

Структурное прореживание моделей глубокого обучения на основе суррогатных моделей

М. О. Иванов

`ivanov.mo@phystech.edu`

Москва, Московский физико-технический институт

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

2025

Цель исследования

Цель

Предложить метод структурного прореживания нейросетей.

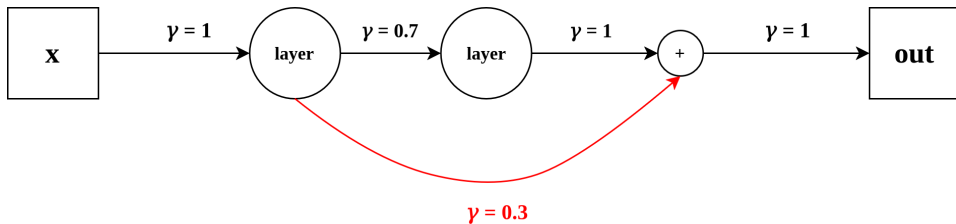
Проблема

Существующие методы работают послойно. Учёт взаимосвязей между слоями модели является вычислительно сложной задачей.

Метод решения

Предлагаемый метод основан на построении суррогатной графовой модели, повторяющей структуру исходной нейросети.

Схема предлагаемого решения



Упрощённая схема residual connection. Каждому ребру сопоставляется γ , по множеству которых граф оптимизируется.

Постановка задачи

- ▶ Архитектура модели представляет собой ориентированный ациклический граф $G = (V, E)$. Каждая вершина $v_i \in V$ соответствует слою нейронной сети, и каждое ребро $e_{ij} \in E$ представляет поток данных из выхода слоя v_i до входа v_j .
- ▶ В данной формулировке удаление вершины означает удаление соответствующего слоя из сети, тогда как удаление ребра к устранению зависимости передачи данных между слоями (например, в случае residual connection).
- ▶ Задана выборка $\mathcal{D} = \mathcal{D}_{\text{train}} \cup \mathcal{D}_{\text{val}}$. Задана функция потерь \mathcal{L} . Пусть \mathbf{w} – параметры модели до прунинга, а \mathbf{w}' – параметры модели после. Задача оптимизации

$$\min_{\mathbf{w}' \in \mathbb{R}^k} |\mathcal{L}(\mathcal{D} \mid \mathbf{w}') - \mathcal{L}(\mathcal{D} \mid \mathbf{w})|.$$

- ▶ Пусть задано $\gamma_{ij} \in [0, 1]$ для каждого ребра $e_{ij} \in E$. Также для графа G задана маска рёбер $\mathbf{M} \in \mathbb{Z}_2^{|E|}$:

$$(\mathbf{M})_{ij} = 1 \quad \Leftrightarrow \quad \gamma_{ij} = 0.$$

- ▶ Тогда оптимизационная задача:

$$\gamma = \arg \min_{\gamma \in [0,1]^{|E|}} \left(E_{\mathbf{M}} \|G(\gamma, \mathbf{M}) - \mathcal{L}(\mathbf{M})\|_2^2 \right).$$

Постановка эксперимента

- ▶ Цель – сравнение качества базовых суррогатных моделей для предложенного метода.
- ▶ В эксперименте проводится прунинг ResNet-50, фиксируется число удаляемых рёбер – 15%.
- ▶ В сравнении участвуют следующие базовые методы:
 - ▶ **(LinReg)**: Линейная регрессия,
 - ▶ **(LogReg)**: Логистическая регрессия,
 - ▶ **(FCN)**: Простая нейросеть.

Результаты эксперимента

- ▶ Рассматриваются задачи классификации изображений из датасета CIFAR10 на 2 класса.
- ▶ Приводится точность предсказаний метрики AUC на контроле после прунинга.

Method	MAPE Test
LinReg	19.6%
LogReg	18.7%
NN	16.8%

Из таблицы видно, что базовые методы решают поставленную задачу, однако точность недостаточна.

Заключение

- ▶ Рассмотрена задача структурного прореживания нейросетей.
- ▶ Предложен метод структурного прореживания нейросетей.
- ▶ Проведены базовые эксперименты для сравнения качества суррогатных моделей.
- ▶ Планируется теоретическое обоснование предложенного метода и его имплементация.