

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ
Декан ЭФ

УТВЕРЖДАЮ /Н.М. Стрельникова/
(Ф.И.О. декана (директора института))

11.03.2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б.1.1.27 Основы работы с большими данными

(код и наименование дисциплины по учебному плану)

Направление подготовки
(специальность)

01.03.05 Статистика

Квалификация выпускника

Бакалавр

(бакалавр/магистр/специалист)

Направленность

Прикладная статистика и анализ данных

Курс 3
Семестр 5

Распределение учебного времени

Трудоемкость по учебному плану	180 / 5	часов/зачетных единиц
Лекции	36	часов
Лабораторные работы	36	часов
Практические занятия	-	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	72	часов
Контактная работа по экзамену	6	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	72	часов
Самостоятельная работа по подготовке к экзамену	30	часов
Экзамен	5	семестр
Зачет	-	семестр
БРК, ДЗ	-	семестр

(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 01.03.05 Статистика

Программу составили:

заведующий кафедрой	ИВС	СОГЛАСОВАНО	Д.В. Морохин
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина
Кафедра информационно-вычислительных систем

(наименование кафедры)		
06.02.2024	протокол №	20
(дата)		

Заведующий кафедрой	СОГЛАСОВАНО	Д.В. Морохин
		(И.О. Фамилия)

Рабочая программа СОГЛАСОВАНА с факультетом (институтом), выпускающей(ими) кафедрой(ами).

СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	Л.В. Смоленникова
		(И.О. Фамилия)

Председатель методической комиссии факультета (института), в который входит выпускающая кафедра

СОГЛАСОВАНО	О.Е. Иванов
	(И.О. Фамилия)

Эксперт(ы): Жубрин Алексей Анатольевич, Помощник генерального директора АО «ММЗ» по информатизации – начальник управления информационных технологий

Рабочая программа проверена и зарегистрирована в УМЦ 12.03.2024 г.

Специалист учебно-методического центра СОГЛАСОВАНО /Т.А. Смирнова/

Раздел 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины является достижение планируемых результатов обучения, соответствующих установленным в ОПОП индикаторам достижения компетенций:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения
1. ОПК-1 Способен осуществлять статистическое наблюдение с использованием стандартных методик и технических средств, включая формирование выборочной совокупности и подготовку статистического инструментария	ОПК-1.1 Осуществляет статистическое наблюдение с использованием стандартных методик и технических средств	знания: Методические документы по формированию входных массивов статистических данных Методики сводки статистических данных умения: Формировать входные массивы статистических данных с использованием стандартных методик и технических средств Осуществлять сводку статистических показателей в соответствии с утвержденными методиками навыки: Способность формировать входные массивы информации баз данных
	ОПК-1.2 Выбирает статистический инструментарий, соответствующий поставленным задачам	знания: Статистический инструментарий формирования выходных массивов статистических данных Статистический инструментарий осуществления логического и арифметического контроля умения: Осуществлять подготовку статистического инструментария для решения поставленных задач Формировать выходные массивы статистической информации с применением статистического инструментария Осуществлять логический и арифметический контроль выходной информации с применением статистического инструментария навыки: Способность формировать выходные массивы информации с применением статистического инструментария

<p>2. ОПК-2 Способен формировать упорядоченные сводные массивы статистической информации и осуществлять расчет сводных и производных показателей в соответствии с утвержденным и методиками, в том числе с применением необходимой вычислительной техники и стандартных компьютерных программ</p>	<p>ОПК-2.1 Формирует упорядоченные сводные массивы статистической информации, в том числе с применением необходимой вычислительной техники и стандартных компьютерных программ</p>	<p>знания: Методики формирования входных массивов статистических данных в соответствии с заданными признаками Методики формирования упорядоченных выходных массивов статистических данных Статистический инструментарий осуществления логического и арифметического контроля Нормативные правовые акты и методические указания по обеспечению сохранности и конфиденциальности статистических данных</p> <p>умения: Формировать входные массивы статистических данных в соответствии с заданными признаками Формировать упорядоченные выходные массивы статистической информации, содержащие группировку единиц статистического наблюдения и групповые показатели, использовать их при подготовке информационно-статистических материалов Осуществлять логический и арифметический контроль выходной информации Обеспечивать сохранность и конфиденциальность статистической информации</p> <p>навыки: Способность формировать выборочную совокупность единиц статистического наблюдения в соответствии с заданными признаками, в том числе с применением необходимой вычислительной техники и стандартных компьютерных программ Способность формировать упорядоченные выходные массивы информации, содержащие группировку единиц статистического наблюдения и групповые показатели, в том числе с применением необходимой вычислительной техники и стандартных компьютерных программ</p>
	<p>ОПК-2.2 Осуществляет расчет сводных и производных показателей в соответствии с утвержденными методиками, в том числе с применением необходимой вычислительной техники и стандартных компьютерных программ</p>	<p>знания: Методики расчета сводных показателей для единиц статистического наблюдения, сгруппированных в соответствии с заданными признаками Методика расчета производных показателей для единиц статистического наблюдения</p> <p>умения: Осуществлять расчет сводных и производных показателей для единиц статистического наблюдения, сгруппированных в соответствии с заданными признаками</p> <p>навыки: Навыки расчета сводных и производных показателей в соответствии с утвержденными методиками, в том числе с применением необходимой вычислительной техники и стандартных компьютерных программ</p>

Дисциплина относится к обязательной части ОПОП.

Дисциплина является обязательной

Для продолжения формирования заявленных компетенций необходимы знания предшествующих дисциплин: Теория статистики (ОПК-1), Теория статистики (ОПК-2), Экономическая статистика (ОПК-2)

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих дисциплинах: Экспериментальное моделирование (ОПК-1), Интеллектуальные системы и технологии управления данными (ОПК-1), Планирование и проведение мероприятий по формированию статистического массива (ОПК-1), Экспериментальное моделирование (ОПК-2), Интеллектуальные системы и технологии управления данными (ОПК-2), Планирование и проведение мероприятий по формированию статистического массива (ОПК-2); государственной итоговой аттестации в форме: Подготовка к процедуре защиты и защита выпускной квалификационной работы (ОПК-1), Подготовка к процедуре защиты и защита выпускной квалификационной работы (ОПК-2)

Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

Для формирования заявленных компетенций используются методологические технологии, реализующие деятельностный, личностно-ориентированный, практико-ориентированный подходы.

Основными стратегическими технологиями являются: лекционные занятия, практические и лабораторные занятия, процедуры самообучения

На достижение конкретных целей обучения направлены применяемые тактические технологии: задания, классическая лекция

Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

5 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
Основы Больших Данных	40	ОПК-1, ОПК-2
Лекция. Предмет курса. Основные понятия, области применения и примеры использования Больших Данных. Архитектура хранилищ Больших Данных. Распределенные базы данных (БД). CAP-теорема и ее следствия для бизнеса. Введение в экосистему Hadoop.	4	
Лабораторная работа. Операторы языка SQL. Разработка SQL-запросов к реляционным сосредоточенным и распределенным БД.	4	
Лекция. Модели данных. NoSQL, базы данных. Операции CRUD в реляционных и NoSQL БД. Технология MapReduce для работы с BigData. Программные продукты для анализа Больших Данных. Аналитическая платформа Loginom. Язык Python как инструмент анализа данных.	4	
Лабораторная работа. Разработка схемы БД Oracle NoSQL MongoDB	2	
Лабораторная работа. Введение в Аналитическую платформу Loginom.	2	

Задания для самостоятельной работы, в том числе выполнение Проработка лекций Подготовка к практическим работам Операторы языка SQL	24	ОПК-1, ОПК-2
Технологии анализа данных	50	
Лекция. Структурированные и неструктурированные данные. Подготовка данных к анализу. Алгоритмы выборки, очистки, сортировки;	4	
Лекция. Поиск закономерностей в данных. Визуализация данных. Статистический анализ данных. Кластерный анализ. Модели временных рядов.	4	
Лекция. Основы языка Python. Библиотеки для анализа данных (NumPy, Matplotlib, scikit-learn)	6	
Лабораторная работа. Алгоритмы цифровой фильтрации данных	2	
Лабораторная работа. Алгоритмы поиска данных.	4	
Лабораторная работа. Алгоритмы кластерного анализа.	4	
Лабораторная работа. Модели временных рядов. Прогнозирование.	4	
Лекция. Смешение и интеграция данных - набор техник интегрировать разнородных данных. Имитационное моделирование – построение моделей имитирующих реальные процессы, разновидность экспериментальных испытаний. Пространственный анализ - класс методов, использующих топологическую, геометрическую и географическую информацию в данных.	2	ОПК-1, ОПК-2
Задания для самостоятельной работы, в том числе выполнение Проработка лекций Подготовка к практическим работам Системы дисковой памяти	20	
Алгоритмы интеллектуального анализа данных	54	
Лекция. Алгоритмы машинного обучения для решения задач регрессии, классификации и кластеризации. Алгоритмы работы рекомендательных систем и прогнозирования. Методы бустинга в задачах распознавания и видеоаналитики. Нейронные сети в задачах классификации. Задача классификации изображений. Оценка качества классификации.	4	
Лекция. Алгоритмы Data mining — глубинный анализ данных. Классификация. Ассоциация. Регрессионный анализ.	8	
Лабораторная работа. Статистическая проверка гипотез.	2	
Лабораторная работа. Байесовский классификатор.	2	
Лабораторная работа. Классификация по методу k ближайших соседей.	2	
Лабораторная работа. Сокращения размерности многомерного пространства признаков	2	
Лабораторная работа. Логистическая регрессия и дискриминационный анализ.	2	
Лабораторная работа. Представление и визуализация	4	

результатов анализа.		
Задания для самостоятельной работы, в том числе выполнение		
Проработка лекций		
Подготовка к практическим работам		
Теория вероятностей и мат. статистика	28	
Иная контактная работа:	0	
Подготовка к экзамену	30	
Проведение экзамена	6	

Раздел 5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Изучение дисциплины рекомендуется начать с ознакомления с рабочей программой, ее структурой и содержанием разделов. Учебный материал структурирован, изучение дисциплины осуществляется в тематической последовательности. **Занятия лекционного типа** дают систематизированные знания по дисциплине, концентрируют внимание на наиболее сложных и важных вопросах. Во время лекционных занятий рекомендуется вести конспектирование учебного материала; обращать внимание на формулировки и категории, раскрывающие суть проблемы, явления или процесса; зафиксировать выводы и практические рекомендации. Подготовка к **занятиям семинарского типа** включает ознакомление с планом лабораторного занятия; работу с конспектом лекций, выполнение домашнего задания, работу с учебной и учебно-методической литературой, научными изданиями и электронными образовательными ресурсами, рекомендованными рабочей программой дисциплины.

Содержание **самостоятельной работы** определяется рабочей программой дисциплины, оценочными и методическими материалами, заданиями и указаниями преподавателя. Самостоятельная работа может осуществляться в аудиторной и внеаудиторной формах. Эффективным средством осуществления самостоятельной работы является электронная информационно-образовательная среда университета, которая обеспечивает доступ к образовательной программе, рабочей программе дисциплины, к электронным библиотечным системам, профессиональным базам данных и информационным справочным системам.

Изучение дисциплины включает выполнение лабораторных работ. Периодичность проведения, формы текущего контроля успеваемости, система оценивания хода освоения дисциплин представлены в рабочей программе.

Раздел 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Учебно-методическое обеспечение

№№ п/п	Список используемой литературы	Количество экземпляров печатных изданий, имеющих в библиотеке, или электронный адрес издания (ресурса) в сети Интернет
УЧЕБНЫЕ, УЧЕБНО-МЕТОДИЧЕСКИЕ И НАУЧНЫЕ ИЗДАНИЯ		
1.	Методы и модели анализа данных: OLAP и Data Mining [Текст] : учеб. пособие по специальности 071900 "Информ. системы и технологии" направления 654700 "Информ. системы" / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. СПб.: БХВ-Петербург, 2004. - 336 с. ISBN 5-94157-522-X. Экземпляры: всего 16.	16

2.	Усков, Андрей Александрович. Интеллектуальные технологии управления [Текст] : искусственные нейронные сети и нечеткая логика / А. А. Усков, А. В. Кузьмин. М.: Горячая линия - Телеком, 2004. - 143 с. ISBN 5-93517-181-3. Экземпляры: всего 10.	10
3.	Сидоркина, Ирина Геннадьевна. Системы искусственного интеллекта [Текст] : [учеб. пособие для вузов по направлению 230100 "Информатика и вычисл. техника"] / И. Г. Сидоркина. М.: Кнорус, 2011. - 245 с. ISBN 978-5-406-00449-4. Экземпляры: всего 88.	88
4.	Цехановский, В. В. Управление данными [Электронный ресурс] : учебник / В. В. Цехановский, В. Д. Чертовской. Санкт-Петербург: Лань, 2022. - 432 с. ISBN 978-5-8114-1853-4.	https://e.lanbook.com/book/212084
5.	Макшанов, А. В. Большие данные. Big Data [Электронный ресурс] : учебник для вузов / Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н.; Журавлев А. Е., Тындыкарь Л. Н. 3-е изд., стер. Санкт-Петербург: Лань, 2023. - 188 с. ISBN 978-5-507-46866-9.	https://e.lanbook.com/book/322664

6.2. Материально-техническая база и программное обеспечение

№№ п/п	Аудитории для проведения учебных занятий, самостоятельной работы и проведения государственной итоговой аттестации	Перечень основного оборудования	Программное обеспечение
1.	518 (III)	Системный блок CEL D-341 FAN/ASUS S-775/512 M/160.0G/DVD+-RW (1), ПК 5 - ICL RAY P222.3 ,клавиат.,мышь.,монитор LG E2251T-BN (14), Сист. блок CE 331/256*2/PC 3200/80 Gb/FDD/DVD-ROM/КЛАВ+МЫШЬ+коврик (1), Комплект учебной мебели (1)	Microsoft Windows Enterprise, Справочная правовая система "Консультант Плюс", Microsoft Office Standard, Агент Dr.Web, Комплект ГАРАНТ-Мастер, Microsoft Access, Microsoft Visio Professional, Microsoft Project Professional, Microsoft Visual Studio Enterprise, Комплект ПО для решения основных пользовательских задач

Раздел 7. ФОРМЫ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ/ ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Критерии оценивания индикаторов достижения компетенций направлены на:

- усвоение теоретического материала (объем знаний, глубина усвоения), предусмотренного рабочей программой;
- умение излагать материал (четкость, грамотность изложения материала, точность и полнота воспроизведения учебного материала);
- умение применять теоретические знания при решении практических заданий.

Шкала оценивания представлена ниже.

Уровень сформированности элементов компетенции	Критерии оценивания	Шкала оценивания
Пороговый уровень	Обучающийся имеет знания основного материала, проявляет умение логично его излагать, но может допускать неточности в изложении материала, недостаточно правильные формулировки, испытывает затруднения в выполнении практических заданий.	удовлетворительно
Продвинутый уровень	Обучающийся твердо знает программный материал, излагает его грамотно и по существу, не допускает существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения	хорошо
Высокий уровень	Обучающийся глубоко и прочно усвоил программный материал, грамотно и логически стройно его излагает, дает исчерпывающие ответы на поставленные вопросы. В ответе тесно увязывается теория с практикой, при этом обучающийся не затрудняется с ответом при видоизменении задания, свободно справляется с задачами, вопросами и другими видами применения знаний, показывает знакомство с монографической литературой, периодическими изданиями, правильно обосновывает принятые решения, свободно владеет разносторонними навыками, приемами выполнения практических работ	отлично

7.1. Текущий контроль успеваемости

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины (модуля) и производится с применением технологии рейтингового контроля в соответствии с технологической картой дисциплины. Порядок составления технологической карты и алгоритм проведения процедуры оценивания видов деятельности обучающихся, направленных на освоение знаний, умений, навыков и/или опыта деятельности, по накопительной системе в баллах устанавливается положением о системе РИТМ в ФГБОУ ВО «ПГТУ»

7.2. Промежуточная аттестация обучающихся

Промежуточная аттестация обучающихся направлена на оценивание результатов обучения по дисциплине (модулю) и проводится с использованием фондов оценочных средств.

Примеры типовых контрольных заданий из базы фонда оценочных средств по образовательной программе.

Билет №0

Основы работы с большими данными

1. Дисковые системы RAID. Способы организации и обеспечения доступности.
2. Принципы информационного поиска.

Тестовые задания из ФОС

1. Именованная совокупность данных, отображающая состояние объектов и их отношений в рассматриваемой предметной области это
 1. **База данных**
 2. Система управления базами данных
 3. Система баз данных
 4. Информационное приложение
2. Средства запросов к БД (добавление, удаление, обновление данных, создание и уничтожение БД) включает в себя
 1. **язык манипулирования данными**
 2. язык описания данных
 3. язык описания хранимых данных
 4. нет правильного ответа
3. Модель предметной области, отражающая объекты и связи между ними (синонимы: модель «сущность - связь», концептуальная модель) это
 1. **ER-модель**
 2. SQL-сервер
 3. СУБД
 4. нет правильного ответа
4. Специалист, обеспечивающий создание, функционирование и развитие БД
 1. **администратор базы данных**
 2. системный администратор
 3. программный администратор
 4. сетевой администратор
5. Наименьшая семантическая единица информации это
 1. **поле**
 2. файл
 3. запись
 4. таблица
6. Позволяет строить БД с древовидной структурой
 1. **иерархическая модель**
 2. сетевая модель
 3. реляционная модель
 4. нет правильного ответа

7. В этой модели возможны связи всех информационных объектов со всеми
1. **сетевой**
 2. объектно-ориентированной
 3. реляционной
 4. иерархической
8. Любая сущность предметной области - это
1. **объект**
 2. сфера
 3. предмет
 4. экземпляр
9. Свойства, характеризующие сущность (столбцы) и степень отношения - количество столбцов это
1. **атрибуты**
 2. признаки
 3. реквизиты
 4. нет правильного ответа
10. Список имен атрибутов
1. **схема отношения**
 2. схема объектов
 3. схема действия
 4. нет правильного ответа
11. Ограниченная по смыслу совокупность значений атрибутов отношения это
1. **домен**
 2. имя
 3. название
 4. тип
12. Совокупность логически связанных полей в базе данных
1. **запись**
 2. связь
 3. поле
 4. заметка
13. Совокупность экземпляров записей одной структуры это
1. **таблица**

2. индекс
 3. конверт
 4. папка
14. Элемент таблицы, по которому можно определить значения других элементов данных этой таблицы
1. **Первичный ключ**
 2. Главный элемент
 3. Внешний ключ
 4. Ведущий элемент
15. Атрибут, уникально идентифицирующий строки отношения это
1. **первичный ключ**
 2. составной ключ
 3. вторичный ключ
 4. альтернативный ключ
16. Математическая структура для моделирования попарных отношений между объектами –
Граф
17. Класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач -
машинное обучение, machine learning
18. Основное средство управления и обращения с запросами к данным, хранящимся в
реляционных базах данных
SQL
19. ... данные поступают в систему при возникновении некоторых событий, а не загружаются в
хранилище данных большими массивами.
Потоковые
20. Статистическая взаимосвязь двух или более случайных величин (либо величин, которые
можно с некоторой допустимой степенью точности считать таковыми). При этом изменения
значений одной или нескольких из этих величин сопутствуют систематическому изменению
значений другой или других величин - ...
Корреляция

Перечень вопросов для проведения промежуточной аттестации

16. В чем состоит принцип распределения в дисковых массивах?
17. В чем состоит принцип зеркалирования в дисковых массивах?
18. Объясните принципы работы RAID-массивов базового уровня 10 и 01.

19. Объясните принципы работы RAID-массивов базового уровня 5.
20. Опишите архитектуру распределенного хранилища данных.
21. Что такое кластер?
22. Какие системы называются масштабируемыми?
23. Что такое «вертикальное» и "горизонтальное" масштабирование.
24. Что такое репликация данных?
25. Опишите технологии репликации «Master-Slave», «Peer-to-Peer».
26. Что называется «шардингом»?
27. В чем различие между «вертикальным» и «горизонтальным» шардингом?
28. Объясните термин NoSQL.
29. Какая система хранения данных называется распределенной?
30. Что утверждает теорема CAP и для каких систем хранения она справедлива?
31. В чем состоит состоятельность данных?
32. Как в системах BigData компенсируется возможная несостоятельность данных?
33. Охарактеризуйте модель данных «Ключ-Значение».
34. Укажите основные этапы вычислений технологии Map-Reduce.
35. Какие действия выполняются на этапе Map?
36. Какова цель этапа Shuffle в технологии Map-Reduce?
37. Какие задачи решаются на этапе Reduce?
38. Какую роль играет Master-node в технологии Map-Reduce?
39. Каковы функции Worker-node в технологии Map-Reduce?
40. Какие задачи решаются методами DataMining?
41. Как определяется задача информационного поиска?
42. Что такое «Инвертированный список»?
43. В чем заключается принцип работы алгоритма «грубой силы» при поиске в тексте заданной строки?
44. Объясните принцип работы алгоритма Бойера — Мура — Хорспула поиска в тексте заданной строки
45. Сформулируйте задачу нечеткого поиска в тексте заданной строки.
46. Как вычисляется расстояние Левенштейна между двумя строками?
47. Объясните принципы логистической регрессии данных.
48. Что такое векторизация слов?
49. Как работают деревья принятия решения?
50. Какую структуру может иметь элемент «Значение» в паре «Ключ-Значение»?
51. Охарактеризуйте графовую модель данных.

52. В каких приложениях целесообразно использовать графовую модель данных?
53. Объясните структуру хранилища данных СУБД MongoDB и соответствие элементов этой структуры элементам традиционных реляционных СУБД.
54. Поясните принципы моделирования данных в документарной модели.
55. Сравните ссылочную модель и модель вложения с реляционной моделью данных.
56. Какие операции понимаются под аббревиатурой CRUD?
57. Каким образом приложения, написанные на разных языках программирования, получают доступ к документам СУБД MongoDB?
58. Как выполнить подключение к серверу СУБД MongoDB из приложения?
59. Как из приложения получить доступ к коллекции документов?
60. Каким образом обеспечивается выбор документов, соответствующих заданным критериям отбора (на примере MongoDB)?
61. Охарактеризуйте байесовские алгоритмы классификации.
62. Что такое "бустинг"? Укажите основные принципы бустинга.
63. Опишите метод ближайшего соседа.
64. По каким алгоритмам строятся решающие деревья?
65. Охарактеризуйте алгоритм ID3.
66. Охарактеризуйте алгоритм CART.
67. Какова задача регрессионного анализа?
68. Что такое "множественная регрессия"?
69. С какой целью выполняется тест Дарбина-Уотсона?
70. Как оценить значимость модели регрессии?
71. Перечислите виды дискриминантного анализа.
72. Какие задачи решает дискриминантный анализ?
73. Охарактеризуйте виды дискриминантного анализа.
74. Алгоритм поиска дискриминантной функции и критерий Фишера. Какую функцию расстояния рекомендуется использовать для оценки качества дискриминантного анализа

An error has occurred while processing HtmlTextBox 'htmlTextBox3': List item () must be in a list (or).