

Поиск архитектуры смеси экспертов с помощью суррогатной модели

Бабкин Пётр

Научный руководитель: к.ф.-м.н. Бахтеев Олег

Кафедра интеллектуальных систем ФПМИ МФТИ
Специализация: Интеллектуальный анализ данных
Направление: 03.04.01 Прикладные математика и физика

2024

План работы

Цель исследования:

Предложить метод поиска архитектуры для смеси экспертов.

Проблема:

Адаптация методов поиска архитектуры нейронной сети для задачи поиска архитектуры смеси экспертов неэффективна: сложность поиска растет линейно с числом экспертов.

Предлагаемый метод:

Предлагается оценивать качество архитектур каждого из экспертов с помощью суррогатной функции. Вспомогательным критерием качества является эмперический риск на результатах кластеризации выборки.

Постановка задачи

Общая задача поиска архитектуры нейронной сети

$$\begin{aligned} & \min_{\alpha \in \mathcal{A}} \mathcal{L}_{\text{val}}(\alpha, \mathbf{w}_\alpha^*), \\ \text{s.t. } & \mathbf{w}_\alpha^* = \arg \min_{\mathbf{w} \in \mathcal{W}_\alpha} \mathcal{L}_{\text{train}}(\alpha, \mathbf{w}), \end{aligned}$$

где \mathcal{A} — пространство архитектур, \mathcal{W}_α — пространство параметров архитектуры α .

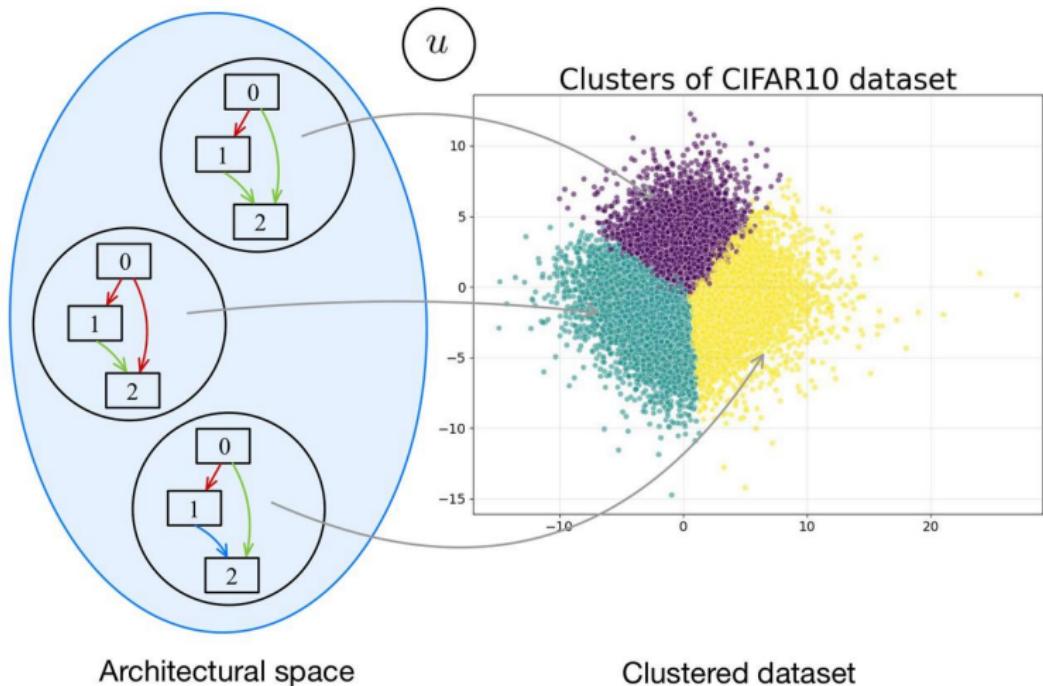
Задача по поиску архитектуры смеси из K экспертов

$$\begin{aligned} & \min_{\alpha \in \mathcal{A}^K} \sum_{k=1}^K \mathcal{L}_{\text{val}}^{\mathcal{D}_k}(\alpha_k, \mathbf{w}_{\alpha_k}^*), \\ \text{s.t. } & \forall k \in \{1, 2, \dots, K\}, \quad \mathbf{w}_{\alpha_k}^* = \arg \min_{\mathbf{w} \in \mathcal{W}_{\alpha_k}} \mathcal{L}_{\text{train}}^{\mathcal{D}_k}(\mathbf{w}, \alpha_k), \end{aligned}$$

где \mathcal{D}_k — кластер датасета, K — количество кластеров.

Схема работы

Архитектуры экспертов подбираются для каждого кластера с помощью суррогатной модели u .



Суррогатная модель для оценки качества экспертов

Мы оцениваем точность работы экспертов по кластерам с помощью суррогатной модели

$$u : \mathcal{A} \times \mathbb{R}^m \rightarrow \mathbb{R}^K,$$

где \mathbb{R}^m — пространство параметров гиперсети.

Результат работы суррогатной модели $v \in \mathbb{R}^K$ — точность архитектуры на K кластерах.

Предлагаемый алгоритм SA-MoENAS

Считаем, что есть датасет обученных архитектур \mathcal{D}_{arch} .

Algorithm SA-MoENAS

Initialize: $K \in \mathbb{N}$, $n \in \mathbb{N}$, \mathcal{D}_{arch} — датасет обученных архитектур, D — первичный датасет

$D' \leftarrow VAE(D)$ — латентное представление датасета

$\{D_1, \dots, D_K\} \leftarrow K\text{-means}$ — кластеризация датасета

$u \leftarrow$ обучение суррогатной модели

Сгенерировать n архитектур

$\{\alpha_1, \dots, \alpha_K\} \leftarrow$ выбрать архитектуры экспертов с помощью u

Цели вычислительного эксперимента

Оценка работы суррогатной модели

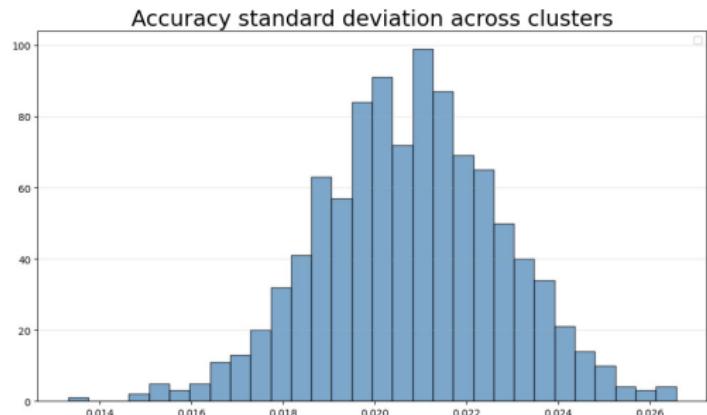
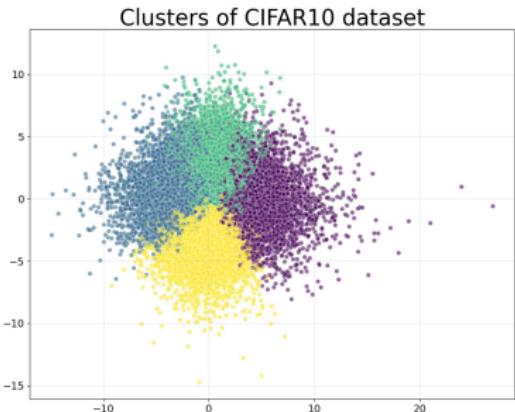
- ▶ Кластеризовать датасет с помощью VAE, проверить эффективность.
- ▶ Проверить, может ли суррогатная модель предсказывать точности архитектур по кластерам.

Оценка качества итоговой смеси экспертов

- ▶ Проверить эффективность работы итоговой смеси экспертов по сравнению с копированием архитектуры.

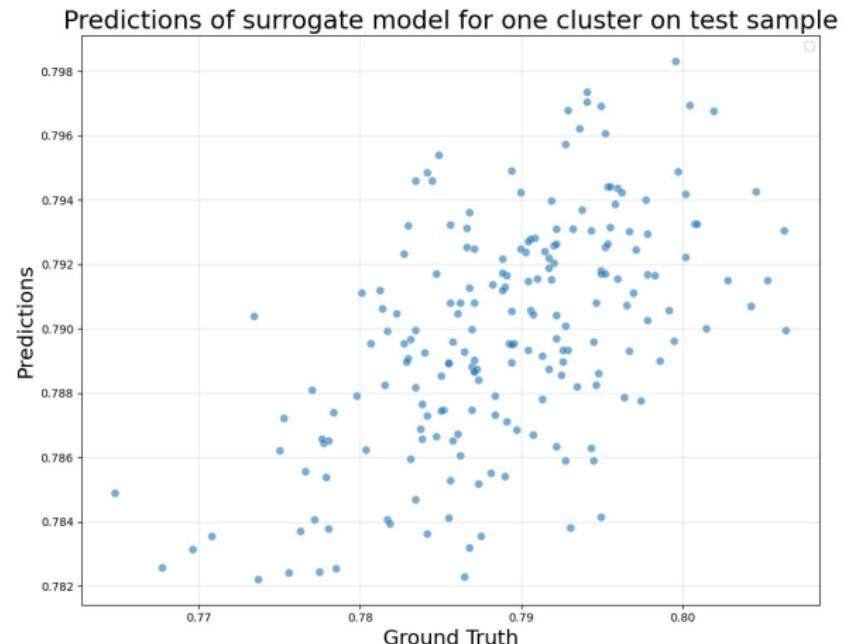
Кластеризация датасета

Существует разброс точности архитектуры по кластерам.



Разброс в среднем составляет 2%. Минимум он составляет 1,2%.

Оценка качества работы суррогатной функции



Суррогатная модель корректно предсказывает точность

Заключение

Достигнутые результаты:

- ▶ Предложен метод для поиска архитектуры смеси экспертов
- ▶ Показана возможность предсказания точности архитектур по кластерам

Дальнейшая работа

- ▶ Планируется подача работы по суррогатным моделям вместе с Александром Уденеевым в феврале
- ▶ Устранить разрыв, возникающий из-за того, что модели в датасете обучаются на всех данных, в то время как в смеси экспертов каждая модель учится на своей части данных
- ▶ Провести больше экспериментов на разных датасетах, сравниться с конкурирующими методами