

УДК: 004.8

*Эль-Курди Ю.Т., бакалавр 3 г.о, Набережночелнинский институт ФГАОУ ВО
«Казанский (Приволжский) федеральный университет», email: yotna@mail.ru;*

РАЗРАБОТКА ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ О ДЕЯТЕЛЬНОСТИ ТРАНСПОРТНОГО ПЕРЕВОЗЧИКА

Аннотация: Транспортные перевозки играют важную роль в обеспечении экономического развития и социального благополучия общества. Однако, с ростом объемов перевозок возникают новые проблемы, связанные с обработкой и анализом данных о деятельности транспортных перевозчиков. Это обуславливает актуальность разработки новых технологий анализа данных в сфере транспорта.

Растущие объемы перевозок порождают большой объем данных, которые требуют систематизации и анализа для принятия взвешенных решений. Разработка интеллектуальной технологии анализа данных позволит автоматизировать процессы сбора, обработки и интерпретации информации, что приведет к повышению эффективности работы транспортных компаний, сокращению затрат и улучшению качества услуг. Такое исследование имеет большое практическое значение и может стать основой для разработки инновационных решений в области транспортной логистики.

Ключевые слова: транспортная компания, обработка данных, технологии, анализ, перевозки.

Транспортная отрасль играет важную роль в мировой экономике и обеспечении потребностей населения. Однако накопление данных не является достаточным условием для эффективного управления и принятия решений. Для того, чтобы выявить скрытые закономерности, тренды и паттерны, которые могут быть полезными для оптимизации бизнес-процессов и улучшения качества услуг, необходимо выполнять анализ данных. Ввиду большого объема циркулируемых данных, такой анализ невозможно выполнить вручную, поэтому для выявления скрытых зависимостей, в том числе при оперативном управлении, используются так называемые методы интеллектуального анализа данных или data mining [1].

На первом шаге аналитик производит подготовку исходных данных, повышая их качество до необходимого уровня. Затем, используя методы

интеллектуального анализа, выявляют закономерности исследуемого массива данных, которые обобщают и сохраняют в виде модели интеллектуального анализа данных [2].

Интеллектуальный анализ данных в транспортной сфере активно применяется различными компаниями для повышения эффективности и оптимизации своих услуг.

В такси-сервисах (Uber, Яндекс.Такси) с помощью методов интеллектуального анализа данных и искусственного интеллекта строятся оптимальные маршруты (на основе входных данных об интенсивности движения, погодных условиях, потребительском спросе, точек локаций автомобиля и клиентов), устанавливаются динамические цены, прогнозируется время прибытия [3].

Логистические компании (DHL, FedEx) оптимизируют маршруты доставки грузов, минимизируя временные и стоимостные затраты, на основе информации о грузах, маршрутах, условиях доставки и подвижном составе [3].

Производители электромобилей (Tesla, Audi) оптимизируют маршруты транспортных средств и прогнозируют время зарядки аккумуляторов на основе информации о расстоянии до зарядных станций, скорости движения, времени суток [3].

Наконец, общественный транспорт оптимизируется на основе анализа данных о пробках, пассажиропотоке и графике движения [3].

При интеллектуальном анализе данных используются алгоритмы машинного обучения. В основе машинного обучения лежит принцип повторяющегося решения поставленной задачи при предъявлении очередного набора признаков из большого массива данных [4]. Модель, построенная с использованием алгоритма машинного обучения на основе собранных данных, впоследствии пригодна для решения задачи для других наборов данных.

Выделяют несколько классов алгоритмов машинного обучения:

- обучение с учителем - модель строится на размеченных данных, для которых известен «ответ», т.е. результирующее поле. Это равносильно предъявлению ученику образца изображения с названием того, что на нем изображено. Построенная модель находит и запоминает закономерности, которые используются при предъявлении нового экземпляра данных. Примером такого случая является предсказание времени в пути до работы исходя из времени суток и погоды [5].

- обучение без учителя – модель строится на основе неконтролируемого обучения на неразмеченных данных. При этом модель на основе алгоритма машинного обучения самостоятельно выделяет «верные» ответы. Примером работы такого типа алгоритмов является выделение типовых видов трафика на основе информации о скорости, местоположении транспорта, плотности трафика, времени суток. Подобная идентификация позволяет оптимизировать маршруты с учетом предсказанных заторов, улучшив управление светофорами и перераспределив транспортные потоки [5].

- обучение с подкреплением — алгоритм обучения на основе «проб и ошибок». В качестве аналогии можно привести процесс обучения ребенка езде на велосипеде. Так и искусственный интеллект за каждую удачу (поездка прямо, осуществленный поворот) получает награду или подкрепление, а за допущенные ошибки (падение) получает штраф в виде отрицательного подкрепления. Обучение с подкреплением используется при функционировании беспилотных автомобилей. Сначала автомобиль обучается в виртуальной среде, где он может "пробовать" разные сценарии на дороге без риска реальных аварий. После виртуального обучения, автомобиль тестируется в реальных условиях на специально оборудованных площадках, а затем - на дорогах общего пользования. Автомобиль учится реагировать на изменения дорожной обстановки (например, другие машины, пешеходы, дорожные знаки), принимая безопасные решения о движении. Именно благодаря обучению с подкреплением автономные транспортные средства могут реагировать на изменения в окружающей среде, соблюдая безопасность и в то же время

достигая поставленной задачи в виде перемещения до заданной точки. Данный тип обучения также рекомендуется применять при оптимизации систем управления трафиком, автоматическом управлении грузоперевозками [5].

Также можно выделить алгоритмы глубокого обучения на основе компьютерного зрения — это форма машинного обучения, которая позволяет компьютерам идентифицировать объекты на изображениях и видео без вмешательства человека. Оно имеет множество применений в мире логистики, например, отслеживание уровня запасов или распознавание поврежденных посылок. Кроме того, алгоритмы распознавания изображений могут использоваться в навигационных системах робототехники, помогая роботам транспортировать товары на складах. Используя эту технологию, компании могут сократить свои эксплуатационные расходы, оптимизировать ресурсы и повысить удовлетворенность клиентов.

В настоящей работе хотелось бы показать, как можно использовать технологии обработки данных, частично упомянутые ранее, в сфере транспортных перевозок. В рамках курсового проектирования были сгенерированы синтетические типовые данные, поступающие и накапливаемые внутри транспортных компаний. Целью исследования была разработка и апробация технологии обработки данных для улучшения работы транспортных компаний.

Для первоначального анализа данных эффективно использование группировки и агрегирования по различным полям, при этом визуально для аналитика удобно отображение полученного результата в виде сводной таблицы или сводной диаграммы.

В качестве одной из гипотез, проверка которой может быть полезна для транспортной компании, можно сформулировать следующую: имеется ли зависимость между классом водителя и количеством перевезенного груза с учетом его характеристик. Возможный результат анализа представлен на рис. 1. Здесь подразумевается, что чем выше класс водителя, тем выше его профессиональное мастерство и более совершенны водительские навыки. Если

получены подобного рода результаты (рис. 1), это может свидетельствовать о том, что опытным водителям требуется меньше времени на доставку груза, следовательно, они выполняют большее количество заказов, крупных по объему в том числе.

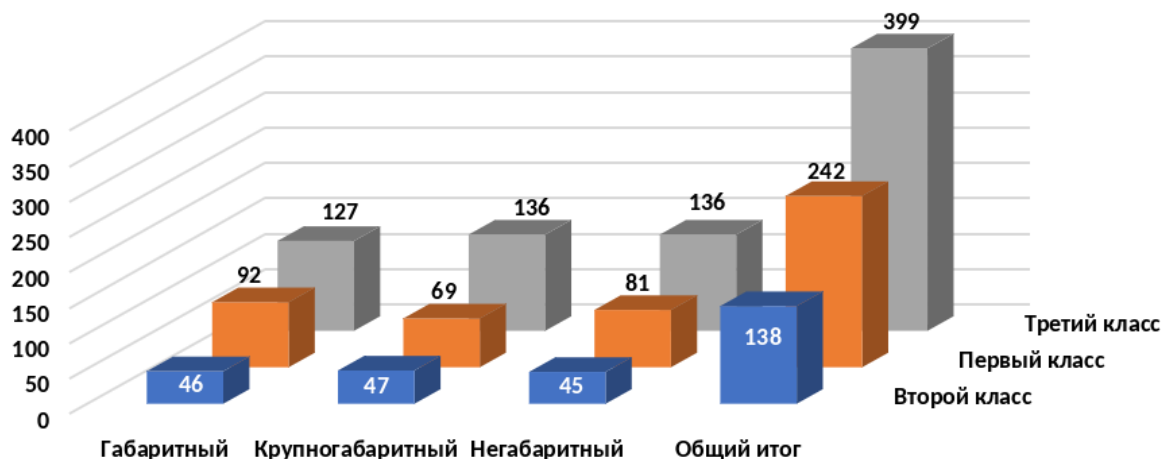


Рис. 1. Сводная диаграмма «Количество выполненных заказов в зависимости от класса водителя»

Также целесообразно проверить гипотезу: влияет ли класс водителя на время доставки груза. Если получен результат, аналогичный рис. 2, это может свидетельствовать о том, что чем выше класс водителя, в том числе его стаж, тем быстрее доставляется заказ. Водители с большим стажем более маневренны и лучше знают топографию города.



Рис. 2. Диаграмма «Длительность доставки заказа водителями в зависимости от класса, мин.»

Третья полезная гипотеза: влияет ли время года на расход топлива. Если получен результат, аналогичный рис. 3, это показывает, что наибольший расход соответствует перевозкам в позднее осеннее и зимнее время года, что обусловлено необходимостью обогрева кабины водителя, а также предварительного разогрева двигателя. На основе анализа данных о расходе топлива и условиях движения можно разработать оптимальные стратегии экономии топлива и снижения выбросов вредных веществ.



Рис. 3. Диаграмма «Расхода топлива в зависимости от сезона, л»

Более сложные технологии анализа данных могут быть основаны на алгоритмах машинного обучения. Нейронные сети - вычислительные модели, имитирующие биологические нейронные сети мозга человека. Они состоят из слоев нейронов, принимающих и обрабатывающих данные. Нейросеть обучается на обучающей выборке для нахождения зависимостей между входными и выходными данными. После обучения она способна предсказывать значения на основе новых входных данных, даже если они не использовались в обучении.

Для транспортной компании целесообразно прогнозировать ожидаемую длительность доставки, задавая в качестве входных влияющих факторов информацию о точках отправления и прибытия, классе водителя, типе груза, месяце, дне недели и времени суток доставки. Данная задача относится к задаче

построения регрессии. Поскольку зависимость результирующего признака от предикторов может быть нелинейной, то использование классической линейной регрессии может не давать нужного качества прогноза. Именно поэтому рекомендуется использовать нейронные сети, которые являются нелинейными аппроксиматорами. В качестве архитектуры такой нейронной сети можно использовать полносвязную сеть с количеством входных нейронов, равным числу влияющих факторов, и одним выходным нейроном, соответствующим результирующему признаку – длительности доставки. Для того, чтобы сеть могла учесть нелинейные зависимости, необходимы один или несколько скрытых слоев.

Таким образом, показано, что структура накапливаемых данных, позволяющая построить представленные сводные диаграммы, дает возможность установить или опровергнуть наличие зависимости между скоростью выполнения доставки грузов и, соответственно, объемом осуществляемых перевозок и классом водителя, а также оценить влияние времени года на расход топлива. Представленная структура нейронной сети предназначена для прогнозирования примерного времени доставки груза, с учетом всех факторов, влияющих на него. Такая модель значительно упростит планирование доставок в сжатые сроки, поможет оптимизировать маршруты и улучшит качество обслуживания клиентов.

Использование выявленных зависимостей и моделей, полученных в ходе анализа, будет полезным для работы транспортной компании. Технологии интеллектуального анализа данных о деятельности транспортного перевозчика, реализованные как посредством стандартных возможностей табличных редакторов, так и аналитических платформ, и языков программирования, играют важную роль в современной логистике и транспортной отрасли. Они позволяют собирать и обрабатывать большие объемы данных, анализировать их и принимать корректные решения для оптимизации процессов перевозки грузов и повышения эффективности работы не только транспортных компаний, но и всей транспортной отрасли.

Список использованных источников

1. Алексеев О. П., Пронин С. В. Применение баз данных для оценки работы пассажирского транспорта // Автомобильный транспорт. – 2009. – №. 24. URL: <https://cyberleninka.ru/article/n/primenenie-baz-dannyh-dlya-otsenki-raboty-passazhirskogo-transporta> . Дата обращения: 16.05.2024
 2. Хорсова А.В. Интеллектуальный анализ данных. [Электронный ресурс]. - URL: <https://cyberleninka.ru/article/n/intellektualnyy-analiz-dannyh-2/viewer> . Дата обращения: 02.05.2024
 3. Применение искусственного интеллекта для оптимизации маршрутов транспорта. [Электронный ресурс] – URL: <https://cyberleninka.ru/article/n/primenenie-iskusstvennogo-intellekta-dlya-optimizatsii-marshrutov-transporta/viewer> . Дата обращения: 12.06.2024.
 4. Флах П. Машинное обучение. М.: ДМК Пресс. 2015. 400 с.
 5. К. Каримов Методы искусственного интеллекта и применение их на транспорте [Электронный ресурс] - URL: <https://cyberleninka.ru/article/n/metody-iskusstvennogo-intellekta-i-primenenie-ih-na-transporte/viewer> . Дата обращения: 21.05.2024.
-

El-Kurdi Y.T., bachelor student, Naberezhnye Chelny Institute, Kazan (Volga Region) Federal University, email: yomna@mail.ru;

DEVELOPMENT OF TECHNOLOGY FOR INTELLIGENT ANALYSIS OF DATA ABOUT THE ACTIVITIES OF A TRANSPORT CARRIER

Abstract: Transport plays an important role in ensuring economic development and social well-being of society. However, with the growth of transportation volumes, new problems arise related to the processing and analysis of data on the activities of transport companies. This determines the relevance of developing new technologies for data analysis in the transport field. Growing volumes of transportation generate a large amount of data that requires systematization and analysis to make informed decisions. The development of intelligent data analysis technology will automate the processes of collecting, processing and interpreting information, which will lead to increased efficiency of transport companies, reduced costs and improved quality of services. Such research is of great practical importance and can become the basis for the development of innovative solutions in the field of transport logistics.

Keywords: transport company, data processing, technology, analysis, transportation.