

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

Рабочая программа дисциплины (модуля)

по дисциплине:	Анализ временных рядов
по направлению:	Прикладная математика и информатика
профиль подготовки:	AI360: Передовые методы искусственного интеллекта Физтех-школа Прикладной Математики и Информатики кафедра машинного обучения и цифровой гуманитаристики
курс:	3
квалификация:	бакалавр

Семестр, формы промежуточной аттестации: 5 (осенний) - Дифференцированный зачет

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 30 час.

Всего часов: 90, всего зач. ед.: 2

Программу составил: Р.Г. Нейчев, старший преподаватель

Программа обсуждена на заседании кафедры машинного обучения и цифровой гуманитаристики 12.02.2024

Аннотация

Курс посвящен анализу временных рядов, уделяя основное внимание методам моделирования и прогнозирования с помощью как статистических, так и основанных на машинном обучении методов. Освещаются основные типы моделей временных рядов, такие как AR, MA, ARMA и ARIMA, а также более сложные подходы, включая векторные авторегрессии (VAR). Отдельное внимание уделяется подходам на основе методов глубокого обучения. Курс также охватывает методы проверки стационарности, различные техники сглаживания и декомпозиции временных рядов. Особое внимание уделяется практическому применению рассматриваемых методов для создания эффективных прогностических моделей. Курс предполагает наличие базовых знаний в области статистики и программирования.

1. Цели и задачи

Цель дисциплины

- познакомить с основными свойствами временных рядов и случайных процессов;
- предоставить практический опыт использования статистических методов для анализа временных рядов;
- развить навыки применения современных подходов для прогнозирования временных рядов.

Задачи дисциплины

- научить студентов оценивать свойства временных рядов (стационарность, периодичность и пр.);
- предоставить набор методов для прогнозирования временных рядов с учетом их свойств;
- обеспечить существенный опыт работы с временными рядами.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-1 Способен применять фундаментальные знания, полученные в области физико-математических и (или) естественных наук и использовать их в профессиональной деятельности	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.2 Способен строить математические модели, производить количественные расчеты и оценки
	ОПК-1.3 Способен определять границы применимости полученных результатов
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.1 Способен применять современные вычислительную технику и сервисы сети Интернет в области (сфере) профессиональной деятельности
	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
	ОПК-2.3 Знает основные требования информационной безопасности

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- постановку задач классификации и регрессии на основе временных данных;
- методы решения этих задач.

уметь:

- формулировать подходящие критерии качества для указанных выше задач;
- реализовывать решение на основе статистических или основанных на машинном обучении методов с учетом ограничений на сложность.

владеть:

- основными программными системами для работы с временными рядами вместе с другими инструментами анализа данных.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Основы временных рядов и стационарность.	6	6		6
2	Прогнозирование и сглаживание.	6	6		6
3	Расширенные модели и интервальные прогнозы.	6	6		6
4	Машинное обучение и глубокое обучение для временных рядов.	6	6		6
5	Трансформеры и обнаружение смены режима.	6	6		6
Итого часов		30	30		30
Подготовка к экзамену		0 час.			
Общая трудоёмкость		90 час., 2 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 5 (Осенний)

1. Основы временных рядов и стационарность.

Вводные понятия временных рядов, включая авторегрессионные (AR) и скользящие средние (MA) процессы. Обсуждение функции автокорреляции (ACF) и стационарности, включая тесты на стационарность, такие как расширенный тест Дики-Фуллера.

2. Прогнозирование и сглаживание.

Различные методы прогнозирования, включая ARIMA и его оценки, а также методы сглаживания, как MA сглаживание и экспоненциальное сглаживание. Включает в себя изучение декомпозиции временных рядов, такой как STL, и сравнение ETS и ARIMA.

3. Расширенные модели и интервальные прогнозы.

Обзор предиктивных интервалов и более сложных моделей, таких как ARCH, VAR и нелинейная авторегрессия. Рассмотрение моделирования структурных изменений и сложных сезонностей с использованием таких инструментов, как Prophet, Neural Pro.

4. Машинное обучение и глубокое обучение для временных рядов.

Применение машинного обучения, включая извлечение признаков, кросс-валидацию и представление обучения специально для временных рядов. Введение в глубокие обучающие модели, такие как RNN, LSTM и CNN, с демонстрациями на PyTorch и других платформах.

5. Трансформеры и обнаружение смены режима.

Подробное изучение трансформеров, включая способы повышения их эффективности, их применение к временным рядам. Также рассматриваются методы обнаружения изменений в распределении данных для временных рядов.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6.Перечень рекомендуемой литературы

Основная литература

Литература на кафедре:

1. Машинное обучение [Текст]/Х. Бринк, Дж. Ричардс, М. Феверолф, Real-World Machine Learning, -СПб., Питер, 2017
2. Python и машинное обучение [Текст] = Python Machine Learning : крайне необходимое издание по новейшей предсказательной аналитике для более глубокого понимания методологии машинного обучения / С. Рашка; пер. с англ. А. В. Логунова .— М. : ДМК Пресс, 2017 .— 418 с.: ил. - Предм. указ.: с. 408-417. - 200 экз. - ISBN 978-5-97060-409-0 (в пер.) .— Полный текст (Доступ из сети МФТИ / Удаленный доступ).

Дополнительная литература

Литература на кафедре:

1. Математические основы машинного обучения и прогнозирования [Текст] / В. В. Вьюгин ; Моск. физ.-техн. ин-т (гос. ун-т), Лаб. структурных методов анализа данных в предсказательном моделировании (ПреМоЛаб), Ин-т проблем передачи информации им. А. А. Харкевича РАН - М.МЦНМО,2013

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

<http://github.com/girafe-ai>

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На лекциях используется компьютер и мультимедийное оборудование (проектор, звуковая система),

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

Успешное освоение дисциплины требует:

- посещения студентом всех видов аудиторных занятий;
- ведения конспекта в ходе лекционных занятий;
- качественной самостоятельной подготовки к практическим занятиям, активной работы на них;
- активной самостоятельной и аудиторной работы студента;
- своевременной сдачи преподавателю заданий по аудиторным видам работ.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Прикладная математика и информатика
профиль подготовки: АІ360: Передовые методы искусственного интеллекта
Физтех-школа Прикладной Математики и Информатики
кафедра машинного обучения и цифровой гуманитаристики
курс: 3
квалификация: бакалавр

Семестр, формы промежуточной аттестации: 5 (осенний) - Дифференцированный зачет

Разработчик: Р.Г. Нейчев, старший преподаватель

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-1 Способен применять фундаментальные знания, полученные в области физико-математических и (или) естественных наук и использовать их в профессиональной деятельности	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.2 Способен строить математические модели, производить количественные расчеты и оценки
	ОПК-1.3 Способен определять границы применимости полученных результатов
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.1 Способен применять современные вычислительную технику и сервисы сети Интернет в области (сфере) профессиональной деятельности
	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
	ОПК-2.3 Знает основные требования информационной безопасности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Анализ временных рядов» обучающийся должен:

знать:

- постановку задач классификации и регрессии на основе временных данных;
- методы решения этих задач.

уметь:

- формулировать подходящие критерии качества для указанных выше задач;
- реализовывать решение на основе статистических или основанных на машинном обучении методов с учетом ограничений на сложность.

владеть:

- основными программными системами для работы с временными рядами вместе с другими инструментами анализа данных.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

1. Слабая и сильная стационарность временных рядов. Тренд, сезонность, циклы.
2. Определение AR процесса, условия стационарности. АКФ, ее графическое представление.
3. Определение MA процесса, условия стационарности.
4. SARIMAX, базовые элементы, первоначальная оценка гиперпараметров с использованием АКФ, ЧАКФ. Выбор модели.
5. Необходимые и желаемые свойства остатков.
6. Экспоненциальное сглаживание: идея, базовые элементы, выбор модели.
7. Модели ARCH и GARCH. Модель VAR. Кусочно-линейная NLAR. Facebook Prophet.
8. Признаки для машинного обучения. Кросс-валидация для временных рядов.
9. Рекуррентные нейронные сети, преимущества и недостатки. Преимущества и недостатки LSTM.
10. Сверточные нейронные сети для временных рядов, преимущества и недостатки, сравнение с RNN.

11. Механизм внимания. Трансформеры, преимущества и недостатки. Стратегии снижения временной и пространственной сложности механизма внимания. Трансформеры для временных рядов.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Что такое AR процесс, условия стационарности?
2. Дайте определение АКФ и её графическое представление.
3. Какие существуют признаки для машинного обучения?
4. Дайте определение для Кросс-валидации для временных рядов.
5. Как определить необходимые и желаемые свойства остатков?
6. Основные свойства временных рядов.
7. Рекуррентные нейронные сети, их преимущества и недостатки.
8. Что такое Трансформеры?
9. Что такое графическое представление АФК?
10. Какие существуют виды свойств остатков?

Критерии оценивания

Оценка "Отлично" (10) - полностью и вовремя решены все задачи без ошибок. Продemonстрирован грамотный подход к решению задач, реализованы оптимальные алгоритмы, код оформлен в едином удобочитаемом стиле.

Оценка "Отлично" (9) - полностью и вовремя решены все задачи без ошибок. Продemonстрирован грамотный подход к решению задач, реализованы оптимальные алгоритмы.

Оценка "Отлично" (8) - полностью и вовремя решены все задачи без ошибок. Продemonстрирован грамотный подход к решению задач.

Оценка "Хорошо" (7) - полностью решены все задачи. Допущены несущественные ошибки.

Оценка "Хорошо" (6) - полностью решено большинство задач. В некоторых задачах допущены и не исправлены ошибки, либо некоторые задачи решены частично.

Оценка "Хорошо" (5) - полностью решено две трети задач. В некоторых задачах допущены и не исправлены ошибки, либо некоторые задачи решены частично.

Оценка "Удовлетворительно" (4) - полностью решено более половины задач. В остальных задачах допущены и не исправлены ошибки, либо некоторые задачи решены частично.

Оценка "Удовлетворительно" (3) - полностью решено более половины задач.

Оценка "Неудовлетворительно" (2) - решено менее половины задач.

Оценка "Неудовлетворительно" (1) - не решено ни одной задачи.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Дифференцированный зачет может проводиться по итогам текущей успеваемости и сдачи заданий и других видов работ, предусмотренных программой дисциплины и (или) путем организации специального опроса, проводимого в устной и (или) письменной форме.

При проведении устного дифференцированного зачета обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося не должен превышать одного астрономического часа.

Во время проведения дифференцированного зачета обучающиеся могут пользоваться программой дисциплины, а также справочной литературой, конспектами лекций или другими материалами.