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


