

УДК 5171.:519.6

КОМП'ЮТЕРНА ПРОГРАМА ДЛЯ ДЕМОНСТРАЦІЇ СПОСОБІВ ВІДБОРУ ОЗНАК МЕТОДАМИ ГЛИБОКОГО НАВЧАННЯ В РЕГРЕСІЙНИХ МОДЕЛЯХ

О.О. Шевченко, студент

Київський національний університет технологій та дизайну

С.М. Краснитський, доктор фіз. - мат. наук, професор

Київський національний університет технологій та дизайну

Ключові слова: комп'ютерна програма, методи глибокого навчання, регресійна модель, Python.

При розробці регресійних моделей часто постає питання щодо можливого скорочення кількості пояснювальних змінних. Існує досить багато методів виконання зазначеної дії. В даній доповіді представлено описання функції комп'ютерної програми, що розроблена мовою Python для демонстрації в учбовому процесі дії та порівняння кількох методів зазначеного типу.

Одним із застосованих із вказаною метою методів є процедура видалення регресорів, для яких стовпці матриці експерименту є лінійно залежними від інших стовпців, що відповідають іншим регресорам. Дана процедура базується на відомому алгоритмі Грама – Шміда ортогоналізації системи векторів евклідового простору. Якщо між стовпцями існує лінійна залежність, то рано чи пізно буде отримано стовпець, повністю складений з нулів, і тоді відповідна змінна виключається з моделі. У матричному вигляді процедура може бути представленою у вигляді

$$Z_{iT} = Z_i - Z(Z'Z)^{-1}Z'Z_i, i = 1, \dots, p,$$

де p – кількість стовпців початкової регресійної матриці, Z – матриця вже перетворених стовпців, Z_i – наступний стовпець регресійної матриці, що перетворюється, Z_{iT} – вектор-результат перетворення, ортогональний стовпцям матриці Z . В сучасних процедурах глибокого навчання часто застосовуються інші методи, в основу яких покладено інші міркування [1]. Назвемо тут два з цих методів: 1) метод Lasso, 2) метод Ridge. Обидва з цих методів засновані на додаванні до функції, що мінімізується за звичайним методом найменших квадратів, штрафної функції, що містить коефіцієнти функції регресії. Мінімізація такої функції втрат зумовлює тенденцію або до зменшення оцінок коефіцієнтів функції регресії, або взагалі до обнуління деяких з них.

Функції втрат до методів Lasso і Ridge мають вигляд, відповідно:

$$L_1(\beta) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p |\beta_j|, L_2(\beta) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p (\beta_j)^2,$$

де y_i – спостережуване значення функції регресії, \hat{y}_i – оціночне значення, β_j – коефіцієнт моделі для j -ї ознаки, λ так званий гіперпараметр.

Список використаних джерел

1. Introduction to Machine Learning with Python, by Andreas C. Müller and Sarah Guido (2017), O'Reilly Media, Inc., 100.