlRefId=info:xri/sid:summon{&}\delimiter"026E30F\\$nhhttp://media.proquest.com.myaccess.library.utoronto.ca/media/pq/classic/doc/3031864731/fmt/pi/rep/NONE?hl={&}cit:auth=Garcia,+Ignacio{&}cit:titl](http://search.proquest.com/myaccess.library.utoronto.ca/docview/1415387618?OpenUrlRefId=info:xri/sid:summon{&}\delimiter).
- [36] M Mccord. Spam Detection on Twitter Using Traditional Classifiers. pages 1–7, 2011.
- [37] Susan Dumais. A Bayesian Approach to Filtering Junk E-Mail. (Cohen), 1996.
- [38] Karl-Michael M Schneider. Techniques for Improving the Performance of Naive Bayes for Text Classification. *Proceedings of CICLing*, pages 682–693, 2005. URL <http://www.phil.uni-passau.de/linguistik/schneider/papers2://publication/uuid/3C12B6FE-3D6B-4FC7-A8AB-5C1A72AD87A7>.
- [39] Justin M Rao and David H Reiley. The Economics of Spam. *Journal of Economic Perspectives*, 26(3):87–110, 2012. ISSN 0895-3309. doi: 10.1257/jep.26.3.87.
- [40] Cristian Lumezanu and Nick Feamster. Observing common spam in tweets and email. *Proceedings of the 2012 ACM conference on Internet measurement conference*, pages 441–466, 2012. doi: 10.1145/2398776.2398824. URL <http://gtnoise.net/papers/2012/lume:imc2012.pdf>.
- [41] E Law and L Ahn. *Human computation*, volume 5. 2011. ISBN 9781608455164. doi: 10.2200/S00371ED1V01Y201107AIM013. URL <http://www.morganclaypool.com/doi/abs/10.2200/S00371ED1V01Y201107AIM013>.
- [42] Edith Law and Luis Von Ahn. Input-Agreement : A New Mechanism for Collecting Data Using Human Computation Games. *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pages 1–10, 2009. doi: 10.1145/1518701.1518881.

- [43] Luis Von Ahn, Benjamin Maurer, Colin Mcmillen, David Abraham, and Manuel Blum. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. *Science*, 321(12 September 2008):1465–1468, sep 2008. ISSN 0036-8075. doi: 10.1126/science.1160379. URL <http://www.sciencemag.org/cgi/doi/10.1126/science.1160379>.
- [44] Dirk Lindebaum. Rhetoric or remedy? A critique on developing emotional intelligence. *Academy of Management Learning and Education*, 8(2):225–237, jan 2009. ISSN 1537260X. doi: 10.1006/ceps.1999.1020. URL <http://linkinghub.elsevier.com/retrieve/pii/S0361476X99910202>.
- [45] Michele D. Dickey. Game design and learning: A conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development*, 55(3):253–273, 2007. ISSN 10421629. doi: 10.1007/s11423-006-9004-7.
- [46] Shawkat Ali and Kate A. Smith. On learning algorithm selection for classification. *Applied Soft Computing*, 6(2):119–138, jan 2006. ISSN 1568-4946. doi: 10.1016/j.asoc.2004.12.002. URL [http://www.sciencedirect.com/science/article/pii/S1568494605000049\\$%5Bdelimiter%5Dhttp://www.sciencedirect.com/science/article/pii/S1568494605000049/pdf?md5=c866bde3670acbe8211f7e5ad80ff9a0%7B%7Dpid=1-s2.0-S1568494605000049-main.pdf](http://www.sciencedirect.com/science/article/pii/S1568494605000049$%5Bdelimiter%5Dhttp://www.sciencedirect.com/science/article/pii/S1568494605000049/pdf?md5=c866bde3670acbe8211f7e5ad80ff9a0%7B%7Dpid=1-s2.0-S1568494605000049-main.pdf).
- [47] Sida Wang and Christopher D Manning. Baselines and bigrams: Simple, good sentiment and topic classification. *ACL 12 Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers - Volume 2*, 2(1): 90–94, 2012.
- [48] D Manning Christopher, Raghavan Prabhakar, Schütze Hinrich, Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schutze. *An Introduction to Information Retrieval*. Number c. 2009. ISBN 0521865719. doi: 10.1109/LPT.2009.2020494.
- [49] Zi Chu, Indra Widjaja, and Haining Wang. Detecting Social Spam Campaigns on Twitter. In *Lecture Notes in Computer Science (including subseries Lecture Notes*

- in Artificial Intelligence and Lecture Notes in Bioinformatics*), volume 7341 LNCS, pages 455–472. 2012. ISBN 9783642312830. doi: 10.1007/978-3-642-31284-7_27. URL http://link.springer.com/10.1007/978-3-642-31284-7_27.
- [50] David H. Wolpert and William G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997. ISSN 1089778X. doi: 10.1109/4235.585893.
- [51] Tg Dietterich. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural computation*, 10(7):1895–1923, 1998. ISSN 1530-888X. doi: 10.1162/089976698300017197. URL <http://www.ncbi.nlm.nih.gov/pubmed/9744903>.
- [52] Sara Dolnicar, Bettina Grun, and Friedrich Leisch. Quick, simple and reliable: forced binary survey questions. *International Journal of Market Research*, 53(2): 233, 2011. ISSN 1470-7853. doi: 10.2501/IJMR-53-2-233-254.
- [53] Robert. Parliamentary Procedure for Meetings. URL <https://goo.gl/hK2dDP>.
- [54] David. Cornish, David. Cornish, and RoseDog Books. *The essential twenty : twenty components of an excellent health care team*. RoseDog Books, 2009. ISBN 1434995550. URL <https://goo.gl/65aR2M>.
- [55] Maeve Duggan. The Demographics of Social Media Users | Pew Research Center, 2015. URL <http://www.pewinternet.org/2015/08/19/the-demographics-of-social-media-users/>.
- [56] Entertainment Software Association. Essential Facts About Computer and Video Game Industry. 2015. URL <http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf>.
- [57] Kathleen Bailey and J D Brown. Learning about Language Assessment: Dilemmas, Decisions, and Directions & New Ways of Classroom Assessment. *Learning*, 4(2), 1999.
- [58] Tom Fawcett. “ In vivo ” spam filtering : A challenge problem for data mining.

-
- [59] Twitter. Twitter Application Management, 2016. URL <https://apps.twitter.com/>.
- [60] R-Studio. Download RStudio, 2016. URL <https://www.rstudio.com/products/rstudio/download/>.
- [61] Robert Ihaka, Ross;Gentleman. The Comprehensive R Archive Network, 2016. URL <https://cran.rstudio.com/>.
- [62] Robert Ihaka, Ross;Gentleman. Httr, 2016. URL <https://cran.r-project.org/web/packages/httr/httr.pdf>.
- [63] Robert Ihaka, Ross;Gentleman. twitteR, 2016. URL <https://cran.r-project.org/web/packages/twitteR/twitteR.pdf>.
- [64] Robert Ihaka, Ross;Gentleman. tm, 2016. URL <https://cran.r-project.org/web/packages/tm/tm.pdf>.
- [65] Robert Ihaka, Ross;Gentleman. qdap, 2016. URL <https://cran.rstudio.com/web/packages/qdap/qdap.pdf>.