UNIVERSITÄT SIEGEN ● Fakultät IV / Ubiquitous Computing ● 57076 Siegen

Prof. Dr. Kristof Van Laerhoven
http://ubicomp.eti.uni-siegen.de
kvl@eti.uni-siegen.de
Tel.: +49 (0) 271/ 740-2312
Hölderlinstraße 3, D-57076 Siegen

Sekretariat: Dagmar Rautmann
Tel.: +49 (0) 271/ 740-3037
Fax: +49 (0) 271/ 740-2532

Siegen, 24.02.2020

Master's Thesis
Data Mining & Natural Language Processing

Supervised by Florian Wolling
florian.wolling@uni-siegen.de
In cooperation with WBS, Johannes Zenkert.

Master's Thesis

Data Mining for the Automated Market Analysis of Wearable Devices

Motivation:
In recent years, innumerable wearables flooded the market of consumer electronics with devices for fitness tracking and health care. All of them advertise with different features and are based on various technologies. To gather an overview of the relevant devices at the market and their development in the last 15 years, an automated market analysis tool is required. It is based on a web crawler that searches for devices and their related description in the internet. Common websites for price comparison and technology review provide all necessary information, but in a continuous text of natural language. Consequently, Natural Language Processing (NLP) methods have to be applied to interpret the description text for Data Mining.

The key research questions are:
1. How many wearable devices reached the market in each year?
2. How did the market entrance price develop in recent years?
3. Which device-specific technologies have been used?

The research and analysis of this thesis concentrates on wearable devices that use the optical measurement principle photoplethysmography (PPG). It enables to non-invasively monitor the pulsating blood volume flow in the microvascular bed of the tissue beneath the skin. In recent years, this technique has undergone a revival and is nowadays standard in wrist-worn devices for ambulant self-monitoring and fitness tracking. While the first devices primarily utilized infrared or red light, modern devices apply green or even yellow and blue light wavelength.

Requirements:
Fundamental knowledge in Natural Language Processing, Data Mining, and Python programming. Independent working and creative thinking are essential.