Log Files Processing by Machine Learning Techniques

M.Tsintsadze, M.Archuadze, M.Khachidze

magda.tsintsadze@tsu.ge, maia.archuadze@tsu.ge, manana.khachidze@tsu.ge

Department of Computer Sciences Faculty of Exact and Natural Sciences

Iv. Javakhishvili Tbilisi State University

University str., 13, Georgia

The paper discusses an example of how machine learning is used to automatically detect system failures. To understand the basic concept of machine learning for log file analysis, an example is discussed where there are three sources of data: 1) system counters, CPUs, hard drives, and network; 2) large number of diluted log files accumulated from system programs; 3) Error logs (application errors etc). For all information from these input sources, the appropriate type is determined using the Bayesian algorithm to categorize the logs into the corresponding solution (learn). As a result, the newly logged logs automatically fall into the appropriate category, making it easy to spot a specific error.

 Referance:

1. Peters, T. (1993). The history and development of transaction log analysis. Library Hi Tech., 42(11), 41−66
2. Anton A. Chuvakin, Kevin J. Schmidt. Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management 1st Edition. 2013
3. Nils J. Nilsson, Introduction to Machine Learning.