ABSTRACT

Bachelor thesis: 81 pages, 4 parts, 6 tables, 4 figures, 2 additions, 10 sources.

DATA MINING, DECISION SUPPORT SYSTEM, MODEL, NAIVE BAYES CLASSIFIER, NEURAL NETWORK KERAS, PATIENT COMPLAINTS, PRIMARY DIAGMOSIS, RANDOM FOREST.

Object of research – a large sample of data, consisting of patient complaints, diagnoses, delivered by doctors and their corresponding codes in ICD-10.

Objective – to perform study and application of methods for the analysis of text and large data sets for diagnosis, based on the complaints of the patient to improve the work of physicians and provide additional services to customers.

Methods of research – Keras neural network, random forest method, naive Bayes classifier.

A software product was developed for analyzing and categorizing a patient's complaint based on a model built using a neural network. The project presents the results of categorizing the patient's complaint using the neural network, the random forest method and the naive Bayes classifier.

The system is implemented using software based RStudio shiny library using the programming language R. Examples of application programs to classify patients' complaints. The ways of possible further improvement of the system are considered.