

## РАЗДЕЛ 3. ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ЗАЩИТА ИНФОРМАЦИИ

### IMPLEMENTATION OF THE CONTEXT-DEPENDENT SENTIMENT DICTIONARY

*V.I. Karyukin, S.B. Yessenzhanova*

**Annotation.** Social networks provide an opportunity to regulate involvement of citizens in conduct of public affairs, discussion of legal norms and public services. News posts and comments of a different nature of statements are divided into categories of positive, negative and neutral assessment. To automatically determine the sentiment of texts and sentences, sentiment dictionaries and machine learning algorithms approaches are used. In the first approach, determination of the sentiment of texts is done with manually designed sentiment dictionaries. In machine learning classification algorithms, training is done on texts initially labelled by experts.

**Methodology.** The use of both mentioned approaches allows to build a single flexible sentiment dictionary with dynamic content filling. The developed model of the dictionary makes it possible to label the sentiment of words initially defined as neutral. It is required to implement a number of important steps for constructing a context-dependent sentiment dictionary: obtaining data from Internet sources, labelling texts by experts, data preprocessing, classification with machine learning algorithms and building a context-dependent sentiment dictionary.

**Realization.** Texts from news portals and social networks were obtained using a web crawler [2] and placed in OMSystem database. The texts were labelled by an expert group on thematic categories: politics, society, economics, finance, culture, sports, technology, etc. and sentiment aspects: positive and negative. Words in the dictionary in OMSystem were divided into positive, negative and neutral ones. We looked for the neutral words in texts and counted a number of positive and negative texts for each word. To determine if a given word in the specific category was positive or negative, we had to compare the difference between positive and negative values. The greater the difference is, the more accurately we can say that this word defines the predominant category of the emotional component.

#### **References:**

1. Kumar, S. Nezhurina, M.I. Sentiment Analysis on Tweets for Trains Using Machine Learning. 10th International Conference on Soft Computing and Pattern Recognition, SoCPaR 2018; Porto, vol. 942, pp. 94-104. Springer Nature, Switzerland (2020).
2. Chaitra P.G., Deepthi V., Vidyashree K.P., Rajini S. (2020) A Study on Different Types of Web Crawlers. Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing, vol 989. Springer, Singapore (2019).
3. Hilman Wisnu, Muhammad Afif and Yova Ruldevyan. Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes. The 8<sup>th</sup> Engineering International Conference, Journal of Physics: Conference Series, Semarang; Indonesia, 2019, vol. 1444. IOP Publishing Ltd (2020).

### AN APPROACH TO BUILDING ESM APPLICATION

*V.I. Karyukin*

**Annotation.** Social networks [1] provide an opportunity to regulate involvement of citizens in the conduct of public affairs, discussion of legal norms, bills, public services and the activities of public authorities [2].

The review of existing analytical systems. There exists a number of monitoring systems of news portals and social networks: iMAS, Alem media monitoring, OMSystem, Brand Analytics, Microsoft Engagement and others [3]. Developed at al-Farabi Kazakh National University, OMSystem