Министерство образования и науки Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

(национальный исследовательский университет)»

Высшая школа электроники и компьютерных наук

Кафедра «Системное программирование»

**Прогнозирование криминогенной ситуации в городе на основе технологий искусственных нейронных сетей**

|  |
| --- |
|  |
| преподаватель:  к.ф.-м.н., доцент каф СП \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_М. Л. Цымблер  «\_\_» \_\_\_\_\_\_\_\_\_ 2019 г. |
| Автор:  Аспирант  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Ф. А. Албо жуайд  «\_\_» \_\_\_\_\_\_\_\_\_ 2019 г. |
|  |

Оглавление

[Введение 4](#_Toc30775588)

[Глава 1. Обзор существующих работ 5](#_Toc30775589)

[1.1 Технологии искусственных нейронных сетей 6](#_Toc30775590)

[1.2 Использование нейронных сетей для борьбы с преступностью 9](#_Toc30775591)

[1.3 Выводы по главе 1 9](#_Toc30775592)

[Глава 2. Предварительная обработка данных 9](#_Toc30775593)

[2.1 Описание данных, используемых для прогнозирования 9](#_Toc30775594)

[2.1.1 Данные о криминогенной ситуации 9](#_Toc30775595)

[2.1.2 Урбанистические данные 9](#_Toc30775596)

[2.1.3 Данные о погоде 9](#_Toc30775597)

[2.2 Трансформация данных 9](#_Toc30775598)

[2.3 Исследование предобработанных данных 9](#_Toc30775599)

[2.3.1 Статистический анализ 9](#_Toc30775600)

[2.3.2 Поиск ассоциативных правил 9](#_Toc30775601)

[2.4 Выводы по главе 2 9](#_Toc30775602)

[Глава 3.Нейронная сеть для прогнозирования криминогенной ситуации 9](#_Toc30775603)

[3.1 Подход к прогнозированию криминогенной ситуации 9](#_Toc30775604)

[3.2 Проектирование нейронной сети 9](#_Toc30775605)

[3.3 Реализация нейронной сети 9](#_Toc30775606)

[3.4 Выводы по главе 9](#_Toc30775607)

[Глава 4. Вычислительные эксперименты 9](#_Toc30775608)

[3.5 Аппаратная платформа и данные для экспериментов 9](#_Toc30775609)

[3.6 Результаты экспериментов 9](#_Toc30775610)

[3.7 Выводы по главе 4 9](#_Toc30775611)

[Заключение 9](#_Toc30775612)

[Литература 10](#_Toc30775613)

# Введение

Преступность - это сложное социальное явление, полностью определяемое противоречиями социальных процессов, взаимодействующих друг с другом, а также криминогенными, анти-криминогенными и социальными явлениями, постоянно меняющимися во времени и пространстве. В качестве систематического этапа борьбы с преступностью, так как преступность на сегодняшний день стала одна из важнейших глобальных проблем, которые постоянно демонстрируют сложные взаимодействия с пространством, временем и средой [1]. Разрабатывая эффективные функции для обнаружения сложных ассоциаций, можно предположить, что где и когда совершаются преступления, которые сейчас становятся актуальной темой для исследователей. Географические информационные системы (ГИС) были первыми и являлись наиболее распространенными аналитическими инструментами для графического представления и анализа пространственно-распределенных данных [2]. ГИС очень полезна для картирования и ретроспективного нахождения связей между криминальными структурами и различными пространственными и социальными условиями, но сама по себе не обеспечивает большой прогностической силы.

До машинного обучения, первым методом просто экстраполировали исторические тенденции, но, очевидно, не могли предсказать отклонения от этих тенденций. Несколько более продвинутых методов, такие как многомерная регрессия, действительно улучшили предсказательную силу, но все еще полагались только на исторические данные о преступности. Существуют и другие факторы, влияющие на уровень преступности, такие как погода и социально экономические факторы [3], [4], и упомянутые методы не учитывали их. Использование дополнительных наборов, данных действительно приводит к значительному увеличению числа функций, а новые методы машинного обучения лучше подходят для обработки большого количества функций и выявления сложных взаимосвязей между этими факторами.

Машинное обучение - это наука о том, что компьютеры принимают решения без вмешательства человека. В последнее время машинное обучение применяется в автомобилях с автоматическим управлением, распознавании речи, поиске в сети и улучшенном понимании генома человека. Это также привело к возможному прогнозированию преступления на основе данных представленных в ссылке.

Классификация - это метод прогнозирования под наблюдением, который позволяет использовать номинальные метки классов. Классификация использовалась во многих областях, включая прогнозирование погоды, медицинское обслуживание, финансы и банковское дело, внутреннюю безопасность и бизнес-аналитику [5].

Анализ преступности на основе машинного обучения обычно включает сбор данных, классификацию, идентификацию паттернов, прогнозирование и визуализацию. Традиционные методы анализа данных - анализ ассоциаций, классификация и прогнозирование, кластерный анализ и анализ выбросов - позволяют идентифицировать закономерности в структурированных данных, в то время как более новые методы идентифицируют закономерности как из структурированных, так и неструктурированных данных [6]. Основной целью моей работы является создание модели прогнозирования, которая может точно прогнозировать преступность. В моём исследовании были применены: априорный алгоритм и искусственные нейронные сети(ANN) для анализа набора данных о преступности, собранных в период с 2015 по 2018 год для города Чикаго.

Диссертация состоит из 4 глав: Обзор существующих работ, Предварительная обработка данных, Нейронная сеть для прогнозирования криминогенной ситуации и вычислительные эксперименты.

# Глава 1. Обзор существующих работ

Соответствующая работа разделяется на Искусственные нейронные сети и прогнозирование преступности: Искусственные нейронные сети и прогнозирование преступности. Искусственные нейронные сети дают обзор методов работы нейронных сетей. Прогнозирование преступности представляют социологический аспект, лежащий в основе прогнозирования преступности и теорий общей преступности.

### Технологии искусственных нейронных сетей

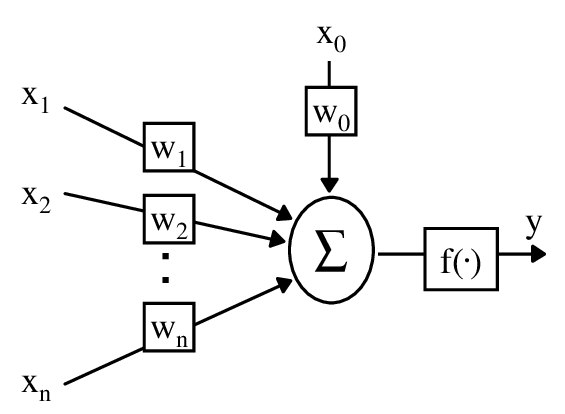
Искусственные нейронные сети (ANN) или системы соединения - это вычислительные методы, разработанные для моделирования того, как человеческий мозг выполняет задачу, используя массивно распределенную параллельную обработку, состоящую из простых блоков обработки. Эти единицы являются только вычислительными элементами, называются нейронами или узлами, которые имеют неврологическую характеристику, хранят практические знания и эмпирическую информацию, чтобы сделать их доступными для пользователя путем регулировки весов. Первый искусственный нейрон, был предложен Warren McCulloch и Walter Pitts в 1943 году. Эта модель была специально предназначена в качестве вычислительной модели "нервной сети" в мозге [7]. На рисунке (1) показана модель для нейрона.

Рис. 1. Модель искусственного нейрона

Выходной сигнал нейрона определяется по формуле (1).

(1)

Где:

Y: это функция активации;

: это передаточная функция;

: это входное значение;

: Эта значение веса.

Первоначально использовалась функция Хевисайда. Позже было предложено использовать другие типы функций активации: квадратный корень [8], логистическая сигмоидальная [9] и другие. Такие функции активации обеспечивали более плавное изменение выходного сигнала нейрона.

У человека также есть входные единицы, связанные с внешним миром, и это его пять чувств, поэтому нейронным сетям нужны входные единицы. И процессоры, в которых вычисления рассчитаны по весам и получены путем соответствующего отклика на каждый вход сети. Для этого выходные сигналы нейрона передаются на вход следующего нейрона (рис. 2). Нейронная сеть состоит из нескольких слоев, каждый из которых может содержать несколько нейронов. Слой, который получает сигналы от внешнего мира, называется входным. Слой, который посылает сигналы внешнему миру, это выходные. Остальные слои называются скрытыми.

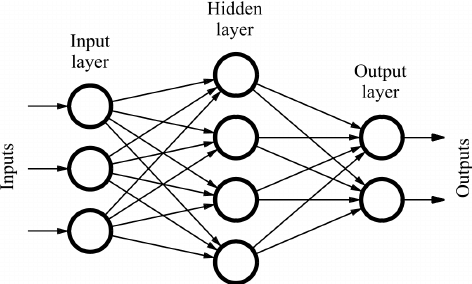
Искусственные нейронные сети делятся на сети прямого распространения сигналов (сети прямой связи), в которых нет циклов, и рекуррентные сети (рекуррентные сети), в которых разрешены циклы.

Рис. 2. Искусственная нейронная сеть

Обучение нейронной сети - это процесс определения веса связей между нейронами, чтобы сеть приближалась к необходимой функции с заданной точностью. Существует три подхода к обучению нейронных сетей [10]: Обучение с учителем, Обучение без учителя и Смешанное обучение.

При Обучение с учителем предоставляются как входы, так и выходы. Затем сеть обрабатывает входные данные и сравнивает полученные результаты с требуемыми выходными данными. Ошибки затем распространяются обратно через систему, заставляя систему регулировать веса, которые контролируют сеть. Этот процесс происходит снова и снова, так как вес постоянно изменяется. Набор данных, который позволяет проводить обучение, называется «обучающим набором». Во время обучения сети один и тот же набор данных обрабатывается много раз, поскольку вес соединения изменяется.

Второй тип обучения называется Обучение без учителя. При Обучение без учителя сеть снабжается входами, но не желаемыми выходами. Затем система сама должна решить, какие функции она будет использовать для группировки входных данных. Это часто называют самоорганизацией или адаптацией.

Третий тип обучения называется «Смешанное обучение» предполагает наличие внешней среды, с которой взаимодействует сеть. Обучение проходит на основе сигналов, полученных из этой среды.

### Использование нейронных сетей для борьбы с преступностью

Надлежащее предсказание преступлений является основной проблемой, которая была принята за основу концентрации в истории, нынешние события в рамках предоставили новую возможность подходов. Методология искусственных нейронных сетей (ANN) является одной из самых важных новых тем в моделировании и анализе. Оценка данных и управление прогнозированием без ссылки на общую статистическую модель или метод. Предварительная диагностика поведения явления, при котором выполняется моделирование, для достижения оптимального и точного представления данных, представляющих явление, близкое к реальности. Может использоваться большую часть времени и ситуации.

В исследовании, проведенном Хуангом и другие , исследователи использовали структуру глубокой преступности, чтобы исследовать уровень предсказаний для совершения преступления. Исследование было начато с разработки подхода с кодировщиком зависимостей, который позволил исследователям предсказать многочисленные виды преступлений, а затем выявить их связь с другими распространенными данными. Исследование было в основном сосредоточено на определении времени и места прогнозирования преступности. В исследовании проводились широкомасштабные эксперименты, основанные на реальных наборах данных, собранных на портале открытых данных New York City (NYC) с (1 января 2014 года по 31 декабря 2014 года). Каждая запись о преступлении имеет формат (широта, долгота, категория преступления, отметка времени). И из данных, использованных Точкой интересов (POI), они собрали 24 031 POI 14 категорий (Искусство, Развлечения и Покупки). и 311 Данные о жалобах на коммунальные услуги: они выбрали 4 ключевые категории жалоб (незаконная парковка, шум, заблокированная дорога и использование здания), которые изучаются в [10]. Анализ результатов показал, что основа глубокой преступности, в основном, фокусируется на точных прогнозах для множества параметров временных интервалов региона.

### Выводы по главе 1

# Глава 2. Предварительная обработка данных

### Описание данных, используемых для прогнозирования

### Данные о криминогенной ситуации

### Урбанистические данные

### Данные о погоде

### Трансформация данных

### Исследование предобработанных данных

### Статистический анализ

### Поиск ассоциативных правил

### Выводы по главе 2

# Глава 3. Нейронная сеть для прогнозирования криминогенной ситуации

### Подход к прогнозированию криминогенной ситуации

### Проектирование нейронной сети

### Реализация нейронной сети

### Выводы по главе

# Глава 4. Вычислительные эксперименты

### Аппаратная платформа и данные для экспериментов

### Результаты экспериментов

### Выводы по главе 4

# Заключение

# Литература

[1] L. Duan, T. Hu, E. Cheng, J. Zhu, and C. Gao, “Deep Convolutional Neural Networks for Spatiotemporal Crime Prediction,” in *In 2017 International Conference on Information and Knowledge Engineering (IKE)*, 2017, pp. 61–67.

[2] E. R. Groff and N. G. La Vigne, “Forecasting the future of predictive crime mapping,” *Crime Prev. Stud.*, vol. 13, pp. 29–57, 2002.

[3] P. Davies, “Women, Crime and an Informal Economy: Female Offending and Crime for Gain,” *Br. Criminol. Conf. Sel. Preceedings*, vol. 2, no. March, pp. 15–19, 1999.

[4] E. G. Cohn, “Weather and crime,” *Br. J. Criminol.*, vol. 30, no. 1, pp. 51–64, 1990.

[5] R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. S. Panahy, and N. Khanahmadliravi, “An experimental study of classification algorithms for crime prediction,” *Indian J. Sci. Technol.*, 2013.

[6] H. Chen, W. Chung, J. J. Xu, G. Wang, Y. Qin, and M. Chau, “Crime data mining: A general framework and some examples,” *Computer (Long. Beach. Calif).*, 2004.

[7] W. S. McCulloch and W. H. Pitts, “originally published in: Bulletin of Mathematical Biophysics, Vol. 5, 1943, p. 115-133,” *Bull. Math. Biophys.*, vol. 5, pp. 115–133, 1943.

[8] B. Carlile, G. Delamarter, P. Kinney, A. Marti, and B. Whitney, “Improving Deep Learning by Inverse Square Root Linear Units (ISRLUs),” vol. 0, no. 1, pp. 1–8, 2017.

[9] D. E. Rumelhart, G. E. Hinton, and J. L. McClelland, “A General framework for Parallel Distributed Processing,” in *Parallel distributed processing: explorations in the microstructure of cognition*, 1986, pp. 45–76.

[10] D. Ravi *et al.*, “Deep Learning for Health Informatics,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 1, pp. 4–21, 2017.