

# Анализ свойств вероятностных моделей в задачах обучения с экспертом

А. И. Базарова<sup>1</sup>, А. В. Грабовой<sup>2</sup>, В. В. Стрижов<sup>3</sup>

**Аннотация:** Данная работа посвящена построению интерпретируемых моделей в машинном обучении. Решается задача аппроксимации заданного набора фигур на контурном изображении. Вводятся предположения о том, что фигуры являются кривыми второго порядка. При аппроксимации фигур используется информация о типе, расположению и виде кривых, а также о множестве их возможных преобразований. Такая информация называется *экспертной*, а метод машинного обучения, основанный на экспертной информации, называется *обучением с экспертом*. Предполагается, что набор фигур аппроксимируется набором *локальных моделей*. Каждая локальная модель на основе экспертной информации аппроксимирует одну фигуру на контурном изображении. Для построения моделей предлагается отобразить кривые второго порядка в признаковое пространство, в котором каждая локальная модель является линейной моделью. Таким образом, кривые высших порядков распознаются при помощи композиции линейных моделей. В качестве прикладной задачи рассматривается задача аппроксимации радужки глаза на контурном изображении.

**Ключевые слова:** смесь экспертов, обучение с экспертом, линейные модели, интерпретируемые модели **DOI:** 00.00000/000000000000000

## 1 Введение

Построения интерпретируемых моделей в машинном обучении [1] является одной из ключевых проблем. Современные решения задачи классификации изображений на основе сетей глубокого обучения ResNet, VGG, Intercept [20] являются плохо интерпретируемыми моделями. В работах [21, 22, 23] показано, что сети глубокого обучения

---

<sup>1</sup>Московский физико-технический институт, bazarova.ai@phystech.edu

<sup>2</sup>Московский физико-технический институт, grabovoy.av@phystech.edu

<sup>3</sup>Московский физико-технический институт, strijov@ccas.ru

чувствительны даже к малому шуму в данных, что связано с их неинтерпретируемостью.

В данной работе предлагается метод *обучения с экспертом*. Данный метод предполагает использование предметных знаний экспертов для повышения качества аппроксимации, а также для получения интерпретируемых моделей машинного обучения. Предметные знания экспертов о выборке назовем *экспертной информацией*. Предполагается, что использование экспертной информации позволяет аппроксимировать выборку простыми интерпретируемыми моделями, например линейными моделями. Методы машинного обучения, которые учитывают экспертные знания при построении моделей, назовем *обучением с экспертом*.

В работе решается задача аппроксимации кривых второго порядка на контурном изображении. Кривые второго порядка выбраны для анализа, так как они легко описываются линейными моделями. При этом данные фигуры требуется восстановить в таких прикладных задачах, как задача распознавания радужки глаза [14, 15, 16], задача описания трека частицы в адронном коллайдере [24]. Экспертная информация о кривой второго порядка позволяет отобразить точки на плоскости в новое признаковое описание, в котором каждая кривая аппроксимируется одной линейной моделью. Модель, которая аппроксимирует одну кривую называется *локальной моделью*. Для аппроксимации всего контурного изображения требуется аппроксимировать несколько кривых второго порядка при помощи нескольких локальных моделей. В данной работе вводятся следующие ограничения на изображения:

- 1) изображение состоит только из кривых второго порядка;
- 2) изображение аппроксимируется малым числом кривых второго порядка;
- 3) число и тип кривых на изображении известно.

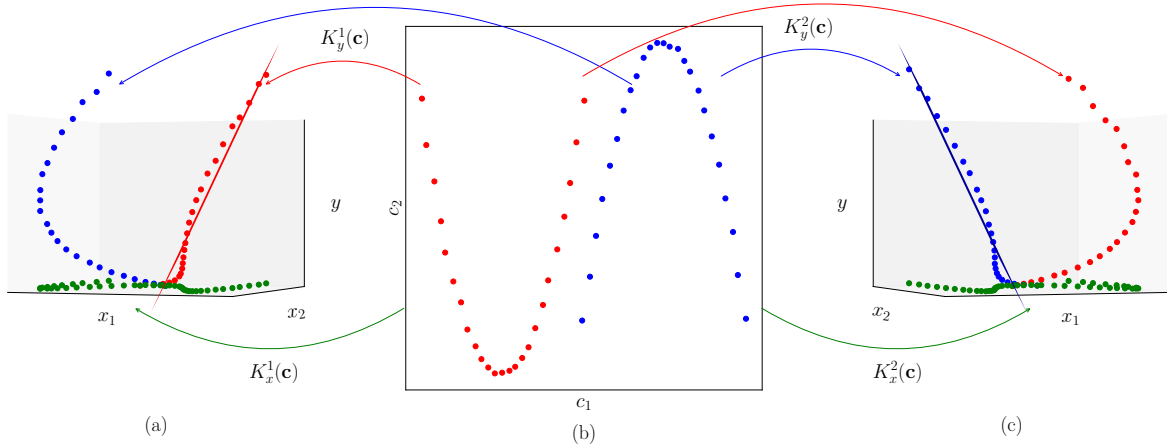


Рис. 1: Пример: а) экспертной информации первого эксперта; б) исходные данные; в) экспертная информация второго эксперта.

На рис. 1 показан пример кривых второго порядка, а также экспертная информация о кривых. Рис. 1а показывает экспертную информацию первого эксперта. При использовании данной информации первая кривая аппроксимируется линейной моделью, а вторая кривая является шумом. Рис. 1б показывает экспертную информацию второго эксперта. При использовании данной информации вторая кривая аппроксимируется линейной моделью, а первая кривая является шумом.

При аппроксимации нескольких кривых на одном контурном изображении строится мультимодель. Примером мультимodelей является случайный лес [3], бустинг деревьев [2], смесь экспертов [4]. В данной работе в качестве мультимodelи рассматривается смесь экспертов. Смесь экспертов — это мультимodelь, которая линейно взвешивает локальные модели, которые аппроксимируют часть выборки. Значения весовых коэффициентов зависят от того объекта, для которого производится предсказание. Для решения задачи смеси экспертов используется вариационный ЕМ-алгоритм [6, 19, 17]. Смесь экспертов имеет множество применений в ряде прикладных задач. В работе [11] решается задача классификации текстов. В работах [8, 9, 5, 12, 13, 18, 7] смесь экспертов используется для предсказания временных рядов в задачах распознавания речи, дневной активности человека и предсказания стоимости ценных бумаг. В работе [10] смесь экспертов рассматривался для решения задачи распознавания рукописных цифр на изображениях.

В качестве примера рассматривается задача аппроксимации изображения радужки глаза. На рис. 2а показан пример изображения, которое требуется аппроксимировать. В данной работе рассматривается обработанное изображение, которое приведено в контурный вид, пример такого изображения показан на рис. 2б. На рис. 2б показаны две локальные модели окружностей, которые аппроксимируют радужку глаза. Окружности являются простым примером кривой второго порядка.

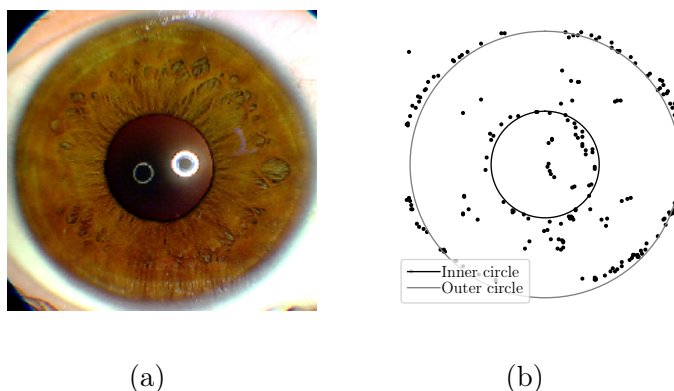


Рис. 2: Пример изображения радужки глаза и его контурного представления: а) изображение радужки глаза; б) контурное изображение радужки и аппроксимирующие данное изображение окружности.

Для задачи аппроксимации радужки глаза используется следующая экспертная информация: радужка глаза аппроксимируется двумя концентрическими окружно-

стями. Экспертная информация используется при построении признакового описания точек плоскости, а также при построении функции оптимизации. Часть функции ошибки для оптимизации, которая использует экспертную информации называется регуляризатором. Таким образом информация о том, что на изображении окружности задается признаковым описанием, а информация о том, что окружности концентрические задается при помощи специального регуляризатора.

В вычислительном эксперименте проводится анализ качества аппроксимации контурного изображения в зависимости от заданной экспертной информации и от уровня шума в синтетически сгенерированных данных. Проводится анализ качества аппроксимации радужки глаза в зависимости от количества экспертной информации, которая использовалась при построении модели. Заметим, что каждое аппроксимируемое изображение это отдельный набор точек, которые требуется аппроксимировать.

## 2 Постановка задачи нахождения параметров кривых второго порядка на изображении

Задано бинарное изображение:

$$\mathbf{M} \in \{0, 1\}^{m_1 \times m_2},$$

где 1 отвечает черной точке изображения, а 0 — белой точке фона. По изображению  $\mathbf{M}$  строится выборка  $\mathbf{C}$ , элементами которой являются координаты  $(x_i, y_i)$  черных точек:

$$\mathbf{C} \in \mathbb{R}^{N \times 2}.$$

Эксперт предполагает, что изображение состоит из кривой второго порядка  $\Omega$ . Пусть для набора точек  $\mathbf{C} \in \mathbb{R}^{N \times 2}$ , образующих кривую  $\Omega$ , задано экспертная информация о фигуре  $E(\Omega)$ . Множество  $E(\Omega)$  состоит из ожидаемого экспертом вида фигуры  $\Omega$  и множества ее допустимых преобразований. На основе экспертного описания введем отображения в новую задачу для аппроксимации:

$$K_x(E(\Omega)) : \mathbb{R}^2 \rightarrow \mathbb{R}^n, \quad K_y(E(\Omega)) : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (2.1)$$

где  $K_x$  отображение объекты в признаковое описание объектов,  $n$  — число признаков, а  $K_y$  отображение в целевую переменную для объекта. Применив отображения  $K_x, K_y$  для выборки  $\mathbf{C}$  поэлементно получаем:

$$K_x(E(\Omega), \mathbf{c}) = \mathbf{x}, \quad K_y(E(\Omega), \mathbf{c}) = y, \quad (2.2)$$

где  $\mathbf{c} = (x_i, y_i)$  — точка из выборки  $\mathbf{C}$ .

Применяя отображения (2.2) к исходному набору точек  $\mathbf{C}$ , получим выборку

$$\mathcal{D} = \{(\mathbf{x}, y) \mid \forall \mathbf{c} \in \mathbf{C} \mathbf{x} = K_x(\mathbf{c}), y = K_y(\mathbf{c})\}. \quad (2.3)$$

Получаем, что исходная задача аппроксимации кривой  $\Omega$  сводится к аппроксимации выборки  $\mathfrak{D}$ . В данной работе предполагается, что выборка  $\mathfrak{D}$  аппроксимируется линейной моделью:

$$g(\mathbf{x}, \mathbf{w}) = \mathbf{x}^\top \mathbf{w}, \quad (2.4)$$

где  $\mathbf{w}$  вектор, параметр, который требуется найти.

Для поиска оптимального вектора параметров  $\hat{\mathbf{w}}$  требуется решить следующую оптимизационную задачу:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathbb{R}^n} \sum_{(\mathbf{x}, y) \in \mathfrak{D}} \|g(\mathbf{x}, \mathbf{w}) - y\|_2^2. \quad (2.5)$$

Таким образом, задача аппроксимации исходной кривой  $\Omega$  сводится к решению задачи линейной регрессии, т. е. нахождения компонент вектора  $\hat{\mathbf{w}}$ , связывающего полученные  $\mathbf{x}$  и  $y$ .

В случае, когда на изображении  $K$  кривых второго порядка  $\Omega_1, \dots, \Omega_K$ , для каждой из которых имеется экспертная информация  $E_k = E(\Omega_k)$ ,  $k \in \{1, \dots, K\}$ , ставится задача построения мультимодели, называемой смесью  $K$  экспертов.

**Определение 2.1.** Назовем мультимодель  $f$  смесью  $K$  экспертов

$$f = \sum_{k=1}^K \pi_k(\mathbf{x}, \mathbf{V}) g_k(\mathbf{w}_k), \quad \pi_k(\mathbf{x}, \mathbf{V}) : \mathbb{R}^{n \times |\mathbf{V}|} \rightarrow [0, 1], \quad \sum_{k=1}^K \pi_k(\mathbf{x}, \mathbf{V}) = 1, \quad (2.6)$$

где  $g_k$  — локальная модель, называемая экспертом,  $\mathbf{x}$  — признаковое описание объекта,  $\pi_k$  — шлюзовая функция, вектор  $\mathbf{w}_k$  — параметры локальной модели, вектор  $\mathbf{V}$  — параметры шлюзовой функции. В данной работе  $g_k$  — линейная модель.

Для каждой кривой второго порядка заданы отображения (2.1). Для удобства введем следующее обозначение:  $K_x^k(\mathbf{c}) = K_x(\Omega_k, \mathbf{c})$  и  $K_y^k(\mathbf{c}) = K_y(\Omega_k, \mathbf{c})$ . Тогда, используя локальные линейные модели, построим универсальную мультимодель, описывающую кривые  $\Omega_1, \dots, \Omega_K$  на изображении  $\mathbf{M}$ :

$$f = \sum_{\mathbf{c} \in \mathbf{C}} \sum_{k=1}^K \pi_k(\mathbf{c}, \mathbf{V}) g_k(K_x^k(\mathbf{c}), \mathbf{w}_k), \quad (2.7)$$

где  $\pi_k$  — шлюзовая функция. В данной работе рассматривается простой случай, когда  $\mathbf{x} = K_x^1(\mathbf{c}) = \dots = K_x^K(\mathbf{c})$ , тогда выражение (2.7) переписывается в следующем простом виде:

$$f = \sum_{\mathbf{c} \in \mathbf{C}} \sum_{k=1}^K \pi_k(\mathbf{x}, \mathbf{V}) g_k(\mathbf{x}, \mathbf{w}_k), \quad (2.8)$$

где шлюзовая функция  $\pi_k$  имеет следующий вид:

$$\pi_k(\mathbf{x}, \mathbf{V}) : \mathbb{R}^{n \times |\mathbf{V}|} \rightarrow [0, 1], \quad \sum_{k=1}^K \pi_k(\mathbf{x}, \mathbf{V}) = 1, \quad (2.9)$$

где  $\mathbf{V}$  — параметры плюсовой функции, а  $g_k$  — локальная модель.

В данной работе

$$\pi(\mathbf{x}, \mathbf{V}) = \text{softmax}(\mathbf{V}_1^\top \sigma(\mathbf{V}_2^\top \mathbf{x})), \quad (2.10)$$

где  $\mathbf{V} = \{\mathbf{V}_1, \mathbf{V}_2\}$  — параметры плюсовой функции,  $\mathbf{V}_1 \in \mathbb{R}^{p \times k}$ ,  $\mathbf{V}_2 \in \mathbb{R}^{n \times p}$ .

Для нахождения оптимальных параметров мультимодели необходимо решить следующую задачу оптимизации:

$$\mathcal{L} = \sum_{(\mathbf{x}, y) \in \mathfrak{D}} \sum_{k=1}^K \pi_k(\mathbf{x}, \mathbf{V})(y - \mathbf{w}_k^\top \mathbf{x})^2 + R(\mathbf{V}, \mathbf{W}, E(\Omega)) \rightarrow \min_{\mathbf{V}, \mathbf{W}}, \quad (2.11)$$

где  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_k]$  — параметры локальных моделей,  $R(\mathbf{V}, \mathbf{W}, E(\Omega))$  — регуляризация параметров, основанная на экспертной информации.

### 3 Построение признакового описания фигур

**Единое пространство для кривых второго порядка.** Произвольная кривая второго порядка, главная ось которой не параллельна оси ординат, задается следующим выражением:

$$x^2 = B'xy + C'y^2 + D'x + E'y + F',$$

где на коэффициенты  $B', C'$  накладываются ограничения, которые зависят от вида кривой. Выражение (2.2) принимает следующий вид:

$$K_x(\mathbf{c}_i) = [x_i y_i, y_i^2, x_i, y_i, 1], \quad K_y(\mathbf{c}_i) = x_i^2,$$

откуда получаем задачу линейной регрессии для восстановления параметров  $B', C', D', E', F'$  по составленной выборке.

**Окружность.** Частным случаем кривой второго порядка, рассмотрим окружность. Пусть  $(x_0, y_0)$  — центр окружности, которую необходимо найти на бинарном изображении  $\mathbf{M}$ , а  $r$  — ее радиус. Элементы выборки  $(x_i, y_i) \in \mathbf{C}$  являются геометрическим местом точек, которое аппроксимируется уравнением окружности:

$$(x_i - x_0)^2 + (y_i - y_0)^2 = r^2. \quad (3.1)$$

Раскрыв скобки, получим:

$$(2x_0) \cdot x_i + (2y_0) \cdot y_i + (r^2 - x_0^2 - y_0^2) \cdot 1 = x_i^2 + y_i^2. \quad (3.2)$$

Тогда отображения (2.2) примут вид:

$$K_x(\mathbf{c}_i) = [x_i, y_i, 1] = \mathbf{x}, \quad K_y(\mathbf{c}_i) = x_i^2 + y_i^2 = y. \quad (3.3)$$

Поставим задачу линейной регрессии (2.3). Компоненты вектора  $\mathbf{w} = [w_0, w_1, w_2]^\top$ , связывающего  $\mathbf{x}$  и  $y$ , восстанавливают параметры окружности:

$$x_0 = \frac{w_0}{2}, \quad y_0 = \frac{w_1}{2}, \quad r = \sqrt{w_2 + x_0^2 + y_0^2}. \quad (3.4)$$

## 4 Композиция фигур

Для построения композиции фигур воспользуемся выражением (2.11), которое принимает следующий вид:

$$\mathcal{L} = \sum_{\mathbf{c} \in \mathbf{C}} \sum_{k=1}^K \pi_k(\mathbf{c}, \mathbf{V}) \left( K_y^k(\mathbf{c}) - \mathbf{w}_k^\top K_x^k(\mathbf{c}) \right)^2 + R(\mathbf{V}, \mathbf{W}, E(\Omega)) \rightarrow \min_{\mathbf{V}, \mathbf{W}}, \quad (4.1)$$

где  $K_x^k, K_y^k$  экспертное представление  $k$ -го эксперта. Предполагая, что все кривые на изображении описываются одним признаковым описанием  $\mathbf{x} = K_x^1(\mathbf{c}) = \dots = K_x^K(\mathbf{c})$ ,  $y = K_y^1(\mathbf{c}) = \dots = K_y^K(\mathbf{c})$ , получаем следующую оптимизационную задачу:

$$\mathcal{L} = \sum_{(\mathbf{x}, y) \in \mathcal{D}} \sum_{k=1}^K \pi_k(\mathbf{x}, \mathbf{V}) \left( y - \mathbf{w}_k^\top \mathbf{x} \right)^2 + R(\mathbf{V}, \mathbf{W}, E(\Omega)) \rightarrow \min_{\mathbf{V}, \mathbf{W}}, \quad (4.2)$$

В качестве регуляризатора  $R$  рассматриваются дополнительные ограничения на вектора параметров моделей. Для решения оптимизационной задачи (4.2) предлагается использовать ЕМ-алгоритм.

## 5 Вычислительный эксперимент

Проведен вычислительный эксперимент для анализа качества моделей кривых второго порядка на изображении. Эксперимент разделен на несколько частей. В первой части проводится эксперимент с несколькими окружностями на изображении. В второй части проводится анализ сходимости метода в зависимости от уровня шума в данных и от заданной экспертной информации. В третьей части проводится эксперимент по аппроксимации радужки глаза.

### 5.1 Эксперимент с окружностями

В данной части эксперимента показан пример обучения мультимодели для аппроксимации нескольких фигур второго порядка одновременно. В качестве данных используется синтетическая выборка, которая получена при помощи генерации трех произвольных непересекающихся окружностей, а также добавления к данным окружностям шума. Шум добавлялся к радиусу окружности для каждой точки, также в выборку были добавлены случайные точки, которые не относятся к окружностям. В эксперименте сравниваются две модели: в первой модели регуляризатор  $R(\mathbf{V}, \mathbf{W}, E(\Omega)) = 0$ , то есть модель без задания регуляризатора, во второй модели регуляризатор:

$$R(\mathbf{V}, \mathbf{W}, E(\Omega)) = - \sum_{k=1}^K \gamma \left( \mathbf{w}_k - \mathbf{w}_k^0 \right)^\top \left( \mathbf{w}_k - \mathbf{w}_k^0 \right),$$

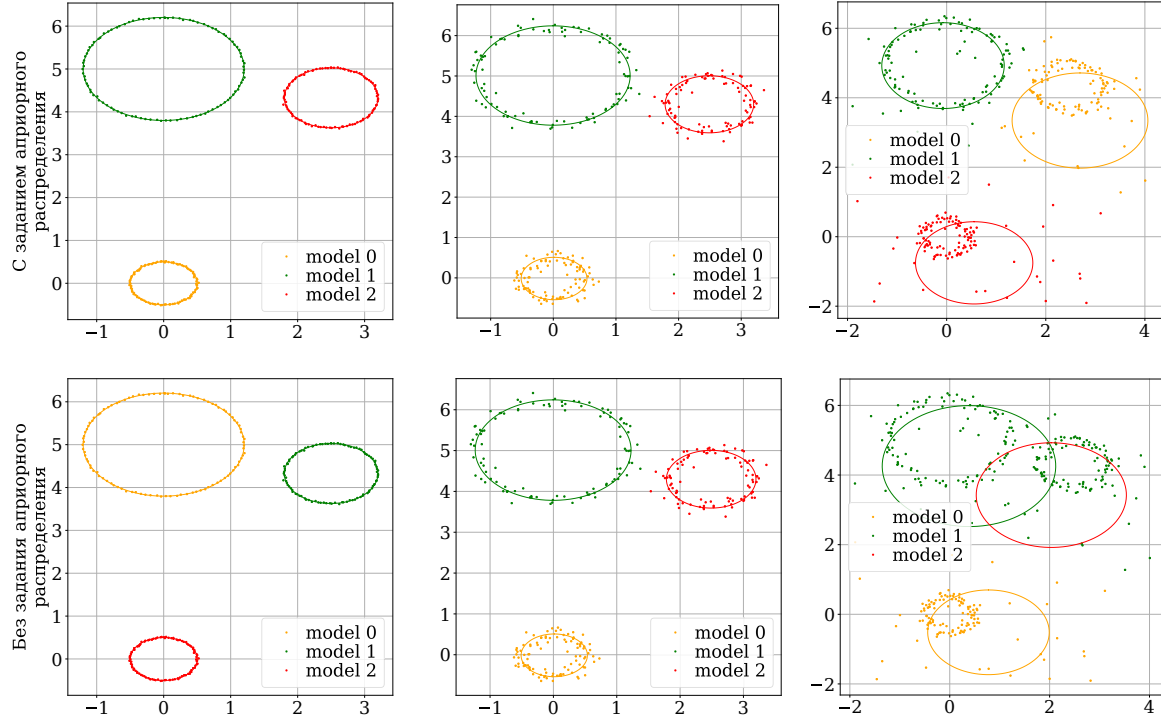


Рис. 3: Мультимодель в зависимости от разных априорных предположений и уровня шума. Сверху вниз: построение с заданием априорного распределения; без задания априорного распределения. Слева на право: окружности без шума; шум в радиусе окружности; шум в радиусе окружности а также произвольные точки по всему изображению.

где  $\mathbf{w}_k^0$  априорные предположения о векторе параметров.

На рис. 3 показан результат построения ансамбля локально аппроксимирующих моделей, которые аппроксимируют выборку. Каждая локальная модель аппроксимирует одну окружность, причем при добавления разного шума, качество аппроксимации падет. На рис. 4 показан график зависимости радиуса окружностей  $r$  и их центров  $(x_0, y_0)$  от номера итерации. Видно, что модель с заданием априорного распределения сходится быстрее чем модель без задания априорного распределения.

## 5.2 Эксперимент с разным уровнем шума и дисперсии априорного распределения

В данной части эксперимента проводится анализ качества аппроксимации  $S$  от уровня шума  $\beta$  в данных и от параметра априорных распределений  $\gamma$ . Выборка получена следующим образом: сначала случайным образом выбирается два вектора параметров  $\mathbf{w}_1^{\text{true}}$  и  $\mathbf{w}_2^{\text{true}}$  — коэффициенты двух парабол. На основе векторов  $\mathbf{w}_1^{\text{true}}$  и  $\mathbf{w}_2^{\text{true}}$  выполняется генерация точек  $x_i$  и  $y_i$  с добавлением нормального шума  $\varepsilon \sim \mathcal{N}(0, \beta)$ .



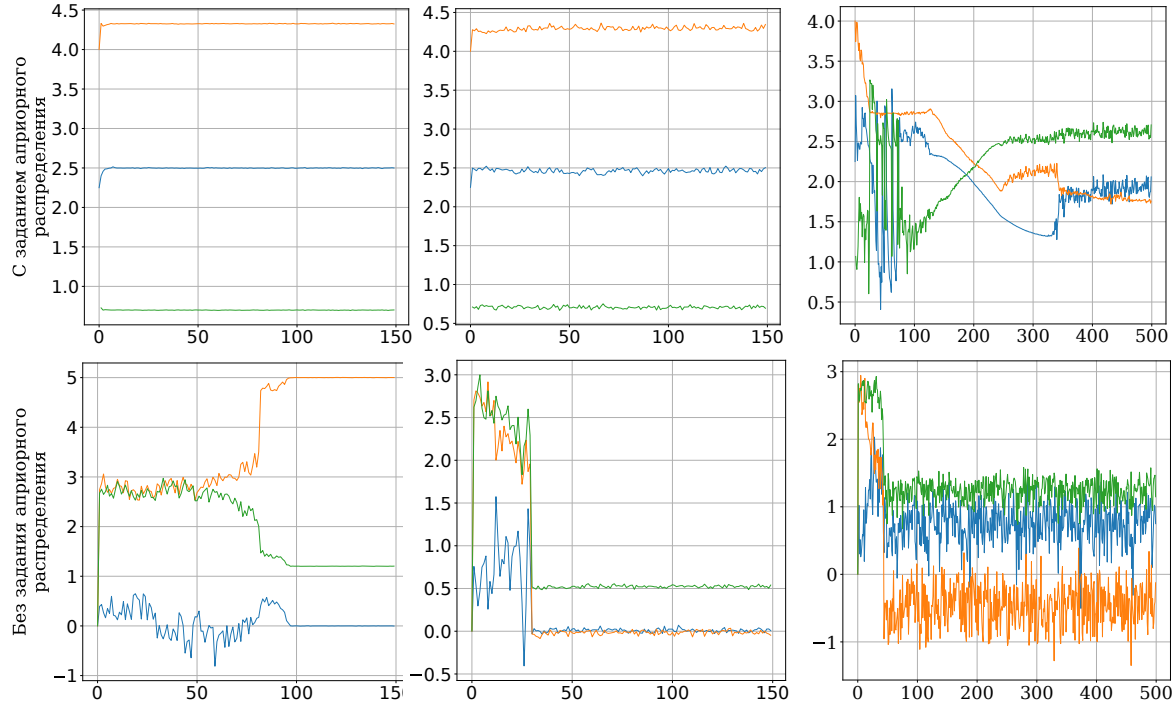


Рис. 4: Зависимость параметров  $r$ ,  $x_0$  и  $y_0$  от номера итерации при разных априорных распределениях. Сверху вниз: построение с заданием априорного распределения; без задания априорного распределения. Слева на право: окружности без шума; шум в радиусе окружности; шум в радиусе окружности а также произвольные точки по всему изображению.

При обучении мультимодели в качестве априорного распределения параметров рассматривается  $\mathbf{w}_1 \sim \mathcal{N}(\mathbf{w}_1^{\text{true}}, \gamma \mathbf{I})$ ,  $\mathbf{w}_2 \sim \mathcal{N}(\mathbf{w}_2^{\text{true}}, \gamma \mathbf{I})$ .

Рассматривается следующий критерий качества:

$$S = \|\mathbf{w}_1^{\text{pred}} - \mathbf{w}_1^{\text{true}}\|_2^2 + \|\mathbf{w}_2^{\text{pred}} - \mathbf{w}_2^{\text{true}}\|_2^2,$$

где  $\mathbf{w}_1^{\text{pred}}$  аппроксимация вектора параметров первой локальной модели, а  $\mathbf{w}_2^{\text{pred}}$  аппроксимация вектора параметров второй локальной модели.

На рис. 6 показана зависимость критерия качества  $S$  от уровня шума  $\beta$  и параметра априорного распределения  $\gamma$ . Из графика видно, что при малом уровне шума  $\beta$  качество аппроксимации не зависит от параметра  $\gamma$ , а при увеличении шума  $\beta$  качество аппроксимации  $S$  падает.

На рис. 6 показан пример работы алгоритма при разных параметрах  $\beta$  и  $\gamma$ . Видно, что в случае отсутствия шума  $\beta$  обе локальные модели аппроксимируют выборку. При увеличении уровня шума качество аппроксимации падает: при  $\beta = 0,2$  при увеличении  $\gamma$  первая локальная модель из параболы переходит в эллипс; при  $\beta = 0,4$  при увеличении  $\gamma$  первая локальная модель из параболы переходит в эллипс, а вторая модель из параболы переходит в гиперболу.

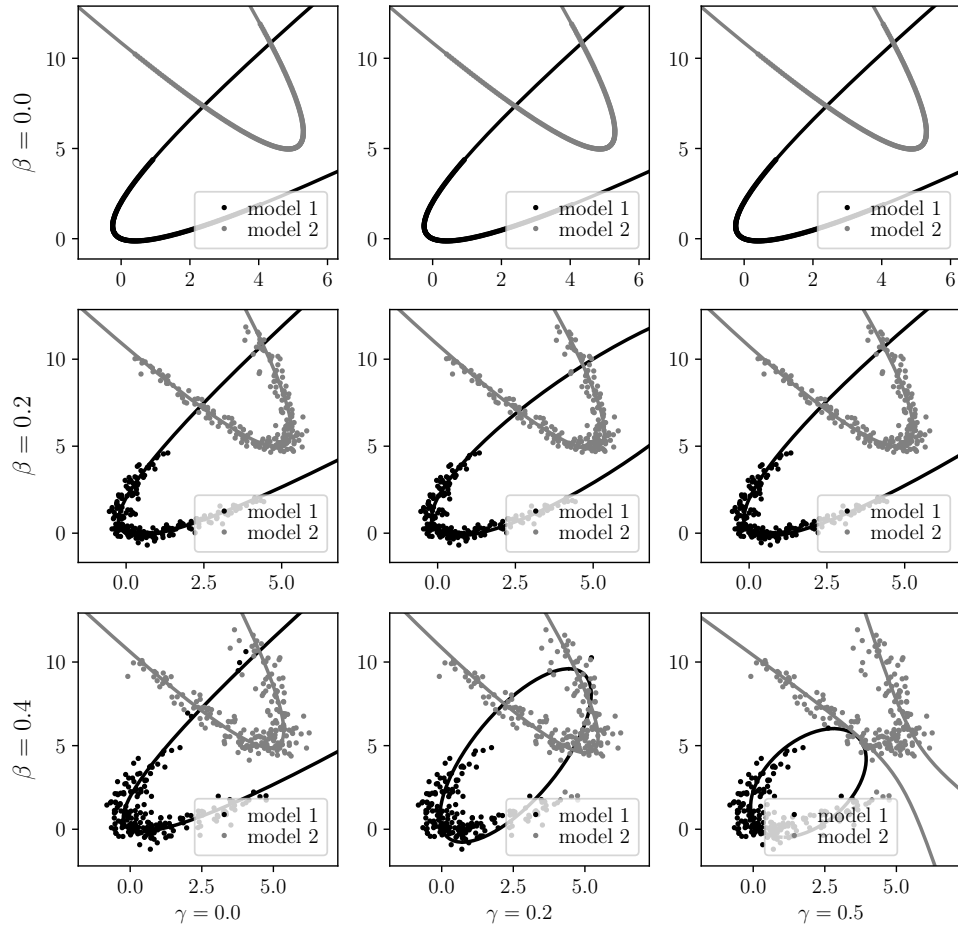


Рис. 5: Результат аппроксимации для данных с разным уровнем шума  $\beta$  и от дисперсии априорного распределения  $\gamma$

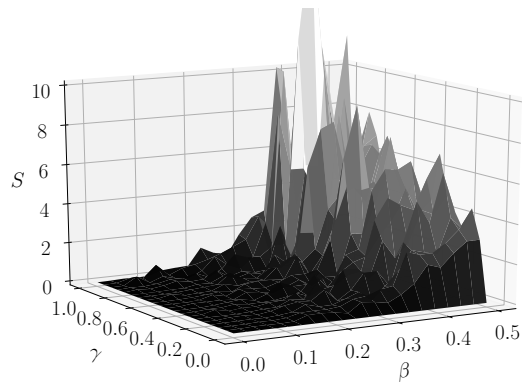


Рис. 6: Зависимость моделей от уровня шума  $\beta$  в данных, а также от дисперсии априорного распределения  $\gamma$

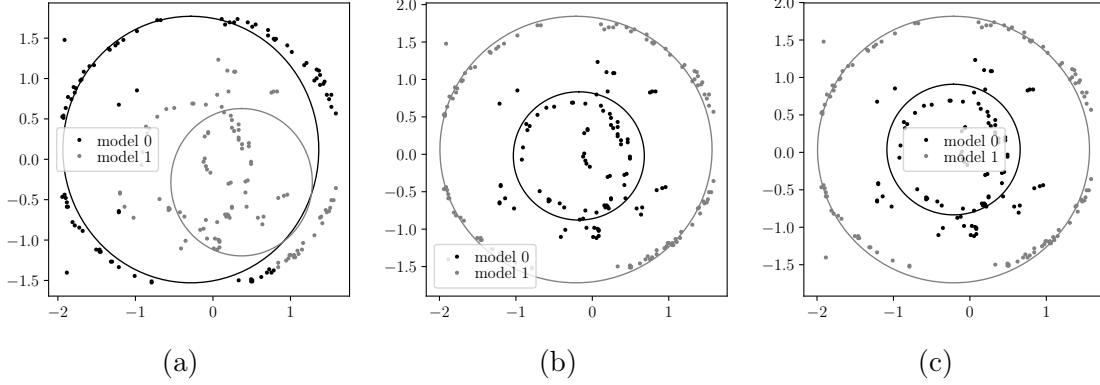


Рис. 7: Визуализация аппроксимации радужки глаза: а) в случае, если задан регуляризатор  $R_0$ ; б) в случае, если задан регуляризатор  $R_1$ ; в) в случае, если задан регуляризатор  $R_2$ .

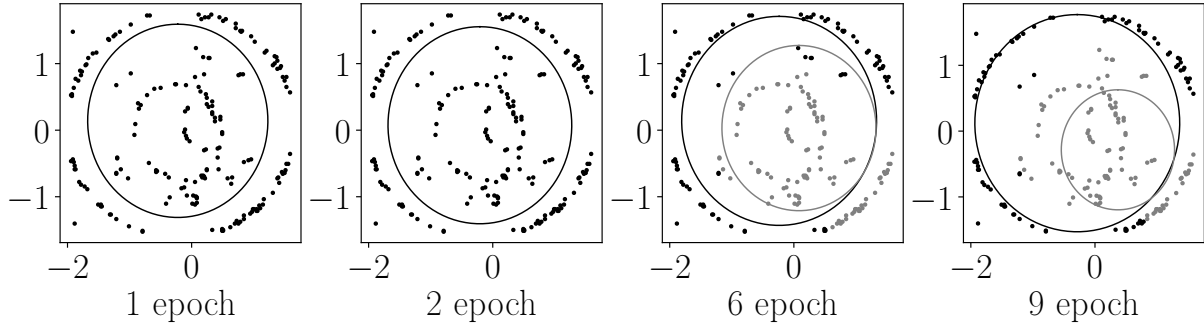


Рис. 8: Визуализация процесса сходимости мультимодели в случае регуляризатора  $R_0$

### 5.3 Аппроксимации радужки глаза

Проводиться анализ качества аппроксимации для задачи аппроксимации радужки глаза на изображении. Радужка глаза состоит из двух concentric окружностей, поэтому рассматривается мультимодель, которая состоит из двух экспертов: каждый эксперт аппроксимирует одну из окружностей. В вычислительном эксперименте сравнивается качество аппроксимации окружностей в случае задания разных регуляризаторов  $R_0, R_1, R_2$ . Регуляризатор  $R_0(\mathbf{V}, \mathbf{W}, E(\Omega)) = 0$ , то есть регуляризатор отсутствует. Регуляризатор:

$$R_1(\mathbf{V}, \mathbf{W}, E(\Omega)) = - \sum_{k=1}^K \mathbf{w}_k^\top \mathbf{w}_k,$$

который поощряет близкие к нулю параметры локальных моделей. Регуляризатор

$$R_2(\mathbf{V}, \mathbf{W}, E(\Omega)) = - \sum_{k=1}^K \mathbf{w}_k^\top \mathbf{w}_k + \sum_{k=1}^K \sum_{k'=1}^K \sum_{j=1}^2 (w_k^j - w_{k'}^j)^2,$$

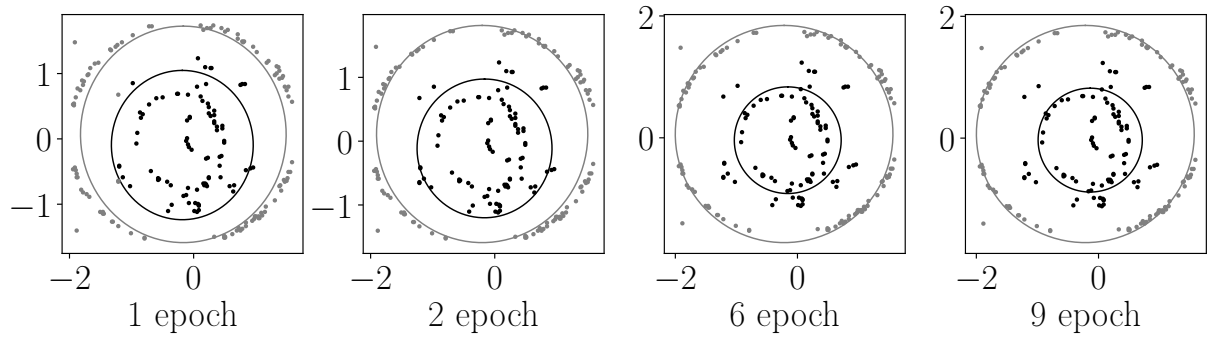


Рис. 9: Визуализация процесса сходимости мультимодели в случае регуляризатора  $R_1$

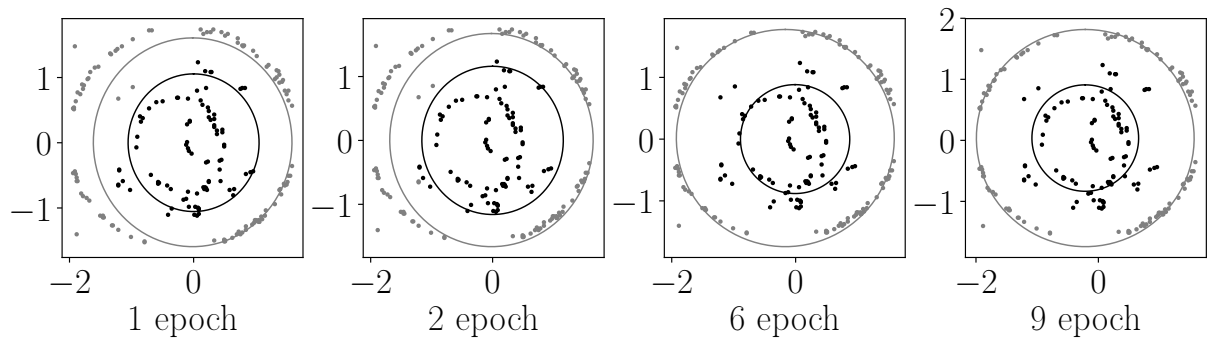


Рис. 10: Визуализация процесса сходимости мультимодели в случае регуляризатора  $R_2$

который поощряет совпадение центров окружностей и близкие к нулю параметры модели. На рис. 7 показан результат работы алгоритма аппроксимации радужки глаза после 10 итераций. Видно, что при отсутствии регуляризатора одна из окружностей находится не верно. В случае, если задан регуляризатор  $R_1$  модель аппроксимирует обе окружности с хорошим качеством, но окружности не являются концентрическими. В случае задания регуляризатор  $R_2$  получаем концентрические окружности на изображении.

На рис. 8-10 показан процесс сходимости мультимodelей в случае задания разных регуляризаторов  $R_0, R_1, R_2$ . Видно, что модели с типом регуляризатора  $R_1$  и  $R_2$  аппроксимируют обе окружности, а мультимodelь с регуляризатором  $R_0$  аппроксимирует только большую окружность.

## 6 Заключение

В данной работе предложен метод для построения интерпретируемых моделей машинного обучения на основе экспертной информации. В качестве задачи рассмотрена задача аппроксимации кривых второго порядка: парабола, гипербола, эллипс. Аппроксимации кривых второго порядка применена в задачи аппроксимации радужки глаза.

Проведен эксперимент, в котором анализируется качество аппроксимации кривых второго порядка в зависимости от начального уровня шума в данных, а также в зависимости от регуляризатора. В ходе эксперимента показано, что при увеличении уровня шума в начальных данных, точность аппроксимации падает: при большом шума вид аппроксимируемой фигуры изменяется с параболы на гиперболу.

Проведен вычислительный эксперимент по аппроксимации радужки глаза при помощи двух концентрических окружностей. В эксперименте показано, что регуляризация, которая основана на экспертной информации улучшает качество аппроксимации.

## Список литературы

- [1] *Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin* "Why Should I Trust You?": Explaining the Predictions of Any Classifier // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining PP. 1135–1144. 2016
- [2] *Chen Tianqi, Guestrin Carlos* XGBoost: A Scalable Tree Boosting System // KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016.
- [3] *Chen Xi, Ishwaran Hemant* Random Forests for Genomic Data Analysis // Genomics. 2012. Issues. 99, No 6. pp. 323–329.

- [4] *Yuksel Seniha Esen, Wilson Joseph N., Gader Paul D* Twenty Years of Mixture of Experts // IEEE Transactions on Neural Networks and Learning Systems. 2012. Issues. 23, No 8. pp. 1177–1193.
- [5] *L. Cao* Support vector machines experts for time series forecasting // Neurocomputing, vol. 51, pp. 321–339, Apr. 2003.
- [6] *A. P. Dempster, N. M. Laird and D. B. Rubin* Maximum Likelihood from Incomplete Data via the EM Algorithm // Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1 pp. 1-38, 1977.
- [7] *M. S. Yumlu, F. S. Gurgen, N. Okay* Financial time series prediction using mixture of experts // in Proc. 18th Int. Symp. Comput. Inf. Sci., 2003, pp. 553–560.
- [8] *Y. M. Cheung, W. M. Leung, and L. Xu* Application of mixture of experts model to financial time series forecasting // in Proc. Int. Conf. Neural Netw. Signal Process., 1995, pp. 1–4.
- [9] *A. S. Weigend, S. Shi* Predicting daily probability distributions of S&P500 returns // J. Forecast., vol. 19, no. 4, pp. 375–392, 2000.
- [10] *R. Ebrahimpour, M. R. Moradian, A. Esmkhani, F. M. Jafarlou* Recognition of Persian handwritten digits using characterization loci and mixture of experts // J. Digital Content Technol. Appl., vol. 3, no. 3, pp. 42–46, 2009.
- [11] *A. Estabrooks, N. Japkowicz* A mixture-of-experts framework for text classification //in Proc. Workshop Comput. Natural Lang. Learn., Assoc. Comput. Linguist., 2001, pp. 1–8.
- [12] *S. Mossavat, O. Amft, B. de Vries, P. Petkov, W. Kleijn* A Bayesian hierarchical mixture of experts approach to estimate speech quality // in Proc. 2nd Int. Workshop Qual. Multimedia Exper., pp. 200–205., 2010
- [13] *C. Sminchisescu, A. Kanaujia, and D. Metaxas* B M3 E: Discriminative density propagation for visual tracking // IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 11, pp. 2030–2044, 2007.
- [14] *I. Matveev* Detection of iris in image by interrelated maxima of brightness gradient projections // Appl.Comput. Math. 9 (2), 252–257, 2010.
- [15] *I. Matveev, I. Simonenko*. Detecting precise iris boundaries by circular shortest path method // Pattern Recognition and Image Analysis. 24. 304-309. 2014.
- [16] *K. Bowyer, K. Hollingsworth, P. Flynn* A Survey of Iris Biometrics Research: 2008–2010.

- [17] *F. Peng, R. A. Jacobs, M. A. Tanner* Bayesian inference in mixtures-of-experts and hierarchical mixtures-of-experts models with an application to speech recognition // J. Amer. Stat. Assoc., vol. 91, no. 435, pp. 953–960, 1996.
- [18] *A. Tuerk* The state based mixture of experts HMM with applications to the recognition of spontaneous speech // Ph.D. thesis, Dept. Eng., Univ. Cambridge, Cambridge, U.K., 2001.
- [19] *Bishop C.* Pattern Recognition and Machine Learning. — Berlin: Springer, 2006. 758 p.
- [20] *Kaiming He* Deep Residual Learning for Image Recognition.
- [21] *Han Xu* Adversarial Attacks and Defenses in Images, Graphs and Text: A Review
- [22] *Xingjun Ma* Understanding Adversarial Attacks on Deep Learning Based Medical Image Analysis Systems
- [23] *Naveed Akhtar* Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey
- [24] *Salamani Dalila* Deep generative models for fast shower simulation in ATLAS