Ph.D. in Information Technology
Thesis Defense

November 29, 2022
at 11:00
Room Alpha

Luca FRITTOLI – XXXIV Cycle

ADVANCED LEARNING METHODS FOR ANOMALY DETECTION IN MULTIVARIATE DATASTREAMS AND POINT CLOUDS

Supervisor: Prof. **Giacomo Boracchi**

Abstract:

Anomaly detection is a challenging problem in several application domains, ranging from industrial quality control to cryptographic attacks. In the literature, several statistical and deep learning models for anomaly detection have been proposed, each underpinning specific assumptions on the nature of the data to be analyzed. This thesis presents new solutions for the anomaly-detection problem in two different settings.

First, we assume that, in normal conditions, data samples are realizations of a random vector. We focus on a change-detection problem, where the goal is to detect permanent changes in the data-generating process by analyzing a datastream acquired over time. We develop QuantTree Exponentially Weighted Moving Average (QT-EWMA), an online and nonparametric change-detection algorithm for multivariate datastreams that can be configured to maintain the target Average Run Length (ARL0), namely the expected time before a false alarm. We extend our work to the concept-drift scenario, where the data samples are the object of a classification problem. We propose Class Distribution Monitoring (CDM), where we employ multiple instances of QT-EWMA to detect changes in the class-conditional distributions of annotated datastreams. As a new change-detection application, we address the problem of detecting errors in sequential cryptographic side-channel attacks. These attacks reconstruct a private key one bit at a time by using a distinguisher involving, e.g., the power consumption of the target device. We propose an error-detection and correction procedure based on a standard change-detection algorithm applied to the univariate datastream formed by distinguisher values.

Then, we address anomaly detection in point clouds, namely lists of the coordinates of points describing, for instance, the surface of a 3D object. We aim to assess whether individual point clouds belong to a certain normal class. The main challenge of handling point clouds is their lack of a grid structure, which prevents the use of traditional Convolutional Neural Networks (CNNs). We propose the composite layer, an effective and flexible operator for point cloud processing in Deep Neural Networks (DNNs). We use our composite layers to implement CompositeNets, which are DNNs for point cloud classification. Most remarkably, we are among the first to address anomaly detection in point clouds by training our CompositeNets in a self-supervised fashion. As an anomaly-detection application, we analyze Wafer Defect Maps (WDMs), i.e., the lists of the 2D coordinates of defects on silicon wafers manufactured by STMicroelectronics. We cast anomaly detection as an open-set recognition problem, where the goal is to correctly classify known defect patterns and detect anomalous patterns in WDMs. The coordinates in WDMs lie on a huge grid, which prevents the use of CNNs. To efficiently process WDMs at full resolution, we train a Submanifold Sparse Convolutional Network (SSCN) on known classes. To detect anomalous patterns, we apply an outlier detector based on a Gaussian Mixture Model (GMM) to the latent representation of the SSCN.

PhD Committee

Prof. **Matteo Matteucci**, DEIB-Politecnico di Milano
Prof. **Alessandro Giusti, Idsia**, Usi-Supsi, Lugano
Dr. **Cristiano Cervellera**, CNR