

1. Общая схема постановки задачи

Важно: рассматриваем только задачу классификации для простоты.

- 1) Как описывается модель : https://github.com/bahleg/tex_slides/blob/master/loi18/slides.pdf (слайд 7).
- 2) Общая идея параметризации берется из DARTS.
- 3) Гиперпараметры — l_2 -регуляризация.
- 4) Задачу ставим как двухуровневую задачу оптимизации (параметры оптимизируем по обучающей выборке, структуру и гиперпараметры — по валидационной). Приблизительно списать можно, например, отсюда : <https://arxiv.org/pdf/1602.02355.pdf> (формула 1 на странице 3).

2. Общие обозначения

- Обучающая выборка: $\mathfrak{D}^{\text{train}} = \{\mathbf{x}_i, y_i\}, \quad i = 1, \dots, m^{\text{train}}.$
- Валидационная выборка: $\mathfrak{D}^{\text{valid}} = \{\mathbf{x}_i, y_i\}, \quad i = 1, \dots, m^{\text{valid}}.$
- Объекты: $[\mathbf{x}_1, \dots, \mathbf{x}_m] = \mathbf{X}, \quad \mathbf{x} \in \mathbb{R}^n.$
- Метки объектов: $[y_1, \dots, y_m] \in \mathbf{y}, \quad y \in \{1, \dots, Z\}.$
- Количество классов: $Z.$
- Модель: $\mathbf{f}(\mathbf{x}, \mathbf{W}).$
- Подмодель: $\mathbf{f}_v(\mathbf{x}, \mathbf{w}_v).$
- Параметры модели: $\mathbf{W}.$
- Ребра графа: $E.$
- Вершины графа: $V.$
- Базовые функции для ребра (i,j) : $\mathbf{g}^{i,j}, \quad |\mathbf{g}^{i,j}| = K^{i,j}.$
- Веса каждой базовой функции: $\gamma^{i,j}.$
- Структура модели: $\mathbf{\Gamma}.$
- Регуляризационное слагаемое: $\mathbf{A}.$
- Функция потерь на обучении: L
- Функция потерь на валидации: $Q.$