Propagation of Uncertainty for Volunteered Geographic Information in Machine Learning

Jin Xing[1] and Renee Sieber[2]
[1]Department of Geomatics, Laval University, Quebec, Quebec, Canada
[2]Department of Geography, McGill University, Montreal, Quebec, Canada

Abstract

Although crowdsourcing drives much of the interest in Machine Learning (ML) in Geographic Information Science (GIScience), the impact of uncertainty of Volunteered Geographic Information (VGI) on ML has been insufficiently studied. This significantly hampers the application of ML in GIScience. In this paper, we briefly delineate five common stages of employing VGI in ML processes, introduce some examples, and then describe propagation of uncertainty of VGI.

1.Background of VGI in Machine Learning

Machine Learning (ML) represents a set of methods that automatically learn from "experience" or training data with respect to given tasks. The learning can be implemented via a huge body of models and algorithms, such as heuristical rules (Swan et al., 2015), decision trees (Paliouras et al., 2000), and cellular automata (Shafizadeh-Moghadam et al., 2017). In Geographic Information Science (GIScience), ML has attracted considerable interest due to its wide applications in place recognition (Zhou et al., 2014), ecology models (Olden, Lawler, and Poff, 2008), remote sensing image classification (Zhang, Zhang, and Du, 2016), transportation pattern discovery (Liu et al., 2017), and gazetteer analysis (Garfinkle et al., 2017). The rapid grow of ML has intensified due to the increasing 'bigness' of geospatial data (Miller and Goodchild, 2015), which describes the exaflood of geographic information at unprecedented volume, velocity, and variety, as well as challenges to veracity (Graham and Shelton, 2013).

Among the diverse sources of big data, Volunteered Geographic Information (VGI) is considered a main provider of input data/services (Goodchild and Glennon, 2010). For example, Openstreetmap {OSM} (Haklay and Weber, 2008), in which individuals have crowdsourced editable web mapping services and content, has become a powerful platform for building, training, and evaluating ML algorithms and models in GIScience (Mnih, 2013). VGI describes the process of obtaining geographic data or services (e.g., rating accuracy of feature labels) from large groups of user (i.e., both physical and virtual users) in an open call that is self-organizing via the Internet (Goodchild, 2007; Hudson-Smith et al., 2009). Uncertainty is innate within VGI, which means data is noisy, containing redundancies, irrelevant content, errors and bias contributed by users, who are often non-experts (O'Neil, 2016). VGI also is disorderly, in which data may be unstructured, incorrectly ordered, mis-formatted (e.g., lacking a header), and possibly poorly geo-registered (Østby, 2016). Finally, users may be unreliable in providing consistent input or inputting within a predefined time period. Noisy, disordered, and unreliable data and service can significantly lower the value of VGI in ML.

Previous work in VGI's uncertainty largely concentrates on the data quality. Researchers focused, for example, on uncertainty regarding the non-expert (e.g., skill levels and motivation), the thematic diversity (scattered focus relative to analysis needs) of input, and the spatial unevenness of contributions (e.g., popularity of places relative to others) (Grira et al., 2010; Budhathoki et al., 2008; Goodchild, 2008; Haklay, 2013; Flanagin and Metzger, 2008; Goodchild and Li, 2012; Roche et al., 2013). In ML, VGI is viewed primarily for its ability to provide data for ML, either as training data or general input data. It also been employed for result evaluation and hyper/parameter tuning of ML (Kanevski, Pozdnukhov, and Timonin, 2008). A worrying trend in GIScience inquiry into ML is its treatment as a big black box, where issues of data uncertainty are treated as I/O problems. We break down the black box of ML into a collection of workflow processes to identify uncertainty from VGI that can occur within the ML as well as in its parameterization and refinement.

Other taxonomies tend to focus on classifying ML methods (e.g., supervised, unsupervised, and reinforcement learning) (Jordan and Mitchell, 2015) and application areas (e.g., computer vision, natural language processing, and speech recognition) (Michalski et al., 2013). The importance of uncertainty and its propagation have not been highlighted. We view the interaction between VGI and ML as five stages throughout the processing of VGI: data collection and cleaning, data distribution, feature/topic detection, model/algorithm selection and training, and evaluation and tuning.

Table. Uncertainty Issues in Applying VGI for ML

ML Process	**Uncertainty Type**	**Examples in VGI**
Data Collection, Annotation, and Cleaning	Data Uncertainty	Inaccurate geolocation; spatial unevenness in data contributions; redundancies; gender, race bias in training data
Data Distribution	Operation	Boundary Vagueness (e.g., artificial boundaries

	Uncertainty	introduced by data splitting); aggregation errors (e.g., heaping error in determining the existence of a traffic jam, binning of VGI point data)
Feature/Topic Detection	Representation Uncertainty	Interpreting location from place (from a well-defined to a poorly defined object, ala Klir and Yuan, 1998)
Model/Algorithm Selection and Training	Decision Uncertainty	Simpler/alternate models than ML may be better like linear regression
Evaluation and Tuning	Service Uncertainty	Biased classification; Inconsistency in grading

2. A General Framework for Integrating Geospatial Crowdsourcing and ML

Our framework (Table) follows the standard ML workflow (data collection and cleaning, splitting of training from testing data, model training, evaluation, parameter tuning) (Pedregosa et al., 2011) and adds components from big data handling (Lee and Kang, 2015) and ML computation (Chu et al., 2007) for de-/re-composition. Since the five stages may occur iteratively (e.g., the evaluation result could be fed back to the training process for accuracy improvement), uncertainty also can propagate without further attention to the origin of the uncertainty.

2.1. Data Collection and Cleaning

The primary utility of VGI in ML is for training and, more generally, input data. Training refers to data used by ML to calculate its parameters/weights so that input data generates expected outputs. Geospatial content is available across a wide range of VGI. It can be raster (landscape photographs) and vector (social checkins, binned aggregations of points); structured (Twitter metadata) and unstructured (Twitter text), explicit (x,y's, placenames in hashtags) and implicit (colloquial names for neighborhood), absolute (latitude/longitude) and relative (concepts of home), passive (geo-fencing) and active (Amazon Mechanical Turk-AMT) (Heipke, 2010). It can be static or dynamic (harvesting of Flickr geotags at point in time or movement data), compensated or voluntary (AMT or VGI) (Kazemi and Shahabi, 2012). Considerable research has been conducts to assess uncertainty with various VGI (cf., Grira, Bédard, and Roche 2010).

Like other crowdsourced content, VGI data contains considerable error, vagueness, and ambiguity (Hsueh, Melville, and Sindhwani, 2009) and is vulnerable to malicious contributions (e.g., via GPS spoofing) (Wang et al., 2014). As suggested above, this is the richest area of current research so this section is admittedly brief. Most research on the negative impact of ML focuses on the issue of algorithmic bias due to input data (O’Neil 2016). Location often serves as a proxy for race so one needs to debias on the basis of primary variable as well as data which functions as its surrogate (Angwin et al. 2016). Often dibiasing requires human intervention (cf., gendered word2vec example in Bolukbasi et al. 2016) so this stage also can utilize crowdsourcing. Geographic unevenness in data contributions can further distort ML output, for example the low OSM participation in Africa (Perkins, 2014) or as the differential accuracy of OSM in urban areas versus rural regions (Haklay, 2010). Privacy protections, like the EU’s General Data Protection Regulation, will increase distortions in VGI as whole swaths of data are removed or masked (Ding et al. 2015). Lastly, much of VGI is streamed, which requires new sampling techniques (e.g., reservoir sampling) to normalize temporal spikes or redundancies.

2.2. Data Distribution

The attraction of VGI to ML is both in its source (geosocial media) and its potential as big data. The latter likely requires de-/re-composition to distribute the computing. Data distribution may incur disorder in VGI because geographic data has its own internal topology and geometry that can be destroyed by arbitrary decomposition or splitting. For example, rectangular decomposition will distort the boundary of geographic objects and increase the uncertainty (De Longueville et al., 2010). Most VGI is point-based and may need to be binned. A more sophisticated polygon, like a hexagon, does not easily alleviate the problem and any aggregation is subject to modifiable areal unit problems (McNamara and Lunzer, 2016) that can alter ML output.

ML can be employed to alleviate uncertainty in data distribution. Felzenszwalb et al. (2010) employed latent support vector machine to decompose the original raster data into multiple object-based rectangles to avoid boundary distortion. Temporal disorder in VGI, such as burstiness of reporting of natural disasters, could be addressed by decomposition with parallel processing (Goodchild, Fu, and Rich, 2007).

2.3. Feature/Topic Detection

ML is designed in large part to recognize patterns, generate rules, approximate functions, and classify data sets. One of the most important VGI services in ML is for feature or topic detection (e.g., forest, alternate route to avoid traffic jam). We lack explicit control over the feature representation in VGI. Users may not provide feature identification as planned or neural networks may fail to extract useful features from noisy VGI (Krishna et al., 2017). For example, uncertainty in placename makes it difficult to infer locations; “downtown nearby” could be interpreted as multiple locations (Flanagin and Metzger, 2008). Although iterative feature/object detection in ML can reduce uncertainty, there is no easy way to clean data to better disambiguate place to a location and location to a place. This resembles the challenge of NLP regarding semantic modeling to disambiguate slang (e.g., “bad”, “hot”, “sick”) in ML (Wolf et al., 2014).

ML in practice is treated as a blackbox, an algorithm amongst many in a software library. Treating ML as a black box means that ML cannot accommodate the geography of VGI, which therefore induces inaccuracies. For example, max pooling, which is a widely used method to pass features from one layer of neural network to another, is considered problematic in convolutional neural network by Hintion et al. (2018) because max pooling lacks topology. In another example, a word embedding algorithm may produce very different vectors to represent “pub” and “bar” due to the surrounding content, which may then require multiple detection iterations.

2.4. Model/Algorithm Selection and Training

Which ML model or algorithm achieves the highest accuracy with a given input dataset and features? What is the best way to calculate the weights or parameters of the ML model/algorithm? Should we rely on a single ML model/algorithm or combine several ones together? These questions are difficult in ML and there are no clear answers (Dietterich, 2000). VGI can potentially assist this selection process with existing knowledge about model/algorithm selection and training strategies (think a wiki of appropriate ML) (Marcus, 2018). However, knowledge contributed via VGI may be unreliable because of a “follow the crowd” mentality with little investigation into alternate approaches (Kamar, Hacker, and Horvitz, 2012). Deep neural network is increasingly popular in ML research but a linear regression may be more appropriate, considering the quality of the data at hand and the ease of an ML implementation (De Albuquerque et al., 2015).

2.5. Evaluation and Tuning

Performance of ML algorithms needs to be evaluated with datasets different from the training process (Gebru et al., 2017). VGI, or rather crowdsourcing, plays a pivotal role in collecting evaluation datasets. To avoid overfitting (i.e., model is too closely fitted to the training data), ML scientists usually employ cross-validation (Kohavi, 1995), which could reduce the influence of uncertainty from VGI training data. The evaluation can also be conducted with crowdsourcing services, such as the translation validation within the Google Translate Community (Kumar, 2002) or Captcha (Chellapilla and Simard, 2005). Here, issues similar to data collection re-emerge, with potential biases introduced by the evaluators, who may be drawn from a particular gender, race, class, or skill level. These issues resemble the social approach to assessing spatial data accuracy in Goodchild and Li (2012), in which the focus shifts from the uncertainty of the contribution to that of the contributor. One may wish to implement ranking or rating systems to improve confidence in the validators.

3. Propagation of Uncertainty in ML and Conclusion

In this paper, we propose a general framework to explore VGI uncertainty in ML. This includes the concrete importance of VGI for training data as well as the use of crowdsourcing for model/algorithm selection and performance evaluation in ML. The propagation of VGI's uncertainty should not be overlooked because it can accumulate throughout the workflow of ML processes.

Uncertainty propagates across the five stages in ML. Uncertainty in data collection could make data distribution more difficult because we do not know the appropriate aggregation size or scale. Without adequate cleaning, noisy data can generate messy features or false positives that will invalidate most ML models and algorithms. Crowdsourcers bring their own bias to the evaluation of ML, which can influence the training of ML for parameter tuning. Disagreements during the cross validations may generate inconsistency in iterations of ML and force us to re-run the process. Therefore, it is critical to identify uncertainty at each stage to minimize the propagation of uncertainty, where possible.

There are pros and cons about the propagation of VGI's uncertainty. On the one hand, uncertainty is innate to geography and will likely occur with any new source of VGI (Lutz, 2015). On the other hand, noise, disorder, and unreliability of VGI require ever more cleaning/validation work and lowers accuracy in ML, possibly by overfitting the results. All of this occurs against the backdrop of increasing sophistication in ML. It is somewhat of a race to see which outpaces which, the efficacy of ML or the messiness of VGI.

Reference

Abad, A., Nabi, M., & Moschitti, A. (2017). Autonomous Crowdsourcing through Human-Machine Collaborative Learning. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 873-876). ACM.

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Kudlur, M. (2016). TensorFlow: A System for Large-Scale Machine Learning. In OSDI (Vol. 16, pp. 265-283).

Ananny, M. (2016). Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Science, Technology & Human Values, 41*(1), 93-117.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias, *ProPublica*. Published May 23, 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.

Batista, G. E., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD explorations newsletter, 6(1), 20-29.

Bishr, M., & Mantelas, L. (2008). A trust and reputation model for filtering and classifying knowledge about urban growth. GeoJournal, 72(3-4), 229-237.

Bolukbasi, T., Chang, K., Zou, J., Saligrama, V. & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. 30th Conference on Neural Information Processing Systems, Barcelona, Spain. 9pp.

Bone, C., & Dragićević, S. (2009). GIS and intelligent agents for multiobjective natural resource allocation: A reinforcement learning approach. Transactions in GIS, 13(3), 253-272.

Budhathoki, N. R., Bruce, B., & Nedovic-Budic, Z. (2008). Reconceptualizing the role of the user of spatial data infrastructure. Geojournal, 72, 149–160.

Chellapilla, K., & Simard, P. Y. (2005). Using machine learning to break visual human interaction proofs (HIPs). In Advances in neural information processing systems (pp. 265-272).

Chen, F., Arredondo, J., Khandpur, R. P., Lu, C. T., Mares, D., Gupta, D., & Ramakrishnan, N. (2012). Spatial surrogates to forecast social mobilization and civil unrests. In Position Paper in CCC Workshop on "From GPS and Virtual Globes to Spatial Computing-2012.

Chen, X., Lin, Q., & Zhou, D. (2013). Optimistic knowledge gradient policy for optimal budget allocation in crowdsourcing. In International Conference on Machine Learning (pp. 64-72).

Chu, C. T., Kim, S. K., Lin, Y. A., Yu, Y., Bradski, G., Olukotun, K., & Ng, A. Y. (2007). Map-reduce for machine learning on multicore. In Advances in neural information processing systems (pp. 281-288).

Colby, M., Yliniemi, L., & Tumer, K. (2016). Autonomous multiagent space exploration with high-level human feedback. Journal of Aerospace Information Systems, 301-315.

Coleman, D., Georgiadou, Y., & Labonte, J. (2009). Volunteered geographic information: The nature and motivation of produsers. IJSDIR, 4(1), 332-358.

Crampton, J. W., Graham, M., Poorthuis, A., Shelton, T., Stephens, M., Wilson, M. W., & Zook, M. (2013). Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb. Cartography and geographic information science, 40(2), 130-139.

David, H. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of Economic Perspectives, 29(3), 3-30.

Davis, J., Arderiu, J., Lin, H., Nevins, Z., Schuon, S., Gallo, O., & Yang, M. H. (2010). The hpu. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on (pp. 9-16). IEEE.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 248-255). IEEE.

De Longueville, B., Ostländer, N., & Keskitalo, C. (2010). Addressing vagueness in Volunteered Geographic Information (VGI)–A case study. International Journal of Spatial Data Infrastructures Research, 5, 1725-0463.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In International workshop on multiple classifier systems (pp. 1-15). Springer, Berlin, Heidelberg.

Ding, S, Lin, L., Wang, G., & Chao, H. (2015). Deep feature learning with relative distance comparison for person re-identification. Pattern Recognition 48, 10, 2993-3003.

Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the world-wide web. Communications of the ACM, 54(4), 86-96.

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object detection with discriminatively trained part-based models. IEEE transactions on pattern analysis and machine intelligence, 32(9), 1627-1645.

Flanagin, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, *72*(3-4), 137-148.

Foody, G. M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sensing of Environment, 103(2), 179-189.

Foster, I., Kesselman, C., Nick, J. M., & Tuecke, S. (2002). Grid services for distributed system integration. Computer, 35(6), 37-46.

Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., ... & Obersteiner, M. (2012). Geo-Wiki: An online platform for improving global land cover. Environmental Modelling & Software, 31, 110-123.

Gaikwad, S. N., Morina, D., Nistala, R., Agarwal, M., Cossette, A., Bhanu, R., ... & Mithal, A. (2015). Daemo: A self-governed crowdsourcing marketplace. In Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (pp. 101-102). ACM.

Gao, S., & Yan, B. (2018). Place2Vec: Visualizing and Reasoning About Place Type Similarity and Relatedness by Learning Context Embeddings. In Adjunct Proceedings of the 14th International Conference on Location Based Services(pp. 225-226). ETH Zurich.

Garfinkle, N. W., Selig, L., Perkins, T. K., & Calfas, G. W. (2017). Geoparsing text for characterizing urban operational environments through machine learning techniques. In Geospatial Informatics, Fusion, and Motion Video Analytics VII (Vol. 10199, p. 101990C). International Society for Optics and Photonics.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. Proceedings of the National Academy of Sciences, 201700035.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural computation*, *4*(1), 1-58.

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. Spatial statistics, 1, 110-120.

Goodchild, M. F. (2011). Scale in GIS: An overview. Geomorphology, 130(1-2), 5-9.

Goodchild, M.F. (2008). Commentary: whither VGI? GeoJournal 72, 239–244.

Goodchild, M.F. (2007). Citizens as Sensors: The World of Volunteered Geography. Geo Journal 69:211-221.

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. International Journal of Digital Earth, 3(3), 231-241.

Goodchild, M. F., Fu, P., & Rich, P. (2007). Sharing Geographic Information: An Assessment of the Geospatial One- Stop. Annals of the Association of American Geographers, 97(2), 250-266.

Graham, M., & Shelton, T. (2013). Geography and the future of big data, big data and the future of geography. Dialogues in Human Geography, 3(3), 255-261.

Gulisano, V., Jimenez-Peris, R., Patino-Martinez, M., Soriente, C., & Valduriez, P. (2012). Streamcloud: An elastic and scalable data streaming system. IEEE Transactions on Parallel and Distributed Systems, 23(12), 2351-2365.

Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. IEEE Pervasive Computing, 7(4), 12-18.

Haklay, M., Singleton, A., & Parker, C. (2008). Web mapping 2.0: The neogeography of the GeoWeb. Geography Compass, 2(6), 2011-2039.

Hall, M. A. (1999). Correlation-based feature selection for machine learning.

Hara, K., Sun, J., Moore, R., Jacobs, D., & Froehlich, J. (2014, October). Tohme: detecting curb ramps in google street view using crowdsourcing, computer vision, and machine learning. In Proceedings of the 27th annual ACM symposium on User interface software and technology (pp. 189-204). ACM.

Heipke, C. (2010). Crowdsourcing geospatial data. ISPRS Journal of Photogrammetry and Remote Sensing, 65(6), 550-557.

Hengl, T., de Jesus, J. M., Heuvelink, G. B., Gonzalez, M. R., Kilibarda, M., Blagotić, A., ... & Guevara, M. A. (2017). SoilGrids250m: Global gridded soil information based on machine learning. PLoS one, 12(2), e0169748.

Hsueh, P. Y., Melville, P., & Sindhwani, V. (2009). Data quality from crowdsourcing: a study of annotation selection criteria. In Proceedings of the NAACL HLT 2009 workshop on active learning for natural language processing (pp. 27-35). Association for Computational Linguistics.

Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sensing, 7(11), 14680-14707.

Hudson-Smith, A., Batty, M., Crooks, A., & Milton, R. (2009). Mapping for the masses: Accessing web 2.0 through crowdsourcing. Social science computer review, 27(4), 524-538.

Ipeirotis, P. G., & Horton, J. J. (2011). The need for standardization in crowdsourcing. In Proceedings of the workshop on crowdsourcing and human computation at CHI.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759.

Kamar, E., Hacker, S., & Horvitz, E. (2012). Combining human and machine intelligence in large-scale crowdsourcing. In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1 (pp. 467-474). International Foundation for Autonomous Agents and Multiagent Systems.

Kanevski, M., Pozdnukhov, A., & Timonin, V. (2008). Machine learning algorithms for geospatial data. Applications and software tools.

Kanhere, S. S. (2011, June). Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. In *Mobile Data Management (MDM), 2011 12th IEEE International Conference on* (Vol. 2, pp. 3-6). IEEE.

Karam, R., & Melchiori, M. (2013). A crowdsourcing-based framework for improving geo-spatial open data. In Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (pp. 468-473). IEEE.

Kaufman, A., King, G., & Komisarchik, M. (2017). How to Measure Legislative District Compactness If You Only Know it When You See it.

Kazemi, L., & Shahabi, C. (2012). Geocrowd: enabling query answering with spatial crowdsourcing. In Proceedings of the 20th international conference on advances in geographic information systems (pp. 189-198). ACM.

Khan, F. (2012). An initial seed selection algorithm for k-means clustering of georeferenced data to improve replicability of cluster assignments for mapping application. Applied Soft Computing, 12(11), 3698-3700.

Kim, J., & Monroy-Hernandez, A. (2016, February). Storia: Summarizing social media content based on narrative theory using crowdsourcing. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 1018-1027). ACM.

Kittur, A., Smus, B., Khamkar, S., & Kraut, R. E. (2011). Crowdforge: Crowdsourcing complex work. In Proceedings of the 24th annual ACM symposium on User interface software and technology (pp. 43-52). ACM.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai(Vol. 14, No. 2, pp. 1137-1145).

Krishna, R., Zhu, Y., Groth, O., Johnson, J., Hata, K., Kravitz, J., ... & Bernstein, M. S. (2017). Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International Journal of Computer Vision*, *123*(1), 32-73.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

Kuderer, M., Gulati, S., & Burgard, W. (2015). Learning driving styles for autonomous vehicles from demonstration. In Robotics and Automation (ICRA), 2015 IEEE International Conference on (pp. 2641-2646). IEEE.

Kumar, S. (2002). Methods for community participation: a complete guide for practitioners.

Kyzirakos, K., Karpathiotakis, M., & Koubarakis, M. (2012). Strabon: a semantic geospatial DBMS. In International Semantic Web Conference (pp. 295-311). Springer, Berlin, Heidelberg.

Lease, M. (2011). On quality control and machine learning in crowdsourcing. Human Computation, 11(11).

Lee, J. G., & Kang, M. (2015). Geospatial big data: challenges and opportunities. Big Data Research, 2(2), 74-81.

Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012). Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 501-510). ACM.

Lin, S. W., & Lo, L. Y. S. (2016, June). Motivation for using the Social Commerce Website in the Sharing Economy: a two-factor Theory Perspective. In PACIS (p. 108).

Ling, W., Dyer, C., Black, A. W., & Trancoso, I. (2015). Two/too simple adaptations of word2vec for syntax problems. In Proceedings of the 2015 Conference of the North

American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 1299-1304).

Liu, J., Yu, X., Xu, Z., Choo, K. K. R., Hong, L., & Cui, X. (2017). A cloud- based taxi trace mining framework for smart city. Software: Practice and Experience, 47(8), 1081-1094.

Lutz, J. (2015). The validity of crowdsourcing data in studying anger and aggressive behavior. *Social Psychology*.

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and planning B: Planning and design*, *37*(4), 682-703.

Marcus, G. (2018). Deep Learning: A Critical Appraisal. arXiv preprint arXiv:1801.00631.

McNamara, A. & Lunzer, A. (2016). Exploring the effects of spatial aggregation. VPRI Research Note RN-2016-001. Viewpoints Research Institute, Los Angeles, CA.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.

Mikolov, T., Le, Q. V., & Sutskever, I. (2013). Exploiting similarities among languages for machine translation. arXiv preprint arXiv:1309.4168.

Miller, H. J., & Goodchild, M. F. (2015). Data-driven geography. GeoJournal, 80(4), 449-461.

Mitchell, T. M. (1997). Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45(37), 870-877.

Mnih, V. (2013). Machine learning for aerial image labeling(Doctoral dissertation, University of Toronto (Canada)).

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529.

Mohamed, A. Z. (2014). A review on crowd sourcing Geo-Social related big data approaches as solution to transportation problem. In Applied Mechanics and Materials(Vol. 663, pp. 622-626). Trans Tech Publications.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3), 436-465.

Ofli, F., Meier, P., Imran, M., Castillo, C., Tuia, D., Rey, N., ... & Joost, S. (2016). Combining human computing and machine learning to make sense of big (aerial) data for disaster response. Big data, 4(1), 47-59.

Olden, J. D., Lawler, J. J., & Poff, N. L. (2008). Machine learning methods without tears: a primer for ecologists. The Quarterly review of biology, 83(2), 171-193.

Østby, G. (2016). Rural–urban migration, inequality and urban social disorder: Evidence from African and Asian cities. *Conflict Management and Peace Science*, *33*(5), 491-515.

O'Neil, C. 2016. Weapons of Math Destruction. New York: Crown Random House

Paliouras, G., Karkaletsis, V., Petasis, G., & Spyropoulos, C. D. (2000). Learning decision trees for named-entity recognition and classification. In ECAI Workshop on Machine Learning for Information Extraction.

Pasquale, F. 2015. The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.

Perkins, C. (2014). Plotting practices and politics:(im) mutable narratives in OpenStreetMap. *Transactions of the Institute of British Geographers*, *39*(2), 304-317.

Roche, S., Propeck-Zimmermann, E., & Mericskay, B. (2013). GeoWeb and crisis management: Issues and perspectives of volunteered geographic information. *GeoJournal*, *78*(1), 21-40.

Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic routing between capsules. In Advances in Neural Information Processing Systems (pp. 3859-3869).

Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., & Taleai, M. (2017). Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. Computers, Environment and Urban Systems, 64, 297-308.

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. Annals of the association of American Geographers, 96(3), 491-507.

Sorokin, A., Berenson, D., Srinivasa, S. S., & Hebert, M. (2010). People helping robots helping people: Crowdsourcing for grasping novel objects. In Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on (pp. 2117-2122). IEEE.

Stephens M (2013) Gender and the GeoWeb: divisions in the production of user-generated cartographic information. *GeoJournal* 78(6): 981–996.

Stockwell, D. (1999). The GARP modelling system: problems and solutions to automated spatial prediction. International journal of geographical information science, 13(2), 143-158.

Swan, A. L., Stekel, D. J., Hodgman, C., Allaway, D., Alqahtani, M. H., Mobasheri, A., & Bacardit, J. (2015). A machine learning heuristic to identify biologically relevant and minimal biomarker panels from omics data. BMC genomics, 16(1), S2.

Tian, F., Liu, B., Sun, X., Zhang, X., Cao, G., & Gui, L. (2017). Movement-based incentive for crowdsourcing. IEEE Transactions on Vehicular Technology, 66(8), 7223-7233.

Truta, T. M., & Vinay, B. (2006). Privacy protection: p-sensitive k-anonymity property. In Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on (pp. 94-94). IEEE.

Vandecasteele, A., & Devillers, R. (2013). Improving volunteered geographic data quality using semantic similarity measurements. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 1(1), 143-148.

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2008). Temporal Evolution of Short- Term Urban Traffic Flow: A Nonlinear Dynamics Approach. Computer- Aided Civil and Infrastructure Engineering, 23(7), 536-548.

Wang, G., Wang, T., Zheng, H., & Zhao, B. Y. (2014). Man vs. Machine: Practical Adversarial Detection of Malicious Crowdsourcing Workers. In USENIX Security Symposium (pp. 239-254).

Wolf, L., Hanani, Y., Bar, K., & Dershowitz, N. (2014). Joint word2vec Networks for Bilingual Semantic Representations. *Int. J. Comput. Linguistics Appl.*, *5*(1), 27-42.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Klingner, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.

Xing, J., Sieber, R., & Kalacska, M. (2014). The challenges of image segmentation in big remotely sensed imagery data. Annals of GIS, 20(4), 233-244.

Yang, C., Raskin, R., Goodchild, M., & Gahegan, M. (2010). Geospatial cyberinfrastructure: past, present and future. Computers, Environment and Urban Systems, 34(4), 264-277.

Zesch, T., & Gurevych, I. (2007). Analysis of the Wikipedia category graph for NLP applications. In Proceedings of the Second Workshop on TextGraphs: Graph-Based Algorithms for Natural Language Processing (pp. 1-8).

Zhai, H., Lingren, T., Deleger, L., Li, Q., Kaiser, M., Stoutenborough, L., & Solti, I. (2013). Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing. Journal of medical Internet research, 15(4).

Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geoscience and Remote Sensing Magazine, 4(2), 22-40. ,.

Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., & Oliva, A. (2014). Learning deep features for scene recognition using places database. In Advances in neural information processing systems (pp. 487-495).

Zhou, Z. H. (2015). Ensemble learning. Encyclopedia of biometrics, 411-416.

Zogaj, S., Bretschneider, U., & Leimeister, J. M. (2014). Managing crowdsourced software testing: a case study based insight on the challenges of a crowdsourcing intermediary. Journal of Business Economics, 84(3), 375-405.