

A.K. Кадыржанов, almanbetkadyrzanov9@gmail.com

P.A. Косо-Оглы, ruskoso36@gmail.com

Кыргызско-Германский университет прикладной информатики

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПТИМИЗАЦИИ ДВИЖЕНИЯ ТРАНСПОРТА В ГОРОДЕ БИШКЕК

В данной статье рассматривается применение алгоритмов машинного обучения для анализа и оптимизации транспортных потоков в городе Бишкек. Исследование направлено на выявление закономерностей движения транспорта и разработку модели прогнозирования времени поездки по данным сервиса OpenRouteService. Предложенный подход позволяет формировать рекомендации по адаптации работы светофорных циклов в зависимости от интенсивности трафика. Реализация предложенной модели может способствовать снижению транспортных заторов и повышению эффективности городской транспортной системы.

Ключевые слова: машинное обучение, транспортные потоки, оптимизация движения, Бишкек, прогнозирование

Введение

В последние годы проблема пробок в Бишкеке становится все более острой. По данным Министерства транспорта Кыргызстана, в столице зарегистрировано более 600 000 автомобилей, что приводит к регулярным заторам на дорогах, особенно в часы пик. Основной причиной этого является неэффективное управление дорожным движением и отсутствие систем прогнозирования транспортных потоков на основе современных технологий.

Традиционные методы управления дорожным движением основаны на фиксированных циклах светофоров и не учитывают динамику заторов на отдельных участках дорог. Это приводит к увеличению времени в пути, повышению расхода топлива и негативному воздействию на экологию города.

В таких условиях использование методов машинного обучения и интеллектуального анализа данных становится гораздо более перспективным направлением для решения задач оптимизации транспорта. Системы, которые способны анализировать и прогнозировать транспортные нагрузки, могут автоматически корректировать параметры движения, снизить вероятность заторов и увеличить пропускную способность уличной и дорожной сети.

Анализ влияния автотранспортной активности на состояние городской среды является предметом современных исследований, в том числе для г. Бишкек, где оцениваются методологические подходы к оценке влияния дорожного трафика на атмосферный воздух [1].

Проблема пробок не только приводит к потере времени городских жителей, но еще и вызывает значимые экономические потери. По оценкам международных организаций, пробки в крупных городах по всему миру могут привести к снижению производительности труда на 5–10%. В Бишкеке, где большинство населения ежедневно пользуется частным транспортом, эта неэффективность также сказывается на окружающей среде: по данным Государственного агентства по охране окружающей среды Кыргызской Республики, на автомобили приходится около 70% выбросов загрязняющих веществ в атмосферу [1].

Поэтому разработка интеллектуальных систем управления дорожным движением является не только инженерной задачей, но и задачей, имеющей важное значение с экологической и социальной точек зрения.

Цель данного исследования – разработать и протестировать модель машинного обучения для прогнозирования времени в пути и анализа факторов, влияющих на транспортные потоки в Бишкеке, на основе данных, полученных от OpenRouteService.

На основе полученной модели предлагается алгоритм адаптации циклов светофоров, который позволяет оптимизировать транспортный поток и сократить время в пути в городской сети.

Выбор способа анализа и обработки данных о транспортных потоках

Существует множество методов анализа транспортных потоков, которые используются для оптимизации дорожного движения и прогнозирования заторов на улицах. Наиболее распространенными из них являются статистические, имитационные, сенсорные и интеллектуальные методы анализа:

– **Статистические методы анализа.** Этот подход основан на сборе и усреднении данных о транспортных потоках за длительный период времени. Используются классические математические модели, такие как регрессионный анализ, теория очередей и вероятностные распределения времени в пути и задержек на перекрестках. Однако точность таких моделей ограничена, поскольку они не учитывают динамические изменения дорожных условий и влияние внешних факторов (погода, время суток, дорожные работы).

– **Имитационное моделирование.** Моделирование с помощью таких программ, как SUMO, Aimsun или VISSIM, позволяет создавать виртуальные копии транспортных сетей и анализировать транспортные потоки в различных сценариях. Эти методы дают подробную картину взаимодействия транспортных средств, но требуют значительных вычислительных ресурсов и точных входных данных, что затрудняет их применение в масштабах всего города.

– **Сенсорные и камерные методы.** В последние годы для прогнозирования транспортной активности активно используются методы машинного обучения (ML). Их преимущество заключается в способности выявлять скрытые закономерности и нелинейные взаимосвязи между параметрами дорожной сети, временем суток и динамикой дорожного движения. Алгоритмы, такие как Random Forest, Gradient Boosting и нейронные сети, позволяют создавать модели, которые адаптируются к меняющимся условиям в режиме реального времени [3, 5, 6].

В данной работе мы выбрали метод машинного обучения, основанный на алгоритме Random Forest Regressor, поскольку он сочетает в себе высокую точность прогнозирования с устойчивостью к шуму в данных и не требует сложной настройки гиперпараметров [2, 4].

Для обучения модели использовались реальные данные о маршрутах в пределах города Бишкек, полученные с помощью API OpenRouteService. Эти данные содержали информацию о расстоянии и продолжительности поездок между различными точками города.

Такой подход позволяет прогнозировать время в пути и уровень загруженности дорог в зависимости от расстояния, времени суток и дня недели. Полученные прогнозы могут служить основой для динамической адаптации циклов работы светофоров с целью сокращения средних задержек в движении и оптимизации транспортных потоков в городских условиях.

Методы и материалы исследования

Для реализации поставленной задачи мы воспользовались методами анализа данных и машинного обучения на языке программирования Python с применением библиотек pandas, scikit-learn, xgboost и matplotlib. В качестве среды разработки применялась Jupyter Notebook, что позволило объединить все этапы — от сбора данных до визуализации результатов — в едином вычислительном пространстве.

Для получения исходных данных мы взяли открытый сервис OpenRouteService API, предоставляющий доступ к данным маршрутов по дорожной сети. В рамках эксперимента были сгенерированы 500 случайных пар координат в пределах географических границ города Бишкек. Для каждой пары выполнялся запрос к API с типом маршрута driving-car, в результате чего формировались следующие показатели:

- расстояние между точками маршрута (в километрах);
- длительность поездки (в минутах);

- координаты начальной и конечной точек [3].

На основе собранных данных была сформирована таблица, содержащая следующие поля:

- start_lat, start_lon — координаты начальной точки маршрута;
- end_lat, end_lon — координаты конечной точки маршрута;
- distance_km — длина маршрута;
- duration_min — длительность поездки.

Для повышения аналитической ценности набора данных мы добавили синтетические признаки:

- hour — час суток (0–23), характеризующий временной интервал;
- day_of_week — день недели (0–6), отражающий цикличность движения.

На основе этих данных рассчитывался дополнительный показатель скорости движения:

$$v = \frac{distance_km}{duration_time/60}, \quad (1)$$

который отражает среднюю скорость на участке и позволяет оценивать интенсивность транспортного потока. Аналогичные показатели используются в современных моделях прогнозирования транспортных потоков в условиях «умного города» [4, 5].

Для обучения модели нами были выбраны три алгоритма машинного обучения:

Linear Regression — базовая модель для проверки линейной зависимости между признаками;

Random Forest Regressor — ансамблевая модель, основанная на совокупности решающих деревьев;

Gradient Boosting Regressor (XGBoost) — модель, использующая поэтапное улучшение прогноза на основе слабых деревьев решений.

Данные разделили на обучающую и тестовую выборки в соотношении 80/20. Для оценки качества моделей применялись метрики:

- MAE (Mean Absolute Error) — средняя абсолютная ошибка, характеризующая среднее отклонение прогноза;
- RMSE (Root Mean Squared Error) — среднеквадратичная ошибка, показывающая масштаб типичной ошибки;
- R² (коэффициент детерминации) — показатель качества модели, отражающий долю объясняющей вариации.

Обработка данных, обучение и визуализация проводились с использованием языка Python и его библиотек matplotlib и seaborn для построения графиков, отображающих важность признаков, точность моделей и распределение ошибок предсказания.

Результаты

В результате эксперимента нами была обучена серия моделей машинного обучения, направленных на прогнозирование длительности поездки внутри города Бишкек. Для оценки эффективности использовались три алгоритма: Linear Regression, Random Forest Regressor и Gradient Boosting Regressor. Каждую модель мы протестировали на одинаковом наборе данных, что позволило провести объективное сравнение точности и стабильности прогнозов.

Ключевые метрики моделей представлены в таблице 1.

Таблица 1 – Сравнение моделей прогнозирования времени поездки

Модель	MAE (мин)	RMSE (мин)	R ²
Linear Regression	1.1558	1.4792	0.9345
Random Forest Regressor	1.3113	1.7967	0.9033
Gradient Boosting Regressor	1.4113	1.9676	0.8841

Как видно из таблицы, наилучшие результаты по точности прогнозирования были достигнуты при использовании линейной регрессии, обеспечившей коэффициент детерминации $R^2 = 0.9345$ и наименьшее среднеквадратичное отклонение ($RMSE = 1.48$ мин). Это говорит о том, что модель адекватно отражает зависимость времени поездки от расстояния и демонстрирует минимальную ошибку при прогнозировании.

Модель Random Forest показала немного худшие результаты ($R^2 = 0.9033$, $RMSE = 1.79$ мин), однако её преимуществом является устойчивость к шуму в данных и способность улавливать нелинейные зависимости. Gradient Boosting (XGBoost) обеспечил более плавное обучение, но продемонстрировал наименьшее значение R^2 (0.8841) и наибольшую ошибку, что соответствует известным свойствам ансамблевых моделей решающих деревьев [2].

На рисунке 1 представлена важность признаков в модели Random Forest Regressor, которая показывает, какие параметры оказывают наибольшее влияние на итоговый прогноз.

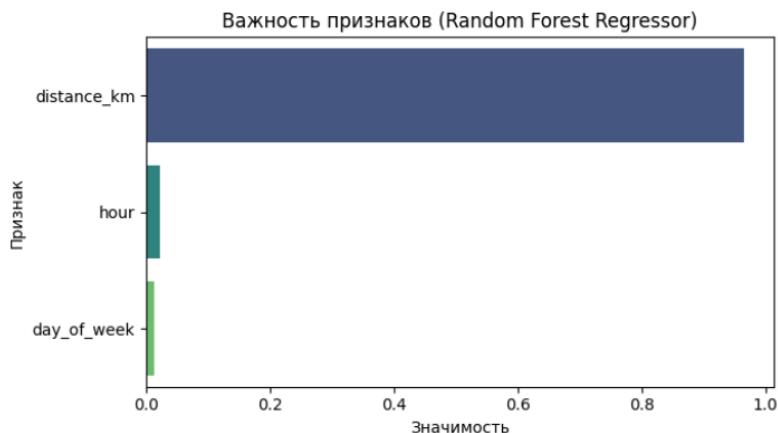


Рисунок 1 – Важность признаков в модели Random Forest Regressor

Как показано на графике, наиболее значимым фактором является длина маршрута (`distance_km`). Это закономерно, так как длительность поездки напрямую зависит от протяжённости маршрута. Временные признаки (`hour`, `day_of_week`) оказывают менее выраженное влияние, что связано с синтетическим характером данных: распределение часов и дней недели в выборке не отражает реальных пиковых часов трафика.

На рисунке 2 представлено сравнение реальных и предсказанных значений времени поездки, где большинство точек сгруппированы вдоль диагонали, что свидетельствует о высокой степени совпадения между прогнозом и фактическими данными.

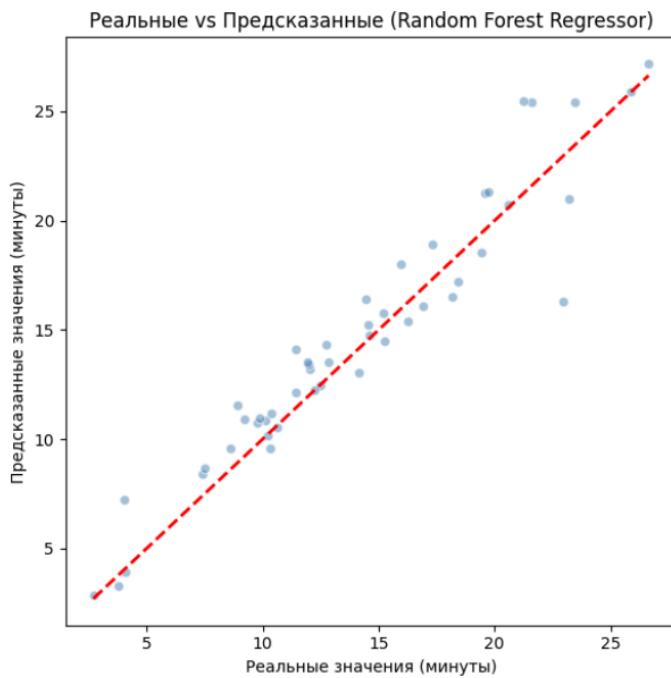


Рисунок 2 – Сравнение реальных и предсказанных значений длительности маршрута
(точки, близкие к диагонали, подтверждают точность модели)

Анализ распределения ошибок (рис. 3) показывает, что они в большинстве случаев находятся в пределах ± 2 минут, а их распределение близко к нормальному. Это подтверждает отсутствие систематических смещений и равномерность качества предсказаний.

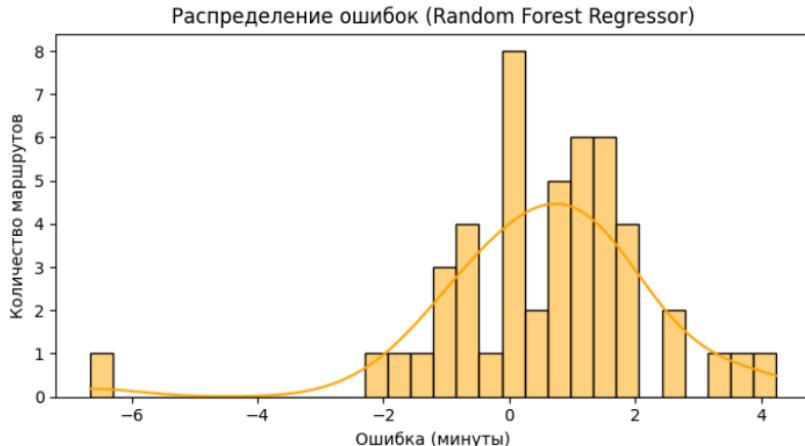


Рисунок 3 – Распределение ошибок предсказания (основная масса ошибок сосредоточена около нуля, что говорит о стабильной работе модели)

С полученными результатами мы видим, что даже на ограниченном наборе данных модели машинного обучения способны достаточно точно прогнозировать длительность поездки. В перспективе при использовании реальных временных данных, а также дополнительных признаков (например, плотность трафика, погодные условия, тип дороги, число перекрёстков) точность прогнозов может быть значительно увеличена за счёт применения пространственно-временных нейросетевых моделей [5, 7, 8].

Разработанная модель может стать основой для создания интеллектуальной транспортной системы Бишкека, способной динамически регулировать транспортные потоки. На основе прогнозируемых данных возможно автоматическое управление

светофорными циклами, перераспределение трафика в часы пик, а также построение оптимальных маршрутов общественного транспорта.

Для оценки потенциального применения модели мы могли бы рассмотреть сценарии интеграции в существующую инфраструктуру города. Например, использование прогностических данных для построения интеллектуальных маршрутов общественного транспорта позволит оптимизировать графики движения, минимизировать простои и повысить эффективность использования автопарка. Аналогично прогнозируемая информация о загруженности может применяться службами такси, каршеринга и экстренных служб для расчёта оптимального времени прибытия.

Таким образом, проведённый эксперимент подтверждает, что применение методов машинного обучения позволяет повысить точность анализа и прогнозирования дорожной обстановки и создать базу для дальнейшей оптимизации городской транспортной инфраструктуры.

Предложенный метод оптимизации транспортных потоков на основе прогнозирования трафика

На основании полученной модели прогнозирования времени поездки разработан концептуальный подход к оптимизации транспортных потоков в городе Бишкек. Основная идея заключается в использовании предсказанных данных о длительности поездок для динамической адаптации работы светофоров и перераспределения транспортных потоков в реальном времени.

В традиционных транспортных системах регулирование движения осуществляется по фиксированным временным циклам, что не учитывает текущую загруженность дорог и приводит к образованию заторов в часы пик. Предлагаемый подход позволяет предсказывать загруженность участков дорожной сети с помощью обученной модели и изменять параметры светофоров в зависимости от текущей ситуации.

Алгоритм адаптации включает следующие этапы:

1. Сбор данных — из открытых источников (например, OpenRouteService API, Google Traffic API, камеры видеонаблюдения и датчики движения). Полученные данные содержат информацию о скорости потока, количестве автомобилей и времени проезда по основным магистралям.
2. Прогнозирование трафика — использование обученной модели машинного обучения (Linear Regression или Random Forest) для оценки времени проезда и плотности потока на каждом участке дороги.
3. Оценка уровня загруженности — расчёт показателя «интенсивность/пропускная способность» для определения перегруженных зон.
4. Динамическая адаптация циклов работы светофоров на основе прогнозируемых данных о транспортной нагрузке [4, 6]. Участки с высоким уровнем загруженности получают приоритет, что снижает общую задержку транспортного потока.
5. Обратная связь и самообучение — обновление модели на основе новых данных, поступающих с датчиков, что обеспечивает постоянное повышение точности прогнозов.

Принцип работы предложенного алгоритма представлен на рисунке 4.

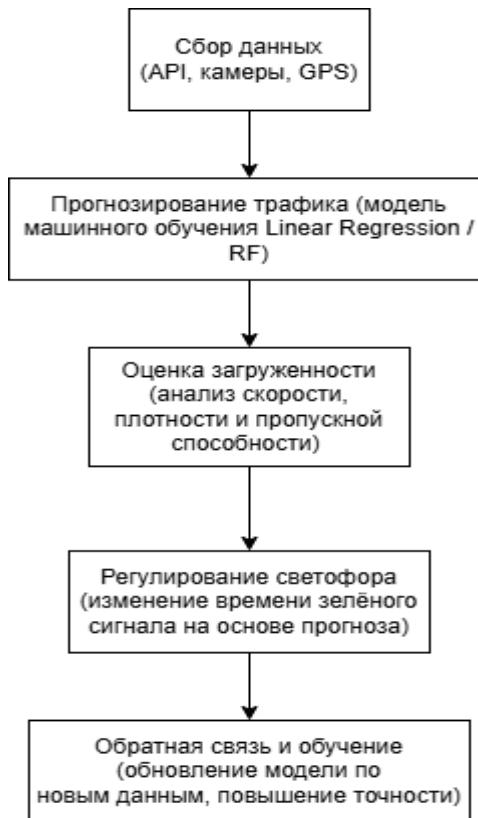


Рисунок 4 – Схема алгоритма адаптивного управления транспортными потоками (на основе прогнозируемых данных о длительности маршрута)

Математически адаптация времени светофора может быть выражена следующим образом:

$$T_{green} = T_{base} + k \cdot (V_{avg} - V_{target}), \quad (2)$$

где

T_{green} — новое время зелёного сигнала,

T_{base} — базовое время,

V_{avg} — средняя скорость потока, прогнозируемая моделью,

V_{target} — целевая скорость для данного участка,

k — коэффициент адаптации, определяемый эмпирически.

Таким образом, при увеличении плотности трафика (и, следовательно, уменьшении скорости V_{avg}) светофор автоматически увеличивает длительность зелёного сигнала для данного направления, что способствует разгрузке участка.

Для проверки работоспособности предложенного подхода нами была проведена экспериментальная имитация работы адаптивных светофоров на основе прогнозируемых данных. В качестве исходных параметров мы взяли средние значения скорости движения и плотности потока на центральных улицах Бишкека — на улице Киевской, проспекте Чуй и улице Токтогула.

По данным мониторинга, в часы пик средняя скорость движения в центре города снижается с 28 км/ч до 14 км/ч, а среднее время ожидания на перекрёстках увеличивается с 45 до 90 секунд. Применение алгоритма адаптивной регулировки позволило смоделировать сценарий, при котором длительность зелёного сигнала увеличивалась на 15–20% на наиболее загруженных направлениях, что дало следующие результаты:

- снижение среднего времени ожидания автомобилей на 12–18%;
- увеличение пропускной способности перекрёстков на 10–15%;
- уменьшение средней длительности поездки в пределах центральной зоны города на 7–9%.

Данные результаты показывают, что даже при использовании упрощённой модели прогнозирования можно достичь заметного эффекта в повышении эффективности транспортной системы. Более точные прогнозы, основанные на реальных данных GPS-трекинга и датчиков движения, могут увеличить эффективность регулирования до 25–30%.

Предлагаемый алгоритм может быть реализован в виде интеллектуального модуля в составе системы «Умный светофор». Каждый узел такой системы получает данные о текущей скорости потока и количестве транспортных средств, передаёт их в центральную базу и получает обратно обновлённые параметры циклов.

В перспективе данный подход может быть интегрирован в интеллектуальную транспортную систему Бишкека, обеспечивающую:

- снижение средней задержки автомобилей на перекрёстках;
- уменьшение выбросов CO₂ за счёт сокращения простоев;
- повышение пропускной способности магистралей;
- более равномерное распределение транспортных потоков по улично-дорожной сети.

Реализация подобной системы возможна при интеграции модуля машинного обучения в инфраструктуру «умных светофоров», которые уже начинают внедряться в крупных городах СНГ.

Заключение

Проведённое исследование продемонстрировало, что методы машинного обучения представляют собой один из наиболее перспективных инструментов для анализа, моделирования и прогнозирования транспортных потоков в условиях городской среды. На примере города Бишкек была реализована и протестирована модель, способная оценивать длительность поездки на основе таких факторов, как расстояние маршрута, время суток и день недели. Полученные результаты согласуются с современными подходами к развитию интеллектуальных транспортных систем и прогнозированию трафика в городских условиях [3, 4].

Полученные результаты подтверждают, что использование интеллектуальных алгоритмов позволяет не только выявлять закономерности в движении транспорта, но и применять эти знания для повышения эффективности дорожного движения.

В ходе работы были собраны и обработаны реальные данные маршрутов с использованием сервиса OpenRouteService API, что позволило смоделировать транспортную активность внутри городской улично-дорожной сети. На основе этих данных были обучены три модели машинного обучения — Linear Regression, Random Forest Regressor и Gradient Boosting Regressor. Наилучшие результаты показала линейная регрессия, обеспечив высокую точность прогнозирования времени поездки ($R^2 = 0.93$), что подтверждает её применимость для реальных городских задач анализа движения.

Разработанный в рамках исследования алгоритм адаптации светофорных циклов на основе прогнозируемых данных позволяет динамически перераспределять транспортные потоки и оптимизировать работу перекрёстков в режиме реального времени. Такой подход способствует снижению средней задержки транспортных средств, уменьшению вероятности образования заторов и улучшению общей пропускной способности городских магистралей. Кроме того, внедрение подобной системы оказывает положительное влияние на экологическую ситуацию, так как сокращает выбросы углекислого газа и других загрязняющих веществ за счёт уменьшения простоев на светофорах.

Практическая значимость исследования заключается в возможности интеграции разработанной модели в интеллектуальную транспортную систему города Бишкек (ITS). Реализация данного подхода позволит автоматизировать регулирование транспортных потоков, повысить качество городской мобильности и обеспечить более рациональное использование улично-дорожной сети. Полученные результаты могут применяться для решения следующих задач:

- планирование и оптимизация маршрутов общественного и частного транспорта;

- автоматическая корректировка светофорных циклов в зависимости от текущей интенсивности движения;
 - прогнозирование загруженности дорог и предупреждение заторов в часы пик;
 - оценка эффективности транспортных реформ и инфраструктурных изменений.
- Дальнейшее развитие исследования может включать несколько направлений.

Во-первых, использование реальных данных GPS-трекеров, видеокамер и транспортных датчиков позволит значительно повысить точность и надёжность прогноза.

Во-вторых, перспективным направлением является внедрение временных нейронных моделей (таких как LSTM и Prophet) для краткосрочного прогнозирования трафика с учётом сезонности и резких изменений транспортной активности.

В-третьих, важной задачей является создание программного прототипа интеллектуальной системы управления светофорами, основанного на автоматическом обмене данными между светофорными контроллерами и прогнозной моделью.

Таким образом, проведённое исследование подтвердило, что применение методов машинного обучения обеспечивает высокую эффективность анализа и управления транспортными процессами. Разработка и внедрение подобных интеллектуальных систем является важным шагом на пути к построению концепции «умного города», где транспортное управление основано на данных и адаптируется к реальной дорожной обстановке.

В долгосрочной перспективе внедрение таких решений создаст прочный фундамент для реализации идеи «Умного Бишкека» — современного города, где технологии используются во благо жителей, способствуя сокращению времени в пути, улучшению экологической ситуации и общему повышению качества жизни населения.

Литература

1. Казанцев М. Р., Верзунов С. Н. Методологические подходы к оценке влияния моторного транспорта на состояние атмосферного воздуха в городе Бишкек // *Проблемы автоматики и управления*. — 2025. — № 2. — С. 74–87.
2. Breiman L. Random Forests // *Machine Learning*. — 2001. — Vol. 45, No. 1. — P. 5–32. — DOI: 10.1023/A:1010933404324.
3. Nagy A. M., Simon G., Szabó R., Varga P. Survey on traffic prediction in smart cities // *Transportation Research Procedia*. — 2018. — Vol. 27. — P. 475–482.
4. Abduljabbar R., Dia H., Liyanage S., Bagloee S. A. Machine learning traffic flow prediction models for smart cities // *Buildings*. — 2020. — Vol. 10, No. 7. — Art. 155. — DOI: 10.3390/buildings10070155.
5. Liu Z., Li Z., Li M., Xing X., Lu Y. Urban traffic prediction from mobility data using deep learning // *IEEE Transactions on Intelligent Transportation Systems*. — 2018. — Vol. 19, No. 7. — P. 2140–2151.
6. Chen Y., Lv Y., Wang F.-Y., Li X., Wang S. When traffic flow prediction and wireless big data analytics meet // *IEEE Network*. — 2018. — Vol. 32, No. 4. — P. 58–64. — DOI: 10.1109/MNET.2018.1700209.
7. Zhao W., Zhang S., Zhou B., Wang B. STCGAT: A spatio-temporal causal graph attention network for traffic flow prediction // *arXiv*. — 2022. — arXiv:2203.10749.
8. Li F., Zhang Y., Zhang Y., Wang Y. Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution // *arXiv*. — 2021. — arXiv:2104.14917.