

А. В. Гончаров¹, В. В. Стрижов²**Построение интерпретируемых моделей глубокого обучения в задаче социального ранжирования***

В работе решается задача построения модели кредитного скоринга. Рассмотрен класс интерпретируемых моделей, доставляющих значение функционала качества на уровне неинтерпретируемых аналогов. Основная идея работы — построение единой процедуры оптимизации с использованием различных интерпретируемых модулей обработки признаков и построения моделей. Особенностью исследования является подход глубокого обучения к построению модели с использованием различных процедур. Предлагается для каждой из процедур заменить оптимизацию локальных критериев качества итерационной оптимизацией в рамках единого функционала качества внутри одной задачи оптимизации. Качество решения задачи предложенным методом оценивается на открытых данных задач социального ранжирования.

Ключевые слова: скоринговая карта; отбор признаков; кредитный скоринг.

1 Введение

В задаче построения моделей кредитного скоринга определяется уровень кредитоспособности заемщика. Эта задача решается крупными банковскими структурами. Аналитические отделы таких структур занимаются построением скоринговых карт, содержащих необходимые для построения модели признаки. Аналитик не может вручную выявить большое число закономерностей, присутствующих в данных. В таком сложном процессе используются технологии, позволяющие автоматизировать процесс выставления баллов клиентам, совмещаются различные скоринговые и риск модели.

Для руководителей подобных организаций важна обоснованность предоставленного скорингового балла во избежании скандальных ситуаций вокруг компании. Это приводит к использованию внутри банковских структур интерпретируемых моделей, которые зачастую имеют более низкое качество по сравнению с неинтерпретируемыми аналогами.

¹Московский физико-технический институт, магистр., E-mail: alex.goncharov@phystech.edu

²Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН, научн. сотр., д.ф.-м.н., E-mail: strijov@ccas.ru

*Работа выполнена при финансовой поддержке РФФИ, проект № 16-07-01158

Активно развиваются два подхода к построению скоринговых карт и моделей. Первый подход использует новые неинтерпретируемые технологии, позволяющие получать высокое качество [1, 2, 3, 4]. Аналитики крупных банков используют эти модели для подтверждения или опровержения основных результатов. В случае несогласованности результатов моделей дело посылают на перепроверку.

Второй подход основан на интерпретируемых ранжирующих и классифицирующих моделях, таких как логистическая регрессия [5, 6, 7, 8, 9]. Развиваются новые методы порождения признаков пространств, их модификации. Полученные модели позволяют выделить необходимые интерпретируемые признаки, которые будут одобрены аналитиками организации.

В работе предложена новая процедура построения интерпретируемых скоринговых моделей методами глубокого обучения. Для определения уровня кредитоспособности заемщика используется анкета заемщика, содержащая признаки как в линейных, так и в ординальных и номинальных шкалах.

Цель работы — создание модели глубокого обучения для построения скоринговой карты как суперпозиции процедур порождения признаков. Необходимо при этом удовлетворить требованиям интерпретируемости и качества модели. Для решения задачи используется эксплуатационный критерий в единой постановке задачи оптимизации.

Процедура включает все шаги, которые аналитик выполняет при построении эксплуатируемой модели. А именно: процедуры группировки [10] и сегментации, создание интерпретируемых комбинаций признаков, построение моделей классификации и вычисление критерия качества.

2 Постановка задачи.

Дана выборка $D = \{(\mathbf{x}_i, y_i) : i = 1, \dots, m\}$, содержащая m объектов. Каждый из них описан линейными, категориальными и бинарными признаками, $\mathbf{x}_i \in \mathbb{R}^l \times \mathbb{C}^c \times \mathbb{B}^b$ и принадлежит одному из двух классов: $y_i \in \{0, 1\}$. Пусть индексы объектов $\{i = 1, \dots, m\} = \mathcal{I}$ поделены $\mathcal{I} = \mathcal{L} \sqcup \mathcal{T}$ на индексы объектов обучения и контроля.

Обозначим матрицу плана $X \in \mathbb{R}^{m \times n}$, $X = (\chi^1, \dots, \chi^n)$, где χ_i — i -ый признак. Зададим множество индексов всех признаков $\mathcal{A} = \mathcal{A}_l \sqcup \mathcal{A}_c \sqcup \mathcal{A}_b$, где \mathcal{A}_l обозначает индексы признаков в линейных шкалах, \mathcal{A}_c — категориальных признаков, а \mathcal{A}_b — бинарных признаков.

Модель классификации F в работе представлена в виде суперпозиции $F = f \circ f_f \circ f_g \circ f_s$, где f_s производит сегментацию линейных признаков, f_g — группировку категорий

для категориальных признаков, f_f — строит новые признаки на основе старых, причем все эти функции переводят объект \mathbf{x} из одного признакового пространства в другое, что вызывает изменение матрицы плана X : суперпозиция $f_f \circ f_g \circ f_s$ преобразует одну матрицу плана в другую.

$$f_f \circ f_g \circ f_s : X \rightarrow \hat{X}, \quad (1)$$

где $\hat{X} = (\hat{\chi}^1, \dots, \hat{\chi}^{\hat{n}})$.

Определение. Функция f называется интерпретируемой, если для любого $\chi_i \in [a - \epsilon, a + \epsilon]$ существует δ : для объекта $\mathbf{x} = \{\dots, \chi_i, \dots\}$ вероятность принадлежности классу $p \in [f(\mathbf{x}) - \delta, f(\mathbf{x}) + \delta]$.

В работе функцией f выбрана логистическая регрессия как интерпретируемая модель классификации над получившимся признаковым пространством.

$$f(\hat{X}, w) = \frac{1}{1 + \exp(-\hat{X}w)}. \quad (2)$$

В качестве целевой функции ошибки $S(w)$ для обучения модели логистической регрессии используем логарифмическую функцию правдоподобия:

$$L(w) = -\ln P(D|w) = -\sum_{i \in \mathcal{L}} (y_i \ln w^T x_i + (1 - y_i) \ln(1 - w^T x_i)). \quad (3)$$

При обучении процедур сегментации, группировки и порождения признаков, а также для оценки качества внешним критерием, согласно банковским требованиям, предлагается использовать площадь под ROC-кривой:

$$R(w) = \frac{\sum_{i \in \mathcal{L}} \sum_{j \in \mathcal{L}} [y_i < y_j] [F(\mathbf{x}_i) < F(\mathbf{x}_j)]}{\sum_{i \in \mathcal{L}} \sum_{j \in \mathcal{L}} [y_i < y_j]} \quad (4)$$

Функция f является параметрической с параметром w , а для f_f , f_g и f_s помимо параметров $P = p_f \sqcup p_g \sqcup p_s$ выделены структурные параметры $H = h_f \sqcup h_g \sqcup h_s$. Задача подбора оптимальных параметров и структурных параметров модели классификации F сводится к решению задачи оптимизации:

$$H, P, w = \underset{H, P, w}{\operatorname{argmin}} R(w). \quad (5)$$

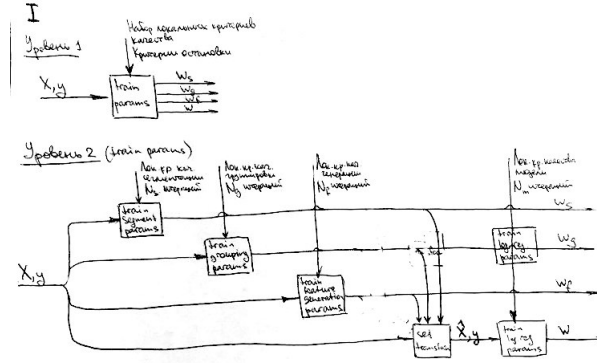


Рис. 1: IDEF0 классического подхода

3 Решение задачи оптимизации

Базовый подход к решению оптимизационной задачи (1) предполагает оптимизацию локальных критериев качества для каждой из процедур и изображен в формате IDEF0 на рисунке 1:

$$h_s, p_s = \operatorname{argmin}_{h_s, p_s} S_s(h_s, p_s), \quad (6)$$

$$h_g, p_g = \operatorname{argmin}_{h_g, p_g} S_g(h_g, p_g), \quad (7)$$

$$h_f, p_f = \operatorname{argmin}_{h_f, p_f} S_f(h_f, p_f), \quad (8)$$

$$w = \operatorname{argmin}_w S(w), \quad (9)$$

где S_s, S_g, S_f для базового подхода описаны далее.

Обучение модели F устроено следующим образом: дана обучающая выборка, для каждой функции из суперпозиции $F = f \circ f_f \circ f_g \circ f_s$ последовательно подбираются параметры, оптимальные для соответствующего целевого функционала. Процедуры поиска независимы, что приводит к неоптимальности решения относительно целевого оптимизируемого функционала.

В работе предлагается использовать единый целевой функционал, который будет использоваться при последовательном обучении каждого отдельного блока. Процедура обучения с учетом этого предложения изображена в формате IDEF0 на рисунке 2а.

Изменение относительно базового подхода мотивировано поиском лучшего, нежели базового, решения задачи (5) относительно глобального критерия качества.

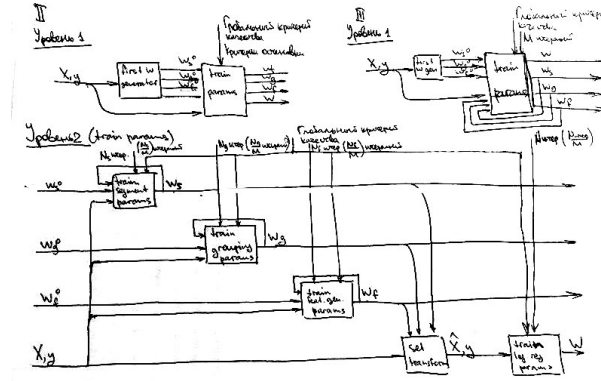


Рис. 2: IDEF0 модифицированного и предлагаемого подхода

При решении задачи (5) на итоговый функционал качества влияет каждая из оптимизируемых переменных H, P, w . Так как процедура оптимизации проводится последовательно для каждой группы переменных, то на каком-то из шагов оптимальное решение может перемтаться быть оптимальным. Предлагается циклично повторять процедуру оптимизации при найденных на предыдущем цикле переменных как начальном приближении, что продемонстрировано на рисунке 2б.

Заметим, что при таком алгоритмическом решении задачи численной оптимизации на каждом цикле могут быть получены различные матрицы \hat{X} , что приводит к неоднозначности выбора начальных параметров функции f при оптимизации цикла.

Предлагается воспользоваться высокой скоростью оптимизации моделей логистической регрессии и проводить полную процедуру оптимизации модели f в каждом цикле.

3.1 Сегментация

Сегментация дает возможность учета сложной кусочно-постоянной зависимости функции ошибки от признака в линейной шкале и упростить интерпретацию итоговой модели. Исходное признаковое пространство дополняется бинарными признаками, обозначающими принадлежность исходного признака заданному сегменту на линейной шкале.

3.1.1 Постановка задачи сегментации

Пусть значения признака χ_i принадлежат отрезку $[a, b]$. Структурным параметром h_s процедуры сегментации является набор количества узлов разбиения для каждого из признаков. Параметрами p_s процедуры сегментации является набор координат узлов разбиения, причем $p_{si}^j \in [a, b], j = 0, \dots, h_{si}; a = p_{si}^0 < \dots < p_{si}^j < \dots < p_{si}^{h_{si}} = b$.

Исходная матрица плана X пополняется новым набором бинарных признаков, являющихся индикаторами принадлежности значения исходного линейного признака каж-

дому сегменту:

$$\{\chi^j\}_{j=1}^{h_{si}} \in \mathbb{B}^{h_{si}} : \chi^j = [\chi \in [p_{si}^{q-1}; p_{si}^q]], q = 1, \dots, h_{si}.$$

Подбор структурных параметров и параметров процедуры сегментации набора линейных признаков с индексами \mathcal{A}_l на обучающей подвыборке выборки D происходит путем решения следующей оптимизационной задачи для целевой функции:

$$h_s, p_s, w = \underset{h_s, p_s, w}{\operatorname{argmin}} R(h_s, p_s, w | D, \mathcal{A}_l, \mathcal{L}),$$

где R - функция качества, описанная в (4).

3.1.2 Решение задачи сегментации

Параметрами задачи сегментации служат узлы разбиения, принадлежащие действительной шкале. В базовой модели для поиска параметров и структурных параметров сегментов используются значения WOE выбранного сегмента q и доля выборки в сегменте q для признака i [?]:

$$\begin{aligned} count_q &= \frac{1}{m} \sum_{j \in \mathcal{L}} [\chi_j^i \in [p_{si}^{q-1}; p_{si}^q]], \\ woe_q &= \log \frac{g_j}{b_j} = \log \frac{\sum_{j \in \mathcal{L}} [\chi_j^i \in [p_{si}^{q-1}; p_{si}^q]] [y_j = 1]}{\sum_{j \in \mathcal{L}} [\chi_j^i \in [p_{si}^{q-1}; p_{si}^q]] [y_j = 0]}, \end{aligned}$$

где g_j и b_j — доли классов 0 и 1 в j -ом сегменте. Для построения базовой модели предлагается делить интерпретируемый линейный признак на максимально возможное число сегментов, а затем объединять их согласно нескольким принципам: доля объектов в сегменте не менее 4% от общего числа выборки, значение woe для соседних сегментов существенно отличается.

Определение. Линейный признак χ будем называть интерпретируемым, если $\chi \in [0, 100]$.

Определение. Процедура сегментации линейного интерпретируемого признака i называется интерпретируемой, если для любого q $count_q > 0.04$, $woe_q - woe_{q-1} > 0.1$.

Определение получено эмпирическим путем исходя из обзора литературы и применяемых методов.

В работе предлагается использовать численные методы оптимизации для поиска параметров сегментации выбранного признака. Целевым функционалом служит $R(w)$, а параметрами оптимизации для каждого признака i — p_{si} .

Структурные параметры будут найдены автоматически. После каждой итерации численной оптимизации производится процедура корректировки параметров сегментации p_{si} : удаляется p_{si}^q , если $p_{si}^q - p_{si}^q - 1 < 2$. Таким образом при консолидации параметров сегментации в одной точке линейного пространства структурный параметр сегментации признака i изменяется.

Определение. Сложностью процедуры сегментации для признака i является число параметров или значение структурного параметра h_{si} .

Определение. Сложностью процедуры сегментации является число параметров сегментации модели или сумма значений структурного параметра $\sum_i h_{si}$.

3.2 Группировка

3.2.1 Постановка задачи группировки

Наличие большого множества категорий для категориального признака резко увеличивает сложность признакового пространства и снижает интерпретируемость построенных моделей. Группировка нескольких категорий в одну снижает сложность модели и повышает интерпретируемость.

Пусть категориальный признак χ_i имеет $|C|$ категорий, где C — множество категорий этого признака. Структурным параметром h_{gi} процедуры группировки признака i является количество групп новых категорий, $|h_{gi}| < |C|$. Параметром группировки является сюръекция $p_{gi} : C \rightarrow h_{gi}$.

Исходное признаковое пространство изменяется: категориальный признак χ заменяется категориальным признаком χ_h .

$$\begin{array}{ccccccccc} \chi & = & 1 & 2 & 3 & \dots & C & & C \text{ — число категорий} \\ & & \downarrow & \downarrow & \downarrow & & \downarrow & & \\ \chi_h & = & \gamma_1 & \gamma_2 & \gamma_3 & \dots & \gamma_C & |h_{gi}| \text{ — число новых категорий, } \gamma_i \in h_{gi} \end{array}$$

Подбор структурных параметров и сюръекций процедуры группировки набора категориальных признаков с индексами \mathcal{A}_c на обучающей подвыборке выборки D происходит путем решения следующей оптимизационной задачи для целевой функции:

$$h_g, p_g, w = \underset{h_g, p_g, w}{\operatorname{argmin}} R(h_g, p_g, w | D, \mathcal{A}_c, \mathcal{L}),$$

где R - функция качества, описанная в (4).

3.2.2 Решение задачи группировки

Параметрами задачи сегментации служат сюръективные отображения. В базовой модели для поиска параметров и структурных параметров сегментов используются значения WOE выбранной категории c и доля выборки, принадлежащей этой категории для признака i :

$$count_c = \frac{1}{m} \sum_{j \in \mathcal{L}} [\mathbf{x}_j^i = c],$$

$$woe_c = \log \frac{g_j}{b_j} = \log \frac{\sum_{j \in \mathcal{L}} [\mathbf{x}_j^i = c][y_j = 1]}{\sum_{j \in \mathcal{L}} [\mathbf{x}_j^i = c][y_j = 0]},$$

где g_j и b_j — доли классов 0 и 1 в j -ом сегменте. Для построения базовой модели предлагается категории, для которых доля объектов мала объединять с ближайшими по значению WOE категориями.

Определение. Процедура группировки категориального признака i называется интерпретируемой, если для любой категории $c \in h_{gi}$ $count_c > 0.04$, а число категорий $|h_{gi}| < 10$.

Определение получено эмпирическим путем исходя из обзора литературы и применяемых методов.

Для решения задачи необходимо подобрать сюръекции, доставляющие оптимум функционалу качества. Численные методы оптимизации не подходят для решения подобных задач. В работе для итерационной оптимизации процедуры группировки предлагается использование генетических алгоритмов.

Определение. Особь a_χ в задаче оптимизации процедуры группировки категориального признака χ_i — последовательность $\{\gamma_1, \dots, \gamma_C\}$, $\gamma_i \in h_{gi}$, задающая произвольное сюръективное отображение.

Определение. Процедура скрещивания двух особей a_{χ_1} и a_{χ_2} — скрещивание соответствующих последовательностей: $\{\gamma_{11}, \dots, \gamma_{1C}\}$ и $\{\gamma_{21}, \dots, \gamma_{2C}\}$. Скрещивание двух таких последовательностей задается случайным бинарным вектором $\mathbf{bin} \in \mathbb{B}^{|C|}$:

$$a_{12_1} = \{\mathbf{bin}_1 \gamma_{11} + (1 - \mathbf{bin}_1) \gamma_{21}, \dots, \mathbf{bin}_C \gamma_{11} + (1 - \mathbf{bin}_C) \gamma_{2C}\}.$$

$$a_{12_2} = \{\mathbf{bin}_1 \gamma_{21} + (1 - \mathbf{bin}_1) \gamma_{11}, \dots, \mathbf{bin}_C \gamma_{21} + (1 - \mathbf{bin}_C) \gamma_{1C}\}.$$

Процедура оптимизации признака χ генетическим алгоритмом построена следующим образом:

1. Порождение поколения, содержащего множество особей.
2. Отбор лучших особей согласно критерию качества.
3. Скрещивание случайных пар особей между собой и добавление их к поколению.

4. Возврат ко 2 шагу, пока критерий качества не сойдется.

Определение. Сложностью процедуры группировки для признака i является число категорий структурного параметра $|h_{gi}|$.

Определение. Сложностью процедуры группировки является число новых категорий модели или сумма значений мощностей структурного параметра $\sum_i |h_{gi}|$.

3.3 Порождение интерпретируемых признаков

3.3.1 Постановка задачи порождения

Пусть множество $F = \{f_i\}$ — множество интерпретируемых порождающих непараметрических функций над признаками, заданных экспертами.

Описание	Формула	in	N in	out
Negate binary	\bar{x}	bin	1	bin
Logarithm	$\log(x)$	lin	1	lin
Logistic sigmoid	$\frac{1}{1+\exp(x)}$	lin	1	lin
Square root	\sqrt{x}	lin	1	lin
Inverse	$\frac{1}{x}$	lin	1	lin
Multiplication	$x * y$	any	2	lin
Sum	$x + y$	any	2	lin

Пусть $\{\chi_i\}, i \in \mathcal{A}_l \sqcup \mathcal{A}_b$ — измеряемые линейные и бинарные признаки. Необходимо найти лучшую комбинацию заданного числа из всех допустимых суперпозиций признаков $\{\chi_i\}$ с использованием функций $\{f_i\}$, после чего вычислить и добавить их в признаковое пространство объектов X .

Определение. Суперпозиция — формула, представляемая в виде дерева, в каждом узле которого находится функция из $\{f_i\}$, а в каждом листе — признак из $\{\chi_i\}$.

Сложность суперпозиции — число использованных в ней элементов из $\{f_i\}$ и $\{\chi_i\}$ с повторениями.

Допустимая суперпозиция — суперпозиция, в которой соблюдены ограничения на сложность, количество переменных и типы переменных в каждом узле.

Из множества всех допустимых суперпозиций Σ необходимо выбрать наилучшую комбинацию суперпозиций $p_f = \{\mathbb{F}_i\}$. Структурными параметрами являются количество суперпозиций в комбинации $h_f = |\{\mathbb{F}_i\}|$ и максимальная сложность суперпозиции d , а параметрами — комбинация $p_f = \{\mathbb{F}_i\}$. Подбор структурных параметров и параметров процедуры порождения на обучающей подвыборке выборки D происходит путем решения следующей оптимизационной задачи для целевой функции:

$$h_f, p_f, w = \underset{h_f, p_f, w}{\operatorname{argmin}} R(h_f, p_f, w | D, \mathcal{A}_l \sqcup \mathcal{A}_b, \mathcal{L}),$$

где R - функция качества, описанная в (4).

3.3.2 Решение задачи порождения

Параметрами задачи порождения признаков служат наборы суперпозиций. В базовой модели для поиска параметров и структурных параметров генерации признаков используется корреляция каждой суперпозиции с целевой переменной:

$$corr_g = \sum_i (1 - corr(\mathbb{F}_i, \mathbf{y})^2)$$

Для построения базовой модели предлагается качество особи оценивать с помощью $corr_g$.

Определение. Процедура порождения признаков называется интерпретируемой, если каждая суперпозиция \mathbb{F}_i имеет сложность не более 10.

Для решения задачи необходимо подобрать набор суперпозиций, доставляющий оптимум функционалу качества. В работе для итерационной оптимизации процедуры порождения признаков предлагается использование генетических алгоритмов.

Определение. Особь a в задаче оптимизации процедуры порождения признаков — комбинация суперпозиций $\{\mathbb{F}_i\}$.

Определение. Процедура скрещивания двух особей a_1 и a_2 — скрещивание соответствующих комбинаций суперпозиций: $\{\mathbb{F}_{1i}\}$ и $\{\mathbb{F}_{2i}\}$. Скрещивание двух таких последовательностей задается случайным бинарным вектором $\mathbf{bin} \in \mathbb{B}^{|\{\mathbb{F}_i\}|}$:

$$a_{12_1} = \{\mathbf{bin}_i \mathbb{F}_{1i}\} \cup \{(1 - \mathbf{bin}_i) \mathbb{F}_{2i}\}.$$

$$a_{12_2} = \{\mathbf{bin}_i \mathbb{F}_{2i}\} \cup \{(1 - \mathbf{bin}_i) \mathbb{F}_{1i}\}.$$

Процедура оптимизации признака генетическим алгоритмом построена следующим образом:

1. Порождение поколения, содержащего множество особей.
2. Отбор лучших особей согласно критерию качества.
3. Скрещивание случайных пар особей между собой и добавление их к поколению.
4. Возврат ко 2 шагу, пока критерий качества не сойдется.

3.4 Объединение блоков

Базовый подход к объединению вышеописанных процедур построения модели подразумевает последовательное выполнение базовых процедур сегментации, группировки и порождения признаков, то есть достижение оптимума локального критерия качества для каждой из процедур. При этом описанные критерии построения процедур не связаны напрямую с целевым функционалом, что приводит к неоптимальности построенного решения.

В работе для реализации предложенной схемы предлагается следующий пошаговый итерационный алгоритм оптимизации:

1. Инициализация начальных точек для градиентных методов и начальных поколений для генетических алгоритмов.
2. Шаги градиентного метода оптимизации для процедуры сегментации для каждого признака
3. Итерация генетической оптимизации для процедуры группировки для каждого признака
4. Итерация генетической оптимизации для процедуры порождения
5. Возврат ко 2 шагу до сходимости функционала качества

Определение. Сложность модели — это сумма сложностей процедур сегментации, группировки, и порождения признаков.

4 Критерии качества

Для исследования применимости предложенного решения и его сравнения с базовыми и альтернативными подходами предлагается использовать различные критерии качества модели.

В задачах кредитного скоринга используют два вида критериев качества: оптимизационные и эксплуатационные. Эксплуатационные критерии качества заданы исходя из потребностей индустрии при решении бизнес-задач. Оптимизационные критерии влияют на настройку параметров модели и нужны для ее построения. Но они не связаны с эксплуатационными критериями.

Эксплуатационные критерии качества:

- AUC — площадь под ROC-кривой (4). Классический критерий качества в задачах бинарной классификации, имеющий физический смысл доли правильно упорядоченных алгоритмом пар объектов. Критерий используется для оптимизации

модели и для ее оценки экспертами. Это основной критерий сравнения моделей.

- Сложность, число признаков. От сложности модели напрямую зависит ее интерпретируемость экспертами, которые и выставляют соответствующие требования. В предложенной модели этот показатель напрямую зависит от значения структурных параметров модели, поэтому в оптимизируемый функционал значение не входит. Происходит сравнение с альтернативными моделями.

Оптимизационные критерии качества:

- Правдоподобие модели. Правдоподобие модели указывает на вероятностный смысл полученного ответа и используется в построении решения для поиска оптимальных параметров логистической регрессии.
- Устойчивость — ортогональность признаков. Для устойчивости модели необходимо отсутствие мультиколлинеарности признаков. Чем разнороднее они будут, тем лучше будет вести себя модель.
- Точность, полнота. В некоторых бизнес-приложениях точность и полнота метода позволяет вычислять финансовые риски используемой модели.

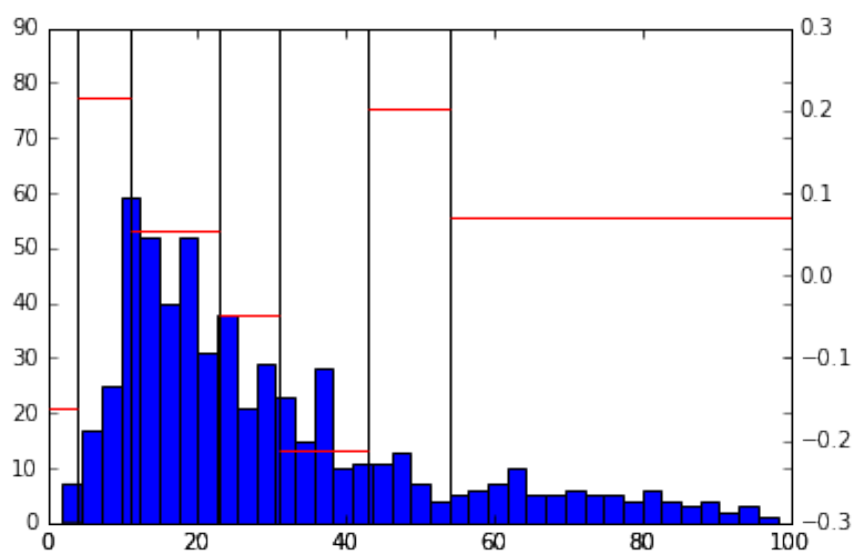
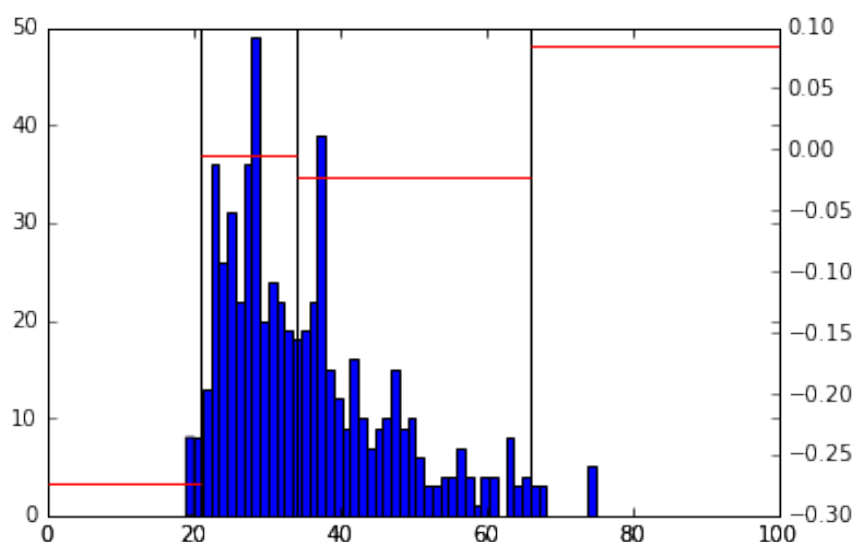
5 Вычислительный эксперимент

Построенная процедура оптимизации всех блоков в единой структуре модели сравнивается как по качеству, так и по сложности модели с другими моделями: нейронные сети, случайные леса и бустинг над деревьями.

Для демонстрации результатов работы модели выбраны классические датасеты по кредитному скорингу: открытые данные соревнования на сайте Kaggle.com (ссылка) и данные репозитория UCI (ссылка).

Результатом вычислительного эксперимента является сравнение качества для моделей, примененных "из коробки". Для устойчивости моделей важно число параметров модели. Для интерпретируемости результатов оцениваются построенные признаки для модели логистической регрессии, так как в остальных случаях модели являются неинтерпретируемыми.

Для проведения вычислительного эксперимента и демонстрации интерпретируемости результатов построенной модели из множества признаков в выборке UCI выбрано подмножество, содержащее: линейные признаки, категориальные признаки и бинарные признаки.

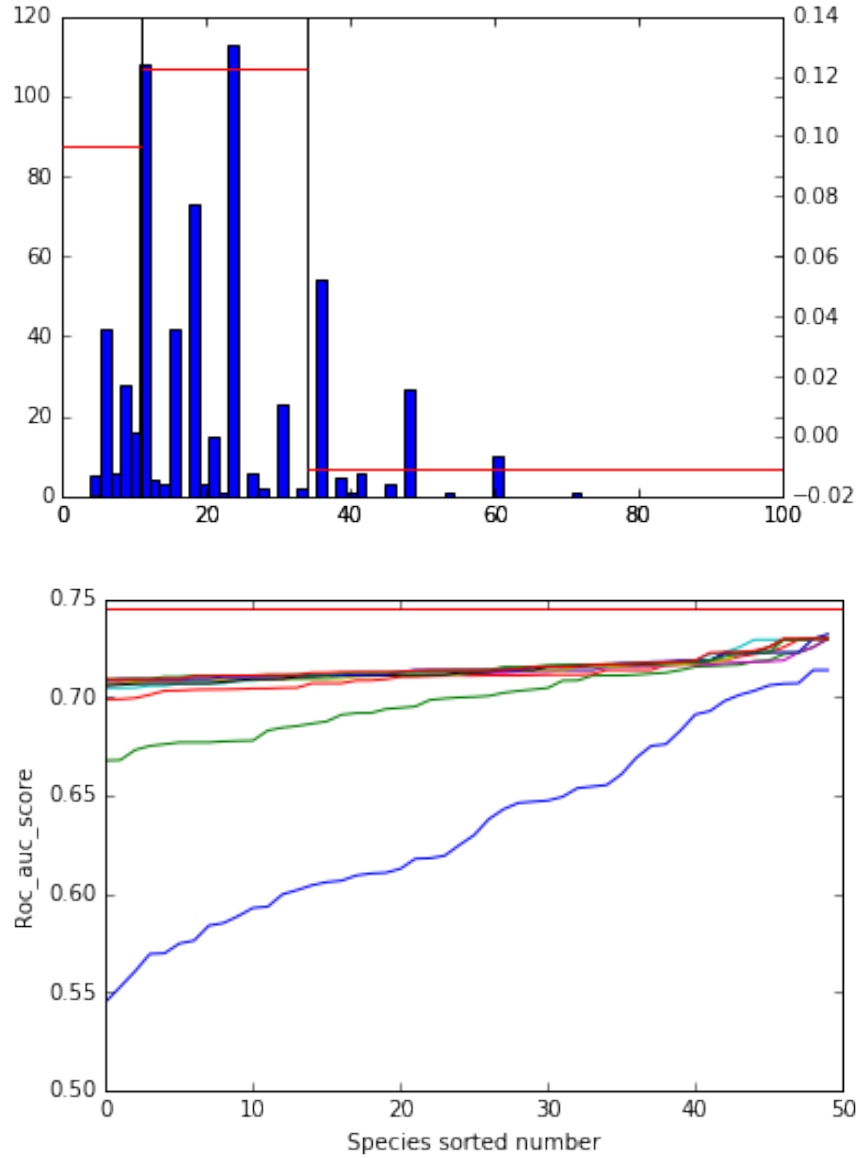


Результаты оптимизации процедуры сегментации для нескольких линейных признаков приведены на графиках ниже. На графиках изображено распределение линейных признаков в виде гистограммы, а также предложенные параметры сегментации признака как узлы сегментации и веса соответствующих бинарных признаков принадлежности сегменту в модели для каждого из сегментов.

Из приведенных графиков видно, что разбиение признаков интерпретируемо. Особенно важно это свойство для хвостов распределения, куда попадает малое число объектов.

WARNING: добавить для каждого сегмента на график значения count и woe.

Для оценки сходимости оптимизационных процедур генетических алгоритмов был проведен ряд экспериментов, внутри каждого из которых проводилась оптимизация блоков: процедуры группировки и процедуры порождения признаков.



Ниже приведены примеры, демонстрирующие построенное решение:

$\chi = 1 \ 2 \ 3 \ 4 \ 5$ 5 — число исходных категорий

↓ ↓ ↓ ↓ ↓

$\chi_h = 1 \ 2 \ 2 \ 3 \ 3 \ 3$ — число новых категорий, $\gamma_i \in h_{gi}$

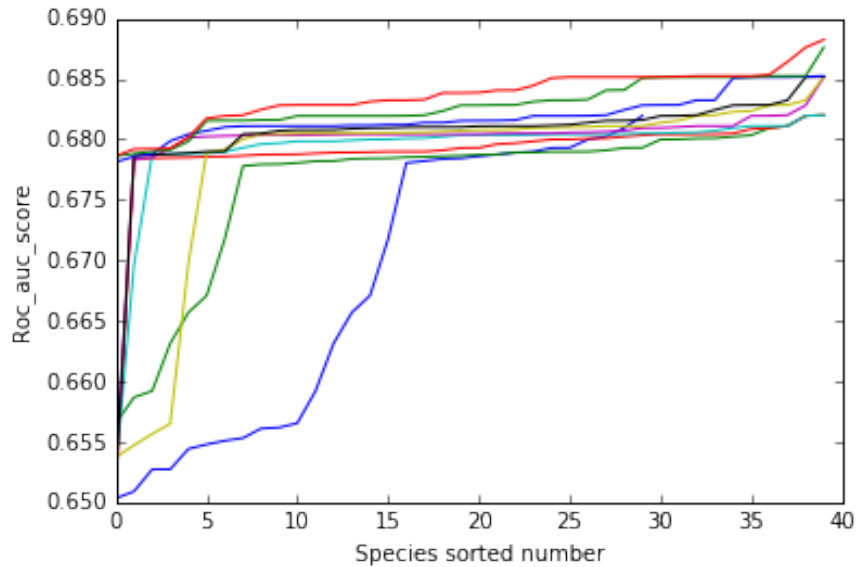
$\chi = 0 \ 1 \ 2 \ 3 \ 4 \ 5$ 6 — число исходных категорий

↓ ↓ ↓ ↓ ↓ ↓

$\chi_h = 3 \ 1 \ 3 \ 2 \ 4 \ 1 \ 4$ — число новых категорий, $\gamma_i \in h_{gi}$

Гиперпараметры модели подбирались автоматически. При генерации поколения присутствовали особи с различным числом категорий. Наилучшее качество достигается оптимальным значением гиперпараметров. На графике представлена зависимость качества для всего поколения на каждом шаге итерационной процедуры оптимизации.

Каждая линия на графике соответствует одному шагу оптимизации. По оси абс-



цисс отложен упорядоченный по качеству номер особи в одном поколении. Видно, что качество решения быстро сходится к оптимальному значению, а наилучшее задает максимальное качество.

Анализируя полученные разбиения исходных признаков видим (привести примеры), что разбиение, на котором достигает максимум функционала качества, является интерпретируемым. Например, признак кредитной истории, содержащий 5 категорий разбивается на две категории: "Просрочил платеж в прошлом", "Критический аккаунт" и "Не брал", "Выплачивал", "Выплачивал до настоящего момента", что является интерпретируемым результатом: негативные примеры объединились в одну категорию, а позитивные и нейтральные - в другую.

Аналогичный эксперимент проведен для стадии порождения признаков. Каждая линия на графе также отвечает одному поколению. Видно, что лучшая особь каждого поколения имеет возрастающее значение функционала качества. Проанализируем построенные признаки, например: $\chi'_1 = \frac{1}{\chi_1 + \chi_2 + \chi_3}$, $\chi'_2 = \sqrt{\chi_2 + \sqrt{\chi_1 + \chi_3}}$. Некоторые из этих признаков являются интерпретируемыми, но в некоторых существует несогласованность операций и размерностей: складываются например кол-во месяцев на текущей работе с кол-вом взятых ранее кредитов.

Решение обладает высокой степенью интерпретируемости. Необходимо сравнить качество и сложность моделей, чтобы убедиться в обоснованности применения нового метода.

Модель	Качество (\mathcal{L})	Качество (\mathcal{T})	Число параметров
	Roc-Auc	Roc-Auc	
Logistic Regression	0.751	0.709	6
End-to-end LR	0.782	0.730	25
XGBoost	0.912	0.702	5000
NN	0.834	0.690	3000

Из результатов вычислительного эксперимента видно, что качество построенного решения при малой сложности признакового пространства и пространства параметров и гиперпараметров остается на высоком уровне, уровне моделей со сложной и неинтерпретируемой структурой. Такой результат был достигнут за счет построенной процедуры порождения и изменения признакового пространства.

СПИСОК ЛИТЕРАТУРЫ

1. L. J. Sanchez-Barrios, G. Andreeva, J. Ansell. Time-to-profit scorecards for revolving credit // European Journal of Operational Research. 2016. Vol. 249, Iss. 2. Pp. 397–406.
2. S. Maldonado, J. Pereza, C. Bravo. Cost-based feature selection for Support Vector Machines: An application in credit scoring // European Journal of Operational Research. 2017. Vol. 261, Iss. 2. Pp. 656–665.
3. C. Luo, D. Wu, D. Wu. A deep learning approach for credit scoring using credit default swaps // Engineering Applications of Artificial Intelligence. 2017. Vol. 65. Pp. 465–470.
4. X. Yufei, L. Chuanzhe, L. Y. Ying, L. Nana. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring // Expert Systems with Applications. 2017. Vol. 78. Pp. 225–241.
5. X. Chen, C. Zhou, X. Wang, Y. Li. The Credit Scoring Model Based on Logistic-BP-AdaBoost Algorithm and its Application in P2P Credit Platform // Proceedings of the Fourth International Forum on Decision Sciences. 2017. Pp. 119–130
6. G. Zeng. Invariant properties of logistic regression model in credit scoring under monotonic transformations // Communications in Statistics - Theory and Methods. 2017. Vol. 46. Iss. 17. Pp. 8791–8807.
7. Y. Wang, J. L. Priestley. Binary Classification on Past Due of Service Accounts using Logistic Regression and Decision Tree // Grey Literature from PhD Candidates. 2017. Vol. 4. <http://digitalcommons.kennesaw.edu/dataphdgreylit/4>

8. S. Walusala W, R. Rimiru, C. Otieno. A hybrid machine learning approach for cregit scoring using PCA and logistic regression // International Journal of Computer. 2017. Vol. 27. Iss. 1.
9. A. Mathew. Cregit scoring using logistic regression // San Jose State University. 2017. Master's Projects. 532. http://scholarworks.sjsu.edu/etd_projects/532
10. R. J. Connor. Grouping for Testing Trends in Categorical Data // Journal of the American Statistical Association. 1972. Vol. 67. Iss. 339.
11. Гончаров А. В., Попова М. С., Стрижов В. В. Метрическая классификация временных рядов с выравниванием относительно центроидов классов // Системы и средства информатики. 2015. **25**. № 4. С. 52–64.
12. Keogh E. J., Ratanamahatana C. A. Exact indexing of dynamic time warping // Knowl. Inf. Syst. 2005. **7**. N 3. P. 358–386.