

# Идентификация сентимента индивидуальных инвесторов анализ финсообществ в Telegram

Твердые цифры - VTB Capital

4 октября 2021 г.

# Знакомство

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?
  - ▶ Fear of forward guidance

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?
  - ▶ Fear of forward guidance
  - ▶ Hard Numbers: Open Consumer Price Database



# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?
  - ▶ Fear of forward guidance
  - ▶ Hard Numbers: Open Consumer Price Database
- [alexisakov @ GitHub](#):

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?
  - ▶ Fear of forward guidance
  - ▶ Hard Numbers: Open Consumer Price Database
- [alexisakov @ GitHub](#):
  - ▶ QPM модель платежного баланса: <https://github.com/alexisakov/reforma>

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?
  - ▶ Fear of forward guidance
  - ▶ Hard Numbers: Open Consumer Price Database
- [alexisakov @ GitHub](#):
  - ▶ QPM модель платежного баланса: <https://github.com/alexisakov/reforma>
  - ▶ реплика базовой модели Банка России модель платежного баланса: <https://github.com/alexisakov/replica>

# Александр Исаков

- главный экономист VTB Capital по России и СНГ - [публичные резюме @ Твердые цифры - VTB Capital](#)
- ранее: Банк России + [к.э.н.](#) + ВТБ
- [Google Scholar](#):
  - ▶ Современная денежно-кредитная политика: обоснованная критика или типичные заблуждения экспертного сообщества?
  - ▶ Fear of forward guidance
  - ▶ Hard Numbers: Open Consumer Price Database
- [alexisakov @ GitHub](#):
  - ▶ QPM модель платежного баланса: <https://github.com/alexisakov/reforma>
  - ▶ реплика базовой модели Банка России модель платежного баланса: <https://github.com/alexisakov/replica>
  - ▶ руководство по работе с твердыми цифрами: <https://github.com/alexisakov/Economica>

# Содержание

# Содержание

- motivation: теория + практика

# Содержание

- motivation: теория + практика
- метод: периметр наблюдения + data pipeline

# Содержание

- motivation: теория + практика
- метод: периметр наблюдения + data pipeline
- первые итоги: стратегия + чувствительность



# Мотивация

# Мотивация: Теория

Два направления исследования:

- поведения розничных инвесторов на финрынке:

# Мотивация: Теория

Два направления исследования:

- поведения розничных инвесторов на финрынке:
  - ▶ **(van der Beck & Jaunin, 2021)**: неэластичный институциональный спрос позволяет розничным инвесторам иметь непропорциональные эффекты для оценок компаний - при доле рынка в 0.2% повышать цены смолл-капов на 25%

# Мотивация: Теория

Два направления исследования:

- поведения розничных инвесторов на финрынке:
  - ▶ **(van der Beck & Jaunin, 2021)**: неэластичный институциональный спрос позволяет розничным инвесторам иметь непропорциональные эффекты для оценок компаний - при доле рынка в 0.2% повышать цены смолл-капов на 25%
  - ▶ **(Kalda et al., 2021)**: тесная связь между платформой и структурой портфеля

# Мотивация: Теория

Два направления исследования:

- поведения розничных инвесторов на финрынке:
  - ▶ **(van der Beck & Jaunin, 2021)**: неэластичный институциональный спрос позволяет розничным инвесторам иметь непропорциональные эффекты для оценок компаний - при доле рынка в 0.2% повышать цены смолл-капов на 25%
  - ▶ **(Kalda et al., 2021)**: тесная связь между платформой и структурой портфеля
- анализ неструктурированных текстовых данных:

# Мотивация: Теория

Два направления исследования:

- поведения розничных инвесторов на финрынке:
  - ▶ **(van der Beck & Jaunin, 2021)**: неэластичный институциональный спрос позволяет розничным инвесторам иметь непропорциональные эффекты для оценок компаний - при доле рынка в 0.2% повышать цены смолл-капов на 25%
  - ▶ **(Kalda et al., 2021)**: тесная связь между платформой и структурой портфеля
- анализ неструктурированных текстовых данных:
  - ▶ **(Costola et al., 2021)**: сентимент новостей о ковиде на основных сайтах (Reuters, MarketWatch и т.п.) опережает рыночные изменения

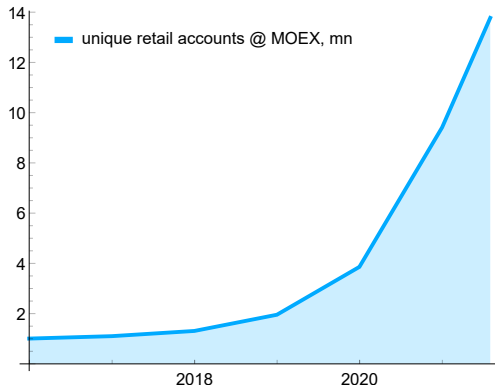
# Мотивация: Теория

Два направления исследования:

- поведения розничных инвесторов на финрынке:
  - ▶ **(van der Beck & Jaunin, 2021)**: неэластичный институциональный спрос позволяет розничным инвесторам иметь непропорциональные эффекты для оценок компаний - при доле рынка в 0.2% повышать цены смолл-капов на 25%
  - ▶ **(Kalda et al., 2021)**: тесная связь между платформой и структурой портфеля
- анализ неструктурированных текстовых данных:
  - ▶ **(Costola et al., 2021)**: сентимент новостей о ковиде на основных сайтах (Reuters, MarketWatch и т.п.) опережает рыночные изменения
  - ▶ **(Rayarel, 2021)**, **(Gjerstad et al., 2021)** - влияние твитов Трампа по темам и компаниям

## Мотивация: Практика

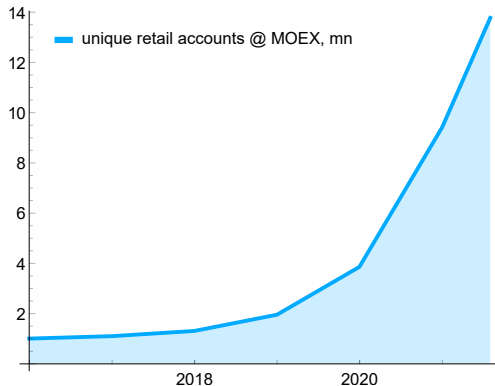
- объем вложений розничных инвесторов по состоянию на 1к21: 6,0 трлн руб. (против 2,6 трлн руб. в 1к19)





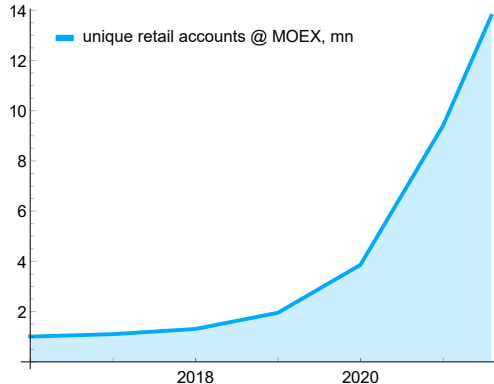
## Мотивация: Практика

- объем вложений розничных инвесторов по состоянию на 1к21: 6,0 трлн руб. (против 2,6 трлн руб. в 1к19)
- по статистике МосБиржи: 14 млн уникальных розничных клиентов, 2.5 млн активных счетов



## Мотивация: Практика

- объем вложений розничных инвесторов по состоянию на 1к21: 6,0 трлн руб. (против 2,6 трлн руб. в 1к19)
- по статистике МосБиржи: 14 млн уникальных розничных клиентов, 2.5 млн активных счетов
- до 15% заявок в ходе последних IPO пришлось на розничных инвесторов



## Метод

# Периметр: Telegram финсообщества

- Telegram - одно из наиболее популярных медиа в России



## Top 10 Apps by Downloads | Q2 2021

RUSSIA

1		▲ +1	Telegram	6		▼ -3	GetContact
2		▼ -1	TikTok	7		▲ +5	State Service
3		▲ +51	Zenly	8		▼ -4	OZON.ru
4		▲ +2	WhatsApp Messenger	9		▼ -2	Wildberries
5			Instagram	10			Avito

Note: Downloads based on combined iOS App Store and Google Play.  
Rank change vs Q1 2021. All estimates from App Annie Intelligence.

## Периметр: Telegram финсообщества

- Telegram - одно из наиболее популярных медиа в России
- сообщества - не аффилированные с брокерами/банками/новостными агентствами



### Top 10 Apps by Downloads | Q2 2021

RUSSIA

1		▲ +1	Telegram	6		▼ -3	GetContact
2		▼ -1	TikTok	7		▲ +5	State Service
3		▲ +51	Zenly	8		▼ -4	OZON.ru
4		▲ +2	WhatsApp Messenger	9		▼ -2	Wildberries
5			Instagram	10			Avito

Note: Downloads based on combined iOS App Store and Google Play.  
Rank change vs Q1 2021. All estimates from App Annie Intelligence.

## Периметр: Telegram финсообщества

- Telegram - одно из наиболее популярных медиа в России
- сообщества - не аффилированные с брокерами/банками/новостными агентствами
- ... более 1000 подписчиков/участников



### Top 10 Apps by Downloads | Q2 2021

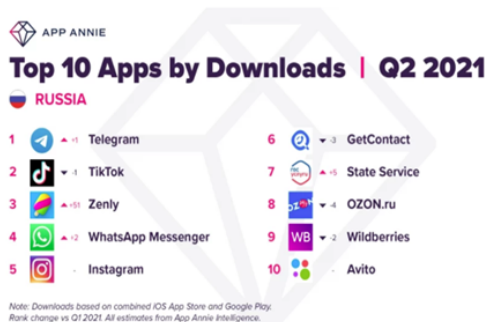
RUSSIA

1		▲ +1	Telegram	6		▼ -3	GetContact
2		▼ -1	TikTok	7		▲ +5	State Service
3		▲ +51	Zenly	8		▼ -4	OZON.ru
4		▲ +2	WhatsApp Messenger	9		▼ -2	Wildberries
5			Instagram	10			Avito

Note: Downloads based on combined iOS App Store and Google Play.  
Rank change vs Q1 2021. All estimates from App Annie Intelligence.

## Периметр: Telegram финсообщества

- Telegram - одно из наиболее популярных медиа в России
- сообщества - не аффилированные с брокерами/банками/новостными агентствами
- ... более 1000 подписчиков/участников
- аудитория  $\approx 1,5$  млн, 75% активных счетов



# Data pipeline

Четыре направления:

- сырой текст



# Data pipeline

Четыре направления:

- сырой текст
- NER - идентификация упоминаний организаций в принципе (natasha, spaCy)

# Data pipeline

Четыре направления:

- сырой текст
- NER - идентификация упоминаний организаций в принципе (natasha, spaCy)
- идентификация эмитентов - из упомянутых организаций (logit, RNN, тикеры)

# Data pipeline

Четыре направления:

- сырой текст
- NER - идентификация упоминаний организаций в принципе (natasha, spaCy)
- идентификация эмитентов - из упомянутых организаций (logit, RNN, тикеры)
- идентификация сентимента - (logit, классификатор над [SBERT от \(SberDevices, 2021\)](#))

## Data pipeline: Сырой текст

*Россия - одно из лучших мест в мире для производства сельхоз продукции. Вчерашнее SPO Русагро в очередной раз показывает, что умные деньги инвестируют в этот сектор: доли в Русагро приобрели западные инвестфонды, увеличили пакеты CEO Русагро Максим Басов (до 7.7%) и Максим Воробьёв (до 10.1%).*

*В России низкая себестоимость производства агропродукции и вторые по площади в мире пашни. В результате Россия заняла в последние годы ведущие позиции в мире по экспорту агрокультур: • пшеницы (1 место, 39 млн т), • сахара (2 место, 45 млн т), • и подсолнечного масла (2 место 3.3 млн т).*

*Максим Воробьев верит в перспективы Русагро: «Фокус компании на новейшие сельскохозяйственные технологии и цифровизацию бизнес-процессов позволит занять достойное место среди ведущих мировых агрохолдингов»*

*«В России растет производство сельхоз продукции. Это связано с улучшением климата и инвестициями в технологии» — считает CEO агрохолдинга Русагро Максим Басов.*

## Data pipeline: NER

*Россия - одно из лучших мест в мире для производства сельхоз продукции. Вчерашнее SPO [Русagro](#) в очередной раз показывает, что умные деньги инвестируют в этот сектор: доли в [Русagro](#) приобрели западные инвестфонды, увеличили пакеты [CEO Русagro](#) Максим Басов (до 7.7%) и Максим Воробьёв (до 10.1%).*

*В России низкая себестоимость производства агропродукции и вторые по площади в мире пашни. В результате Россия заняла в последние годы ведущие позиции в мире по экспорту агрокультур: • пшеницы (1 место, 39 млн т), • сахара (2 место, 45 млн т), • и подсолнечного масла (2 место 3.3 млн т).*

*Максим Воробьев верит в перспективы [Русagro](#): «Фокус компании на новейшие сельскохозяйственные технологии и цифровизацию бизнес-процессов позволит занять достойное место среди ведущих мировых агрохолдингов»*

*«В России растет производство сельхоз продукции. Это связано с улучшением климата и инвестициями в технологии» — считает CEO агрохолдинга [Русagro](#) Максим Басов.*

# Data pipeline: распознавание упоминаний

- Русагро → AGRO

# Data pipeline: распознавание упоминаний

- Русагро → AGRO
- CEO → не эмитент

# Data pipeline: распознавание упоминаний

- Русагро → AGRO
- CEO → не эмитент
- модели: тикеры, logreg, RNN



# Data pipeline: распознавание сентимента

Шкала 1..5:

- ❶ Справедливая стоимость **ПОНИЖЕНА**. ?? ORG: справедливая стоимость 205 руб., угроза падения -25%.

# Data pipeline: распознавание сентимента

Шкала 1..5:

- 1 **Справедливая стоимость ПОНИЖЕНА.** ?? ORG: справедливая стоимость 205 руб., угроза падения -25%.
- 2 **ORG ДИВИДЕНДЫ ORG УПАЛИ НА 86% В 2015 ГОДУ.** Вот что происходит с бизнесом и дивидендами во время девальвации, падения нефти

# Data pipeline: распознавание сентимента

Шкала 1..5:

- ❶ **Справедливая стоимость ПОНИЖЕНА.** ?? ORG: справедливая стоимость 205 руб., угроза падения -25%.
- ❷ **ORG ДИВИДЕНДЫ ORG УПАЛИ НА 86% В 2015 ГОДУ.** Вот что происходит с бизнесом и дивидендами во время девальвации, падения нефти
- ❸ **Новости компаний** — обзоры прессы перед открытием рынка ??Минприроды: запасов нефти в России хватит на 59 лет, газа — на 103 года ??В России появился новый ретейлер под руководством экс-главы ORG

# Data pipeline: распознавание сентимента

Шкала 1..5:

- 1 **Справедливая стоимость ПОНИЖЕНА.** ?? ORG: справедливая стоимость 205 руб., угроза падения -25%.
- 2 **ORG ДИВИДЕНДЫ ORG УПАЛИ НА 86% В 2015 ГОДУ.** Вот что происходит с бизнесом и дивидендами во время девальвации, падения нефти
- 3 **Новости компаний** — обзоры прессы перед открытием рынка ??Минприроды: запасов нефти в России хватит на 59 лет, газа — на 103 года ??В России появился новый ретейлер под руководством экс-главы ORG
- 4 **ORG имеет несколько сильных драйверов роста.** ORG уже повысил эффективность, что отразилось на хорошей отчетности за 1П 2020 года. Новая стратегия развития. В этом году ORG представит обновленную стратегию развития

# Data pipeline: распознавание сентимента

Шкала 1..5:

- 1 **Справедливая стоимость ПОНИЖЕНА.** ?? ORG: справедливая стоимость 205 руб., угроза падения -25%.
- 2 **ORG ДИВИДЕНДЫ ORG УПАЛИ НА 86% В 2015 ГОДУ.** Вот что происходит с бизнесом и дивидендами во время девальвации, падения нефти
- 3 **Новости компаний** — обзоры прессы перед открытием рынка ??Минприроды: запасов нефти в России хватит на 59 лет, газа — на 103 года ??В России появился новый ретейлер под руководством экс-главы ORG
- 4 **ORG имеет несколько сильных драйверов роста.** ORG уже повысил эффективность, что отразилось на хорошей отчетности за 1П 2020 года. Новая стратегия развития. В этом году ORG представит обновленную стратегию развития
- 5 **ORG : СПРАВЕДЛИВАЯ ЦЕНА - МИНИМУМ 23400 РУБЛЕЙ.** #оценка ORG  
Фундаментальная цена акции ORG при различных сценариях дивидендных выплат составляет от 23400 руб. до 28300 руб. Таким образом, даже в самом худшем случае фундаментальная цена акции составляет 23400 руб, апсайд **11.8%**.

# Data pipeline: распознавание сентимента

- seed sample size  $\approx 2k$

# Data pipeline: распознавание сентимента

- seed sample size  $\approx 2k$
- общий сентимент - средняя по оценкам:

# Data pipeline: распознавание сентимента

- seed sample size  $\approx 2k$
- общий сентимент - средняя по оценкам:
  - ▶ BERT - классификатор над SBERT от (SberDevices, 2021)



# Data pipeline: распознавание сентимента

- seed sample size  $\approx 2k$
- общий сентимент - средняя по оценкам:
  - ▶ BERT - классификатор над SBERT от (SberDevices, 2021)
  - ▶ logreg - регуляризированный logit на TF-IDF

## Результаты

# Первые результаты

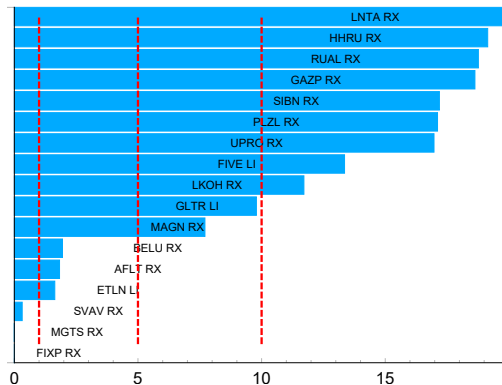
- 1 данные показывают, что для 1/3 российских акций публичный sentiment может быть причиной изменений котировок

# Первые результаты

- ❶ данные показывают, что для 1/3 российских акций публичный sentiment может быть причиной изменений котировок
- ❷ публичный sentiment выглядит перспективным инструментом генерации идей - особенно по коротким позициям

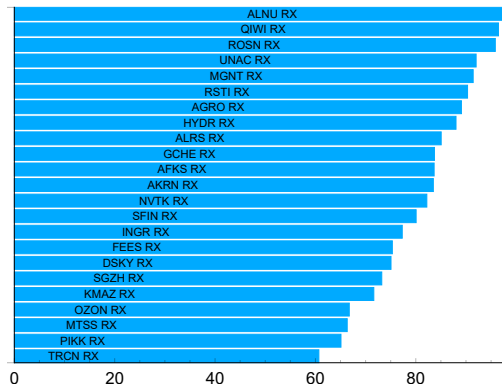
# Публичный sentiment может быть причиной изменений котировок

Средний p-value для Грейнджер теста по лагам 1-5



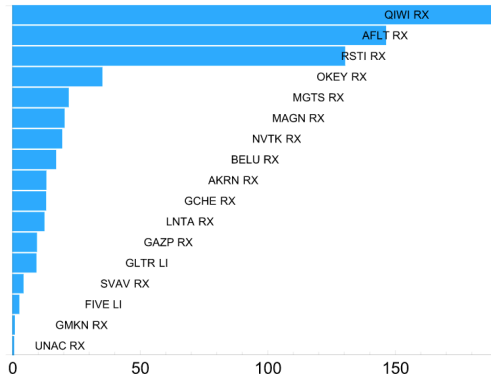
# Публичный sentiment может быть причиной изменений котировок

Средний p-value для Грейнджер теста по лагам 1-5



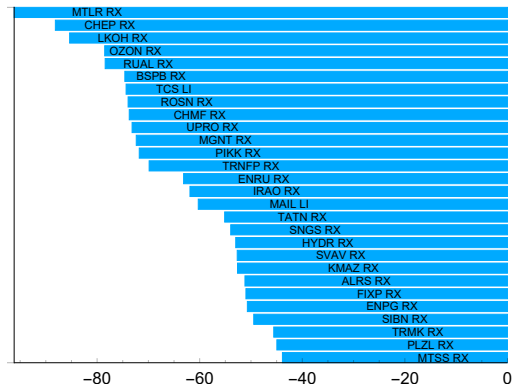
# Перспективный инструмент генерации идей

Разница в размере портфеля против buy+hold за янв'19-авг'21:



# Перспективный инструмент генерации идей

Разница в размере портфеля против buy+hold за янв'19-авг'21:





# На связи

- Твердые цифры - VTB Capital

# На связи

- Твердые цифры - VTB Capital
- ответим

