

# Few Data Science and Machine Learning Research Areas

## **Machine Learning Foundation and Theory:**

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1), 1-47.

Handelman, G. S., Kok, H. K., Chandra, R. V., Razavi, A. H., Huang, S., Brooks, M., ... & Asadi, H. (2018). Peering Into the Black Box of Artificial Intelligence: Evaluation Metrics of Machine Learning Methods. *American Journal of Roentgenology*, 1-6.

## **Classification Algorithms**

Joachims, T. (2006, August). Training linear SVMs in linear time. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 217-226). ACM.

Razavi, A. H., Inkpen, D., Uritsky, S., & Matwin, S. (2010, May). Offensive language detection using multi-level classification. In *Canadian Conference on Artificial Intelligence*(pp. 16-27). Springer, Berlin, Heidelberg.

Toutanova, K., Klein, D., Manning, C. D., & Singer, Y. (2003, May). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1* (pp. 173-180). Association for Computational Linguistics.

## **Topic Modeling and Clustering**

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.

Chaney, A. J. B., & Blei, D. M. (2012, March). Visualizing Topic Models. In *ICWSM*.

Sinoara, R. A., Scheicher, R. B., & Rezende, S. O. (2017, November). Evaluation of latent dirichlet allocation for document organization in different levels of semantic complexity. In *Computational Intelligence (SSCI), 2017 IEEE Symposium Series on* (pp. 1-8). IEEE.

Wilamowska, K., & Manic, M. (2001). Unsupervised pattern clustering for data mining. In *Industrial Electronics Society, 2001. IECON'01. The 27th Annual Conference of the IEEE*(Vol. 3, pp. 1862-1867). IEEE.

## **Digital Humanities:**

Saeedi, P., Faili, H., & Shakery, A. (2014). Semantic role induction in Persian: An unsupervised approach by using probabilistic models. *Literary and Linguistic Computing*, 31(1), 181-203.

Herbelot, A. (2014). The semantics of poetry: A distributional reading. *Digital Scholarship in the Humanities*, 30(4), 516-531.

Mahmoudi, M. R., & Abbasalizadeh, A. (2018). On comparing and clustering the alternatives of love in Saadi's lyric poems (Ghazals). *Digital Scholarship in the Humanities*.

### **Knowledge Representations and Graph Queries**

Zhu, G., & Iglesias Fernandez, C. A. (2015). Sematch: semantic entity search from knowledge graph.

Najmi, E., Malik, Z., Hashmi, K., & Rezgui, A. (2016, April). ConceptRDF: An RDF presentation of ConceptNet knowledge base. In *Information and Communication Systems (ICICS), 2016 7th International Conference on* (pp. 145-150). IEEE.

Speer, R., & Havasi, C. (2012, May). Representing General Relational Knowledge in ConceptNet 5. In *LREC* (pp. 3679-3686).

### **Visualization**

Haidar, A. A., Yang, B., & Ganascia, J. G. (2016, July). Visualizing the First World War Using StreamGraphs and Information Extraction. In *Information Visualisation (IV), 2016 20th International Conference* (pp. 290-293). IEEE.

Yang, B., & Ganascia, J. G. (2017). Creating knowledge maps using Memory Island. *International Journal on Digital Libraries*, 18(1), 41-57.

### **Deep Learning**

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).

Goodfellow, I. (2016). NIPS 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*.

Salehi, M., Pourzaferani, M., & Razavi, S. A. (2013). Hybrid attribute-based recommender system for learning material using genetic algorithm and a multidimensional information model. *Egyptian Informatics Journal*, 14(1), 67-78.