

— СТУДЕНЧЕСКАЯ НАУЧНАЯ КОНФЕРЕНЦИЯ — 2026

Сравнение методов прогнозирования суточного спроса на городской велопрокат

АВТОР

Соколов Д. А., студент 3 курса

НАУЧНЫЙ РУКОВОДИТЕЛЬ

Лаврова И. В., канд. техн. наук, доцент

СЕКЦИЯ

Прикладная математика и информатика, Казань, 2026

Актуальность: зачем прокату прогноз на завтра

- ◆ Недооценка спроса в утренний и вечерний пик оборачивается пустыми стойками и потерянными поездками пользователей.
- ◆ Переоценка спроса — лишними велосипедами на станциях и лишними перевозками при перебалансировке.
- ◆ Ошибка прогноза ведёт к простоям техники, сорванным поездкам и росту затрат на ручную перебалансировку.

ПРОБЛЕМА ИССЛЕДОВАНИЯ

Прокат не может мгновенно перераспределить велосипеды: решение о развозке по станциям и плане ремонта принимается накануне, а ошибка прогноза напрямую превращается в простой техники, сорванные поездки и рост затрат на ручную перебалансировку.

— ЦЕЛЬ И ЗАДАЧИ

Цель работы и три задачи

1

Собрать и подготовить данные

Данные городского проката за 2023–2025 годы: 1 096 суточных записей с признаками «день недели», «праздник», «температура», «осадки».



2

Реализовать методы и схему проверки

Три метода и единая схема: обучение на 2023–2024 годах, проверка на данных 2025 года, скользящая проверка по пяти блокам.



3

Сравнить методы и найти слабые места

Сравнение по метрикам MAE и MAPE и определение условий, где прогноз ошибается сильнее всего.

Открытый набор наблюдений городского велопроката, 2023–2025

1 096

Суточных записей

По одной строке на каждый день наблюдений

3 года

Период наблюдений

Данные 2023–2025 годов

4

Признака-предиктора

День недели, праздник, температура, осадки

1

Целевая переменная

Число поездок за сутки

≈3%

Доля праздничных дней

Оценка по календарю выборки

Учебный вычислительный эксперимент на открытых данных; название проката и имена вымышленные.

Учебный набор наблюдений городского велопроката, 2023–2025 (данные и результаты — учебный эксперимент)

Методика вычислительного эксперимента

1

Обучение моделей

Данные 2023–2024 годов: день недели, праздник, температура, осадки. Настройка параметров каждого метода.



2

Проверка на 2025 годе

Отложенная выборка 2025 года — модель не видела эти наблюдения при обучении.



3

Скользящая проверка

Пять блоков последовательной проверки: обучение на прошлом, прогноз на следующем интервале.



4

Расчёт ошибок

Отклонение прогноза от факта: абсолютная ошибка в поездках (MAE) и относительная в процентах (MAPE).

Три сравниваемых метода прогнозирования

01

Скользящее среднее за 7 дней

Базовый метод без признаков: прогноз равен среднему спросу за прошлую неделю. Прост в расчёте и служит точкой отсчёта. Ограничение — не учитывает погоду, праздники и тренд, поэтому сглаживает пики и провалы.

02

Линейная регрессия с признаками

На входе — температура, осадки, день недели и признак праздника. Модель прозрачна, коэффициенты объяснимы. Ограничение — улавливает лишь линейные связи.

03

Градиентный бустинг

Ансамбль деревьев решений, улавливающий нелинейные зависимости спроса от погоды и календаря. На входе те же признаки, что и у регрессии. Ограничение — риск переобучения и меньшая прозрачность модели.

Все три метода обучались на одних данных 2023–2024 годов и проверялись на 2025 годе.

Метрики качества прогноза: MAE и MAPE

MAE измеряется в поездках, MAPE — в процентах, что позволяет сравнивать методы между собой.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

y_i	Фактический суточный спрос	поездки
$\hat{y}_i$	Прогноз модели на тот же день	поездки
n	Число наблюдений в проверке	365
MAPE	Средняя абсолютная ошибка в процентах	%

MAPE не определена при $y_i = 0$; в выборке таких суток нет.

Сравнение трёх методов прогнозирования на данных 2025 года

Метод	MAPE, %	MAE, поездок
Скользящее среднее за 7 дней	18,4	214
Линейная регрессия с признаками погоды	14,1	163
Градиентный бустинг	9,7	112

Градиентный бустинг лучше скользящего среднего на 8,7 п.п. по MAPE и почти вдвое по MAE.

Где прогноз промахивается сильнее всего

МИНИМАЛЬНАЯ ОШИБКА

Обычный день + стабильная погода

Минимальная ошибка: модель работает устойчиво, режим поездок предсказуем и хорошо представлен в обучающей выборке 2023–2024 годов.

УМЕРЕННАЯ ОШИБКА

Обычный день + резкая смена погоды

Умеренный рост ошибки: похолодание или дождь смещают спрос, но будний профиль поездок частично компенсирует отклонение.

ЗАМЕТНЫЙ ПРОМАХ

Праздничный день + стабильная погода

Заметный промах: нетипичный профиль праздника слабо представлен в данных, модель недооценивает спрос даже при хорошей погоде.

МАКСИМАЛЬНАЯ ОШИБКА

Праздничный день + резкая смена погоды

Максимальный MAPE: сочетание редкое в обучении, поведение пользователей нетипично, модель сильнее всего недооценивает спрос.

ПОГОДА →

ТИП ДНЯ →

Границы применимости полученных результатов

01

Один город

Эксперимент проведён на данных одного города. Переносить найденные закономерности на другие города без повторной проверки нельзя: структура поездок и климат там иные.

02

Три года наблюдений

В распоряжении 2023–2025 годы — 1 096 суточных записей. Редких сочетаний «праздник + плохая погода» в выборке мало, поэтому именно на них модель обучена слабее всего.

03

Фактическая погода вместо прогнозной

В модель подавались фактические температура и осадки. В реальном режиме «прогноз на завтра» доступен только прогноз погоды, поэтому ошибка будет выше полученных 9,7–18,4 %.

04

Ограниченный набор признаков

Учтены только календарь и погода. Не вошли тарифы, ремонты станций, городские события и загруженность соседних станций — часть ошибок объясняется именно этим.

Учебный эксперимент, не официальная статистика; название проката и имена вымышлены.

Учебный вычислительный эксперимент, данные 2023–2025 гг.

Выводы, планы на 2027 год и литература

- 1. Градиентный бустинг даёт лучший результат: MAPE 9,7 % и MAE 112 поездок против 14,1 % / 163 у линейной регрессии и 18,4 % / 214 у скользящего среднего.
- 2. Календарно-погодные признаки и нелинейность модели дают основной выигрыш в точности прогноза суточного спроса.
- 3. Добавить второй город, перейти на прогнозную погоду вместо фактической и расширить набор признаков.
- 4. Открытый набор наблюдений городского велопроката за 2023–2025 годы, 1 096 суточных записей (учебный эксперимент).
- 5. Работа по градиентному бустингу: описание алгоритма и настройки гиперпараметров для задач регрессии.
- 6. Руководство по метрикам качества прогноза MAE и MAPE: формулы и интерпретация значений.
- 7. Исследование по прогнозированию спроса на городской велопрокат с учётом погоды и календаря.

Контакт: кафедра прикладной информатики

Вопросы: Благодарю за внимание! Готов ответить на вопросы.