

ВЫБОР ОПТИМАЛЬНОЙ СТРУКТУРЫ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ

Потанин Марк Станиславович

Аспирант

Факультет ФИВТ МФТИ, Долгопрудный, Россия

E-mail: mark.potinin@phystech.edu

Научный руководитель — Стрижов Вадим Викторович

Исследуется проблема выбора оптимальной структуры модели. Моделью является суперпозиция обобщенных линейных моделей, элементами которой являются линейная регрессия, логистическая регрессия, метод главных компонент, автоэнкодер и нейросеть. Исследуется зависимость ошибки от вида суперпозиции автокодировщиков [1] и многослойной нейронной сети. Функция ошибки состоит из двух слагаемых: ошибки восстановления элементов выборки после кодирования и ошибки восстановления зависимых переменных. Слагаемые используют одни и те же независимые переменные, но разные зависимые переменные. Для автокодировщика зависимая переменная — это сами признаки объекта, для нейронной сети, следующей за ним, зависимая переменная — ответ y на объекте. Точка разделения — это место в суперпозиции, где автокодировщик, имеющий оптимальные параметры, передает преобразованный вектор признаков в нейросеть. Необходимо найти оптимальное расположение разделения автокодировщика и сети, которое минимизирует ошибку аппроксимации выборки. Под структурой такой модели понимаются величины, задающие вид итоговой суперпозиции — то есть число слоев автокодировщика и нейросети, а также число нейронов в слоях. Процедура минимизации ошибки аппроксимации выборки следующая: сначала максимизируется точность реконструкции кодировщиков, затем оптимизируются параметры нейросети. Исследуются зависимость точности, сложности и устойчивости модели от способа задания структуры.

Вместе с точностью оптимизируется сложность модели. Под сложностью понимается *структурная сложность модели* — это число параметров модели, с учетом их области определения. Альтернативой этому определению является *статистическая сложность модели* — минимальная длина описания, т.е. минимальное количество информации, которое требуется для передачи информации о модели и о выборке [2].

Предложен алгоритм выбора оптимальной структуры нейронной

сети. Задается множество случайных начальных значений структурных параметров. Затем вычисляется значение функции ошибки аппроксимации, которое характеризует качество модели в наборе. Согласно этой функции выбираются модели, которые обмениваются структурными параметрами, образуя новую структуру. Многократное повторение этой операции позволяет получить оптимальную структуру модели. Проведен вычислительный эксперимент с использованием реальных и синтетических данных. В результате эксперимента существенно понижена структурная сложность моделей с сохранением точности аппроксимации.

Таблица 1: Результат применения генетического алгоритма для прореживания сети

Выборка \mathcal{D}	Ошибка сети с прореживанием	Ошибка сети без прореживания	Сложность без прореживания	Сложность после прореживания
Credit Card	0.3204 ± 0.0032	0.2681 ± 0.0034	68	25
Protein	4.4968 ± 0.0238	4.4968 ± 0.0238	16	1
Airbnb	35.0773 ± 0.5909	33.9163 ± 0.5978	32	12
Wine quality	0.5818 ± 0.0147	0.5941 ± 0.0149	20	4
Synthetic, 10^{-3}	0.3005 ± 0.0081	0.303 ± 0.0079	60	12

Литература

1. *Hinton G. E., Salakhutdinov R. R.* Reducing the dimensionality of data with neural networks // Science, 2006. Vol. 313. No. 5786. P. 504–507.
2. *Бахтеев О. Ю., Стрижов В. В.* Выбор моделей глубокого обучения субоптимальной сложности // Автомат. и телемех., 2018. Вып. 8. С. 129–147.