Abstract

mgr Katarzyna Wójcik

Opinion mining is a field of research that can have a significant impact on today's business. Many consumers make decisions after searching the Internet for other users' opinions. From the business management point of view, the knowledge derived from such opinions is important and can influence the decisions made. Thanks to the opinion mining, managers can easily and quickly obtain information about how the company and its products are perceived by consumers. To analyze consumer opinions, as a specific type of texts expressing the subjective attitude of the author of a statement to its subject, methods of exploratory text analysis are used, including those based on the measurement of similarity or distance between text documents.

The dissertation consists of four chapters. The first one is devoted to defining and characterizing consumer opinions in the context of the information society. This chapter presents the areas of science for which information is the subject of research and economic activity, in which information has practical application. Next, the definitions of data, information and knowledge are cited as successive concepts in the hierarchical model of information. The last part of the chapter is devoted to the role of consumer opinion as a carrier of information in the information society, and as a subject of research in broadly understood management.

The second chapter is focuses on the issues of text documents and the measurement of similarity between them. The concept of a text document and the concepts of text and text file are defined there. Then, the areas of text mining and natural language processing are also defined and briefly characterized. In the further part of the chapter, the concepts of similarity and distance are defined in the context of data analysis. The last part of the chapter is devoted to discussing various methods of representing text documents in computer systems and determining the similarity between them in this context.

The third chapter of the dissertation deals with the opinion mining. This part of the paper discusses the process of opinion mining along with various approaches and research methods used in its implementation. Ways of taking external knowledge into account in the opinion mining are mentioned here. The next part of the chapter is devoted to the presentation of the author's model of automatic analysis of consumer opinions. The last part of this chapter concerns the assessment of the usefulness of text documents similarity measures in the opinion mining.

The fourth and last chapter of the dissertation is a description of the empirical research carried out. It includes both a description of the research procedure and a discussion of the results obtained. The whole work ends with conclusions and verification of the implementation of the objectives of the dissertation and the hypotheses put forward in it.