Bibliography

[1] Wendy Bohte and Kees Maat. “Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands”. In: *Transportation Research Part C: Emerging Technologies* 17.3 (2009), pp. 285–297.

[2] Kevin Gingerich, Hanna Maoh, and William Anderson. “Classifying the purpose of stopped truck events: An application of entropy to GPS data”. In: *Transportation Research Part C: Emerging Technologies* 64 (2016), pp. 17–27.

[3] Xia Yang et al. “Urban freight delivery stop identification with GPS data”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2411 (2014), pp. 55–61.

[4] Leonardo Sarti et al. “Stop Purpose Classification from GPS Data of Commercial Vehicle Fleets”. In: *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE. 2017, pp. 280–287.

[5] Stephen Greaves and Miguel Figliozzi. “Collecting commercial vehicle tour data with passive global positioning system technology: Issues and potential applications”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2049 (2008), pp. 158–166.

[6] P Gonzalez et al. “Automating mode detection using neural networks and assisted GPS data collected using GPS-enabled mobile phones”. In: *15th World congress on intelligent transportation systems*. 2008, pp. 16–20.

[7] Bryce W Sharman and Matthew J Roorda. “Analysis of freight global positioning system data: clustering approach for identifying trip destinations”. In: *Transportation Research Record* 2246.1 (2011), pp. 83–91.

[8] Lei Gong et al. “Identification of activity stop locations in GPS trajectories by density-based clustering method combined with support vector machines”. In: *Journal of Modern Transportation* 23.3 (2015), pp. 202–213.

[9] Russel Aziz et al. “Identifying and characterizing truck stops from GPS data”. In: *Industrial Conference on Data Mining*. Springer. 2016, pp. 168–182.

[10] “Details Of The GPS Position Calculation”. In: *Courses.Psu.Edu* (2019). URL: https://www.courses.psu.edu/aersp/aersp055_r81/satellites/gps_details.html.

[11] “Scania takes lean production efficiency into mining”. In: *Scania Press releases* (2016). URL: https://www.scania.com/group/en/scania-takes-lean-production-efficiency-into-mining/ (visited on 04/10/2019).

[12] Jason Brownlee. “Discover feature engineering, how to engineer features and how to get good at it”. In: *Machine Learning Process* (2014). URL: https://machinelearningmastery.com/discover-feature-engineering-how-to-engineer-features-and-how-to-get-good-at-it/.

[13] Isabelle Guyon et al. *Feature extraction: foundations and applications*. Vol. 207. Springer, 2008.

[14] Rohin Mohanadas. *Discerning Truck Stop Semantics through Latent Space Clustering*. 2018. URL: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-240598.

[15] Lei Gong, Toshiyuki Yamamoto, and Takayuki Morikawa. “Identification of activity stop locations in GPS trajectories by DBSCAN-TE method combined with support vector machines”. In: *Transportation Research Procedia* 32 (2018), pp. 146–154.

[16] Isabelle Guyon and André Elisseeff. “An introduction to variable and feature selection”. In: *Journal of machine learning research* 3.Mar (2003), pp. 1157–1182.

[17] Xiaojin Zhu. *Introduction to semi-supervised learning*. eng. San Rafael, Calif.]: Morgan Claypool, 2009. Chap. 2, pp. 9–15. ISBN: 1-59829-547-0.

[18] Fabio G Cozman, Ira Cohen, and Marcelo C Cirelo. "Semi-supervised learning of mixture models". In: *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*. 2003, pp. 99–106.

[19] Xiuming Liu, Dave Zachariah, and Johan Wågberg. "Reliable Semi-Supervised Learning when Labels are Missing at Random". In: *arXiv preprint arXiv:1811.10947* (2018).

[20] Aarti Singh, Robert Nowak, and Jerry Zhu. "Unlabeled data: Now it helps, now it doesn't". In: *Advances in neural information processing systems*. 2009, pp. 1513–1520.

[21] Dong-Hyun Lee. "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks". In: *Workshop on Challenges in Representation Learning, ICML*. Vol. 3. 2013, p. 2.

[22] Samuli Laine and Timo Aila. "Temporal ensembling for semi-supervised learning". In: *arXiv preprint arXiv:1610.02242* (2016).

[23] Antti Tarvainen and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results". In: *Advances in neural information processing systems*. 2017, pp. 1195–1204.

[24] Takeru Miyato et al. "Virtual adversarial training: a regularization method for supervised and semi-supervised learning". In: *IEEE transactions on pattern analysis and machine intelligence* (2018).

[25] Alexey Kurakin, Ian Goodfellow, and Samy Bengio. "Adversarial machine learning at scale". In: *arXiv preprint arXiv:1611.01236* (2016).

[26] Avital Oliver et al. "Realistic evaluation of deep semi-supervised learning algorithms". In: *Advances in Neural Information Processing Systems*. 2018, pp. 3235–3246.

[27] Hoang Minh. "How to Handle Imbalanced Data in Classification Problems". In: *James Tech Blog* (2018). URL: https://medium.com/james-blogs/handling-imbalanced-data-in-classification-problems-7de598c1059f (visited on 04/04/2019).

[28] Vaishali Ganganwar. "An overview of classification algorithms for imbalanced datasets". In: *International Journal of Emerging Technology and Advanced Engineering* 2.4 (2012), pp. 42–47.

[29] Mohammad Hossin and MN Sulaiman. "A review on evaluation metrics for data classification evaluations". In: *International Journal of Data Mining & Knowledge Management Process* 5.2 (2015), p. 1.

[30] Wikipedia contributors. *Student's t-test — Wikipedia, The Free Encyclopedia*. [Online; accessed 14-March-2021]. 2021. URL: https://en.wikipedia.org/w/index.php?title=Student%5C%27s_t-test&oldid=1011607328.

[31] Matei Zaharia et al. "Apache Spark: A Unified Engine for Big Data Processing". In: *Commun. ACM* 59.11 (Oct. 2016), pp. 56–65. ISSN: 0001-0782. DOI: 10.1145/2934664. URL: http://doi.acm.org/10.1145/2934664.

[32] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[33] Adam Paszke et al. "Automatic differentiation in PyTorch". In: *NIPS-W*. 2017.

[34] Michael R Berthold et al. *Guide to intelligent data analysis: how to intelligently make sense of real data*. Springer Science & Business Media, 2010. Chap. 4, pp. 33–79.

[35] Jure Leskovec, Anand Rajaraman, and Jeffrey David Ullman. *Mining of massive datasets*. Cambridge university press, 2014. Chap. 7, pp. 228–266.

[36] Nitesh V Chawla et al. "SMOTE: synthetic minority over-sampling technique". In: *Journal of artificial intelligence research* 16 (2002), pp. 321–357.

[37] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[38] Sebastian Ruder. "An overview of gradient descent optimization algorithms". In: *arXiv preprint arXiv:1609.04747* (2016).

[39] Yi Sun et al. "Deep learning face representation by joint identification-verification". In: *Advances in neural information processing systems*. 2014, pp. 1988–1996.

[40] Frank Seide et al. "Feature engineering in context-dependent deep neural networks for conversational speech transcription". In: *2011 IEEE Workshop on Automatic Speech Recognition & Understanding*. IEEE. 2011, pp. 24–29.

[41] Emma Strubell, Ananya Ganesh, and Andrew McCallum. "Energy and Policy Considerations for Deep Learning in NLP". In: *arXiv preprint arXiv:1906.02243* (2019).

[42] C Bradford Barber, David P Dobkin, and Hannu Huhdanpaa. “Qhull: Quickhull algorithm for computing the convex hull”. In: *Astrophysics Source Code Library* (2013).

[43] Songrit Maneewongvatana and David M Mount. “It’s okay to be skinny, if your friends are fat”. In: *Center for Geometric Computing 4th Annual Workshop on Computational Geometry*. Vol. 2. 1999, pp. 1–8.