

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ

Магистратура ЭФ

Отчет по учебной практике, научно-исследовательской работе (получению первичных навыков научно-исследовательской работы), научному семинару в 3-м семестре студента

22712 группы

Хорунженко Аркадий Сергеевич

(ФИО)

1. Магистерская программа “Цифровые технологии и искусственный интеллект в финансах и бизнес-аналитике”
2. Тема диссертации “Исследование эффективности нейронных сетей для прогнозирования финансовых временных рядов”
3. Научный руководитель Макушев Василий Леонидович

Дата

Подпись научного руководителя

Подпись студента

## ПЛАН РАБОТЫ

Введение .....	3
1. Применение модели ARMA для прогнозирования временных рядов.....	4
2. Применение модели ARIMA для прогнозирования временных рядов .....	6
Заключение .....	9
Список использованных источников .....	10

## ВВЕДЕНИЕ

Выбранной темой магистерской диссертации является «Исследование эффективности применения нейронных сетей для прогнозирования временных рядов».

Прогнозирование временных рядов является одним из важнейших современных инструментов для исследования и анализа во многих областях науки и практики. В разных сферах исследований применяются различные модели для построения прогнозов. Анализ моделей временных рядов является, таким образом, важнейшей частью построения прогноза. Принципиальное различие в структурах пространственных и временных выборок требует рассмотрения для них отдельных методов и подходов к оцениванию. При построении эконометрических регрессионных моделей для временных рядов в различных задачах возникает ряд особенностей, которые необходимо учесть.

При обработке информации о поведении финансовых временных рядов необходимо учитывать, что методы анализа нестационарных случайных процессов существенно отличаются от приемов работы со стационарными случайными временными рядами. Однако в рамках системы фондового рынка существует множество приемов, описывающих его динамику, которые обладают так называемой однородной нестационарностью и могут быть описаны при помощи подходов, применимых к стационарным рядам. К числу таких методов относится применение линейных стохастических моделей авторегрессии и проинтегрированного скользящего среднего (AutoRegressive Integrated Moving Average, ARIMA) и модель авторегрессии и скользящего среднего (ARMA).

Для построения этих моделей, как правило, достаточно использовать информацию, содержащуюся в самих анализируемых данных временного ряда.

В данном реферате рассматриваются несколько статей, описывающих особенности реализации прогноза неизвестных параметров временного ряда с помощью этих моделей.

# **1 ПРИМЕНЕНИЕ МОДЕЛИ ARMA ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ**

Модель авторегрессионного скользящего среднего (ARMA) является одной из самых общих моделей для прогнозирования временных рядов. Она может объяснить взаимосвязь временного ряда как со случайным шумом (часть скользящего среднего), так и с самим собой на предыдущем шаге (часть авторегрессии). Исследованию границ применимости данной модели посвящено множество статей российских и зарубежных авторов.

В данном разделе работы рассмотрим статью «Моделирование и прогнозирование временного ряда суммой логистической, линейной и гармонической компонент на основе ARMA-модели», авторов Семенычева В.К., Семенычева В.В. и Павлова В.Д., опубликованную в 2009 году. [4]

В этой статье рассмотрен новый подход к моделированию и прогнозированию временных рядов суммой логистической, линейной и гармонической компонент на основе параметрической модели авторегрессии-скользящего среднего. В качестве основного логистического тренда использовалась логиста Рамсея, по своему виду близкая к функции Верхулста, но при этом имеющая более простое аналитическое выражение.

Авторами был предложен метод моделирования и прогнозирования на сравнительно коротких выборках временных рядов, содержащих сумму логисты Рамсея, линейный тренд, гармонику и стохастическую компоненту.

В статье были проведены исследования точности рассматриваемой модели, как на тестовых выборках, так и на реальных статистических данных.

Исследования на тестовых выборках показали высокую точность моделирования и прогнозирования в широком диапазоне значений параметров модели, соотношения «шум/сигнал».

Тестирование на реальных данных подтвердили высокую точность и возможность использования метода для различных экономических процессов и явлений.

Полученные авторами результаты показали высокую точность моделирования и прогнозирования в широком диапазоне значений параметров модели, соотношения «шум/сигнал», а тестирование на реальных данных показало возможность использования предложенной модели и метода ее параметризации для различных экономических процессов.

Еще одна интересная статья, касающаяся использования модели ARMA при прогнозировании финансовых временных рядов – «Применение динамической модели ARMA-GARCH для прогнозирования динамики курса акций» авторов Шемякиной М.А. и Клейменкина Д.В., опубликованная в электронном научном журнале «Дневник науки» в 2021 году. [6]

Данная статья посвящена исследованию эффективности прогнозирования динамики курса акций ПАО «Сбербанк» (SBER) за период с января 2005 года по март 2021 года, для которых были проанализированы различные комбинации моделей ARMA и GARCH. Было выявлено, что распределение финансовых временных рядов отличается от нормального.

Для реализации и тестирования динамической модели ARMA-GARCH был выбран язык программирования Python. В качестве оценки точности прогнозирования доходности были использованы три показателя: средняя абсолютная ошибка (MAE); среднеквадратичная ошибка (RMSE); средняя абсолютная масштабированная ошибка (MASE). В результате моделирования авторами был сделан вывод об отсутствии какой-либо закономерности в автокорреляциях акций для различных лагов. Для прогнозирования доходности были рассмотрены комбинации моделей ARMA и GARCH. В результате моделирования показатель точности MASE показал, что динамические модели ARMA-GARCH являются достаточно эффективными и подходящими моделями для изучения рядов доходности.

Таким образом, на основе ошибки прогнозирования MASE можно сделать вывод о том, что модели ARMA-GARCH дают более точный результат по сравнению с наивным методом предсказания.

## **2 ПРИМЕНЕНИЕ МОДЕЛИ ARIMA ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ**

Модель ARIMA означает интегрированное скользящее среднее с авторегрессией. Эта модель предоставляет набор функций, которые являются очень мощными и гибкими для выполнения любых задач, связанных с прогнозированием временных рядов. Вопросы ограничений при применении модели АРИМА для прогнозирования временных рядов рассматриваются многими авторами в различных контекстах.

Так, например, данная модель успешно применялась в исследованиях, посвященных прогнозированию динамики развития эпидемии COVID-19. В статье Макаровских Т.А., Аботалев М.С.А. «Автоматический подбор параметров модели ARIMA для прогноза количества случаев заражения и смерти от Covid-19», опубликованной в «Вестнике ЮУрГУ» в 2021 году [3]. В этой работе исследовано применение модели ARIMA прогнозирования временных рядов для анализа открытых данных о распространении коронавирусной инфекции в ряде регионов Российской Федерации. В частности, рассмотрена возможность применения существующих методов и алгоритмов языка программирования Python, приводятся алгоритмы подбора параметров модели ARIMA.

В статье наглядно продемонстрировано, что параметры модели различны для временных рядов разной длины, кроме того, параметры модели могут меняться с течением времени. Также показано, что существуют наборы данных, для которых нет возможности получить параметры модели, дающие наименьшую погрешность. Исследована частота переобучения модели, приведены данные об изменении параметров модели для временных рядов разной длины.

В практической части работы проведено сравнение прогнозов, полученных в конце октября 2020 г. и актуальных данных на середину ноября 2020 г. Данное сравнение позволило сделать вывод, что полученный прогноз позволяет

достаточно точно предсказать суммарное число заражений и летальных исходов на 7–10 дней

Модель ARIMA также успешно используется, например, при прогнозировании сетевого трафика. Об этом речь идет в статье Крюкова Ю.А. и Чернягина Д.В. «ARIMA – модель прогнозирования значений трафика», опубликованной в журнале «Информационные технологии и вычислительные системы» в 2011 году. [2]

Прогнозирование сетевого трафика представляет значимый интерес в таких областях как отслеживание перегрузок в сети, контроль потоков данных и сетевое управление. Тщательно подобранная модель трафика способна выявить и предсказать важнейшие характеристики сетевого трафика: кратковременно и долговременно зависимые процессы, самоподобность на больших временных масштабах. В рассматриваемой статье представлена модель ARIMA с минимальным числом параметров, имеющая адекватный прогноз. Предложена процедура оценки параметров модели ARIMA и выбора модели с минимальным числом параметров. Приведены сравнения оценок качества прогноза для полученных моделей.

В практической части данной статьи рассмотрен алгоритм прогнозной части системы динамического управления трафиком, которая позволит своевременно предупреждать о последующих значительных выбросах сетевого трафика. Как важный ее компонент, найдена модель ARIMA с минимальным числом параметров, позволяющая совершать достоверные краткосрочные прогнозы. Показано, что прогноз более двух шагов вперед хорошо осуществляется с помощью модели  $ARIMA(1,1,0)$ .

Таким образом, можно сделать вывод, что модель ARIMA имеет широкое распространение при решении задач прогнозирования временных рядов в различных отраслях науки. В связи с чем наработан широкий инструментарий для ее применения. Тем не менее, можно отметить, что существуют определенные сложности для применения этой модели в экономических исследованиях.

Принципиальное различие в структурах пространственных и временных выборок требует рассмотрения для них отдельных методов и подходов к оцениванию. При построении эконометрических регрессионных моделей для временных рядов возникает ряд особенностей, которые необходимо учесть:

- упорядоченность во времени (хронологический порядок);
- зависимость от прошлого (“память”, серийная или авто- корреляция);
- различаются краткосрочные и долгосрочные зависимости и модели;
- часто встречается феномен “ложной регрессии”;
- возможны случаи, когда имеется лишь небольшое число наблюдений, которое невозможно увеличить (т.к. изменяется вид или структура зависимости).

Тем не менее, всеми авторами упоминается перспективность данной модели при построении прогноза временного ряда.



## ЗАКЛЮЧЕНИЕ

В результате проведенного изучения литературных источников можно сделать следующие выводы.

Модели, основанные на анализе временных рядов, в частности ARIMA-модели и ARMA-модели, являются сложно настраиваемыми при проведении полного анализа, однако дают практически всегда хороший результат там, где требуется качественный прогноз на среднесрочный и краткосрочный период. В рассмотренных в данном реферате работах, как и во многих других, авторы вручную подбирают параметры для имеющихся на момент публикации временных рядов и проводят расчеты в широко используемых для статистического анализа программах для небольших временных промежутков в рамках решения узкоспециализированных задач.

Основным недостатком большинства представленных работ является отсутствие объяснения выбора соответствующей спецификации моделей, а также отсутствие «объяснения» подбора гиперпараметров моделей прогнозирования.

Таким образом, получить результаты высокой точности можно только используя динамическое оценивание изменяющихся коэффициентов. Поэтому решение данной задачи является актуальным. Проведенный анализ показывает, что большинство работ по прогнозированию временных рядов в разных отраслях науки посвящено исследованию ручного подбора параметров модели ARIMA для данных по одному или нескольким фиксированным регионам; опубликовано очень мало работ, в которых предлагается использовать современные программные средства для прогнозирования временных рядов и в которых описывались бы библиотеки для анализа финансовых временных рядов (в частности, касающихся котировок ценных бумаг).

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Кратович П.В. Нейронные сети и модели ARIMA для прогнозирования котировок // Программные продукты и системы. 2012. № 1(89). С. 132–134.
2. Крюков Ю.А., Чернягин Д.В. ARIMA – модель прогнозирования значений трафика // Информационные технологии и вычислительные системы. 2011. №2.
3. Макаровских Т.А., Аботалёб М.С.А. Автоматический подбор параметров модели ARIMA для прогноза количества случаев заражения и смерти от Covid-19 // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2021. Т. 10, № 2. С. 20–37. DOI: 10.14529/cmse210202
4. Семёнычев В.К., Павлов В. Д., Семёнычев В. В. Моделирование и прогнозирование временного ряда суммой логистической, линейной и гармонической компонент на основе ARMA-модели // Journal of new economy. 2009. №1 (23). URL: <https://cyberleninka.ru/article/n/modelirovanie-i-prognozirovanie-vremennogo-ryada-summoy-logisticheskoy-lineynoy-i-garmonicheskoy-komponent-na-osnove-arma-modeli> (дата обращения: 22.11.2023).
5. Трегуб А. В., Трегуб И. В. Методика построения модели ARIMA для прогнозирования динамики временных рядов // Вестник МГУЛ – Лесной вестник. 2011. №5. URL: <https://cyberleninka.ru/article/n/metodika-postroeniya-modeli-arima-dlya-prognozirovaniya-dinamiki-vremennyh-ryadov> (дата обращения: 01.12.2023).
6. Шемякина М.А., Клейменкин Д.В. Применение динамической модели ARMA-GARCH для прогнозирования динамики курса акций // электронный научный журнал «Дневник науки». 2021. №7. – URL: [www.dnevniknauki.ru](http://www.dnevniknauki.ru) (дата обращения: 22.11.2023).