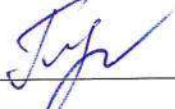


Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Санкт-Петербургский государственный лесотехнический университет
имени С.М. Кирова»


СОГЛАСОВАНО

Начальник ОЦНПК


_____/Л.Я. Громская/
13 мая 2022г.

УТВЕРЖДАЮ

Руководитель программы аспирантуры


_____/С.А. Иванов/
13 мая 2022г.

РАБОЧАЯ ПРОГРАММА

дисциплины 2.1.7.1. Интеллектуальный анализ данных при проведении научных исследований

(шифр по учебному плану; наименование)

уровень образования: подготовка кадров высшей квалификации

по научной специальности 2.3.1. Системный анализ, управление и обработка информации, статистика

(шифр и наименование научной специальности)

Кафедра информационных систем и технологий

(наименование кафедры)

Объем дисциплины – 3 з.е.

Форма контроля – зачет

Рабочая программа составлена на основании Федеральных государственных требований

Составители:

1. _____ - _____ Иванов Сергей Александрович
(ученое звание) (должность) (Ф.И.О. полностью)

Рабочая программа обсуждена на заседании кафедры _____
Информационных систем и технологий _____

протокол № 08 от « 13 » мая 20 22 г.

Заведующий кафедрой, к.т.н., доцент _____ /Вагизов М.Р. /
(ученое звание, подпись, Ф.И.О. полностью)

Проверено

ООПиКО

_____ /
(подпись)

1 _____ /
(расшифровка подписи)

1. ЦЕЛЬ И ЗАДАЧИ ДИСЦИПЛИНЫ, ЕЁ МЕСТО В ПРОГРАММЕ АСПИРАНТУРЫ

1.1. Цели и задачи дисциплины

Цель дисциплины: формирование теоретических знаний и практических навыков по сбору, обработке и интерпретации данных в рамках научно-исследовательской деятельности.

Задачи дисциплины:

- изучение задач классификации и кластеризации больших объемов данных;
- изучение интеллектуальных систем для решения аналитических задач при проведении научных исследований;
- формирование навыков работы с большими данными;
- изучение технологий и программных средств обработки больших данных и методов машинного обучения для решения научно-исследовательских задач.

1.2. Место дисциплины в структуре программы

Дисциплина «Интеллектуальный анализ данных при проведении научных исследований» является элективной дисциплиной.

Дисциплина основывается на результатах освоения программы курса магистратуры, а также дисциплины «Статистический анализ данных».

Изучение дисциплины необходимо для прохождения практики по получению профессиональных умений и опыта профессиональной деятельности (педагогическая практика), подготовки диссертации на соискание ученой степени кандидата наук.

1.3. Объем дисциплины, виды учебной работы и форма аттестации

Вид учебных занятий	Часов / з.е.	Курс, семестр
Всего по дисциплине	108 / 3	II, 4
Контактная работа обучающихся с преподавателем, всего	20	
в том числе,		
лекции	20	
практические занятия (семинары)	-	
лабораторные работы	-	
Самостоятельная работа	88	
Форма промежуточной аттестации	зачет	II, 4

2. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Планируемые результаты изучения дисциплины (модуля) (знания, умения и навыки, опыт деятельности в данной области):

Знать:

- виды программного обеспечения для обработки больших объемов данных;
- понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных;
- методы интеллектуального анализа больших данных и машинного обучения, принципы поиска, сбора, очистки, хранения, обработки и анализа и визуализации больших данных для решения научно-исследовательских задач.

Уметь:

- применять методы обработки больших объемов данных и использовать программные средства с технологией Big Data при решении научно-исследовательских задач;
- применять модели машинного обучения для решения научно-исследовательских задач в различных предметных областях.

Владеть:

- навыками работы с программным обеспечением для обработки больших объемов данных;
- понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов;
- профессиональной терминологией в области больших данных и машинного обучения;
- навыками практического применения методов интеллектуального анализа больших данных и машинного обучения.

3. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1. Наименование тем (разделов), их содержание, объём в часах лекционных занятий

Темы (разделы) дисциплины и их содержание	Трудо-емкость, час	Результаты обучения по дисциплине (модулю)
1. Интеллектуальный анализ данных, большие данные, машинное обучение. Методы и задачи интеллектуального анализа данных, машинного обучения и обработки больших данных. Области применения методов и технологий интеллектуального анализа данных, машинного обучения и обработки больших данных. Примеры задач машинного обучения: поиск информации в интернете, распознавание изображений, лиц, эмоций, пола, возраста, распознавание речи, языка, эмоциональной окраски текстов, прогнозирование продаж, прогнозирование оттока клиентов, кредитный скоринг, рекомендательные системы и др. Основные характеристики больших данных и их влияние на сбор, хранение, обработку и анализ данных. Критерии аналитических задач, решение которых предпочтительно с использованием технологий BigData.	2	<p>Знать:</p> <ul style="list-style-type: none">- виды программного обеспечения для обработки больших объемов данных;- понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных;- методы интеллектуального анализа больших данных и машинного обучения, принципы поиска, сбора, очистки, хранения, обработки и анализа и визуализации больших данных для решения научно-исследовательских задач. <p>Уметь:</p> <ul style="list-style-type: none">- применять методы обработки больших объемов данных и использовать программные средства с технологией Big Data при решении научно-исследовательских задач;- применять модели машинного обучения для решения научно-исследовательских задач в различных предметных областях. <p>Владеть:</p> <ul style="list-style-type: none">- навыками работы с программным обеспечением для обработки больших объемов данных;- понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов;- профессиональной терминологией в области больших данных и машинного обучения;- навыками практического применения методов интеллектуального анализа больших данных и машинного обучения.
2. Классификации. Матрицы данных. Гипотезы компактности и скрытых факторов. Структура матрицы данных и задачи обработки. Измерение признаков. Основные типы шкал. Проблема адекватности. Основные задачи анализа и интерпретации данных в научно-исследовательской деятельности. Пошаговый дискриминантный анализ. Ошибки клас-	6	<p>Знать:</p> <ul style="list-style-type: none">- виды программного обеспечения для обработки больших объемов данных;- понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных;- методы интеллектуального анализа больших данных и машинного обучения, принципы поиска, сбора, очистки, хранения, обработки и анализа и визуализации больших данных для решения научно-исследовательских задач. <p>Уметь:</p> <ul style="list-style-type: none">- применять методы обработки больших объе-

Темы (разделы) дисциплины и их содержание	Трудоемкость, час	Результаты обучения по дисциплине (модулю)
сификации. Непараметрическое оценивание.		<p>мов данных и использовать программные средства с технологией Big Data при решении научно-исследовательских задач;</p> <ul style="list-style-type: none"> - применять модели машинного обучения для решения научно-исследовательских задач в различных предметных областях. <p>Владеть:</p> <ul style="list-style-type: none"> - навыками работы с программным обеспечением для обработки больших объемов данных; - понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов; - профессиональной терминологией в области больших данных и машинного обучения; - навыками практического применения методов интеллектуального анализа больших данных и машинного обучения.
<p>3. Кластерный анализ и снижение размерности</p> <p>Основные типы задач кластер-анализа. Меры подобия и функции расстояния. Использование методов теории графов в задачах кластеризации. Анализ матриц исходных данных. Метод главных компонент. Корреляционная матрица и её свойства. Модели факторного анализа. Многомерное шкалирование. Метрический и неметрические подходы. Многомерное шкалирование неметрических данных. Многомерные развертки.</p>	6	<p>Знать:</p> <ul style="list-style-type: none"> - виды программного обеспечения для обработки больших объемов данных; - понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных; - методы интеллектуального анализа больших данных и машинного обучения, принципы поиска, сбора, очистки, хранения, обработки и анализа и визуализации больших данных для решения научно-исследовательских задач. <p>Уметь:</p> <ul style="list-style-type: none"> - применять методы обработки больших объемов данных и использовать программные средства с технологией Big Data при решении научно-исследовательских задач; - применять модели машинного обучения для решения научно-исследовательских задач в различных предметных областях. <p>Владеть:</p> <ul style="list-style-type: none"> - навыками работы с программным обеспечением для обработки больших объемов данных; - понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов; - профессиональной терминологией в области больших данных и машинного обучения; - навыками практического применения методов интеллектуального анализа больших данных и машинного обучения.
4. Методы и инструментальные	6	Знать:

Темы (разделы) дисциплины и их содержание	Трудоемкость, час	Результаты обучения по дисциплине (модулю)
<p>средства интеллектуального анализа данных Классификация методов прогнозирования. Методы скользящего среднего. Экспоненциальное сглаживание. Регрессионный анализ и прогнозирование. Прогнозирование на основе параметрических моделей. Понятие об интеллектуальных системах анализа и интерпретации данных. Предметно-ориентированные аналитические системы, статистические пакеты, нейронные сети, деревья решений, обнаружение логических закономерностей, генетические алгоритмы, системы визуализации многомерных данных, полученных в ходе научных исследований.</p>		<ul style="list-style-type: none"> - виды программного обеспечения для обработки больших объемов данных; - понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных; - методы интеллектуального анализа больших данных и машинного обучения, принципы поиска, сбора, очистки, хранения, обработки и анализа и визуализации больших данных для решения научно-исследовательских задач. <p>Уметь:</p> <ul style="list-style-type: none"> - применять методы обработки больших объемов данных и использовать программные средства с технологией Big Data при решении научно-исследовательских задач; - применять модели машинного обучения для решения научно-исследовательских задач в различных предметных областях. <p>Владеть:</p> <ul style="list-style-type: none"> - навыками работы с программным обеспечением для обработки больших объемов данных; - понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов; - профессиональной терминологией в области больших данных и машинного обучения; - навыками практического применения методов интеллектуального анализа больших данных и машинного обучения.
Итого часов лекций:	20	

3.2. Практические (семинарские) занятия

Практические (семинарские) занятия - не предусмотрены учебным планом.

3.3. Лабораторные занятия

Лабораторные занятия - не предусмотрены учебным планом.

3.4. Курсовой проект (работа)

Курсовой проект (работа) - не предусмотрены учебным планом.

3.5. Самостоятельная работа и контроль успеваемости

В рамках часового фонда самостоятельной работы данной дисциплины

предусматривается выполнение следующих видов учебных занятий:

Вид самостоятельной работы	Примерная трудоёмкость, час
проработка лекционного материала по конспекту и учебной литературе	18
самостоятельное изучение отдельных вопросов тем дисциплины, не рассматриваемых на лекциях	10
поиск, изучение и презентация информации по заданной проблеме, анализ научных публикаций по заданной теме	10
исследовательская работа, участие в конференциях, семинарах	10
анализ данных по заданной теме	10
подготовка к промежуточной аттестации (контроль)	30
Итого:	88

- **темы, выносимые для самостоятельного изучения:**

В рамках тем дисциплины аспиранты должны изучить дополнительный материал по следующим вопросам:

Тема	Примерная трудоёмкость, час
Использование кластеризации признаков для снижения размерности	2
Методы оценивания моделей авторегрессии, скользящего среднего и смешанных моделей	2
Табличные процессоры и базы данных в задачах обработки данных	4
Метод непараметрического оценивания на основе K-ближайшего соседства	2
Итого:	10

Вопросы для самоконтроля

1. Гипотезы компактности и скрытых факторов.
2. Измерение признаков.
3. Основные типы шкал.
4. Решающие поверхности и дискриминантные функции.
5. Основные типы задач кластер-анализа.
6. Анализ матриц исходных данных.
7. Модели факторного анализа.
8. Классификация методов прогнозирования.
9. Регрессионный анализ и прогнозирование.

10. Понятие об интеллектуальных системах анализа и интерпретации данных.

11. Требования к статистическим пакетам общего назначения.

Текущий контроль проводится в форме контрольного опроса (КО).

Промежуточная аттестация проводится в форме зачета.

3.6. Распределение часов по темам и видам занятий

№ темы	Наименование темы дисциплины (модуля)	Объем работы аспиранта, ч					Оценоч. ср-ва /Форма контроля
		лекции	практ. занятия	лаборат. занятия	самост. работа	всего	
1	Интеллектуальный анализ данных, большие данные, машинное обучение.	2	-	-	20	22	КО
2	Классификации	6	-	-	22	28	КО
3	Кластерный анализ и снижение размерности	6	-	-	16	22	КО
4	Методы и инструментальные средства интеллектуального анализа данных	6	-	-	20	26	КО
	Подготовка к промежуточной аттестации	-	-	-	10	10	Вопросы для зачета / зачет
	ВСЕГО по дисциплине	20	-	-	88	108	зачет

3.7. Образовательные технологии

Изучение дисциплины построено на использовании традиционных технологий (лекций) в сочетании с самостоятельной работой обучающегося. Предусматривается применение инновационных форм учебных занятий, с применением информационных потоково-групповых лекций, проблемных лекций, активных лекций (с элементами лекции-гипотезы, лекции-консультации, лекции-дискуссии), а также использование современных подходов к оценке знаний

обучающихся. В лекционных занятиях предусматривается широкое использование мультимедийных технологий.

3.7. Образовательные технологии

Изучение дисциплины построено на использовании традиционных технологий (лекций) в сочетании с самостоятельной работой обучающегося. Предусматривается применение инновационных форм учебных занятий, с применением информационных потоково-групповых лекций, проблемных лекций, активных лекций (с элементами лекции-гипотезы, лекции-консультации, лекции-дискуссии), а также использование современных подходов к оценке знаний обучающихся. В лекционных занятиях предусматривается широкое использование мультимедийных технологий.

4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

4.1. Основная литература

1. Кумратова, А. М. Методы хранения и анализа данных : учебное пособие / А. М. Кумратова, И. И. Василенко. — Краснодар : КубГАУ, 2021. — 183 с. — ISBN 978-5-907474-28-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/254297>
2. Ямашкин, С. А. Системный анализ пространственных данных : учебное пособие / С. А. Ямашкин, А. А. Ямашкин. — Саранск : МГУ им. Н.П. Огарева, 2020. — 44 с. — ISBN 978-5-7103-4050-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/204680>



4.2. Дополнительная литература

1. Афанасьева, Т. В. Введение в проектирование систем интеллектуального анализа данных : учебное пособие / Т. В. Афанасьева. — Ульяновск : УлГТУ, 2017. — 64 с. — ISBN 978-5-9795-1686-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165064>
2. Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семёнова, Е. С. Четвертакова, С. С. Вожов. — Новоси-

бирск : НГТУ, 2017. — 74 с. — ISBN 978-5-7782-3183-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/118287>



4.3. Учебно-методическое обеспечение самостоятельной работы

1. Кузьмич, Р. И. Модификации метода логического анализа данных для задач классификации : монография / Р. И. Кузьмич, И. С. Масич. — Красноярск : СФУ, 2018. — 180 с. — ISBN 978-5-7638-3698-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/157748>
2. Методы и модели эконометрики : учебное пособие / О. И. Бантикова, В. И. Васянина, Ю. А. Жемчужникова, А. Г. Реннер. — Оренбург : ОГУ, [б. г.]. — Часть 1 : Анализ данных — 2015. — 574 с. — ISBN 978-5-7410-1331-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/98113>

4.4. Ресурсы сети «Интернет»

1. Электронно-Библиотечная Система издательства «Лань»
<http://e.lanbook.com>
2. Сайт журнала «Известия Санкт-Петербургской лесотехнической академии»
<http://spbftu.ru/science/pub/izvest/>
3. Сборники трудов молодых ученых СПбГЛТУ
<http://spbftu.ru/science/pub/young/>
4. Программы научно-технических конференций
<http://spbftu.ru/science/program/>
5. Сайт Российской Национальной библиотеки <http://www.nlr.ru/>
6. Виртуальная справочно-правовая система компании КонсультантПлюс
<http://www.consultant.ru/>
7. Всемирная электронная база данных научных изданий
<http://www.sciencedirect.com/>
8. Научная электронная библиотека <http://elibrary.ru/defaultx.asp>

9. Электронные книги <http://eknigi.org>
10. Электронные книги <http://razym.ru>
11. Электронно-библиотечная система «Университетская библиотека онлайн»
www.biblioclub.ru
12. Электронная библиотека диссертаций <http://diss.rsl.ru>
13. Президентская библиотека им Б. Н. Ельцина <http://www.prlib.ru>
14. Российское образование Федеральный портал <http://www.edu.ru>

4.5. Информационные технологии

1. Пакет прикладных программ «Microsoft Office»
2. «Интернет» ресурсы.
3. Справочно-правовая система «КонсультантПлюс» <http://www.consultant.ru>.
4. ЭБС Издательство Лань ЭБС <http://e.lanbook.com>.

5. ОЦЕНОЧНЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

5.1. Текущий контроль

Контрольный опрос (КО)

Типовые вопросы для контрольного опроса (КО)

КО-1

1. Эволюция методов обработки и анализа данных. Переход от статистической обработки к методам OLAP и Data Mining.
2. Принципы, лежащие в основе методов OLAP и Data Mining.
3. Общая характеристика методов интеллектуального анализа данных. Обзор основных задач, решаемых методами Data Mining.
4. Алгоритмы предварительной обработки данных. Очистка данных: факторный анализ, корреляционный анализ, фильтрация.
5. Трансформация данных: метод скользящего окна, квантование, сортировка и группировка, слияние данных.
6. Понятие и основные принципы «извлечения знаний».
7. Основные задачи, решаемые методами Data Mining.
8. Автокорреляция. Линейная и логистическая регрессия.
9. Дерево решений: принципы построения, использование, интерпретация результатов.
8. Метод поиска ассоциативных правил, области его применения.
10. Нейросетевые методы анализа данных. Принципы функционирования искусственных нейронных сетей. Основные области их применения.

КО-2

1. Как составляются матрицы данных?
2. Назовите гипотезы компактности и скрытых факторов?
3. Опишите структуру матрицы данных, что является задачей обработки.
4. Перечислите основные типы шкал.
5. В чем заключается проблема адекватности?
6. Перечислите основные задачи анализа и интерпретации данных в науч-

но-исследовательской деятельности.

КО-3

1. Перечислите основные типы задач кластер-анализа.
2. Какие меры подобия и функции расстояния Вы знаете?
3. Как используются методы теории графов в задачах кластеризации?
4. В чем заключается метод главных компонент?
5. Назовите модели факторного анализа.
6. В чем заключается многомерное шкалирование?
7. В чем заключаются метрический и неметрические подходы?

КО-4

1. Приведите классификацию методов прогнозирования.
2. Приведите примеры применения методов скользящего среднего.
3. В чем заключается метод экспоненциального сглаживания?
4. Для чего используются регрессионный анализ и прогнозирование?
5. Как выполняется прогнозирование на основе параметрических моделей?

Критерии оценивания

№ п/п	Критерии оценки	Оценка	Оценка в баллах
1	Правильность ответа на вопрос	- отвечено правильно	1
		- отвечено частично или не правильно	0

Оценивается каждый ответ. Максимум - 1 балл

Шкала оценивания

Баллы по критерию оценки	0	1
Оценка	Не зачтено	Зачтено

В рамках контролируемых тем аудитории задаются вопросы. При наличии желающих дать ответ, опрашиваются обучающиеся до момента получения правильной формулировки, использующей необходимые понятия, категории и законы. В случае отсутствия желающих ответить, обучающиеся опрашиваются по усмотрению преподавателя до получения правильной формулировки ответа. Время опроса ограничено – 10–15 мин (2–3 мин на вопрос).

5.2. Промежуточная аттестация (зачет)

5.2.1. Типовые вопросы для зачета

1. Эволюция методов обработки и анализа данных. Переход от статистической обработки к методам OLAP и Data Mining.
2. Принципы, лежащие в основе методов OLAP и Data Mining.
3. Общая характеристика методов интеллектуального анализа данных. Обзор основных задач, решаемых методами Data Mining.
4. Алгоритмы предварительной обработки данных. Очистка данных: факторный анализ, корреляционный анализ, фильтрация.
5. Трансформация данных: метод скользящего окна, квантование, сортировка и группировка, слияние данных.
6. Понятие и основные принципы «извлечения знаний».
7. Основные задачи, решаемые методами Data Mining.
8. Автокорреляция. Линейная и логистическая регрессия.
9. Дерево решений: принципы построения, использование, интерпретация результатов.
8. Метод поиска ассоциативных правил, области его применения.
10. Нейросетевые методы анализа данных. Принципы функционирования искусственных нейронных сетей. Основные области их применения.
11. Как составляются матрицы данных?
12. Назовите гипотезы компактности и скрытых факторов?
13. Опишите структуру матрицы данных, что является задачей обработки.
14. Перечислите основные типы шкал.
15. В чем заключается проблема адекватности?
16. Перечислите основные задачи анализа и интерпретации данных в научно-исследовательской деятельности.
17. Перечислите основные типы задач кластер-анализа.
18. Какие меры подобия и функции расстояния Вы знаете?
19. Как используются методы теории графов в задачах кластеризации?

20. В чем заключается метод главных компонент?
21. Назовите модели факторного анализа.
22. В чем заключается многомерное шкалирование?
23. В чем заключаются метрический и неметрические подходы?

5.2.2. Критерии оценки освоения дисциплины (зачет)

С целью оценки уровня освоения дисциплины на зачете используется система «зачтено / не зачтено».

Оценка	Критерии
Зачтено	Аспирант показал творческое отношение к обучению, в совершенстве или в достаточной степени овладел теоретическими вопросами дисциплины, показал все требуемые знания, умения и навыки
Не зачтено	Аспирант имеет пробелы по отдельным теоретическим разделам специальной дисциплины и не владеет как минимум основными умениями и навыками.

Зачёт проводится в устной форме по вопросам к зачёту. Преподаватель задаёт аспиранту 2–3 вопроса по разным темам, охваченным дисциплиной. При необходимости преподаватель задаёт уточняющие (в рамках уже заданных) или дополнительные вопросы. Решение принимается по совокупности ответов на все заданные вопросы.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Сведения об обеспеченности образовательного процесса оборудованием

№ п/п	№ аудитор	Перечень основного оборудования, которым оснащены аудитории
		для проведения лекций
1	310-2	Сеть компьютеров, проектор, плакаты, видеоматериалы

Учебная аудитория для проведения занятий лекционного типа – оснащена оборудованием и техническими средствами обучения.

Учебная аудитория для групповых и индивидуальных консультаций– оснащена оборудованием и техническими средствами обучения.

Учебная аудитория для текущего контроля и промежуточной аттестации – оснащена оборудованием и техническими средствами обучения.

Помещение для самостоятельной работы – оснащено компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечено доступом в электронную информационно-образовательную среду организации.

7. АННОТАЦИЯ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

«Интеллектуальный анализ данных при проведении научных исследований»

Объем дисциплины – 3 з.е.

Форма контроля – зачет.

1. Цель изучения дисциплины

формирование теоретических знаний и практических навыков по сбору, обработке и интерпретации данных в рамках научно-исследовательской деятельности.

2. Задачи изучения дисциплины

- изучение задач классификации и кластеризации больших объемов данных;
- изучение интеллектуальных систем для решения аналитических задач при проведении научных исследований;
- формирование навыков работы с большими данными;
- изучение технологий и программных средств обработки больших данных и методов машинного обучения для решения научно-исследовательских задач.

3. Содержание

Тема 1. Интеллектуальный анализ данных, большие данные, машинное обучение.

Тема 2. Классификации

Тема 3. Кластерный анализ и снижение размерности

Тема 4. Методы и инструментальные средства интеллектуального анализа данных

4. Требования к предварительной подготовке аспирантов

Дисциплина основывается на результатах освоения программы курса магистратуры, а также дисциплины «Статистический анализ данных».

5. Требования к результатам освоения

В результате изучения дисциплины аспирант должен:

знать:

- виды программного обеспечения для обработки больших объемов данных;

- понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных;

- методы интеллектуального анализа больших данных и машинного обучения, принципы поиска, сбора, очистки, хранения, обработки и анализа и визуализации больших данных для решения научно-исследовательских задач.

уметь:

- применять методы обработки больших объемов данных и использовать программные средства с технологией Big Data при решении научно-исследовательских задач;

- применять модели машинного обучения для решения научно-исследовательских задач в различных предметных областях.

владеть:

- навыками работы с программным обеспечением для обработки больших объемов данных;

- понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов;

- профессиональной терминологией в области больших данных и машинного обучения;

- навыками практического применения методов интеллектуального анализа больших данных и машинного обучения

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Дисциплина «Интеллектуальный анализ данных при проведении научных исследований» относится к элективной дисциплине учебного плана подготовки аспирантов по программе аспирантуры «Системный анализ, управление и обработка информации» по научной специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Интернет-адрес сайта курса: <https://edu.spbftu.ru>.

Дисциплина «Интеллектуальный анализ данных при проведении научных исследований» осваивается аспирантами на лекционных занятиях, а также в ходе самостоятельной работы.

В ходе лекций обучающимся рекомендуется:

- вести конспектирование учебного материала;
- обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации по их применению;
- задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Дисциплина предусматривает самостоятельную работу, в рамках которой следует изучить часть вопросов дисциплины самостоятельно, а также выполнить подготовку к промежуточной аттестации – зачету. Самостоятельная работа аспирантов по изучению отдельных тем дисциплины включает проработку печатных изданий и интернет-источников, анализ теоретического материала, подготовку к контролю знаний.

Текущий контроль знаний аспирантов по дисциплине проводится в виде контрольного опроса.

Результат промежуточной аттестации по дисциплине аспирант может получить в процессе сдачи зачета.

В процессе подготовки к зачету аспирантам следует проработать материалы лекций и рекомендуемую литературу. В ходе приема зачета оцениваются

обобщенные результаты обучения по дисциплине: владение теоретическими вопросами дисциплины, оценка умений и навыков, приобретенных в ходе освоения дисциплины.

Зачет сдается в устной форме. Предлагаемые вопросы соответствуют вопросам, подготовленным преподавателем для промежуточной аттестации.

Критерии оценки ответа аспиранта на зачете, а также форма его проведения доводятся преподавателем до сведения аспирантов до начала зачета.

**СВЕДЕНИЯ О ПЕРЕУТВЕРЖДЕНИИ
РАБОЧЕЙ ПРОГРАММЫ**

на 20 ____ / 20 ____ учебный год

Рабочая программа переутверждена без изменений / с изменениями на заседании кафедры _____

протокол № ____ от « ____ » _____ 20 ____ г.

(протокол изменений на 20 ____ / 20 ____ учебный год прилагается).

Заведующий кафедрой _____ / _____ /
(подпись) (расшифровка подписи)

на 20 ____ / 20 ____ учебный год

Рабочая программа переутверждена без изменений / с изменениями на заседании кафедры _____

протокол № ____ от « ____ » _____ 20 ____ г.

(протокол изменений на 20 ____ / 20 ____ учебный год прилагается).

Заведующий кафедрой _____ / _____ /
(подпись) (расшифровка подписи)

на 20 ____ / 20 ____ учебный год

Рабочая программа переутверждена без изменений / с изменениями на заседании кафедры _____

протокол № ____ от « ____ » _____ 20 ____ г.

(протокол изменений на 20 ____ / 20 ____ учебный год прилагается).

Заведующий кафедрой _____ / _____ /
(подпись) (расшифровка подписи)