

Ансамблевый классификатор на основе рекомендательной системы

Ю. Кашницкий, А. Лоптев, Д.И. Игнатов

1 Постановка задачи

- Есть входной контекст вида «объекты-признаки», где один из признаков — целевой (то, что будем предсказывать), и набор классификаторов cl_1, \dots, cl_N ;
- Построен многозначный классификационный контекст «объекты-классификаторы» с оценками качества классификации (как вариант, «зазор» классификации = вероятность правильной классификации минус максимальная вероятность неправильной классификации) в процессе кросс-валидации;
- Есть тестовые объекты, для которых такие оценки получить не можем, так как, естественно, не знаем целевого класса тестовых объектов;
- Задача — применить рекомендательную систему и каждому тестовому объекту «рекомендовать» свой классификатор. Таким образом составить ансамблевый классификатор RMCS (Recommender-based Multiple Classifier System).

2 Игрушечный пример

Контекст данных (объекты – признаки) и классификационный контекст (объекты – классификаторы):

	m_1	m_2
g_1	1	2
g_2	2	3
g_3	3	1

	cl_1	cl_2	cl_3
g_1	0.5	0.2	0.6
g_2	0.3	0.8	0.4
g_3	?	?	?

Обозначения ¹:

- Θ — признаки объектов;
- X — признаки классификаторов (будут найдены в процессе обучения рекомендательной системы);
- Y — матрица оценок качества классификаций на объектах. Это матрица классификационного контекста, только для тестовых объектов, где оценок нет, заведены переменные, например, $y_1^{(3)}, y_2^{(3)}, y_3^{(3)}$;
- μ — вектор средних значений по строкам матрицы Y (нужен для нормализации оценок);
- J — функция потерь, которую надо минимизировать;
- λ — коэффициент регуляризации.

$$\Theta = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, X = \begin{bmatrix} 1 & 1 & 1 \\ x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \end{bmatrix}, Y = \begin{bmatrix} 0.5 & 0.3 & y_1^{(3)} \\ 0.2 & 0.8 & y_2^{(3)} \\ 0.6 & 0.4 & y_3^{(3)} \end{bmatrix}$$

Этапы:

1. Нормализация: $Y \rightarrow Y - \mu = Y - \begin{bmatrix} 0.4 \\ 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.1 & -0.1 & (y_1^{(3)} - 0.4) \\ -0.3 & 0.5 & (y_2^{(3)} - 0.5) \\ 0.1 & -0.1 & (y_3^{(3)} - 0.5) \end{bmatrix}$;
2. Инициализация $x^{(1)}$ и $y^{(3)}$ малыми случайными величинами;
3. Минимизация $J(x^{(1)}, y^{(3)}) = \frac{1}{2} * \sum_{i=1}^3 \sum_{j=1}^3 (\Theta^{(j)T} X^{(i)} - y_j^{(i)})^2 + \frac{\lambda}{2} * \sum_{i=1}^3 \sum_{k=1}^3 (x_k^{(i)})^2 + \frac{\lambda}{2} * \sum_{i=1}^3 (y_i^{(3)})^2$.

¹Согласно видео-лекциям профессора Andrew NG про рекомендательные системы http://www.youtube.com/watch?v=saXRzxcgFN0o&list=PL_nY1DYXHPT-3dorG7Em6d18P4JRFDvH