Syllabus for Anomaly Detection course

Lecturer: **Dr. Eliahu Khalastchi**

Goal:

The goal of this course is to introduce the anomaly detection problem, key challenges and applications, and to survey the different anomaly detection approaches and real-world problems. In addition, the course provides the elementary skill of creating an anomaly detector and investigate its achievements.

Prerequisites: introduction to AI (89-570)

Recommended to take in **parallel**: machine learning (89-511), Statistical methods in computer science (89-362)

Assignments and grading:

The students are required to present their project at the end of the semester. The project includes choosing an anomaly detection problem and a suitable a data set, code an anomaly detector that fits this problem, and investigate its accuracy and performance. The grade is given according to the presentation: Describe the data (20%), Key challenges (10%), Approach outline (50%), Results (20%).

Schedule:

	Subject	**Relevant papers**
1	Introduction to Anomaly Detection • What are anomalies • Applications of anomaly detection • Related problems • Key challenges (and online vs. offline) • Types of input data • Types of anomalies • Taxonomy of approaches • Evaluation of an anomaly detector	Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey." ACM computing surveys (CSUR) 41.3 (2009): 15.
Anomaly Detection Approaches		
2	Classification based • Rule based • Neural Networks • SVM	Görnitz, N., Kloft, M. M., Rieck, K., & Brefeld, U. (2013). Toward supervised anomaly detection. Journal of Artificial Intelligence Research. Agrawal, S., & Agrawal, J. (2015). Survey on anomaly detection using data mining techniques. Procedia Computer Science, 60, 708-713.

3-4	Nearest Neighbor based • Density based vs. Distance based • k-NN Global Anomaly Score • Local Outlier Factor (LOF) • Connectivity based Outlier Factor (COF) • Local Outlier Probability (LoOP) • Influenced Outlierness (INFLO) • Local Correlation Integral (LOCI) Clustering based • Cluster based Local Outlier Factor (CBLOF) • Local Density Cluster based Outlier Factor (LDCOF)	Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey." ACM computing surveys (CSUR) 41.3 (2009): 15. Amer, M., & Goldstein, M. (2012). Nearest-neighbor and clustering based anomaly detection algorithms for rapidminer. In Proc. of the 3rd RapidMiner Community Meeting and Conference (RCOMM 2012) (pp. 1-12).
5	Statistical • Parametric • Non-parametric • Outlier detection on temporal data	Gupta, M., Gao, J., Aggarwal, C. C., & Han, J. (2014). Outlier detection for temporal data: A survey. IEEE Transactions on Knowledge and Data Engineering, 26(9), 2250-2267. Hodge, V., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial intelligence review*, *22*(2), 85-126.
6	Model based • FDI • DX	Travé-Massuyès, L. (2014). Bridging control and artificial intelligence theories for diagnosis: A survey. Engineering Applications of Artificial Intelligence, 27, 1-16.
7	Graph based • Static graphs, Dynamic graphs • Feature, proximity, community based • Rational learning based • decomposition based • real-world applications • research opportunities	Akoglu, L., Tong, H., & Koutra, D. (2015). Graph based anomaly detection and description: a survey. Data Mining and Knowledge Discovery, 29(3), 626-688.
8	Special • Information Theory based • Spectral decomposition based • Visualization based • Sonification based(?)	Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey." ACM computing surveys (CSUR) 41.3 (2009): 15.
Real-world problems		
9	Network intrusion detection • Types of cyber-attacks on networks • Network Anomaly Detection techniques	Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2014). Network anomaly detection: methods, systems and tools. Ieee communications surveys & tutorials, 16(1), 303-336.

		Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, *60*, 19-31.
10	Anomaly Detection in BIG DATA • The nature of big data • Key challenges • Relevant technologies • Relevant approaches and examples	Camacho, J., Macia-Fernandez, G., Diaz-Verdejo, J., & Garcia-Teodoro, P. (2014, April). Tackling the big data 4 vs for anomaly detection. In IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2014 (pp. 500-505). IEEE.
11	Anomaly detection for autonomous robots • Key challenges • Offline learning • Online learning	Khalastchi, E., Kalech, M., Kaminka, G. A., & Lin, R. (2015). Online data-driven anomaly detection in autonomous robots. Knowledge and Information Systems, 43(3), 657-688.
12-13	Students present their projects 10min each • Describe the data (20%) • Key challenges (10%) • Approach outline (50%) • Results (20%)	