Doctoral School of Information and Biomedical Technologies
Polish Academy of Sciences (TIB PAN)

SUBJECT: Information Theory and Machine Learning based approach for effective biomedical signal Classification and Feature extraction

SUPERVISOR:

Ph. D., DSc, Eng. Agnieszka Pregowska (https://www.ippt.pan.pl/en/staff/?osoba=aprego)
Pawinskiego 5B, 02–106 Warsaw, PL
tel.: +48 22 828 53 73/+48 22 826 12 81 ext. 412
e-mail: aprego@ippt.pan.pl

Name of the institute in which the topic will be realized: **Institute of Fundamental Technological Research**

Scientific discipline: **Information and communication technology**

PROJECT DESCRIPTION

Currently, substantial efforts are being developed for the enrichment of medical imaging applications using these algorithms to diagnose the errors in disease diagnostic systems which may result in extremely ambiguous medical treatments [1]. Biomedical signals (EEG, EKG, MEG, MRI, etc.) can be analyzed in several ways using different mathematical theories and tools (Statistical analysis, Fourier analysis, Wavelet analysis). Recently, Information Theory is strongly exploited. Concepts derived from **Information Theory** such as *entropy, discrete entropy, Renyi entropy, permutation entropy, mutual information, complexity, permutation complexity, and discrete Lyapunov exponents* allow us a deeper insight into the nature of biosignals. This is because they take into account simultaneously the probabilistic nature and the structures of signals (e.g. internal patterns of signals [2–5]). Besides, these techniques have a unique feature that enables measuring the level of randomness of signals and distinguishing deterministic traits from random features. This is extremely important in the classification of such complex signals as the ECG, EEG, MEG, and MRI especially in the early stages of anomaly formation when detection of subtle differences is crucial [6, 7]. Based on our previous information-theoretic research [8, 9] we set the following research hypothesis that Information Theory can be successfully applied as a **clinical tool for vital signs abnormalities classification**. Before applying the Information Theory-based tools, biosignals have to be converted into a discrete sequence of symbols. It is known that the effectiveness of the signals classifications method strongly depends on the signal digitalization applied [10, 11]. Codification (digitalization) can take place in various ways. A **successful encoding method** of biomedical signals (e.g. a sequence of measurements performed with some sampling frequency) into sequences of symbols from a given finite alphabet is one of the important issues. Specifically, in choosing a biosignal digitization method, the challenge is also to adopt the most effective parameters, and here the use of the Machine Learning approach is the most promising idea.

The Doctoral Thesis aims to develop and implement **new effective classification algorithms based on Information Theory concepts and** support these algorithms with **Machine Learning techniques**. Such algorithms should allow to **analysis and classification effectively biomedical signals online**. This software will be responsible for analyzing patient data from an electrocardiogram, heartbeat sensor, and other relevant medical data, classifying subjects' exam results, and providing a suggestion about his/her health status.

Diagnostics algorithms will be validated on:

- simulated signals modeling experimental recordings and on the *in vivo* recordings, e.g. electrical heart activity and brain signals (individual sensory neurons) and
- on signals coming from an experimental database of ***Mount Sinai Hospital New York, USA, IDIBAPS Institut d'investigacions Biomèdiques August Pi i Sunyer Barcelona, Spain,*** and ***Peacs BV Netherland***.

The results obtained will support **intelligent monitoring of patients** and help deliver **targeted, and personalized medicine** while providing smooth communication and high productivity in medical units.

BIBLIOGRAPHY

[1] **May M,** Eight ways machine learning is assisting medicine. Nat Med. 27(1):2-3, 2021.
[2] **Shannon C**, A mathematical theory of communication, Bell Labs Tech. J 27,379–423, 623–656,1948.
[3] **Cover TM, Thomas JA**, Elements of Information Theory, A Wiley-Interscience Publication, New York, USA, 1991.
[4] **Bossomaier T, Barnett L, Harre M, Lizier JT**, An introduction to transfer entropy. Springer,2016.
[5] **Amigó JM, Small M**, Mathematical methods in medicine: neuroscience, cardiology and pathology. Philos. Trans. R. Soc. A 375(2016),20170016,2017.
[6] **Shumbayawonda E, Fernández A, Hughes MP, Abásolo D**, Permutation entropy for the characterisation of brain activity recorded with magnetoencephalograms in healthy ageing, Entropy 19(4),141,2017.
[7] **Faust O, Hagiwara Y, Hong TJ, Lih OS, AcharyaUR**, Deep learning for health- care applications based on physiological signals: a review, Comput. Meth. Prog. Bio. 161, 1–13,2018.
[8] **Pregowska A, Kaplan E, Szczepanski J,** How Far can Neural Correlations Reduce Uncertainty? Comparison of Information Transmission Rates for Markov and Bernoulli Processes, Int. J. Neural. Syst. 29, 1950003–1–13,2019.
[8] **Pregowska A, Proniewska K, van Dam P, Szczepanski** J, Using Lempel-Ziv complexity as effective classification tool of the sleep-related breathing disorders, Comput. Meth. Prog. Bio. 182, 105052–1–7, 2019.
[10] **Borst A, Theunissen FE**, Information theory and neural coding, Nature Neuroscience 2(11), 947–957, 1999.
[11] **Gerstner W, Kistler WM, Naud R, Paninski L**, Variability of spike trains and neural codes, in Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition, 168-201, Cambridge,2014.