



Факультет социальных наук

Прикладная политология

Москва, 2024

Как предсказать динамику нападений? DSS для муниципалитетов

Выполнил: Олег Адамович



Цель:

Предотвращение преступлений и контроль преступной деятельности в будущем.

Предсказание роста уровня преступлений поможет муниципалитетам:

- 1) координировать действия правоохранительных органов в долгосрочной перспективе
- 2) планировать расширение / сужение штата полиции
- 3) рассчитывать выделение денег на меры, помогающие бороться с преступлениями





Проблема

Современные муниципалитеты сталкиваются с **необходимостью прогнозирования преступной активности** для более эффективного управления правоохрнительными ресурсами. Существующие модели либо недооценивают сезонные и пространственные факторы, либо требуют больших объемов данных, которые не всегда доступны.

Неясно, насколько **модели временных рядов, такие как SARIMA**, могут быть адаптированы для предсказания динамики преступности, особенно на примере конкретных видов преступлений, как, например, **нападений в Чикаго**, где число совершаемых преступлений этого вида растет на фоне снижения среди других преступлений.





Гипотезы

Общая :

- Модели временных рядов (SARIMA) могут эффективно предсказывать динамику преступности на муниципальном уровне, учитывая сезонные колебания и тренды

Частные:

- Суточная и недельная сезонность являются значимыми факторами для предсказания уровня нападений
- Добавление пространственных данных (географических меток) увеличит точность модели, позволив выявить криминогенные районы
- Общий тренд нападений связан с изменением числа сотрудников правоохранительных органов на 1000 горожан
- Модель SARIMA превосходит другие подходы по точности прогнозирования для коротких временных горизонтов (месяц, полгода)





Predictive Policing Systems

- 1) рост публичной доступности + прогресс вычислительных мощностях => **рост точности PPS** + интерес исследователей к области (*Perry 2013*)
- 2) Анализ **пространственной концентрации** преступности (*Brantingham and Brantingham 1981; Anselin et al. 2000; Weisburd, Bernasco, and Bruinsma 2008; Bernasco and Elffers 2010; Gorr and Harries 2003*)
- 3) Часть пространственных исследований фокусируются на **hotspots** (*Eck et al. 2005; Sherman, Gartin, and Buerger 1989*)
- 4) Теория **устойчивых моделей** преступности (*Brantingham and Brantingham 1984*) и теория **рутинной деятельности** (*L. E. Cohen and Felson 1979*): важная теор. основа для понимания связи между пространством и преступлениями

DSS



Prediction policing system



Crime rate prediction



Machine Learning для PPS

5) Последние годы в predictive policing systems идет активное внедрение алгоритмов машинного обучения (**ML**): в ход идут нейронные сети, случайные леса, деревья решений и др. (*Hassani et al. 2016; Shamsuddin, Ali, and Alwee 2017; Kounadi et al. 2020; Sarzaeim, 2024*)

6) **Суть:** много статистической информации нужно для обучения и проверки производительности модели. Пример – прогноз транспортных угонов в Тайване (*Lin, Yen, and Yu et al. 2018*)

7) **Плюс ML:** хорошо работает вне выборки. **Минусы ML:** успех зависит от доступности больших объемов данных + разные алгоритмы имеют разную архитектуру и подходят для разных прогнозов (*Goodfellow, Bengio, and Courville 2016*) + некоторые методы имеют трудности с количественной оценкой достоверности своих прогнозов (*Gawlikowski et al. 2021*)



Time Series как альтернатива

8) **Стат. модели из области анализа временных рядов** – надежный выбор, если рассматриваем одно пространство в течение времени: они не требуют больших объемов данных, и точность их прогнозов легче оценить (*Shumway, Stoffer, and Stoffer 2000; Brockwell and Davis 2009; Hamilton 2020*)

9) Литература имеет богатую док. базу, доказывающую сезонный или циклический характер некоторых типов преступлений (*Cohen 1941; Landau, Fridman 1993; Gorr, Olligschlaeger, Thompson 2003; Borowik, Wawrzyniak, and Cichosz 2018*); TS учитывают **фактор сезонности**, что выгодно их выделяет

10) У (*Chen, Yuan, and Shu 2008*) ARIMA показала лучшие результаты; у (*Cesario, Catlett, and Talia 2016*) прогнозы ARIMA были на 80% точными на тестовых данных; ARIMA была успешно применена и для прогнозирования уровня преступности в Северном Минданао, на Филиппинах (*Gono, Murcia, 2020*)



Теоретическое обоснование

- Модели TS обладает рядом преимуществ при прогнозировании преступности: особенно они эффективны для отражения временной динамики, в то время как модели ML отличаются гибкостью и способностью учитывать широкий спектр переменных.
- В расчете на долгосрочную перспективу мы полагаем, что следует отдать предпочтение именно TS. Мы делаем ставку на авторегрессионную интегрированную скользящую среднюю (ARIMA), зарекомендовавший себя в других исследованиях, и, в частности, на ее сезонный вариант – SARIMA.



Исследовательский вопрос

Глобальный:

Насколько хорошо можно прогнозировать уровень преступности на муниципальном уровне методом временных рядов в рамках модели SARIMA?

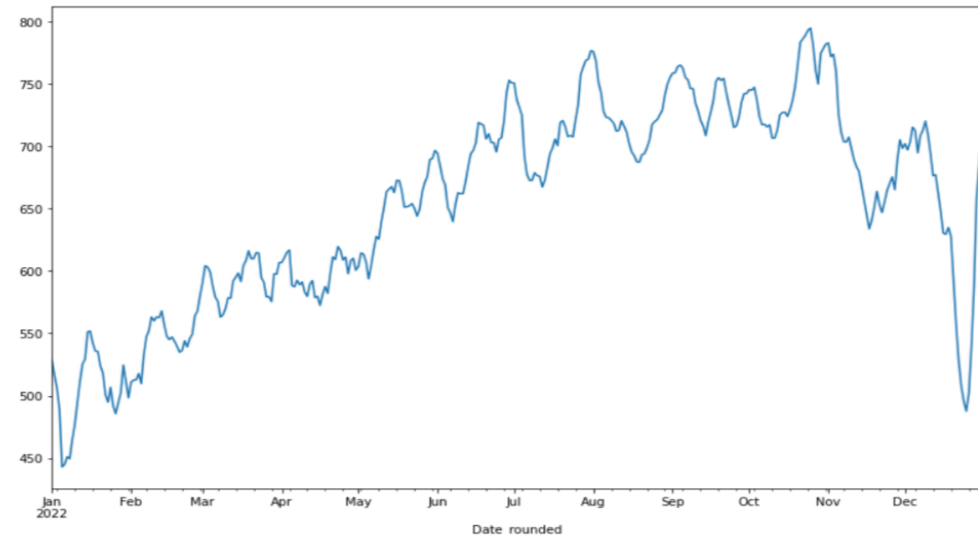
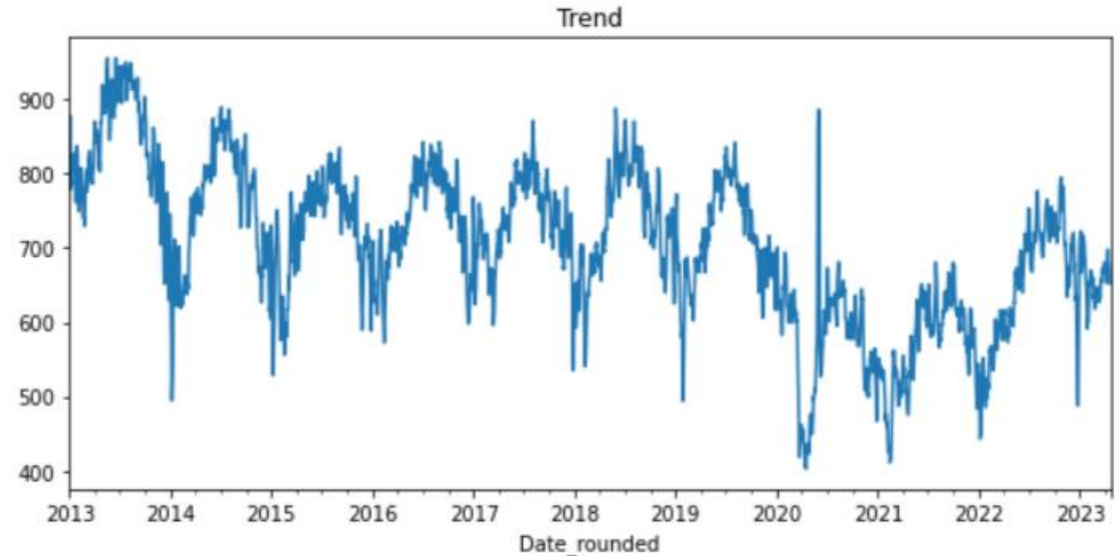
Ситуативный:

Как рост уровня нападений в Чикаго зависит от факторов места и сезонности?



Динамика преступности в Чикаго за 10 лет

- Общее снижение уровня преступности за последние 10 лет.
- Годовая сезонность.
- Пик преступности в период с августа по начало декабря.
- Всплеск преступности в период пандемии.
- Падение уровня преступности в декабре.
- Это общая статистика, мы не знаем, как меняется уровень разных типов преступлений.

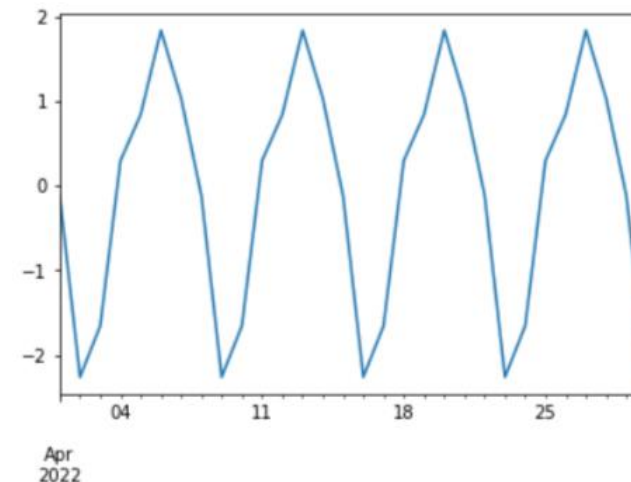
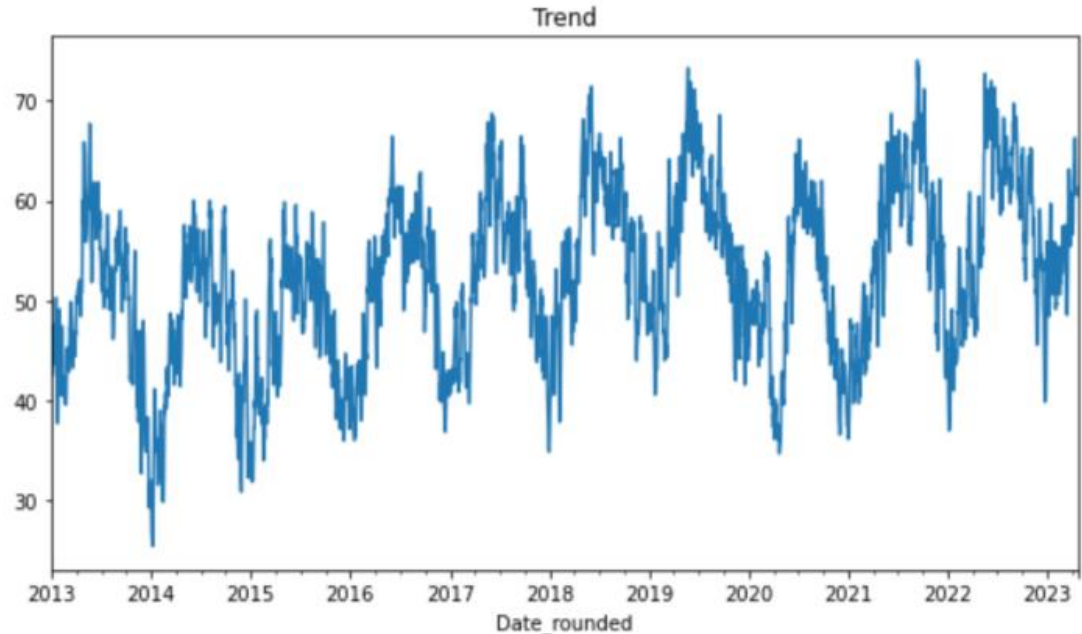




Статистика по нападениями (assault)

Assault – это угроза или попытка насилия над человеком, не приводящая к серьезным повреждениям.

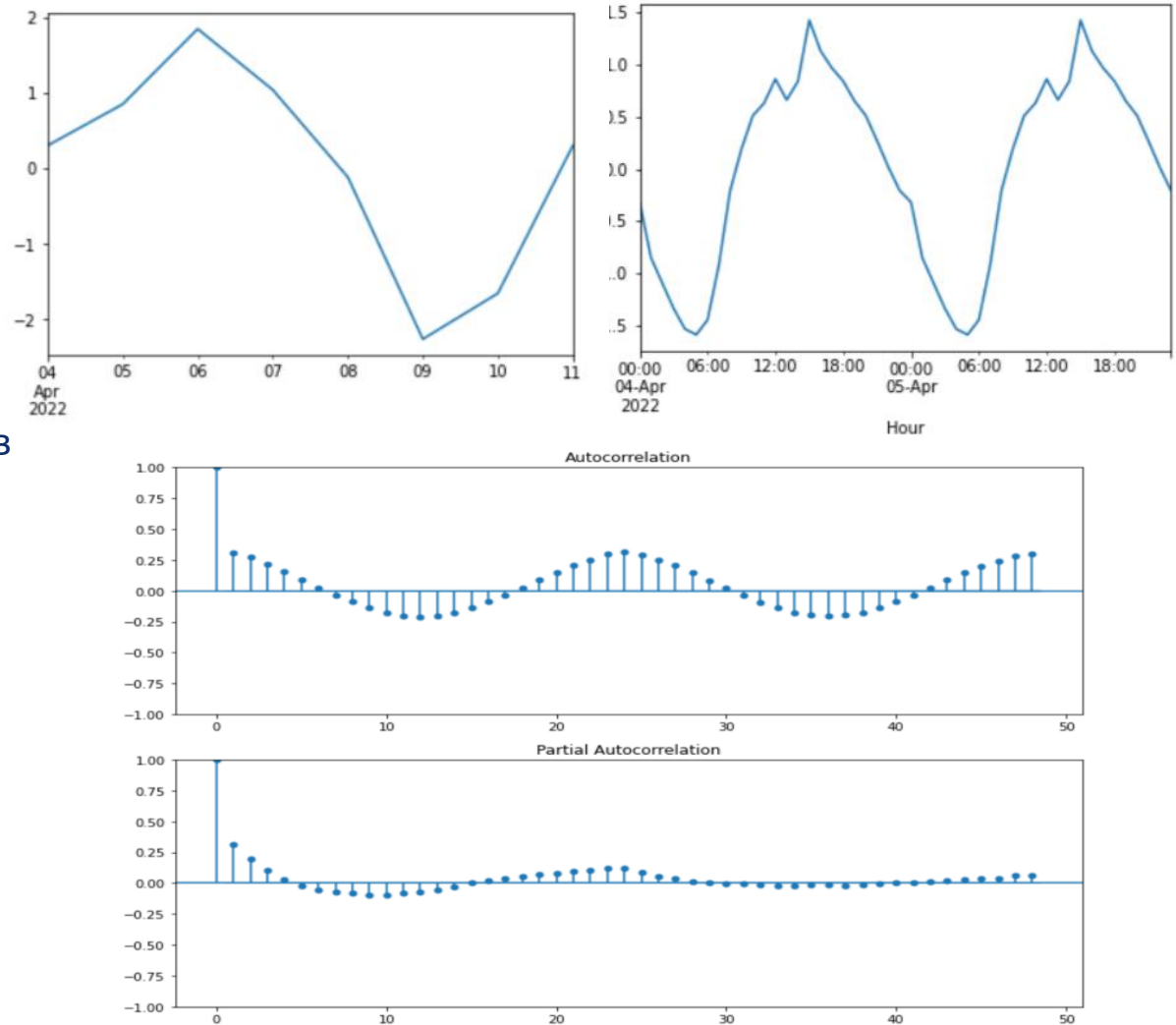
- При общем снижении преступности нападений стало больше.
- В пределах месяца видна недельная цикличность.





Статистика по нападениями (assault)

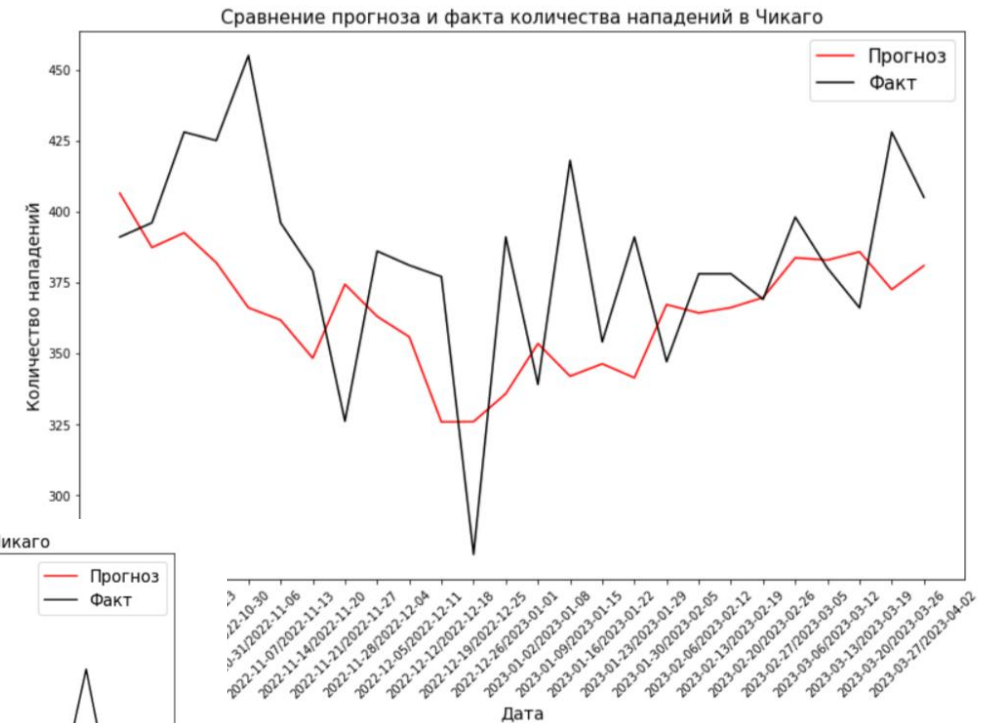
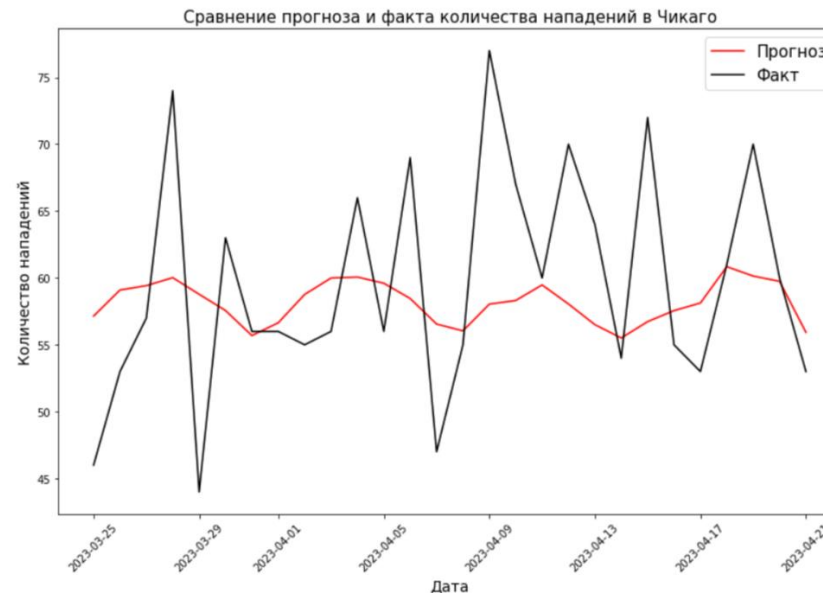
- Больше всего нападений совершается по вторникам и меньше всего по субботам.
- Пик нападений приходится на 17:00, а меньше всего их в 6 утра.
- На графике ACF видна волнообразная структура с периодом около 24 лагов, что указывает на сильную суточную сезонность.
- На первых нескольких лагах графика PACF присутствуют значимые пики, что может указывать на зависимость от недавних часов.





DSS для борьбы с нападениями (assault)

- Используем модель SARIMA
- Учитываем тренды и сезонные колебания
- Сделан прогноз на последние пол года и месяц
- Наша модель дает более сглаженные данные, так как в прогнозе использует средние данные о преступлениях за всю неделю





Какая точность моделей?

- Чем меньше горизонт прогнозирования, тем точнее модель.
- На месяц вперед прогноз делается с $MASE = 0,74$.
- На пол года вперед $MASE = 0,86$
- Значение ниже 1 указывает на то, что наша модель справляется с прогнозированием лучше, чем наивная (случайное блуждание), что можно считать хорошим результатом.

Требуют дальнейшего изучения...

- Несмотря на общее снижение уровня преступности, количество нападений из года в год растет.
- Эти тренды требуют дальнейшего изучения, чтобы лучше понимать городскую картину преступности.
- Дополнительного изучения требуют сезонность нападений – пик приходится на вторник, а минимум преступлений на субботу.



Как повысить точность и расширить функционал нашей DSS?

- Если к данным о нападениях добавить географические метки, то мы лучше поймем, какие района города наиболее проблемны.
- Анализ социального портрета преступников и конкретных мест, где происходят нападения, поможет лучше понять условия и причины роста нападений.
- Добавление в модель других временных рядов может улучшить точность прогноза на длительное время. К примеру, если будет обнаружена обратная зависимость между количеством нападения и количеством правоохранителей на тысячу горожан.

Как наша DSS поможет муниципалам?

- Понимание динамики преступности, основанное на прогнозе нашей модели, поможет властям города предсказать криминогенную обстановку в будущем.
- Власти смогут уделить больше внимания борьбе с теми видами преступлений, которые по прогнозу должны расти.
- Одновременно можно будет поощрять те меры, которые уже работают и приводят к снижению других видов преступлений.

