

Экстраполяция кривых обучения в поиске нейронных архитектур (NAS)

Презентация по теме NAS

На основе современных исследований

9 ноября 2025 г.

Основная идея

- Полное обучение каждой архитектуры в NAS слишком дорого.
- Экстраполяция кривых обучения использует ранние эпохи для прогнозирования итоговой производительности.
- Позволяет **преждевременно прекращать обучение** слабых архитектур и сосредоточиться на перспективных.

Суть метода

Частичная тренировка + прогноз → оценка качества архитектуры без полного обучения.

① Экстраполяция кривых обучения:

- Аппроксимация первых эпох кривой обучения.
- Использование регрессии, Gaussian Process или нейросетей.
- Прогноз конечной точности.

② Суррогатные модели:

- Используют характеристики архитектуры (слои, фильтры, тип операций).
- Прогнозируют качество модели без полного обучения.

③ Преждевременное прекращение обучения (Early Stopping):

- Архитектуры с низкой прогнозируемой производительностью отбрасываются.

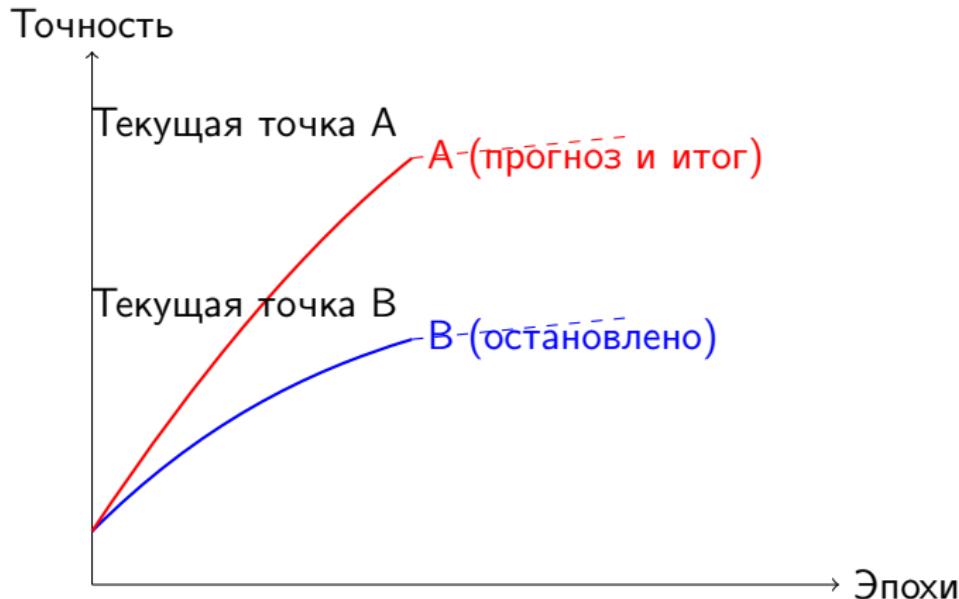
Преимущества

- Снижение вычислительных затрат.
- Возможность ранжирования большого числа архитектур на раннем этапе.
- Фокусировка поиска на перспективных кандидатах.
- Совместимость с другими подходами NAS (One-Shot, Hypernetworks).

Сложности и ограничения

- Необходимость точных прогнозов — ошибка экстраполяции может привести к потере перспективных архитектур.
- Разная скорость обучения архитектур усложняет прогнозирование.
- Выбор модели прогнозирования (регрессия, нейросеть, GP) критически важен.

Иллюстрация: Экстраполяция кривых обучения



- Кривые обучения A и B строятся за первые несколько эпох.
- Экстраполяция позволяет прогнозировать конечную точность.
- Архитектура B прекращается досрочно, архитектура A продолжается.

Список литературы I

-  Baker et al. (2017). *Designing Neural Network Architectures Using Reinforcement Learning.*
-  Domhan et al. (2015). *Speeding up Automatic Hyperparameter Optimization of Deep Neural Networks by Extrapolation of Learning Curves.*
-  Klein et al. (2017). *Learning Curve Prediction with Bayesian Neural Networks.*
-  Liu et al. (2018). *Progressive Neural Architecture Search.*
-  Raval et al. (2019). *Neural Architecture Search via Learning Curve Extrapolation.*
-  Swersky et al. (2014). *Predictive Entropy Search for Bayesian Optimization with Unknown Constraints.*