

Дифференцируемый алгоритм поиска архитектуры модели с контролем её сложности

Константин Дмитриевич Яковлев

Московский физико-технический институт

Курс: Автоматизация научных исследований
(практика, В. В. Стрижов)/Группа 874

Эксперт: В. В. Стрижов

Консультант: О. С. Гребенькова, О. Ю. Бахтеев

2021

DARTS и линейная гиперсеть

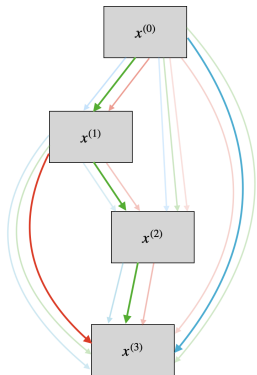


Рис.: Структура ячейки

- ▶ Значения в узлах:
$$x^{(j)} = \sum_{i < j} o^{(i,j)}(x^{(i)})$$
- ▶ Смешанная операция:
$$\hat{o}^{(i,j)}(x) = \sum_{o \in \mathcal{O}} \alpha_o^{(i,j)} o(x)$$
- ▶ Линейная гиперсеть:
$$G_{\text{linear}}(\lambda) = \lambda b_1 + b_2$$
$$\alpha = G_{\text{linear}}(\lambda)$$
- ▶ Получение дискретной архитектуры:
$$o^{(i,j)} = \arg \max_{o \in \mathcal{O}} \alpha_o^{(i,j)}$$