



# Машинное обучение в науках о Земле

Михаил Криницкий

К.Т.Н., Н.С.

Институт океанологии РАН им. П.П. Ширшова

Лаборатория взаимодействия океана и атмосферы и  
мониторинга климатических изменений (ЛВОАМКИ)

# Михаил Криницкий

- к.т.н., н.с. ИО РАН, доцент Физтеха
- разработка и анализ алгоритмов машинного обучения и глубокого обучения в применении к задачам наук о Земле



# Правила оффлайн-занятий

- Чем ниже посещаемость студента, тем более пристрастным будет зачет и диф.зачет

# Правила онлайн-занятий

- Камера включена у всех
- Микрофон выключен у всех за исключением говорящего
- Дискуссия приветствуется. Для ее начала следует включить микрофон

# Темы 1го курса

- Введение в МО: классификация задач МО, способы решения задач анализа данных
- Задачи типа «обучение с учителем»: формулировка, способы решения
- Технические средства анализа данных
- Линейная регрессия: вероятностная постановка, решение в подходе оптимизации правдоподобия
- Оценка неопределенности в задачах типа «обучение с учителем»
- Задача классификации: формулировка, способы решения
- Различные модели для решения задач регрессии и классификации
- Искусственные нейронные сети: формулировка, особенности оптимизации

# Домашние задания (1 курс)

- ДЗ №1: постановка задач машинного обучения для избранных проблем наук о Земле
- ДЗ №2: вероятностная постановка задачи линейной регрессии
- ДЗ №3: оценка неопределенностей в задаче линейной регрессии
- ...
- ДЗ №4: мультиномиальная логистическая регрессия и оценка качества моделей классификации
- ДЗ №5: свойства функции потерь логистической регрессии
- ДЗ №6: Курсовой проект (решение реальной геофизической задачи)
- ...

# Темы 2го курса

- Технические средства построения и оптимизации ИНС
- Полносвязные нейронные сети
- Сверточные нейронные сети
- Избранные задачи, решаемые с применением ИНС: автокодировщик, варианты сегментации, задача обучения метрике
- Задачи типа «обучение без учителя»: сокращение размерности, кластеризация
- Рекуррентные нейронные сети и задача обработки последовательностей
- Порождающие модели, порождающие состязательные сети

# Домашние задания (1 курс)

- ДЗ №1: Построение и оптимизация полносвязной ИНС
- ДЗ №2: Построение и оптимизация сверточной ИНС
- ДЗ №3: Задача семантической сегментации
- ДЗ №4: Задача регрессии на двумерной сетке
- ДЗ №5: Автокодировщик для обработки геофизических данных
- ДЗ №6: Задача обучения метрике
- ДЗ №7: Задача сокращения размерности
- ДЗ №8: Кластеризация
- ДЗ №9: Обработка временных рядов с применением РНС
- Курсовой проект

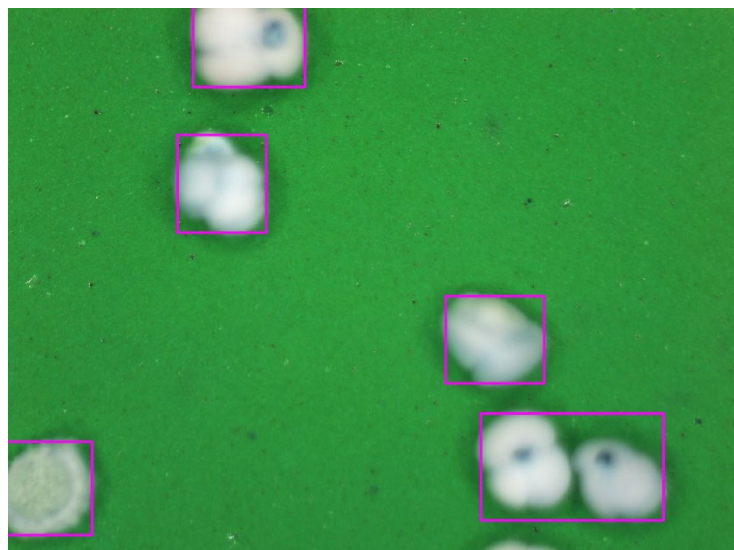


# Михаил Криницкий

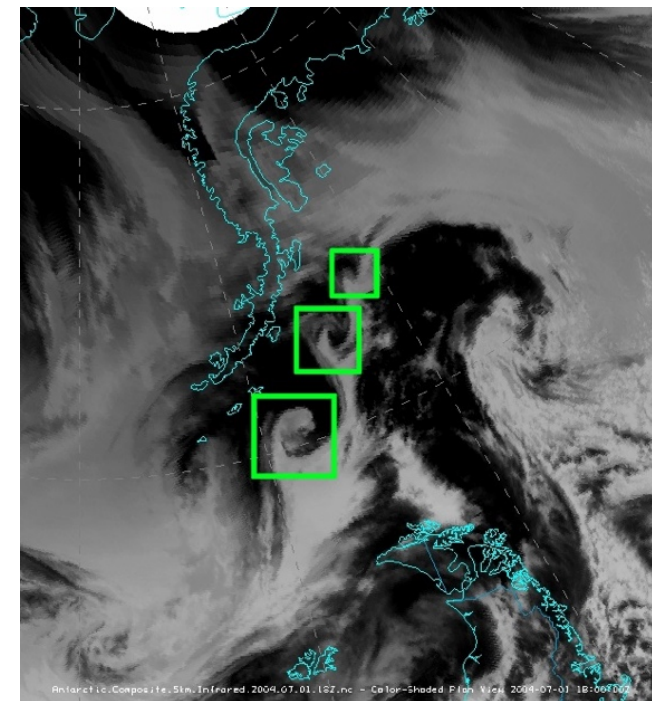
- к.т.н., н.с. ИО РАН, доцент Физтеха
- разработка и анализ алгоритмов машинного обучения и глубокого обучения в применении к задачам наук о Земле



Обработка метеоснимков



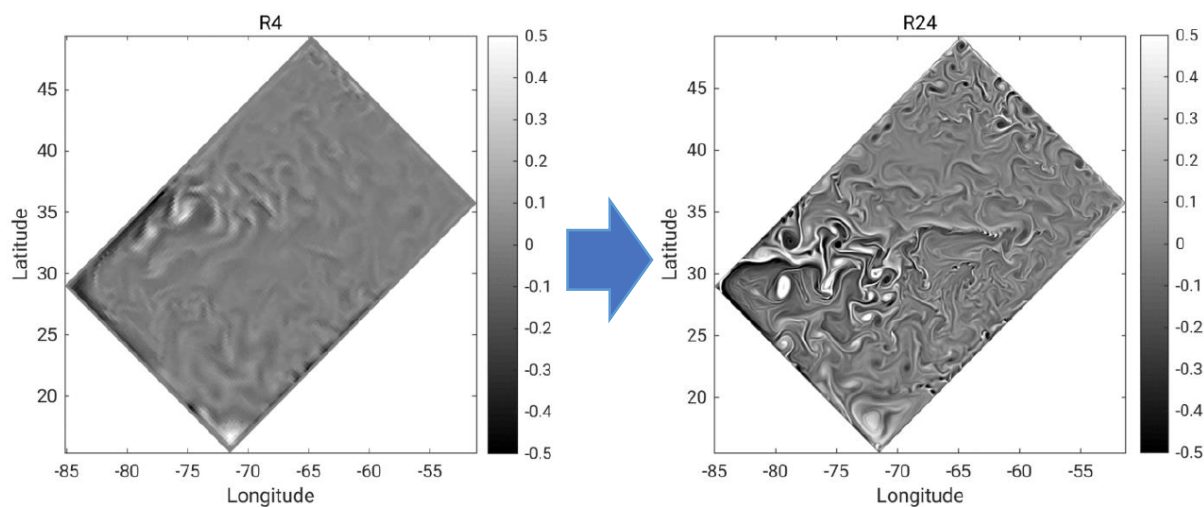
Анализ геологических микроснимков



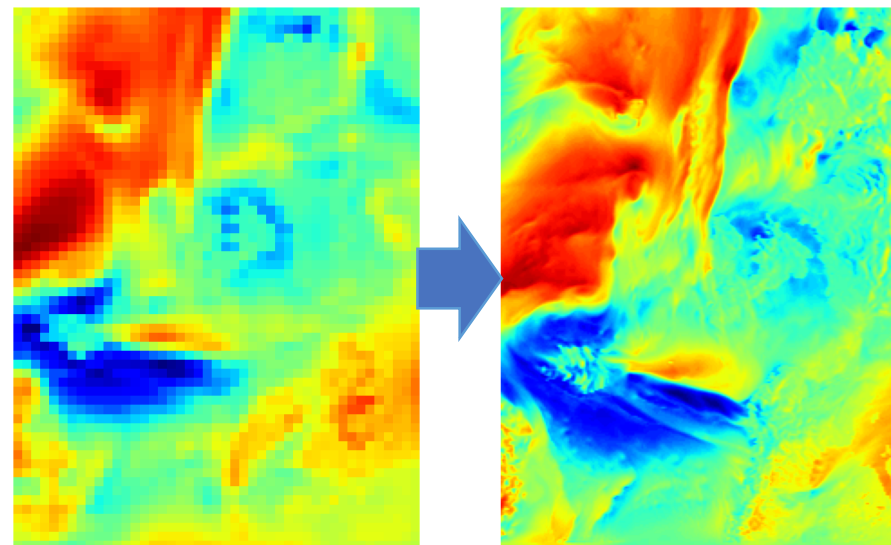
Анализ спутниковых снимков

# Михаил Криницкий

- к.т.н., н.с. ИО РАН, доцент Физтеха
- разработка и анализ алгоритмов машинного обучения и глубокого обучения в применении к задачам наук о Земле



Масштабирование динамики течений



Масштабирование скорости ветра

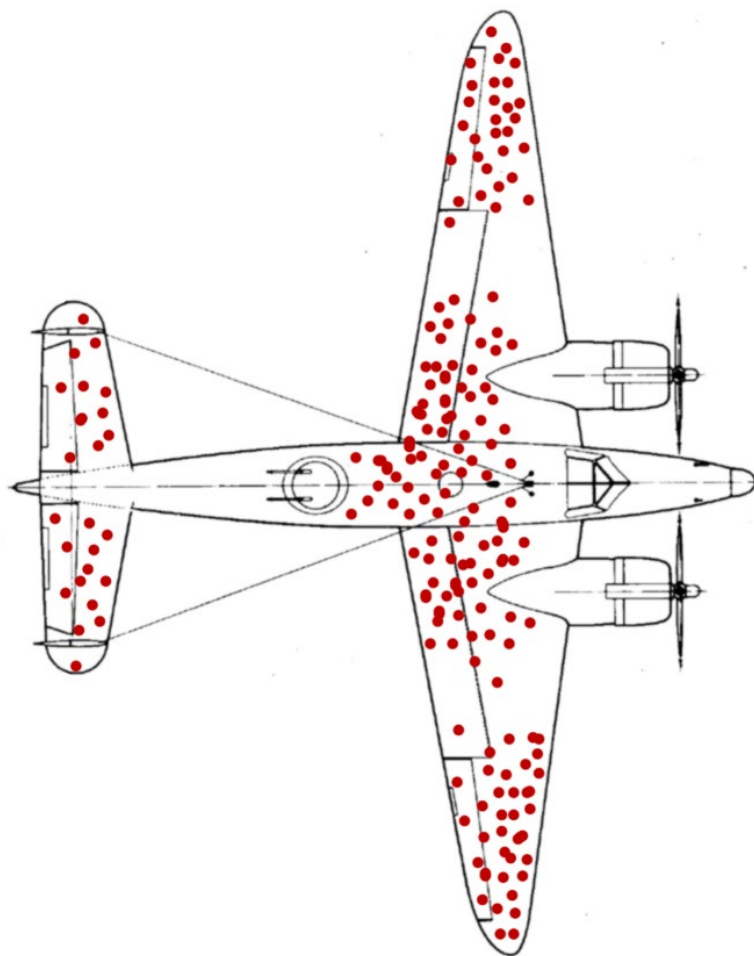
# Курс: Машинное обучение в науках о Земле

# Ошибка выживших



Абрахам Вальд

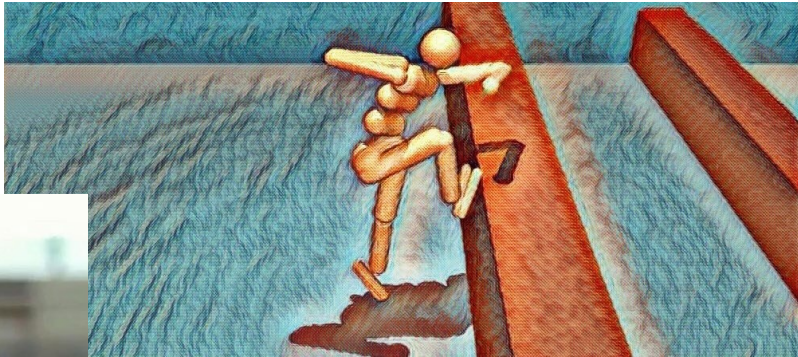
# Ошибка выживших



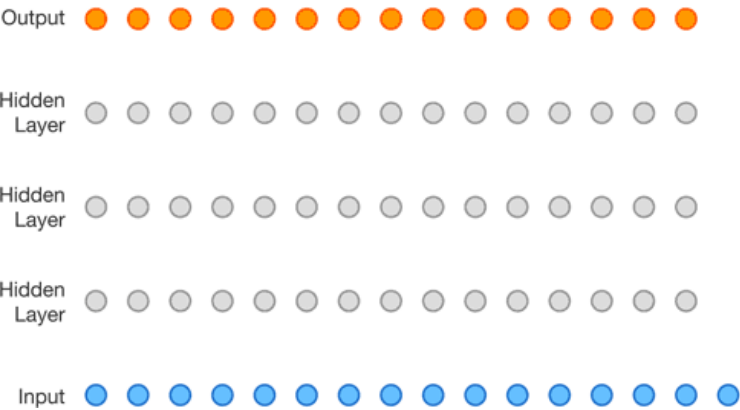
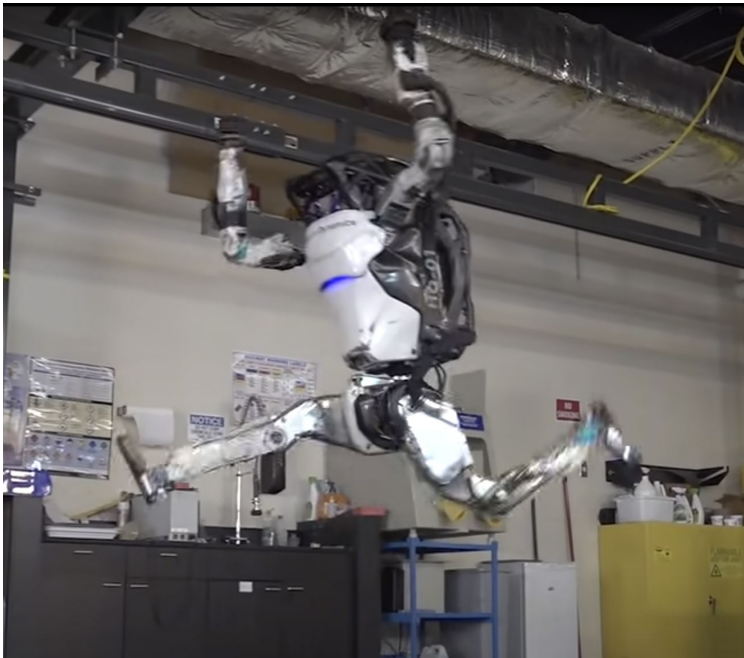
Абрахам Вальд



# Ошибка выживших



Абрахам Вальд



# ML4ES

Цель курса:

Повысить шансы на «выживание».

Дать понимание, «куда воевать» в деле обработки геофизических данных

Дать понимание, «куда копать», если модель не работает или работает не так, как хочется.

# ML4ES

Тебе нужен этот курс, если:

- Хочешь разбираться в особенностях современной технологии, которая все чаще применяется в задачах геофизики
- Хочешь применять МО в своей научной деятельности
- Хочешь настроить «детектор кошки» на своем участке
- Хочешь понимать особенности современных технологий распознавания лиц
- Хочешь аргументированно обосновать свою точку зрения по вопросу приватности своих биометрических персональных данных
- ...



# Организационные вопросы

- Материалы курса, 1й год:
  - <https://github.com/MKrinitskiy/ML4ES1-F2020-S2021>
- Материалы курса, 2й год:
  - <https://github.com/MKrinitskiy/ML4ES2-F2020-S2021>

# Организационные вопросы

- Материалы курса, 1й год:
  - <https://github.com/MKrinskiy/ML4ES1-F2020-S2021>
- Материалы курса, 2й год:
  - <https://github.com/MKrinskiy/ML4ES2-F2020-S2021>
- ДЗ: [krinitsky.ma@phystech.edu](mailto:krinitsky.ma@phystech.edu)

# Организационные вопросы

- Материалы курса, 1й год:
  - <https://github.com/MKrinitский/ML4ES1-F2020-S2021>
- Материалы курса, 2й год:
  - <https://github.com/MKrinitский/ML4ES2-F2020-S2021>
- ДЗ: [krinitский.ma@phystech.edu](mailto:krinitский.ma@phystech.edu)
- Расписание 1-го курса
- Расписание 2-го курса



# Введение

Михаил Криницкий

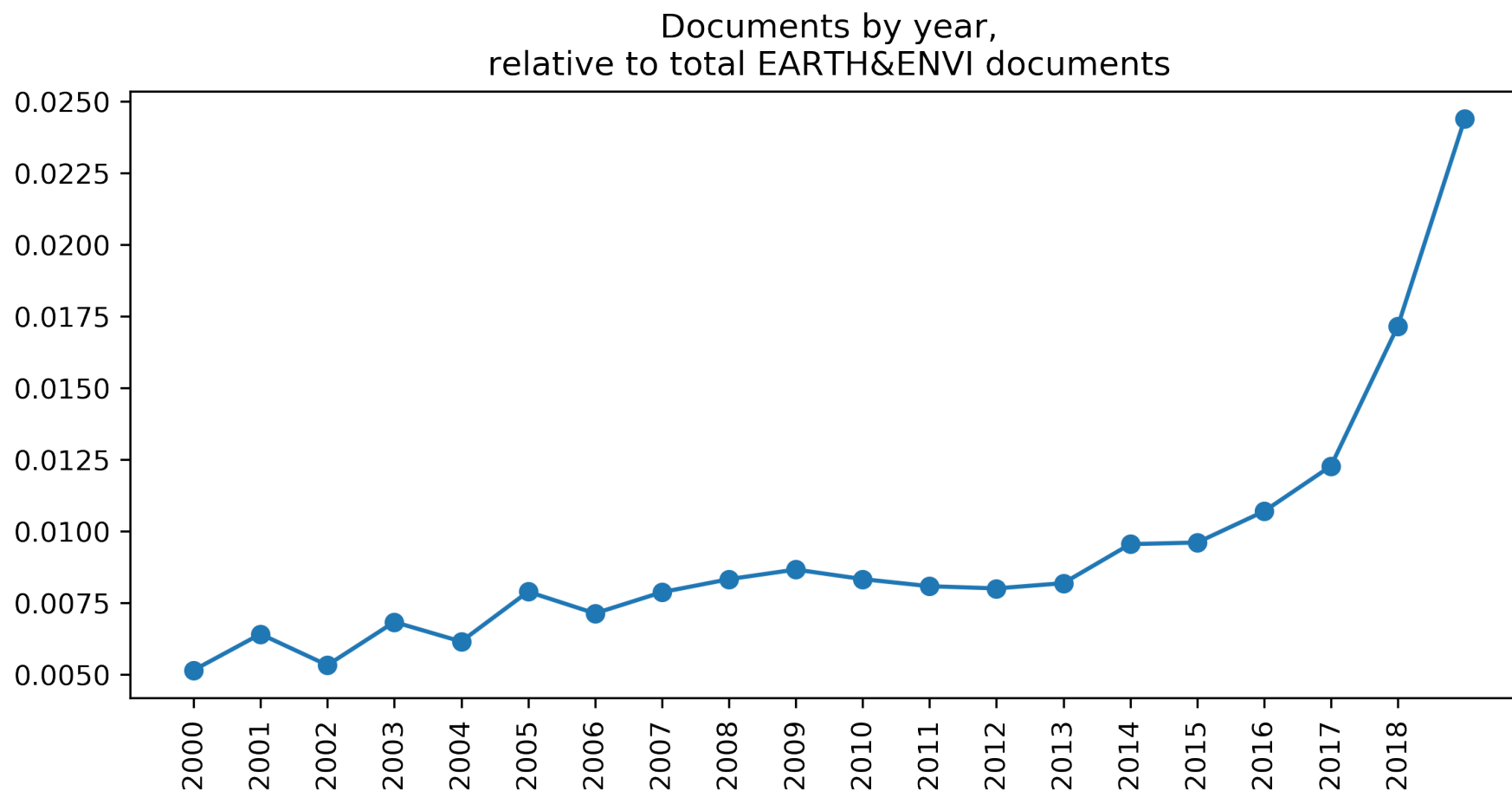
К.Т.Н., Н.С.

Институт океанологии РАН им. П.П. Ширшова

Лаборатория взаимодействия океана и атмосферы и  
мониторинга климатических изменений (ЛВОАМКИ)

# МО в науках о Земле

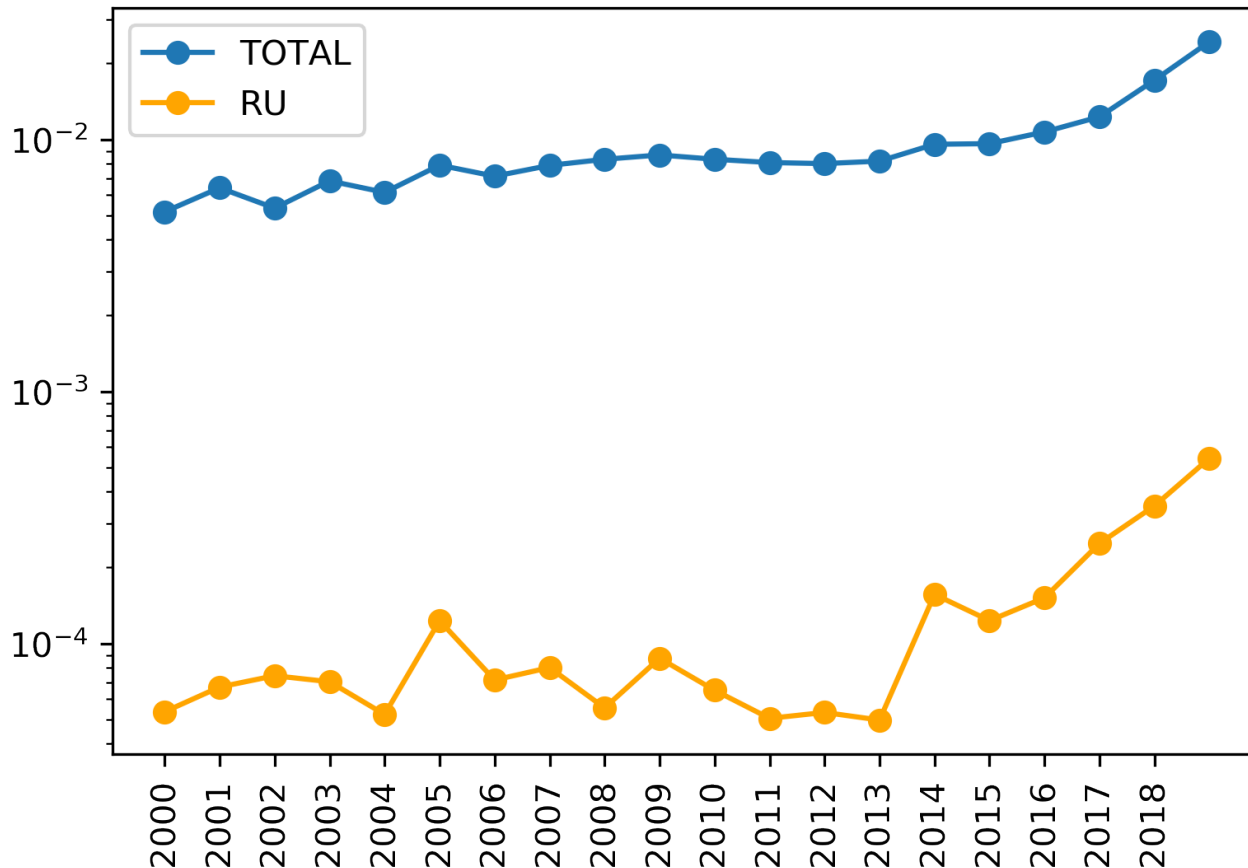
Статьи в рецензируемых журналах по тематике “Environmental sciences” и “Earth sciences” с применением методов машинного обучения. По данным Scopus.



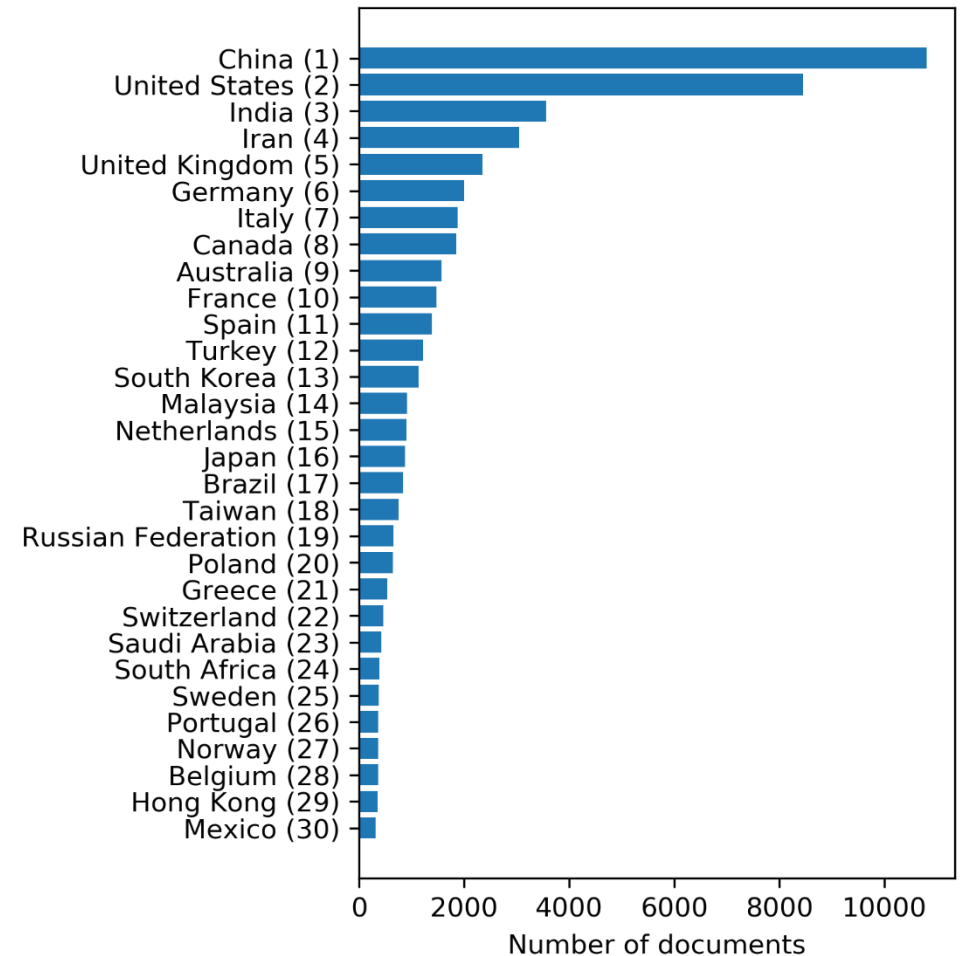
# МО в науках о Земле

Статьи в рецензируемых журналах по тематике “Environmental sciences” и “Earth sciences” с применением методов машинного обучения. По данным Scopus.

Documents by year,  
relative to total EARTH&ENVI documents



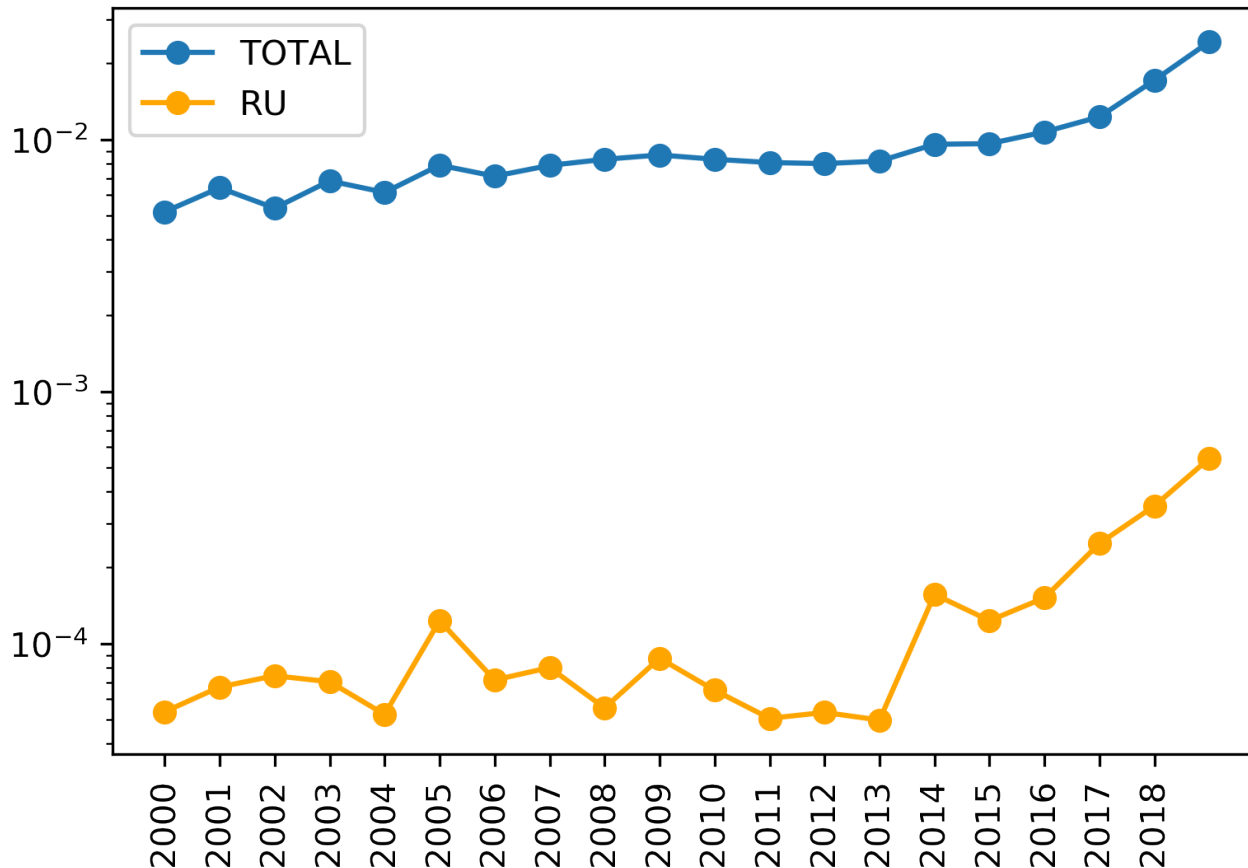
Documents, 2000-2019  
AI&ML in EART&ENVI



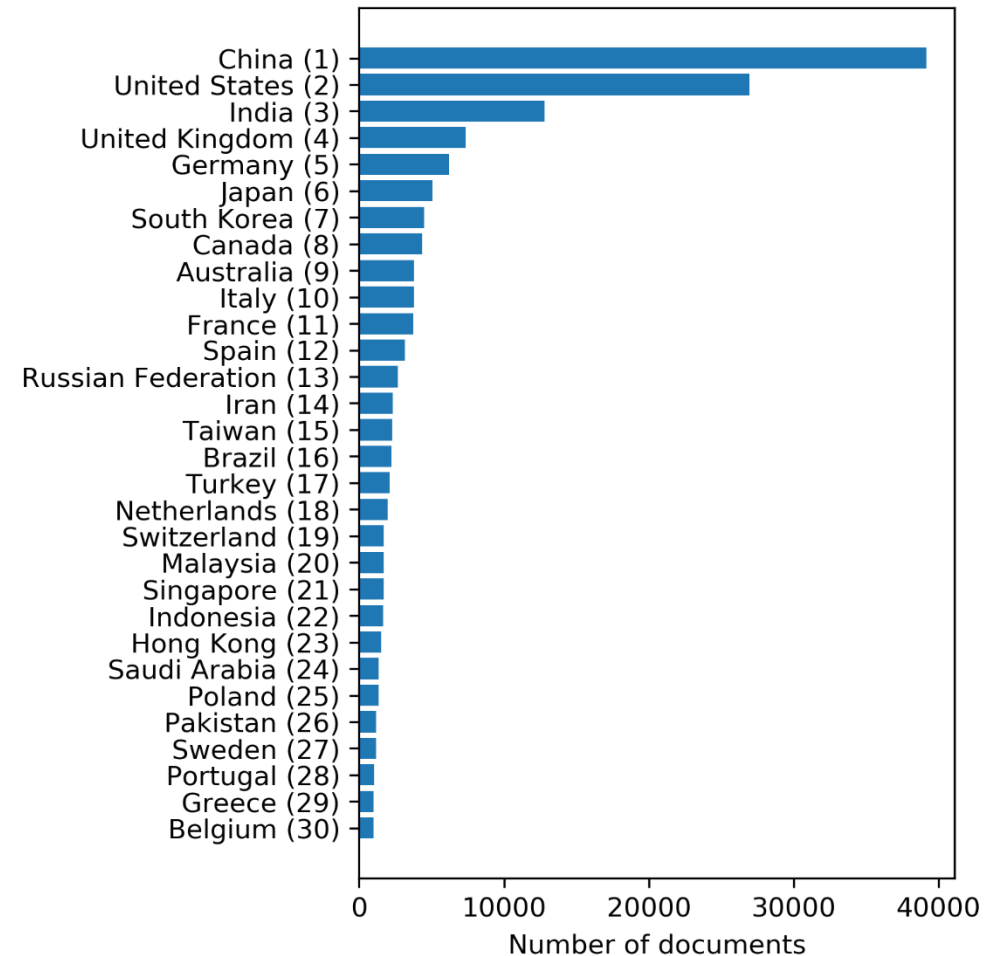
# МО в науках о Земле

Статьи в рецензируемых журналах по тематике “Environmental sciences” и “Earth sciences” с применением методов машинного обучения. По данным Scopus.

Documents by year,  
relative to total EARTH&ENVI documents



Documents, 2019  
AI&ML in EART&ENVI



# План лекции:

- Очень краткое введение в методы машинного обучения (ММО)
  - ММО: что это и зачем это нужно
  - Если я знаю, чего я хочу: «Обучение с учителем»
  - Если я еще не знаю, чего хочу: «Обучение без учителя»
- Примеры задач наук о Земле, решаемые с применением ММО:
  - Идентификация и построение траекторий полярных мезоциклонов
  - Кластеризация состояний стратосферного полярного вихря
  - Науकाстинг атмосферных осадков...
  - Масштабирование геофизических полей



# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

## ЧЕМ ЗАНИМАЮТСЯ ФИЗИКИ?

атмосферы

океана

взаимодействия океана и атмосферы

цунами

ядерщики, высоких энергий

био-

...

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

ЧЕМ ЗАНИМАЮТСЯ ФИЗИКИ?

ИЩУТ ИСТИНУ (об устройстве Вселенной) ?

ОТКРЫВАЮТ ЗАКОНЫ МИРОЗДАНИЯ ?

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

## ЧЕМ ЗАНИМАЮТСЯ ФИЗИКИ?

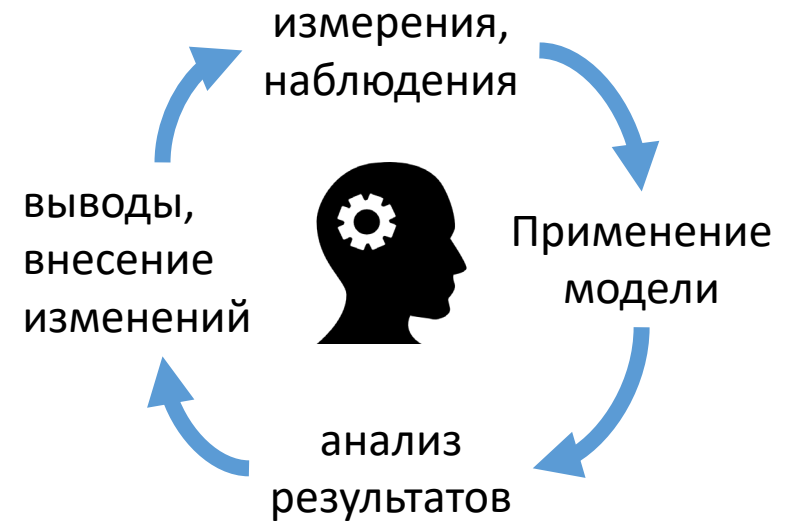
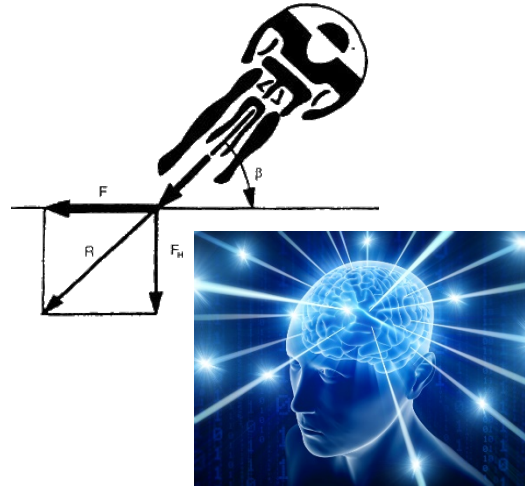
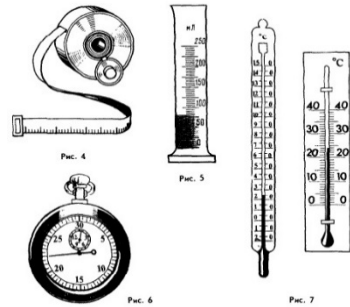
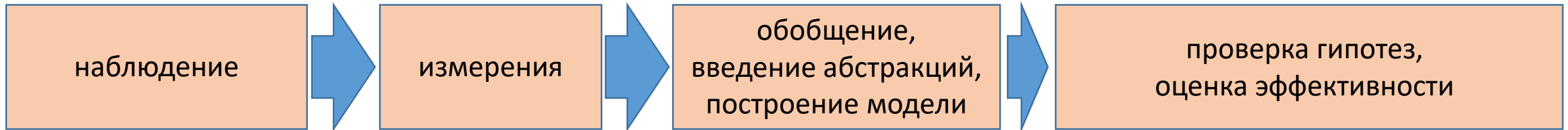
ПРЕДЛАГАЮТ МОДЕЛИ наблюдаемых процессов (**какие?..**)

ПРОВЕРЯЮТ - ОЦЕНИВАЮТ ЭФФЕКТИВНОСТЬ (**как?..**)

ПРИМЕНЯЮТ НА ПРАКТИКЕ

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

## КАК проводятся физические исследования?



Настоящая наука начинается с тех пор, как начинают измерять.  
Точная наука немыслима без меры.  
Д.И. Менделеев

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит

**но делать-то что-то надо**

наблюдение

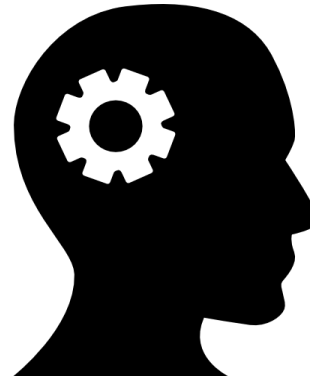
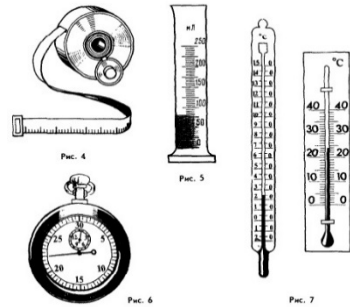
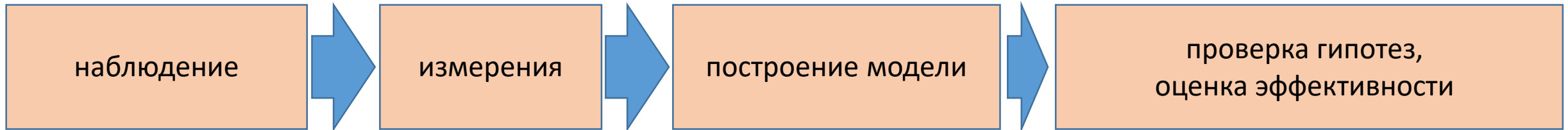


by Frits Ahlefeldt

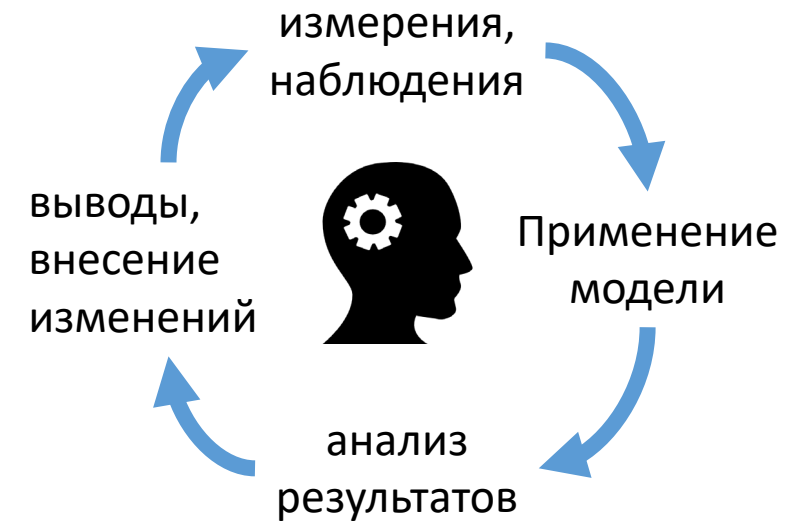
# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит

все равно строим модель



обобщение ?  
введение абстракций ?



# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит

все равно строим модель

- Для чего? Какова цель?
- Что у нас для этого есть?
- Какого рода модель?
- Какая должна быть модель?

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит  
все равно строим модель

- Для чего? Какова цель?
- Что у нас для этого есть?
- Какого рода модель?
- Какая должна быть модель?
- Оценить неизвестную(ые) величину(ы)  $\{y_i\}$
- 
- 
-



# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит  
все равно строим модель

- Для чего? Какова цель?
- Что у нас для этого есть?
- Какого рода модель?
- Какая должна быть модель?
- Оценить неизвестную(ые) величину(ы)  $\{y_i\}$
- Данные измерений  $\{x_i\}$  (желательно релевантных задаче)

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит  
все равно строим модель

- Для чего? Какова цель?
- Что у нас для этого есть?
- Какого рода модель?
- Какая должна быть модель?
- Оценить неизвестную(ые) величину(ы)  $\{y_i\}$
- Данные измерений  $\{x_i\}$
- $\mathcal{F}: \mathbb{X} \rightarrow \mathbb{Y}$

# ОЧЕНЬ КРАТКОЕ ВВЕДЕНИЕ В МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Когда (человеку) непонятно, что происходит  
все равно строим модель

- Для чего? Какова цель?
- Что у нас для этого есть?
- Какого рода модель?
- Какая должна быть модель?
- Оценить неизвестную(ые) величину(ы)  $\{y_i\}$
- Данные измерений  $\{x_i\}$
- $\mathcal{F}: \mathcal{X} \rightarrow \mathcal{Y}$
- Обобщающая. Достоверная (в каком смысле?)

Применимая.