

Как прогнозировать вероятности важных событий?



*Марат Салихов
22 марта 2023 года*

Пара слов обо мне



- Asst. Prof. операций и data science
 - МШУ Сколково и РЭШ
- PhD INSEAD'19, постдок Yale'21
- Преподаю: бизнес-аналитику в Совместном Бакалавриате ВШЭ-РЭШ
- Исследую: экспертное прогнозирование вероятностей событий



Принятие решений в условиях неопределенности

Нужен ли будет в пятницу зонт?



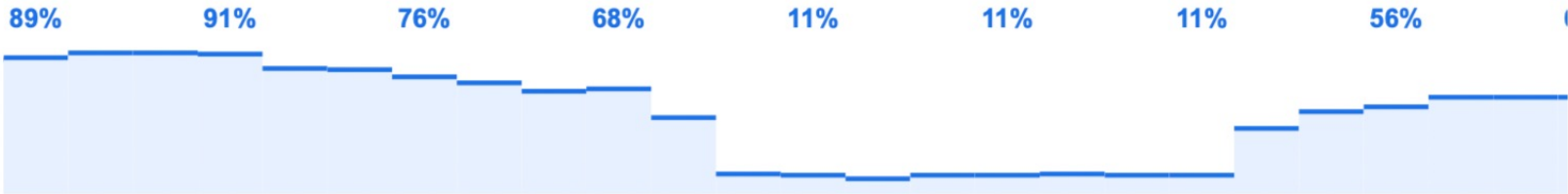
9 °C | °F

Вероятность осадков: 80%
Влажность: 73%
Ветер: 7 м/с

Погода

пятница
Дожди

Температура | Вероятность осадков | Ветер



02:00

05:00

08:00

11:00

14:00

17:00

20:00

23:00

ПН

ВТ

СР

ЧТ

ПТ

СБ

ВС

ПН



6° -1°



7° 2°



6° 3°



8° 6°



9° 3°



11° 3°



8° 2°



8° 0°

Принятие решений в условиях неопределенности

Нужен ли будет в пятницу зонт?

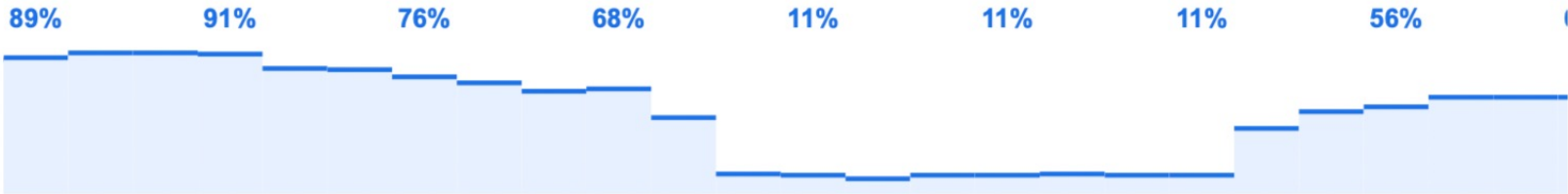


9 °C | °F

Вероятность осадков: 80%
Влажность: 73%
Ветер: 7 м/с

Погода
пятница
Дожди

Температура | Вероятность осадков | Ветер



02:00 05:00 08:00 11:00 14:00 17:00 20:00 23:00


ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС	ПН
6° -1°	7° 2°	6° 3°	8° 6°	9° 3°	11° 3°	8° 2°	8° 0°

Варианты

	4	-1
	1	2

Принятие решений в условиях неопределенности

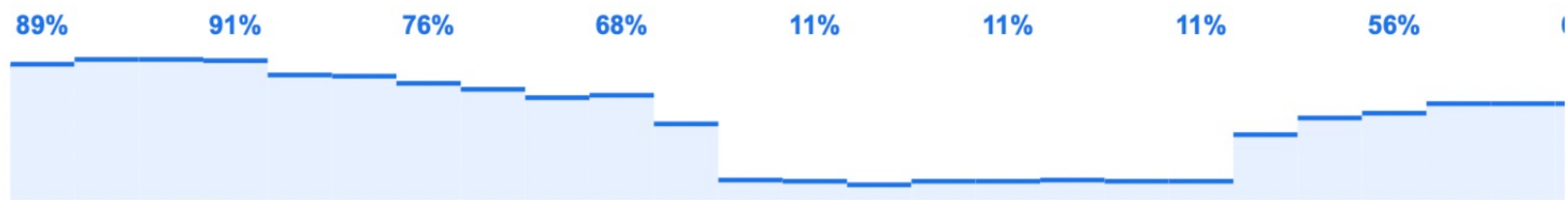
Нужен ли будет в пятницу зонт?


9 °C | °F

 Вероятность осадков: 80%
 Влажность: 73%
 Ветер: 7 м/с

Погода
пятница
Дожди

Температура | Вероятность осадков | Ветер



02:00 05:00 08:00 11:00 14:00 17:00 20:00 23:00

ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС	ПН
6° -1°	7° 2°	6° 3°	8° 6°	9° 3°	11° 3°	8° 2°	8° 0°

weather.com

Варианты

	4	-1
	1	2

Вероятность дождя p

Взять зонт, если:

$$4(1-p) - p < 1 - p + 2p$$

То есть если $p > 50\%$

Зонт не нужен с 11-30 до 19-30

Бизнес-решения

Скоринг кредитов

Вероятность дефолта

Цепочки поставок

Вероятность нарушения

Новые продукты

Вероятность неудачи

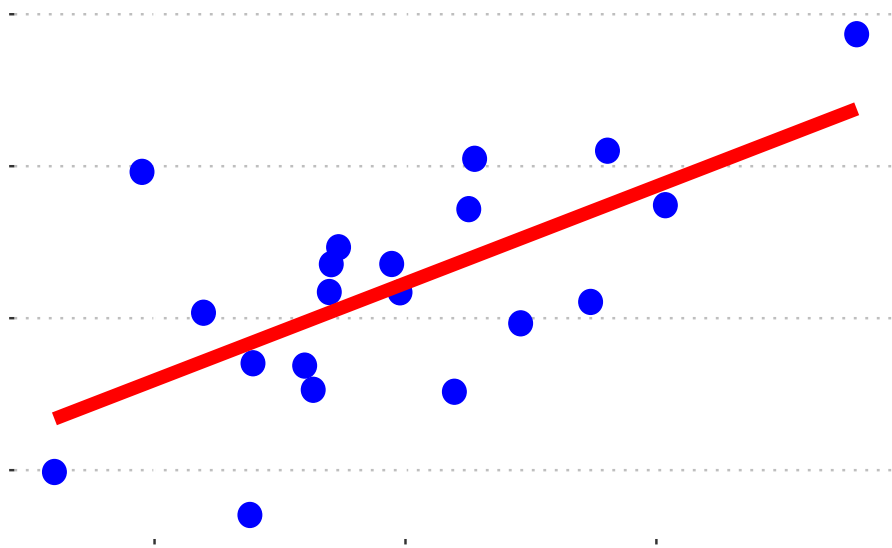


Чтобы сделать рациональное решение, необходимо:

1. Оценить издержки и выгоды решения при каждом варианте
2. И вероятности каждого варианта







Как вычисляются прогнозы

Статистические данные



- Примеры: регрессия, временные ряды
- Предсказание стабильных процессов
- Примеры
 - Спрос на «черные носки»
 - Прогноз погоды
- Искусственный интеллект

Субъективные прогнозы

			
	30	150	80
	90	40	50
	300	80	170

- Пример: экспертные панели
- Когда данных мало
- Например
 - Спрос на новые продукты
 - **Политические прогнозы**

ИИ предсказывает политику (июль 2022)

Исследователи в области ИИ предложили алгоритм, основанный на GPT-подобных моделях и использующий большой датасет новостей

<https://arxiv.org/pdf/2206.15474.pdf>

Forecasting Future World Events with Neural Networks

Andy Zou
UC Berkeley

Tristan Xiao
UC Berkeley

Ryan Jia
UC Berkeley

Joe Kwon
MIT

Mantas Mazeika
UIUC

Richard Li
UC Berkeley

Dawn Song
UC Berkeley

Jacob Steinhardt
UC Berkeley

Owain Evans
University of Oxford

Dan Hendrycks
UC Berkeley

Результаты

Точность: **65%** у их алгоритма и **92%** у экспертов-людей
(но у каких экспертов?)

Политическое прогнозирование

Произойдет ли вооруженный конфликт между Индией и Пакистаном в 2023 году?

По мнению профессора Белокреницкого, предпринятые шаги могут свидетельствовать о положительной динамике развития ситуации в регионе: «Отношения между Индией и Пакистаном напоминают маятник: он сдвигается то в сторону улучшения, то в сторону ухудшения, а иногда замирает на средней позиции. И вот сейчас они движутся от очень плохих отношений к более нормальным. Это не значит, что будет достигнут полноценный мир, но какие-то подвижки **можно ожидать**».

Многие прогнозы формулируются вербально:

“можно ожидать”, “возможно”, “почти неизбежно”

Невозможно превратить в точные вероятности: сложно принимать решения

Можем ли мы получить вероятности?

Тетлок: нужно вести счет



*Профессор
факультета психологии
Пенсильванского
Университета (США)*

Экспертиза:
Геополитическое
прогнозирование

Тетлок: нужно вести счет



*Профессор
факультета психологии
Пенсильванского
Университета (США)*

Экспертиза:
Геополитическое
прогнозирование

Как измерить
качество прогноза?

**Метрика Брайера
(Brier score)**

Квадрат разности между
исходом и прогнозом

Лучшее значение: 0

Худшее значение: 1

Значение для $p = 0.5$: 0.25

Тетлок: нужно вести счет



*Профессор
факультета психологии
Пенсильванского
Университета (США)*

Экспертиза:
Геополитическое
прогнозирование

Как измерить
качество прогноза?
**Метрика Брайера
(Brier score)**
Квадрат разности между
исходом и прогнозом

Лучшее значение: 0
Худшее значение: 1
Значение для $p = 0.5$: 0.25

	Иван	Исход	МБ
Сб. 1	1.0	1	0.00
Сб. 2	0.9	1	0.01
Сб. 3	0.5	0	0.25
Сб. 4	0.7	0	0.49
Сб. 5	0.0	1	1.0
СМБ			0.35

Тетлок: нужно вести счет



Профессор
факультета психологии
Пенсильванского
Университета (США)

Экспертиза:
Геополитическое
прогнозирование

Как измерить
качество прогноза?
**Метрика Брайера
(Brier score)**
Квадрат разности между
исходом и прогнозом

Лучшее значение: 0
Худшее значение: 1
Значение для $p = 0.5$: 0.25

	Иван	Исход	МБ
Сб. 1	1.0	1	0.00
Сб. 2	0.9	1	0.01
Сб. 3	0.5	0	0.25
Сб. 4	0.7	0	0.49
Сб. 5	0.0	1	1.0
СМБ			0.35

1. Подобрать несколько событий для прогнозирования
2. Запросить прогнозы вероятностей этих событий у экспертов
3. Вычислить метрики Брайера для каждого эксперта и события
4. Упорядочить экспертов по значению средней метрики Брайера (СМБ)
5. Наилучший эксперт – тот, у кого СМБ **минимальна**.

Исследования Тетлока



*Профессор
факультета психологии
Пенсильванского
Университета (США)*

Экспертиза:
Геополитическое
прогнозирование

Ранние исследования: **Expert Political Judgment (EPJ)**

- Проводились в 1980х годах
- Эксперты: экономисты, политологи, военные аналитики
- Результаты: СМБ близко к 0.25
- Не лучше подбрасывания монетки
- Или «обезьяны, метаящей дротики»
- Но лучше, чем обычные люди

Исследования Тетлока



*Профессор
факультета психологии
Пенсильванского
Университета (США)*

Экспертиза:
Геополитическое
прогнозирование

Ранние исследования: **Expert Political Judgment (EPJ)**

- Проводились в 1980х годах
- Эксперты: экономисты, политологи, военные аналитики
- Результаты: СМБ близко к 0.25
 - Не лучше подбрасывания монетки
 - Или «обезьяны, метящей дротики»
- Но лучше, чем обычные люди

Нынешние исследования: **Good Judgment Project (superforecasters)**

- Турнир по прогнозированию (с 2011 по 2014 годы)
- **Открытое участие:** талантливые любители
 - Но отобранные по IQ и политическим знаниям
- Выступили на 70% лучше ожидаемой планки
- СМБ лучшего прогнозиста: 0.07
 - Гораздо лучше обезьяны
- Секрет успеха: отбор верхних 2%, работа в командах, агрегирование прогнозов

Детали о «суперпрогнозистах»

- **События**

- Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года? Да/Нет
- Четко определенные исходы
- Вопросы задаются от 3 до 24 месяцев до времени, когда результат будет известен
- Очень сложно предсказать исходы на более продолжительном горизонте
- За четыре года турнира были предсказаны 493 события

- **Прогнозисты**

- 7202 прогнозистов участвовало за 4 года...
- ...и сделали 3,143,360 прогнозов
- Кто был среди лучших прогнозистов?
 - В основном программисты, но также ученые, журналисты, и многие другие
- Платили ли им за прогнозы?
 - Нет, только за участие (подарочный сертификат Амазона на \$250)

Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



Дедлайн

Исход события
будет известен.

Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



Дедлайн

Исход события
будет известен.

Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



0.75

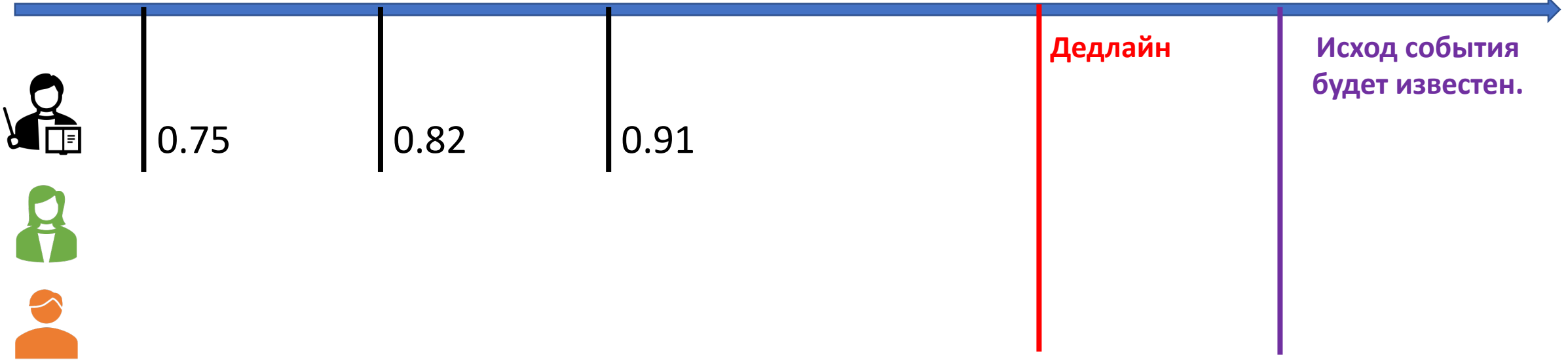


Дедлайн

Исход события
будет известен.

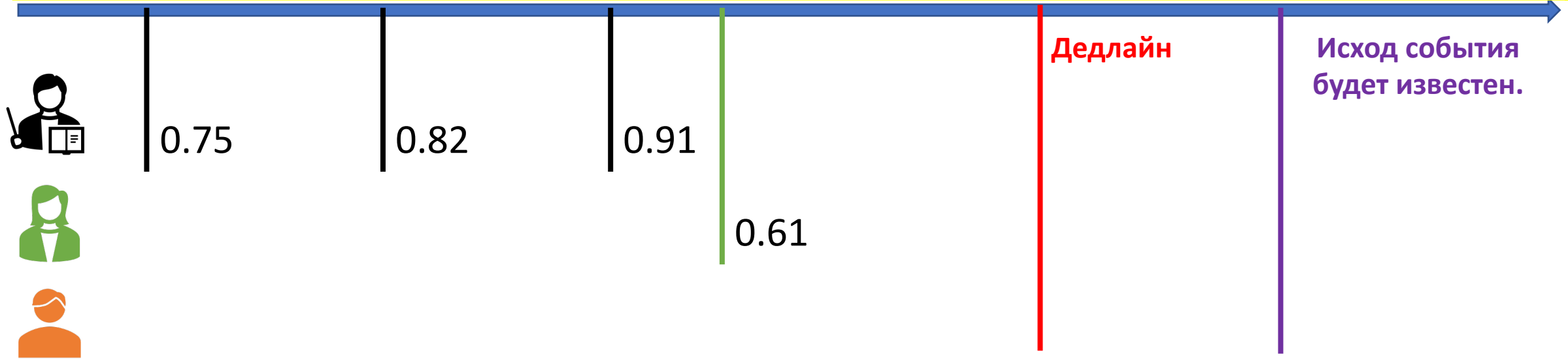
Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



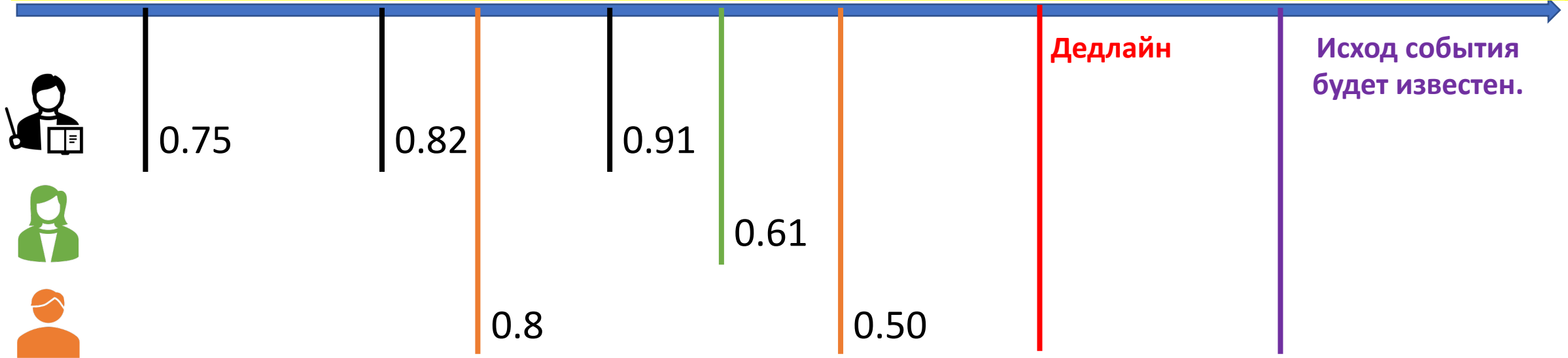
Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



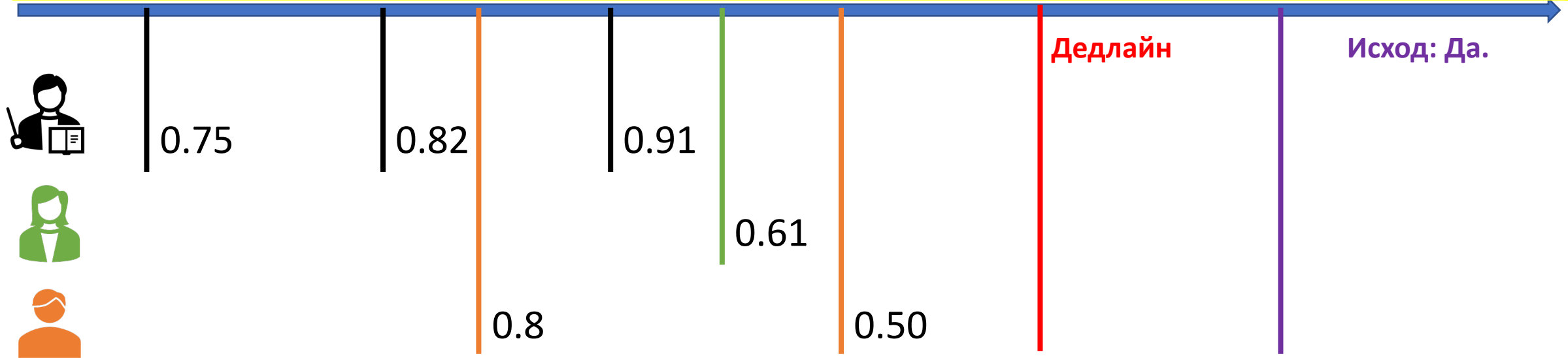
Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



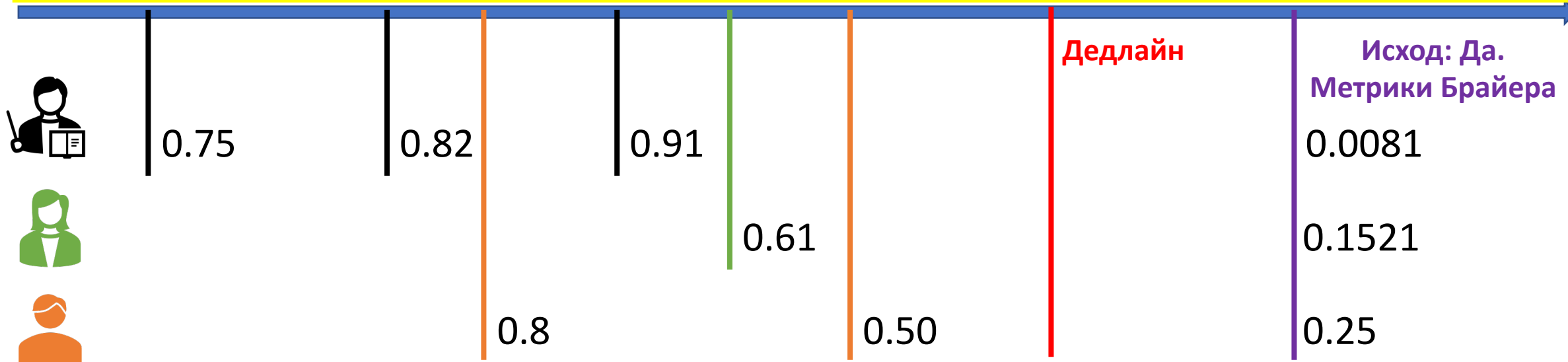
Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



Как был структурирован турнир?

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



Как улучшить качество прогнозов?

Подготовка

улучшить качество индивидуальных прогнозов

Работа в командах

установить контакты между экспертами

Отбор лучших

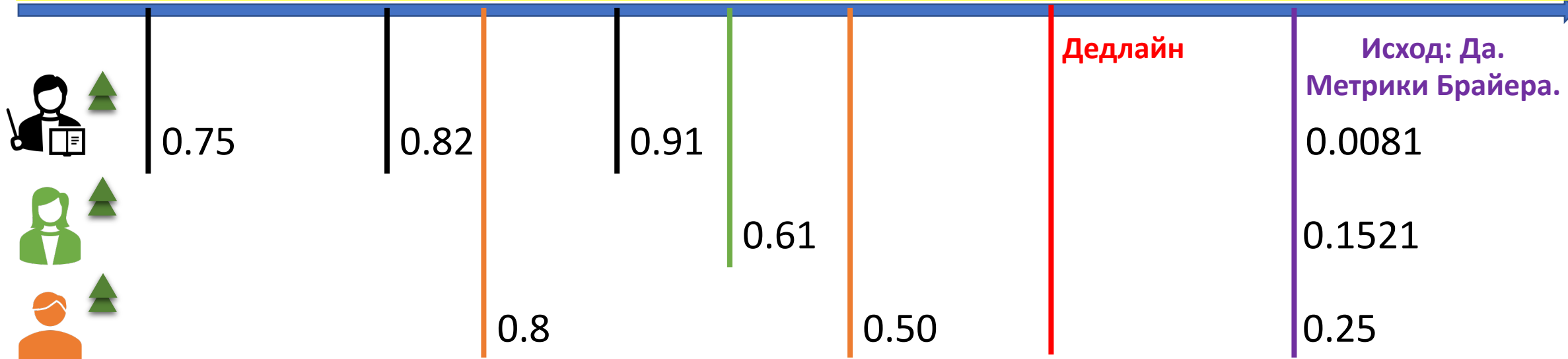
отбор лучших прогнозистов и объединение их в команды

Агрегация

комбинирование прогнозов от нескольких экспертов

Подготовка

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



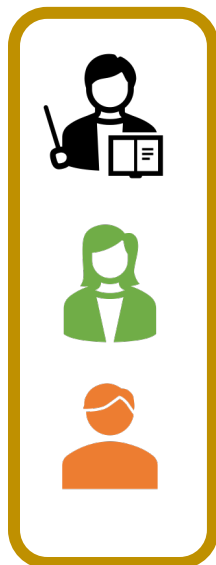
Повысить уровень каждого прогнозиста

Обучить основам теории вероятностей и статистике

Объяснить, как устроены когнитивные искажения

Работа в командах

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



0.75

0.82

0.91

0.61

0.8

0.50

Дедлайн

Исход: Да.
Метрики Брайера.

0.0081

0.1521

0.25

Все прогнозисты могут обсуждать свои прогнозы в командах

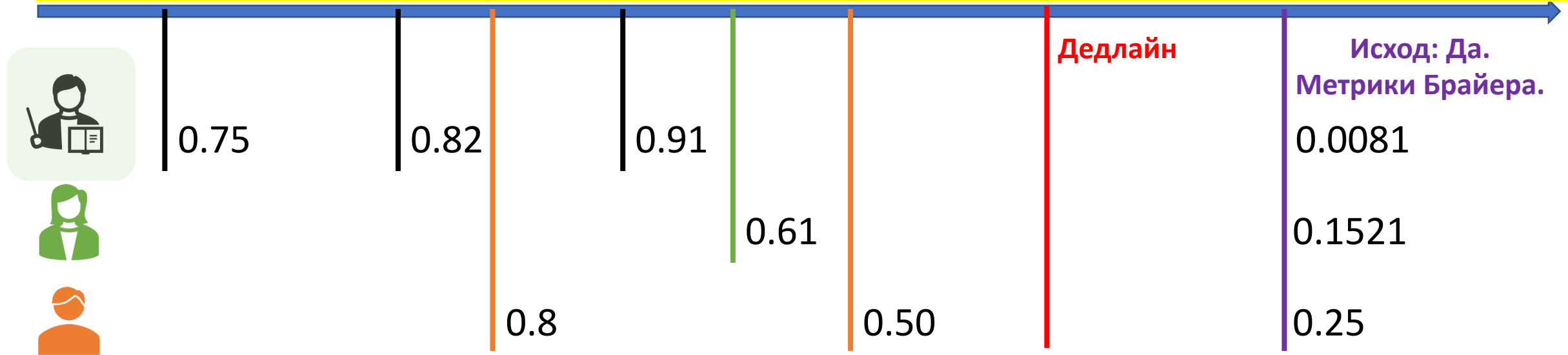
Но прогнозы делаются индивидуально

Плюсы: обмен информацией

Минусы: снижение оригинальности прогнозов

Отбор лучших

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



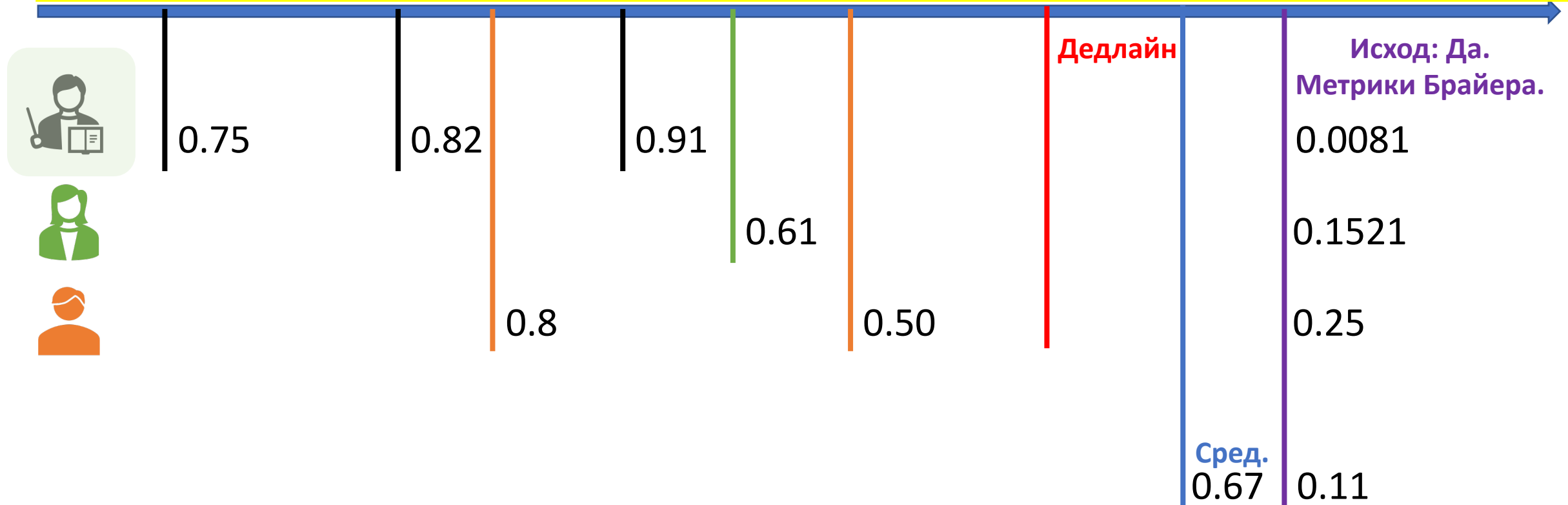
Выбираем лучшие 2 процента по СМБ в первом году.

В следующие три года они продолжают быть лучшими.

Их назвали “суперпрогнозистами” и объединили в команды.

Агрегирование

Будет ли Даниэль Ортега переизбран на следующий президентский срок в Никарагуа на выборах в конце 2011 года?



Например, можно усреднить три прогноза.

В нашем примере работает лучше, чем двое из троих прогнозистов.

Агрегирование «убирает» шумы и искажения.

Если брать среднее по многим событиям, то работает лучше, чем лучшие.

Агрегирование

Смещение среднего к более экстремальным значениям часто улучшает кач-во прогноза.

	Иван	Сергей	Даша	Среднее	Экстрем.	Исход
Событие 1	0.8	0.7	0.6	0.7	0.8	1
Событие 2	0.9	0.9	0.6	0.8	0.9	1
Событие 3	0.6	0.3	0.5	0.47	0.4	0
Событие 4	0.6	0.2	0.5	0.43	0.4	0
Событие 5	0.4	0.1	0.6	0.37	0.3	1
СМБ	0.226	0.208	0.196	0.188	0.172	

Также можно улучшить качество прогнозирования взвешиванием прогнозистов

Насколько сильно улучшения происходят?

10-15% от отбора прогнозистов
по интеллекту и открытости

10-20% от совместной работы в командах или
конкуренции в рамках предсказательных рынков

10%: индивидуальная подготовка прогнозистов

15-30%: агрегирование. Взвешенное усреднение.
Большой вес лучшим прогнозистам.
Экстремизация результирующих средних.

Как предсказывают лучшие?

- Гранулярные предсказания
 - Эксперты, у которых более широкий спектр прогнозов, предсказывают лучше
- Частые обновления прогнозов
- Открытость новому
- Когнитивная рефлексия
 - Умение «заметить подвох»

Где найти более актуальные прогнозы?

<http://www.metaculus.com>

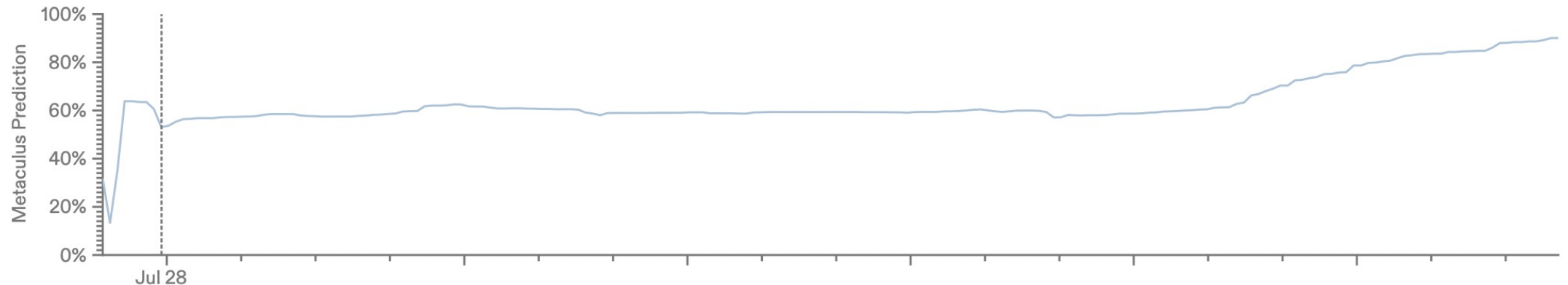
Агрегирует опросы участников

Talibans capture Afghan Palace by 2026-09-11

Created by [alwaysrinse](#) . Opened Jul 5, 2021 • resolved Aug 15, 2021.

^ 35 v Resolved : Yes

Forecast Timeline



Total Predictions **29**

Total Forecasters **14**

Community Prediction **57%**

Metaculus Prediction **53%**

M Metaculus

Где найти более актуальные прогнозы?

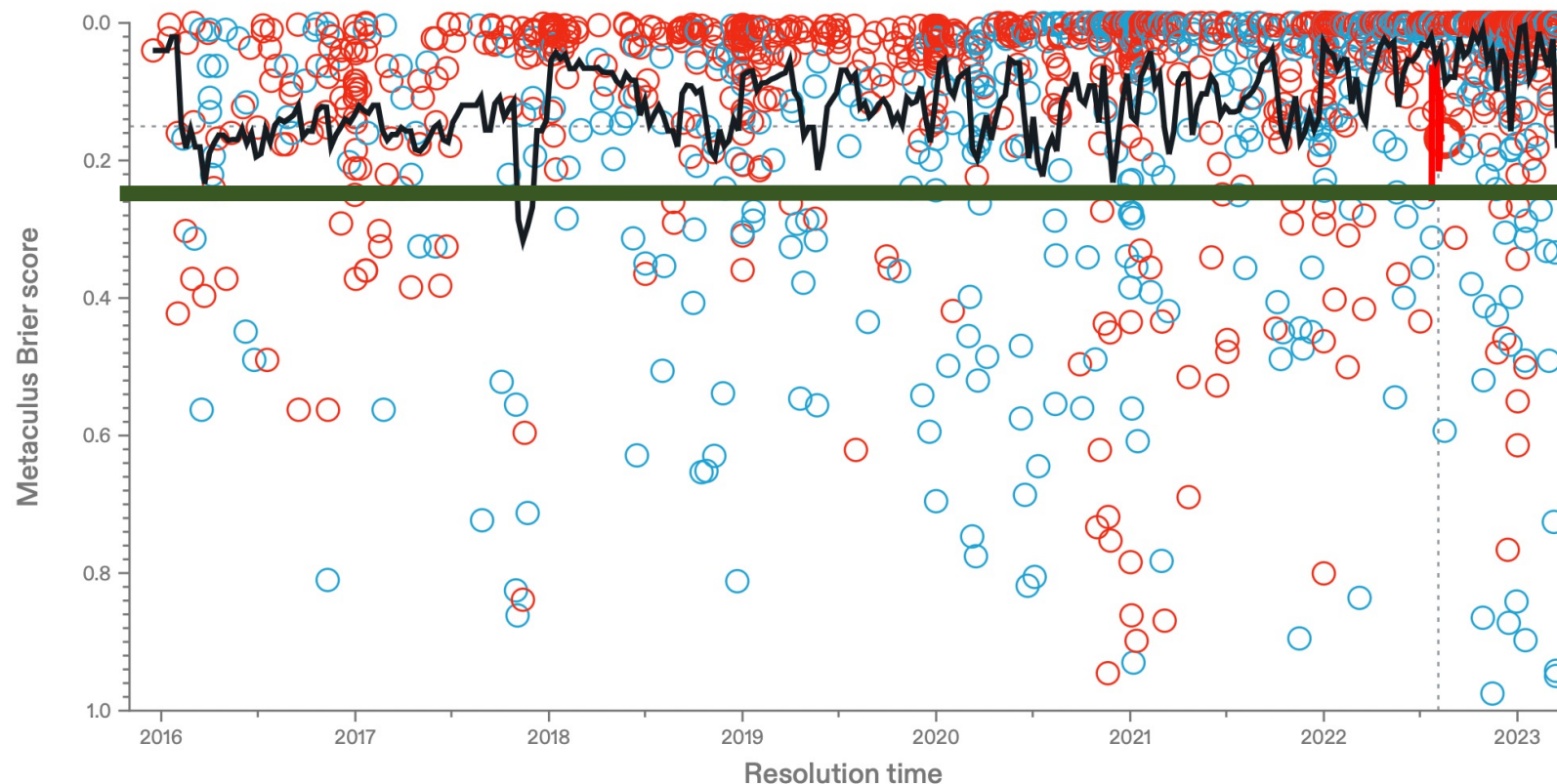
<https://www.metaculus.com/questions/track-record/>

Предсказания довольно точные

Total questions: 1706

Average score: 0.086

Maximum Uncertainty Score: 0.250



Каждая точка – это вопрос

Все прогнозы
выше этой линии
лучше случайных
догадок

Who will win the 2022 Kenyan Presidential election? (Raila Odinga).

Resolution: **No**

Community prediction: **47%**

Metaculus prediction: **41%** (43% postdiction)

Рынки предсказаний (prediction markets)

- Виртуальные акции, цены привязаны к исходам событий
- Например: держатель 1 акции получает 1 доллар, если премьер Японии посещает Северную Корею до 1 января 2023 года
- Участники рынка могут продавать и покупать акции
- Цены соответствуют вероятностям
 - Если актив недооценен, участники рынка будут покупать
 - Если переоценен, то продавать
- Такие рынки обычно агрегируют информацию эффективнее опросов

Polymarket

polymarket.com



Will Germany or United Kingdom report a higher 7-day COVID-19 case average per capita on November 24?

New

Volume

\$810

Germany

\$0.29

United King...

\$0.71



Will Squid Game be the #1 Netflix show worldwide on Week 46, 2021 (Nov 15 - Nov 21)?

New

Volume

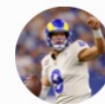
\$802

Yes

\$0.12

No

\$0.88



Will Matthew Stafford have the most passing yards in the 2021 NFL season on November 27th?

New

Volume

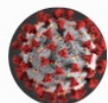
\$763

Yes

\$0.01

No

\$0.99



Will 231 million Americans have received at least one dose of an approved Covid-19 vaccination by November 29?

New

Volume

\$727

Yes

\$0.95

No

\$0.05



Will NASA have confirmed more than 4580 existing exoplanets by November 27?

New

Volume

\$706

Yes

\$0.24

No

\$0.76



Will 'House of Gucci' get 86% or higher Audience Score on Rotten Tomatoes?

New

Volume

\$658

Yes

\$0.41

No







\$0.59



Kalshi

<https://kalshi.com/events/economics>

24 results

			Yes	No	Volume
	Fed interest rates March 2023 meeting	Above 4.75%	76¢	26¢	1,395,741
		Above 5.00%	3¢	98¢	
	US inflation In March 2023	Above 0.2%	37¢	66¢	44,268
		Above 0.1%	75¢	29¢	
	US GDP growth In 2023 Q1	Above 2.4%	54¢	47¢	20,540
		Above 2.9%	33¢	72¢	
	US core inflation Mar 2023	Above 0.6%	99¢	99¢	12,814
		Above 0.4%	65¢	37¢	
	Jobs numbers In March 2023	Above 200,000	40¢	61¢	16,015
		Above 100,000	78¢	27¢	
	Fed Funds rate	5.50% to 5.00%	32¢	74¢	

Insight Prediction

insightprediction.com



Will the NBER Declare a Recession in 2023?

Resolution Date: Apr 1, 2024

Volume

\$1,525

Yes

42%

No

72%



Who Will Be the Senate-confirmed Treasury Secretary on..

Resolution Date: Jun 2, 2023

Volume

\$250

Janet Yellen

96%

Wally...

50%



Will There Be a NY-03 Special Election in 2023?

Resolution Date: Jan 2, 2024

Volume

\$2

Yes

98%

No

-



Will Trump Tweet Before April 2023?

Resolution Date: Apr 2, 2023

Volume

\$4,716

Yes

-

No

99%

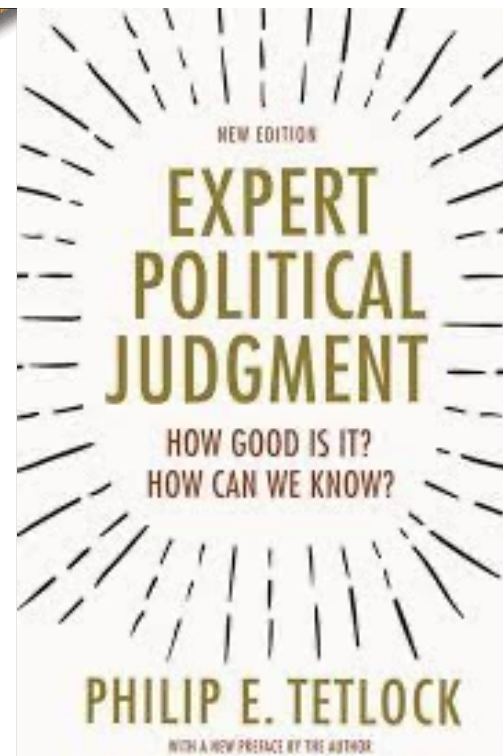
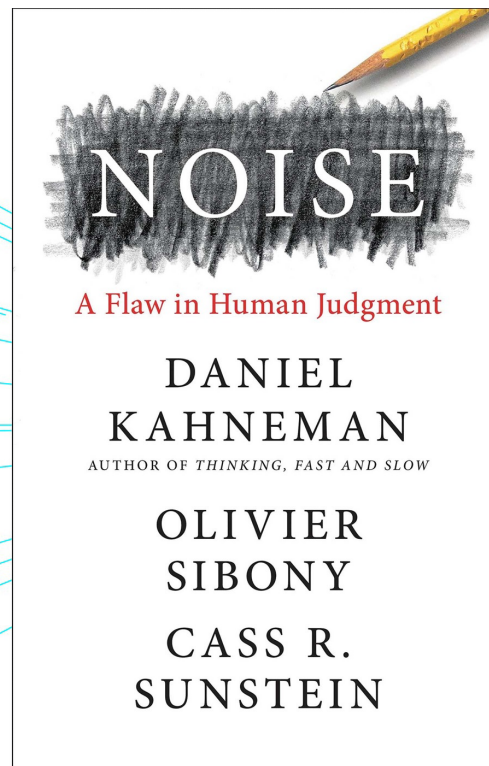
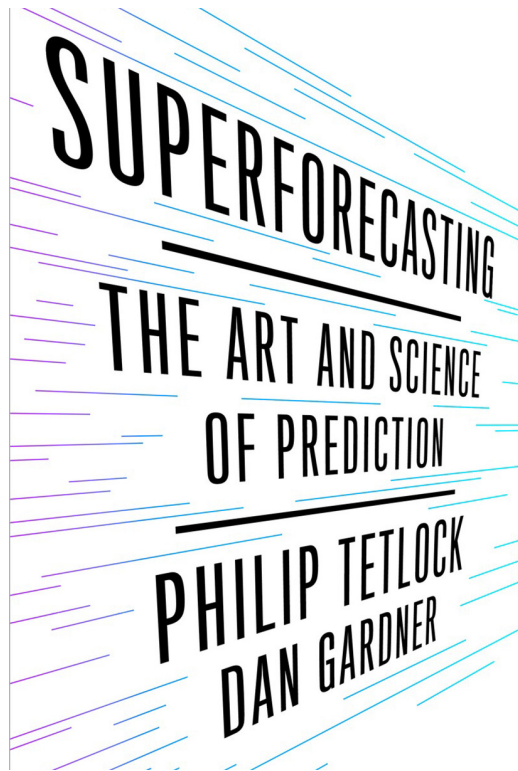
Вопросы, требующие дальнейших исследований

- По каким характеристикам нужно отбирать потенциальных прогнозистов?
- Как стимулы внутри организаций влияют на качество экспертных прогнозов?
- Как применить эти методы прогнозирования для решения других задач?
- Как правильно мотивировать экспертов? Платить ли за результат?
- Какой метод агрегирования более эффективен: рынки или опросы?
- Как наиболее эффективно агрегировать прогнозы?
- Какие вопросы нужно задавать экспертам для наилучшего качества прогноза?
- Как комбинировать статистические и экспертные прогнозы?

Выводы

Для принятия решений нужны точные вероятности событий.

Чтобы получить эти вероятности, мы можем готовить, выбирать и опрашивать экспертов, после чего агрегировать их прогнозы.



Заключение

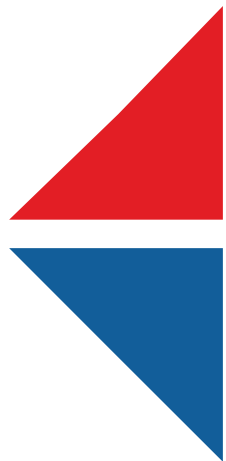
Если Вам интересно узнать больше, пишите: msalikhov@nes.ru

Будем рады вас видеть в РЭШ! Приглашаем на следующие события:

**День открытых дверей
Совместного бакалавриата НИУ ВШЭ и РЭШ**
9 апреля 2022 в 12-00, кампус РЭШ
онлайн-трансляция



Школьный конкурс РЭШ
онлайн
С 1 по 30 апреля 2022
Регистрация уже открыта



РЭШ

Российская
экономическая
школа