BaNAS

Поиск нейросетевых архитектур (NAS) стал важной задачей в машинном обучении. Предложенный метод BaNAS решает эту задачу через комбинацию байесовской оптимизации и нейросетевых предсказателей, что позволяет эффективно исследовать пространство архитектур.

1. Алгоритм:

1. Выбираются случайных архитектур из пространства поиска.

2. Итерационно проводится обучение ансамбля мета-нейронных сетей на выбранных архитектурах. Каждая сеть ансамбля является сетью прямой связи с полностью связанными слоями, каждому слою даётся случайная инициализация весов и случайный порядок обучающего набора. Используемая функция ошибки — вариация MAPE.

3. Далее формируется набор архитектур-кандидатов посредством случайных изменений лучших архитектур после обучения.

4. Для каждой архитектуры-кандидата определяется значение переданной на вход функции сбора независимой выборки Томпсона.

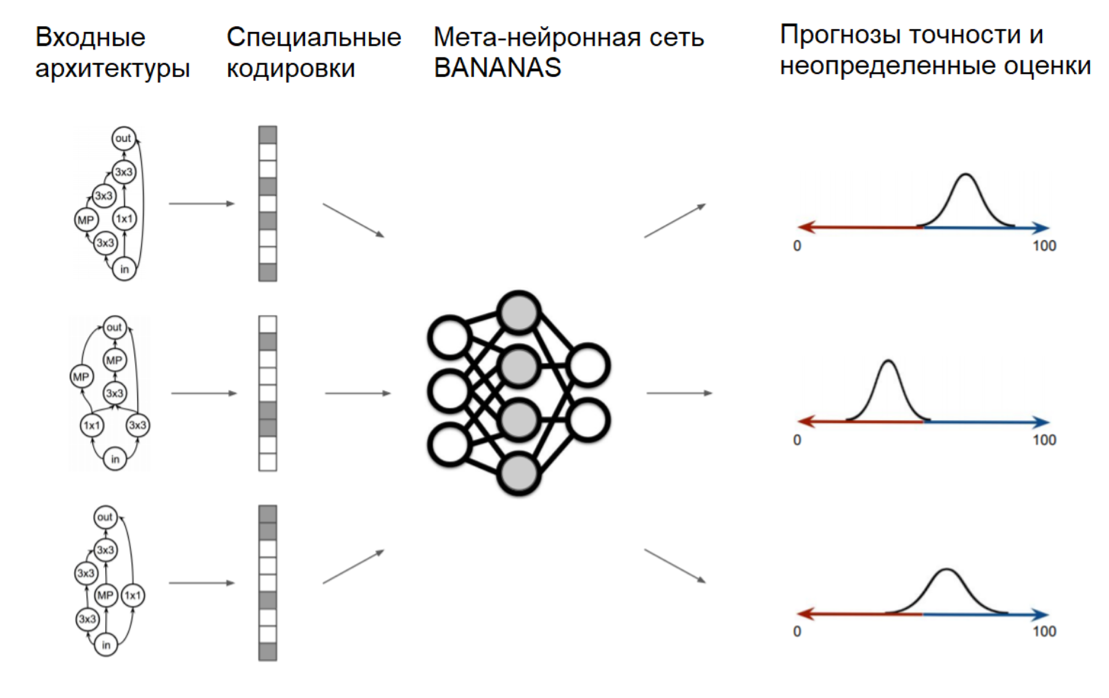
5. Для архитектуры-кандидата с минимальным значением функции сбора определяется значение целевой вероятностной функции.

Рисунок 1 - Алгоритм работы BaNAS

1. Реализация алгоритма

1. Инициализация: Случайный выбор начального набора архитектур

2. Обучение предсказателя: Ансамбль обучается на оцененных архитектурах

3. Генерация кандидатов: Создание новых архитектур через мутацию

4. Выбор архитектур: Использование ITS для выбора перспективных кандидатов

5. Обновление: Сохранение лучшей архитектуры и повторение процесса

1. Оптимизационная задача

Цель: найти архитектуру нейронной сети α∈A, которая или максимизирует кастомную метрику производительности:

где:

A — пространство возможных архитектур,

L(α) — метрика производительности

1. Результаты (в BaNAS – 5 слоев)

