Data Science of the People, for the People, by the People: A Viewpoint on an Emerging Dichotomy

Kush R. Varshney
Mathematical Sciences and Analytics Department
IBM Thomas J. Watson Research Center
1101 Kitchawan Road, Yorktown Heights, New York
krvarshn@us.ibm.com

ABSTRACT

This paper presents a viewpoint on an emerging dichotomy in data science: applications in which predictions of data-driven algorithms are used to support people in making consequential decisions that can have a profound effect on other people's lives and applications in which data-driven algorithms act autonomously in settings of low consequence and large scale. An example of the first type of application is prison sentencing and of the second type is selecting news stories to appear on a person's web portal home page. It is argued that the two types of applications require data, algorithms and models with vastly different properties along several dimensions, including privacy, equitability, robustness, interpretability, causality, and openness. Furthermore, it is argued that the second type of application cannot always be used as a surrogate to develop methods for the first type of application. To contribute to the development of methods for the first type of application, one must really be working on the first type of application.

Categories and Subject Descriptors

H.0 [**Information Systems**]: General

General Terms

Algorithms, Human Factors

Keywords

causality, discrimination, interpretability, open data, privacy, robustness, sociotechnical systems

1. INTRODUCTION

Although perhaps overly hyped, it cannot be denied that developments in big data, analytics, and data science are transforming industries across all sectors. As this field progresses, a dichotomy of application domains is emerging. On one hand, we have applications such as medical diagnosis [22], loan approval [32] and prison sentencing [9] in which data-driven predictions are used to support people in making consequential decisions that can have a profound effect on people. On the other hand, we have applications in which data-driven algorithms automatically take actions without having people in the decision-making loop, usually on a large scale and for decisions of a less consequential nature; examples include streaming services deciding on the compression level of video packets to transmit to subscribers every few seconds [19, 25], web portals deciding which news story to show on top [1], and inside speech recognition systems [29]. For ease of reference, let the applications of consequence to people with human decision makers in the loop be Type A applications and the others be Type B.

Summarizing these two types of applications, Rayid Ghani, the chief data scientist for the Obama presidential campaign, recently said that [45] "the power of data science is typically harnessed in a spectrum with the following two extremes: helping humans in discovering new knowledge that can be used to inform decision making, or through automated predictive models that are plugged into operational systems and operate autonomously."

In this paper, it is argued that Type A and Type B applications of data science are fundamentally different from each other with different desiderata for data, processing, models, and analysis. As such, it is also argued that to really advance techniques for Type A applications, one must truly be working on Type A applications; Type B applications are often not suitable surrogates or sandboxes in which to test methods intended for Type A applications. This perspective is in conflict with the one, e.g., espoused by Claudia Perlich, the chief scientist of the computational advertising firm Dstillery, who has said that advertising [3] "is really the ultimate opportunity to try different things and find out what works in data science and what doesn't. . . . I find that I can disseminate some of the things that I learn and help bring my findings to medicine and other life-touching fields." In the remainder of the paper, taking a cue from John Wycliffe and Abraham Lincoln, various aspects that differentiate Type A applications from Type B applications are discussed under three broad categories, nominally *of the people*, *for the people*, and *by the people*.

2. DATA SCIENCE OF THE PEOPLE

In most Type A applications, the decisions to be made are decisions on individual people. Consequently the data required is data *of the people*, i.e. each data point is an individual person's record. If the application is of consequence, then usually at least some of the attributes in the required data are sensitive, e.g. fields related to health status, personal finances and educational attainment. Maintaining pri-

Bloomberg Data for Good Exchange Conference.
28-Sep-2015, New York City, NY, USA.

vacy and anonymity is the key concern when dealing with sensitive data of the people [16].

It is precisely because Type A applications are of consequence that privacy laws have been enacted in these domains, e.g. the Health Insurance Portability and Accountability Act (HIPAA), the Gramm-Leach-Bliley Act (GLBA), and the Family Educational Rights and Privacy Act (FERPA) in the United States. Usually in Type B applications, there is much less of an incentive to strictly protect privacy than in Type A applications because of the lack of such laws. Thus, although recognized as an issue, there are few existing methods and little research and development on privacy protection for Type B workloads.

We recently studied a Type A problem of predicting health care costs of people in a market that a provider does not serve, which required us to use importance sampling to account for covariate shift [56]; a very similar problem occurs in the Type B problem of targeted advertising [11]. However, we had to consider anonymity because we were working on the health care application, and to properly perform importance sampling, we had to develop a novel privacy-preserving data transformation that preserves the probability distribution of the data. This example illustrates the fact that if we had not been working on a Type A application, we would not have been pushed to develop a new data science method. Privacy is, by now, a well-studied topic in data science [2], but there will continue to be new avenues for research that will only be illuminated when developing solutions to real-world Type A problems.

A different issue that arises when working with and making decisions based on data of the people is fairness and discrimination. Just as there are sensitive attributes when considering privacy, there are different attributes in data of the people, such as race and gender, against which we would like predictive models to be equitable (in certain settings). As examples, we might want a loan approval classification algorithm to have similar acceptance rates for male and female applicants, and a data-driven model that aids in prison sentencing decisions to present similar predictions for black and white convicts. In such problems, which typically arise only in Type A applications, we may be willing to sacrifice generalization accuracy for equitability.

Human decision makers are known to be discriminatory due to bounded rationality [52], but algorithms can be too, for the following reason. The naïve approach of simply not including the protected attributes like race and gender as features in a predictive model fails to recognize that other included features may be highly correlated with them [28]. The legal standard known as disparate impact says that a decision-making process, even if it is not explicitly discriminatory at face value, is illegal if it has a disproportionately adverse effect on members of a protected class, and this is quite possible when using features correlated with attributes defining protected groups [7]. We note, however, that disparate impact is only illegal in Type A applications such as housing and employment, not in Type B applications like video streaming quality of service. Disparate impact can be stated mathematically in the language of detection theory based on the so-called 80% rule [21], and prevented using newly-developed data science methods [26, 21]. (There are even interesting mathematical connections between privacy and equitability [27, 43].)

The study of discrimination due to data analytics has only arisen because of interactions between machine learning researchers and the law and policy communities. It is another illustration of an area of study and advancement in data science that would not have been imagined simply by working on Type B applications (where we do not typically have any quality criterion other than accuracy); working with subject matter experts in Type A applications led to this area.

Let us consider a third issue when offering predictions based on data of the people: robustness. When we defined Type A applications in Section 1, we said that the decisions being made about people can have a profound effect on their lives, but that this is not the case with Type B applications. Moreover, in Type B applications, we have the opportunity to make millions or billions of predictions, but in Type A applications, we often only get one shot to make a prediction per person. For example, whereas severely mispredicting the level of engagement a person will have with a news article is undesirable, severely mispredicting the response a person will have to a medical treatment can be catastrophic.

Together, these considerations imply that while average or expected error across all of the millions or billions of predictions is a sensible criterion for a decision rule in Type B applications, alternatives optimizing worst-case error, known as robust formulations, make more sense for Type A applications. The main robustness considered in classification is against uncertainty in the class frequencies or relative costs of different types of errors [37, 17], but robustness against other uncertainties are also considered [49, 57]. When predictive models trained under expected error make gross mistakes, they are usually on people with rare characteristics; there are learning approaches to deal with rare classes [48], but there is little work on data points with rare features [58, 5]. (Robust regression methods prevent outliers from having outsized impact on estimation procedures, but this is not the type of robustness of interest here.)

Software packages for and real-world applications of learning decision functions that are robust to uncertainties in future data are not prevalent, but there is clearly a need in Type A applications because every prediction matters and worst-case perfomance matters. There is little need for such robustness if it sacrifices average performance in Type B applications. This is a topic that has not yet transitioned from theory to widespread practical use in data science. If and when it does, the advance will come as a result of the interplay between theory and Type A application.

3. DATA SCIENCE FOR THE PEOPLE

In Section 2, we examined differences between Type A and Type B applications that manifest because the data samples on which decisions are made represent people. In this section, we switch gears and examine differences between Type A and Type B applications that manifest because of the consumer of the predictions. In Type A applications, models are learned *for the people* to look at, understand,

and use to aid their decision making, but there is no person in the loop in Type B applications.

For people to use the predictions of an analytics model in their decision making, they must trust the model. To trust a model, they must understand how it makes its predictions, i.e. the model should be interpretable for the people [23, 41]. The model cannot be a black box that acts in ways that seem mysterious to the user.

Many of the older methods in the artificial intelligence literature, including decision lists and decision trees are interpretable [12, 14, 38, 39]. In contrast, newer developments such as ensemble methods, kernel methods, and deep neural networks are not interpretable. The older interpretable model learning algorithms are generally greedy or heuristic in nature and usually have inferior predictive performance to the newer uninterpretable approaches. However, recent work is revisiting the problem of interpretable learning through principled optimization formulations and achieving predictive performance approaching that of uninterpretable methods [40, 24, 18, 10, 34, 30, 55, 33, 50, 20].

We recently studied a Type A problem of predicting which IBM employees will voluntarily resign from the company, which required us to develop a classification algorithm to be placed within a larger decision-making system involving human decision makers [47]. In this problem, after several iterations with the business users, we decided on interpretable classification rule sets because of the trust and justification reasons mentioned earlier. In fact, this real-world problem was a key motivator for us to pursue the interpretable learning algorithm proposed in [34]. The choice of classifier taken in this project contrasts with the uninterpretable random forest classifier we used for a (social good) satellite image analysis problem in which the classifications were used autonomously without a human in the loop [51]. Together, these examples illustrate that Type A problems have unique requirements in terms of model interpretability that do not arise in Type B problems, and without the impetus from Type A applications, new data science approaches such as [34] would not be developed.

Causality is another issue that comes up when a predictive model is for people to consume and act upon. Typically in Type B applications, as long as a feature adds predictive value in generalization, regardless of whether it has any causal relationship with the outcome, it is included in a model [4]. However, in Type A applications, the user of the analytics wants to gain understanding into phenomena of interest, especially an understanding of what inputs or features cause the outcome being studied [36]. Causality allows one to understand what levers can be pulled to change the outcome, e.g. in the employee voluntary attrition problem, if lack of job promotion can be shown to cause resignation then promotions are a way to retain employees.

Precisely because of this understanding that comes with causal modeling is classical causal inference so popular in Type A applications in social sciences, epidemiology, public health, etc. The randomized controlled trial is the gold standard experimental design for causal inference, but is often difficult to carry out in those fields. Therefore, a body of techniques has been developed to try to tease out causal effects from various other experimental designs.

Randomized controlled trials known as A/B testing are used extensively in Type B applications [31], but, with a handful of exceptions [15, 6], there is not much active development of new causal inference techniques spurred by Type B applications. In contrast, there continue to be new data science methods developed for Type A applications, most recently incorporating ideas from high-dimensional machine learning, because causality is so critical to many Type A applications that it has to be a part of the inference objective no matter what [8, 4]. Thus again, we see Type A applications pushing data science in a different direction than Type B applications.

4. DATA SCIENCE BY THE PEOPLE

In Section 2, the general public constituted the data points and in Section 3, people making decisions were the consumers and users of data-driven models. In this final body section, devoted to data science *by the people*, we examine the role of the general public in applying data science methods in Type A and Type B problems.

First, we consider the question of why the general public would even want to conduct data science. Model interpretability, as discussed in Section 3, lends transparency, trust, and adoption to machine learning methods in Type A applications, but there is no greater transparency and way to develop trust than to have someone carry out the entire process himself or herself. Even if people do not actually do so, knowing that they could, gives a great sense of trust. Such openness allows for the possibility of audit and accountability that is necessary in many consequential Type A applications. On the other hand, Type B applications are less likely to require the possibility of audit because of their lower consequence, lower criteria for fairness, and lower need for robustness. Moreover, the set of potential users is more diverse in Type A applications.

Data science cannot be done without data, but data is an invaluable resource that most organizations keep locked up. The open data movement, however, is changing the equation by unlocking data and making it freely available to use, reuse and redistribute by the people. Due to executive mandates starting about six years ago, governments around the world —national and municipal—are leading the opening up of data through web portals they have established [44, 46].

Although there are several instances of Type A applications that rely solely on data internal to an organization or entity, many Type A applications rely on the type of government data, appropriately anonymized, that is now available on open data portals, sometimes in combination with internal data. In our health provider example, we could use either a provider's member cost data or open government medical cost survey data in combination with open government demographic data to solve the problem of interest. Open government data tends not to be useful in Type B applications.

Beyond governments, few other organizations currently release their internal data with all of the requirements needed to be labeled as open data. Data philanthropy is an emerging trend for corporations to make their data available for general use; the data sets opened in this way tend to have a social good angle and are done so in one shot rather than in continually maintained portals, e.g. Orange releasing anonymized mobile phone call detail records in Africa and Bitly releasing data on clicks to their URL shortener [35]. Other corporate data sets are opened, suitably anonymized and again on a one time basis, for competitions and challenges. Also, academics often release small-scale and large-scale data sets used in machine learning and data mining research. Both Type A and Type B applications can benefit from such data, but these avenues are less developed and codified than open government data.

In terms of software, there are best-of-breed open source packages available for a wide variety of data science algorithms applicable to Type A problems, Type B problems, or both. As discussed earlier in the section, there is a greater need for non-expert general public users to be able to run the software in Type A applications, so there is greater need for ease of use in Type A applications: another form of consumability. More generally, the data science algorithms in Type A applications tend to be embedded within complex sociotechnical systems, whereas the data science algorithms in Type B applications tend to be embedded within complex technical-only systems. The sociotechnical aspect of Type A applications introduces additional requirements for software and algorithms beyond what is currently available in those open source packages and what is needed for Type B applications. This is another point of differentiation between Type A and Type B for future development.

Finally, the recent years have seen the lowering of the barrier of entry to data science by the people due to cloud services that allow anyone to easily obtain sufficient computational resources. However, the scale, especially in the size of data sets, of Type B applications tends to result in a much higher computational burden and need for distributed and other advanced infrastructure requirements than Type A applications. Type B applications are more often in the truly big data regime than Type A applications, and thus not so easily approached by the general public. Overall, in comparing data science by the people for Type A and Type B applications in terms of data, software and computation, we see a bifurcation with open data ecosystems and user experience on the A side and truly big data on the B side.

5. CONCLUSION

In this viewpoint, we have examined two categories of data science applications, which we have named Type A and Type B. Type A applications are ones in which algorithmic outputs are used to support people in making decisions about other people that are consequential to them. Type B applications are ones in which the outputs drive autonomous machine processing or actions without humans in the decision-making loop that bear little consequence to people. We have studied how these two applications differ along several dimensions grouped into the categories *of the people*, *for the people*, and *by the people*: specifically privacy, discrimination, robustness, interpretability, causality, open data, and criteria for algorithms and infrastructure. The differences in the two types of applications are stark enough to have different requirements for data, processing, analysis, and infrastructure. Because of the differences, advances in one type of application will not necessarily transfer into advances in the other.

We are not the first to study these issues; for example, the recent Fairness, Accountability, and Transparency in Machine Learning workshops have highlighted several issues in data science of the people and for the people [54], and [13] has examined issues in data science by the people. As has also been recognized by others, we believe the primary reason for this emerging dichotomy is the recognition that to make impact in certain applications of data science, the accuracy of the results is not the only criterion that matters [53, 42]. There are important tradeoffs to consider in terms of accuracy vs. anonymity, accuracy vs. equitability, expected accuracy vs. worst-case accuracy, accuracy vs. interpretability, predictive accuracy vs. causal inference, and accuracy vs. integrability into sociotechnical systems.

This emerging dichotomy presents an opportunity to consider Type A applications and Type B applications separately rather than as a single pursuit. With that recognition, people and organizations working on Type A applications can focus their research and development energies on the myriad considerations besides pure accuracy described in this paper and other considerations of a similar flavor. Similarly, people and organizations working on Type B applications can maintain a singular focus on accuracy and the technical systems that lead to improved accuracy.

6. ACKNOWLEDGMENTS

The author thanks Vijay Iyengar, Been Kim, Samory Kpotufe, Ben Letham, Dmitry Malioutov, Aleksandra Mojsilović, Karthikeyan Natesan Ramamurthy, Guolong Su, Lav Varshney, Suresh Venkatasubramanian, Dennis Wei, and Bo Zhang for conversations that helped the author clarify his thinking on this subject.

7. REFERENCES

[1] D. Agarwal, B.-C. Chen, P. Elango, and R. Ramakrishnan. Content recommendation on web portals. *Comm. ACM*, 56(6):92–101, June 2013.

[2] C. C. Aggarwal and P. S. Yu, editors. *Privacy-Preserving Data Mining: Models and Algorithms*. Springer, New York, NY, 2008.

[3] A. Anisin. Data science interview: Claudia Perlich, chief scientist Dstillery. http://blog.dominodatalab.com/data-science-interview-claudia-perlich-chief-scientist-dstillery, May 2015.

[4] S. Athey and G. W. Imbens. Machine learning methods for estimating heterogeneous causal effects. http://arxiv.org/pdf/1504.01132.pdf, July 2015.

[5] J. Attenberg, P. Ipeirotis, and F. Provost. Beat the machine: Challenging humans to find a predictive model's "unknown unknowns". *ACM J. Data Inf. Qual.*, 6(1):1, Mar. 2015.

[6] J. Barajas, R. Akella, A. Flores, and M. Holtan. Estimating ad impact on clicker conversions for causal attribution. In *Proc. SIAM Int. Conf. Data Min.*, pages 640–648, Vancouver, Canada, Apr.–May 2015.

[7] S. Barocas and A. D. Selbst. Big data's disparate impact. *California Law Rev.*, 104, 2016.

[8] A. Belloni, V. Chernozhukov, and C. Hansen. High-dimensional methods and inference on structural and treatment effects. *J. Econ. Persp.*, 28(2):1–23, Spring 2014.

[9] R. Berk and J. Hyatt. Machine learning forecasts of risk to inform sentencing decisions. *Federal Sentencing Reporter*, 27(4):222–228, Apr. 2015.

[10] D. Bertsimas, A. Chang, and C. Rudin. An integer optimization approach to associative classification. In *Adv. Neur. Inf. Process. Syst. 25*, pages 269–277. 2012.

[11] S. Bickel, C. Sawade, and T. Scheffer. Transfer learning by distribution matching for targeted advertising. In *Adv. Neur. Inf. Process. Syst. 21*, pages 145–152. 2009.

[12] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Chapman & Hall, Boca Raton, FL, 1984.

[13] S. Chou, W. Li, and R. Sridharan. Democratizing data science: Effecting positive social change with data science. In *KDD at Bloomberg*, New York, NY, Aug. 2014.

[14] W. W. Cohen. Fast effective rule induction. In *Proc. Int. Conf. Mach. Learn.*, pages 115–123, Tahoe City, CA, Jul. 1995.

[15] B. Dalessandro, C. Perlich, O. Stitelman, and F. Provost. Causally motivated attribution for online advertising. In *Proc. Workshop Data Min. Online Advert. Internet Econ.*, page 7, Beijing, China, Aug. 2012.

[16] J. P. Daries, J. Reich, J. Waldo, E. M. Young, J. Whittinghill, A. D. Ho, D. T. Seaton, and I. Chuang. Privacy, anonymity, and big data in the social sciences. *Comm. ACM*, 57(9):56–63, Sept. 2014.

[17] M. A. Davenport, R. G. Baraniuk, and C. D. Scott. Tuning support vector machines for minimax and Neyman-Pearson classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(10):1888–1898, Oct. 2010.

[18] K. Dembczyński, W. Kotłowski, and R. Słowiński. ENDER: A statistical framework for boosting decision rules. *Data Min. Knowl. Disc.*, 21(1):52–90, Jul. 2010.

[19] R. Eisinger, R. A. F. Romero, and R. Goularte. Machine learning techniques applied to dynamic video adapting. In *Proc. Int. Conf. Mach. Learn. Appl.*, pages 819–822, San Diego, CA, Dec. 2008.

[20] A. Emad, K. R. Varshney, and D. M. Malioutov. A semiquantitative group testing approach for learning interpretable clinical prediction rules. In *Proc. Signal Process. Adapt. Sparse Struct. Repr. Workshop*, Cambridge, UK, July 2015.

[21] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubramanian. Certifying and removing disparate impact. In *Proc. ACM SIGKDD Conf. Knowl. Discov. Data Min.*, Sydney, Australia, Aug. 2015.

[22] K. R. Foster, R. Koprowski, and J. D. Skulfca. Machine learning, medical diagnosis, and biomedical engineering research - commentary. *Biomed. Eng. Online*, 13:94, July 2014.

[23] A. A. Freitas. Comprehensible classification models – a position paper. *SIGKDD Explorations*, 15(1):1–10, June 2013.

[24] J. H. Friedman and B. E. Popescu. Predictive learning via rule ensembles. *Ann. Appl. Stat.*, 2(3):916–954, Sep. 2008.

[25] N. Govind. Optimizing the Netflix streaming experience with data science. http://techblog.netflix.com/2014/06/optimizing-netflix-streaming-experience.html, June 2014.

[26] S. Hajian and J. Domingo-Ferrer. A methodology for direct and indirect discrimination prevention in data mining. *IEEE Trans. Knowl. Data Eng.*, 25(7):1445–1459, July 2013.

[27] S. Hajian, J. Domingo-Ferrer, and O. Farràs. Generalization-based privacy preservation and discrimination prevention in data publishing and mining. *Data Min. Knowl. Discov.*, 28(5–6):1158–1188, Sept. 2014.

[28] M. Hardt. How big data is unfair: Understanding sources of unfairness in data driven decision making. https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de, Sept. 2014.

[29] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Process. Mag.*, 29(6):82–97, Nov. 2012.

[30] B. Kim, C. Rudin, and J. A. Shah. Bayesian case model: A generative approach for case-based reasoning and prototype classification. In *Adv. Neur. Inf. Process. Syst. 27*, pages 1952–1960. 2014.

[31] R. Kohavi, R. Longbotham, D. Sommerfield, and R. M. Henne. Controlled experiments on the web: Survey and practical guide. *Data Min. Knowl. Disc.*, 18(1):140–181, Feb. 2009.

[32] S. Lessmann, H.-V. Seow, B. Baesens, and L. C. Thomas. Benchmarking state-of-the-art classification algorithms for credit scoring: A ten-year update. *Eur. J. Oper. Res.*, 247(1):124–136, Nov. 2015.

[33] B. Letham, C. Rudin, T. McCormick, and D. Madigan. Building interpretable classifiers with rules using Bayesian analysis. *Ann. Appl. Stat.*, forthcoming.

[34] D. M. Malioutov and K. R. Varshney. Exact rule learning via Boolean compressed sensing. In *Proc. Int. Conf. Mach. Learn.*, pages 765–773, Atlanta, GA, June 2013.

[35] A. Pawelke and A. R. Tatevossian. Data philanthropy: Where are we now? http://www.unglobalpulse.org/data-philanthropy-where-are-we-now, May 2013.

[36] J. Peters, J. M. Mooij, D. Janzing, and B. Schölkopf. Causal discovery with continuous additive noise models. *J. Mach. Learn. Res.*, 15:2009–2053, June 2014.

[37] F. Provost and T. Fawcett. Robust classification for imprecise environments. *Mach. Learn.*, 42(3):203–231, Mar. 2001.

[38] J. R. Quinlan. Simplifying decision trees. *Int. J. Man-Mach. Studies*, 27(3):221–234, Sep. 1987.

[39] R. L. Rivest. Learning decision lists. *Mach. Learn.*, 2(3):229–246, Nov. 1987.

[40] U. Rückert and S. Kramer. Margin-based first-order rule learning. *Mach. Learn.*, 70(2–3):189–206, Mar. 2008.

[41] C. Rudin. Algorithms for interpretable machine learning. In *Proc. ACM SIGKDD Conf. Knowl. Discov. Data Min.*, page 1519, New York, NY, Aug. 2014.

[42] C. Rudin and K. L. Wagstaff. Machine learning for science and society. *Mach. Learn.*, 95(1):1–9, Apr. 2014.

[43] S. Ruggieri. Using t-closeness anonymity to control for non-discrimination. *Trans. Data Privacy*, 7(2):99–129, Aug. 2014.

[44] A. Sahuguet, J. Krauss, L. Palacios, and D. Sangokoya. Open civic data: Of the people, by the people, for the people. *Bull. Tech. Comm. Data Eng.*, 37(4):15–26, Dec. 2014.

[45] C. Shan. How data science can be used for social good. http://www.carlshan.com/2015/01/08/data-science-social-good.html, Jan. 2015.

[46] E. Shaw. Improving service and communication with open data: A history and how-to. Technical report, Ash Center, Harvard Kennedy School, June 2015.

[47] M. Singh, K. R. Varshney, J. Wang, A. Mojsilović, A. R. Gill, P. I. Faur, and R. Ezry. An analytics approach for proactively combating voluntary attrition of employees. In *Proc. IEEE Int. Conf. Data Min. Workshops*, pages 317–323, Brussels, Belgium, Dec. 2012.

[48] Y. Tang, Y.-Q. Zhang, N. V. Chawla, and S. Krasser. SVMs modeling for highly imbalanced classification. *IEEE Trans. Syst., Man, Cybern. B*, 39(1):281–288, Feb. 2009.

[49] T. Tulabandhula and C. Rudin. Robust optimization using machine learning for uncertainty sets. In *ECML/PKDD PhD Sess. Proc.*, pages 121–128, Nancy, France, Sept. 2014.

[50] B. Ustun and C. Rudin. Supersparse linear integer models for optimized medical scoring systems. *Mach. Learn.*, forthcoming.

[51] K. R. Varshney, G. H. Chen, B. Abelson, K. Nowocin, V. Sakhrani, L. Xu, and B. L. Spatocco. Targeting villages for rural development using satellite image analysis. *Big Data*, 3(1):41–53, Mar. 2015.

[52] K. R. Varshney and L. R. Varshney. Quantization of prior probabilities for hypothesis testing. *IEEE Trans. Signal Process.*, 56(10):4553–4562, Oct. 2008.

[53] K. L. Wagstaff. Machine learning that matters. In *Proc. Int. Conf. Mach. Learn.*, pages 529–536, Edinburgh, United Kingdom, Jun.–Jul. 2012.

[54] H. Wallach. Big data, machine learning, and the social sciences: Fairness, accountability, and transparency. https://medium.com/@hannawallach/big-data-machine-learning-and-the-social-sciences-927a8e20460d, Dec. 2014.

[55] F. Wang and C. Rudin. Falling rule lists. In *Proc. Int. Conf. Artif. Intell. Stat.*, pages 1013–1022, San Diego, CA, May 2015.

[56] D. Wei, K. N. Ramamurthy, and K. R. Varshney. Health insurance market risk assessment: Covariate shift and k-anonymity. In *Proc. SIAM Int. Conf. Data Min.*, pages 226–234, Vancouver, Canada, Apr.–May 2015.

[57] D. Wei and K. R. Varshney. Robust binary hypothesis testing under contaminated likelihoods. In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pages 3407–3411, Brisbane, Australia, Apr. 2015.

[58] G. M. Weiss. Mining with rarity: A unifying framework. *SIGKDD Explor. Newsletter*, 6(1):7–19, June 2004.