

# Лекция 1

Введение

---

Екатерина Тузова

email: [machine.teaching@gmail.com](mailto:machine.teaching@gmail.com)

web: [http://mit.spbau.ru/sewiki/index.php/Машинное\\_обучение\\_2017](http://mit.spbau.ru/sewiki/index.php/Машинное_обучение_2017)

# Правила игры

---

# Правила игры

- 13 лекций
- 12 опросов по 5 баллов в начале лекции
- 8 домашних заданий по 20 баллов при сдаче в первую неделю, 10 баллов при сдаче во вторую неделю
- Экзамен 180 баллов
  
- Оценки за курс: 300 баллов – отлично, 250 баллов – хорошо, 200 баллов – удовлетворительно

# Машинное обучение

---

# Что такое машинное обучение?



Arthur Samuel (1959)

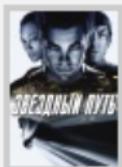
Field of study that gives computers the ability to learn without being explicitly programmed.

Чем отличается задача найти кратчайший путь в графе от антиспам фильтр?

# Применение машинного обучения

---

Если вам понравился этот фильм, не пропустите



[Звездный путь](#)  
Star Trek



[Стражи Галактики](#)  
Guardians of the  
Galaxy



[Стартрек:  
Возмездие](#)  
Star Trek Into  
Darkness



[Миссия  
«Серенити»](#)  
Serenity



[Изгой-один:  
Звёздные войны.](#)  
[Истории](#)  
Rogue One

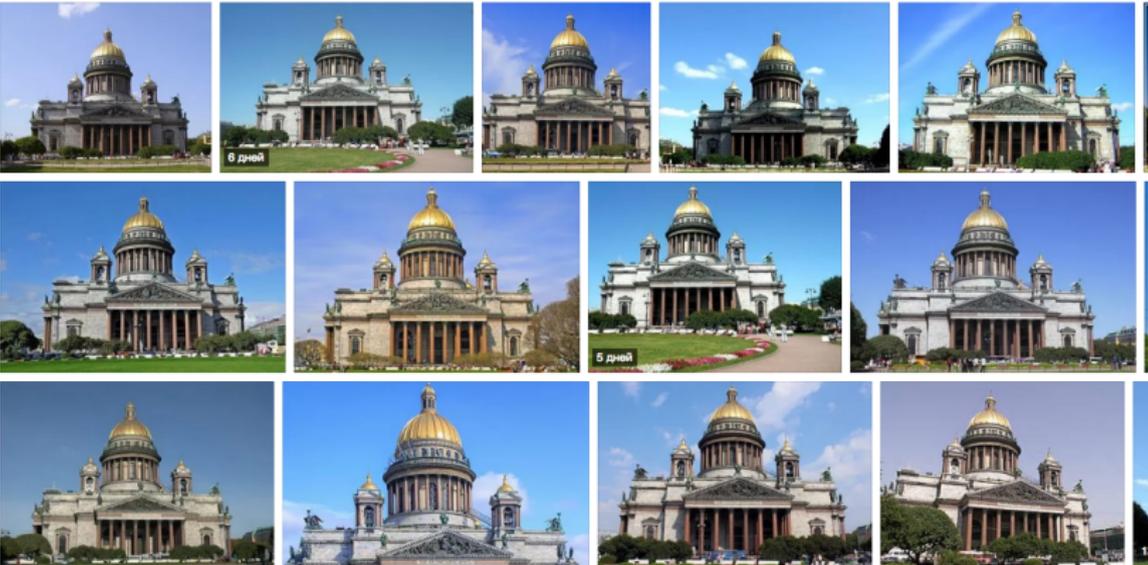
# Информационный поиск

Яндекс

Загруженная картинка X

Найти

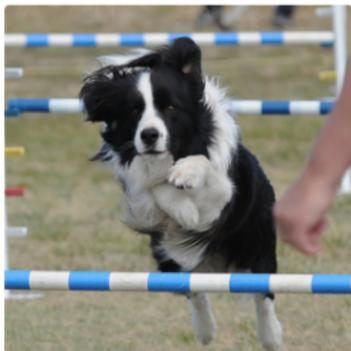
ПОИСК КАРТИНКИ ВИДЕО КАРТЫ МАРКЕТ ЕЩЕ



6 дней

6 дней

# Подпись изображений



"black and white dog jumps over bar."

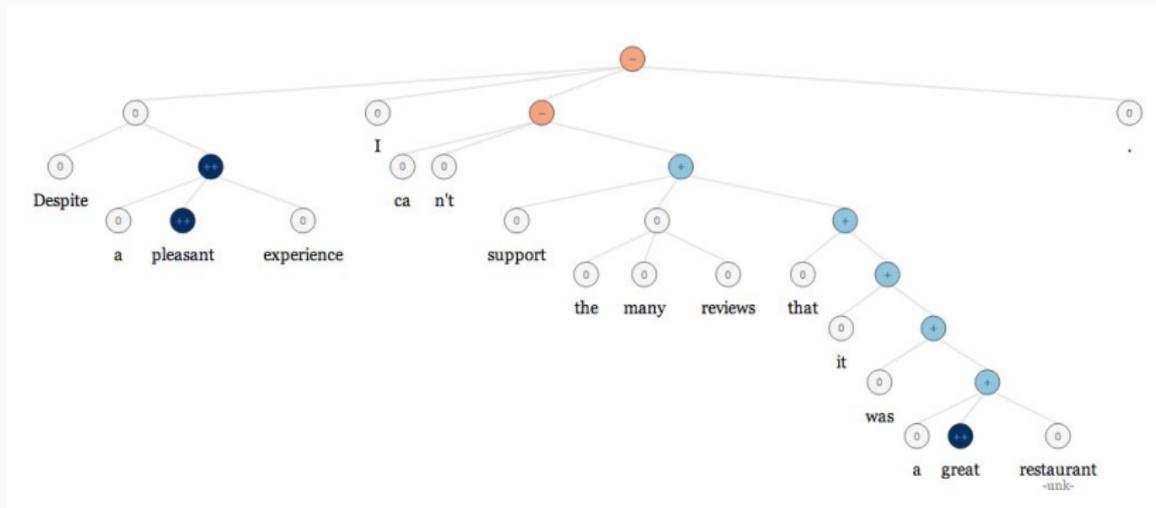


"young girl in pink shirt is swinging on swing."



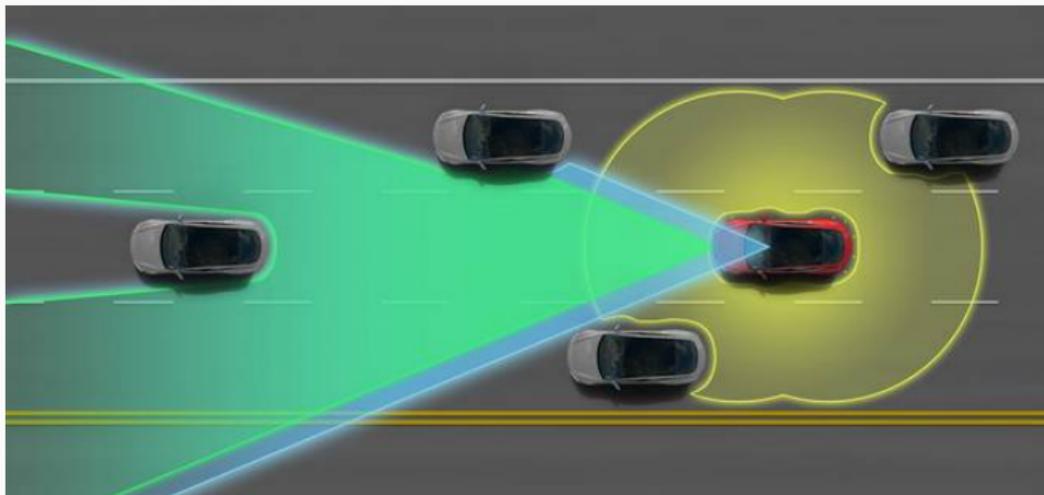
"man in blue wetsuit is surfing on wave."

# Анализ тональности текста





# Беспилотный автомобиль



Какая математика  
понадобится

---

# Какая математика понадобится?

– Линейная алгебра

# Какая математика понадобится?

- Линейная алгебра
- Теория вероятности

# Какая математика понадобится?

- Линейная алгебра
- Теория вероятности
- Математическая статистика

# Какая математика понадобится?

- Линейная алгебра
- Теория вероятности
- Математическая статистика
- Дискретная математика

# Какая математика понадобится?

- Линейная алгебра
- Теория вероятности
- Математическая статистика
- Дискретная математика
- Методы оптимизации

Кто уже использовал методы  
машинного обучения?

---

Достаточно ли знать  
алгоритмы ML и математику?

---

# Что еще надо понимать

- Когда надо применять ML

## Что еще надо понимать

- Когда надо применять ML
- Как сформулировать задачу в терминах ML

## Что еще надо понимать

- Когда надо применять ML
- Как сформулировать задачу в терминах ML
- Как выбрать подходящий класс алгоритмов

## Что еще надо понимать

- Когда надо применять ML
- Как сформулировать задачу в терминах ML
- Как выбрать подходящий класс алгоритмов
- Где посмотреть существующие решения

## Что еще надо понимать

- Когда надо применять ML
- Как сформулировать задачу в терминах ML
- Как выбрать подходящий класс алгоритмов
- Где посмотреть существующие решения
- Как настроить алгоритм

## Что еще надо понимать

- Когда надо применять ML
- Как сформулировать задачу в терминах ML
- Как выбрать подходящий класс алгоритмов
- Где посмотреть существующие решения
- Как настроить алгоритм
- Как оценить результаты

# Немного истории

---

1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"
- 1986 David Rumelhart и Ronald Williams заново открыт и популяризирован алгоритм обратного распространения ошибки

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"
- 1986 David Rumelhart и Ronald Williams заново открыт и популяризирован алгоритм обратного распространения ошибки
- 1997 Компьютер Deep Blue обыграл чемпиона мира по шахматам Гарри Каспарова

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"
- 1986 David Rumelhart и Ronald Williams заново открыт и популяризирован алгоритм обратного распространения ошибки
- 1997 Компьютер Deep Blue обыграл чемпиона мира по шахматам Гарри Каспарова
- 2006 Geoffrey Hinton ввел в обиход термин «Deep learning»

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"
- 1986 David Rumelhart и Ronald Williams заново открыт и популяризирован алгоритм обратного распространения ошибки
- 1997 Компьютер Deep Blue обыграл чемпиона мира по шахматам Гарри Каспарова
- 2006 Geoffrey Hinton ввел в обиход термин «Deep learning»
- 2011 Суперкомпьютер IBM Watson одержал победу в телевикторине Jeopardy!

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"
- 1986 David Rumelhart и Ronald Williams заново открыт и популяризирован алгоритм обратного распространения ошибки
- 1997 Компьютер Deep Blue обыграл чемпиона мира по шахматам Гарри Каспарова
- 2006 Geoffrey Hinton ввел в обиход термин «Deep learning»
- 2011 Суперкомпьютер IBM Watson одержал победу в телевикторине Jeopardy!
- 2014 Facebook изобрел алгоритм DeepFace для распознавания лиц

- 1955 Arthur Samuel создает первую самообучающуюся шашечную программу для IBM 701
- 1958 Frank Rosenblatt создает первую искусственную нейронную сеть
- 1963 Larry Roberts сформулировал тезисы компьютерного зрения
- 1973 James Lighthill. "Искусственный интеллект: Общий обзор"
- 1986 David Rumelhart и Ronald Williams заново открыт и популяризирован алгоритм обратного распространения ошибки
- 1997 Компьютер Deep Blue обыграл чемпиона мира по шахматам Гарри Каспарова
- 2006 Geoffrey Hinton ввел в обиход термин «Deep learning»
- 2011 Суперкомпьютер IBM Watson одержал победу в телевикторине Jeopardy!
- 2014 Facebook изобрел алгоритм DeepFace для распознавания лиц
- 2016 Программа AlphaGo обыграла чемпиона мира по игре в го Lee Se-dol в четырех партиях из пяти

– Проблема комбинаторного взрыва

- Проблема комбинаторного взрыва
- Низкая производительность компьютеров

- Проблема комбинаторного взрыва
- Низкая производительность компьютеров
- Проблема представлений знаний “здорового мысла”

- Проблема комбинаторного взрыва
- Низкая производительность компьютеров
- Проблема представлений знаний “здорового мысла”
- Парадокс Моравека

# Типы машинного обучения

---

1. С учителем
2. Без учителя
3. С частичным привлечением учителя
4. Обучение с подкреплением
5. Активное обучение

$X$  - множество объектов

$Y$  - множество ответов

Обучающая выборка:  $X^l = (x_i, y_i)_{i=1}^l$

**Задача:** Построить алгоритм  $a: X \rightarrow Y$

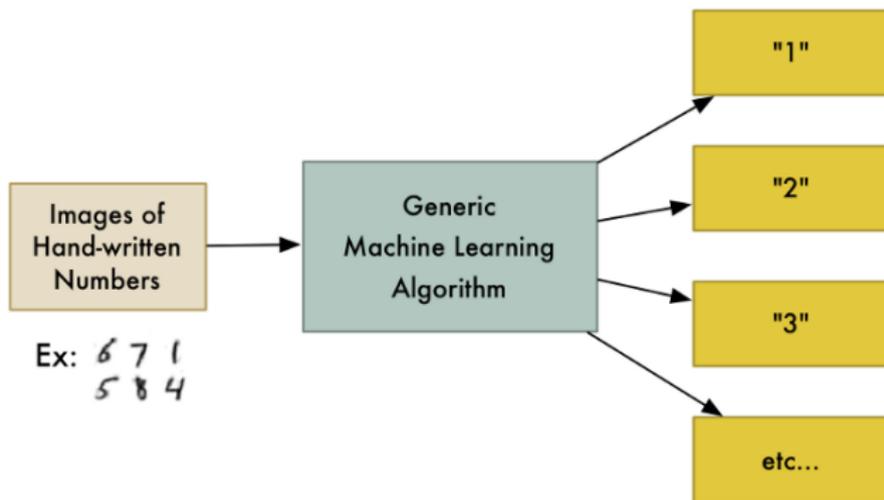
$X$  - множество объектов

$Y$  - множество классов

Обучающая выборка:  $X^l = (x_i, y_i)_{i=1}^l$

**Задача:** Построить алгоритм  $a: X \rightarrow Y$ , способный классифицировать произвольный объект  $x \in X$ .

# Обучение с учителем. Классификация



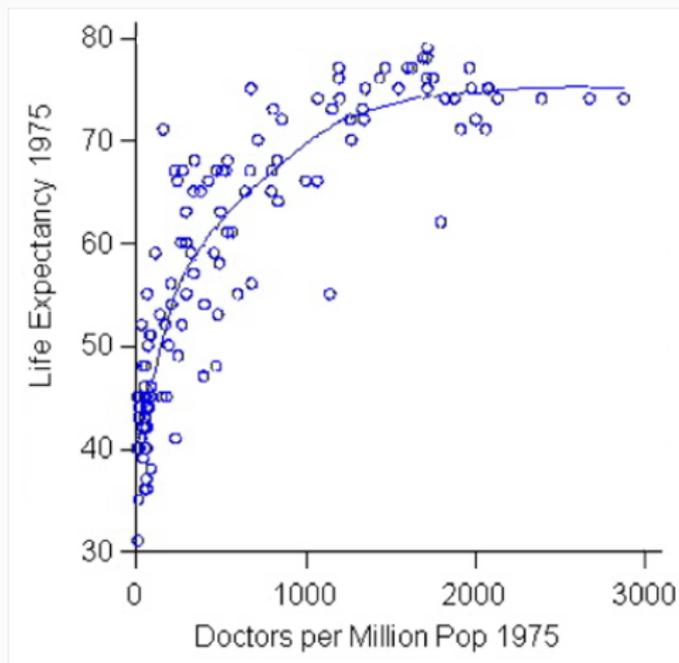
$X$  - множество объектов

$Y$  - множество ответов ( $Y \in \mathbb{R}$ )

Обучающая выборка:  $X^l = (x_i, y_i)_{i=1}^l$

Существует неизвестная целевая зависимость  $y^* : X \rightarrow Y$

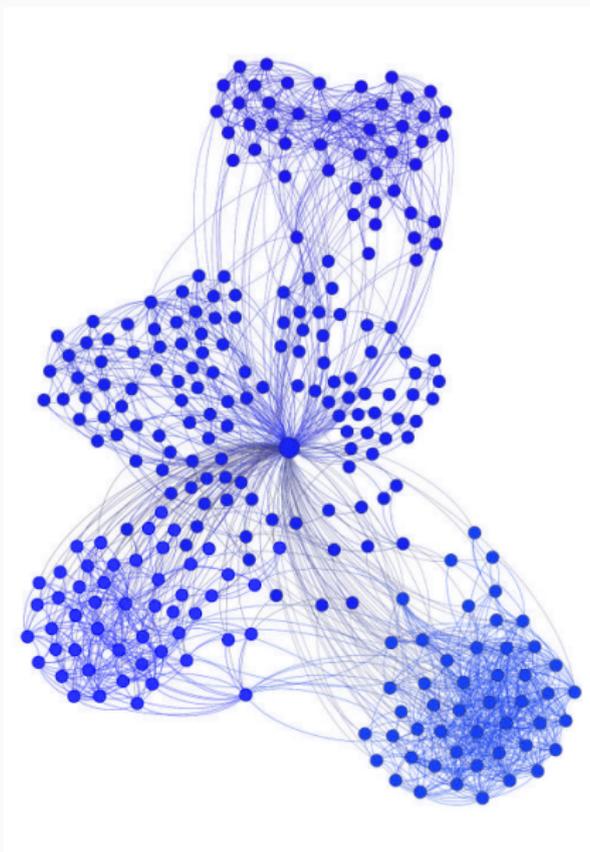
**Задача:** Построить алгоритм  $a : X \rightarrow Y$ , аппроксимирующий целевую зависимость  $y^*$



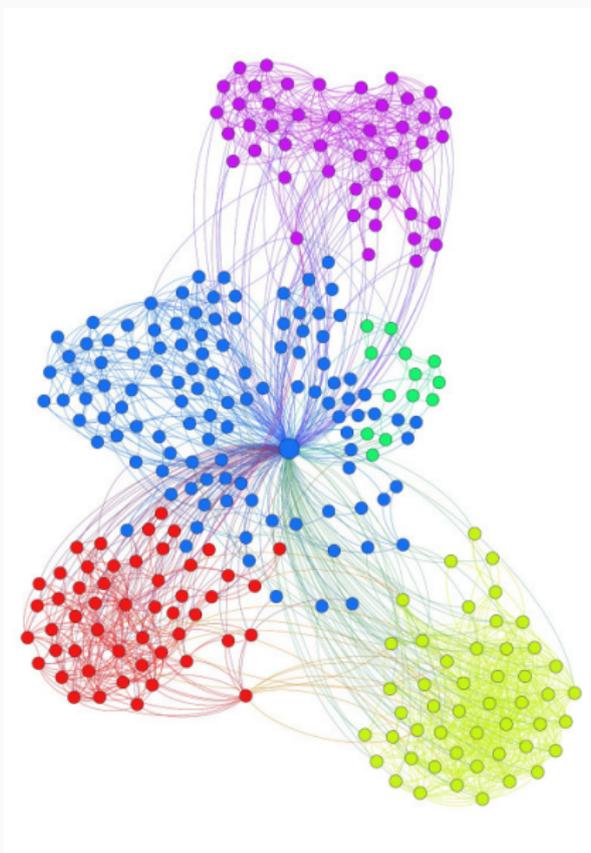
# Обучение без учителя

---

# Обучение без учителя



# Обучение без учителя



$X$  - множество объектов

Обучающая выборка:  $X^l = \{x_i\}_{i=1}^l$

**Задача:** Обнаружить внутренние взаимосвязи, зависимости, закономерности, существующие между объектами.

# Обучение с частичным привлечением учителя

---

$X$  - множество объектов

$Y$  - множество ответов

Обучающая выборка:  $X^l = (x_i, y_i)_{i=1}^l$ ,  $X^u = \{x_i\}_{i=1}^u$

**Задача:** Построить алгоритм  $a: X \rightarrow Y$

# Обучение с подкреплением

---

$S$  – множество состояний окружения

$A$  – множество действий агента

$R$  – множество "выигрышей"

**Задача:** Построить алгоритм  $a: S \rightarrow A$ , который максимизирует общий "выигрыш"



# Активное обучение

---

$X$  - множество объектов

$Y$  - множество классов

Обучающая выборка:  $X^l = (x_i, y_i)_{i=1}^l$

Получение дополнительных ответов  $y_i$  **дорого**.

**Задача:** Построить алгоритм  $a: X \rightarrow Y$ , используя как можно меньше дополнительных элементов.



# Типы признаков

---

$$f : X \rightarrow D_f$$

– Бинарные ( $D_f = \{0, 1\}$ )

$$f : X \rightarrow D_f$$

- Бинарные ( $D_f = \{0, 1\}$ )
- Номинальные ( $D_f$  – конечное множество)

$$f : X \rightarrow D_f$$

- Бинарные ( $D_f = \{0, 1\}$ )
- Номинальные ( $D_f$  – конечное множество)
- Порядковые ( $D_f$  – конечное упорядоченное множество)

$$f : X \rightarrow D_f$$

- Бинарные ( $D_f = \{0, 1\}$ )
- Номинальные ( $D_f$  – конечное множество)
- Порядковые ( $D_f$  – конечное упорядоченное множество)
- Количественные ( $D_f = \mathbb{R}$ )

## Типы признаков (примеры)

- Бинарные (Пол, наличие боли в спине, в сознании ли пациент)

## Типы признаков (примеры)

- Бинарные (Пол, наличие боли в спине, в сознании ли пациент)
- Номинальные (Тип боли: колющая, режущая, ноющая)

## Типы признаков (примеры)

- Бинарные (Пол, наличие боли в спине, в сознании ли пациент)
- Номинальные (Тип боли: колющая, режущая, ноющая)
- Порядковые (Общее состояние больного: удовлетворительное, средней тяжести, тяжелое, крайне тяжелое)

## Типы признаков (примеры)

- Бинарные (Пол, наличие боли в спине, в сознании ли пациент)
- Номинальные (Тип боли: колющая, режущая, ноющая)
- Порядковые (Общее состояние больного: удовлетворительное, средней тяжести, тяжелое, крайне тяжелое)
- Количественные (Температура тела, пульс, артериальное давление)

– Академическое

- Академическое
- Практическое

Курс К. В. Воронцова

Видеолекции ШАД

Christopher M. Bishop "Pattern Recognition and Machine Learning" G.

James, D. Witten, T. Hastie, R. Tibshirani: "An Introduction to  
Statistical Learning"

Вопросы?

## На следующей лекции

- Метод ближайших соседей
- Гипотеза компактности
- Обобщенный метрический классификатор
- Проклятие размерности
- Отбор эталонов