Project in Applied Bioinformatics

Anomaly detection from biomedical signals using unsupervised machine learning

Professor Jan Komorowski
Department of Cell and Molecular Biology

Patients care is at the core of public health services. Each patient that attends a medical consult is routinely physically assessed by recording their medical signs, which in turn is a direct measurement of the patient state. These recordings, from large cohorts, or time series from single individuals, become valuable pieces of information to extract patterns.

Substantial algorithmic advances in the field of machine learning during the last years have been driven largely by a specific class of models known as deep neural networks. Deep neural networks are computational models consisting of multiple processing layers, with each layer being able to learn increasingly abstract, higher-level representations of the input data relevant to perform specific tasks. These models have dramatically improved the state of the art in speech recognition (Amodei et al. 2016), image recognition (He et al. 2015), strategy games such as Go (Silver et al. 2016), and in medical applications (Gulshan et al. 2016; Esteva et al. 2017). The ability of deep neural networks to recognize patterns and learn useful features from raw input data without requiring extensive data preprocessing, engineering or handcrafted rules makes them particularly well suited to interpret biomedical signals. Furthermore, since deep neural networks performance tend to increase as the amount of training data increases, this approach is well positioned to take advantage of the widespread digitization of biomedical records.

Therefore, we propose to implement an unsupervised algorithm to identify anomalies applied to biomedical signals, for instance, electromyography, evoked auditory potentials, among others. The interpretation and automatic evaluation of medical signals through machine learning could potentially improve the sensitivity and specificity of diagnosis on different medical fields. Moreover, this methodology could serve as a non-invasive method for early diagnosis of a wide range of diseases, such as Alzheimer's.

Candidates will require strong programing skills in languages commonly used for data science, R, python or Julia, as well as interest in machine learning algorithms. Experts in computation, medicine and machine learning will support and supervised the proposed project.

Contact:
Daniel Rivas: daniel.rivas@icm.uu.se
Jan Komorowski: jan.komorowski@icm.uu.se

Amodei, Dario, Sundaram Ananthanarayanan, Rishita Anubhai, Jingliang Bai, Eric Battenberg, Carl Case, Jared Casper, Bryan Catanzaro, Qiang Cheng, Guoliang Chen, Jie Chen, Jingdong Chen, Zhijie Chen, Mike Chrzanowski, Adam Coates, Greg Diamos, Ke Ding, Niandong Du, Erich Elsen, Jesse Engel, Weiwei Fang, Linxi Fan, Christopher Fougner, Liang Gao, Caixia Gong, Awni Hannun, Tony Han, Lappi Johannes, Bing Jiang, Cai Ju, Billy Jun, Patrick LeGresley, Libby Lin, Junjie Liu, Yang Liu, Weigao Li, Xiangang Li, Dongpeng Ma, Sharan Narang, Andrew Ng, Sherjil Ozair, Yiping Peng, Ryan Prenger, Sheng Qian, Zongfeng Quan, Jonathan Raiman, Vinay Rao, Sanjeev Satheesh, David Seetapun, Shubho Sengupta, Kavya Srinet, Anuroop Sriram, Haiyuan Tang, Liliang Tang, Chong Wang, Jidong Wang, Kaifu Wang, Yi Wang, Zhijian Wang, Zhiqian Wang, Shuang Wu, Likai Wei, Bo Xiao, Wen Xie, Yan Xie, Dani Yogatama, Bin Yuan, Jun Zhan, and Zhenyao Zhu. 2016. "Deep Speech 2 : End-to-End Speech Recognition in English and Mandarin." In *Proceedings of The 33rd International Conference on Machine Learning*, edited by Balcan Maria Florina and Q. Weinberger Kilian, 173--82. Proceedings of Machine Learning Research: PMLR.

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. 'Dermatologist-level classification of skin cancer with deep neural networks', *Nature*, 542: 115-18.

Gulshan, Varun, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, Ramasamy Kim, Rajiv Raman, Philip C. Nelson, Jessica L. Mega, and Dale R. Webster. 2016. 'Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs', *JAMA*, 316: 2402.

He, K., X. Zhang, S. Ren, and J. Sun. 2015. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification." In *2015 IEEE International Conference on Computer Vision (ICCV)*, 1026-34.

Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. 'Mastering the game of Go with deep neural networks and tree search', *Nature*, 529: 484-89.