Intelligent system of formation of text descriptions of investment activities.

The aim of the project was to develop a prototype system for investment activities with the ability to automate the writing of comments.

The target audience of this system is investment agents and professionals in the field of finance.

The system has a specific subject area, for its creation the most popular analogues were studied. The user interface of such systems is quite complex, their main features were highlighted: the addition of new tools, management, taking into account the types of holdings, portfolio optimization, reporting.

The subject area of finance is also very broad, but the basis of our system consisted of concepts:
Portfolio - a set of investments held by an individual or organization.
Holding - a type of business entity that owns shares of other companies.
Strategy - algorithm of actions on investment and withdrawal of funds.
Commentary - documents containing analysis and recommendations regarding investments.

Much attention was paid to the function of report generation, which will allow investment agents to save time on the routine work of writing comments and reports on the performance of assets.

The database model was built according to the basic concepts of the subject area, as well as taking into account the roles of the system users (administrator, content manager, fund manager, portfolio manager). The entities portfolio, holding, user, comment, strategy, the appropriate attributes were added.

The architecture is a client-server. Accordingly, client modules for each of the user categories and the server part, which implements the business logic of the application.

Part of the comment generation is implemented as an API, in Python language. The main tasks of these modules were paraphrasing and paraphrasing the text. Paraphrasing - Replacing words in text with synonyms. (T0 model) In our system, it is used to be able to rewrite a comment of a child into a comment that is synonymous, not identical to the given comment.
Abstract Retelling - Phrases and sentences that emphasize key points of the original text are put together to create a clear and concise retelling. (GPT-3 model) Used to generate a summary of each portfolio holding's performance (so that the client can quickly navigate the performance without having to read through a huge reporting document)

Both language models have a transformer architecture (use a self-awareness mechanism, weighing the importance of each part of the input data). In such models, the text is tokenized

(broken down into smaller parts) and each token is represented as an embedding (a vector value describing its meaning as a sequence of numbers). This allows you to quickly detect synonymous embedding and determine the "weight" of the word in the provided context. The model of transformer neural network gives an opportunity to memorize the context of a large amount of data provided on the input, which allows making a high-quality retelling without loss of meaning.

As a result, a system was developed with functions: control of holdings and portfolios, data management, generation of portfolio reports and comments for holdings, saving and loading data.

Among the highlighted ways of development: Fine-tuning of models on examples written by content managers; Personalization of texts generated by models; Automation of language model retraining on corrected texts.