Contribution ID: **57** Type: **not specified**

Active Learning for Budget-Constrained Labeling in Data Stream Monitoring

Modern industrial systems generate high-dimensional data streams often used for statistical process monitoring (SPM), i.e., distinguishing between multiple in-control and out-of-control (OC) states. While supervised SPM methods benefit from labeled data in assessing the process state, label acquisition is often expensive and infeasible at large scale. This work proposes a novel stream-based active learning framework for SPM that optimally selects data points to label under a constrained budget. Unlike traditional active learning methods, which assume independent data, our approach explicitly models temporal dependencies by integrating partially hidden Markov models to combine labeled and unlabeled information. The proposed method addresses both class imbalance and the emergence of previously unseen OC states. A dual criterion is developed to balance exploration (i.e., discovering new OC conditions) and exploitation (i.e., improving classification accuracy on known states). The labeling strategy operates in real-time, providing decisions for each incoming data point. Through a simulation study and a case study on resistance spot welding in the automotive industry, the proposed method is demonstrated to improve SPM performance, especially when labeling resources are scarce.

Acknowledgements: The research activity of C. Capezza and A. Lepore was supported by Piano Nazionale di Ripresa e Resilienza (PNRR) - Missione 5 Componente 2, Investimento 1.3-D.D. 1551.11-10-2022, PE00000004 within the Extended Partnership MICS (Made in Italy - Circular and Sustainable). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Special/ Invited session

Classification

Both methodology and application

Keywords

Statistical Process Monitoring, Hidden Markov Model, Sequential Data Analysis, Imbalanced Classification

Primary authors: CAPEZZA, Christian (Department of Industrial Engineering, University of Naples Federico II); LEPORE, Antonio (Department of Industrial Engineering, University of Naples Federico II); PAYNABAR, Kamran (School of Industrial and Systems Engineering, Georgia Tech)

Presenter: CAPEZZA, Christian (Department of Industrial Engineering, University of Naples Federico II)

Track Classification: Statistical Process Monitoring