

Прудкий Александр Сергеевич

канд. пед. наук, и.о. заведующего кафедрой, доцент

Жевлакова Кристина Евгеньевна

студентка

Кукарцев Владислав Викторович

канд. техн. наук, доцент

ФГБОУ ВО «Российский государственный

аграрный университет – МСХА им. К.А. Тимирязева»

г. Москва

АДАПТИВНОЕ МАШИННОЕ ОБУЧЕНИЕ В РЕАЛЬНОМ ВРЕМЕНИ: МЕТОДЫ БОРЬБЫ С ОТКЛОНЕНИЕМ КОНЦЕПТОВ В ПОТОКОВЫХ ДАННЫХ

Аннотация: статья исследует современные подходы к адаптивному машинному обучению в условиях реального времени, акцентируя внимание на проблеме отклонения концептов в потоковых данных. Рассматриваются основные методы обнаружения и коррекции отклонения данных, позволяющие моделям сохранять точность при работе с нестационарными информационными потоками. Анализируются алгоритмы онлайн-обучения, адаптивные пороговые механизмы и техники инкрементального обновления моделей. Обсуждаются перспективы внедрения этих методов в системы мониторинга, прогнозирования и автоматизации, а также вызовы, связанные с вычислительной сложностью и задержками обработки.

Ключевые слова: адаптивное машинное обучение, потоковые данные, отклонение концептов.

Традиционный подход к машинному обучению предполагает статичность данных. Мы собираем набор данных, обучаем модель и разворачиваем ее в продакшн. Однако реальный мир динамичен. Потребительские предпочтения меняются, рыночные условия флуктуируют, поведение злоумышленников эволюционирует. В таких условиях статическая модель неизбежно деградирует. Ее

прогнозы становятся менее релевантными, ошибки растут. Это явление называется концептуальным дрейфом или отклонением концептов.

С ростом объемов потоковых данных, будь то датчики интернета вещей или финансовые транзакции, потребность в моделях, способных адаптироваться на лету, стала критической. Современные системы аналитики больше не могут позволить себе роскошь переобучения раз в месяц. Им необходима способность подстраиваться под изменения немедленно, сохраняя точность предсказаний.

Важно различать источник изменений. Дрейф данных происходит, когда меняется распределение входных признаков, но связь между признаками и целевой переменной остается прежней. Истинный дрейф означает, что меняется сама зависимость между признаками и результатом, как при изменении экономических условий, влияющих на кредитный скоринг.

Первый этап адаптации — обнаружение дрейфа. Мониторинг распределения данных сравнивает статистические характеристики прошлых и текущих данных с помощью тестов вроде Колмогорова-Смирнова. Преимущество в том, что не требуется целевая переменная, но метод видит только дрейф данных. Мониторинг ошибок модели отслеживает точность прогнозов. Алгоритмы CUSUM и Page Hinkley фиксируют выход ошибки за контрольные границы. ADWIN использует адаптивное скользящее окно для обнаружения значимых изменений в средней ошибке.

Когда дрейф обнаружен, система должна адаптироваться. Периодическое переобучение запускает процесс обучения заново каждую ночь или неделю. Это надежно, но не спасает при внезапных изменениях. Инкрементальное обучение позволяет модели обновляться по мере поступления каждого нового образца, плавно следуя за изменениями без хранения всей истории. Ансамблевые методы держат пул моделей, обученных на разных отрезках времени, и динамически взвешивают их голоса, отдавая предпочтение тем, что лучше работают на свежих данных.

Внедрение этих методов требует соответствующей архитектуры. Современный конвейер строится вокруг Kafka для надежного поступления событий и движков потоковой обработки вроде Apache Flink или Spark Streaming. Ключевой компонент петля обратной связи, когда обнаружение дрейфа автоматически запускает обновление модели.

Практическое применение охватывает множество индустрий. В промышленности адаптивные модели предсказывают отказы оборудования при изменении режимов работы. В информационной безопасности они необходимы для обнаружения атак, так как паттерны трафика постоянно меняются. В финансах онлайн скоринг и выявление мошенничества требуют немедленной реакции на новые схемы злоумышленников.

Существуют и вызовы. Проблема ложных срабатываний требует тонкой настройки детекторов. Вычислительная сложность онлайн алгоритмов создает нагрузку на инфраструктуру. Получение размеченных данных в реальном времени часто требует использования проксиметрик.

В современном мире, где данные текут непрерывно, адаптивность перестает быть роскошью и становится необходимым условием для поддержания качества MLem. Модели, не умеющие учиться на лету, обречены на устаревание. Дальнейшее развитие связано с интеграцией адаптивных механизмов в AutoML и развитием федеративного обучения, позволяющего адаптировать модели без централизации чувствительных данных.

Список литературы

1. Информационные системы и технологии в АПК: учебник / А.В. Бабкина, И.Е. Быстренина, М.И. Горбачев [и др.]. – М.: Российский государственный аграрный университет, 2025. – 615 с. EDN ZJCHJU

2. Обнаружение аномалий в данных промышленных сенсорных систем добычи на основе алгоритмов кластеризации / С.О. Курашкин, Л.В. Красовская, М.В. Журавлев, Т.В. Бирюкова // Известия Тульского государственного университета. Науки о Земле. – 2025. – №4. – С. 382–388. EDN WMCTCH

3. Проектирование модуля информационной системы как фактор повышения качества сервисных услуг предприятия / Л.В. Красовская, С.В. Кукарцева, М.В. Журавлев, А.В. Низамеева // Право, экономика и управление: теория и практика: материалы III Всероссийской научно-практической конференции с международным участием. – Чебоксары: Среда, 2025. – С. 191–194. EDN CHXXVN

4. Функциональное моделирование бизнес-процессов как основа проектирования информационных систем предприятия / М.В. Журавлев, В.Ю. Панченко, С.В. Кукарцева, А.Р. Глинская // Социально-экономические процессы современного общества: материалы Всероссийской научно-практической конференции с международным участием. – Чебоксары: Среда, 2026. – С. 163–167. EDN LVNZDY

5. Экономическое влияние цифровой зрелости и ИИ на производительность: прогнозирование, факторный анализ и выводы / Р.А. Нажмутдинов, Д.Р. Ишмухаметова, А.С. Прудкий, С.В. Кукарцева // Современные тренды управления, экономики и предпринимательства: от теории к практике: сборник материалов III Всероссийской научно-практической конференции с международным участием (Ульяновск, 11–13 ноября 2025 года). – Чебоксары: Среда, 2025. – С. 160–164. EDN IZFNCS