

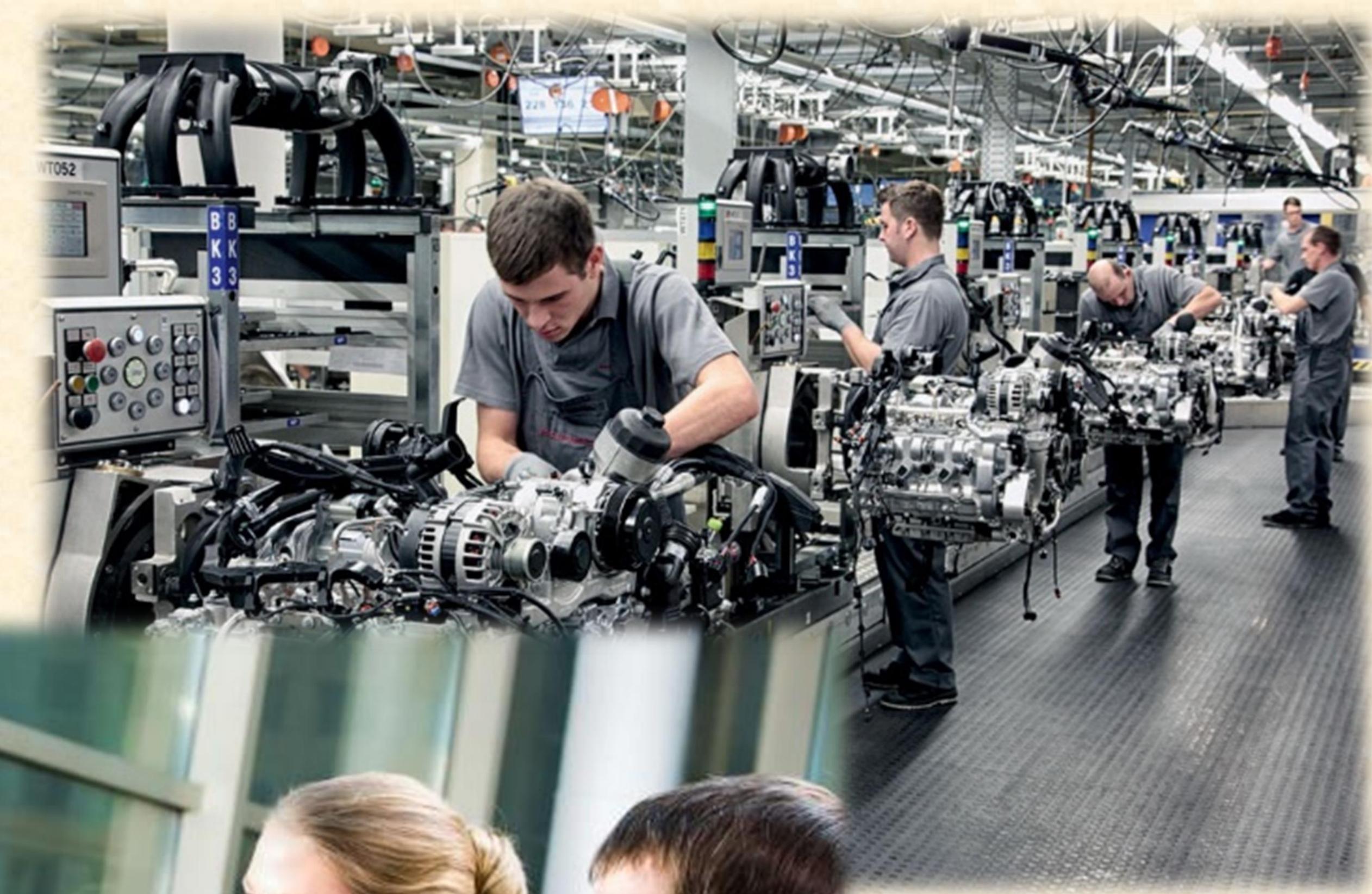
Яндекс

Яндекс

Образование в сфере анализа данных

Александр Авдюшенко
куратор академических программ







«Из будущего можно убежать только в одно место – в прошлое»

Машинное обучение

В основе – математика: численные методы,
методы оптимизации, теория вероятностей...

Нет предзаданной модели с уравнениями.

Машинное обучение

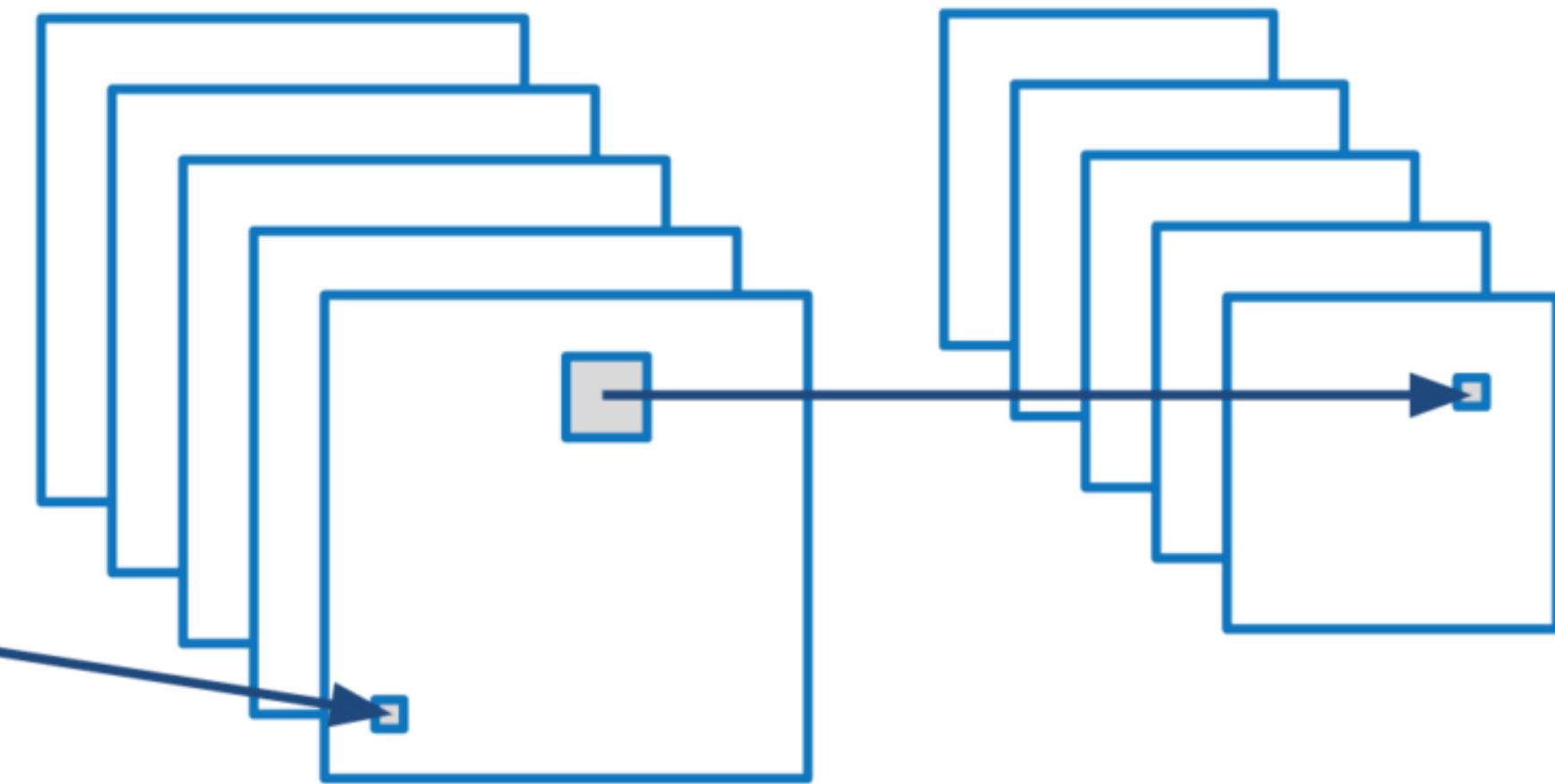
В основе – математика: численные методы, методы оптимизации, теория вероятностей...

Нет предзаданной модели с уравнениями.

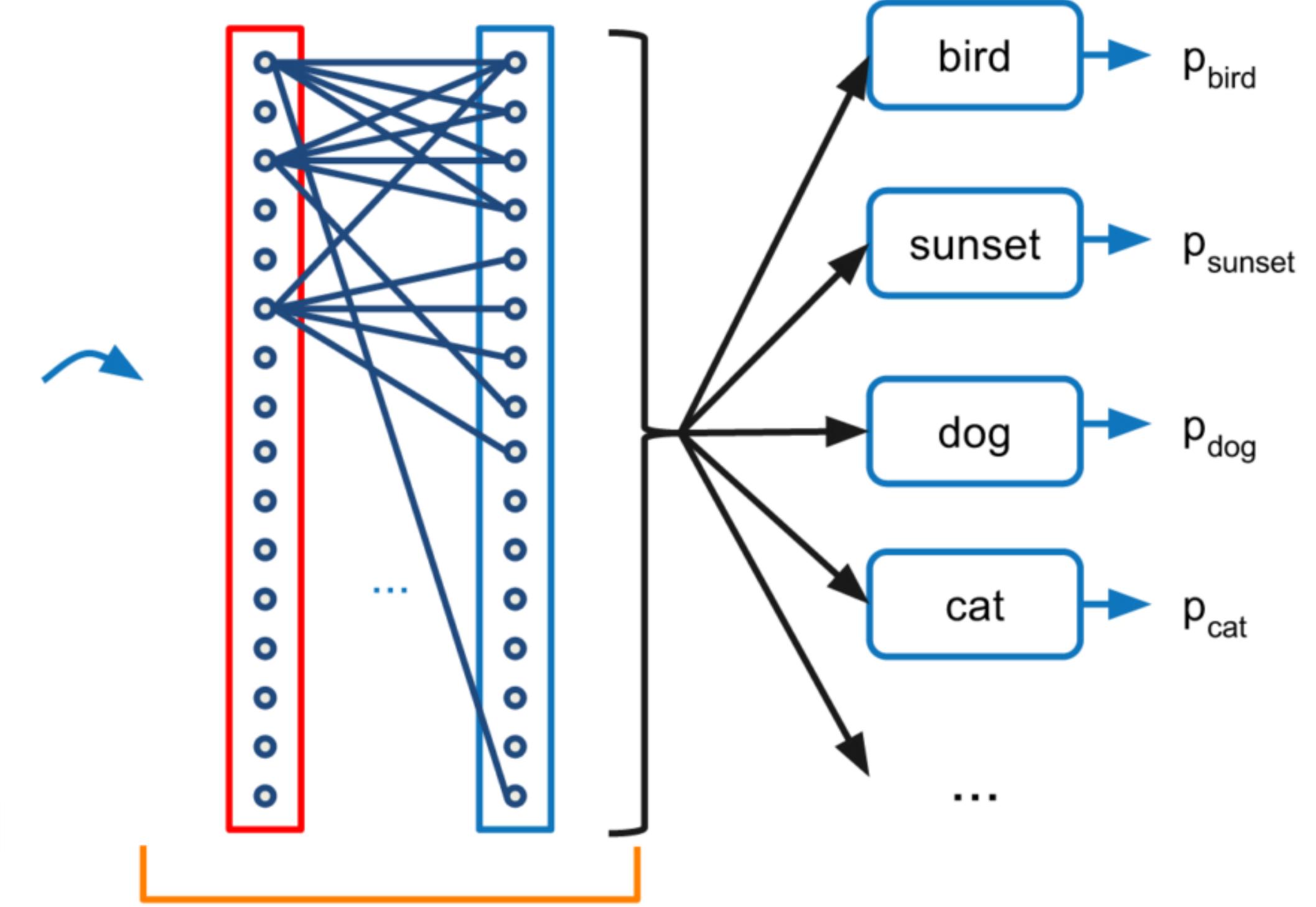




свёртка +
нелинейность



несколько таких блоков



полносвязные
слои

классификация
на N классов

Примеры хорошо решённых задач





Oom
Alden
Eastman
Alfred
Arthur
Luke
Fletcher
Matthew
Theodore
Richard
Shirmer
and
Hisswald

Чтение по губам

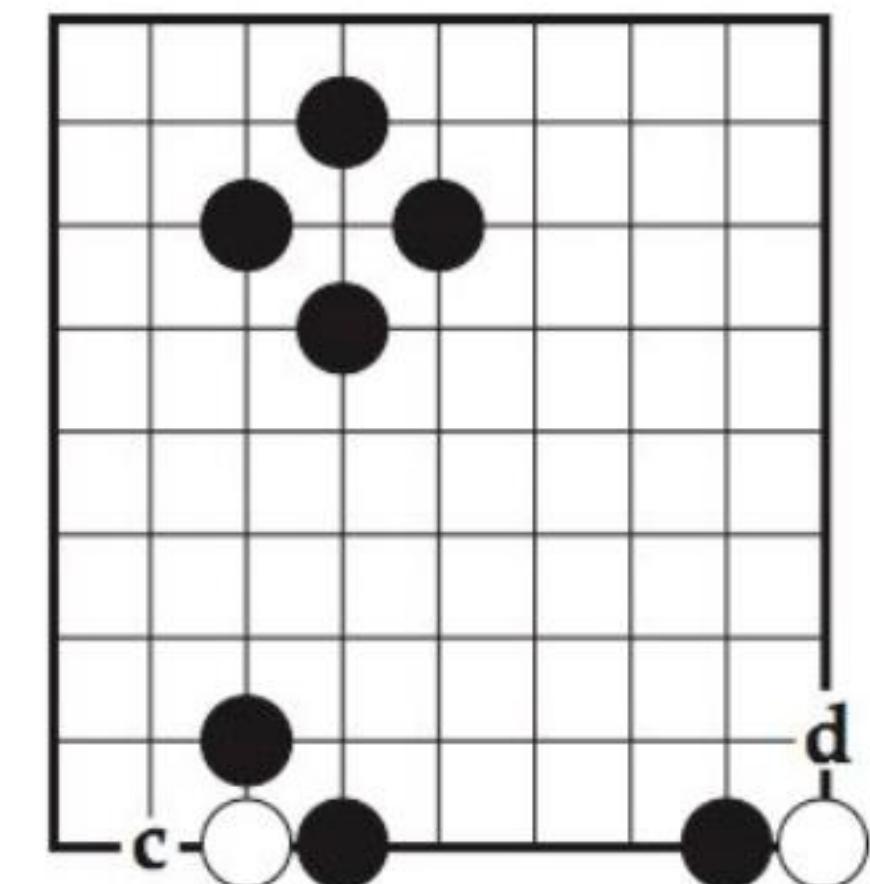
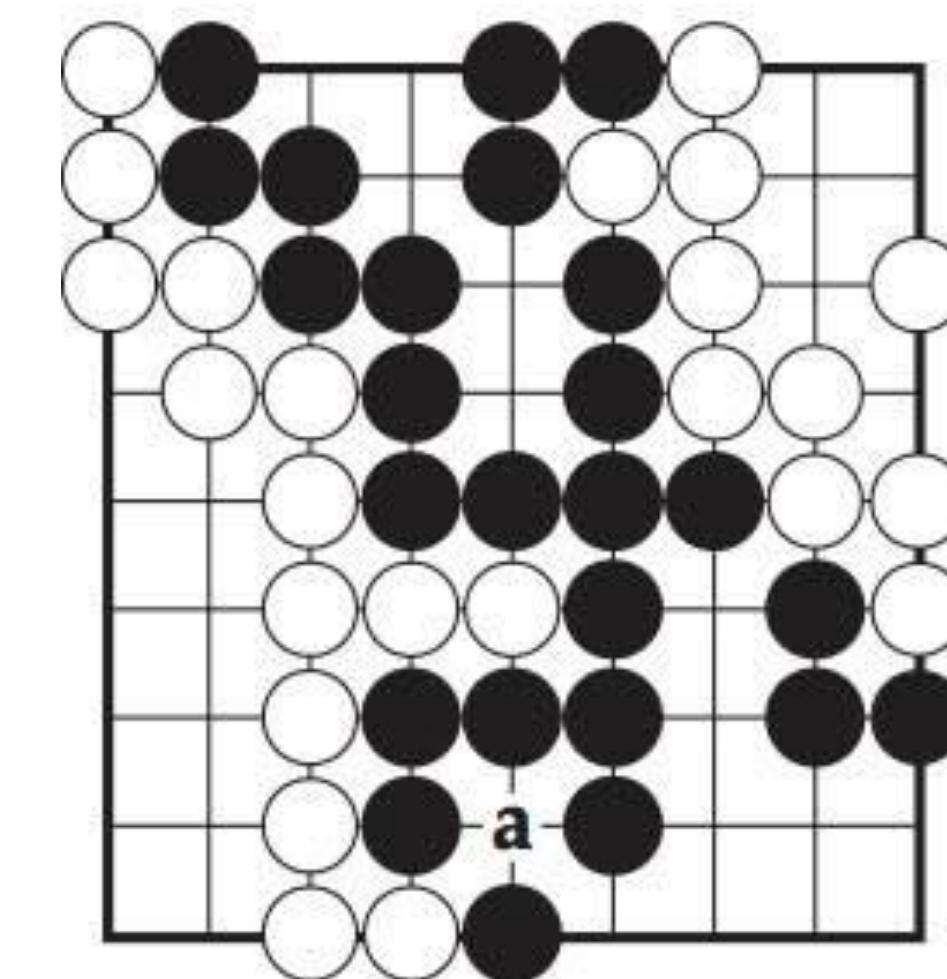
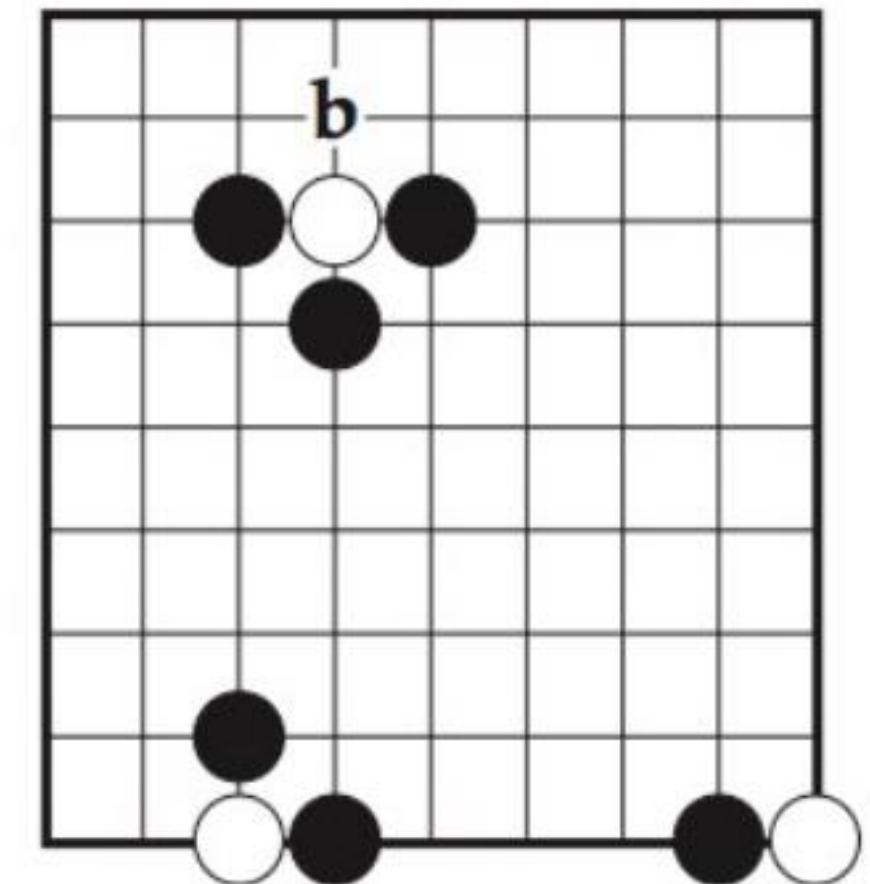
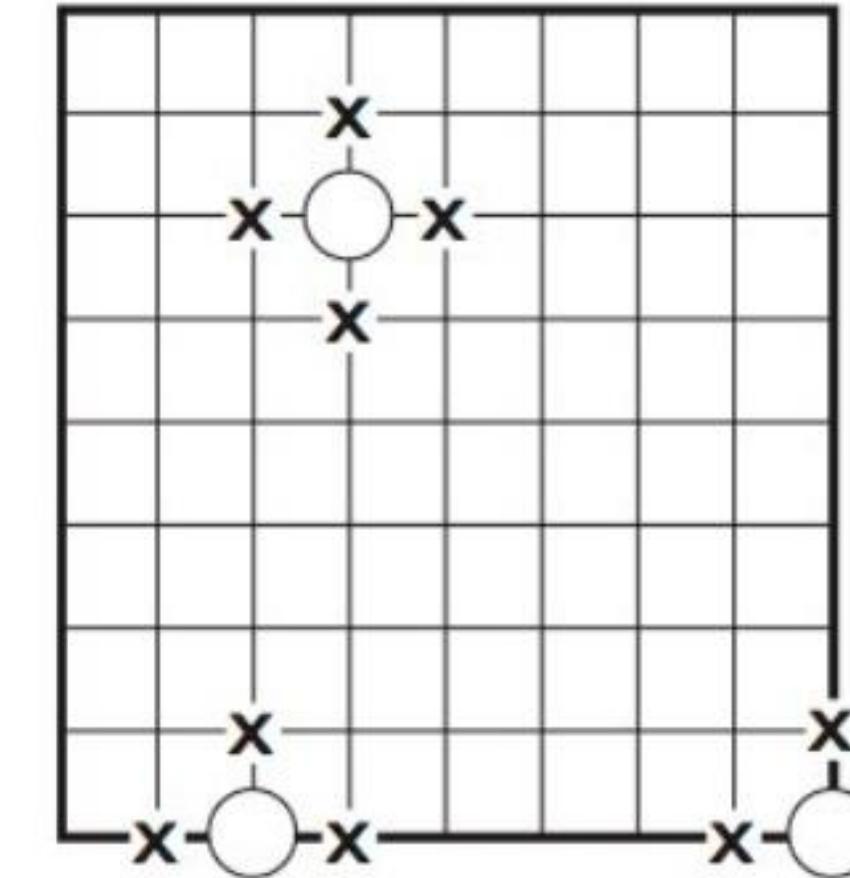


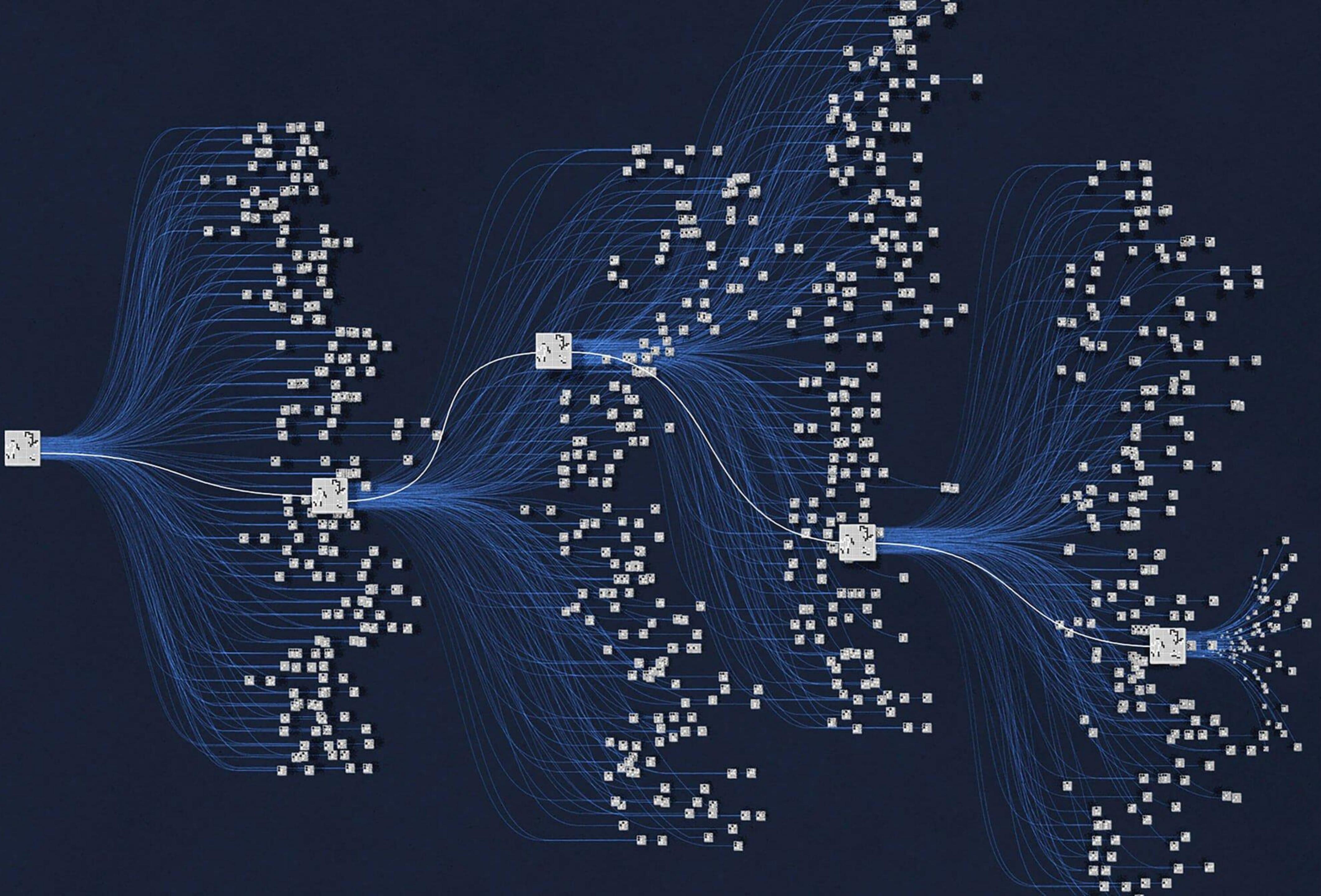
Ноябрь 2016, University of Oxford & Google DeepMind

Доля правильных слов	%
Профессиональный чтец по губам	26.2
Модель* (LSTM + CNN)	46.8

* <https://arxiv.org/pdf/1611.05358v1.pdf>

Игра в го





AlphaGo vs Lee Sedol



Март 2016

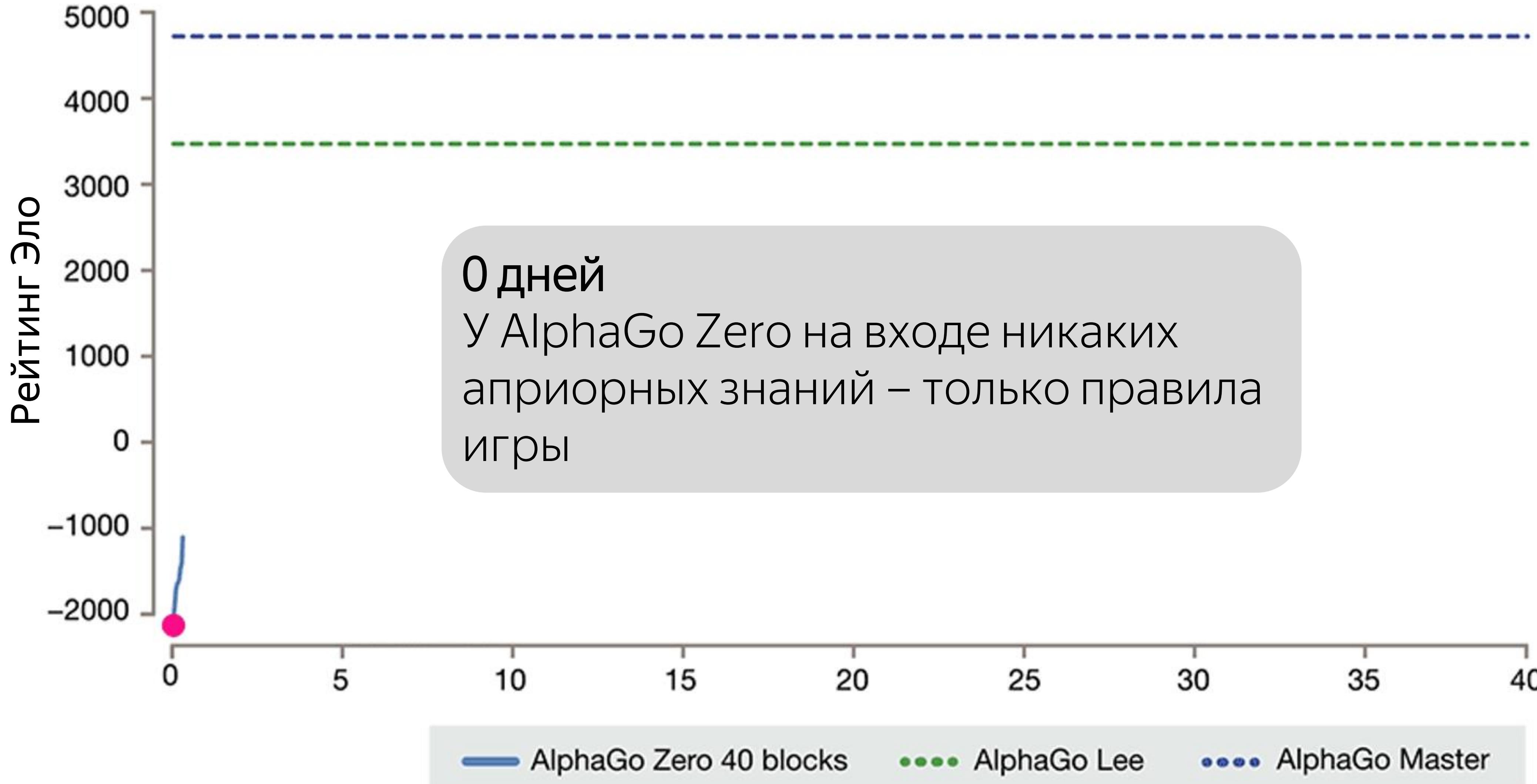
AlphaGo vs Lee Sedol

4 : 1

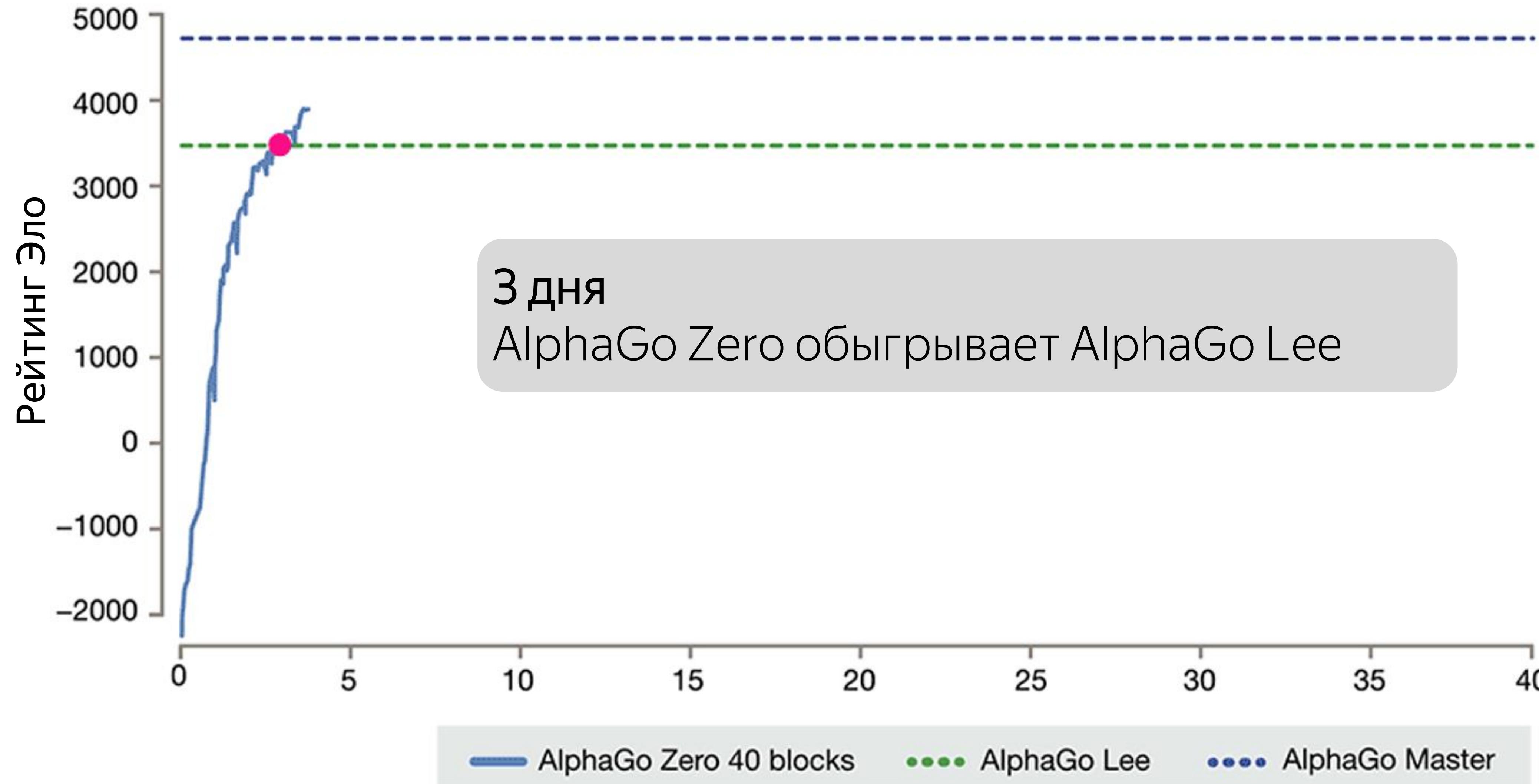


Март 2016

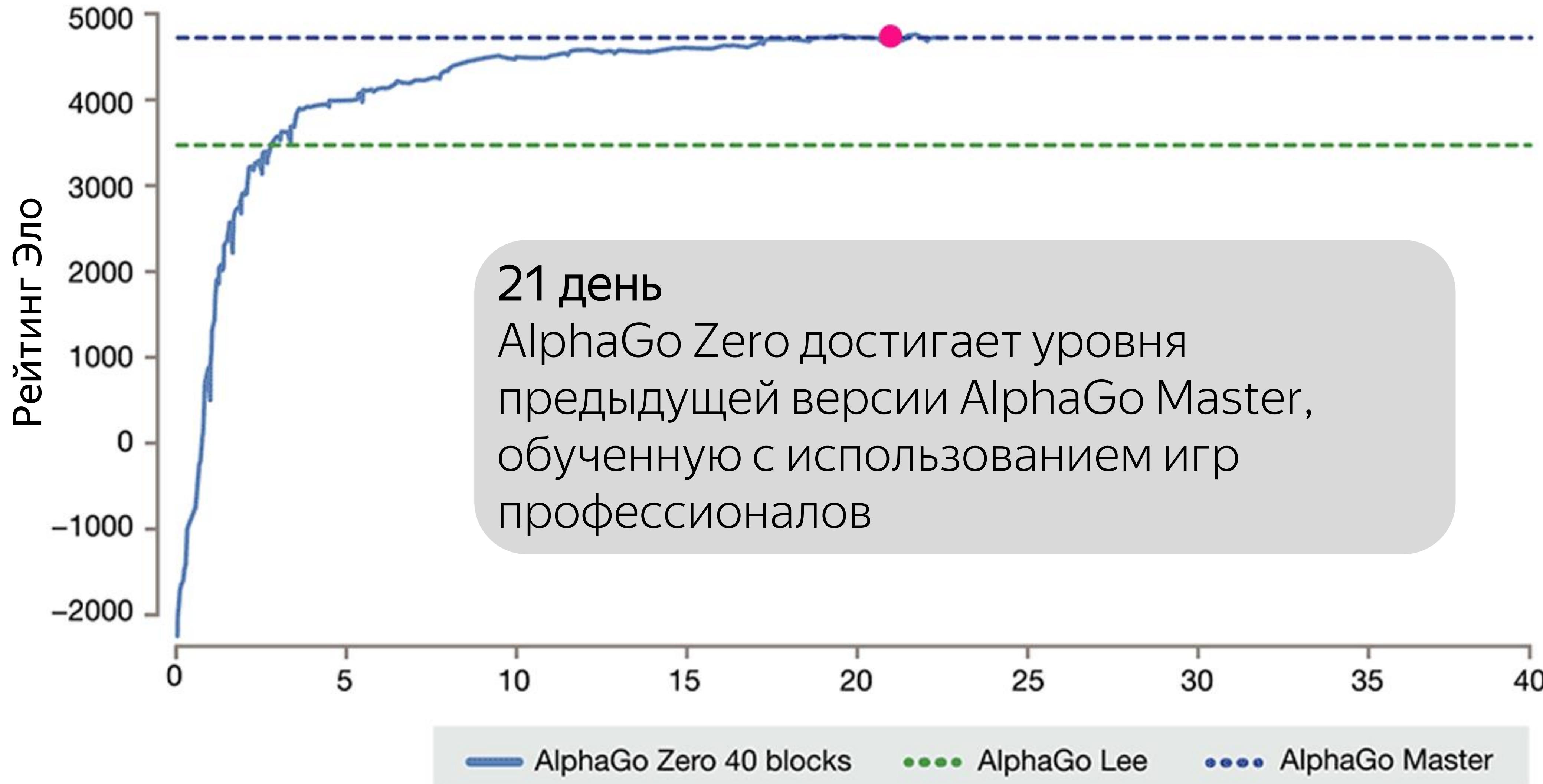
AlphaGo Zero



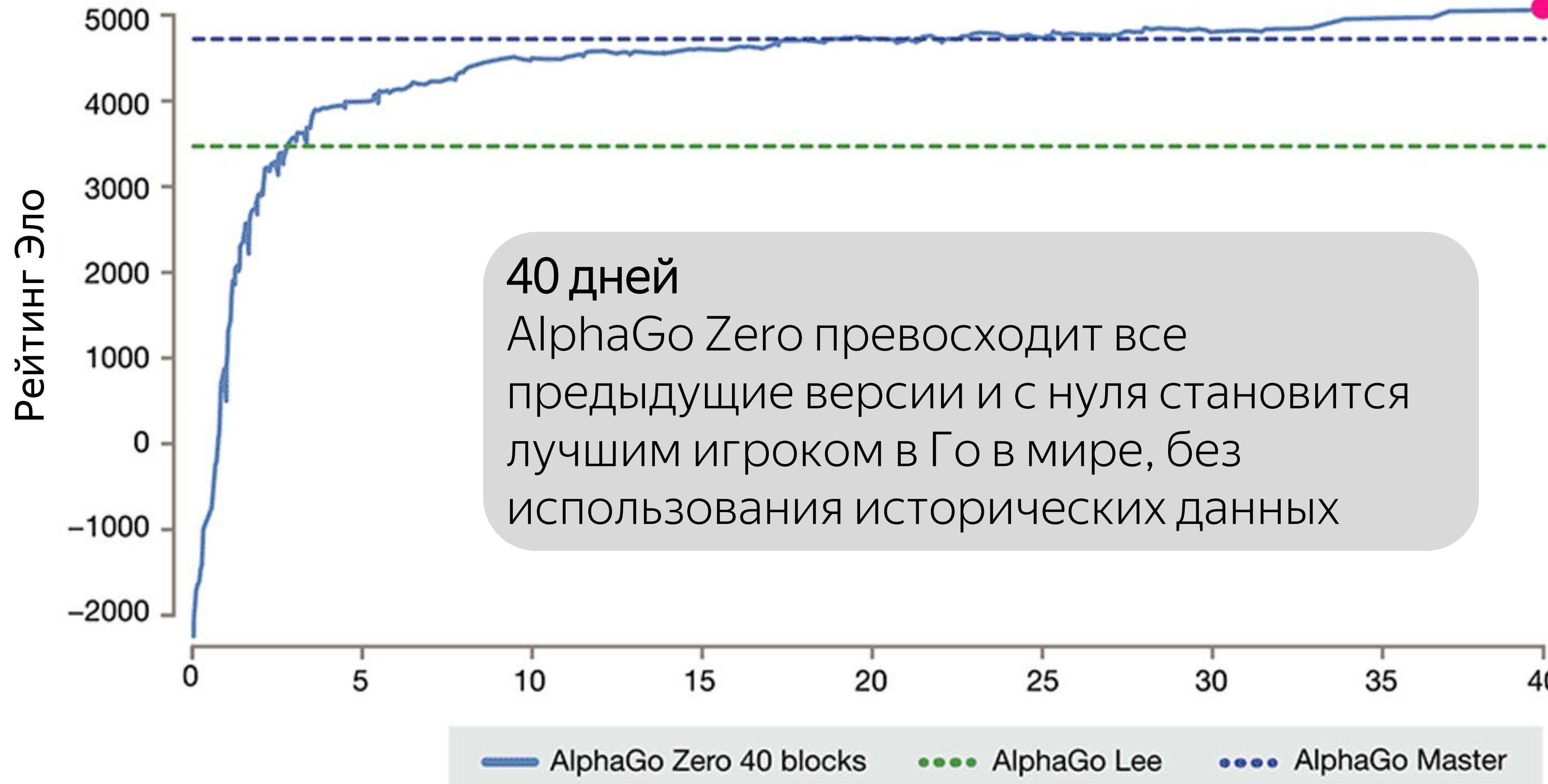
AlphaGo Zero



AlphaGo Zero



AlphaGo Zero



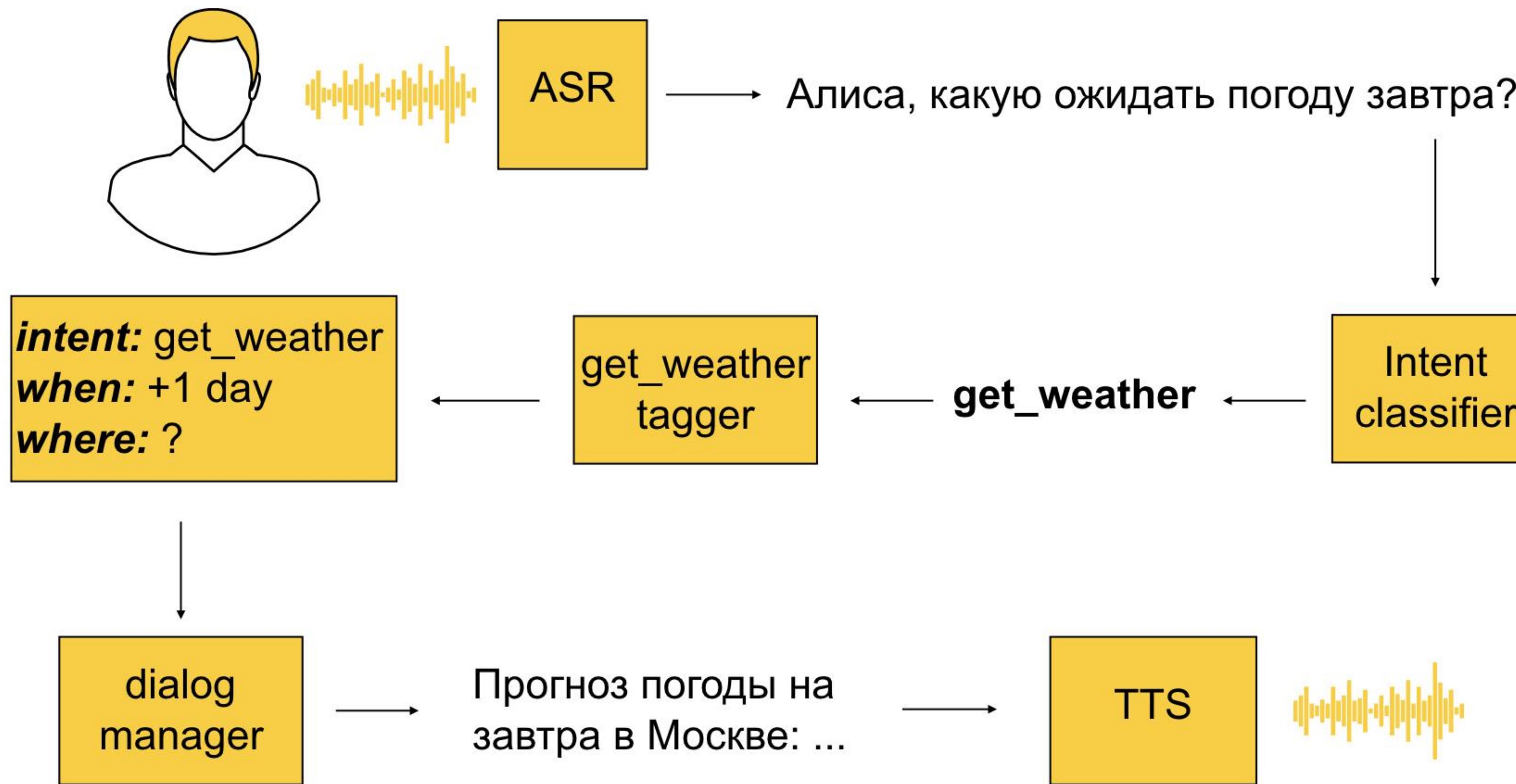


Привет, я Алиса!

Я готова помочь

- Определить песню > Узнать, что на фото > Включить сказку >
- Одеться по погоде > Поиграть > Построить маршрут >
- Вызвать такси > Найти нужное место > Управлять компьютером >

Алиса, вид сверху



Влияние интернета на образование



О себе



2003
медаль

2007 2009
бакалавр магистр



2014
к.ф.-м.н.



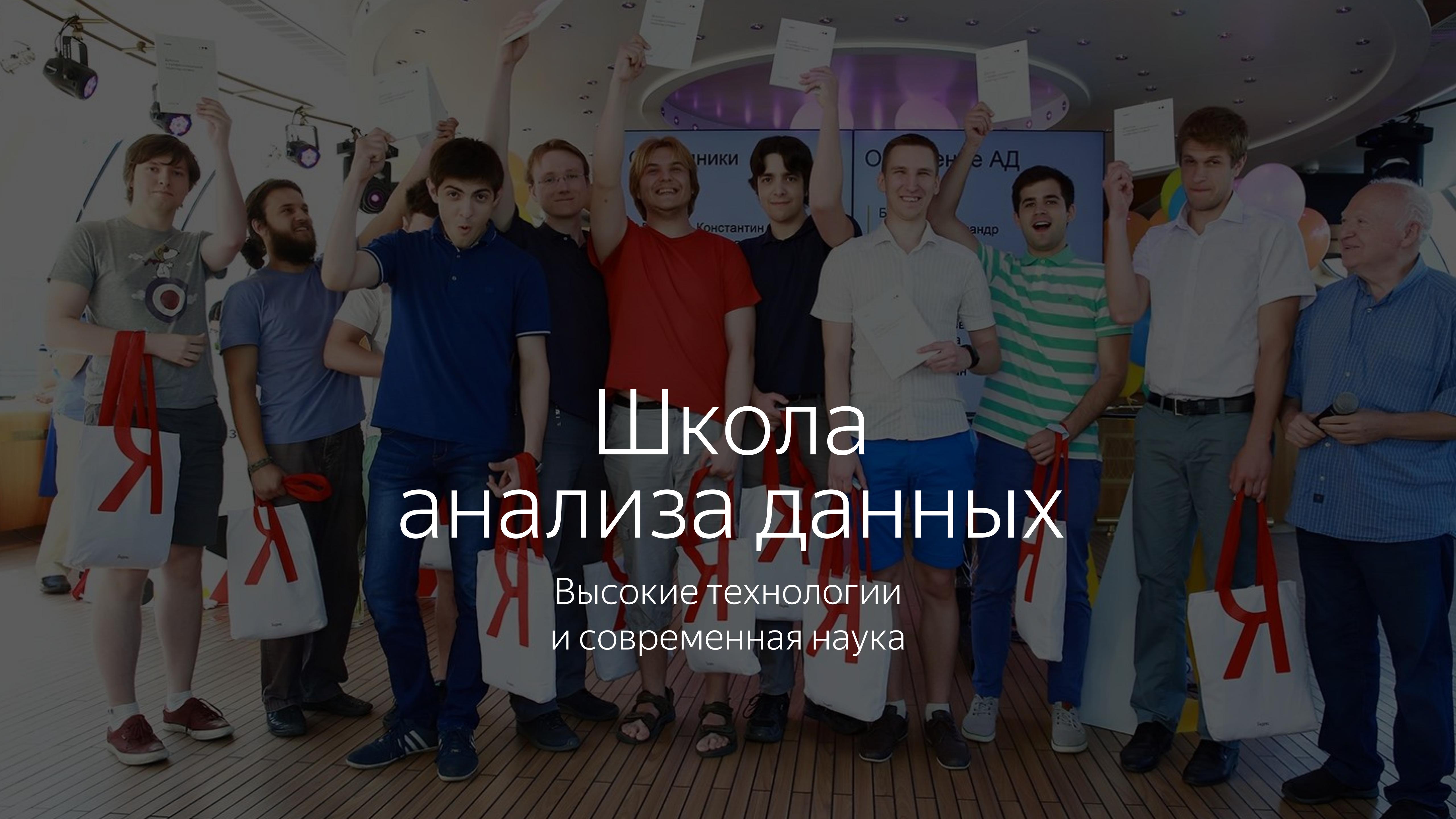
2016
ШАД



2015 - 2018
аналитик



2016 - ...
куратор

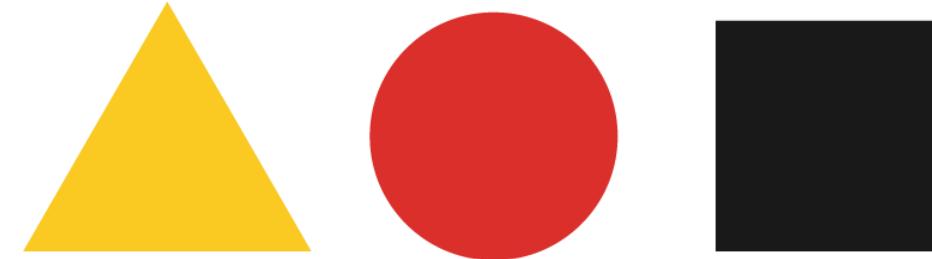


Школа анализа данных

Высокие технологии
и современная наука

CS центр

Совместный проект Школы анализа данных Яндекса и JetBrains

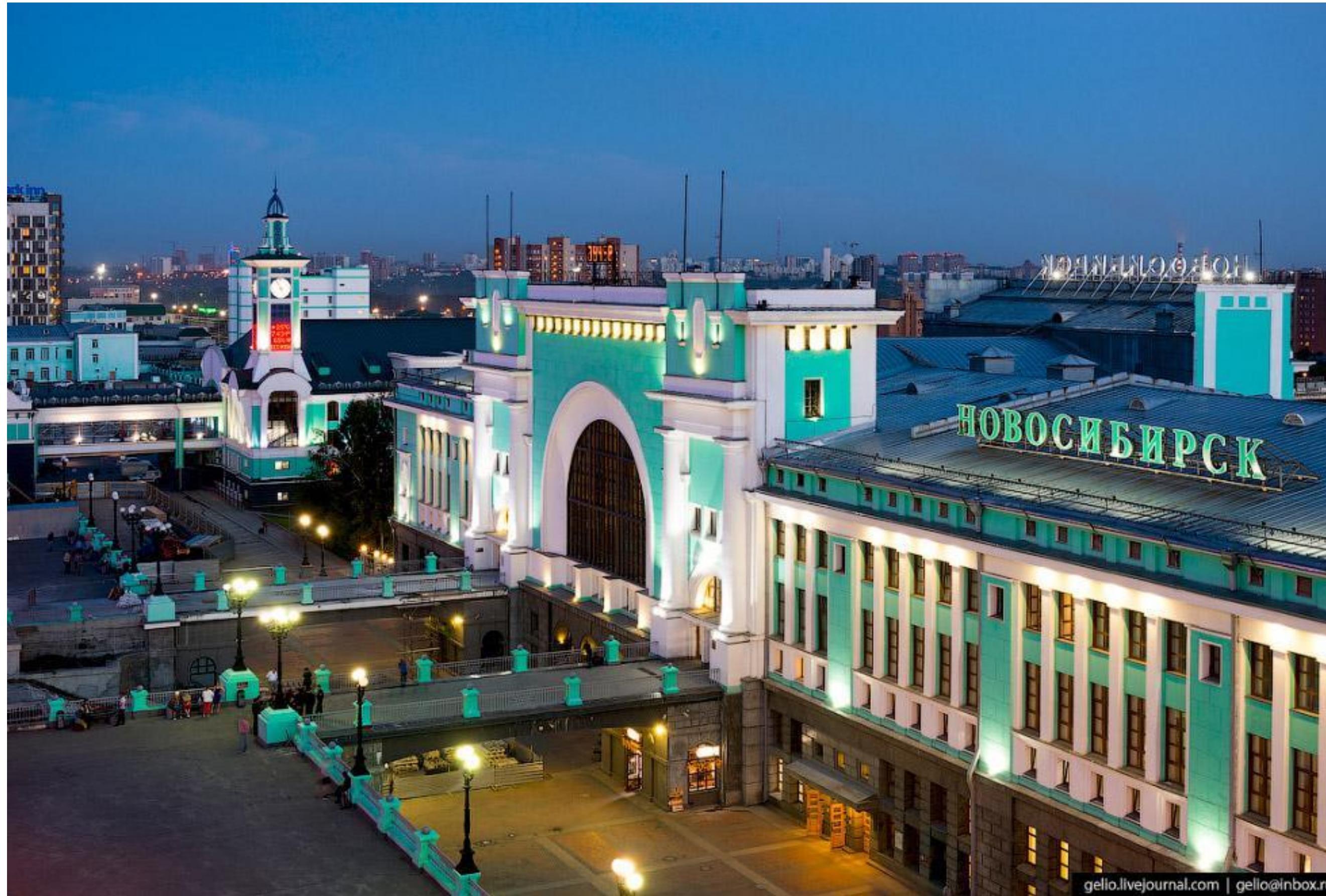


ШКОЛА АНАЛИЗА ДАННЫХ



- Курсы по двух- или трёхлетней бесплатной очной программе
 - *Анализ данных*
 - *Разработка ПО*
- Практики от сотрудников ИТ-компаний и институтов СО РАН

Очные занятия



По будням с 18:10 и по субботам



Заочное обучение

Платное заочное обучение длительностью один год

Базовые курсы в первом семестре,
специализация и практика во втором

Базовые курсы направлений

Анализ данных

Разработка ПО

Алгоритмы и структуры данных, части 1 и 2

Математика: дискретный анализ/теория вероятностей/линейная алгебра

Программирование: C++/Java/Python/Kotlin

Машинное обучение, часть 1

Функциональное программирование

Машинное обучение, часть 2

Языки программирования и компиляторы

Спецкурсы

- Основы проектирования приложений
- Программирование под Android
- Сетевые технологии
- Программирование на Rust
- Теория игр
- Введение в лингвистику
- Теория информации
- Параллельное программирование
- Компьютерное зрение
- Deep Learning на пальцах
- ...



Не менее 12 курсов

Практики



Сопоставление изображений одежды для поиска похожих

Практики

```
for (i in 0..10)
{
    println(i)
}
```



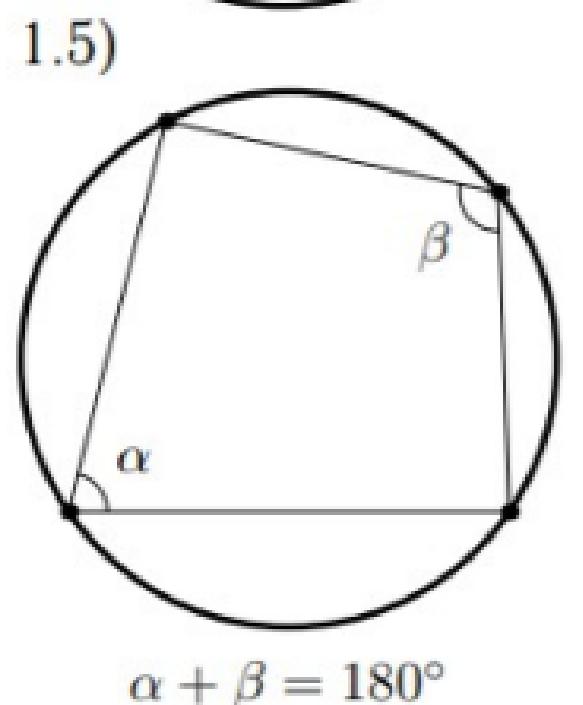
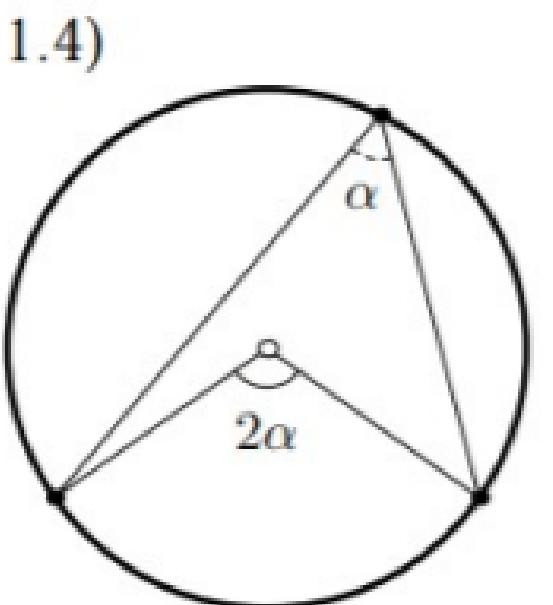
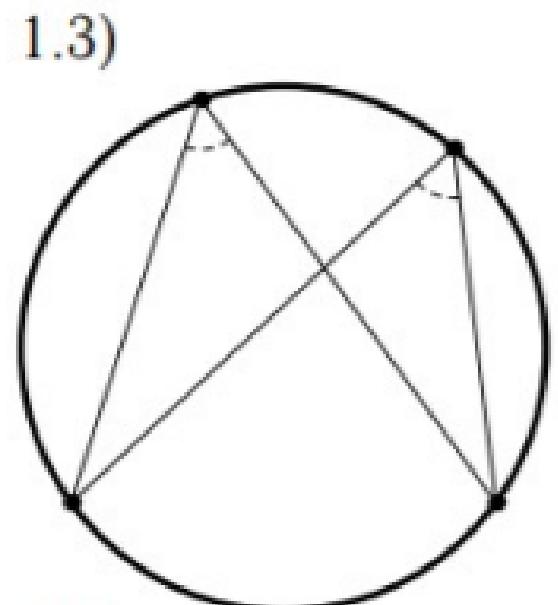
```
val range = IntRange(0, 10)
val iterator =
range.iterator()
while (iterator.hasNext()) {
    val i = iterator.next()
    println(i)
}
```



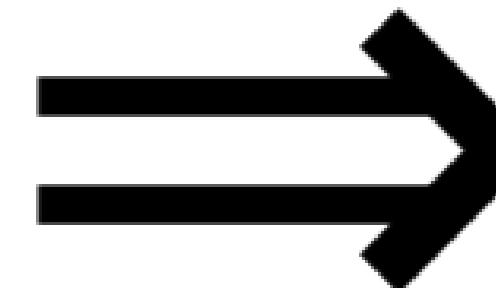
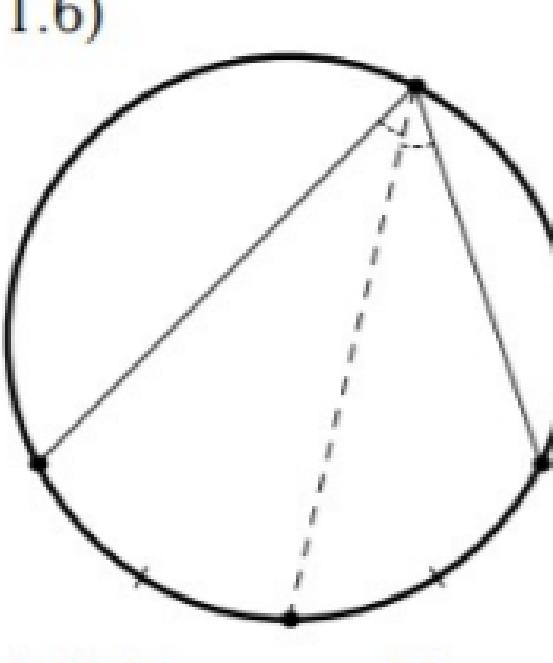
```
var i = 0
// + some checks
while (i <= 10) {
    println(i)
    i++
}
```

Оптимизация циклов `for` и операторов `in` в
компиляторе Kotlin/Native

Практики



$$\alpha + \beta = 180^\circ$$



ЛАТЭКС

```
\begin{tikzpicture}
    \draw (62.718, 51.188) circle (37.203);
    \draw (29.838, 26.303) circle (19.146);
    \draw (65.765, 55.720) -- (72.861, 78.025);
    \draw (18.896, 50.798) -- (35.069, 6.104);
\end{tikzpicture}
```

Конвертер чертежей в код

Онлайн-курсы

Погружение в СУБД

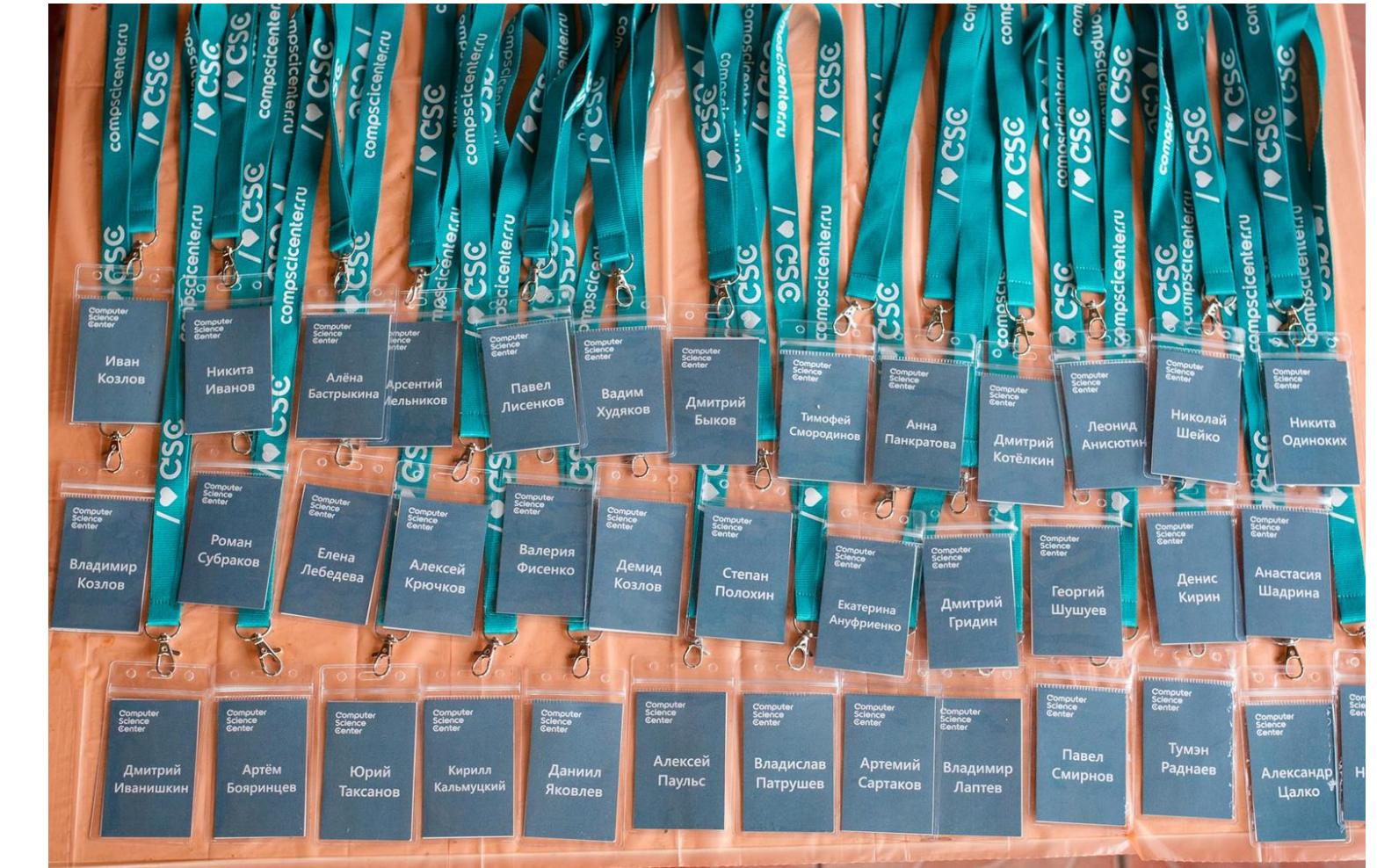
знатокам будет неинтересно

```
SELECT table_name, TRUNC(sum(bytes)/1024/1024) Meg
FROM
(SELECT segment_name table_name, bytes FROM dba_segments
WHERE segment_type = 'TABLE' UNION ALL
SELECT i.table_name, s.bytes
FROM dba_indexes i, dba_segments s
WHERE s.segment_name = i.index_name
AND s.segment_type = 'INDEX' UNION ALL
SELECT l.table_name, s.bytes
FROM dba_lobs l, dba_segments s
WHERE s.segment_name = l.segment_name
AND s.segment_type = 'LOBSEGMENT' UNION ALL
SELECT l.table_name, s.bytes
FROM dba_lobs l, dba_segments s
WHERE s.segment_name = l.index_name
AND s.segment_type = 'LOBINDEX'
GROUP BY table_name
/* Ignore really small tables */
HAVING SUM(bytes)/1024/1024 > 1000
```



▶ 🔍 -0:31 1.25x 1080px ::

Свободное время



Подробнее о нас

- На сайте compscicenter.ru
- В соц. сетях CS центра
 - vk.com/compscicenter
 - facebook.com/CSCenter
 - t.me/compscicenter_ru
 - instagram.com/compscicenter/
- Видео лекций на YouTube
 - youtube.com/c/compscicenterru
- Пишите на почту
 - info@compscicenter.ru

The background of the image features a complex, abstract network structure composed of numerous small, semi-transparent red and pink triangles. These triangles are interconnected by a web of thin red lines, creating a sense of depth and connectivity. The overall effect is reminiscent of a molecular model or a complex data visualization.

Open Data Science

Open Data Science (ODS.ai)

- › Открытое сообщество в Slack: 30К участников, 4М сообщений
- › 23 часовых пояса, 150+ публичных каналов
- › 130+ data-завтраков, 50+ встреч, 30+ ML-тренировок
- › Мотивирует
- › Обучает:

2017 – MO-1 <https://habr.com/company/ods/blog/322626/>

2018 – ML-1 <https://mlcourse.ai>

2019 – «DeepLearning на пальцах» <https://dlcourse.ai>

The Home of Data Science & Machine Learning

Kaggle helps you learn, work, and play

[Create an account](#)

or

[Host a competition](#)

jobs
board >



Competitions

Climb the world's most elite machine learning leaderboards

Datasets

Explore and analyze a collection of high quality public datasets

Kernels

Run code in the cloud and receive community feedback on your work

ODS + kaggle + 1-2 года = junior data scientist

The Home of Data Science & Machine Learning

Kaggle helps you learn, work, and play

[Create an account](#)

or

[Host a competition](#)

jobs
board >



Competitions

Climb the world's most elite
machine learning
leaderboards

Datasets

Explore and analyze a
collection of high quality
public datasets

Kernels

Run code in the cloud and
receive community feedback
on your work

Курс «DeepLearning на пальцах» в НГУ



Нейросети

2012-2015

» 2018

ML-тренировки в НГУ

- › запустились с октября
- › за осень три серебра на Kaggle и третье призовое место в конкурсе от ЦФТ
- › за весну два серебра и пять бронз, третье место в Changellenge >> Cup
- › получили доступ к кластеру НГУ: 40 CPU, 755Gb RAM, **8 GPU** NVIDIA Tesla V100

Что же делать?

Учиться («Программирование – вторая грамотность»*):

- › CS центр <https://compscicenter.ru>
- › ШАД <https://yandexdataschool.ru>
- › Академия Яндекса <https://academy.yandex.ru>
- › Stepik <https://welcome.stepik.org/ru>
- › Deep Learning «на пальцах» <https://dlcourse.ai>
- › Курсера <https://ru.coursera.org>

Онлайн-образование и интернет
не заменят живого общения.

* http://ershov.iis.nsk.su/ru/second_literacy/article '1980



Что же делать?

Читать книги:

- › Вы, конечно, шутите, мистер Фейнман!
<http://www.ozon.ru/context/detail/id/26100045/>
- › Саймон Сингх, «Великая теорема Ферма».
http://modernlib.ru/books/singh_saymon/velikaya_teorema_ferma/read
- › Роберт Грин, «Мастер игры».

Ищите учителей!

Ссылка на эту презентацию <https://bit.ly/2VL3BTx>

Александр Авдюшенко



avalur@yandex-team.ru



@ovalur

