ДЛЯ ЧЕГО ТРАНСПОРТНОЙ ОТРАСЛИ «ОБЪЯСНИМЫЙ» ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

В Институте проблем транспорта им. Н.С. Соломенко (ИПТ) РАН разработан итеративный алгоритм интерпретируемой классификации для анализа транспортных рисков.

Когда речь идёт о диагностике технического состояния подвижного состава, оценке аварийности, распознавании транспортных средств по видеопотоку или классификации рисков на объектах транспортной инфраструктуры, решение автоматической системы должно быть не только точным, но и объяснимым. Инженер, эксперт-расследователь ДТП, диспетчер или регулирующий орган обязаны понимать, на основании чего сделан тот или иной вывод.

Современные модели машинного обучения часто демонстрируют высокую точность, но работают как «чёрный ящик»: они не дают ответа на вопрос, почему принято именно такое решение. Для транспортной отрасли это серьёзное ограничение. Автоматизированные системы всё активнее используются для мониторинга технического состояния инфраструктуры, анализа причин инцидентов и оценки рисков на маршрутах и объектах. Цена ошибки в подобных задачах может быть слишком высока.

Вячеслав Дюк, и. о. заведующего лабораторией интеллектуальных транспортных систем, главный научный сотрудник ИПТ РАН, предлагает подход, в котором объяснимость закладывается в саму архитектуру алгоритма, а не добавляется постфактум внешними средствами интерпретации, которые лишь приближённо описывают уже принятое моделью решение.

В основе предложенного подхода — идея, интуитивно близкая работе опытного инженера или эксперта: новый случай оценивается не по абстрактной формуле, а через сопоставление с ограниченным набором показательных прецедентов — «якорных» объектов, представляющих типичные и характерные примеры каждого класса. Для каждого такого прецедента строится собственная, локальная мера сходства (метрика), в которой значимость признаков определяется заново, применительно именно к этой области пространства данных.

Для автоматического построения таких локальных метрик Вячеслав Дюк разработал детерминированный итеративный алгоритм ILMR (Iterative Local Metric Regression). Развивая линию исследований, которую коллектив лаборатории ведёт применительно к задачам распознавания транспортных средств методами технического зрения, ILMR добавляет к этим наработкам ключевое качество — объяснимость финального решения, что особенно значимо при сертификации автоматизированных систем мониторинга, расследовании инцидентов и обосновании выводов перед контролирующими органами.

Алгоритм имеет линейную вычислительную сложность и не требует специализированного оборудования (графических ускорителей), что делает его экономичным с точки зрения энергопотребления и пригодным для развёртывания на периферийных устройствах — бортовых системах, дорожных контроллерах, диспетчерских комплексах. Кроме того, метод совместим с современными нейросетевыми представлениями данных, что открывает путь к его применению поверх систем компьютерного зрения, уже используемых для распознавания транспортных средств и объектов инфраструктуры.