

Generative Artificial Intelligence

Project Report

Individual Educational Plan

Rafail Venediktov, Damir Nurtudinov, Andrey Starodumov

I. Introduction

A proposal was received to develop AI model to build an individual educational plan for users of the company INNOPROG at the request of this company [last page of this article].

A. Few Words about Customer

INNOPROG is an online educational service in Telegram for studying IT disciplines, where an individual learning plan is built for each student. To increase efficiency and involvement in the learning process, we use an interactive training system, which includes capacious informative videos and a variable selection of tasks, which is adjusted based on the level of knowledge of each student using artificial intelligence. In addition, we offer a flexible office hours system, which consists of individual lessons with a teacher at a time convenient for the student, as well as 24-hour support from our mentors via chat. By purchasing a monthly subscription, our student gets access to all educational modules and 240 minutes of individual lessons with a teacher.

B. Work Plan

It was decided to implement two models that could help the company automate and improve the interaction with the users within their educational plans.

1) Task Complexity Prediction: The objective here is to create a model that can accurately determine the appropriate level of task difficulty for each user. This decision will be based on the user's history of problem-solving success rates and their preferred level of challenge. By analyzing past performance data and stated preferences, the model can assign tasks that are neither too easy nor too difficult, thereby optimizing the learning curve and keeping users engaged and challenged.

2) Chat Assistant: The purpose of the Chat Assistant is to provide real-time support to users. This assistant should be capable of handling inquiries related to both the educational tasks and organizational aspects of the user's plans. It needs to understand and process various question types and deliver accurate and helpful responses. This involves training the model on a wide range of topics and ensuring it has access to up-to-date and relevant information.

II. Related Work

A. Chat Assistant

Since our goal is automation the answer the users' questions, we need to build a language model that can understand which problem the user asks about and give an appropriate answer. And this behavior is inherent in the Russian T5 model. The Russian T5 is the model that can answer the questions. It is already a solution for the chat assistance. [?]

III. Methodology

Both models mentioned in Section I are considered here as separate tasks.

A. Task Complexity Prediction

The first task facing us is the ability to predict the possible future tasks for the users.

1) Dataset: The problem here is that the initial task complexity that was given by an instructor does not consider the level of the students (users). Due to the different levels of students, tasks may appear both more difficult and easier than they actually are. That's why it was decided to consider the users' subjective estimation of the task complexity that was solved by them.

[Tab. I] shows which feature was taken to train the model. The dataset includes the time spent on solving the problem. And this time is also estimated by the users themselves.

Feature 1	User's estimation of a task complexity
Feature 2	Time spent on solving the problem
Feature 3	Instructor's task complexity

Таблица 1
Dataset for task complexity

2) Model: The model is the Neural Network with embeddings and LSTM layer. Here you can find the model structure in initialization function:

```
def __init__(self,
              input_dim,
              embedding_dim,
              hidden_dim,
              output_dim,
              n_layers,
              bidirectional,
              dropout):

    super(Model, self).__init__()

    self.word_embeddings = nn.Embedding(input_dim, embedding_dim, padding_idx=1)
    self.lstm = nn.LSTM(embedding_dim, hidden_dim, bidirectional = bidirectional)
    self.hidden2tag = nn.Linear(
        hidden_dim * 2 if bidirectional else hidden_dim,
        output_dim
    )
```

B. Chat Assistant

This model is a more complex task, since here the Large Language Model should be trained to answer the user's question in the chat.

1) Dataset: The task involves developing a chat assistant using a language model trained on a dataset of conversations. The dataset comprises 300 question-and-answer pairs with users, which is relatively small for effective training.

2) Model: To overcome the data limitation, the decision was made to utilize the Russian T5 model. This model is leveraged due to its architecture, which is suitable for understanding and generating human-like text. Training the T5 model on the available conversations aims to equip the chat assistant with the ability to respond accurately to user inquiries.

The choice of the Russian T5 model may also be influenced by its pre-training on a diverse text corpus, making it potentially more adept at handling the nuances of language with limited training data. However, the success of the model will largely depend on the quality and relevance of the dataset it is trained on.

3) Conclusion: Furthermore, it's important to consider the possibility of data augmentation or the use of transfer learning techniques, where a model pre-trained on a large corpus is fine-tuned on a smaller, task-specific dataset. This approach can help in improving the performance of the model despite the dataset's size constraints.

The ultimate goal is to create a chat assistant that can seamlessly engage with users, providing them with accurate and contextually relevant answers, thereby enhancing the user experience. The effectiveness of the T5 model in this application will need to be evaluated through testing and user feedback.

IV. GitHub link

The code part of the work is placed in the following GitHub repository (link).

V. Experiments and Evaluation

A. Task Complexity Prediction

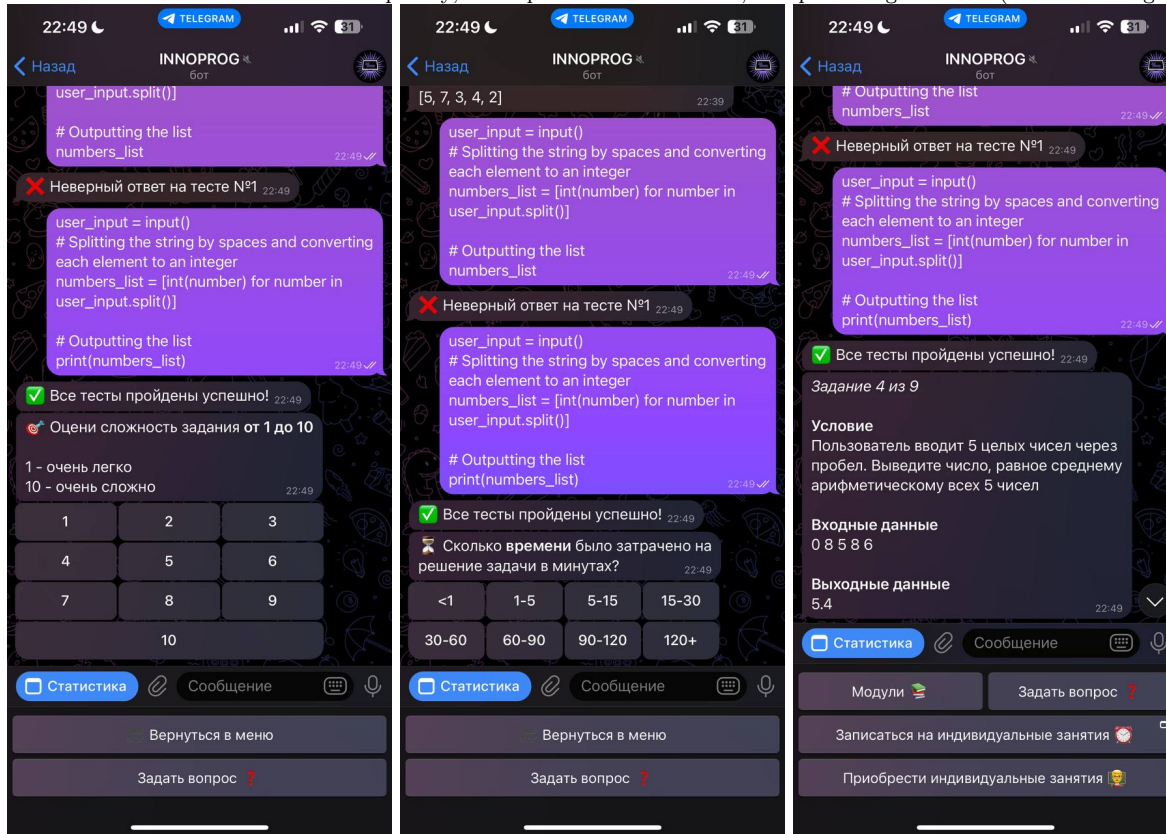
The model that predicts the possible tasks that are suitable to users shows good results. That's why it was decided to integrate the model into the bot for testing it in real time. This model will help the users to get tasks that are not so hard for them, but still develop their skills and increase their knowledge.

B. Chat Assistant

To address the issue of limited computing resources and extended training times associated with a 60-epoch schedule on a small dataset, an optimized approach has been applied. The resultant output, though derived from this constrained environment, successfully generates the anticipated results. This can be observed in the provided figure which visually represents the effectiveness of the solution under these restrictive conditions.

[Fig. 1] shows how the process of the collecting the information for the predicting the next tasks is organized.

Рис. 1. Users' estimation of task complexity, time spent to solve the task, and producing new task (from left to right)



VI. Analysis and Observations

A. Chat Assistant

The model built from scratch and trained on small data give the results that can only be considered as bad ones. That's why this model will be implemented in future with the help of already existing model from OpenAI, and the work will held.

For example, on the question "Как установить PyCharm?" we got the following answer:

"Charm — это проект, который помогает справиться с задачей. Если файлов таких нет, то нужно создать в PyCharm самостоятельно - любую файловую программу. Например, pyCharm понимает, что такое PyCharm и ему нужно использовать, чтобы создать файл на устройстве. Чтобы понять, где искать файл, достаточно нажать правой кнопкой мыши на название открытого файла или сам файл в боте И далее запустить в боте название открытого файла или сам файл, чтобы открыть файл в папке. Чтобы открыть папку, Вам потребуется создать файл в папке. Там же вы и найдёте сам файл pyCharm, где Вы и найдёте открытый файл".

VII. Conclusion

We've successfully developed a model that assigns tasks to users, facilitating their learning process. This model is instrumental in guiding users through a series of educational activities designed to ensure they achieve success in their learning objectives. The accomplishment of this phase signifies that the project's goals regarding user engagement and education have been met.

However, the project's scope extends beyond this initial success. The next phase involves the development and implementation of a separate model that will function within INNOPROG as a chat assistant. This new model will require careful integration and testing to ensure that it enhances the user experience by providing interactive, responsive

assistance. The work on this model is ongoing, and it will involve refining algorithms to interpret and respond to user queries effectively, as well as ensuring the chat assistant can guide users through the educational content seamlessly.

The completion of this second model is crucial for the overall project, as it aims to provide a comprehensive and interactive learning platform. The integration of the chat assistant will not only augment the educational framework but also offer a more personalized and adaptive learning experience. This two-pronged approach ensures that the project not only delivers content but also engages users interactively, thus fulfilling the dual objectives of the project.

INNOPROG, the client for whom the educational model and chat assistant have been developed, has expressed satisfaction with the delivered products. Their feedback indicates that the educational tasks model has been seamlessly integrated into their existing systems, enhancing the learning experience for users. The success of this implementation is a testament to the project's alignment with INNOPROG's vision for an innovative educational platform. The positive reception of the educational model has set a strong foundation for the continued development and integration of the chat assistant component. INNOPROG's endorsement of the work done so far provides additional motivation and validation for the project team as they move forward with the next phase, ensuring that the final product will meet the high standards and expectations that have been established.



Общество с ограниченной ответственностью
«ИННОПРОГ»

420500 РТ, Верхнеуслонский р-он, г.Иннополис,
ул. Университетская, д.5, пом.115, м.15/2

Тел. +79393954195

ИНН/КПП 1683011286/168301001

р/с 40702810401500170591

ООО "Банк Точка"

к/с 30101810745374525104

БИК 044525104

E-mail: r.venediktov@innopro.ru

<https://innopro.ru/>

№ 15-09-23 от «15» сентября 2023г.

Студенту 4 курса Университета Иннополис
Нуртдинову Дамиру Наилевичу

**Письмо на создание модели искусственного интеллекта для составления
индивидуального трека обучения.**

Просим на безвозмездной основе разработать модель искусственного интеллекта в рамках курса «Генеративный искусственный интеллект», которая даст возможность для учеников платформы INNOPROG реализовать индивидуальный трек обучения.

Генеральный директор ООО «ИННОПРОГ»



Венедиктов В.В.