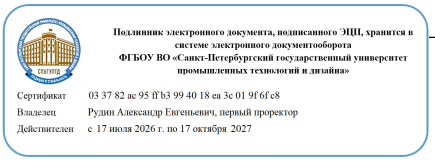
Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Санкт-Петербургский государственный университет промышленных технологий и дизайна»
(СПбГУПТД)

УТВЕРЖДАЮ
Первый проректор, проректор по УР




_____А.Е. Рудин

# Рабочая программа дисциплины

**Б1.В.12** Технологии анализа больших данных

Учебный план: 2026-2027 38.03.01 ИЭСТ Фин тех и цифр ин в биз ОО №1-1-185.plx

Кафедра: 33 Цифровых и аддитивных технологий

Направление подготовки: (специальность) 38.03.01 Экономика

Профиль подготовки: (специализация) Финансовые технологии и цифровые инновации в бизнесе

Уровень образования: бакалавриат

Форма обучения: очная

**План учебного процесса**

| Семестр (курс для ЗАО) | | Контактная работа обучающихся | | Сам. работа | Контроль, час. | Трудоёмкость, ЗЕТ | Форма промежуточной аттестации |
|---|---|---|---|---|---|---|---|
| | | Лекции | Практ. занятия | | | | |
| 5 | УП | 32 | 32 | 79,75 | 0,25 | 4 | Зачет |
| | РПД | 32 | 32 | 79,75 | 0,25 | 4 | |
| Итого | УП | 32 | 32 | 79,75 | 0,25 | 4 | |
| | РПД | 32 | 32 | 79,75 | 0,25 | 4 | |

Санкт-Петербург
2025

Рабочая программа дисциплины составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 38.03.01 Экономика, утверждённым приказом Минобрнауки России от 12.08.2020 г. № 954

Составитель (и):

| | | |
|---|---|---|
| Доцент | _____ | Лебедева Светлана Викторовна |
| От кафедры составителя:<br>Заведующий кафедрой цифровых и аддитивных технологий | _____ | Сошников Антон Владимирович |
| От выпускающей кафедры:<br>Заведующий кафедрой | _____ | Климова Наталья Сергеевна |

Методический отдел:

_____

## 1 ВВЕДЕНИЕ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

**1.1 Цель дисциплины:** Сформировать компетенции обучающегося в области теоретических и практических основ технологий анализа больших данных, получение знаний об особенностях анализа данных для решения разнообразных бизнес-аналитических задач, возникающих в профессиональной деятельности.

**1.2 Задачи дисциплины:**

Освоение методологии обработки больших данных;

Владение методами и технологиями работы с данными и знаниями;

Иметь представление о структуризации данных и технологиях Knowledge Discovery in Databases и Data Mining;

Освоение основ статистического анализа данных.

**1.3 Требования к предварительной подготовке обучающегося:**

Предварительная подготовка предполагает создание основы для формирования компетенций, указанных в п. 2, при изучении дисциплин:

Системы искусственного интеллекта

Информационные технологии

## 2 КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

| **ПК-2: Способен формировать возможные решения на основе разработанных для них целевых показателей** |
|---|
| **Знать:** Основные понятия и терминологию в области технологий обработки и анализа больших данных; методы организации взаимодействия аналитических платформ с распределенными хранилищами информации. |
| **Уметь:** Анализировать и оценивать различные варианты решений, исходя из поставленных целевых показателей деятельности организации, а также разрабатывать и адаптировать алгоритмы и модели, соответствующие специфике задач анализа больших данных. |
| **Владеть:** Навыками работы с системой интерактивного исследования и визуализации больших данных. |

## 3 РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

| Наименование и содержание разделов, тем и учебных занятий | Семестр (курс для ЗАО) | Контактная работа | | СР (часы) | Инновац. формы занятий | Форма текущего контроля |
|---|---|---|---|---|---|---|
| | | Лек. (часы) | Пр. (часы) | | | |
| Раздел 1. Технологии анализа данных | 5 | | | | | О |
| Тема 1. Введение в анализ больших данных. Методология анализа данных. Практические занятия: Аналитический и информационный подходы к моделированию. | | 4 | 4 | 10 | ИЛ | |
| Тема 2. Инженерия данных. Эволюция инженерии данных. Инженерия данных и наука о данных. Инженерия жизненного цикла данных. Основные фоновые процессы. Управление данными. Практические занятия: Этапы моделирования, формы представления данных. Подготовка данных к анализу | | 4 | 4 | 4 | ИЛ | |
| Раздел 2. Технологии анализа больших объемов данных | | | | | | О |
| Тема 3. BIG DATA Ключевые тренды Big Data. Главные характеристики Big Data. Фазы жизненного цикла. Практические занятия: Подготовка данных к анализу. Очистка данных. | | 4 | 4 | 10 | ИЛ | |
| Тема 4. Технологии аналитики больших данных. Data Mining, Машинное обучение, Business Intelligence. Основные понятия. Жизненный цикл проекта Data Science. Дескриптивная аналитика. Диагностическая аналитика. Предиктивная аналитика.Прескриптивная аналитика. Принципы анализа данных. Практические занятия: Подготовка данных к анализу. Преобразование данных. Извлечение и визуализация данных. | | 8 | 8 | 10 | ИЛ | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Тема 5. Современные технологии хранения данных.<br>DWH, Data Lake и Data Swamp. Современные платформенные решения. хранения и анализа больших данных.<br>Практические занятия: Разделение исходного набора данных на таблицы фактов и таблицы измерений. | | 4 | 4 | 19,75 | ИЛ | |
| Тема 6. Методы многомерного моделирования.<br>Подходы Ральфа Кимбалла и Билла Инмона. Архитектура хранилищ данных. Архитектура Lakehouse. Модели данных хранилища данных. Модель для витрин данных. Модель на основе денормадизованного подхода.<br>Практические занятия: разработка модели хранилища данных. Извлечение и визуализация данных. | | 8 | 8 | 26 | ИЛ | |
| Итого в семестре (на курсе для ЗАО) | | 32 | 32 | 79,75 | | |
| Консультации и промежуточная аттестация (Зачет) | | 0,25 | | | | |
| **Всего контактная работа и СР по дисциплине** | | 64,25 | | 79,75 | | |

## 4 КУРСОВОЕ ПРОЕКТИРОВАНИЕ

Курсовое проектирование учебным планом не предусмотрено

## 5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

### 5.1 Описание показателей, критериев и системы оценивания результатов обучения

#### 5.1.1 Показатели оценивания

| Код компетенции | Показатели оценивания результатов обучения | Наименование оценочного средства |
|---|---|---|
| ПК-2 | Формулирует сущность основных методов обработки и анализа больших данных на основе использования аналитических платформ с распределенными хранилищами информации.<br>Строить алгоритмы анализа и обработки больших данных.<br>Демонстрирует результат визуализации больших данных. | |

#### 5.1.2 Система и критерии оценивания

| Шкала оценивания | Критерии оценивания сформированности компетенций | |
|---|---|---|
| | Устное собеседование | Письменная работа |
| Зачтено | Отвечает на теоретический вопрос по материалам лекций, возможно допуская несущественные ошибки. Обучающийся своевременно выполнил практические задания. Учитываются баллы, накопленные в течение семестра. | |
| Не зачтено | При ответе на вопрос допускает существенные ошибки. Своевременно не выполняет (выполнил частично) практические задания. Не учитываются баллы, накопленные в течение семестра. | |

### 5.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности

#### 5.2.1 Перечень контрольных вопросов

| № п/п | Формулировки вопросов |
|---|---|
| Семестр 5 | |
| 1 | Big Data, определение, предпосылки, эволюция развития. |
| 2 | Big Data, определение, источники и виды данных. |
| 3 | Инженерия данных, определение, эволюция развития основных технологий . |

| | |
|---|---|
| 4 | Инженерия данных и наука о данных, сравнительный анализ. |
| 5 | Инженерия жизненного цикла данных. Генерация, хранение, поглощение. |
| 6 | Инженерия данных. Аналитика данных. |
| 7 | Инженерия данных. Обратный ETL. |
| 8 | Инженерия данных. Основные фоновые процессы. |
| 9 | Применение Big Data, основные технологические тренды. |
| 10 | Основные характеристики Big Data. 4 V. |
| 11 | Основные характеристики Big Data. 7 V. |
| 12 | Современные технологии управления большими данными. MapReduce и Hadoop. |
| 13 | Технологии аналитики больших данных. Методология OSEMN. |
| 14 | Технологии аналитики больших данных. Дескриптивная аналитика |
| 15 | Технологии аналитики больших данных. Предиктивная аналитика. |
| 16 | Технологии аналитики больших данных. Прескриптивная аналитика |
| 17 | Технологии аналитики больших данных. Диагностическая аналитика. |
| 18 | Современные технологии хранения данных. Data Lake и Data Swamp. |
| 19 | Современные технологии хранения данных. Lakehouse. |
| 20 | Методы многомерного моделирования. Модель для витрин данных. |
| 21 | Методы многомерного моделирования. Модель на основе денормадизованного подхода. |
| 22 | Современные технологии хранения данных. DWH. |

**5.2.2 Типовые тестовые задания**

Не предусмотрено.

**5.2.3 Типовые практико-ориентированные задания (задачи, кейсы)**

Типовые тестовые задания находятся в Приложении к данной РПД

**5.3 Методические материалы, определяющие процедуры оценивания знаний, умений, владений (навыков и (или) практического опыта деятельности)**

**5.3.1 Условия допуска обучающегося к промежуточной аттестации и порядок ликвидации академической задолженности**

Проведение промежуточной аттестации регламентировано локальным нормативным актом СПбГУПТД «Положение о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся»

**5.3.2 Форма проведения промежуточной аттестации по дисциплине**

| Устная | + | Письменная | | Компьютерное тестирование | | Иная | |
|---|---|---|---|---|---|---|---|

**5.3.3 Особенности проведения промежуточной аттестации по дисциплине**

•время на подготовку к устному собеседованию составляет 30 минут;
•выполнение кейс-задания осуществляется на компьютере за 60 минут.

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

**6.1 Учебная литература**

| Автор | Заглавие | Издательство | Год издания | Ссылка |
|---|---|---|---|---|
| **6.1.1 Основная учебная литература** | | | | |
| Васильев, Е. П., Орешков, В. И. | Интеллектуальный анализ данных в технологиях принятия решений | Рязань: Рязанский государственный радиотехнический университет | 2023 | https://www.iprbookshop.ru/134854.html |
| Татарникова, Т. М. | Интеллектуальный анализ данных | Москва, Вологда: Инфра-Инженерия | 2024 | https://www.iprbookshop.ru/143351.html |
| Данилова, Л. Ф., Полетайкин, А. Н. | Интеллектуальный анализ данных на языке Python для обучающихся по направлениям подготовки 09.03.03 Прикладная информатика, 38.03.05 Бизнес-информатика | Новосибирск: Сибирский государственный университет телекоммуникаций и информатики | 2024 | https://www.iprbookshop.ru/149511.html |
| Курносов, М. Г. | Введение в методы машинной обработки данных | Новосибирск: Автограф | 2020 | http://www.iprbookshop.ru/102117.html |

| 6.1.2 Дополнительная учебная литература | | | | |
|---|---|---|---|---|
| Мельниченко, А. С. | Математическая статистика и анализ данных | Москва: Издательский Дом МИСиС | 2018 | http://www.iprbookshop.ru/78563.html |
| Пальмов, С. В. | Интеллектуальный анализ данных | Самара: Поволжский государственный университет телекоммуникаций и информатики | 2017 | http://www.iprbookshop.ru/75376.html |
| Адлер, Ю. П., Черных, Е. А. | Статистическое управление процессами. «Большие данные» | Москва: Издательский Дом МИСиС | 2016 | http://www.iprbookshop.ru/64199.html |
| Жуковский, О. И. | Информационные технологии и анализ данных | Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент | 2014 | http://www.iprbookshop.ru/72106.html |
| Шифф В. К. | Корреляционный, регрессионный и дисперсионный анализ | СПб.: СПбГУПТД | 2019 | http://publish.sutd.ru/tp_ext_inf_publish.php?id=201950 |

**6.2 Перечень профессиональных баз данных и информационно-справочных систем**

Информационная система «Единое окно доступа к образовательным ресурсам. Раздел. Информатика и информационные технологии» [Электронный ресурс]. URL: http://window.edu.ru/catalog/?p_rubr=2.2.75.6

Электронно-библиотечная система IPRbooks [Электронный ресурс]. URL: http://www.iprbookshop.ru/

**6.3 Перечень лицензионного и свободно распространяемого программного обеспечения**

MicrosoftOfficeProfessional

Microsoft Windows

Deductor Academic

**6.4 Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине**

| Аудитория | Оснащение |
|---|---|
| Лекционная аудитория | Мультимедийное оборудование, специализированная мебель, доска |
| Компьютерный класс | Мультимедийное оборудование, компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду |

**Приложение**

рабочей программы дисциплины Технология анализа больших данных
*наименование дисциплины*

по направлению подготовки 38.03.01 Экономика
наименование ОП (профиля): Финансовые технологии и цифровые инновации в бизнесе

**5.2.3 Типовые практико-ориентированные задания (задачи, кейсы)**

<table>
<tr><th>№ п/п</th><th>Условия типовых практико-ориентированных заданий (задач, кейсов)</th></tr>
<tr><td>1</td><td>Провести регрессионный анализ зависимости выручки от числа торговых точек.<br>Таблица 1<br>Исходные данные
<table>
<tr><th>№</th><th>Выручка, тыс. руб.</th><th>Число торговых точек</th></tr>
<tr><td>1</td><td>1598</td><td>2</td></tr>
<tr><td>2</td><td>2644</td><td>5</td></tr>
<tr><td>3</td><td>2197</td><td>4</td></tr>
<tr><td>4</td><td>1959</td><td>5</td></tr>
<tr><td>5</td><td>1052</td><td>3</td></tr>
<tr><td>6</td><td>1922</td><td>3</td></tr>
<tr><td>7</td><td>2385</td><td>5</td></tr>
<tr><td>8</td><td>2581</td><td>5</td></tr>
<tr><td>9</td><td>3105</td><td>5</td></tr>
<tr><td>10</td><td>3896</td><td>4</td></tr>
<tr><td>11</td><td>1510</td><td>4</td></tr>
<tr><td>12</td><td>1880</td><td>2</td></tr>
<tr><td>13</td><td>3620</td><td>5</td></tr>
<tr><td>14</td><td>5002</td><td>6</td></tr>
<tr><td>15</td><td>2819</td><td>5</td></tr>
</table></td></tr>
<tr><td>2</td><td>Провести корреляционный анализ зависимости выручки от числа торговых точек.<br>Таблица 1<br>Исходные данные
<table>
<tr><th>№</th><th>Выручка, тыс. руб.</th><th>Число торговых точек</th></tr>
<tr><td>1</td><td>1598</td><td>2</td></tr>
<tr><td>2</td><td>2644</td><td>5</td></tr>
<tr><td>3</td><td>2197</td><td>4</td></tr>
<tr><td>4</td><td>1959</td><td>5</td></tr>
<tr><td>5</td><td>1052</td><td>3</td></tr>
<tr><td>6</td><td>1922</td><td>3</td></tr>
<tr><td>7</td><td>2385</td><td>5</td></tr>
<tr><td>8</td><td>2581</td><td>5</td></tr>
<tr><td>9</td><td>3105</td><td>5</td></tr>
<tr><td>10</td><td>3896</td><td>4</td></tr>
<tr><td>11</td><td>1510</td><td>4</td></tr>
<tr><td>12</td><td>1880</td><td>2</td></tr>
<tr><td>13</td><td>3620</td><td>5</td></tr>
<tr><td>14</td><td>5002</td><td>6</td></tr>
<tr><td>15</td><td>2819</td><td>5</td></tr>
</table></td></tr>
</table>