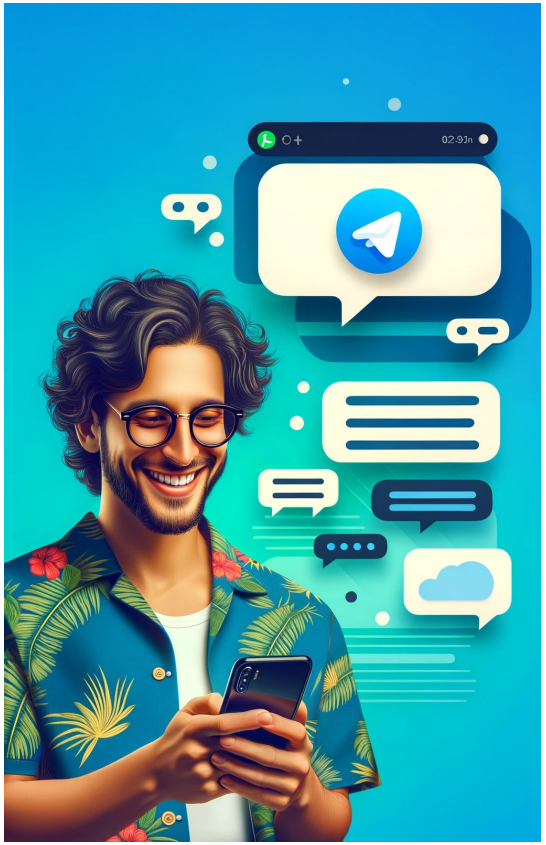


Summarcha: Chat Summarizer using Machine Learning

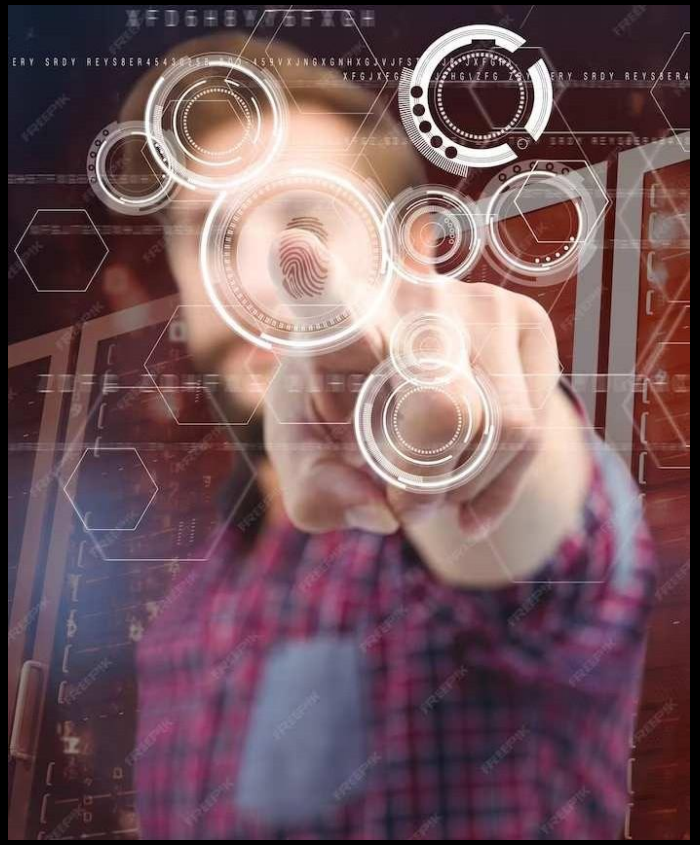
Introduction

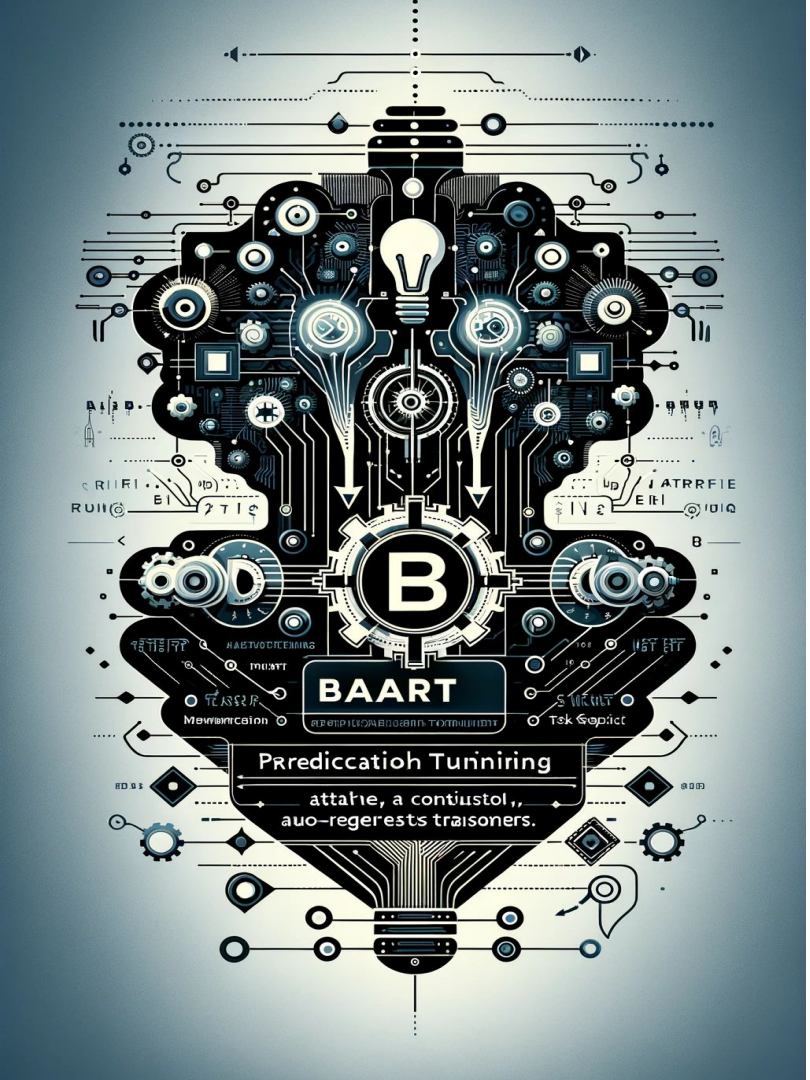
The exponential growth of digital communication has led to information overload, making it challenging to process and extract key insights from lengthy chat conversations.

/person_s_dr
/exam_s_dr
/etc_s_dr

Machine Learning Approach

Summarcha leverages advanced machine learning algorithms to analyze and extract essential information from chat data. By understanding context and relevance, the system tries to generate accurate and meaningful summaries, enhancing productivity and decision-making.





Solution 1 - base Bart model with prefix tuning

This solution explores the use of the base BART (Bidirectional and Auto-Regressive Transformers) model, with a focus on prefix tuning. Prefix tuning involves pre-appending a learnable continuous task-specific vector to the input, allowing the model to adapt to summarization tasks with minimal changes to its parameters. This approach is efficient in terms of computation and retains the general knowledge of the pre-trained BART model.

Inference API

Summarization

Examples

summarize:
Ivan Domrachev: I just think that since the participation in flipped classes is not mandatory, it's okay that they are overlapping
Vladimir Ivanov: All the students?
Ivan Domrachev: Yep, but the attendance rate is quite low anyway
Vladimir Ivanov: Anyway. I suggest you visit the lecture as it is mandatory.
Lev Kozlov: So, will we experience another rescheduling soon?
Vladimir Ivanov: No. Individual talks may swap, but it is only possible if a speaker decides to do so.
Vladimir Ivanov: Actually, I have to ask prof. Zlatanov…
Lev Kozlov: We can invite him to the flipped classes and call it the last lecture of ICV course
Lev Kozlov: Or invite him to CV part only, it will be more appropriate I think

Compute

Computation time on Intel Xeon 3rd Gen Scalable cpu: 3.856 s

Ivan Domrachev: I just think that since the participation in flipped classes is not mandatory, it's okay that they are overlapping. I suggest you visit the lecture as it is mandatory. Individual talks may swap, but it is only possible if a speaker decides to do so.

Inference API

Summarization

Examples

summarize:
Иван Домрачев: Я просто думаю, что, поскольку участие в параллельных занятиях не обязательно, это нормально, что они перекрываются
Владимир Иванов: Все ученики?
Иван Домрачев: Да, но посещаемость в любом случае довольно низкая
Владимир Иванов: В любом случае. Я предлагаю вам посетить лекцию, поскольку она обязательна.
Лев Козлов: Итак, скоро у нас будет еще одно перенесение?
Владимир Иванов: Но. Отдельные выступления могут меняться местами, но это возможно только в том случае, если спикер решит это сделать.
Владимир Иванов: На самом деле, я должен спросить проф. Златанов…
Лев Козлов: Мы можем пригласить его на перевернутые занятия и назвать это последней лекцией курса ICV
Лев Козлов: Или пригласить его только на часть резюме, я думаю, это будет более уместно

Compute

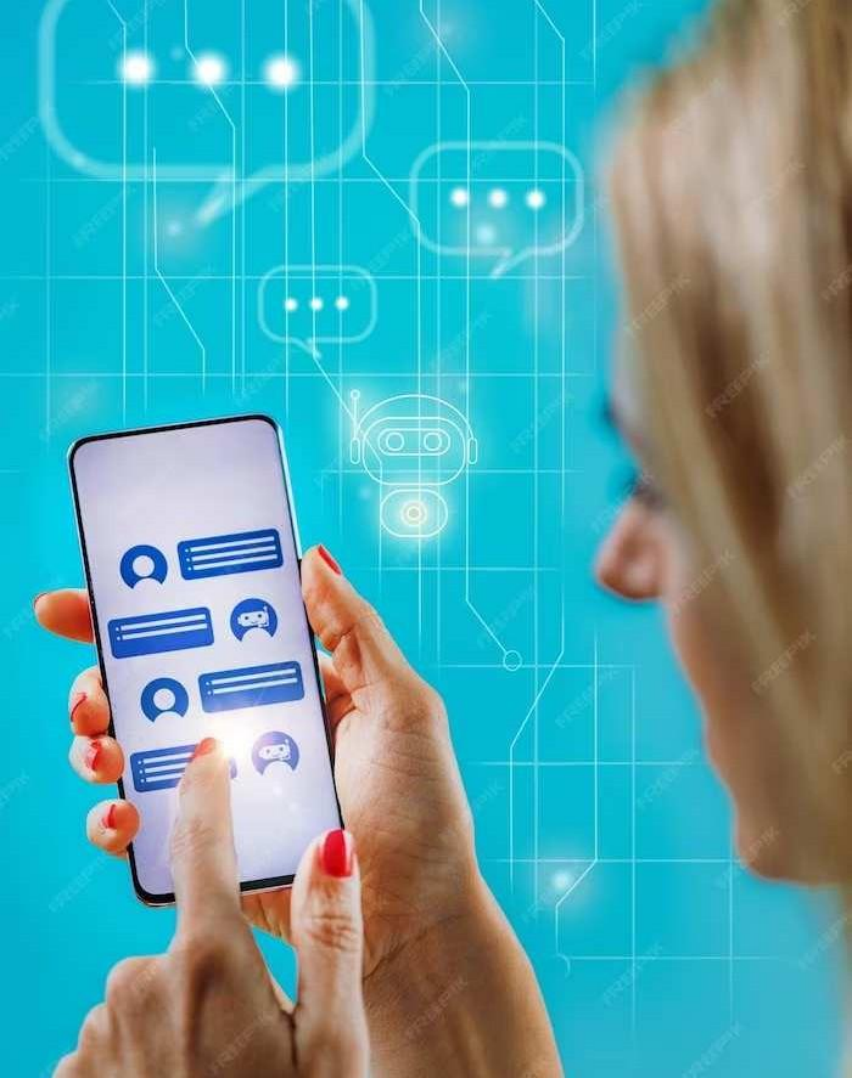
Computation time on Intel Xeon 3rd Gen Scalable cpu: 11.944 s

На перекцией ‘енять’ это они так, рещит ще обязательна. summarize: ‘‘’. “””. ”’ ”.” ””, “Н’, ‘ ,””. ”;. ”;. ,

Solution 2 - Bart-model with Translation

In this approach, the BART model is further enhanced with a summarization-specific fine-tuning and integrated with the Telegram Translation API for real-time multi-language support. This allows the model to not only summarize chats but also translate and summarize conversations in various languages, making it highly effective in multilingual chat environments.


Vladimir Boikov
С сортировками что-то (возможно только я) упустили важн...
Столько элементов часто обычно не сортят, так что nlogn хватит, можно подождать
22:43
Forwarded from Vladimir Boikov
Нашел кстати на habr код на java и статью. Нужно оставить на всякий случай.
22:43
Forwarded from Pavel Stupak
спасибо!
22:43
Forwarded from Mj
А в 5м тесте что ?) У меня RE ложится, видать какие-то крайние не вижу..
22:43
Forwarded from Mikhail Efimov
А кто-нибудь решил на жабе C задачу из последнего контеста? дейкстра с хипой дает TL, решил через SPFA
22:43
Forwarded from Mikhail Efimov
Mj
А в 5м тесте что ?) У меня RE ложится, видать какие-то край...
В 9 переполнение если что, для экономии времени работяг
22:43
Forwarded from Vladimir Boikov
Mikhail Efimov
А кто-нибудь решил на жабе C задачу из последнего контес...
А на чем решаешь? Там Питону дали вроде много сек.
22:43
Forwarded from Mikhail Efimov
Vladimir Boikov
А на чем решаешь? Там Питону дали вроде много сек.
Последний контест про графы, там 1 сек у всех, решаю на Java
22:43
GF
/summarize
22:44
Mikhail Efimov, Vladimir Boikov, Pavel Stupak, Igor Glushatov, Alexander Glazunov, Lesika, Иван K and Nyamerka) are taking part in a contest about graphs. Mikhail Efimov decides on Java. Vladimir Boikov found a habr code for java and an article. Alexander Glazunov forgot to change language. 5 out of 6 tests failed.
edited 22:44



Solution 3 - ru-T5 model fine-tuned on the dataset

The ru-T5 model, a Russian variant of the T5 model, is fine-tuned on a curated dataset consisting of Russian language chat logs. This fine-tuning aims to adapt the model to understand and summarize conversations in Russian more accurately. The dataset includes a mix of formal and informal chat logs to ensure the model's effectiveness across different styles of communication.

Final solution and results

Our team's final solution was the development of a custom Telegram client featuring an innovative 'Summarize' button. This feature, powered by the ru-T5 model, allows users to effortlessly summarize chat conversations. It offers flexibility in summarization, enabling users to summarize either a specific number of messages or only the unread messages in a chat. The integration of this feature into the Telegram Web client has been well-received for its user-friendly interface and practicality, making it a valuable tool for both personal and professional communication within the Russian-speaking community.

Thanks!

You can test the summarcha by scanning the QR and loginning into our telegram web application