Application of data mining methods in detecting of bacteria proliferation syndrome in the small intestine

Arnis Kirshners[1] and Sergei Parshutin[1]

Riga Technical University, Institute of Information Technology
arnis.kirsners@rtu.lv, sergei.parshutin@rtu.lv

Abstract. The paper presents a data mining approach as a possible solution for a medical problem of detecting bacteria proliferation syndrome in the small intestine by analysing glucose test results and patient survey data about intestinal tract. The proposed approach initially determines groups of similar objects - clusters, by analysing short time series given by the glucose test results; and then merges it with patient survey data having only most informative attributes. The obtained data set was used to find relationships between clustering results and the descriptive parameters using classification algorithms. The obtained classifiers and patient self-assessment data served as a basis to determine whether the patient has to be assigned an advanced bacteria proliferation syndrome test in the small intestine or not. The article presents analysis of the acquired classifier evaluation results, as well as provides practical recommendations for medical experts that would use the proposed approach to detect bacteria proliferation syndrome in the small intestine.

References

LUPASCU, A., GABRIELLI, M., LAURITANO, E.C., et. al. (2005): Hydrogen glucose breath test to detect small intestinal bacterial overgrowth: a prevalence case-control study in irritable bowel syndrome. *Aliment Pharmacol Ther, 22: 1157-1160.*

POSSERUD, I., STOTZER, P.O., BJORNSSON, E.S., ABRAHAMSSON, H. and SIMREN, M. (2007): Small intestinal bacterial overgrowth in patients with irritable bowel syndrome. *Gut, 56: 802-808.*

THOMASSEY, S. and FIORDALISO, A. (2006): A hybrid sales forecasting system based on clustering and decision trees. *Decision Support Systems, Issue 1, Volume 42, 408-421.*

Keywords

SHORT TIME SERIES, CLUSTERING, CLASSIFICATION, BACTERIA PROLIFERATION SYNDROME