

## Программа учебной дисциплины «Прогнозирование временных данных»

Утверждена  
Академическим советом ООП  
Протокол № от «27» июня 2018 г.

Автор	Соколовский Е.И, esokolovskii@yandex.ru
Число кредитов	5
Контактная работа (час.)	68
Самостоятельная работа (час.)	122
Курс	1
Формат изучения дисциплины	Без использования онлайн курса

### I. ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ

Целями освоения дисциплины «Прогнозирование временных данных» являются овладение студентами основными концепциями и навыками работы с данными, имеющими структуру временных рядов.

В результате освоения дисциплины студент должен:

#### **знать:**

- основные понятия разделов машинного обучения и прикладной статистики, направленные на моделирование временных рядов;
- основные методы и алгоритмы моделирования, прогнозирования, интерполяции, сглаживания, сегментации одномерных и многомерных временных рядов и их характеристик;
- основные методы построения симуляционных моделей.

#### **уметь:**

- применять основные понятия прикладной статистики для постановки задач обработки временных рядов;
- применять основные методы и алгоритмы обработки временных рядов;
- оценивать степень достоверности результатов, полученных с помощью основных методов обработки временных рядов;
- работать с программным обеспечением и библиотеками для обработки временных рядов с применением широко используемых языков программирования.

#### **владеть:**

- навыками использования подходов машинного обучения и статистического моделирования в важнейших практических приложениях;
- навыками обработки и интерпретации результатов эксперимента;
- методами теоретического исследования основных математических (вероятностно-статистических) моделей временных рядов.

Изучение дисциплины «Прогнозирование временных данных» базируется на следующих дисциплинах:

- «Вероятность», «Математическая статистика», «Финансовая математика» и «Случайные процессы» в объеме бакалавриата;
- «Машинное обучение» в объеме бакалавриата; допускается параллельное прослушивание курса по машинному обучению и настоящей дисциплины.

Для освоения учебной дисциплины студенты должны владеть следующими знаниями и компетенциями:

- знать основные понятия теории вероятностей, математической статистики, случайных процессов, машинного обучения;
- знать простейшие методы решения задач по теории вероятностей, математической статистики, машинного обучения;
- обладать навыками работы с языком программирования python.

## II. СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

### 1. Введение в теорию случайных процессов (2 лекции, 2 семинара)

Случайный процесс, его основные характеристики. Непрерывность случайного процесса в среднем квадратичном, по вероятности и с вероятностью единица. Процессы с независимыми приращениями. Пуассоновский и винеровский процессы. Гауссовские случайные векторы, их линейные преобразования. Стационарный случайный процесс. Эргодичность случайного процесса. Спектральное представление случайного процесса.

### 2. Марковские случайные процессы (1 лекция, 1 семинар)

Марковский случайный процесс и дискретная марковская цепь. Асимптотическое поведение дискретной марковской цепи.

### 3. Некоторые вопросы статистического анализа финансовых данных (1 лекция, 1 семинар)

Статистика “тиков” эмпирических данных. Статистическое описание логарифмов относительных изменений цен: “вытянутость”, “тяжелые хвосты”. Статистика волатильности, корреляционная структура, эффекты “кластерности” в ценах. Фрактальность статистических данных.

### 4. Элементарные модели и простейшие алгоритмы обработки временных рядов (1 лекция, 1 семинар)

Обработка сигналов: очистка от шума, сглаживание с адаптивными весами, выделение тренда и оценка дисперсии. Экспоненциальное сглаживание временных рядов: оценки среднего значения и параметра масштаба, случай неравномерного расположения наблюдений, метод Хольта-Винтерса, робастный вариант экспоненциального сглаживания.

### 5. Стохастические авторегрессионные и условно-гауссовские модели динамики рыночных цен (1 лекция, 1 семинар)

Моделирование временных рядов на основе моделей типа ARMA. Оценивание коэффициентов модели. Уравнения Юла-Уолкера. Прогнозирование в авторегрессионных моделях. Интегральная модель ARIMA. Условно-гауссовские модели ARCH, GARCH. Критерии выбора порядка модели.

### 6. Сегментация временных данных на основе скрытых марковских моделей (1 лекция, 1 семинар).

Марковские модели цен с переключением режимов. Скрытая марковская модель. Максимизация функции правдоподобия скрытой марковской модели. Алгоритм Витерби.

7. Линейная фильтрация в экономике и финансовой математике. Фильтр Калмана (1 лекция, 1 семинар).

Модель пространства состояний. Уравнения Калмановской фильтрации в дискретном времени. Вопросы применения уравнений линейной фильтрации в задачах оценки динамических характеристик финансовых инструментов.

8. Основные задачи теории принятия решений: ретроспективная и последовательная постановки (2 лекции, 2 семинара)

Фундаментальная лемма Неймана-Пирсона. Последовательные статистические тесты. Последовательный критерий отношения правдоподобия. Вальдовская схема наблюдений и принятия решения.

9. Задачи обнаружения разладок временных рядов (2 лекции, 2 семинара)

Задача «о разладке». Некоторые широко используемые статистики в проблеме скорейшего обнаружения разладки. Статистика “контрольных карт” Шухарта. Статистика кумулятивных сумм. Статистика Ширяева-Робертса. Модификация стандартных статистик для выявления разладки в дисперсии и ковариационной структуре.

10. Прогнозирование временных рядов на основе рекуррентных нейронных сетей (1 лекция, 1 семинар)

Нейронные сети как модели последовательностей. Рекуррентная модель RNN и модель долгой краткосрочной памяти LSTM. Алгоритм обратного распространения ошибки сквозь время. Моделирование и прогнозирование многокомпонентных (с трендом, периодических) временных рядов.

11. Моделирование и прогнозирование волатильности на финансовом рынке (1 лекция, 1 семинар)

Дискретные модели стохастической волатильности. Оценивание волатильности финансовых инструментов с помощью подходов линейной фильтрации.

12. Моделирование и прогнозирование процессов с длинной памятью на финансовом рынке (1 лекция, 1 семинар)

Эффект последствия в ценах. Фрактальное броуновское движение и фрактальный гауссовский шум как модели сильного последствия. Моделирование и оценивание параметров временных рядов с длинной памятью. Статистический R/S анализ временных рядов.

13. Модель Блека-Шоулза (1 лекция, 1 семинар)

Модели стоимости акций в непрерывном времени. Линейная модель Л. Башелье. Диффузионная (B,S) модель. Рациональная стоимость колл-опционов европейского типа.

### III. ОЦЕНИВАНИЕ

Результирующая оценка по дисциплине рассчитывается по формуле:

$$O_{итог} = 0.8 O_{накопл} + 0.2 O_{экз};$$

Накопленная и итоговая оценки округляются арифметически.

Накопленная оценка рассчитывается по формуле:

$$O_{\text{накопл}} = 0.6 O_{\text{проекты}} + 0.2 O_{\text{дз}} + 0.1 O_{\text{коллоквиум}} + 0.1 O_{\text{тесты}}$$

Оценка за проекты выставляется по результатам выполнения 3х командных проектов курса. Оценка за домашнее задание рассчитывается как среднее арифметическое всех оценок за домашние работы. Оценка за промежуточные тесты рассчитывается как среднее арифметическое всех оценок за промежуточные тесты.

Студенту, получившему отличную накопленную оценку, данная оценка может быть выставлена в качестве итоговой при условии, что им набрано не менее 10 дополнительных баллов. Дополнительные баллы выставляются за выполнение дополнительных частей домашних заданий, имеющих повышенную сложность и не участвующих при выставлении накопленной оценки.

#### **IV. ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ**

Оценочные средства для текущего контроля студента: домашние работы, тесты.

Оценочные средства для промежуточной аттестации: коллоквиум.

Примеры заданий промежуточной аттестации:

1. Основные понятия теории случайных процессов. Основные классы случайных процессов. Гауссовские случайные векторы. Непрерывность случайного процесса. Характеристики случайных процессов.
2. Марковские случайные процессы, дискретные Марковские цепи.
3. Модель авторегрессии–скользящего среднего.
4. Экспоненциально взвешенное скользящее среднее. Метод Хольта-Винтерса.
5. Скрытые Марковские модели. Алгоритм Витерби.
6. Фильтр Калмана.
7. Лемма Неймана-Пирсона. Вальдовская и Нейман-Пирсоновская схемы принятия решения.
8. Основные статистики в задаче обнаружения разладок временных рядов.
9. Основные методы обнаружения аномалий временных рядов.
10. Условно-гауссовские модели ARCH, GARCH в финансовой математике.

#### **V. РЕСУРСЫ**

##### **5.1 Основная литература**

1. Ширяев А. Н. Вероятностно-статистические методы в теории принятия решений. – М.: МЦНМО, 2015.

2. Hyndman, R.J., Athanasopoulos G. Forecasting: principles and practice. — OTexts, 2016.  
<https://www.otexts.org/book/fpp>

## 5.2 Дополнительная литература

1. Gustafsson F. Adaptive filtering and change detection. – New York : Wiley, 2000.  
<https://eden.dei.uc.pt/~tbohnert/math/ap-cp.pdf>

## 5.1 Программное обеспечение

№ п/п	Наименование	Условия доступа
1.	Anaconda	<i>Свободно распространяемое ПО</i>

## 5.2 Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

№ п/п	Наименование	Условия доступа
	<i>Интернет-ресурсы (электронные образовательные ресурсы)</i>	
1.	Открытое образование	URL: <a href="https://openedu.ru/">https://openedu.ru/</a>
2.	Coursera	URL: <a href="https://www.coursera.org">https://www.coursera.org</a>

## 5.3 Материально-техническое обеспечение дисциплины

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

- ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
- мультимедийный проектор с дистанционным управлением.

Учебные аудитории для лабораторных и самостоятельных занятий по дисциплине оснащены ПЭВМ (операционная система, офисные программы), с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.