

Построение интерпретируемых моделей глубокого обучения в задаче социального ранжирования

Гончаров Алексей Владимирович

Московский физико-технический институт
Кафедра интеллектуальных систем, ФУПМ.
Научный руководитель: д.ф.-м.н. В. В. Стрижов.

ММРО-2017

Предложить решение задачи социального скоринга и систему, создающую эксплуатируемую модель без участия аналитика (end-to-end). Модель строится по принципу глубокой сети как интерпретируемая суперпозиция.

Проблема

- В индустрии модель строится поэтапно, каждый этап оптимизируется отдельно

Требования

- Интерпретируемость модели оценивается экспертами в предметной области
- Оптимизация набора эксплуатационных критериев качества

Предлагается

- Единая процедура оптимизации структуры и параметров

Наблюдается быстрый рост сложности при медленном росте качества в задаче социального ранжирования:

Модель	Качество (\mathcal{L})	Качество (\mathcal{T})	Число параметров
Logistic Regression	55%	53%	12
Deep Neural Nets	59%	57%	240
Random Forest	61%	57%	1400
XGBoost	63%	58%	12000

Мотивация: предложить процедуру с ограниченной сложностью, но высоким качеством.

Дано

- $\mathcal{D} = \{(\mathbf{x}_i, y_i) : i = 1, \dots, m\}$, $\mathbf{x}_i \in \mathbb{R}^l \times \mathbb{C}^o \times \mathbb{B}^b$, $y_i \in \{0, 1\}$
- Индексы объектов $\{i = 1, \dots, m\} = \mathcal{I} = \mathcal{L} \sqcup \mathcal{T}$
- Индексы признаков $\mathcal{A} = \mathcal{A}_l \sqcup \mathcal{A}_n \sqcup \mathcal{A}_o \sqcup \mathcal{A}_b$
- Матрица плана $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{X} = (\chi^1, \dots, \chi^n)$.
- Модель классификации $f(\mathbf{X}, \mathbf{w})$ и критерий качества S

Требуется найти оптимальные параметры модели на обучающей выборке:

$$\mathbf{w} = \underset{\mathbf{w}}{\operatorname{argmin}} S(\mathbf{w} | \mathcal{D}, \mathcal{A}, \mathcal{L}).$$

В задачах социального ранжирования (скоринга) используют следующие критерии качества:

Эксплуатационные критерии качества:

- AUC — площадь под ROC-кривой
- Сложность, число признаков
- Устойчивость во времени

Оптимизационные критерии качества:

- Правдоподобие модели
- Устойчивость — ортогональность признаков
- Точность, полнота

Многокритериальная оптимизация

$$S(\mathbf{w}) = \sum_j \alpha_j S_j(\mathbf{w}),$$

где α_j заданы при помощи экспертной оценки.

Процедура построения модели:

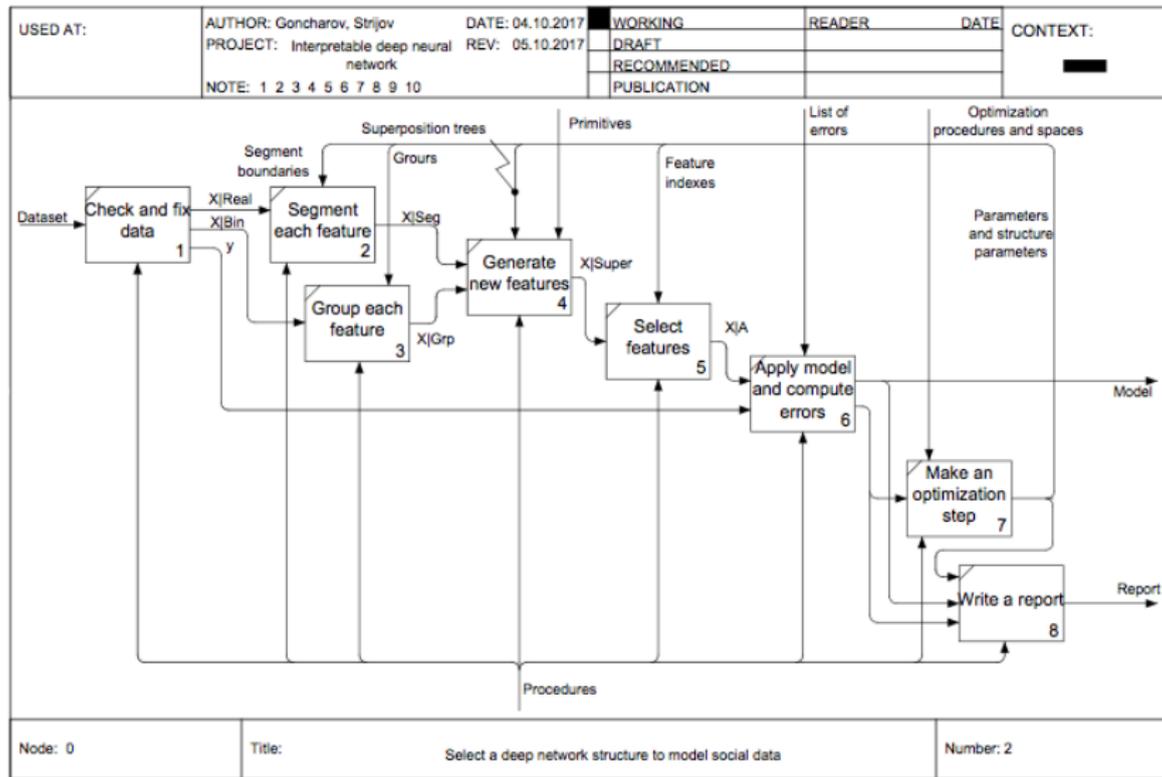
- Сегментация
- Группировка
- Порождение признаков
- Построение линейной модели

Стандартный подход

Оптимизации в каждой из этих задач локальных критериев качества, которые не связаны с целевым критерием.

Предлагается

Единая процедура оптимизации: пошаговая итерационная оптимизация целевого заданного критерия качества.



Пусть линейный признак $x \in [a, b]$. Структурным параметром Θ процедуры сегментации является количество узлов разбиения.

Параметрами θ являются координаты узлов разбиения, причем $\theta^j \in [0, 1], j = 0, \dots, K_1; a = \theta^0 < \dots < \theta^j < \dots < \theta^\Theta = b$.

Исходному признаковому пространству в таком случае добавляется новое множество признаков $x_\theta \in \mathbb{B}^\Theta : x_\theta^i = [x \in [\theta^{i-1}, \theta^i)], i = 1, \dots, \Theta$.

Для каждого сегментируемого признака решается следующая оптимизационная задача:

$$\Theta, \mathbf{w}_\Theta, \theta, \mathbf{w}_\mathcal{A} = \operatorname{argmin}_{\Theta, \mathbf{w}_\Theta, \theta, \mathbf{w}_\mathcal{A}} S(\Theta, \mathbf{w}_\Theta, \theta, \mathbf{w}_\mathcal{A} | \mathcal{D}, \mathcal{A}).$$

Процедура группировки

Пусть категориальный признак x имеет K категорий.
Структурным параметром процедуры группировки является количество групп, $|\Gamma| < K$. Параметром группировки является сюръекция $h : \mathbb{B}^K \rightarrow \mathbb{B}^{|\Gamma|}$.

Исходное признаковое пространство изменяется:
категориальный признак x изменяется на категориальный признак x_h .

x	$=$	1	2	3	...	K	K — число категорий
		\downarrow	\downarrow	\downarrow		\downarrow	
x_h	$=$	γ_1	γ_2	γ_3	...	γ_K	$ \Gamma $ — число групп, $\gamma_i \in \Gamma$

Для каждого группируемого признака решается следующая оптимизационная задача:

$$|\Gamma|, \mathbf{w}_{|\Gamma|}, h, \mathbf{w}_{\mathcal{A}} = \operatorname{argmin}_{|\Gamma|, \mathbf{w}_{|\Gamma|}, h, \mathbf{w}_{\mathcal{A}}} S(|\Gamma|, \mathbf{w}_{|\Gamma|}, h, \mathbf{w}_{\mathcal{A}} | \mathcal{D}, \mathcal{A}).$$

Множество интерпретируемых порождающих функций G

Описание	Формула	in	N	in	out
Conjunction	$x \wedge y \wedge \dots$	bin	2-6		bin
Disjunction	$x \vee y \vee \dots$	bin	2-6		bin
Negate binary	\bar{x}	bin	1		bin
Logarithm	$\log(x)$	lin	1		lin
Logistic sigmoid	$\frac{1}{1+\exp^x}$	lin	1		lin
Inverse	$\frac{1}{x}$	lin	1		lin
Multiplication	$x * y * \dots$	lin	2-3		lin
Polynomial transformation	$\sum_x a_x x^{\alpha_x}$	lin	1-3		lin

- Порожденные признаки $\mathcal{F} = \{f_i\}$ — это допустимые суперпозиции порождающих функций
- Сложность модели — количество используемых функций и исходных признаков в суперпозиции

- Измеряемые признаки $\mathbf{X} = \{\chi\}$
- Заданные экспертами порождающие функции $G = \{g(b, \chi)\}$
- Правила порождения: $\mathcal{G} \ni G$, суперпозиция $g_k \circ g_l \in \mathcal{G}$ построена с учетом ограничений на сложность, типы входных и выходных переменных.
- Правила упрощения: g_u не принадлежит \mathcal{G} , если существует правило $g_u \rightarrow g_v \in \mathcal{G}$

Набор порожденных признаков $F(X) = \{f_i(X)\}$ находится внутри процедуры оптимизации:

$$F(X), \mathbf{w}_F, \mathbf{w}_A = \operatorname{argmin}_{F(X), \mathbf{w}_F, \mathbf{w}_A} S(F(X), \mathbf{w}_{F(X)}, \mathbf{w}_A | D, A).$$

Необходимо оптимизировать функцию потерь S . Оптимальные параметры процедур построения модели получаем при решении следующей оптимизационной задачи:

$$\begin{aligned} & \{\Theta\}, \{\theta\}, \{|\Gamma|\}, \{h\}, \{\mathbf{w}_\Theta\}, \{\mathbf{w}_{|\Gamma|}\}, F(X), \mathbf{w}_F, \mathbf{w}_A = \\ & = \operatorname{argmin}_{\dots} S(\dots | D, \mathcal{A}), \end{aligned}$$

где $\{.\}$ обозначают множества параметров для каждой целевой группы признаков.

Сегментация

$\theta_i \in [a, b], i = 1, \dots, \Theta$ — непрерывные/дискретные. Задача решается методом Нелдера-Мида.

Группировка

$h : \mathbb{B}^K \rightarrow \mathbb{B}^{|\Gamma|}$ — сюръективное отображение между двумя конечными множествами. Задача решается генетическим алгоритмом.

Порождение признаков

$F(X) = \{f_i(X)\}$ — множество суперпозиций элементарных функций. Задача поиска оптимального набора решается генетическим алгоритмом

Построение модели

w_d — непрерывные веса модели. Задача решается методом градиентного спуска.

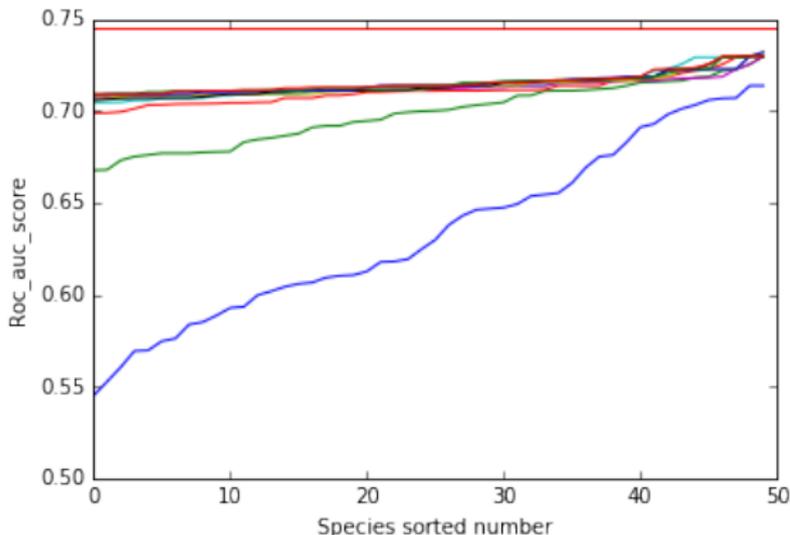
Далее приводится алгоритм решения единой оптимизационной задачи

- 1 Инициализация стартовых точек для Нелдера-Мида и градиентного спуска, инициализация первого поколения для генетических алгоритмов.
- 2 Шаг Нелдера-Мида для задачи сегментации
- 3 Шаг генетического алгоритма группировки
- 4 Шаг генетического алгоритма порождения признаков
- 5 Шаг градиентного спуска построения модели
- 6 Повторять 2-5 до сходимости функции ошибки

Вычислительный эксперимент

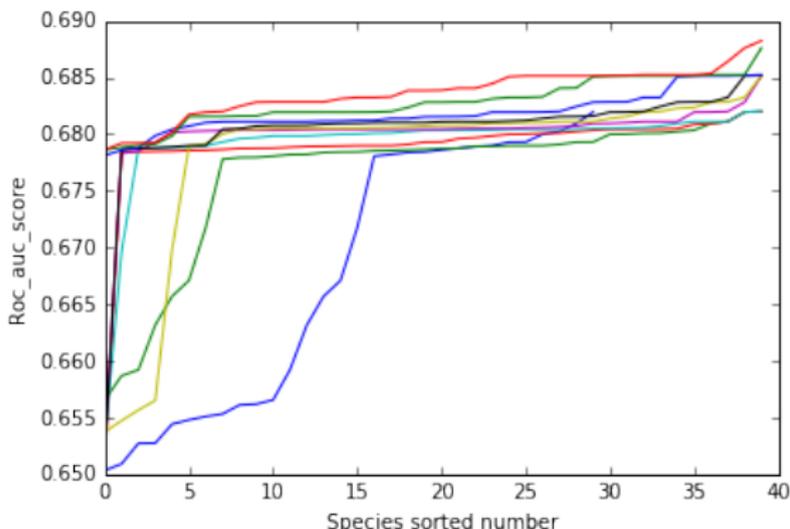
Вычислительный эксперимент проводился на выборке "Statlog (German Credit Data)", UCI.

Оптимизация процедуры группировки



Признак: кредитная история. Исходных категорий: 5, новых: 2
1: "Просрочил платеж в прошлом", "Критический аккаунт"

Оптимизация процедуры порождения признаков



Пример порожденных признаков:

$$\chi'_1 = \frac{1}{\chi_1 + \chi_2 + \chi_3}$$

$$\chi'_2 = \sqrt{2\chi_2} + \sqrt{\chi_1 + \chi_2}$$

Вычислительный эксперимент проводился на выборке "Statlog (German Credit Data)", UCI.

Модель	Качество (\mathcal{L})	Качество (\mathcal{T})	Число параметров
	Roc-Auc	Roc-Auc	
Logistic Regression	0.76	0.73	9
End-to-end LR	0.80	0.77	50
Random Forest	0.100	0.76	1200
XGBoost	0.92	0.77	10000

- Предложена система создания эксплуатируемой модели без участия аналитика (end-to-end)
- Модель построена по принципу глубокой сети как интерпретируемая суперпозиция
- Получено интерпретируемое решение
- Решена проблема поэтапной оптимизации локальных критериев качества

Планы

- Усовершенствовать имеющуюся процедуру оптимизации модели
- Изменить процедуру порождения признаков