UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

Dottorato di Ricerca in Ingegneria dell'Informazione

Doctoral Course

Generative Models for Intrusion Detection and Cybersecurity Analytics

Dr. Siphesihle Sithungu

University of Johannesburg – South Africa

E-mail address: siphesihles@uj.ac.za

Short Abstract: This course introduces generative modelling approaches for anomaly detection in cybersecurity systems. Modern intrusion detection increasingly relies on modelling the probability distribution of normal system behaviour and identifying deviations from this distribution. Generative models (including probabilistic models, autoencoders, variational approaches, and adversarial learning) provide a principled framework for learning such representations from complex data.

The course will cover the theoretical foundations of generative modelling, the role of probability distributions in anomaly detection, and practical approaches to applying generative models to cybersecurity data. Students will gain both conceptual understanding and practical experience implementing generative models for intrusion detection.

Course Contents in brief:

- Foundations of anomaly detection and intrusion detection systems
- Probabilistic modelling of normal system behaviour
- Classical generative models for anomaly detection
- Autoencoders and deep generative architectures
- Variational Autoencoders and latent variable models
- Generative Adversarial Networks
- Evaluation of generative intrusion detection systems

Total # of hours of lecture: 16

References:

[1] Christopher M. Bishop. Pattern Recognition and Machine Learning. Springer, 2006.

[2] Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press, 2012.

[3] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016.

[4] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. International Conference on Learning Representations (ICLR), 2014.

[5] Ian Goodfellow et al. Generative Adversarial Networks. NeurIPS, 2014.

[6] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly Detection: A Survey. ACM Computing Surveys, 2009.

[7] Robin Summer and Paxson and Vern Paxson. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. IEEE Symposium on Security and Privacy, 2010.

[8] Nicolás Moustafa and Jill Slay. UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems. Military Communications and Information Systems Conference, 2015.

[9] Lukas Ruff et al. Deep One-Class Classification. International Conference on Machine Learning (ICML), 2018

CV of the Teacher

SIPHESIHLE P. SITHUNGU received his BSc Honours degree in Computer Science and Informatics from University of Johannesburg, South Africa, in 2018, his Master's degree in Computer Science in 2020, and his Ph.D. degree in Computer Science in 2024. He is currently a Senior Lecturer and Deputy Head of Department at the University of Johannesburg. He is a member of the International Federation for Information Processing (IFIP) Working Group 12.9 (Computational Intelligence) and BCS – The Chartered Institute for IT. His current research interests include gradient-free generative modelling (how machines can learn and create beyond backpropagation), biologically inspired AI and AI-driven cybersecurity.

Final Exam: method of final examination

- Students design and implement a generative-model-based anomaly detection system on a cybersecurity dataset (e.g. network traffic, system logs or benchmark IDS dataset).
- The project will include a short report discussing the modelling approach, training process and anomaly detection results.

Room and Schedule

Room: **Aula Riunioni del Dipartimento di Ingegneria dell'Informazione, Via G. Caruso 16, Pisa – Ground Floor e Aula Riunioni Piano 6 - L. Lazzarino**

Schedule:

Day 1 (**Monday, 13th April, 9.30 – 12.30**): Foundations of Intrusion Detection and Anomaly Detection (3 hours) - **Via Caruso**

Day 2 (**Tuesday, 14th April, 9.30 – 12.30**): Generative Modelling Foundations (3 hours) - **Via Caruso**

Day 3 (**Wednesday, 15th April, 9.30 – 12.30**): Deep Generative Models for Anomaly Detection (3 hours) - **L. Lazzarino**

Day 4 (**Thursday, 16th April, 12.30 – 15.30**): Advanced Generative Approaches (3 hours) - **L. Lazzarino**

Day 5 (**Friday, 17th, 9.00 – 13.00**): Practical Applications (4 hours) **Via Caruso**