Project No:

Department of Communication Systems

Project Title: Data Mining of Biomedical Spectroscopy data

Degree Scheme: CCS ☒ ITM ☒ ECS ☒ TCS ☒

Level: BSc/BEng ☒   MSci/MEng ☐

Supervisor: Dr. P Angelov   2nd Supervisor: Dr. F Martin
Assistant: Mr. J Trevisan

Project Description:

This project will focus on the use of classification algorithms (written in Matlab) for automatic classification of spectra produced by Fourier transformed signals from Infrared (IR) microscopy. The data are produced by the researchers at the Department of Biological Studies in collaboration with the company Uniliver, USA. The aim is to automatically classify and identify the types of cell based on the spectra of their absorption of the infrared microscopy. Different types of classifiers (such as k nearest neighbours, kNN; hierarchical, fuzzy classifier, support vector machines will be used) and results compared on a range of metrics (classification rate, time, complexity).

Classification has been used in bio-medical data mining for some time. The recent developments in microscopy made possible to extract more information about the composition of the cells (including stem cells, cancer affected cells etc.) than what can be done using optical microscopy (image). Different cells absorb differently the infra-red (IR) waves. The spectrum of each cell can be processed and digitized by considering different intervals of the frequency. These results can be fed into a classifier that will automatically classify the spectrum (and respectively the cell) into one of the two or three categories that have been pre-defined. The experiments have been performed and data collected. The student will use the software that has been developed to the real data collected recently.

This work is novel, and of interest to industry (Unilever Co, USA) which is an excellent opportunity for a possible career in research or related industry.

Skills required:

Data management, Matlab, data handling, processing and mining, classification.

References:

[1] Grude, O., Hammiche, a., Pollock, H., Bentley, A., Walsh, M., Martin, F. and Fullwood, N.: 2007, Near-field photo-thermal micro-spectroscopy for adult stem-cell identification and characterization, *Journal of Microscopy* 228, 366–372.
[2] *Kelly, J., Angelov, P., M. J. Walsh, H. M. Pollock, M. A. Pitt, P. L. Martin-Hirsch, F. Martin, A Self-Learning Fuzzy Classifier with Feature Selection for Intelligent Interrogation of mid-IR Spectroscopy Data Derived from Different Categories of Exfoliative Cervical Cytology, International Journal on Computational Intelligence Research, special issue on the future of fuzzy systems research, invited paper, to appear 2009* (available on request)
[3] Angelov, P., X. Zhou, F. Klawonn, Evolving Fuzzy Rule-based Classifiers, *First 2007 IEEE International Conference on Computational Intelligence Applications for Signal and Image Processing*, April 1-5, 2007, Honolulu, Hawaii, USA, pp.220-225.