

## Обзор на статью Learning to Stop While Learning to Predict

<https://arxiv.org/pdf/2006.05082.pdf>

В последнее время растет интерес к изучению связей между моделями глубокого обучения и традиционными алгоритмами машинного обучения. В то время, как в традиционных алгоритмах важной деталью является критерий останова, который может представлять из себя условие сходимости или правило ранней остановки, в моделях глубокого обучения используется фиксированное число используемых слоёв для работы со всеми экземплярами задачи. Подобно критерию останова глубина глубокой архитектуры может быть различной для разных входов. В данной статье для решения проблемы фиксированной глубины предложена архитектура, где глубокая модель обратной связи и критерий останова изучаются вместе, чтобы последовательно определить оптимальное количество слоёв для каждого входного экземпляра. (Fig. 1)

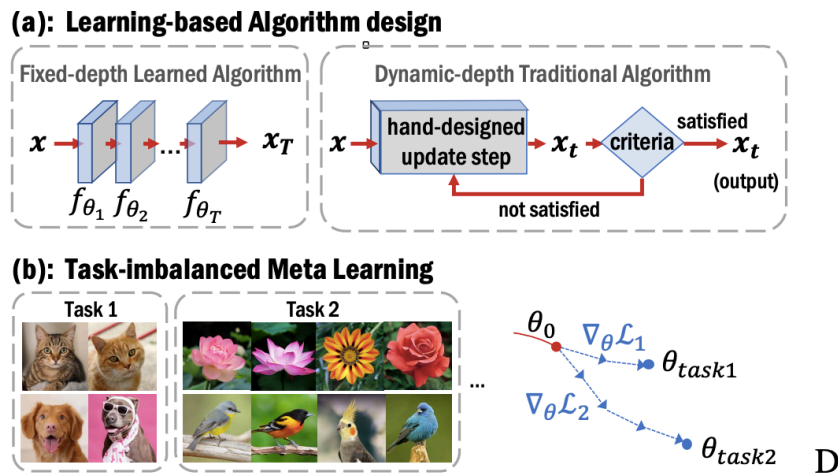


Figure 1. Motivation for learning to stop.

Предложенная авторами модель состоит из следующих частей:

- предсказательной модели  $F$ , которая преобразует входные данные в признаки (или состояния)  $x_1, \dots, x_T$ ;
- критерия останова  $\pi_\phi : (x, x_t) \rightarrow \pi_t \in [0,1]$ , который последовательно просматривает состояния и определяет вероятность останова вычисления на слое  $t$ .

Данные две компоненты позволяют последовательно предсказывать следующее целевое состояние, одновременно определяя, когда следует остановиться.

Предсказательная модель  $F_T$  представляет из себя глубокую модель с  $T$  слоями, на каждом из которых вычисляются состояния  $x_t = f_{\theta_t}(x_{t-1})$ ,  $t = 1, \dots, T$ , где  $x_0$  определяется входным  $x$ . С помощью критерия останова  $\pi_\phi$  на каждом слое  $t$  вычисляется своя вероятность останова вычислений  $\pi_t$ . Тогда распределение времени останова выглядит следующим образом:

$$\begin{cases} q_\phi(t) = \pi_t \prod_{\tau=1}^{t-1} (1 - \pi_\tau) & \text{if } t < T, \\ q_\phi(T) = \prod_{\tau=1}^{T-1} (1 - \pi_\tau) & \text{else.} \end{cases}$$

Время останова  $t$  - дискретная случайная величина с распределением  $q_\phi(t)$ . Учитывая всевозможные стоп-позиции функционал потерь запишем в виде:

$$\mathcal{L}(\theta, q_\phi; x, y) = \mathbb{E}_{t \sim q_\phi} l(y, x_t; \theta) - \beta H(q_\phi),$$

где  $l(y, x_t; \theta)$  - функция потерь,  $H(q_\phi) = - \sum_t q_\phi(t) \log(q_\phi(t))$  - энтропийный регуляризатор,  $\beta$  - коэффициент регуляризации.

При заданном наборе данных  $D = \{(x, y)\}$  параметр предсказательной модели  $\theta$  и критерия останова  $\phi$  могут быть найдены как

$$\min_{\theta, \phi} \frac{1}{|D|} \sum_{(x, y) \in D} \mathcal{L}(\theta, q_\phi; x, y).$$

Для обучения такой модели было предложено декомпозировать задачу на 2 части (Fig. 3):

1. На этапе обучения предсказательной модели нужно найти оптимальное  $\theta$  - параметр предсказательной модели при распределении времени останова oracle

$$q_\theta^*(t | y, x) = \frac{p_\theta(y | t, x)^{\frac{1}{\beta}}}{\sum_{t=1}^T p_\theta(y | t, x)^{\frac{1}{\beta}}} =$$

$$= \frac{\exp(-\frac{1}{\beta} l(y, x_t; \theta))}{\sum_{t=1}^T \exp(-\frac{1}{\beta} l(y, x_t; \theta))},$$

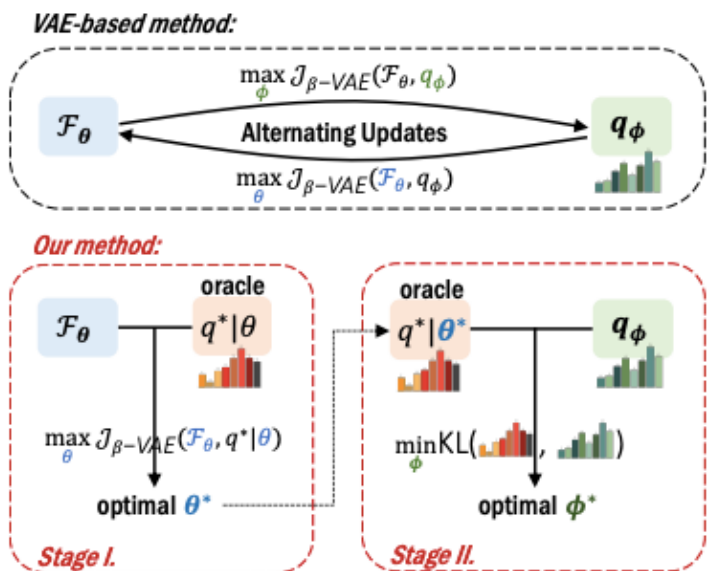


Figure 3. Two-stage training framework.

т.е. нужно найти оптимум

$$\max_{\theta} \frac{1}{|D|} \sum_{(x,y) \in D} \sum_{t=1}^T q_{\theta}^*(t|y, x) \log p_{\theta}(y|t, x)$$

2. Зафиксировав оптимальный  $\theta$  найдём такой критерий останова  $\pi_{\phi}$ , распределение которого бы было максимально похожим на распределение  $q_{\theta}^*(t)$ . Для этого используется обратная дивергенция Кульбака-Лейблера:

$$KL(q_{\phi} || q_{\theta}^*) = - \sum_{t=1}^T q_{\theta}^*(t|y, x) \log q_{\phi}(t) - H(q_{\theta}^*)$$

Приведенные в работе эксперименты показывают эффективность как модели, так и алгоритма обучения в широком диапазоне применений. Стоит отметить результат, полученный при распознавании изображений: применение данной архитектуры повысило точность предсказательной модели, однако, как было отмечено авторами статьи, имитировать поведение распределения времени очень сложно. Несмотря на это результаты данного исследования имеют направление для потенциального прорыва.