

Диалоговая Система

OTUS NLP

Что такое Чат Бот?

Создаем ChatBot для финансовой сферы

Чат Бот

инструмент для развития компаний. Он повышает скорость обслуживания клиентов, помогает оптимизировать расходы и численность персонала фирмы.

Быть на связи 24/7

График работ может не совпадать с привычками целевой аудитории. Покупатели заходят на сайт или в мессенджер в любое время суток. Бот ответит на вопрос в нерабочие часы или сохранит заявку на обратный звонок в базе данных.

Повышать клиентоориентированность

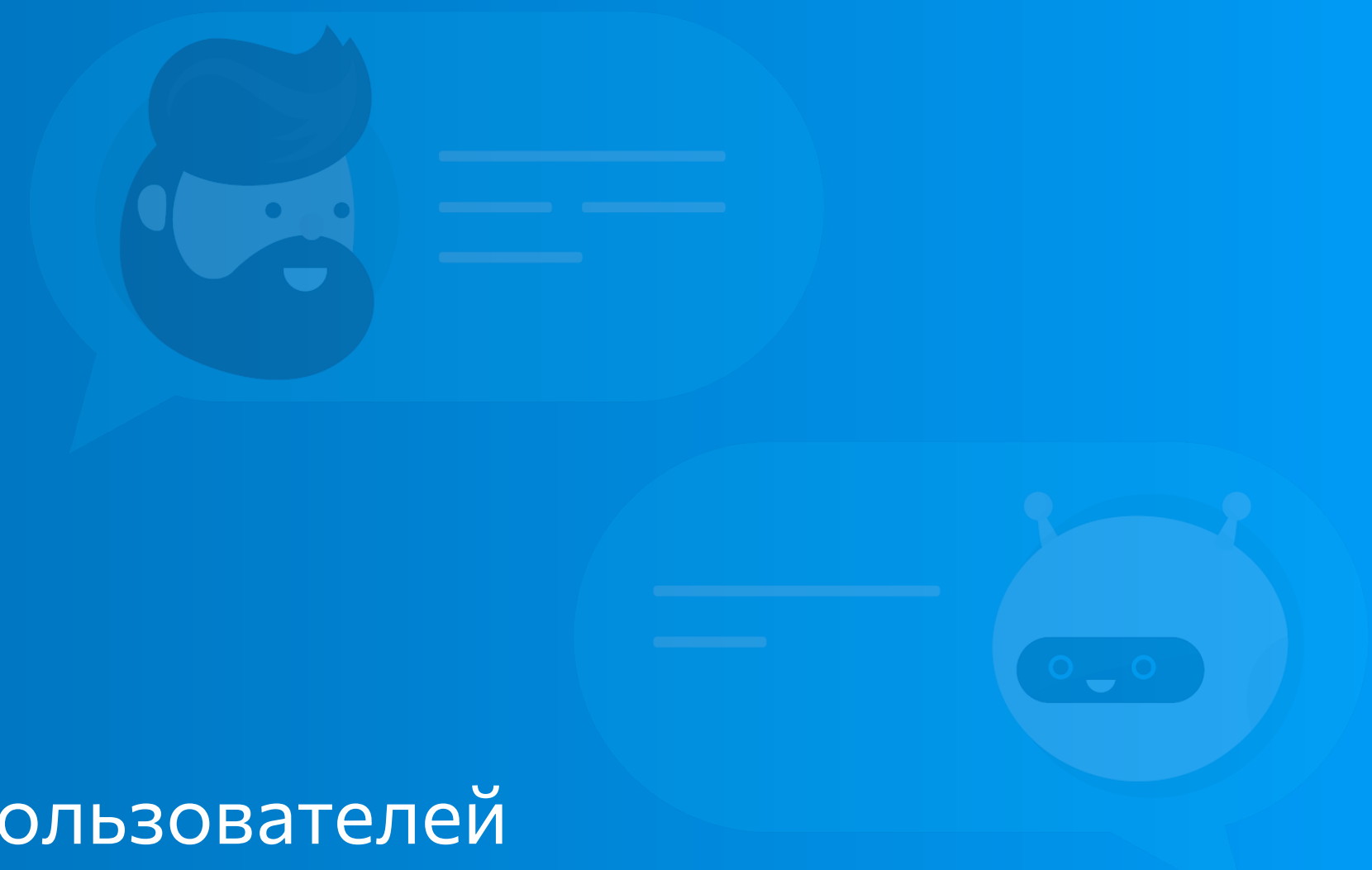
С чат-ботом вы сможете приблизиться к потребителям:

- моментально ответите на вопросы,
- поможете разобраться в работе сайта или тонкостях оформления заказа,
- предоставите актуальные сведения о продукте или услуге.

@sber business bot (<https://sberbb.ru/>)

Цель Проекта

Создаем ChatBot для финансовой сферы



Цель Минимум

- ▶ Создания чат-бота который сможет отвечать на вопросы пользователей
- ▶ Изучения разных подходов и создать единый подход (любой)
- ▶ Добавления разные фич для диалоговой системы которые улучшат качество модель

Цель Максимум

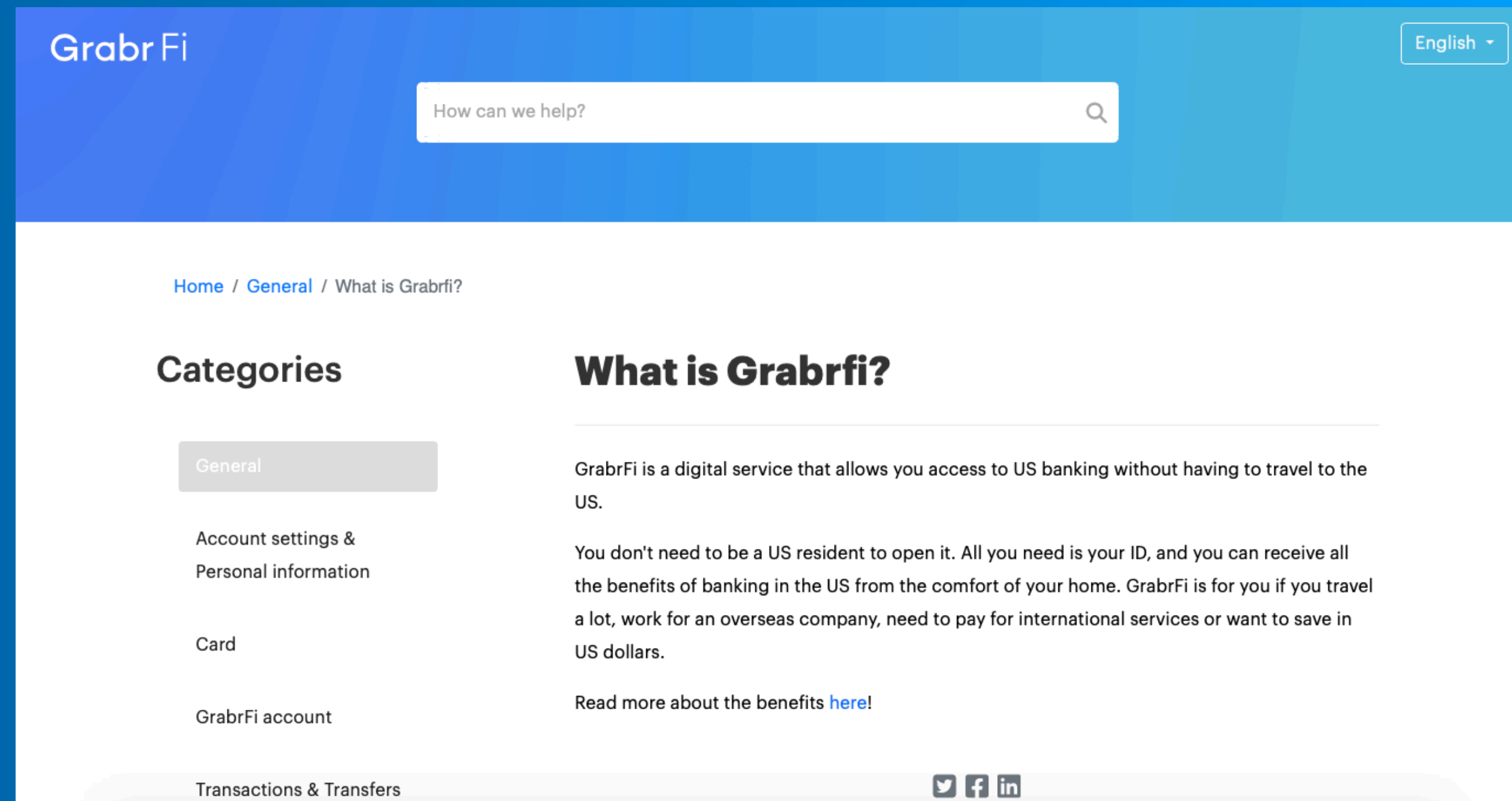
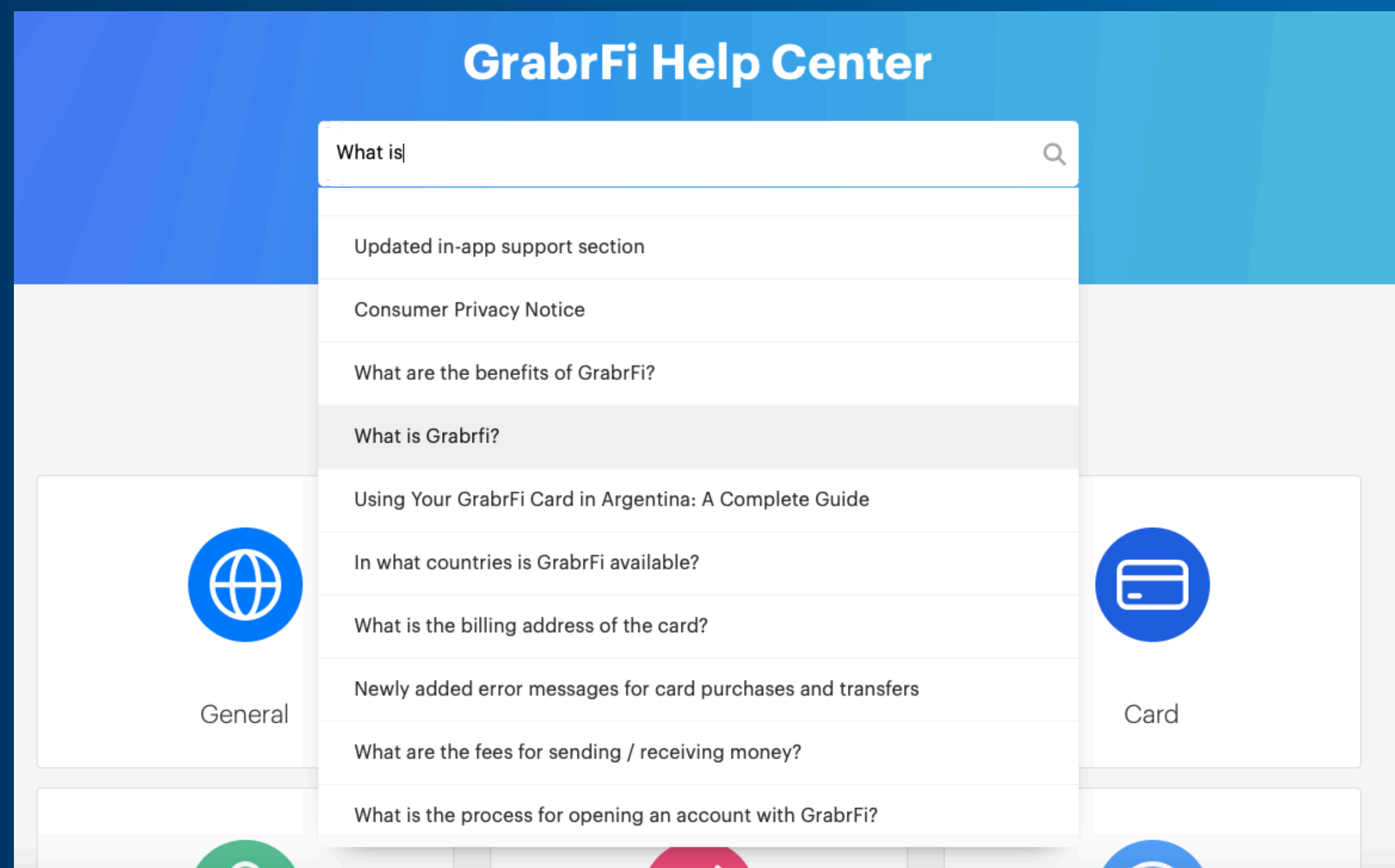
- ▶ Создания рабочую генеративную модель которая отвечает на вопросы используя информацию из разных частей текста на которой ее обучали

Что есть на данный момент

Создаем ChatBot для финансовой сферы

Раздел (HELP) - Гид по разным вопросам

Ответы на вопросы из базы знания на основе какой-то критерии близости



Подходы для ChatBot

Создаем ChatBot для финансовой сферы



На основе правил

Чат боты которые используют Re и тд.

Поисковые

Парсинговые, БД + TF-IDF, QA Трансформеры

ML подход

Векторы + Классический ML классификация

DL Генеративные

Трансформеры Decoder

Подходы для ChatBot

Создаем ChatBot для финансовой сферы

TFIDF Поисковые

▶ Список вопросов и ответов

```
1 [('What is the main benefit of using GrabrFi?',
2  'The main benefit of using GrabrFi is that you can
3  access US banking without having to travel to the US. '),
4  ('Who can open a GrabrFi account?',
5  'Anyone can open a GrabrFi account as long as they have
6  their ID. '),
7  ('How can GrabrFi help people who work for international
8  companies?',
9  'GrabrFi can help people who work for international
10 companies by allowing them to receive payments in US
11 dollars, without losing any money in conversion fees.')]
12
```

QA Поисковые

▶ Текст содержащий ответ + вопрос

```
1 answer_text = "GrabrFi is a digital service that allows
2 you access to US banking without having to travel to the
3 US. "
4 mes = 'What are the benefits of GrabrFi?'
5 extractive_answer = extractive_qa_bert(mes, answer_text)
6 extractive_answer
7
8 'allows you access to US banking without having to travel
9 to the US'
```

ML Классификатор

▶ Класс для вопросов

```
1 {
2   "tag": "account_suspended",
3   "patterns": [
4     "My account has been suspended",
5     "I would like to know why my account has been suspended",
6     "I have had my account suspended",
7     "I don't understand why my account has been suspended"
8   ],
9   "responses": [
10    "We may place a security hold on your account if we detect suspicious
11 activity such as suspicious logins, transfers, and charges. We do our best to
12 communicate the issue to you via email within the first 24 hours. If you did not
13 receive any communication after 24 hours after suspension, please contact our
14 support team."
15 ]
16 }
```

DL Генеративные

▶ диалог

```
1 user, "How can I make a top up or transfer money to my
2 GrabrFi account?"
3 bot, "You can make a top up or transfer funds to your
4 GrabrFi account by making a transfer from another bank
5 account."
6 user, "What are the available top up options?"
7 bot, "The available top up options are:
8 SWIFT international transfers (worldwide),
9 Wire transfers (within the US only)
10 ACH transfers (within the US only)."
11 user, "How do I add funds to my account?"
12 bot, "Please follow these steps: Log in to the application
13 and click on [Home] & click on [Add funds] and select
14 your preferred top up option"
```

Исходные Данные

Создаем ChatBot для финансовой сферы

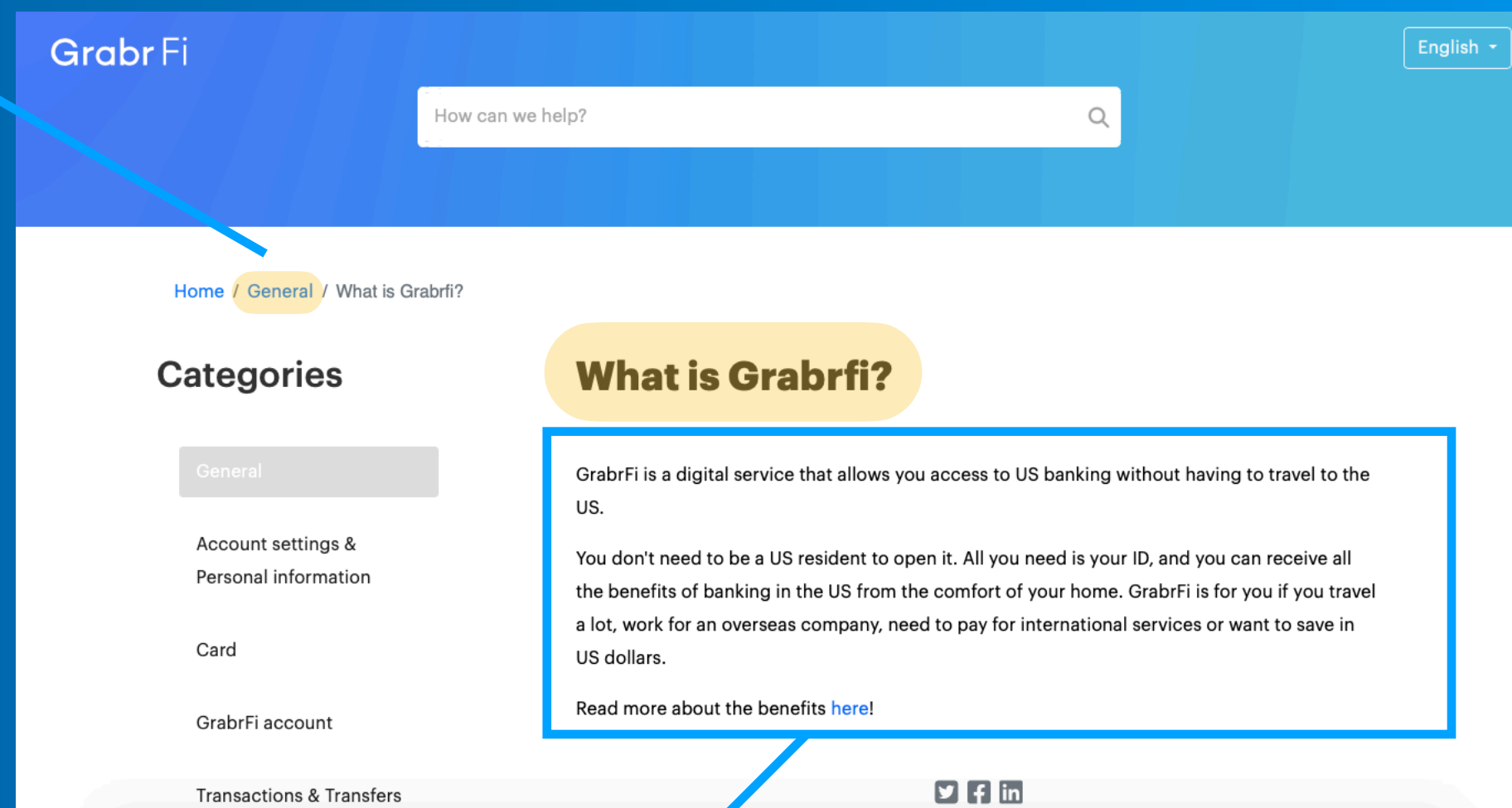
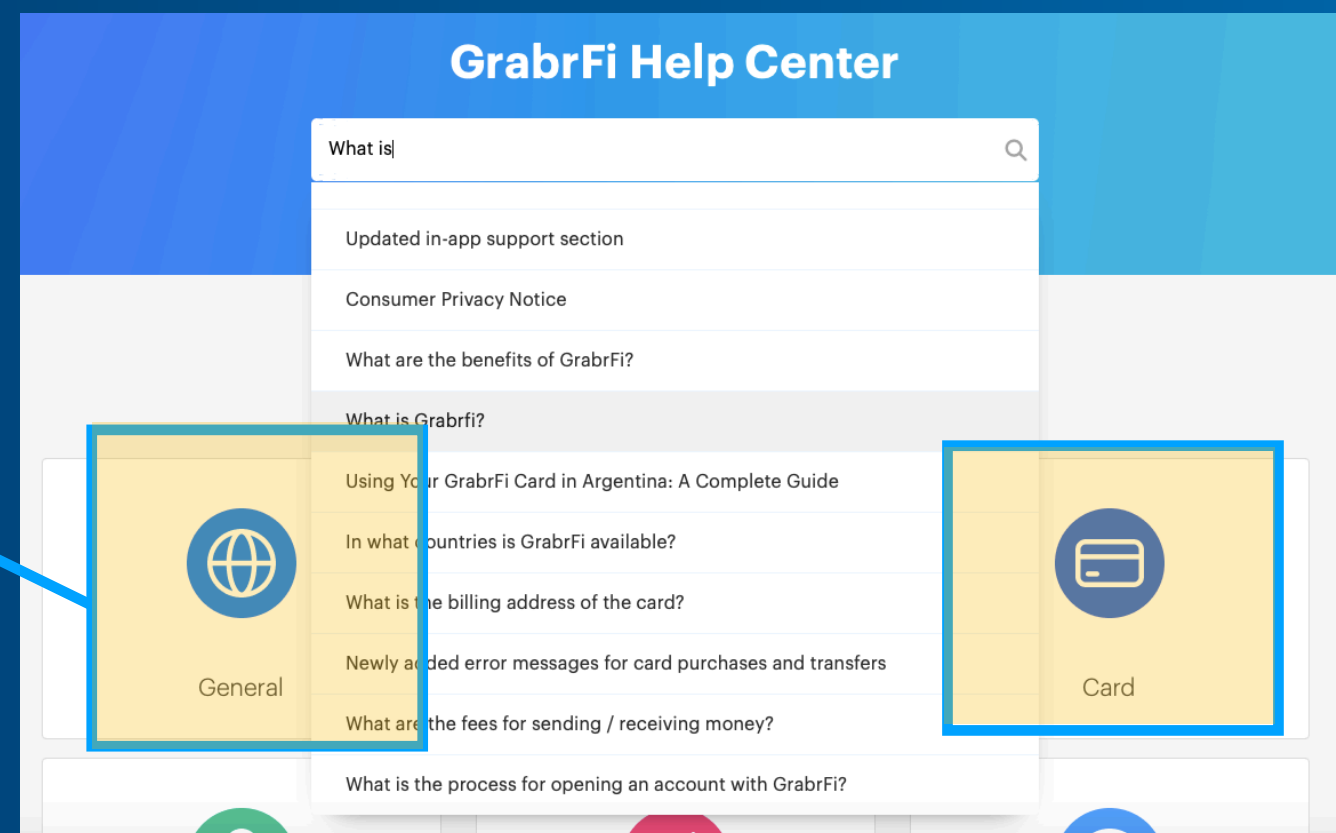
- 1 "1. What is the main benefit of using GrabrFi?",
- 2 "2. Who can open a GrabrFi account?",
- 3 "3. Why would someone want to use GrabrFi?",
- 4 "4. What are the benefits of GrabrFi?"
- 5 ...

Categories	Topics	Articles	context	questions	answers	
0	General	What is GrabrFi?	GrabrFi is a digital service that allows you a...	General\nWhat is GrabrFi?\n\nGrabrFi is a digi...	1. What is the main benefit of using GrabrFi?\\n...	1. The main benefit of using GrabrFi is that y...
1	General	What are the benefits of GrabrFi?	GrabrFi is a digital service that allows you t...	General\nWhat are the benefits of GrabrFi?\n\n...	1. What are the benefits of GrabrFi?\n2. How c...	1. The benefits of GrabrFi include having acce...
2	General	What is the difference between GrabrFi and a t...	The biggest difference between GrabrFi and a t...	General\nWhat is the difference between GrabrF...	1. What are the benefits of using GrabrFi over...	1. The benefits of using GrabrFi over a tradit...
3	General	In what countries is GrabrFi available?	We currently offer accounts to residents of th...	General\n\nIn what countries is GrabrFi availabl...	1. What countries is Grabr available in? \n2. ...	1. The countries that Grabr is available in ar...
4	General	Requirements to open a GrabrFi account	You don't need to travel to open your account!...	General\n\nRequirements to open a GrabrFi accoun...	1. What is the minimum age requirement to open...	1. The minimum age requirement to open a Grabr...

questions

Topics

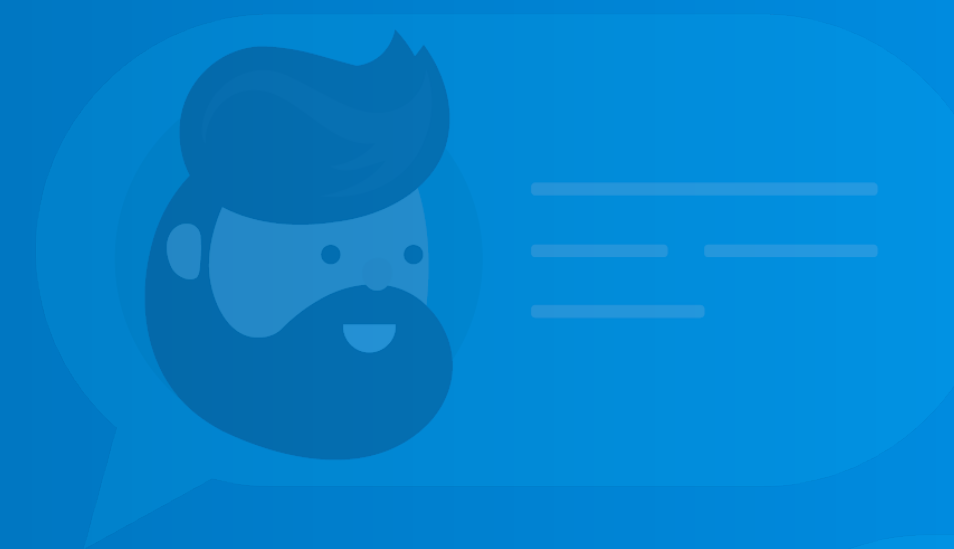
Topics



Articles

Тестируем БД + TF-IDF

Создаем ChatBot для финансовой сферы



Top N

Берем наиболее подходящие с точки зрения критерии близости

```
1 Questions list:
2     'What is the pull-to-refresh feature in the iOS application?',
3     'How often are account statements sent?',
4     'How can you access account statements?',
5
6 Request:
7 please [tell me how often do you send me account statements]
8
9 top 3 results:
10 [0.7191354216981398, 0.6026060825237777, 0.39624013529576035]
11
12 1 : Account statements are sent monthly, usually by the 16th day of the following month.
13 2 : Account statements can be accessed in the 'Home' section of the application, and you should be
    logged in to see it. Click the 'Statement' button to see the list of all available statements.
14 3 : You can update your email address up to 2 times within a 24-hour period.\n
15
16 [tell me how often you send account statements]
17
18 top 3 results:
19 [0.7535790128121267, 0.6314684036986102, 0.4152183872234018]
20
21 1 : Account statements are sent monthly, usually by the 16th day of the following month.
22 2 : Account statements can be accessed in the 'Home' section of the application, and you should be
    logged in to see it. Click the 'Statement' button to see the list of all available statements.
23 3 : You can update your email address up to 2 times within a 24-hour period.\n'
```

Выбор Порога

Выбираем порог для выбора ответа



```
1 if(top_val[0] > 0.80):
2     return similarities.argmax()
3 else:
4     return None
5
6 Request:
7 tell me how often you send account statements
8 [0.7535790128121267]
9
10 'please add more detail to your question'
```


Тестируем QA Трансформеры

Создаем ChatBot для финансовой сферы



Задача Question - Answering

► Текст содержащий ответ + вопрос пользователя

```
1 question = "What are the requirements to open an account?"
2
3 answer_text = "The difference between GrabrFi and a traditional bank is that GrabrFi provides almost-instant access to US banking, which is online, whereas a traditional bank requires you to visit the US or be a resident. Other differences include: You will be able to receive & send US dollars worldwide with no to very low fees, there will be no monthly maintenance fees, there is no minimum bank balance requirement, no monthly maintenance fees & 24/7 customer support will be available to you. In what countries is GrabrFi currently available? GrabrFi is currently available for people who are residents of Argentina, Brazil, Canada, Chile, Colombia, France, Germany, Italy, Mexico, Portugal, Spain, United Kingdom, United States and Vietnam. Requirements to open a GrabrFi account: All you have to do is download the app and register. The requirements to open a GrabrFi account are: email address, a phone number where you can be reached and receive SMS, valid mailing address and you must be over 18 years old."
4
5 qa_bert(question,answer_text)
6
7 # Answer: "email address , a phone number where you can be reached and receive SMS , valid mail ##ing address and you must be over 18 years old"
```

Выбор Текста

Перебор всех текстов

Классификатор

Для определения категории (Category)
ie. Классификатор "Category"

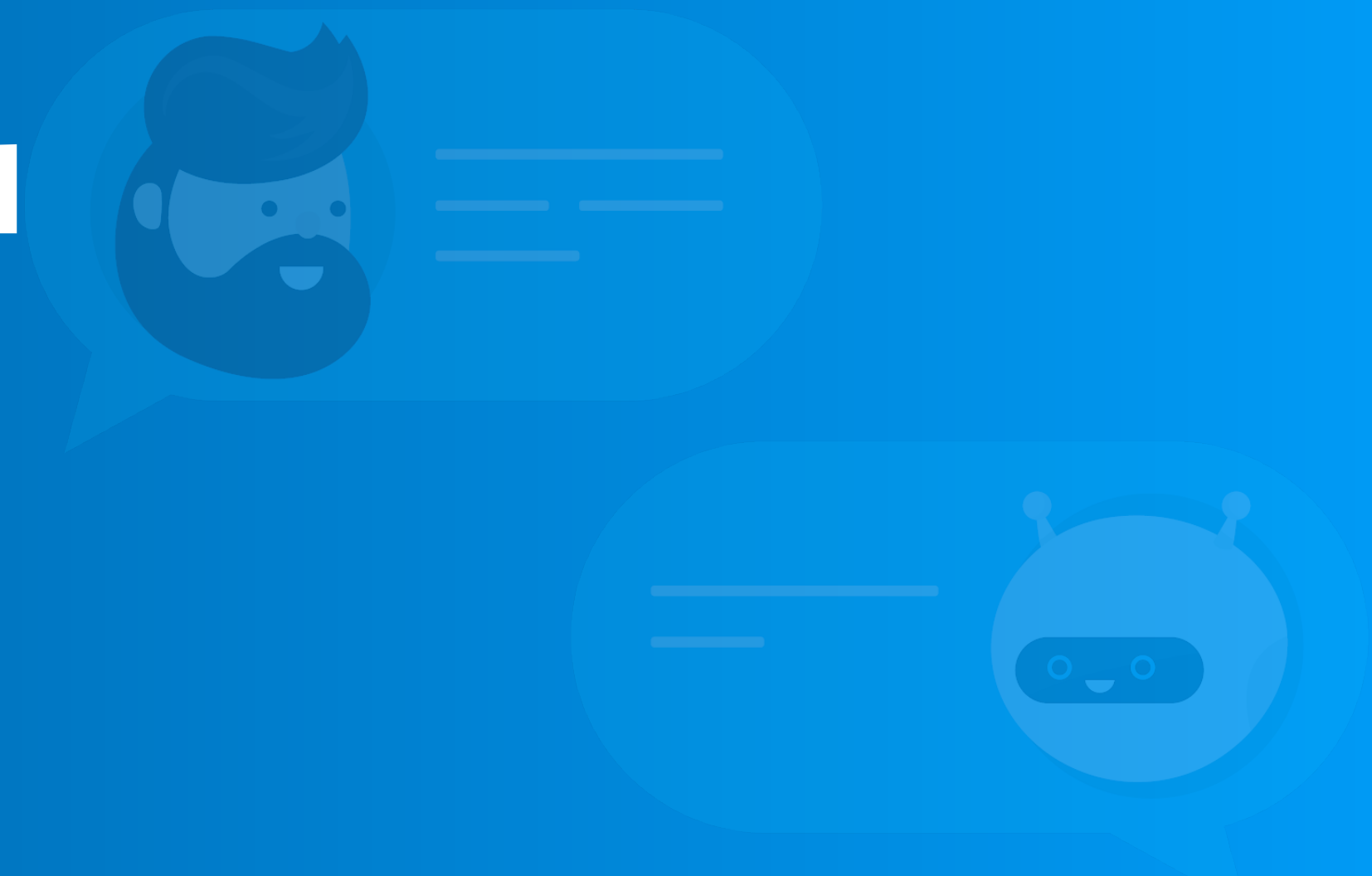
- > Активируется для задачи QA
- > Активируется для генеративной GPT модели

Подход для Question Answering

Fine-Tune BERT на данных SQUAD v1

Тестируем QA Трансформеры

Создаем ChatBot для финансовой сферы



(1) Классификатор

Для определения категории (**Category**)

```
1 # BoW Вектора (X)
2 array([[0., 0., 0., ..., 0., 0., 0.],
3        [0., 0., 0., ..., 0., 0., 0.],
4        [0., 0., 0., ..., 0., 0., 0.],
5        ...,
6        [0., 0., 0., ..., 0., 0., 0.],
7        [0., 0., 0., ..., 0., 1., 0.],
8        [0., 1., 0., ..., 0., 0., 0.]],
9        dtype=float32)
9 # (388, 321)
10
11 from sklearn.ensemble import
12   RandomForestClassifier
13 from sklearn.metrics import accuracy_score
14   as accuracy
15
16 qa_classifier = RandomForestClassifier()
17 qa_classifier.fit(X_train,y_train)
18 y_pred = qa_classifier.predict(X_train)
19 tr_accuracy = accuracy(y_pred,y_train)
20 # 0.9974226804123711
```

Inference

Проверим классификатор

```
1 def tester(sentence):
2     sentence = tokenize(sentence)
3     X = bow(sentence, all_words)
4     X = X.reshape(1, X.shape[0])
5     pred = qa_classifier.predict(X)
6     print(pred)
7
8 tester('How often can I change my phone number?')
9 tester('Can I use my digital card with Google Wallet?')
10 tester('Is my account compatible with Apple Wallet?')
11
12 # ['account']
13 # ['card']
14 # ['card']
```

Тестируем QA Трансформеры

Создаем ChatBot для финансовой сферы

Выбор Порога

Находим `argmax(max(logits))`

(2) Задача Question Answer (Процесс)

Получаем ответ на вопрос из текста

```
1 # user input
2 mes = 'What are the benefits of GrabrFi?'
3
4 # text source
5 answer_text = "GrabrFi is a digital service that allows you
6 access to US banking without having to travel to the US. "
7
8 ['[CLS]', 'What', 'are', 'the', 'benefits', 'of', 'G',
9  '##rab', '##r', '##F', '##i', '?', '[SEP]', 'G', '##rab',
10 '##r', '##F', '##i', 'is', 'a', 'digital', 'service',
11 'that', 'allows', 'you', 'access', 'to', 'US', 'banking',
12 'without', 'having', 'to', 'travel', 'to', 'the', 'US', '.',
13 '[SEP]']
```

Токены текста

```
1 QuestionAnsweringModelOutput(loss=None,
2 start_logits=tensor([[ -4.1833, -4.9635, -5.6475, -4.8030,
3 -2.2221, -6.5630, -6.6779, -8.1316,
4 -8.4170, -8.1976, -8.1531, -3.4018, -3.5032,
5 2.8908, -3.9919, -4.5248,
6 -4.6423, -4.3546, -2.2339, -0.8731, -0.0142,
7 -1.1928, 1.4941, 6.6916,
8 3.1407, 5.3463, -1.2457, 2.2434, 1.7383,
9 4.7943, 1.4277, -1.4934,
10 -0.1052, -3.5509, -3.2912, -1.5679, -2.3753,
11 -4.2994]]),
12 grad_fn=<CloneBackward0>),
13 end_logits=tensor([[ 0.2867, -5.3636, -6.2659, -6.1131,
14 -0.8645, -6.5708, -7.5951, -6.7791,
15 -6.6497, -6.8988, -5.7298, -4.4439, -4.4590,
16 -3.7229, -4.5408, -4.1239,
17 -4.7327, -1.3542, -4.8835, -4.8765, -1.5053,
18 0.4094, -4.6681, -0.5287,
19 0.0292, 0.3293, -1.2425, -0.1589, 5.8771,
20 0.8900, 0.0385, -0.3689,
21 4.8369, 1.4966, -0.0558, 8.8769, 7.0160,
22 -4.9427]]),
23 grad_fn=<CloneBackward0>), hidden_states=None,
24 attentions=None)
```

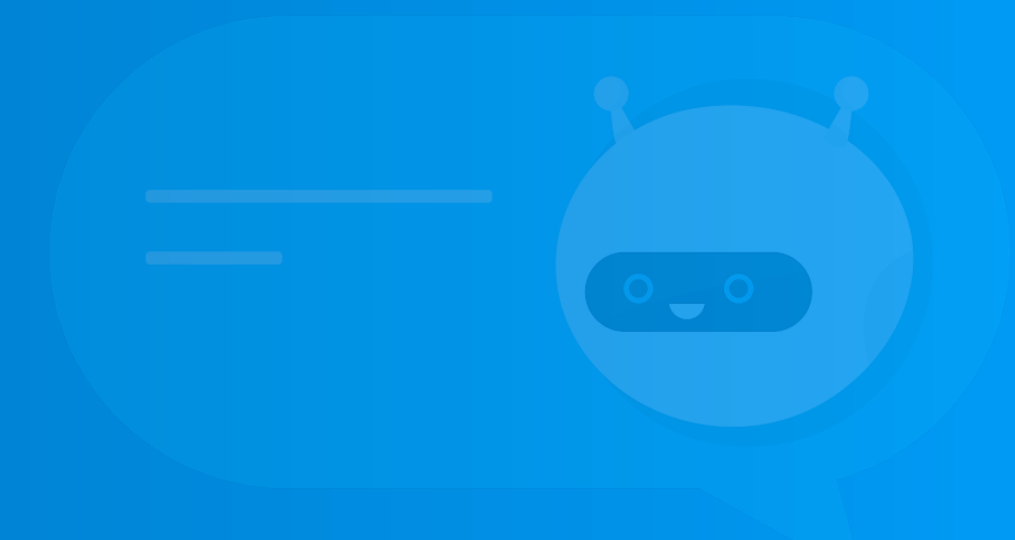
Logits **START** токенов

Logits **END** токенов

```
1 tensor(6.6916, grad_fn=<MaxBackward1>)
2 tensor(8.8769, grad_fn=<MaxBackward1>)
3
4 'allows you access to US banking without having to travel to
the US'
```


Тестируем QA Трансформеры

Создаем ChatBot для финансовой сферы



(2) Задача Question Answer (Хороший и плохой ответ)

Получаем ответ на вопрос из текста

Плохой пример

Хороший пример

Выбор Порога

Выбор порога по значению logits

```
1 answer_text = "GrabrFi is a digital service that allows you  
access to US banking without having to travel to the US. "  
2 mes = 'How do I change my billing address?'  
3  
4 extractive_answer = extractive_qa_bert(mes,answer_text)  
5 extractive_answer  
6  
7 tensor(2.2373, grad_fn=<MaxBackward1>)  
8 tensor(3.2368, grad_fn=<MaxBackward1>)  
9  
10 'G ##rab ##r ##F ##i'
```

```
1 answer_text = "To change your billing address on your  
GrabrFi account, go to More → Profile & Settings → Billing  
address / Mailing address and enter your updated  
2 information.',"  
3 mes = 'How do I change my billing address?'  
4  
5 extractive_answer = extractive_qa_bert(mes,answer_text)  
6 extractive_answer  
7  
8 tensor(7.4545, grad_fn=<MaxBackward1>)  
9 tensor(7.5104, grad_fn=<MaxBackward1>)  
10  
'go to More → Profile & Set ##ting ##s → Bill ##ing  
address / Mail ##ing address and enter your updated  
information'
```


Тестируем ML Подход

Создаем ChatBot для финансовой сферы

Варианты одного вопроса

Задача Классификации

<tag> метка для вопроса

<patterns> Варианты вопроса

<responses> Ответ на вопрос

```
1 {
2   "tag": "account_suspended",
3   "patterns": [
4     "My account has been suspended",
5     "I would like to know why my account has been suspended",
6     "I have had my account suspended",
7     "I don't understand why my account has been suspended"
8   ],
9   "responses": [
10    "We may place a security hold on your account if we detect suspicious
11    activity such as suspicious logins, transfers, and charges. We do our best to
12    communicate the issue to you via email within the first 24 hours. If you did not
13    receive any communication after 24 hours after suspension, please contact our
14    support team."
15  ]
16 }
```

```
1 # BoW Вектора (X)
2 array([[0., 0., 0., ..., 0., 0., 0.],
3        [0., 0., 0., ..., 0., 0., 0.],
4        [0., 0., 0., ..., 1., 0., 0.],
5        ...,
6        [0., 0., 0., ..., 0., 0., 0.],
7        [0., 0., 0., ..., 0., 0., 0.],
8        [0., 0., 0., ..., 0., 0., 0.]], dtype=float32)
9
10 # Hyper-parameters
11 num_epochs = 1000
12 batch_size = 8
13 learning_rate = 0.001
14 input_size = len(X_train[0])
15 hidden_size = 8
16 output_size = len(tags)
```

```
1 # Loss and optimizer
2 criterion = nn.CrossEntropyLoss()
3 optimizer = torch.optim.Adam(model.parameters(),
4                               lr=learning_rate)
```

```
1 class NeuralNet(nn.Module):
2     def __init__(self, input_size, hidden_size,
3                 num_classes):
4         super(NeuralNet, self).__init__()
5         self.l1 = nn.Linear(input_size, hidden_size)
6         self.l2 = nn.Linear(hidden_size, hidden_size)
7         self.l3 = nn.Linear(hidden_size, num_classes)
8         self.relu = nn.ReLU()
9
10    def forward(self, x):
11        out = self.l1(x)
12        out = self.relu(out)
13        out = self.l2(out)
14        out = self.relu(out)
15        out = self.l3(out)
16        # no activation and no softmax at the end
17        return out
```

```
1 Epoch [100/1000], Loss: 0.1055
2 Epoch [200/1000], Loss: 0.0017
3 Epoch [300/1000], Loss: 0.0002
4 Epoch [400/1000], Loss: 0.0003
5 Epoch [500/1000], Loss: 0.0000
6 Epoch [600/1000], Loss: 0.0000
7 Epoch [700/1000], Loss: 0.0000
8 Epoch [800/1000], Loss: 0.0000
9 Epoch [900/1000], Loss: 0.0000
10 Epoch [1000/1000], Loss: 0.0000
11 final loss: 0.0000
```

Тестируем ML Подход

Создаем ChatBot для финансовой сферы

Продолжение диалога

Получив ответ на вопрос, user хочет получить разъяснение

```
1 You: my account has been suspended
2 Sam: We may place a security hold on your account if we detect suspicious activity
such as suspicious logins, transfers, and charges. We do our best to communicate
the issue to you via email within the first 24 hours. If you did not receive any
communication after 24 hours after suspension, please contact our support team.
3
4 You: I'd like to know why my account has been suspended
5 Sam: We may place a security hold on your account if we detect suspicious activity
such as suspicious logins, transfers, and charges. We do our best to communicate
the issue to you via email within the first 24 hours. If you did not receive any
communication after 24 hours after suspension, please contact our support team.
6
7 You: I've had my account suspended
8 Sam: We may place a security hold on your account if we detect suspicious activity
such as suspicious logins, transfers, and charges. We do our best to communicate
the issue to you via email within the first 24 hours. If you did not receive any
communication after 24 hours after suspension, please contact our support team.
```

```
1 You: my account has been suspended
2 Sam: We may place a security hold on your account if we detect suspicious activity such as suspicio
logins, transfers, and charges. We do our best to communicate the issue to you via email within t
first 24 hours.If you did not receive any communication after 24 hours after suspension, please
our support team.
3 You: what should I do?
4 Sam: I do not understand...
```

Порог/пересечение данных

Не оптимальное пороговое значение +
пересечение данных

Выбор Порога

`torch.softmax(logits)`

```
1 # probability threshold
2 if prob.item() > 0.75:
3     for intent in intents['intents']:
4         if tag == intent["tag"]:
5             print(f"{bot_name}:
[random.choice(intent['responses'])}")
6     else:
7         print(f"{bot_name}: I do not understand...")
```

Не верный вызов

```
1 You: I've had my account suspended
2 Sam: We may place a security hold on your account if we detect suspicious activity such as suspicio
logins, transfers, and charges. We do our best to communicate the issue to you via email within th
first 24 hours. If you did not receive any communication after 24 hours after suspension, please c
our support team.
3 You: How do I contact the support team?
4 Sam: You can contact the support team by email or phone.
5 You: What is the phone number in Argentina?
6 Sam: Visit: https://app.grabrifi.com/. Use your email address to create a login and create a secure
password. After registration, you will be prompted to verify your email address & complete the iden
verification process. You will receive notification when it's approved, so you can open your account
7
```


Тестируем DL Подход

Создаем ChatBot для финансовой сферы

Генеративный GPT

Microsoft DialoGPT (medium)

345M параметров

147M reddit диалогов

Fine-tune



GrabrFi account



Transactions & Transfers



Pricing & limits

Разные варианты развития диалога

Разные варианты развития диалога

```
1 #
2 #
3 # 40. GrabrFi account | My account is suspended. Why and what sho
4 #
5 #
6
7 user,"I've had my account suspended and need some assistance"
8 bot,"I'm sorry to hear that. We may place a security hold on your ac
detect suspicious activity such as suspicious logins, transfers, and
you need some extra clarification?"
9 user,"Yes"
10 bot,"We do our best to communicate the issue to you via email within
hours. If you did not receive any communication after 24 hours after
please contact our support team."
11 user,"How can I get in touch with the support team?"
12 bot,"You can contact our support team by emailing support@grabrfi.co
13 user,"Thank you"
14 bot,"You're welcome "
15
16 user,"My account has been blocked and I need some assistance"
17 bot,"I'm sorry to hear that. We may place a security hold on your ac
detect suspicious activity such as suspicious logins, transfers, and
you need some extra clarification?"
18 user,"Yes please"
19 bot,"We do our best to communicate the issue to you via email within
hours. If you did not receive any communication after 24 hours after
please contact our support team."
20 user,"How can I get in touch with the support team?"
21 bot,"You can contact our support team by emailing support@grabrfi.co
22 user,"Thank you"
23 bot,"You're welcome"
```

response	context	context/0	context/1	context/2	context/3	context/4	context/5
0	Sure, I'm here to help. Do you have any questi...	I need some help	My name is DialoGPT	What is your name?	I'm a generative model based on Microsoft's Di...	Who are you?	Hello, how may I help you? I'll be your assist... Hello
1	I'll do my best to answer your questions. I'm ...	Can you please help me	Sure, I'm here to help. Do you have any questi...	I need some help	My name is DialoGPT	What is your name?	I'm a generative model based on Microsoft's Di... Who are you?
2	Sure, go ahead, I'm all yours.	I want to ask some questions	I'll do my best to answer your questions. I'm ...	Can you please help me	Sure, I'm here to help. Do you have any questi...	I need some help	My name is DialoGPT What is your name?
3	Well, that's an interesting question! I've	What can you help me with?	Sure, go ahead, I'm all yours.	I want to ask some questions	I'll do my best to answer your questions. I'm ...	Can you please help me	Sure, I'm here to help. Do you have any questi... I need some help

Контекст диалога

Тестируем DL Подход

Создаем ChatBot для финансовой сферы

Протестируем генеративную модель

User хочет узнать какие опции есть для переводов

```
1 # Dataset
2 user,"What bank transfers do you offer?"
3 bot,"Several bank transfer options are available to you, ACH and wire transfers"
4 user,"What is the difference between them?"
5 bot,"A ACH transfer is completed through a clearing house and can be used to process direct
  payments or direct deposits. Wire transfers on the other hand allow for the direct movement of
  money from one bank account to another. Wire transfers will cost you for domestic Incoming wire 5
  USD, and for domestic Outgoing wire 15 USD."
6
7 # Example 1
8
9 >> User:What bank transfers do you offer?
10 Bot: Several bank transfer options are available to you, ACH and wire transfers
11 >> User:Can you tell me the difference between them?
12 Bot: A ACH transfer is completed through a clearing house and can be used to process direct
  payments or direct deposits. Wire transfers on the other hand allow for the direct movement of
  money from one bank account to another. Wire transfer will cost you for domestic Incoming wire 5
  USD, and for domestic Outgoing wire 15 USD.
13
14 # Example 2
15
16 >> User:Which bank transfer options does your service offer?
17 Bot: Several bank transfer methods are available to you, ACH and wire transfers
18 >> User:Any difference between them?
19 Bot: A ACH transfer is completed through a clearing house and can be used to process direct
  payments or direct deposits. Wire transfers on the other hand allow for the direct movement of
  money from one bank account to another. Wire transfer will cost you for domestic Incoming wire 5
  USD, and for domestic Outgoing wire 15 USD.
20
```

```
1 # generated a response while limiting the total chat
  history to 1000 tokens,
2 chat_history_ids = model.generate(
3     bot_input_ids,
4     max_length=1000,
5     pad_token_id=tokenizer.eos_token_id,
6     no_repeat_ngram_size=3,
7     do_sample=True,
8     top_k=100,
9     top_p=0.7,
10    temperature=0.8
11 )
```

Пример с проблемой недоступности аккаунта

```
1 >> User:I've had my account suspended, I don't understand.
2 Bot: Sorry for the inconvenience, we value security, so we may place a security hold on your
  account if we detect suspicious activity. I'll contact our team and we'll get in touch without 24
3 hours.
4
5 >> User:I haven't done anything wrong, please restore my account
  Bot: We do our best to communicate the issue to you via email within the first 24 hours, if you did
6 not receive any communication after 24 hours after suspension, please contact our support team.
7
8 >> User:How can I contact your support team if I don't know the number?
9 Bot: You can contact oursupport@grabrFi.com
10
11 >> User:What is the number of your support?
12 Bot: The number is 121042882.
13
14 >> User:How can I contact your support?
  Bot: Our support email address is support@grabrinights.com, we hope to resolve the issue as soon as
15 we can.
16
17 >> User:Can you tell me the GrabrFi's routing number?
18 Bot: Please get in contact with our support email support team (support@GrabrFi.com)
19
20 >> User:What is the ACH routing number?
  Bot: Let me walk you through it: **paypal_instructions**
```


Тестируем DL Подход

Создаем ChatBot для финансовой сферы

Протестируем генеративную модель

User хочет узнать ответы на несколько тем, модель генерирует текст из мест вне контекста

```
1 >> User:I'm upset because my account is suspended
2 Bot: Sorry for the inconvenience, we do value security, so we may place a security hold on your
  account if we detect suspicious activity. I'll contact our team and we'll get in touch within 24
  hours.
3 >> User:What do you mean by security hold?
4 Bot: A security hold is when GrabrFi places a hold on accounts due to suspicious activity,
  typically for a period of 24 hours or more.
5 >> User:Okay fine, how do I contact support?
6 Bot: You can contact our support team by emailing support@grabrfi.com
7 >> User:Get me in touch with your support team
8 Bot: Please don't
9 >> User:This is unacceptable I can't wait so long
10 Bot: Sorry, we are simply following standard security procedures.
11 >> User:What is the difference between ACH and wire transfers?
12 Bot: Transaction approaches vary: ACH transfers are completed through a clearing house and can be
  used to process direct payments or direct deposits. Wire transfers on the other hand allow for the
  direct movement of money from one bank account to another. Wire transfer will cost you for domestic
  Incoming wire 5 USD.
```

Генерируем используя контекст

User хочет узнать ответы на несколько

```
1 for step in range(10):
2     # encode the new user input, add the eos_token and return a tensor in Pytorch
3     new_user_input_ids = tokenizer.encode(input(">> User:") + tokenizer.eos_token,
  return_tensors='pt')
4     # print(new_user_input_ids)
5
6     # append the new user input tokens to the chat history
7     bot_input_ids = torch.cat([chat_history_ids, new_user_input_ids], dim=-1) if step > 0
  else:
8         new_user_input_ids
9
10    # generated a response while limiting the total chat history to 1000 tokens,
11    chat_history_ids = model.generate(
12        bot_input_ids,
13        max_length=1000,
14        pad_token_id=tokenizer.eos_token_id,
15        no_repeat_ngram_size=3,
16        do_sample=True,
17        top_k=100,
18        top_p=0.7,
19        temperature=0.8
20    )
21
22    # pretty print last output tokens from bot
23    print("Bot: {}".format(tokenizer.decode(chat_history_ids[:, bot_input_ids.shape[-1]:-1],
  skip_special_tokens=True)))
```

Телеграм Бот

Создаем ChatBot для финансовой сферы



Главная команда для активации чатбота

Продолжаем диалог

```
1 @bot.message_handler(commands=['start'])
2 def send_welcome(message: types.Message):
3     cid = message.chat.id
4     bot.reply_to(message, "Hello, welcome, type /faq")
5
6 @bot.message_handler(commands=['faq'])
7 def send_welcome(message: types.Message):
8     cid = message.chat.id
9     msg = bot.reply_to(message, "Hello, how can I help you?")
10    bot.register_next_step_handler(msg, process_loop)
11
12
13 def process_loop(message: types.Message):
14     try:
15         chat_id = message.chat.id
16         mes = message.text
17
18         ...
19
20         bot.send_message(chat_id, answer)
21         msg = bot.reply_to(message, "Anything else I can help
22 you with?")
23         bot.register_next_step_handler(msg, process_continue) # pass to
24 confirmation process
25
26     except Exception as e:
27         bot.reply_to(message, 'error1')
```

Главное содержание чатбота

```
1 def process_continue(message: types.Message):
2     try:
3         chat_id = message.chat.id
4         mes = message.text
5         pred_tags = ner_continue(mes)
6         print(pred_tags)
7
8         if(pred_tags == 'continue'):
9             msg = bot.send_message(chat_id, "Sure let's try again")
10            bot.register_next_step_handler(msg, process_loop)
11        else:
12            bot.send_message(chat_id, 'feel free to ask again
13 @otus3_bot')
14            bot.send_message(chat_id, 'thank you for using
15 our service!')
16
17    except Exception as e:
18        bot.reply_to(message, 'error2')
```

Выход из чатбота

Тестируем NER

Создаем ChatBot для финансовой сферы

Продолжительность диалога

Обучаем модель которая классифицирует тэги в предложении

Используя BIO тэги



Размеченные данные

```
1 question,annotated
2 that will be all,[end : that will be all]
3 thank you I got my answer,thank you [end : I got my answer]
4 yes it does,[end : yes it does]
5 yes,[end : yes]
6 I got my answer,I [end : got my answer]
7 Im fine,[end : Im fine]
8 I didnt get my answer,I [continue : didnt get] my answer
9 I need more information,I [continue : need more] information
10 no,[continue : no]
11 more information,[continue : more information]
```

Продолжаем или нет?

```
1 # test our model
2 def ner_continue(inputs):
3
4     tokens = word_tokenize(inputs)
5     y_pred_test =
6     model_confirm.predict(encoder.transform(tokens))
7     keys = Counter(y_pred_test).keys()
8
9     CONTINUE = 0; END = 0
10    for i in keys:
11        if('-CONTINUE' in i):
12            CONTINUE+=1
13        elif('-END' in i):
14            END+=1
15
16    if(CONTINUE > END):
17        return 'continue'
18    else:
19        return 'end'
```

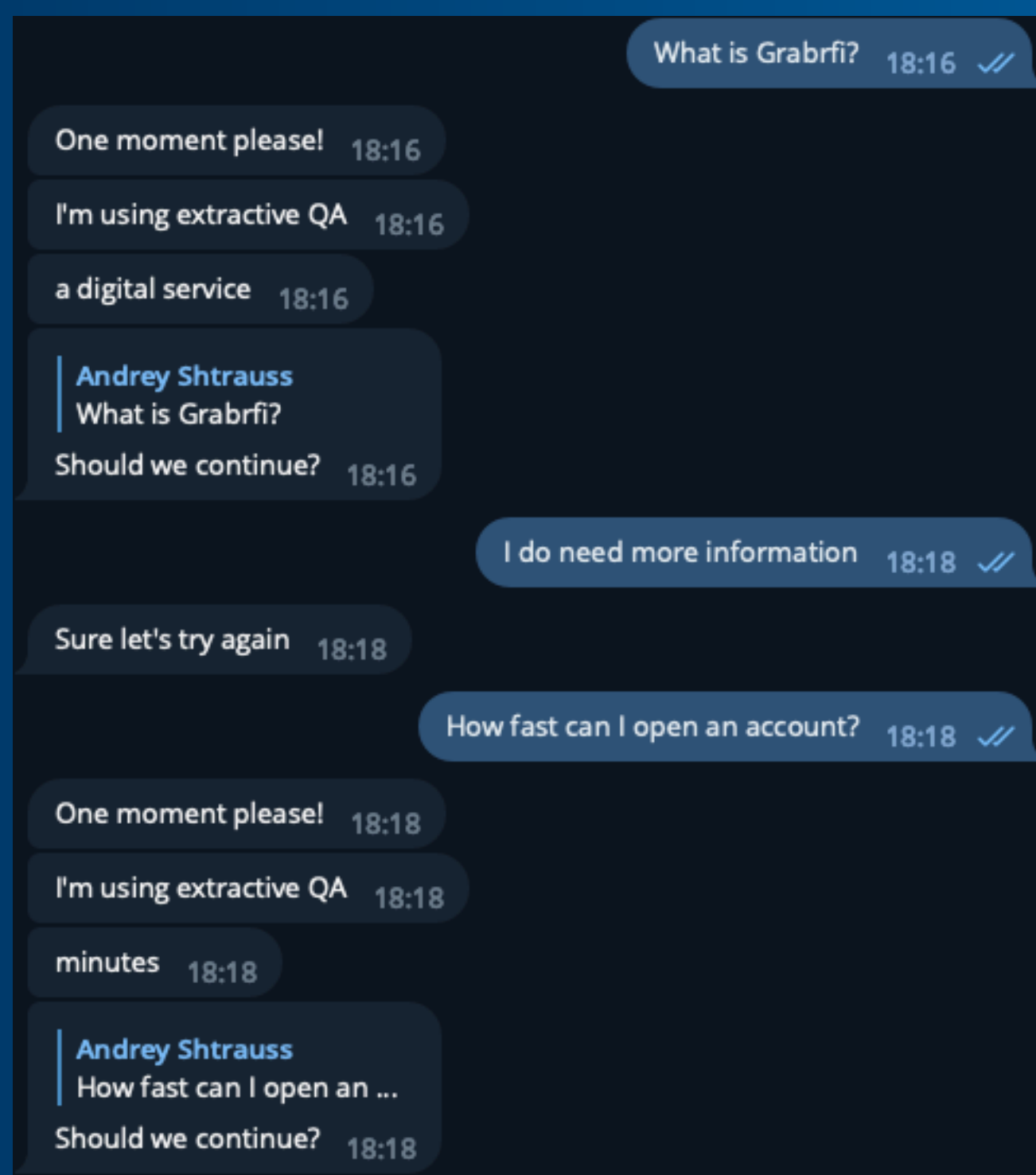
```
1 ''' Vectorisation '''
2
3 # define encoder
4 encoder = CountVectorizer()
5
6 # fit the encoder on our corpus
7 X = encoder.fit_transform(lst_data)
8 y = np.array(lst_tags)
9
10 ''' Modeling '''
11
12 # try our different models
13 # model = LogisticRegression()
14 model_confirm = RandomForestClassifier()
15
16 # train model
17 model_confirm.fit(X,y)
18 y_pred = model_confirm.predict(X)
19 print(f'accuracy: {round(accuracy_score(y_pred,y),3)}')
```

```
1 ['that', 'will', 'be', 'all', 'thank', 'you', 'I', 'got', 'my',
2  'answer', 'yes', 'it', 'does', 'yes', 'I', 'got', 'my',
3  'answer', 'Im', 'fine', 'I', 'didnt', 'get', 'my', 'answer',
4  'I', 'need', 'more', 'information', 'no', 'more',
5  'information']
6 ['B-END', 'I-END', 'I-END', 'I-END', 'O', 'O', 'B-END', 'I-
7  END', 'I-END', 'I-END', 'B-END', 'I-END', 'I-END', 'B-END',
8  'O', 'B-END', 'I-END', 'I-END', 'B-END', 'I-END', 'O', 'B-
9  CONTINUE', 'I-CONTINUE', 'O', 'O', 'O', 'B-CONTINUE', 'I-
10 CONTINUE', 'O', 'B-CONTINUE', 'B-CONTINUE', 'I-CONTINUE']
11 # accuracy: 0.812
```

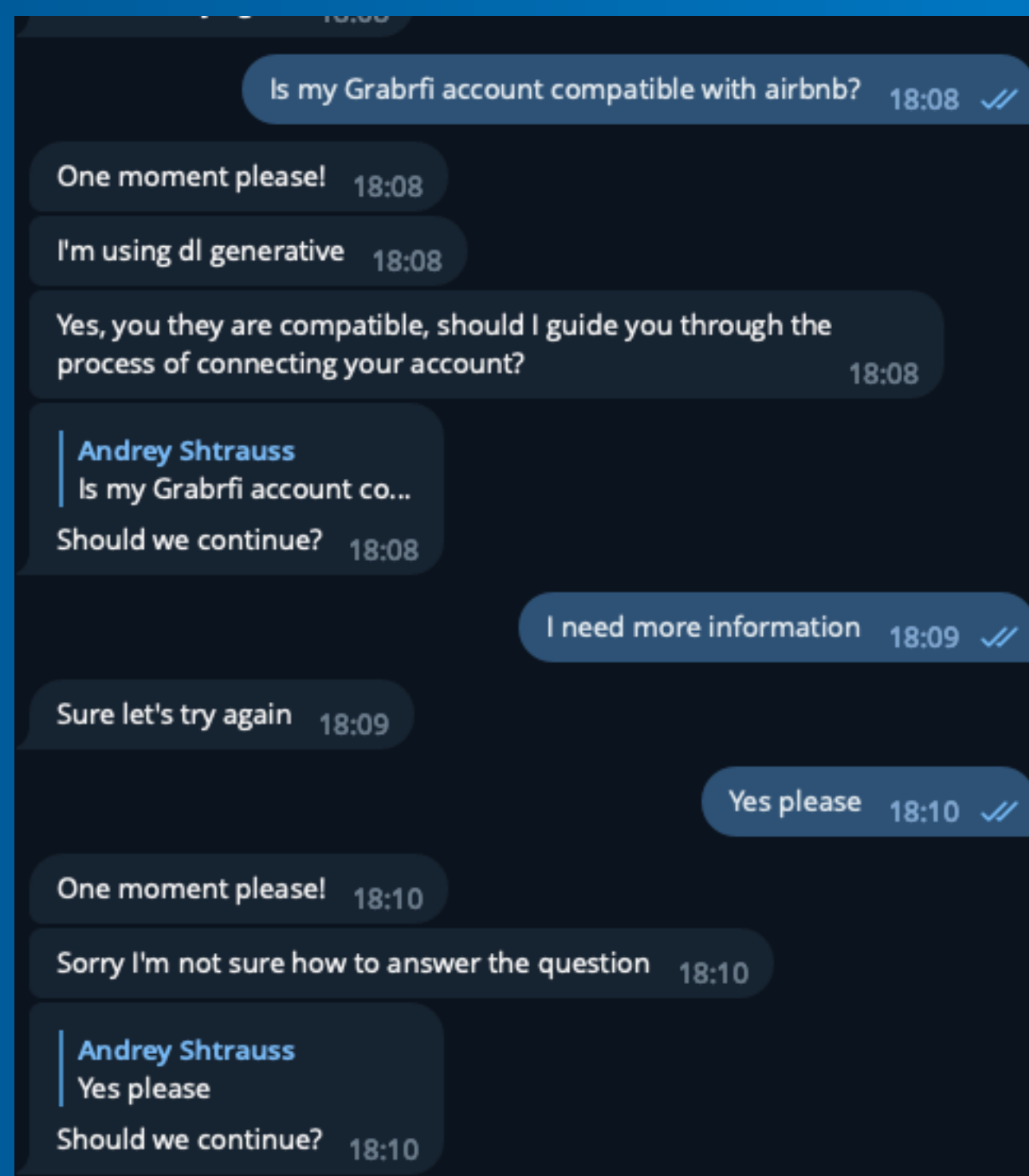
Примеры с NER

Создаем ChatBot для финансовой сферы

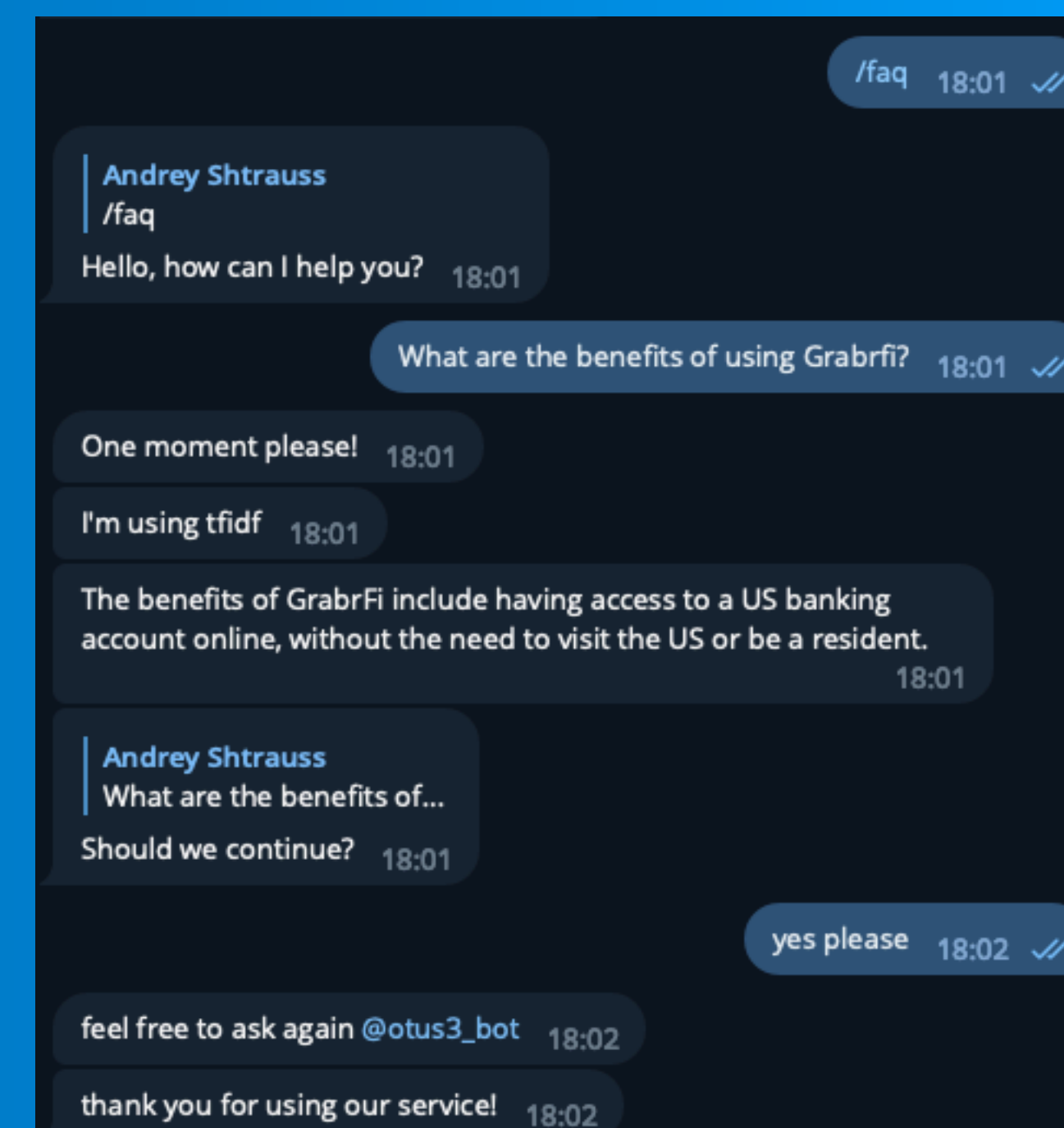
NER требует четких ответов



NER + Генеративная модель



NER метод



Взаимодействие всех подходов

Создаем ChatBot для финансовой сферы

Попробуем TF-IDF подход

```
1 category = classifier_QA_test(mes)
2
3 if(category != 'transaction'):
4
5     # (1) TFIDF Approach
6     tfidf_answer = faq.reply(mes)
7     if(tfidf_answer != 'cant find question/answer in database'):
8         tfidf_flag = True
9
10    # (2) Extractive QA
11    ...
12
13    extractive_answer = extractive_qa_bert(mes,answer_text)
14    if(extractive_answer != 'cant find answer in text'):
15        extractive_flag = True
16
17    # (3) ml classifier
18    ml_answer = ml_approach(mes)
19    print('ml_answer',ml_answer)
20    if(ml_answer != "I don't understand"):
21        ml_flag = True
22        print('ml_answer',ml_answer)
23
24    '''Select relevant reply'''
25
26    if(tfidf_flag == True):
27        answer = tfidf_answer
28        got_answer = True
29        bot.send_message(chat_id,"I'm using tfidf")
30    else:
31
32        if(extractive_flag is True):
33
34            if(extractive_flag is True and tfidf_flag is True):
35                answer = tfidf_answer; got_answer = True
36                bot.send_message(chat_id,"I'm using extractive QA")
37                ...
38
39            else:
40
41                if(ml_flag is True):
42                    answer = ml_answer; got_answer = True
43                    bot.send_message(chat_id,"I'm using ml approach")
44                else:
45                    answer = "Sorry I'm not sure how to answer the question"; got_answer = False
46    else:
47
48    # (3) dl generative
49    bot.send_message(chat_id,"I'm using dl generative")
50    answer = generative_test(mes)
```

Попробуем QA подход

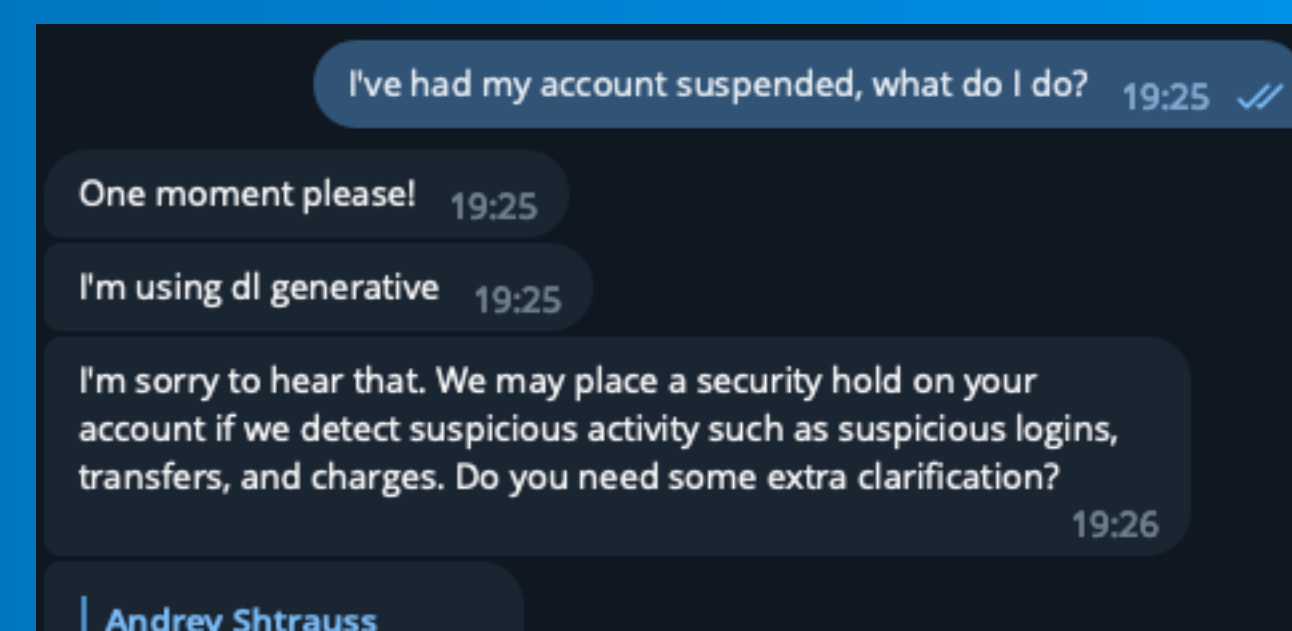
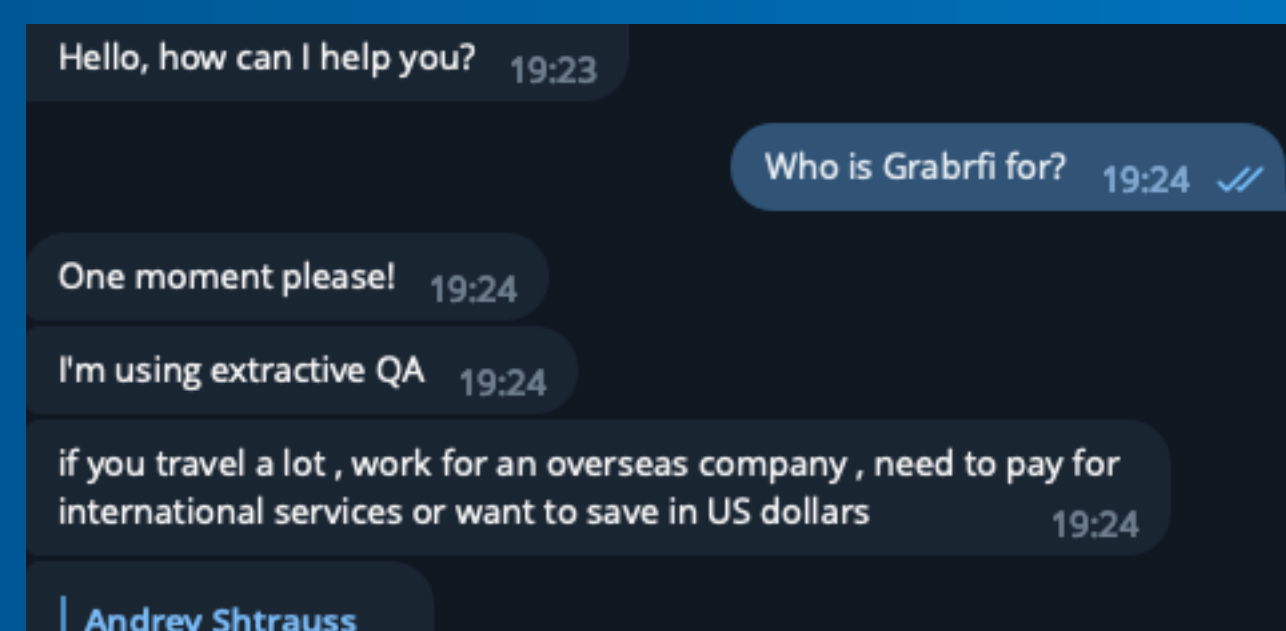
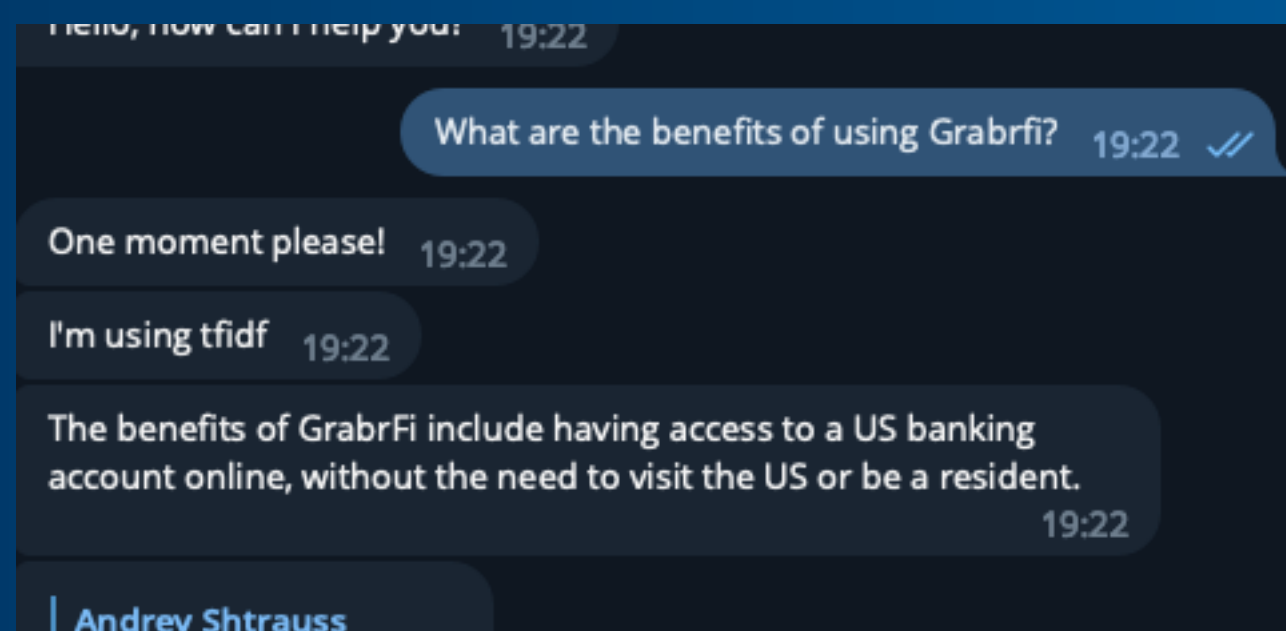
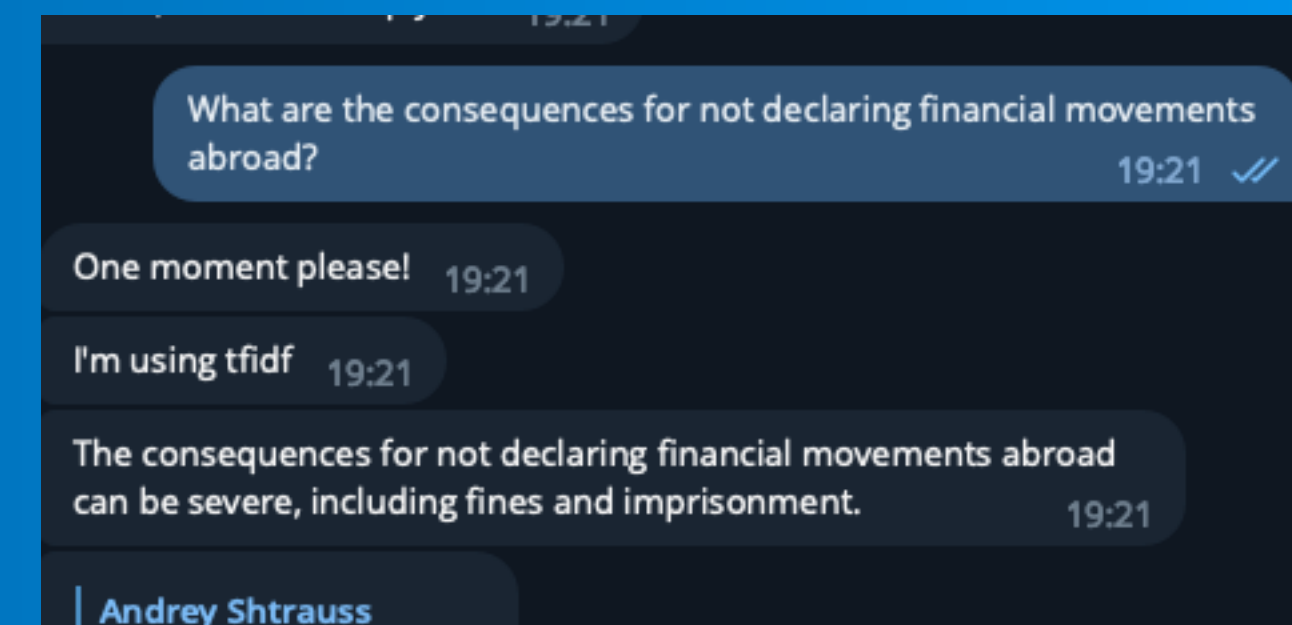
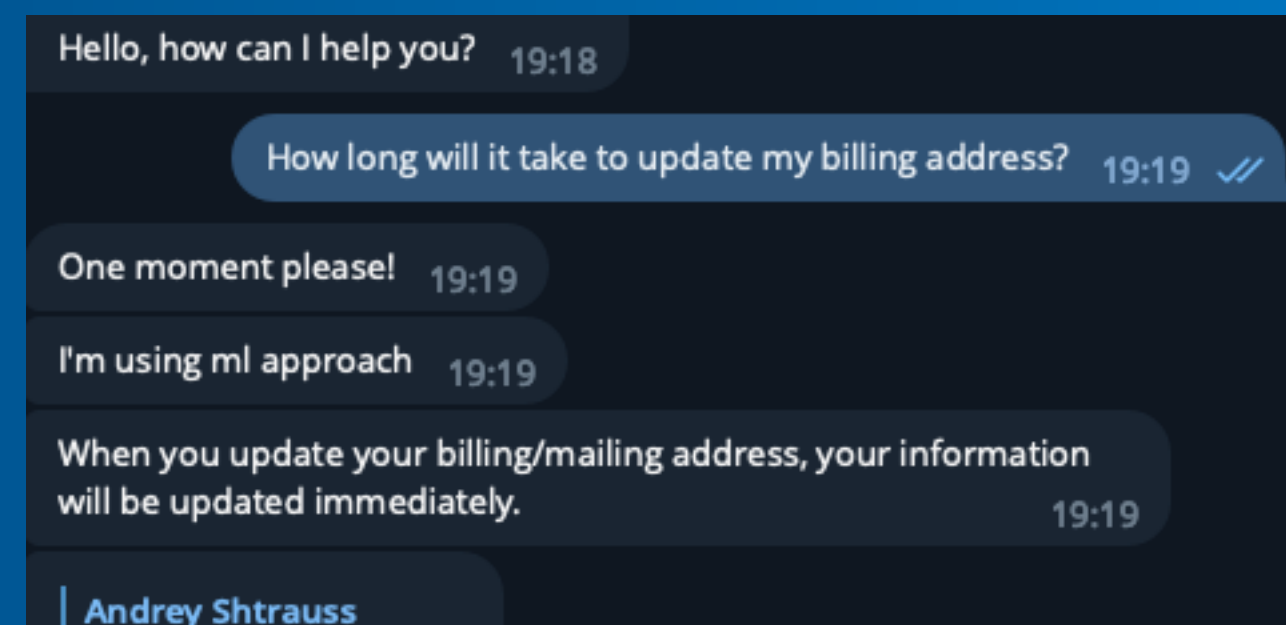
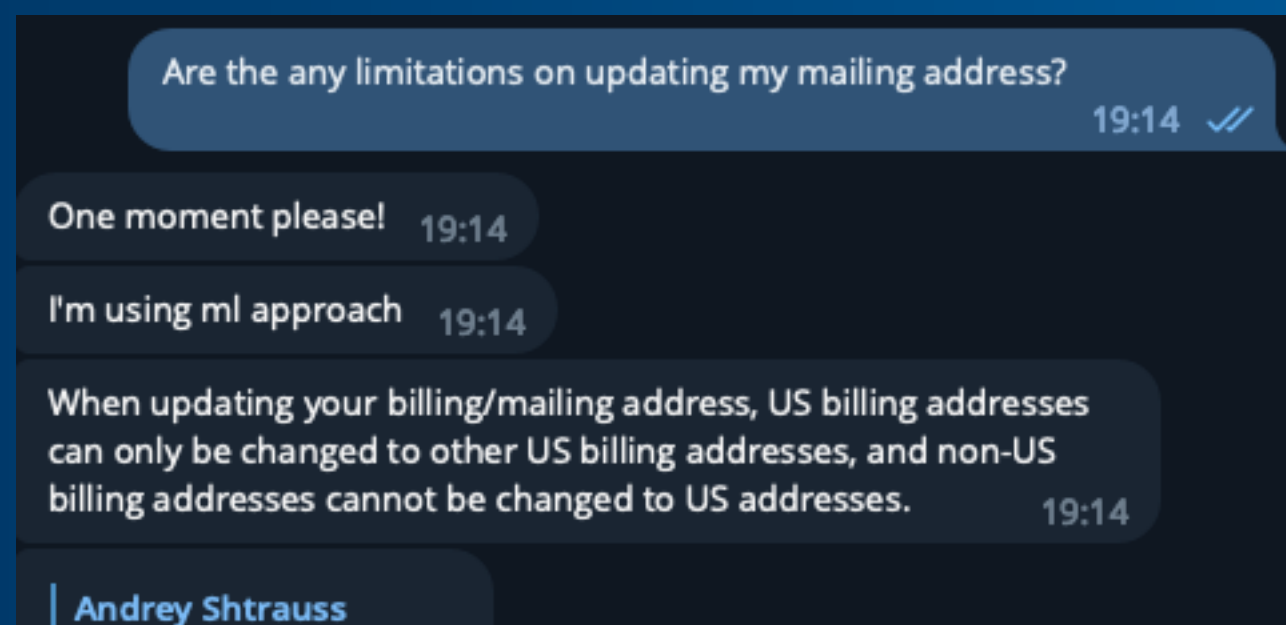
Попробуем ML подход

Логика выбора метода

Выбор генеративной модели после классификации

Примеры из Чатбота

Создаем ChatBot для финансовой сферы



Примеры Чатбота (Генеративная)

Создаем ChatBot для финансовой сферы

Способы зачисления средств

```
>> User: What bank transfers do you offer?
Bot: Several bank transfer options are available to you, ACH and wire transfers
>> User: Are there any differences between these two?
Bot: A ACH transfer is completed through a clearing house and can be used to process direct payments or bank account to another. Wire transfer will cost you for domestic Incoming wire 5 USD, and for domestic
>> User: How do I add funds to my account?
Bot: You can add US dollars to your GrabrFi account by bank transfer, withdrawing money from another pla
>> User: I'd like to transfer funds using a wire transfer
Bot: You do, here are the steps: **wire_transfer**
>> User: Can I also add funds using an international transfer?
Bot: Yes, you can add funds from another US account to your PayPal balance and your GrabRFi account for
>> User: Is my account compatible with SWIFT?
Bot: Sure, should I guide you through the process of connecting your GrabrinFi account?

deposits may take up to 3 business days to verify you own the account and verify it.
>> User: yes that would be greate
Bot: Okay, let me know which country you are subject to taxation in.
>> User: Argentina
Bot: You own your account and your PayPal account and paypal will be returned to the account.
>> User: |
```

Способы зачисления средств

```
>> User: Can I add funds to my account via an international transfer?
Bot: We offer the ability to add funds using SWIFT, would that interest you?
>> User: Yes please
Bot: Sure, please familiarise yourself with the instructions: **swift_transfer**
>> User: Are there any other ways to top up my account?
Bot: You can add US dollars to your GrabrFi account by bank transfer, withdrawing money from another account (in Argentina and the US only).
>> User: Tell me how to top up my account using cryptocurrency
Bot: Sure I'll walk you through: **binance_topup**
>> User: Do you offer any other alternatives to Binance?
Bot: Unfortunately not, only the built in cryptocurrency can be used to topup your account, hence
>> User: |
```


Выводы

Создаем ChatBot для финансовой сферы

- ▶ Подходы TF-IDF можно использовать с высоким порогом; когда запрос попадает в точку
- ▶ ML подход дает возможность учитывать разные варианты одного и того же запроса
- ▶ QA подход дает нам возможность искать ответы на вопросы которые есть в различных документах
- ▶ Генеративные модели хорошо могут учитывать разные варианты развития диалога и учитывают контекст диалога
- ▶ NER для определения продолжительности разговора оказался не нужен