Czech Technical University in Prague
Faculty of Electrical Engineering

Department of Cybernetics

BACHELOR PROJECT ASSIGNMENT

Student: Tomáš K a l a

Study programme: Open Informatics

Specialisation: Computer and Information Science

Title of Bachelor Project: Event Detection from Text Data

Guidelines:

1. Get familiar with the topic of event detection from potentially large text collections.
2. Reimplement the method of He et al. in Python and test it on a dataset provided by the thesis supervisor.
3. Propose and implement modifications of this algorithm. Consider changes in the cost function, utilization of document clustering with consequent topic dependent event detection and application of word/document embedding.
4. Extend the algorithm to be able to annotate the individual events and organize them.
5. Compare the results reached in the previous steps. Employ the list of real events given by the thesis supervisor to make the comparison as objective as possible.

Bibliography/Sources:

[1] He, Qi, Kuiyu Chang, and Ee-Peng Lim. "Analyzing feature trajectories for event detection." Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2007.
[2] Fung, Gabriel Pui Cheong, et al. "Parameter free bursty events detection in text streams." Proceedings of the 31st international conference on Very large data bases. VLDB Endowment, 2005.
[3] Mikolov, Tomas, Chen, Kai, Corrado, Greg, and Dean, Jeffrey. "Efficient Estimation of Word Representations in Vector Space". In Proceedings of Workshop at ICLR, 2013.
[4] Zhong, Shi. "Efficient online spherical k-means clustering." In Proceedings of the IEEE International Joint Conference on Neural Networks, 2005., vol. 5, pp. 3180-3185, 2005.
[5] Atefeh, Farzindar, and Wael Khreich. "A survey of techniques for event detection in twitter." Computational Intelligence 31.1, pp. 132-164, 2015.

Bachelor Project Supervisor: doc. Ing. Jiří Kléma, Ph.D.

Valid until: the end of the summer semester of academic year 2017/2018

L.S.

prof. Dr. Ing. Jan Kybic
Head of Department

prof. Ing. Pavel Ripka, CSc.
Dean

Prague, January 12, 2017