



## МНОГОУРОВНЕВАЯ АНАЛИТИЧЕСКАЯ СИСТЕМА МОНИТОРИНГА И ПРОГНОЗИРОВАНИЯ ПРЕСТУПНОСТИ

*Камалова Нилуфар Илхомовна – PhD, доцент Бухарского  
государственного университета.*

*Шавкатова Шаҳинабону Шухрат кизи – студентка 2 курса Бухарского  
государственного университета.*

**Аннотация.** В данном исследовании разработана многоуровневая аналитическая система для автоматизации мониторинга статистики преступности. Система построена на основе интеграции реляционных баз данных, статистического анализа и моделей машинного обучения. Концептуальная ER-модель обеспечивает структурную целостность данных.

Статистический анализ позволил выявить взаимосвязи между территориальными и демографическими показателями. На этапе прогнозирования использовалась модель линейной регрессии для прогнозирования возраста преступника на основе возраста жертвы.

**Ключевые слова:** Статистика преступности, реляционная база данных, статистический анализ, аналитическая система поддержки принятия решений, модель машинного обучения, визуализация, язык программирования Python.

### **Введение**

Растущий уровень преступности в современном обществе требует комплексного подхода к обеспечению безопасности государства и общества. Если бы статистика преступности оценивалась только на основе количественных показателей, то сегодня в научных исследованиях все чаще подчеркивается недостаточность такого подхода. Поэтому использование современных методов анализа данных приобретает все большее значение для



углубленного изучения преступности, выявления ее причин и прогнозирования.

Научная работа «Анализ преступности с использованием анализа данных», опубликованная IGI Global, обосновывает необходимость систематического анализа статистики преступности на основе больших объемов данных. В исследовании изучалась база данных преступлений, сформированная по времени, территории и видам преступлений с использованием статистических и аналитических моделей. Авторы подчеркивают важность не только регистрации преступлений, но и выявления их динамики и закономерностей. В статье показано, что использование современных аналитических подходов, включая регрессионный анализ, кластеризацию, визуализацию и алгоритмы машинного обучения, наряду с традиционными статистическими методами, в анализе статистики преступности является эффективным. Эти методы позволяют определить уровень преступности и ее связь с социально-экономическими факторами. В результате можно выявить основные факторы, влияющие на преступность, и принимать научно обоснованные решения по ее снижению [1].

Помимо изучения уровня преступности на основе традиционных статистических методов, использование технологий искусственного интеллекта, в частности алгоритмов машинного обучения и глубокого обучения, может значительно улучшить процесс прогнозирования преступности. [2] – в статье систематически анализируются существующие научные исследования по прогнозированию преступности и рассматриваются возможности и ограничения этих подходов на научной основе. В исследовании рассмотрено более 150 научных статей, каждая из которых содержит подробную информацию о том, как методы машинного обучения и глубокого обучения используются для прогнозирования будущих тенденций преступности. Эти подходы позволяют выявлять сложные закономерности в



данных о преступности, генерировать точные прогнозы с учетом временных и пространственных характеристик. Кроме того, проанализированные исследования предоставляют критерии оценки эффективности моделей прогнозирования преступности, используемых наборов данных, а также преимущества и ограничения используемых алгоритмов машинного обучения/глубокого обучения.

Возможности прогнозирования преступлений заранее путем анализа данных о динамическом поведении человека возрастают. Одно из исследований, основанных на этом подходе, — научная статья «Однажды произошло преступление: к прогнозированию преступности на основе демографических и мобильных данных», в которой предлагается новая методология, объединяющая агрегированные данные мобильных телефонов и демографическую информацию для прогнозирования преступлений. В исследовании авторы объединяют анонимизированные данные о поведении людей из инфраструктуры мобильных сетей с демографическими показателями, в отличие от методов, обычно используемых для прогнозирования преступности. Этот подход превращает задачу выявления районов с высокой вероятностью совершения преступлений в задачу бинарной классификации и использует алгоритмы на основе деревьев решений, такие как случайный лес.

В статье «Социально-экономические, градостроительные и транспортные условия, связанные с преступностью» авторы рассматривают три основные группы факторов при анализе уровня преступности:

1. Социально-экономические факторы — уровень дохода, уровень безработицы, уровень образования.
2. Городская среда (градостроительная среда) — плотность застройки, дороги и транспортная система, планировка населенных пунктов.



3. Транспортные условия — перемещение населения и транспортные маршруты, потоки людей и их распределение по территории. Результаты исследования показывают, что сочетание вышеперечисленных факторов обладает высокой прогностической способностью в определении уровня преступности. Например, в густонаселенных районах, районах с низким уровнем дохода или районах с высокой мобильностью вероятность совершения преступлений значительно возрастает. Также отмечается возможность использования геопространственных и статистических моделей в анализе преступности. С помощью этих моделей можно определить взаимосвязь между социально-экономическими показателями и показателями мобильности и уровнем преступности, создать карты рисков для различных регионов и разработать стратегии профилактики[3].

В статье «Изучение роли статистического анализа в криминологии» авторы выделяют три основные функции статистического анализа в изучении преступности:

- Выявление трендов — мониторинг и прогнозирование изменений уровня преступности во времени.
- Выявление корреляций — анализ взаимосвязи между социальными, экономическими, урбанистическими и демографическими характеристиками и уровнем преступности.
- Поддержка принятия решений — разработка стратегий профилактики и политики безопасности на основе статистических результатов.

В статье также подчеркивается эффективное использование регрессионных моделей, кластеризации, анализа временных рядов и других статистических методов в анализе преступности. Эти подходы позволяют выявлять основные факторы, влияющие на преступность, анализировать



пространственно-временное распределение преступлений и эффективно распределять ресурсы.

Результаты исследования показывают, что статистический анализ важен для понимания социально-экономических факторов, влияющих на преступность, и повышения эффективности мер по предотвращению преступности. При этом в качестве основных факторов, определяющих точность результатов, подчеркиваются качество, полнота и систематичность данных, а также правильный выбор статистических методов[4].

Анализ преступности. В руководстве Р. П. Макколи «Анализ преступности» процесс анализа преступности описывается в трех этапах:

1. Сбор данных - формирование систематической базы данных о типах преступлений, территориальном распределении, времени и других характеристиках.

2. Анализ и визуализация - выявление закономерностей преступности, определение очагов преступности и анализ взаимосвязей с использованием статистических методов.

3. Принятие решений и профилактика - разработка мер по предотвращению преступности и эффективное распределение ресурсов на основе результатов анализа.

В руководстве подробно рассматриваются практические аспекты анализа преступности, включая геопространственный анализ, анализ временных рядов и методы выявления взаимосвязей между показателями. Эти методы показаны как эффективные инструменты для разработки стратегий прогнозирования и предотвращения преступности. В статье также подчеркиваются преимущества использования статистических программных средств и аналитических моделей в анализе преступности. Такой подход позволяет быстро обрабатывать данные, визуализировать их в виде графиков и карт, а также оптимизировать процесс принятия решений.



## Методология

В данном исследовании была разработана информационная система с многоуровневой аналитической архитектурой для мониторинга статистики преступлений. Исследование основано на практико- экспериментальном подходе и включает интеграцию реляционной базы данных, методов статистического анализа и моделей машинного обучения (Рисунок 1).

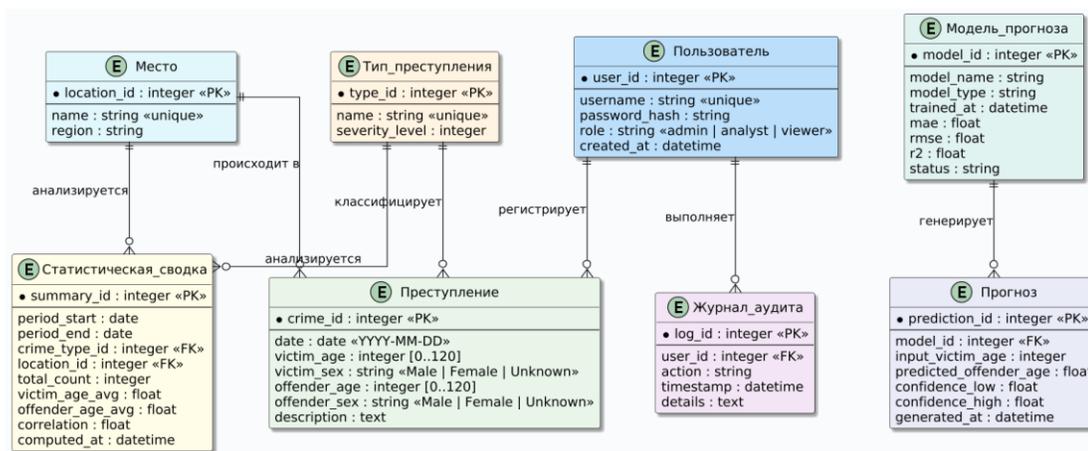


Рисунок 1. Концептуальная ER-модель системы мониторинга статистики преступлений.

### Архитектура системы

Система концептуально состоит из трех слоев:

1. Основной слой
2. Статистический слой
3. Слой машинного обучения

Слой транзакций обеспечивает сбор и хранение данных в реальном времени. Этот слой содержит основные сущности, такие как Пользователь, Преступление, Тип преступления, Местоположение и Журнал аудита. Таблица «Преступления» — это центральная таблица фактов, в которой хранятся атрибуты преступления, такие как дата, район, жертва и преступник. Для



обеспечения целостности данных используются первичные и внешние ключи, а также ограничения домена.

На аналитическом слое генерируются статистические показатели. На этом этапе данные агрегируются, определяются распределения и рассчитываются корреляции. Прогностический слой выполняет прогнозирование на основе модели машинного обучения.

Система работает по замкнутому аналитическому циклу:

$$Д \rightarrow С \rightarrow М \rightarrow В \rightarrow \Phi \rightarrow Д$$

где Д — исходные данные, С - статистические результаты, М - модель машинного обучения, В - визуализация, а  $\Phi$  представляет собой процесс принятия решений и обратной связи.

*Сбор и обработка данных*

Данные хранятся в реляционной базе данных SQLite. Каждая запись содержит следующие атрибуты:

- тип преступления,
- дата,
- регион,
- возраст и пол жертвы
- возраст и пол преступника,
- дополнительное описание.

Введенные данные проверяются с помощью механизма валидации. Возрастные показатели ограничены диапазоном 0–120, дата хранится в формате ISO. На следующем этапе данные преобразуются в DataFrame pandas и подготавливаются для статистического анализа.

Методы статистического анализа. В статистическом модуле были выполнены следующие операции:

- Описательная статистика (среднее, медиана, дисперсия).
- Частотный анализ.



- Пространственная агрегация.
- Расчет коэффициента корреляции Пирсона.

Модель машинного обучения

На этапе прогнозирования использовалась простая модель линейной регрессии. Модель выражается следующим образом:

$$\hat{y} = Wx + b$$

где  $x$  — возраст жертвы,  $\hat{y}$  — прогнозируемый возраст преступника, а  $W$  и  $b$  — параметры модели.

Модель была обучена с использованием библиотеки Python. В качестве функции потерь была выбрана среднеквадратичная ошибка (MSE). Параметры были оптимизированы с помощью алгоритма градиентного спуска. Модель обучалась в течение 300 итераций, и использовался метод разделения на обучающую и тестовую выборки.

В качестве критериев оценки использовались MSE, MAE и коэффициент детерминации ( $R^2$ ).

#### *Визуальная аналитика*

Результаты были визуализированы с помощью matplotlib. Были сгенерированы столбчатые диаграммы по типам преступлений, круговые диаграммы регионального распределения и графики регрессионных линий. Визуальные результаты служат для поддержки процесса принятия решений.

#### **Результаты**

Разработанная система позволила автоматизировать мониторинг статистики преступности. Данные анализировались по регионам и видам преступлений, и в некоторых регионах наблюдалась высокая концентрация определенных видов преступлений.

Была обнаружена положительная корреляция между возрастом жертвы и преступника. Эта взаимосвязь была статистически значимой, что указывает на то, что возраст является важным параметром в прогностической модели.



Модель линейной регрессии дала удовлетворительные результаты в прогнозировании возраста преступника. Показатели ошибок модели были низкими, а значение  $R^2$  показало, что модель может объяснить значительную часть дисперсии данных.

Благодаря замкнутой архитектуре системы существует возможность переобучения модели с добавлением новых данных. Это делает систему адаптивной и масштабируемой платформой.

В целом, разработанная система мониторинга оказалась эффективным инструментом для углубленного анализа динамики преступности, оценки уровня территориальной безопасности и разработки управленческих решений на основе данных.

### **Обсуждение**

Результаты данного исследования показали эффективность организации мониторинга статистики преступности на основе автоматизированного и многоуровневого аналитического подхода. Разработанная система позволяет не только централизованно хранить данные, но и анализировать их на статистическом и прогностическом уровне. Это позволяет делать более широкие и глубокие аналитические выводы по сравнению с традиционными формами отчетности.

Результаты, полученные в территориальном разрезе, показали неравномерное распределение видов преступлений. Было установлено, что в некоторых регионах наблюдается высокая концентрация определенных видов преступлений, что позволяет предположить, что это может быть связано с социально-экономическими факторами, демографическими показателями или инфраструктурой. Таким образом, система служит аналитической основой для разработки региональной стратегии безопасности.

Выявленная корреляция между возрастом жертвы и преступника подтверждает, что это важный фактор для прогностической модели. Однако



показатель возраста не следует интерпретировать как единственный определяющий фактор. Преступления представляют собой многофакторный процесс и также связаны с социальной средой, экономическими условиями, уровнем образования и другими факторами. Поэтому в будущем целесообразно использовать многомерные модели.

Модель линейной регрессии обеспечила простой и понятный подход на начальном этапе. Показатели оценки модели продемонстрировали ее достаточную точность для практического применения. В то же время сравнение с более сложными алгоритмами (например, случайным лесом, градиентным бустингом или нейронными сетями) является актуальным направлением для будущих исследований.

Одним из важных преимуществ системы является ее замкнутая архитектура. Механизм обратной связи позволяет обновлять параметры модели на основе новых данных. Это делает систему динамичной и гибкой платформой. В этом отношении разработанный подход можно рассматривать не только как инструмент мониторинга, но и как аналитическую систему, поддерживающую принятие решений.

### **Заключение**

В рамках данного исследования была разработана многоуровневая аналитическая система для мониторинга статистики преступности. Система позволяет проводить всесторонний анализ данных о преступности путем интеграции реляционной базы данных, методов статистического анализа и моделей машинного обучения.

Архитектура, построенная на концептуальной ER-модели, обеспечивает целостность данных, структурную точность и масштабируемость. Статистический модуль позволяет проводить углубленный анализ региональных и демографических показателей. Модель машинного обучения



вводит элементы прогнозной аналитики, предсказывая возраст правонарушителя.

Результаты показывают, что система может использоваться для выявления динамики преступности, оценки зон риска и эффективного распределения ресурсов. Замкнутая аналитическая архитектура делает систему гибкой и модернизируемой платформой.

В будущем целесообразно обогатить систему многомерными моделями, внедрить анализ временных рядов и расширить ее функциональные возможности путем интеграции с географическими информационными системами. В целом, разработанная система мониторинга служит для автоматизации анализа статистики преступности и улучшения процесса принятия научно обоснованных решений.

#### **Использованная литература:**

1. Dayara T. et al. Crime analyses using data analytics //International Journal of Data Warehousing and Mining (IJDWM). – 2022. – Т. 18. – №. 1. – С. 1-15.
2. Mandalapu V. et al. Crime prediction using machine learning and deep learning: A systematic review and future directions //Ieee Access. – 2023. – Т. 11. – С. 60153-60170.
3. Bogomolov A. et al. Once upon a crime: towards crime prediction from demographics and mobile data //Proceedings of the 16th international conference on multimodal interaction. – 2014. – С. 427-434.
4. Tadeu P. Exploring the Role of Statistical Analysis in Criminology from an Educational Point of View //International e-Journal of Educational Studies. – 2024. – Т. 8. – №. 18. – С. 224-233.
5. McCauley R. P., Policano T., Renshaw B. Crime analysis. – University of Louisville, School of Police Administration, 1970.