

**ОБУЧЕНИЕ МОДЕЛИ МАШИНЫ РЕЛЕВАНТНЫХ
ВЕКТОРОВ ПО БОЛЬШИМ ОБЪЕМАМ ДАННЫХ**

Молчанов Дмитрий Александрович

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: dmo1ch111@gmail.com

В данной работе рассматривается машина релевантных векторов (Relevance Vector Machine, RVM) — модель для решения задачи регрессии с возможностью автоматического отбора признаков.

Зададим вероятностную модель машины релевантных векторов. Пусть $x \in \mathbb{R}^d$ — объект, $t \in \mathbb{R}$ — его ответ, $(X, T) = (x_i, t_i)_{i=1}^N$ — обучающая выборка. Рассматривается задача линейной регрессии с большим числом объектов N и небольшим числом признаков d . Модель RVM:

$$p(t_i | x_i, w, \beta) = \mathcal{N}(t | w^T x_i, \beta^{-1}),$$
$$p(w | A) = \mathcal{N}(w | \vec{0}, A^{-1}), \text{ где } A = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_d)$$

Настройка гиперпараметров данной модели проводится путем максимизации так называемой *обоснованности*:

$$E(A, \beta) = \int p(T | X, w, \beta) p(w | A) dw \rightarrow \max_{A, \beta}$$

Максимизация обоснованности модели RVM позволяет произвести отбор признаков. При настройке гиперпараметров α_i часть из них уходит в плюс бесконечность. Это означает, что априорное распределение для соответствующей компоненты вектора весов вырождается в дельта-функцию, эта компонента вектора весов становится обязана равняться нулю, а это означает, что соответствующий признак исключается из модели.

Существует и давно известна эффективная процедура обучения данной модели, основанная на методе простой итерации. Тем не менее, с ростом числа объектов и признаков эта процедура становится крайне неэффективной.

В данной работе был получен новый метод для обучения этой модели, основанный на стохастической оптимизации обоснованности. Экспериментально показано, что предложенный метод позволяет осуществлять обучение этой модели на больших выборках (миллионы объектов и более).

В дальнейшие планы исследований по данной задаче входит обоб-

щение предложенного метода на случай большого числа признаков. Также планируется обобщение метода на случай, когда обоснованность модели и ее градиенты не могут быть вычислены аналитически. Работа в этом направлении позволит, например, создавать эффективные методы оптимизации структуры нейронных сетей с большим числом параметров.

Литература

1. Tipping M. E. Sparse Bayesian learning and the relevance vector machine // Journal of Machine Learning Research, 2001, No 1. P. 211–244.