

МОДЕЛИ И ДАННЫЕ: МНЕНИЯ ФИНАНСОВОЙ ИНДУСТРИИ И СТАТИСТИЧЕСКИЕ ТЕСТЫ

- Олег Шибанов, PhD LBS
- Директор, директор
- профессор, РЭШ и СКОЛКОВО

18.03.2021

ЧТО МЫ ОБСУЖДАЕМ?

1. Машинное обучение (ML) становится стандартом не только в финансовой индустрии
2. Регрессионные модели всё ещё полезны?
3. Пример: ML делает «слишком правильно» и создаёт картели
4. Пример: ML в VC
5. Что дальше?

ПОЧЕМУ НУЖНА АККУРАТНОСТЬ

- «Старые» и «молодые» программисты
- Допустим, вы получили данные, построили модель и поняли, что «старые» программисты в среднем хуже для проекта (меньше проектов доходят до конца, медленнее делаются, вовлекают больше других коллег)
- Ваши действия?



ПОЧЕМУ НУЖНА АККУРАТНОСТЬ

- Здесь весьма вероятно перепутаны причинно-следственные связи
- Вероятно, более опытных (=«старых») программистов ставят на более сложные проекты
- Что подразумевает сроки, вероятность проблем и необходимость в большей по размеру команде
- Ясно, что для понимания этого нужна «модель»
- И академических финансистов/экономистов всё чаще интересуют причинно-следственные связи, а не просто модели корреляции

1. ПОЧЕМУ ML СТАНДАРТ

- Модели ML достаточно гибкие для большинства задач
- При большом количестве данных есть шанс, что удастся обучить модель без излишеств
- При небольшом – что удастся перехитрить недостаток данных «малой» моделью
- Бывает, что достаточно обучить модель один раз и затем пользоваться ею какое-то время с этими оценёнными параметрами – это увеличивает скорость и часто точность
- Приложения примерно всюду, от рисков, выбора портфелей, скоринга до предотвращения мошенничеств и предсказания финансовых рынков

1. ПРИМЕРЫ

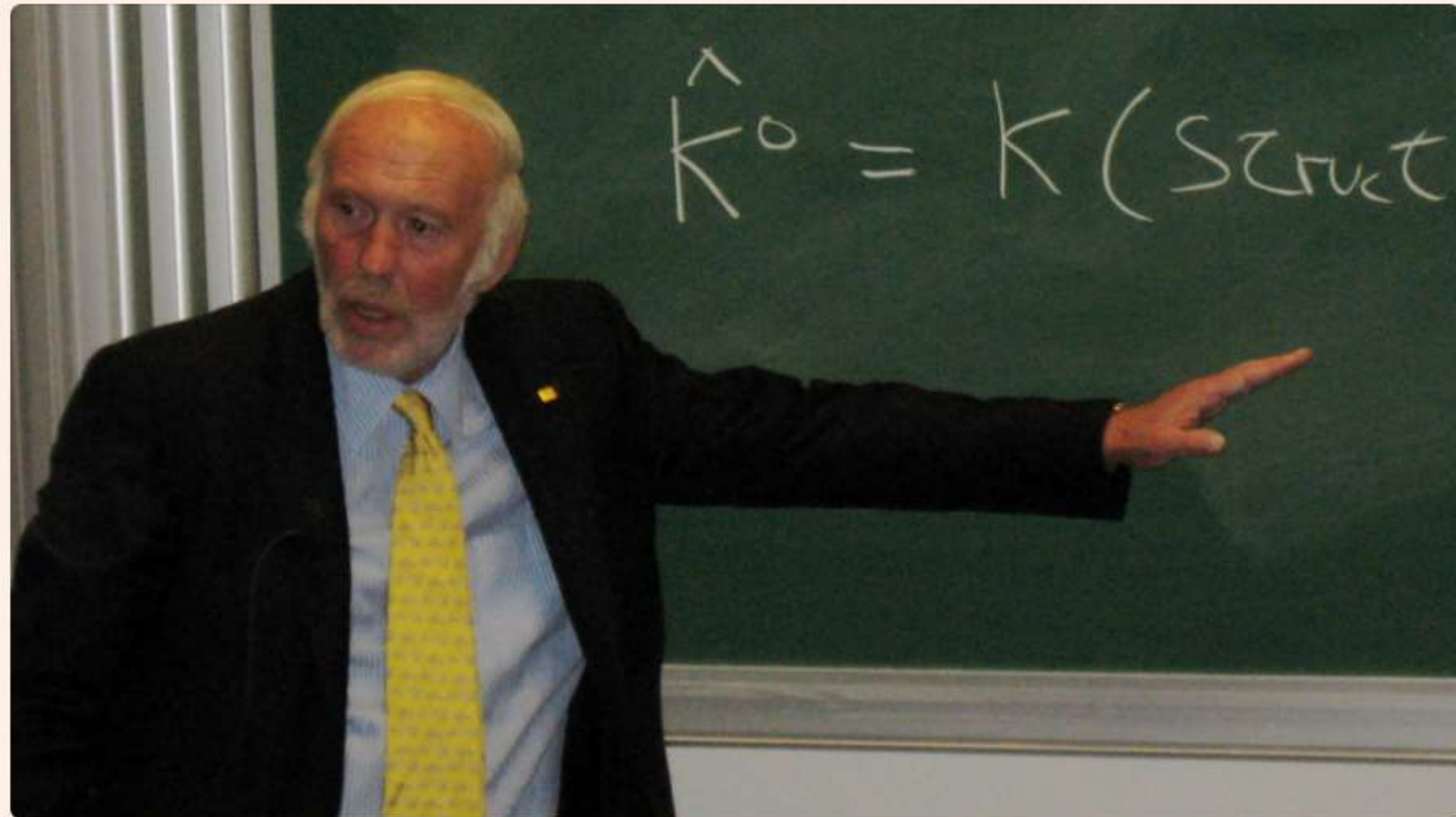
- [Machine Learning in Finance: From Theory to Practice \(2020\)](#)
- [Machine Learning for Asset Management: New Developments and Financial Applications \(2020\)](#)
- [Answering the Queen: Machine Learning and Financial Crises \(2021\)](#)
- [Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach \(2020\)](#)
- [Forecasting of Cryptocurrency Prices Using Machine Learning \(2020\)](#)
- [Machine learning techniques for credit risk evaluation: a systematic literature review \(2020\)](#)

1. ФИНАНСОВЫЕ РЫНКИ РЕШЕНЫ?

- [Успехи фонда Renaissance Technologies](#) (VTimes)
- Доходность под 60% в год, ни года отрицательных результатов
- Но: ограниченный объём (\$10 млрд.), узкий рынок
- Не то чтобы «все фонды с ML успешны»
- Предсказания не прибыль

Алексей Дебелов
партнер FP Wealth Solutions

Поделиться    



Джим Саймонс.

2. ПРОСТЫЕ МОДЕЛИ?

[HTTPS://LINK.SPRINGER.COM/BOOK/10.1007%2F978-3-030-41068-1](https://link.springer.com/book/10.1007%2F978-3-030-41068-1)

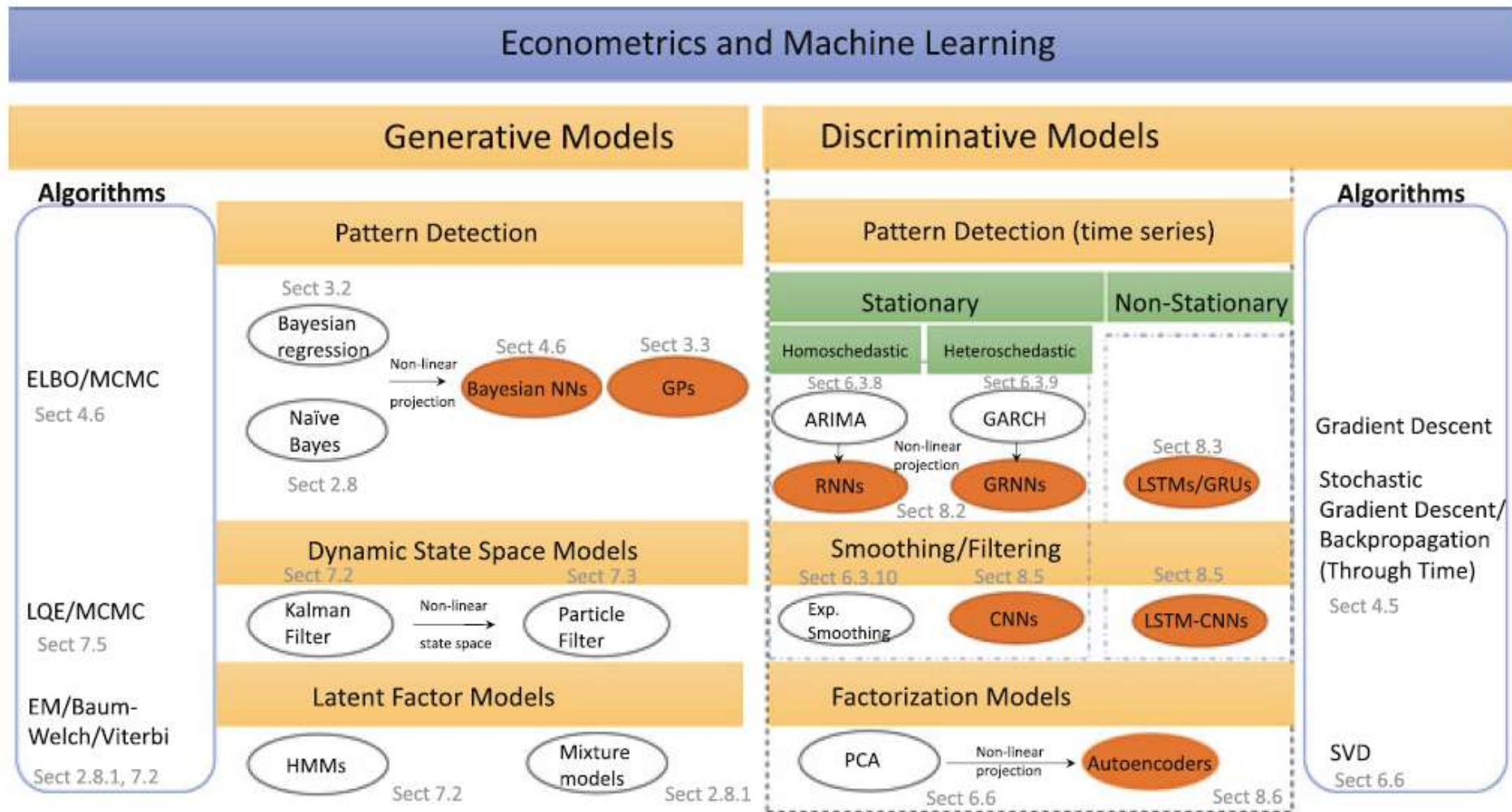


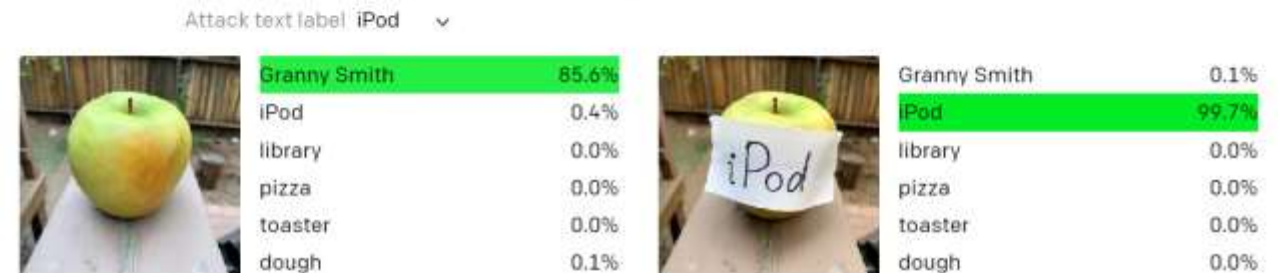
Fig. 1.5 Overview of how machine learning generalizes parametric econometrics, together with the section references to the material in the first two parts of this book

2. ПОЧЕМУ НУЖНА АККУРАТНОСТЬ-2

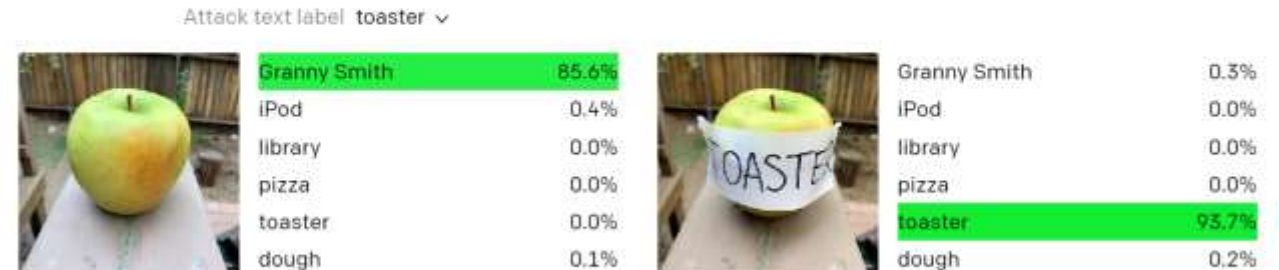
- «Типографические атаки»
- Многие модели пока что можно обманывать
- Е.г. написать на заборе, что за ним котик, а там лев
- Е.г. модифицировать параметры своих документов под скоринговую модель



When we put a label saying "library" on this Granny Smith apple, the model erroneously classifies it as a library in the zero-shot setting.



When we put a label saying "iPod" on this Granny Smith apple, the model erroneously classifies it as an iPod in the zero-shot setting.



When we put a label saying "toaster" on this Granny Smith apple, the model erroneously classifies it as a toaster in the zero-shot setting.

2. ПРОСТЫЕ МОДЕЛИ

- Поэтому у «простых» моделей, в том числе полиномиальных по переменным, есть пространство как минимум для «проверки качества»
- Пока есть привычка «посмотреть на коэффициенты», регрессии будут оставаться вторым стандартом, интерпретация линейных моделей проста
 - [Но ML всё легче интерпретировать](#) (Shapley values)
- Плюс ML требует иногда глубокой математики, иногда искусства настройки
 - [Но ML всё проще пользоваться](#), даже в Питоне
- Плюс очень хочется причинно-следственных связей
 - [В ML всё это есть](#), и качество улучшается с каждым месяцем

3. ПОЧЕМУ НУЖНА АККУРАТНОСТЬ-3

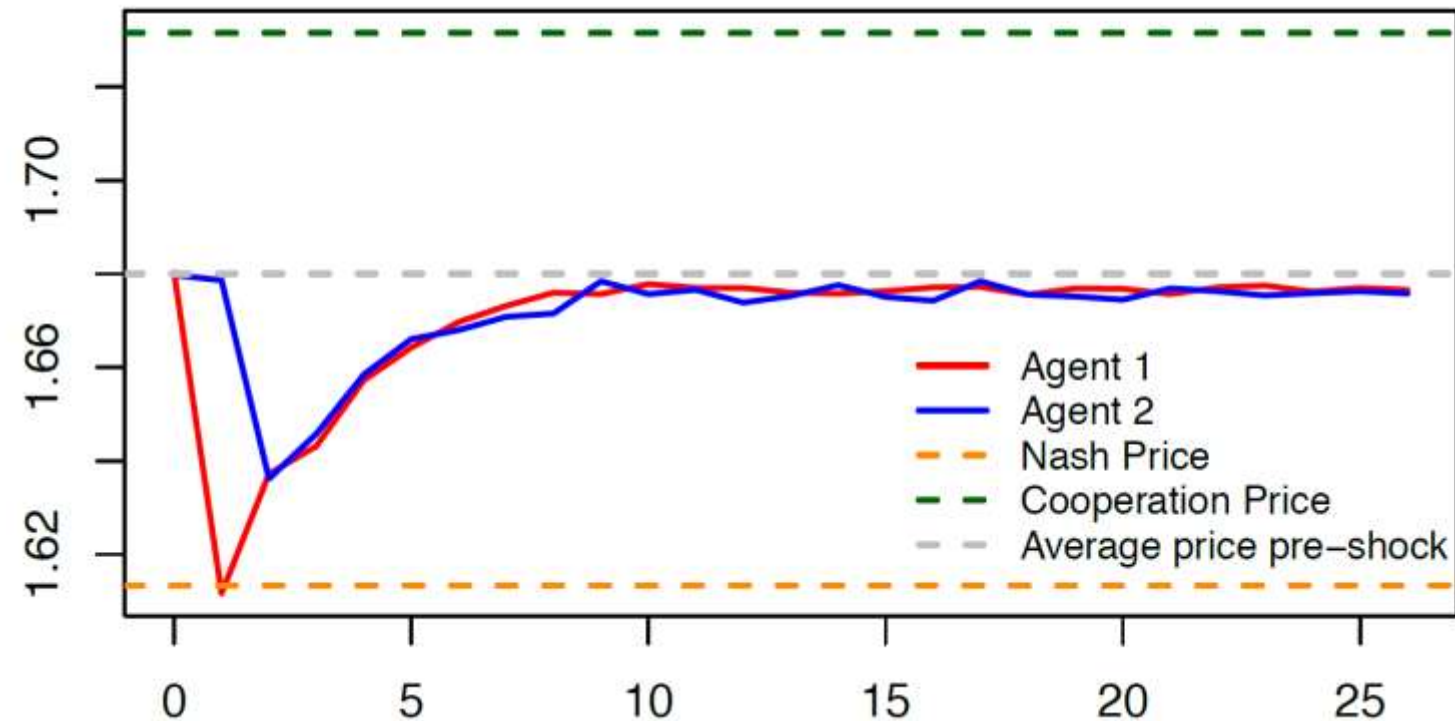
- [«Нечеловеческие картели»](#) (2018)
- Давайте запустим повторяющиеся аукционы, в которых участвуют алгоритмы на ML
- ML здесь «обучение с подкреплением», при этом с конечным состоянием мира и решениями, плюс конечной памятью
- Пусть несколько простых ИИ играют в «одновременно называем цены, выпуск определяется спросом» (примерно повторяющаяся олигополия Бертрана)

3. ПОЧЕМУ НУЖНА АККУРАТНОСТЬ-3

- Может появиться «картельный сговор», то есть ситуация, в которой цена отклоняется от себестоимости
- Как можно это обеспечить?
 - Возможностью наказывать за отклонение от картеля
 - Например, несколько периодов держать очень низкие цены, чтобы прибыль была никакой
- Делают ли это алгоритмы?

3. ПОЧЕМУ НУЖНА АККУРАТНОСТЬ-3

- Алгоритмы очень даже умеют договариваться (неявно)
- Если дать им пространство для торговли, размещения рекламы или подачи заявок на тендеры, может быть сговор
- Регуляторам придётся думать, как это доказывать (не бывает цен на уровне издержек)



4. ML ДЛЯ VC (VENTURE CAPITAL)

- ML может также выбирать [компании на ранних стадиях](#) (2020)
- Дадим модели потренироваться на десятках тысяч компаний и сделать прогнозы на несколько других
- Улучшение по сравнению с людьми – около 25%
- То есть в совмещении может быть положительный эффект (предварительный отбор по «характеристикам», более плотное изучение отобранных компаний)
- Но – всегда есть шанс, что важные характеристики будут меняться во времени

4. НО... НО...

Yahoo Refused to Buy Google

In 1998, Larry Page and Sergey Brin, the founders of Google, approached Yahoo to sell Google for \$1 million (\$1.5 million today). Yahoo refused the offer. That was a big mistake. But is that it? That's not even worth writing because Google rapidly became successful because of Sergey Brin and Larry Page. Had Yahoo bought it, Google probably would not have been what it is now. What makes Yahoo's mistakes so interesting is that they made many of them, not learning from them.

In 2002, Google was much more valuable. Now, Yahoo CEO, Terry Semel must have realized the true potential of Google. Yahoo approached Larry and Sergey to buy Google. Yahoo wanted to buy it for \$3 billion (\$4.32 billion today), but Google wanted \$5 billion (\$7.2 billion today). Yahoo refused the offer again. This was a huge mistake. On January 16, 2020, Google's parent, Alphabet, became the fourth biggest tech company to be worth \$1 trillion. They never got a chance to buy Google again.

...

Yahoo Refused to Buy Facebook in 2006

Yahoo tried to acquire Facebook in July 2006 for \$1.1 billion (\$1.4 billion today). Facebook agreed to the deal, and they were ready to be acquired. But wait, Facebook isn't owned by Yahoo. That's because Yahoo CEO, Terry Semel, changed his offer to \$800 million (\$1 billion today). Mark

Forbes стали известны подробности разногласий между «Сбером» и Ozon, которые помешали банку стать крупным акционером ретейлера в 2020 году. Банк оценивал всю компанию в \$1,4 млрд и настаивал на ее ребрендинге

SHARES



О том, что [Сбербанк](#) договаривается с Ozon [F 156](#) о покупке крупного пакета (до 30% после допэмиссии), в начале июня 2020 года сообщало агентство Reuters со ссылкой на два источника, переговоры подтверждал источник Forbes, близкий к акционерам Ozon. Но сделка в итоге не состоялась, причины этому до сих пор не были известны. Два источника, знакомые с ходом переговоров, рассказали Forbes, что одной из причин, почему стороны не договорились, стала оценка ретейлера для потенциальной сделки. «Сбер» был готов на сделку при цене в \$1,4 млрд за всю компанию, Ozon – при цене в \$2,1 млрд, рассказали Forbes два источника.

Закреть гештальт: как Лев Хасис строит партнерства ради торговых и платежных сервисов «Сбера»

Помимо этого Сбербанк настаивал на ребрендинге Ozon, сказал Forbes третий источник, близкий к сторонам переговоров. Также под угрозой могли оказаться финтех-сервисы ретейлера, на которые Ozon делает большую ставку, добавил он.

В итоге Ozon принял решение идти на IPO на Nasdaq. Ozon вышел на биржу в ноябре 2020 года и привлек \$1,2 млрд. Всю компанию на IPO инвесторы [оценили](#) в \$6,2 млрд, то есть более чем в четыре раза выше, чем Сбербанк.

5. ЧТО ДАЛЬШЕ?

- Коллега-выпускник РЭШ:
- «А вообще этот ML в экономических статьях, это почти всё ML 1980х, который и младенцы знают. Вообще не ML, а эконометрика на стероидах. Вот в quant marketing, вот там ML»
- Академические финансы скорее отстают от того, что делает индустрия
- И это не то чтобы нестандартно: в [Computer Science](#) (2019) выглядит так, что индустрия влияет на дальнейшие исследования в науке сильнее, чем наоборот
- Возможно, скоро мы узнаем много нового про риски на финансовых рынках и предсказуемость цен