



НЕЙРОСЕТЕВЫЕ МОДЕЛИ РАНЖИРОВАНИЯ
ДЛЯ ЗАДАЧ МЕТААВТОМАТИЧЕСКОГО
МАШИННОГО ОБУЧЕНИЯ

Мардышкин Ростислав Романович
Университет ИТМО, ПИиКТ

Шиков Егор Николаевич, к. т. н.

СТРУКТУРА ПРЕЗЕНТАЦИИ

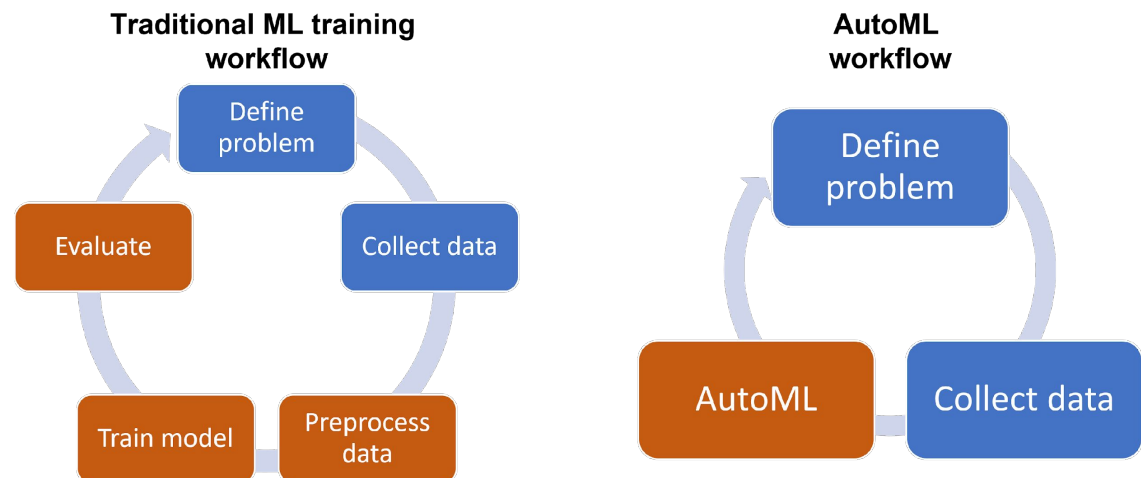
1. Введение в метаавтоматическое машинное обучение
2. Актуальность
3. Описание существующего решения и предложение нововведений
4. Описание внедряемого решения
5. Эксперименты
6. Результаты
7. Заключение



ВВЕДЕНИЕ В МЕТААВТОМАТИЧЕСКОЕ МАШИННОЕ ОБУЧЕНИЕ

AutoML — это процесс создания пайплайна машинного обучения, который включает в себя подготовку данных, конструирование признаков, генерацию моделей и их оценку.

Мета-обучение позволяет найти оптимальный ML-пайплайн за счет использования суррогатных моделей на мета-признаках.



<https://learn.microsoft.com/ru-ru/dotnet/machine-learning/automated-machine-learning-mlnet>



АКТУАЛЬНОСТЬ

AutoML

- Выбор модели
- Проектирование
- Оптимизация

оценка большого количества
различных моделей и конфигураций

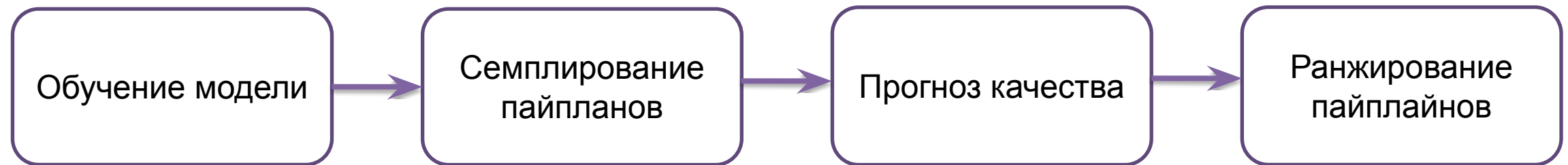
сложный и трудоемкий процесс

Решение:

использование уже известной информации о ранее обученных моделях на аналогичных данных

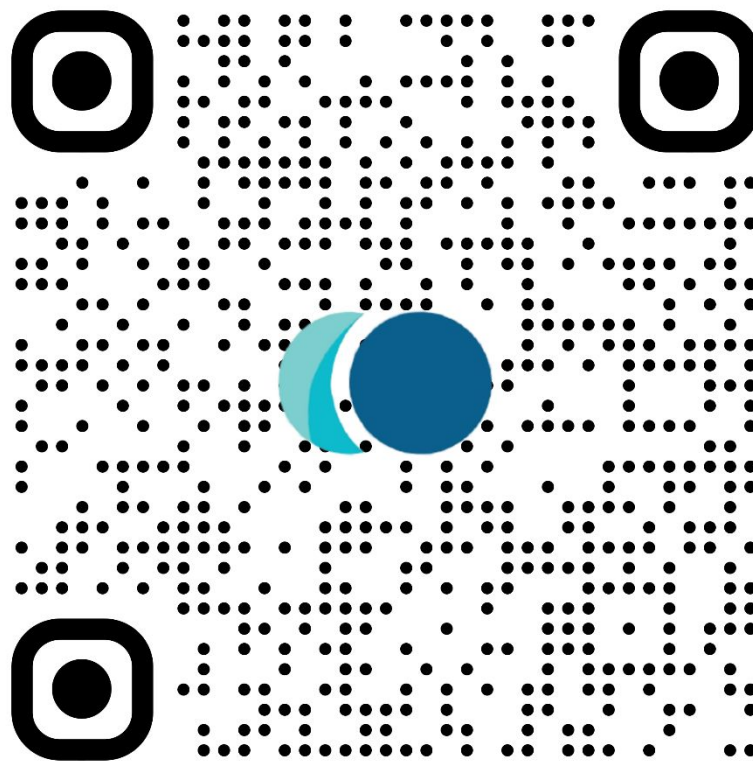
СУРРОГАТНАЯ МОДЕЛЬ

Суррогатная модель, которая будет оценивать качество пайплайнов



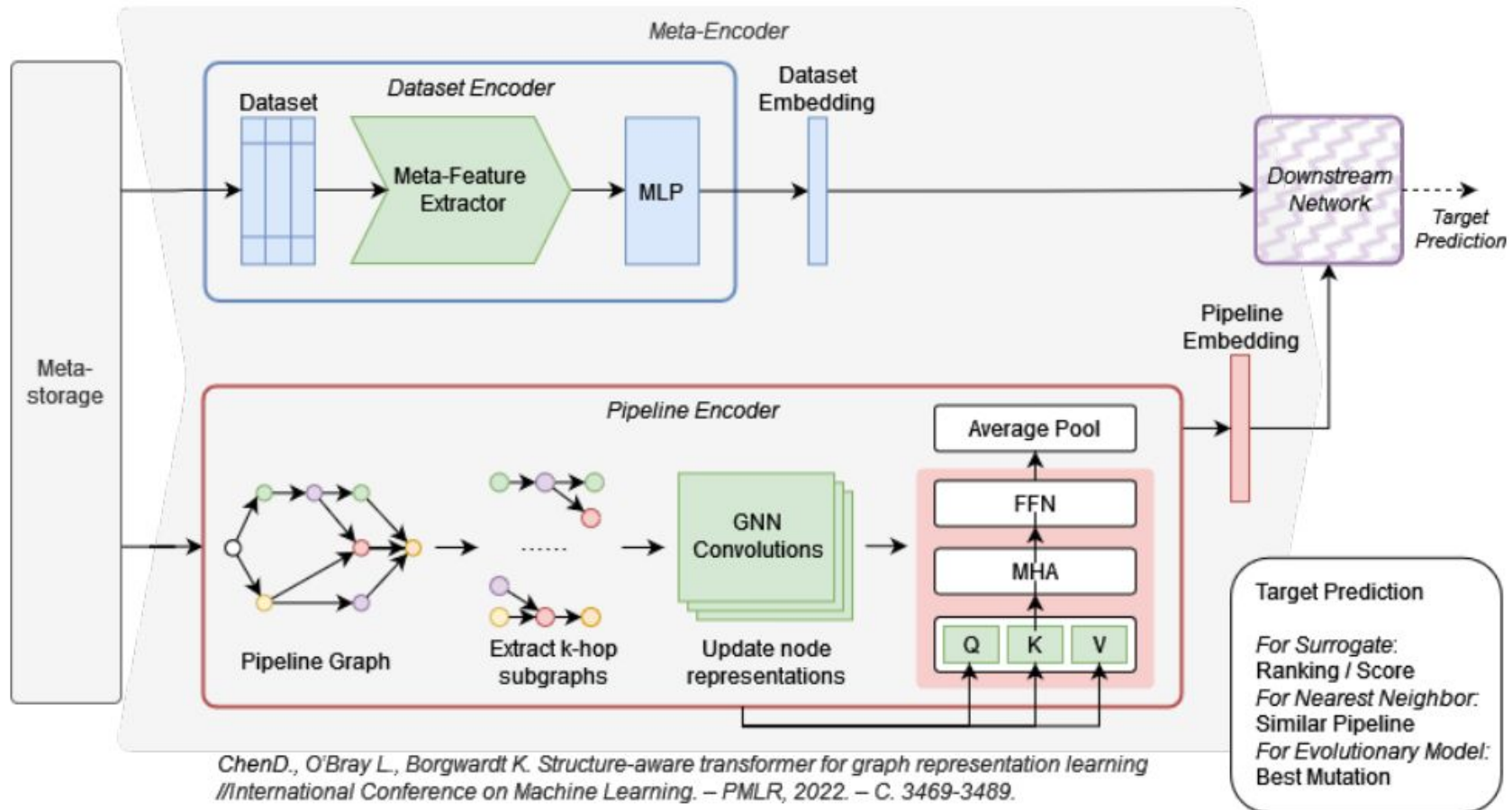
GAMLET

GAMLET - это открытая платформа для разработки в области метаобучения в AutoML



Лицензия BSD-3-Clause

СУРРОГАТНАЯ МОДЕЛЬ



- нейросеть для обработки датасетов
- графовая нейросеть для обработки пайпланов ML
- конечная модель для прогнозирования результата пайплайна на датасете

ПОСТАНОВКА ЗАДАЧИ МО

Как обучать суррогатную модель?

1) Регрессия

Для снижения разброса проводим нормализацию (pipeline, task)

2) Бинарная классификация

Попадает ли пайплайн в ТОПк для данного датасета?

Если да - «1», если нет - «0».

3) Ранжирование. Попарный подход

Рассматриваем пары строк в ранжированном списке
Модель строится сведением к минимизации количества пар, в которых был предсказан неправильный порядок.

$$\sum_{i=1}^l (a(q_i, d_i) - y(q_i, d_i))^2 \rightarrow \min$$

$$a: X \rightarrow \mathbb{R} \quad \sum_{x_i < x_j} [a(x_j) - a(x_i) < 0] \rightarrow \min \quad \longrightarrow \quad \begin{aligned} & \sum_{x_i < x_j} L(a(x_j) - a(x_i)) \rightarrow \min \\ & L(M) = \log(1 + e^{-M}) \end{aligned}$$

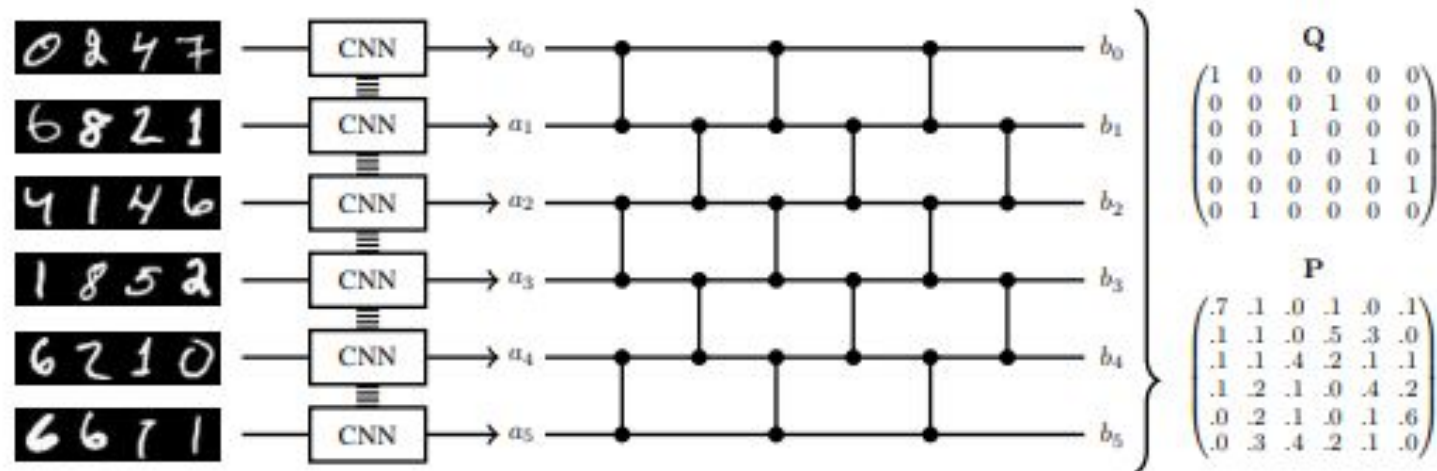


ПОСТАНОВКА ЗАДАЧИ МО

Предлагаемый подход:

Нейронная сортировка. В нейронной сортировке используется архитектура, объединяющая две сети: сверточная нейронная и нечетно-четная сортирующая сеть.

Обучить можно с помощью кросс-энтропии между полученной матрицей перестановок и истинной матрицей перестановок.

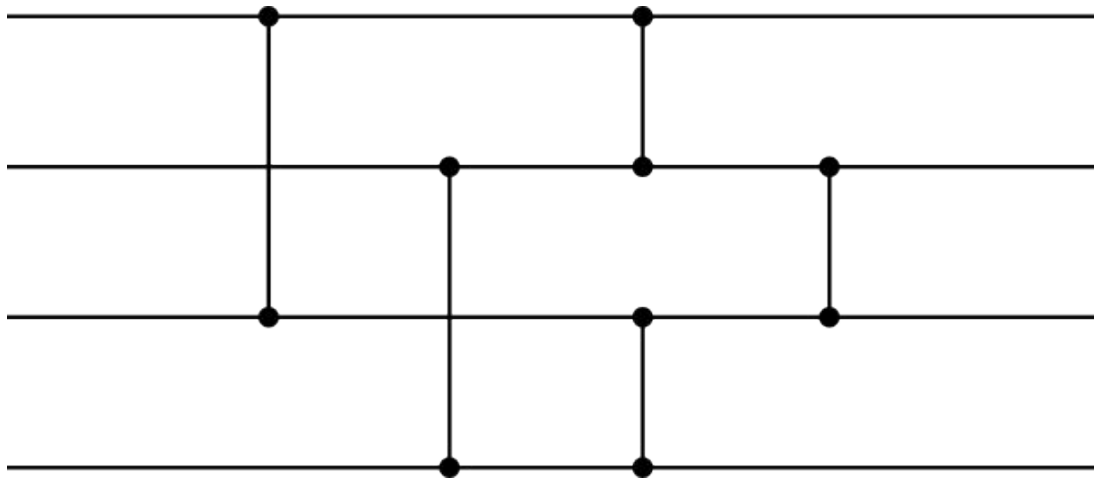


F. Petersen, C. Borgelt, H. Kuehne, and O. Deussen, "Monotonic differentiable sorting networks," in International Conference on Learning Representations (ICLR), 2022.

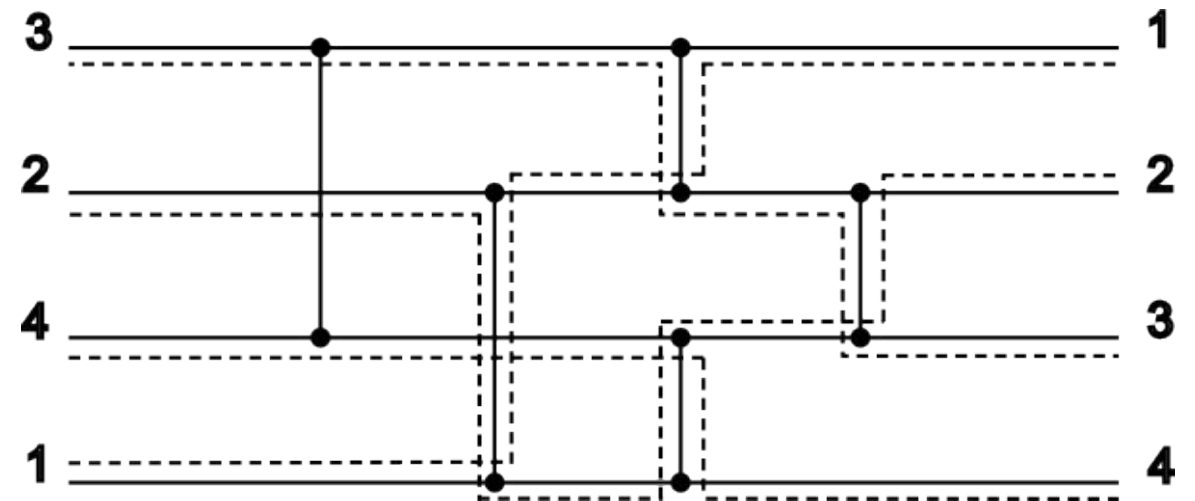


СОРТИРУЮЩАЯ СЕТЬ

Сортирующая сеть (англ. *Sorting network*) — метод сортировки, основанный только на сравнениях данных.



Схематическое изображение сортирующей сети для последовательности из 4 чисел. Глубина сети: 4. Размер сети: 5



Процесс сортировки числовой последовательности (3, 2, 4, 1)

https://neerc.ifmo.ru/wiki/index.php?title=Сортирующие_сети



ЭКСПЕРИМЕНТЫ

FEDOT мета-датасет

Кол-во уникальных элементов пайплайна	>50
Количество датасетов	43
Количество пар датасет-пайплайн	880314
Среднее кол-во элементов в пайплайне	5.3
Среднее кол-во связей в пайплайне	6.7

Метрики качества:

Насколько релевантно отранжировали список пайплайнов по качеству для тестовых датасетов?

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i},$$

$$Hits @ N = \frac{1}{|T|} \sum_{i=1}^{|T|} q(rank_i, N), \quad q(x, N) = \begin{cases} 0, & x > N \\ 1, & x \leq N \end{cases}.$$

$$NDCG@K = \frac{1}{|Users|} \sum_{u=1}^{|Users|} \frac{\sum_{k=1}^K \frac{2^{grade_k - 1}}{\log_2(k+1)}}{\sum_{k=1}^{|REL_K|} \frac{2^{grade_k - 1}}{\log_2(k+1)}},$$

РЕЗУЛЬТАТЫ

Подход		NDCG	MRR	HITS
Одинарный (регрессия метрики)		0,4165	0,6128	0,1433
Попарный		0,5007	0,6709	0,2520
Нейронная сортировка	<i>bitonic k=3</i>	0,5149	0,6721	0,2680
	<i>odd_even k=3</i>	0,5138	0,6884	0,2751
	<i>bitonic k=4</i>	0,5304	0,6766	0,2780
	<i>odd_even k=4</i>	0,5130	0,6700	0,2511
	<i>bitonic k=5</i>	0,5088	0,6955	0,2463
	<i>odd_even k=5</i>	0,5095	0,6818	0,2596

ЗАКЛЮЧЕНИЕ

- Реализован модуль нейронной сортировки и внедрен в текущую архитектуру GAMLET;
- Внедренный метод нейронной сортировки дает прирост метрики $NDCG@k$ на 6-7% по сравнению с попарным подходом;
- Лучшая выявленная конфигурация для нашей архитектуры на предложенных данных: bitonic (сеть Бетчера) и 4 пайплайна;





СПАСИБО ЗА ВНИМАНИЕ!

iT's **MO** *re than a*
UNIVERSITY

Ваши контакты

