# РАСПОЗНАВАНИЕ ОБРАЗОВ

**ЭКЗАМЕН:**

**Допуск:** метод ближайшего соседа + индивидуальное задание
**Экзамен:** 2 вопроса + задача

## Вопросы

1. Распознавание образов и машинное обучение. Задачи распознавания образов. Математические модели. Вектор признаков. Факторы, затрудняющие построение систем распознавания образов.

2. Метод ближайшего соседа. Применение для задачи распознавания растровых изображений. Варианты используемых мер близости. Выбросы и прототипы.

3. Проклятие размерности и построение алгоритмов решения задач распознавания образов. Примеры влияния на сложность решаемых задач.

4. Распределение ошибок. Нормальное распределение. Генерирование нормально распределённой случайной величины.

5. Метод наименьших квадратов распознавания функциональной зависимости.

6. Переобучение. Регуляризация. Кросс-валидация. Параметры кросс-валидации и варианты её реализации.

7. Принцип максимального правдоподобия. Метод наименьших квадратов как результат применения принципа максимального правдоподобия для распознавания функциональной зависимости при нормальном распределении ошибки.

8. Байесовская теория решений. Байесовское решающее правило. Распознавание с минимизацией ошибки при заданных априорных и условных вероятностях. Условный риск. Функция издержек. Распознавание с минимизацией риска.

9. Разделяющие функции. Многомерное нормальное распределение и разделяющие функции для его частных случаев.

10. Использование нейронных сетей для решения задач распознавания и классификации. Перцептрон. Теорема о сходимости алгоритма обучения перцептрона.

11. Многослойный перцептрон. Обучение нейронной сети с помощью метода наискорейшего спуска.

12. Метод обратного распространения ошибки. Вычислительная сложность обучения нейронных сетей.

# Задачи

1. Докажите

   **Утверждение.** *Пусть $\xi$ распределена равномерно на отрезке $[a, b]$, а $\mathsf{F}(x)$ — некоторая функция распределения. Тогда случайная величина $\mathsf{F}(\xi)^{-1}$ распределена в соответствии с $\mathsf{F}$.*

2. Докажите формулу Байеса. Покажите, что сумма апостериорных вероятностей всех предпосылок, рассчитанных по ней, равна 1.
3. Докажите, что случайная величина $\eta$ распределённая нормально с параметрами $\mu$ и $\sigma$ может быть получена из случайной величины со стандартным нормальным распределением $\xi$ по следующей формуле:
   $$\eta = \mu + \sigma\xi.$$
4. Покажите, что для $N$ равномерно распределённых на $[0, 1]$ случайных величин $\xi_i$ имееем
   $$\mathsf{M}S_n = \frac{1}{2}N, \quad \mathsf{D}S_n = \frac{1}{12}N.$$
5. Покажите, что если признаки $\xi_i$ и $\xi_j$ независимы, то $\sigma_{ij} = 0$.
6. Нормальная плотность распределения значений для случайной величины, представляемой случайным $m$-мерным вектором признаков $x$, задаётся формулой
   $$\mathsf{p}(x) = \frac{1}{(2\pi)^{m/2}\det(\Sigma)^{1/2}}\exp\bigl[-\frac{1}{2}(x-\mu)^{\top}\Sigma^{-1}(x-\mu)\bigr], \qquad (1)$$
   где $\mu$ — вектор средних значений признаков: $\mu_i = \mathsf{M}x_i$, $\Sigma = (\sigma_{ij})$ — *ковариационная матрица* вектора признаков. Покажите, что если все признаки взаимно независимы, то плотность $\mathsf{p}(x)$ равна произведению одномерных нормальных плотностей компонент $x$.
7. Пусть для прогнозирования типа потока* требуется иметь $M$ значений векторов признаков $x^{(i)} \in \mathbb{R}^{20}$ в одной ячейке регулярной сетки. Каков должен быть размер выборки $N$ для решения задачи классификации вектора признаков $x$ предложенным выше способом, если известно, что любое возможное значение $x$ принадлежит $[a_1, b_1] \times \ldots \times [a_{20}, b_{20}]$?

8. (*Задача — один из вариантов а) – г).*) Предположим, что нам необходимо оценить плотность распределения некоторой случайной величины в единичном гиперкубе в $\mathbb{R}^m$, основываясь на выборке из $N$ элементов.
   а) Покажите, что почти все точки выборки находятся около границы $m$-мерного гиперкуба, используя для этого подход, применённый при рассмотрении $m$-мерной сферы.
   б) Найдите $l_m(p)$ — длину ребра гиперкуба в $m$-мерном пространстве, который содержит долю $p$ точек от всех точек выборки $(0 \leq p \leq 1)$. Оцените $l_5(0.01)$, $l_5(0.1)$, $l_{20}(0.01)$, $l_{20}(0.1)$.
   в) Покажите, что почти все евклидовы расстояния между точками выборки являются достаточно большими и примерно одинаковыми в $\mathbb{R}^m$, и что их окрестности, для того, чтобы содержать даже небольшое количество точек выборки, должны иметь большой радиус.
   г) Покажите, что расстояния Чебышёва между почти всеми точками выборки достаточно велики.
9. При решении задачи линейной регрессии мы получаем следующую систему уравнений.

$$\begin{cases} \left(\sum\limits_{i=1}^{N} x_i^2\right)A + \left(\sum\limits_{i=1}^{N} x_i\right)B = \sum\limits_{i=1}^{N} t_i x_i, \\ \left(\sum\limits_{i=1}^{N} x_i\right)A + NB = \sum\limits_{i=1}^{N} t_i. \end{cases} \tag{2}$$

   Решая систему уравнений (2), находим значения $A_0$ и $B_0$, дающие минимум $E(A, B)$.

   Укажите, когда разрешима система уравнений (2), и объясните, почему решение единственно в этом случае. Какой выборке соответствует ситуация, когда определитель системы уравнений равен нулю?
10. Найдите коэффициенты $w$, минимизирующие функцию ошибки $E(w)$ в случае полиномиальной регрессии.
11. Важен ли порядок, в котором пронумерованы векторы признаков из обучающей выборки для алгоритма обучения перцептрона, предложенного в доказательстве теоремы о сходимости перцептрона?
12. Доказательство сходимости алгоритма обучения перцептрона было

проведено нами в предположении, что $x(n) \in \mathcal{H}_1$ и классификация производится неверно на всех итерациях вплоть до $(n+1)$-ой. Покажите, что при других возможных результатах работы алгоритма обучения перцептрона к $(n+1)$-ой итерации (имеются итерации, на которых классификация производится верно, выбираемые из выборки в соответствии с введённой нумерацией векторы признаков принадлежат классу $\mathcal{H}_2$), проведённые нами рассуждения достаточны для того, чтобы утверждать о сходимости алгоритма обучения.

13. Покажите, что если в качестве функции активации используется линейная функция, то преобразование, задаваемое такой нейронной сетью будет линейным.
14. Покажите, что в случае использования линейных функций активации, если количество скрытых узлов нейронной сети меньше количества входов или выходов сети, то, линейное преобразование, задаваемое такой нейронной сетью, имеет меньшую размерность, чем чем размерность входа или выхода.

## Индивидуальные задания

1. Задание 3, стр. 42. Определение выбросов и прототипов.
2. МНК. Кросс-валидация параметров модели — параметры $M$, $\lambda$. Сравнение результатов кросс-валидации для разбиений исходной выборки $\{x_i, f(x_i)\}_{i=1}^N$ на подвыборки:
   а) в случайном порядке,
   б) в порядке возрастания $x_i$,
   в) свой вариант.
3. МНК. Кросс-валидация параметров модели — параметры $M$, $\lambda$. Сравнение результатов кросс-валидации для разбиений исходной выборки $\{x_i, f(x_i)\}_{i=1}^N$ на подвыборки:
   а) размер блока равен 1,
   а) размер блока больше 1 — перебрать варианты,
   а) свой вариант.
4. МНК. Кросс-валидация параметров модели — параметры $M$, $\lambda$. Сравнение результатов кросс-валидации для разбиений исходной выборки $\{x_i, f(x_i)\}_{i=1}^N$ на подвыборки:
   а) блоки одной мощности,
   б) блоки формируются исходя из разбиения интервала
   $$[\min\{x_i\}, \max\{x_i\}]$$
   на равные подинтервалы,
   в) свой вариант.
5. Нейронные сети. Кросс-валидация параметров модели — количество слоёв в перцептроне, количество нейронов в слое. Сравнение результатов кросс-валидации для разбиений исходной выборки $\{x_i, f(x_i)\}_{i=1}^N$ на подвыборки:
   а) в случайном порядке, б) в порядке возрастания $x_i$, в) свой вариант.
6. Нейронные сети. Кросс-валидация параметров модели — количество слоёв в перцептроне, количество нейронов в слое. Сравнение результатов кросс-валидации для разбиений исходной выборки $\{x_i, f(x_i)\}_{i=1}^N$ на подвыборки:
   а) размер блока равен 1,
   б) размер блока больше 1 — перебрать варианты,
   в) свой вариант.

7. Нейронные сети. Кросс-валидация параметров модели — количество слоёв в перцептроне, количество нейронов в слое. Сравнение результатов кросс-валидации для разбиений исходной выборки $\{x_i, f(x_i)\}_{i=1}^{N}$ на подвыборки:
   а) блоки одной мощности,
   б) блоки формируются исходя из разбиения интервала

   $$[\min\{x_i\}, \max\{x_i\}]$$

   на равные подинтервалы,
   в) свой вариант.
8. Нейронные сети. Использование различных функций активации: логистический сигмоид, гиперболический тангенс, ReLU. Исследование влияния выбора функции активации на скорость обучения.