ALGORITHMIC SUPPORT OF A PERSONAL VIRTUAL ASSISTANT FOR AUTOMATING THE PROCESSING OF CLIENT REQUESTS

Konstantin Dobratulin[1], Marina Nezhurina[1]

INTRODUCTION

This article describes creating algorithmic support for the functioning of a personal virtual assistant, which allows automating the processing of customer requests. The study aims to reduce errors and processing time for a client request in business systems - text chats or voice channels using a text transcription system. The results of the development of algorithmic support and an assessment of the quality of work on synthetic data presented.

The development of personal virtual assistants has made it possible to expand the range of possibilities for using assistants in solving business problems, to ensure communication between business and the end user as if communication were built between two people. The performance of not only routine daily tasks, but also complex business logic, assigned to personal virtual assistants, has created a demand for application in all areas of business.

MODELING THE PROCESS OF OBTAINING AND ANALYZING DATA

Imagine the model of receiving and analyzing data in an automated system as a chain of processes going from the stage of receiving a request from the user to sending a response to the user. First, the primary source of data from the end user is a user request to the system, transmitted in one channel of communication with a personal virtual assistant, whether it be a voice or text request. If communication takes place in a voice channel, it must present the data for further algorithmic processing in a form understandable for algorithmic support, in some machine representation. Speaking of such a machine representation, we can mean the translation of information in the natural language of the user into a textual or indicative form. Receiving information in this form from the user, it becomes possible to respond to the user's request - to decide in an automated system that the user wanted to request and provide him with information or provide a service as part of a business service.

For example, if a user requested from a personal virtual assistant an action from a business system in natural language, showing a specific action ("turn on", "say", "I have a question about ..."), the personal virtual assistant must answer the request or perform an action (turn on a specific service, answer a question, delegate decision making to another business system). Fig. 1 shows the processing a user request by a personal virtual assistant.

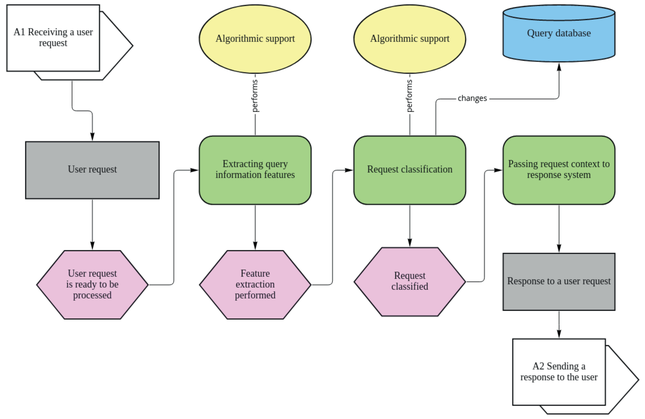

Fig. 1. Processing a user request by a personal virtual assistant.

SELECTION OF SOFTWARE COMPONENTS FOR ALGORITHMIC SUPPORT

Many large companies use the scikit-learn library to create software that uses machine learning algorithms and is also used in the scientific community for research and experimentation. When working with user requests, it becomes possible to build algorithmic support that uses information about the request, for example, by processing the user's natural speech in a textual representation. As an open-source library for working with natural language, we will use the fastText library developed by Facebook. In the set of technologies, it is necessary to formalize the format for receiving and transmitting data, as well as the possibility of transferring data to the database. Fig 2. represent scheme of interaction of algorithmic support.

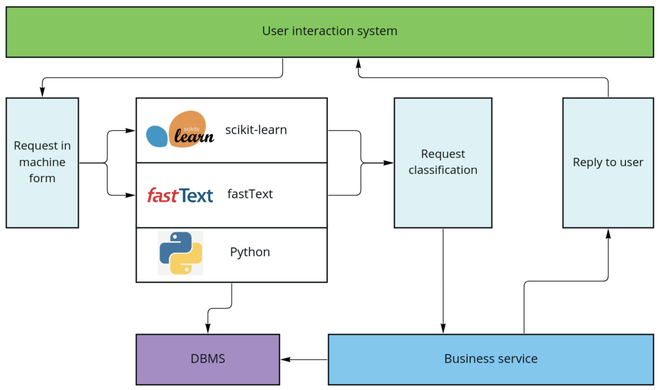

Fig. 2. Scheme of interaction of algorithmic support.

RESULTS

The results of the work described creating algorithmic support for the functioning of a personal virtual assistant. We obtain the results of assessing the quality of the classification of queries on The 20 Newsgroups data set. When running the two machine learning models, the weighted average value of the F1 metric was 82,5%. Table I show the results of classification in The 20 Newsgroups data set. It also obtained the running time of machine learning models. The running time of the logistic regression models of 189 ms and the fastText model of 214 ms. These results show the applicability of the selected models in an industrial environment.

TABLE I

RESULTS OF CLASSIFICATION IN THE 20 NEWSGROUPS DATA SET

Class	Precision	Recall	F1-score	Support
0	0,752	0,734	0,743	319
1	0,680	0,781	0,727	389
2	0,740	0,731	0,736	394
3	0,716	0,735	0,725	392
4	0,828	0,816	0,822	385
5	0,845	0,729	0,783	395
6	0,764	0,897	0,825	390
7	0,905	0,891	0,898	396
8	0,940	0,945	0,942	398
9	0,881	0,932	0,906	397
10	0,947	0,945	0,946	399
11	0,931	0,889	0,910	396
12	0,764	0,784	0,774	393
13	0,895	0,838	0,866	396
14	0,898	0,919	0,908	394
15	0,788	0,925	0,851	398
16	0,711	0,901	0,795	364
17	0,964	0,854	0,906	376
18	0,780	0,594	0,674	310
19	0,814	0,470	0,596	251
Accuracy	0,825	0,825	0,825	–
Macro avg	0,827	0,815	0,817	7532
Weighted avg	0,829	0,825	0,823	7532

CONCLUSIONS

As a result of the tasks of developing algorithmic support for a personal virtual assistant in a system for automated processing of client requests, a description of the process of collecting and storing data, a description of the principles of processing initial data, as well as principles for preparing initial data for algorithmic support are given. In accordance with the above principles, the preparation of initial data for algorithmic support was carried out. In the process of developing algorithmic support and a prototype of a personal virtual assistant, a set of technologies was formed and their choice for use in development was justified, as well as the interaction of components of the algorithmic support of a personal virtual assistant was formalized.

MISIS National University of Science and Technology

[1]NUST MISIS,
4, Leninskiy Prospekt,
Moscow, Russia (https://misis.ru/)