# Пример слайдов по работе курса

Эдуард Анатольевич Владимиров

Московский физико-технический институт

22 февраля 2022 г.

# Слайд об исследованиях

```
Исследуется проблема ....
```

Цель исследования предложить метод ....

#### Требуется предложить

- метод ...,
- 2) метод ...,
- 3) метод ....

#### Решение

Для ....

#### Постановка задачи ...

Заданы

3) ....

1) признаки ...,

2) целевая переменная ...,

. . .

Требуется выбрать модель ... из множества

$$\mathfrak{G} = \{ g | g : \mathbb{R}^n \to \mathbb{Y}' \}.$$

Оптимизационная задача . . . :

$$g = \arg\min_{g \in \mathfrak{G}} \mathcal{L}\big(\ldots\big),$$

где  $\mathcal{L}$  — функция ошибки.

<sup>&</sup>lt;sup>1</sup>Lopez-Paz D., Bottou L., Scholkopf B., Vapnik V. Unifying distillation and privileged information // ICLR, 2016.

<sup>&</sup>lt;sup>2</sup> Hinton G., Vinyals O., Dean J. Distilling the knowledge in a neural network // NIPS, 2015.

# Предложенный метод ...

Заданы

- 1) ...,
- 2) ....

Параметрические семейства:

$$\mathfrak{F} = \left\{ f | f = \mathsf{softmax} \big( v \big( x \big) / \mathcal{T} \big), \quad v : \mathbb{R}^n \to \mathbb{R}^K \right\},$$

$$\mathfrak{G} = \left\{ \mathsf{g} | \mathsf{g} = \mathsf{softmax} \left( \mathsf{z} (\mathsf{x}) / T \right), \quad \mathsf{z} : \mathbb{R}^n \to \mathbb{R}^K \right\},$$

где ....

Функция ошибки

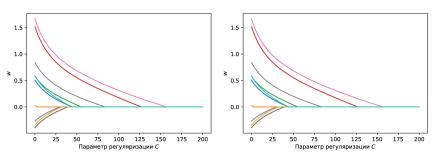
$$\mathcal{L}(\mathbf{g}) = -\sum_{i=1}^{m} \underbrace{\sum_{k=1}^{K} y_i^k \log \mathbf{g}(\mathbf{x}_i)\big|_{T=1}}_{\text{исходная функция потерь}} - \sum_{i=1}^{m} \underbrace{\sum_{k=1}^{K} \mathbf{f}(\mathbf{x}_i)\big|_{T=T_0} \log \mathbf{g}(\mathbf{x}_i)\big|_{T=T_0}}_{\text{слагаемое дистилляции}},$$

где . . . .

Оптимальная модель выбирается из класса,  $\hat{\mathbf{g}} = \arg\min_{\mathbf{g} \in \mathfrak{G}_{\mathbf{d}}} \mathcal{L}(\mathbf{g}).$ 

### Анализ предложенного метода . . .

На графике показана зависимость значения параметров $w_i$  в зависимости от параметра  $I_1$ -регуляризации C.



С увеличением параметра регуляризации C число ненулевых параметров  $w_i$  уменьшается.

# Выводы

1. Предложен . . . . 2. Доказаны теоремы ..., — . . . , <del>-</del> .... 3. Предложен метод . . . **—** ..., **—** .... 4. Предложены методы . . . **—** ..., — .... 5. Предложена вероятностная интерпретации . . . .