

Пример слайдов по работе курса

И. О. Фамилия

Московский физико-технический институт

10 февраля 2022 г.

Слайд об исследованиях

Исследуется проблема

Цель исследования —

предложить метод

Требуется предложить

1) метод,

2) метод,

3) метод

Решение

Для

Постановка задачи ...

Заданы

- 1) признаки ... ,
- 2) целевая переменная ... ,
- 3)

...

Требуется выбрать модель ... из множества

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

Оптимизационная задача ...:

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

где \mathcal{L} — функция ошибки.

¹Lopez-Paz D., Bottou L., Scholkopf B., Vapnik V. Unifying distillation and privileged information // ICLR, 2016.

²Hinton G., Vinyals O., Dean J. Distilling the knowledge in a neural network // NIPS, 2015.

Предложенный метод ...

Заданы

1) ...,

2)

Параметрические семейства:

$$\mathfrak{F} = \{ \mathbf{f} | \mathbf{f} = \text{softmax}(\mathbf{v}(\mathbf{x})/T), \quad \mathbf{v} : \mathbb{R}^n \rightarrow \mathbb{R}^K \},$$

$$\mathfrak{G} = \{ \mathbf{g} | \mathbf{g} = \text{softmax}(\mathbf{z}(\mathbf{x})/T), \quad \mathbf{z} : \mathbb{R}^n \rightarrow \mathbb{R}^K \},$$

где

Функция ошибки

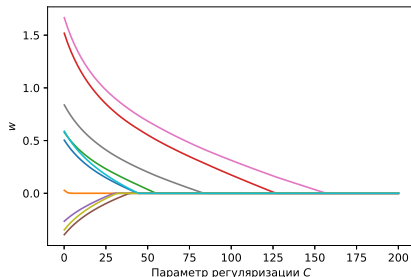
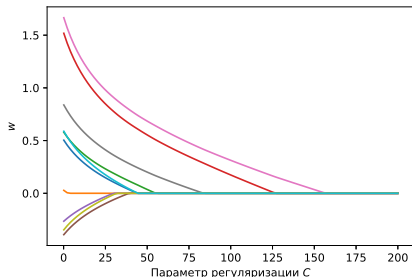
$$\mathcal{L}(\mathbf{g}) = - \underbrace{\sum_{i=1}^m \sum_{k=1}^K y_i^k \log \mathbf{g}(\mathbf{x}_i) \Big|_{T=1}}_{\text{исходная функция потерь}} - \underbrace{\sum_{i=1}^m \sum_{k=1}^K \mathbf{f}(\mathbf{x}_i) \Big|_{T=T_0} \log \mathbf{g}(\mathbf{x}_i) \Big|_{T=T_0}}_{\text{слагаемое дистилляции}},$$

где

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

Анализ предложенного метода ...

На графике показана зависимость значения параметров w_i в зависимости от параметра l_1 -регуляризации C .



С увеличением параметра регуляризации C число ненулевых параметров w_i уменьшается.

Выводы

1. Предложен
2. Доказаны теоремы ...,
— ...,
—
3. Предложен метод ...
— ...,
—
4. Предложены методы ...
— ...,
—
5. Предложена вероятностная интерпретации