**Санкт-Петербургский государственный университет**

**Р А Б О Ч А Я П Р О Г Р А М М А**

**УЧЕБНОЙ ДИСЦИПЛИНЫ**

Статистическое обучение для моделирования и прогнозирования временных рядов

Statistical learning for time series modelling and forecasting

**Язык(и) обучения**

русский

Трудоемкость в зачетных единицах: 3

Регистрационный номер рабочей программы: 064825

2020

**Раздел 1. Характеристики учебных занятий**

**1.1. Цели и задачи учебных занятий**

Целью данной дисциплины является ознакомление обучающихся с методами моделирования и прогноза временных рядов в рамках методов статистического моделирования.

**1.2. Требования подготовленности обучающегося к освоению содержания учебных занятий (пререквизиты)**

Дисциплина изучается магистрантами по направлению подготовки 01.04.02 «Прикладная математика и информатика» в четвертом семестре обучения.

Обучающиеся должны иметь навыки научно-исследовательской работы и прослушать курс “Статистическое и машинное обучение”, читаемый в 3 семестре.

**1.3. Перечень результатов обучения (learning outcomes)**

В результате освоения дисциплины обучающиеся должны научиться строить эффективные алгоритмы для моделирования и прогнозирования временных рядов, а также уметь оценивать точность их применения.

Дисциплина участвует в формировании компетенций обучающихся по образовательной программе, установленных учебным планом для данной дисциплины:

ПКП-11 Способен применять современные методы анализа и прогноза многомерных данных и временных рядов для решения прикладных задач с использованием языков программирования высокого уровня, предназначенных для решения статистических задач.

ПКП-9 Способен использовать современные стохастические методы вычислений, строить статистические модели и анализировать данные.

**1.4. Перечень и объём активных и интерактивных форм учебных занятий**

Семинары 24 часа.

Курс подразумевает интерактивное обсуждение рассматриваемых задач и их решений.

**Раздел 2. Организация, структура и содержание учебных занятий**

**2.1. Организация учебных занятий**

**2.1.1 профиль Статистическое моделирование**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Трудоёмкость, объёмы учебной работы и наполняемость групп обучающихся | | | | | | | | | | | | | | | | | | |
| Код модуля в составе дисциплины,  практики и т.п. | Контактная работа обучающихся с преподавателем | | | | | | | | | | | | Самостоятельная работа | | | | Объём активных и интерактивных  форм учебных занятий | Трудоёмкость |
| лекции | семинары | консультации | практические  занятия | лабораторные работы | контрольные работы | коллоквиумы | текущий контроль | промежуточная  аттестация | итоговая аттестация | под руководством преподавателя | в присутствии  преподавателя | сам. раб. с использованием  методических материалов | текущий контроль (сам.раб.) | промежуточная аттестация (сам.раб.) | итоговая аттестация  (сам.раб.) |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | | | | | | | | | | | | | |
| Форма обучения: очная | | | | | | | | | | | | | | | | | | |
| Семестр 4 |  | 24 |  |  |  |  |  |  | 2 |  | 36 |  | 37 |  | 9 |  | 24 | 3 |
|  |  | 1-25 |  |  |  |  |  |  | 1-25 |  | 1-25 |  | 1-1 |  | 1-1 |  |  |  |
| ИТОГО |  | 24 |  |  |  |  |  |  | 2 |  | 36 |  | 37 |  | 9 |  | 24 | 3 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Виды, формы и сроки текущего контроля успеваемости и промежуточной аттестации | | | | | | |
| Код модуля в составе дисциплины, практики и т.п. | Формы текущего контроля успеваемости | | Виды промежуточной аттестации | | Виды итоговой аттестации  (только для программ итоговой аттестации и дополнительных образовательных программ) | |
| Формы | Сроки | Виды | Сроки | Виды | Сроки |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | |
| Форма обучения очная | | | | | | |
| Семестр 4 |  |  | зачёт, устно, традиционная форма | по графику промежуточной аттестации |  |  |

**2.2. Структура и содержание учебных занятий**

**Основной курс** **Основная траектория** **Очная форма обучения**

Период обучения (модуль): **Семестр 4**

|  |  |  |  |
| --- | --- | --- | --- |
| № п/п | Наименование темы (раздела, части) | Вид учебных занятий | Количество часов |
| 1 | Обсуждение методов решения реальных задач в области анализа данных типа временных рядов методами статистического обучения | семинары | 8 |
| Под руководством преподавателя | 18 |
| по методическим материалам | 18 |
| 2 | Доклады с результатами применения методов. | семинары | 16 |
| Под руководством преподавателя | 18 |
| по методическим материалам | 19 |

**Более подробное содержание**

Обучающимся предлагаются различные реальные и модельные задачи, предлагаются варианты их решения, происходит их обсуждение.

Методы анализа данных:

1. Общий подход – формирование матрицы из скользящих отрезков временного ряда.
2. Общий подход с train и test прогнозами.
3. Специально разрабатываемые методы под конкретную модель временного ряда.
4. Методы, в которых формальный подбор параметров идет методами машинного обучения (SVR, Random Forest, SSA).
5. Применение классических методов без обучения (ARIMA, экспоненциальное сглаживание).

Данные:

1. Данные изменения климата, по месяцам, в которых есть тренда, сезонность и шум.
2. Данные, имитирующие поведение цен акций, состоящие из суммы кусочно-линейного тренда и броуновского движения

**Раздел 3. Обеспечение учебных занятий**

**3.1. Методическое обеспечение**

**3.1.1 Методические указания по освоению дисциплины**

Успешное освоение дисциплины возможно благодаря посещению семинаров, участию в обсуждении докладов, хорошей подготовке обучающихся к своим докладам, самостоятельной работе, включающей в себя чтение специальной литературы по разделам темы.

**3.1.2 Методическое обеспечение самостоятельной работы**

Список литературы по изучаемым темам.

**3.1.3 Методика проведения текущего контроля успеваемости и промежуточной аттестации и критерии оценивания**

Для получения зачета обучающийся должен участвовать в in-class соревнованиях на Kaggle и получить результат лучше, чем заданная граница, а также написать отчет о применении заданного ему семейства методов на R или Python для построения прогноза.

При выполнении этих требований ставится «зачтено». Иначе – «не зачтено».

A – все задания выполнены в срок.

B – все задания выполнены, но одно из заданий не выполнено в срок.

C – все задания выполнены, но два задания не выполнены в срок.

D – все задания выполнены, но три задания не выполнены в срок.

E – все задания выполнены, но больше трех заданий не выполнено в срок.

F – хотя бы одно из заданий не выполнено.

**3.1.4 Методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации (контрольно-измерительные материалы, оценочные средства)**

**Пример реальной задачи:**

Предсказание поведения цен акций, модель – кусочно-линейный тренд плюс случайный шум типа броуновского движения. Подобрать метод, который будет угадывать направление движения цены с вероятностью больше 0.5 на модельных данных. Предполагается использование методов машинного обучения для временных рядов.

**3.1.5 Методические материалы для оценки обучающимися содержания и качества учебного процесса**

Для оценки содержания и качества учебного процесса может применяться анкетирование или опрос в соответствии с методикой и графиком, утверждаемым в установленном порядке.

**3.2. Кадровое обеспечение**

**3.2.1 Образование и (или) квалификация штатных преподавателей и иных лиц, допущенных к проведению учебных занятий**

К преподаванию дисциплины могут быть допущены преподаватели, имеющие высшее образование по соответствующему направлению.

**3.2.2 Обеспечение учебно-вспомогательным и (или) иным персоналом**

Не предусмотрено.

**3.3. Материально-техническое обеспечение**

**3.3.1 Характеристики аудиторий (помещений, мест) для проведения занятий**

Учебные аудитории для проведения учебных занятий, оснащенные стандартным оборудованием, используемым для обучения в СПбГУ в соответствии с требованиями материально-технического обеспечения

**3.3.2 Характеристики аудиторного оборудования, в том числе неспециализированного компьютерного оборудования и программного обеспечения общего пользования**

Стандартное оборудование, используемое для обучения в СПбГУ. MS Windows, MS Office, Mozilla FireFox, Google Chrome, Acrobat Reader DC, WinZip, Антивирус Касперского.

**3.3.3 Характеристики специализированного оборудования**

Не требуется.

**3.3.4 Характеристики специализированного программного обеспечения**

R, Python.

**3.3.5 Перечень и объёмы требуемых расходных материалов**

Белая бумага формата А4 для печати на принтере, 5 листов на обучающегося.

**3.4. Информационное обеспечение**

**3.4.1 Список обязательной литературы**

1.Ермаков, С. М. Метод Монте-Карло в вычислительной математике : Вводный курс - М. : БИНОМ. Лаборатория знаний ; СПб. : Невский диалект, 2009, 192 с

2.Балдин, Евгений Михайлович. Компьютерная типография LATEX (+ CD) - СПб: БХВ-Петербург, 2008. (есть в ВЦ)

3. С. И. Николенко, А. Л. Тулупьев. Самообучающиеся системы. - М. : Издательство МЦНМО, 2009. - 287 с.

**3.4.2 Список дополнительной литературы**

1. Ермаков, Сергей Михайлович. Статистическое моделирование: учебное пособие. Ч. 1-3. Санкт-Петербургский государственный университет. Научно-исследовательский институт математики и механики им. акад. В.И. Смирнова. - СПб. :СПбГУ, 2006.

2. Методы Монте-Карло в финансовой математике: Сборник работ аспирантов и студентов/ Санкт-Петербургский государственный университет. Кафедра статистического моделирования ; ред. : С. М. Ермаков, Ю. Н. Каштанов . - СПб. : СПбГУ, 2006.

**3.4.3 Перечень иных информационных источников**

* Сайт Научной библиотеки им. М. Горького СПбГУ:
* <http://www.library.spbu.ru/>
* Электронный каталог Научной библиотеки им. М. Горького СПбГУ:
* <http://www.library.spbu.ru/cgi-bin/irbis64r/cgiirbis_64.exe?C21COM=F&I21DBN=IBIS&P21DBN=IBIS>
* Перечень электронных ресурсов, находящихся в доступе СПбГУ:
* <http://cufts.library.spbu.ru/CRDB/SPBGU/>
* Перечень ЭБС, на платформах которых представлены российские учебники, находящиеся в доступе СПбГУ:
* <http://cufts.library.spbu.ru/CRDB/SPBGU/browse?name=rures&resource_type=8>
* ACM Digital Library: <http://cufts.library.spbu.ru/CRDB/SPBGU/resource/12>
* Institute of Electrical and Electronics Engineers (IEEE): <http://cufts.library.spbu.ru/CRDB/SPBGU/resource/375>
* MathSciNet - электронная коллекция Американского математического сообщества (AMS): <http://cufts.library.spbu.ru/CRDB/SPBGU/resource/415>
* O’Reilly: <http://cufts.library.spbu.ru/CRDB/SPBGU/resource/483>
* **Zentralblatt MATH:** <http://cufts.library.spbu.ru/CRDB/SPBGU/resource/86>

• <http://statmod.ru/wiki/>

**Раздел 4. Разработчики программы**

Голяндина Нина Эдуардовна, к.ф.-м.н., доцент Кафедры статистического моделирования СПбГУ, n.golyandina@spbu.ru, тел. +7(812) 428-42-31