УДК 004.85:519.24

Обнаружение дрейфа данных в промышленных конвейерах машинного обучения

Карпов Р.А., Мартынова М.О.

Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ

г. Казань, Российская Федерация

Модель, выведенная в промышленную эксплуатацию, обучена на конечной выборке и наследует свойства породившей её совокупности. Стоит характеристикам входных данных сместиться — и обобщения модели теряют состоятельность, хотя ни код, ни веса никто не менял. Д.Е. Намиот и Е.А. Ильюшин называют такое изменение правилом эксплуатации, а не исключением из него [1].

Разграничивают три ситуации: ковариантный сдвиг меняет распределение признаков, сдвиг априорных вероятностей — распределение меток, дрейф концепции — саму зависимость между ними [2]. Различение не академическое: первые два случая распознаются по входу, третий без разметки не наблюдаем.

Отсюда два класса детекторов. Первый сравнивает эксплуатационное распределение признаков с эталонным и разметки не требует: сигнал приходит рано, но сообщает лишь об изменении входа. Второй опирается на качество модели — DDM следит за частотой ошибок и поднимает тревогу при выходе за границу, построенную по биномиальной аппроксимации [3], ADWIN разрезает окно переменной длины, едва его половины разойдутся значимо [4]. Точность выше, но нужна разметка, а в скоринге она приходит через недели.

Расхождение эмпирических функций распределения обучающей и эксплуатационной выборок для непрерывного признака измеряет статистика Колмогорова — Смирнова [5]:

$$D_{n,m} = \sup_x |F_n(x) - G_m(x)| \quad (1)$$

Гистограммное сравнение закрепилось за индексом стабильности популяции, где выборки разбиты на k интервалов, а p и q — доли наблюдений i-го интервала на обучении и в эксплуатации:

$$\mathrm{PSI} = \sum_{i=1}^{k} (p_i - q_i) \ln\left(\frac{p_i}{q_i}\right) \quad (2)$$

Привычные пороги (до 0,1 — сдвига нет, от 0,1 до 0,25 — умеренный, свыше 0,25 — нужно вмешательство) выросли из практики, а не из контроля ошибок первого рода: Б. Юрдакул показал асимптотическую связь индекса с распределением хи-квадрат и предложил пороги, зависящие от объёма выборки [6].

Покомпонентная проверка упирается в множественность гипотез: при пятидесяти признаках и уровне значимости 0,05 вероятность хотя бы одной ложной тревоги превышает 0,9, и дежурная смена приучается её игнорировать; поправка Бонферрони срезает вместе с ложными тревогами и чувствительность к слабому, зато реальному сдвигу. С. Рабансер с соавторами показали, что понижение размерности с проверкой в сжатом пространстве нередко надёжнее перебора признаков [7].

Мы предлагаем развести мониторинг на два уровня. Нижний считается на каждом суточном срезе и собран из дешёвых критериев — Колмогорова — Смирнова для непрерывных признаков, хи-квадрат для категориальных, индекс стабильности как сводка по витрине; его дело не вынести вердикт, а сузить область внимания. Верхний включается только после срабатывания нижнего и проверяет целевую метрику на той доле выборки, где разметка уже пришла. Переобучение запускается при согласии обоих уровней. Сработал один нижний — это смещение входа без последствий для качества, и таков самый частый исход. Упало качество при стабильном входе — налицо дрейф концепции, и нужен пересмотр признакового описания, а не дообучение [3].

Цена схемы очевидна: появления разметки она не ускоряет и там, где обратная связь запаздывает сильнее допустимого срока реакции, не спасает. Зато чувствительность нижнего уровня можно держать высокой, не превращая дежурство в поток ложных тревог. Дальше схему предстоит оценить на открытых потоковых наборах данных с размеченным моментом дрейфа.

Список использованной литературы

1. Намиот Д.Е. *Мониторинг сдвига данных в моделях машинного обучения* / Д.Е. Намиот, Е.А. Ильюшин // International Journal of Open Information Technologies. – 2022. – Т. 10, № 12. – С. 84–93.

2. Gama J. *A survey on concept drift adaptation* / J. Gama [et al.] // ACM Computing Surveys. – 2014. – Vol. 46, № 4. – P. 1–37. – DOI 10.1145/2523813.

3. Gama J. *Learning with drift detection* / J. Gama [et al.] // Advances in Artificial Intelligence – SBIA 2004. Lecture Notes in Computer Science. – 2004. – Vol. 3171. – P. 286–295. – DOI 10.1007/978-3-540-28645-5_29.

4. Bifet A. *Learning from time-changing data with adaptive windowing* / A. Bifet, R. Gavaldà // Proc. of the 2007 SIAM Int. Conf. on Data Mining. – 2007. – P. 443–448. – DOI 10.1137/1.9781611972771.42.

5. Massey F.J. *The Kolmogorov-Smirnov test for goodness of fit* / F.J. Massey // Journal of the American Statistical Association. – 1951. – Vol. 46, № 253. – P. 68–78. – DOI 10.1080/01621459.1951.10500769.

6. Yurdakul B. *Statistical properties of population stability index*: doctoral dissertation / B. Yurdakul. – Kalamazoo: Western Michigan University, 2018. – URL: https://scholarworks.wmich.edu/dissertations/3208 (дата обращения: 29.09.2026).

7. Rabanser S. *Failing loudly: an empirical study of methods for detecting dataset shift* / S. Rabanser, S. Günnemann, Z.C. Lipton // Advances in Neural Information Processing Systems. – 2019. – Vol. 32. – P. 1396–1408.