

ВЫСШАЯ ИНЖИНИРИНГОВАЯ ШКОЛА

ОДОБРЕНО УМС ВИШ

Протокол № 132/15-12-22

от 15.12.2022 г.

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**  
**ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ (DATA MINING)**

Направление подготовки  
(специальность)

[1] 09.04.02 Информационные системы и  
технологии

Семестр	Трудоемкость, кред.	Общий объем курса, час.	Лекции, час.	Практич. занятия, час.	Лаборат. работы, час.	В форме практической подготовки/В СРС, час.	КСР, час.	Форма(ы) контроля, экс./зач./КР/КП
4	2	72	10	10	0	52	0	3
Итого	2	72	10	10	0	52	0	

## АННОТАЦИЯ

В курсе «Интеллектуальный анализ данных (Data mining)» представлены технологии интеллектуальной обработки данных (т.н. “ Data Mining ”) и возможности их применения для анализа документов. Теоретическая часть курса лежит в области инновационных методов анализа сложных и разнородных данных. В курсе рассмотрены история методов анализа разнородных данных, проблематика и подходы к анализу графических и текстовых данных, “Big Data” и перспектива анализа больших массивов данных, существующие методы Data Mining, этапы анализа данных, применяемые техники: классификация, кластеризация, прогнозирование и визуализация. Практическая значимость курса представляется в возможности применения техник Data Mining и специализированных инструментов (программного обеспечения) в исследованиях.

### 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Целью курса «Интеллектуальный анализ данных» является формирование представления о типах задач в области интеллектуального анализа данных (Data Mining) и методах их решения. У магистров формируются навыки решения практических задачи анализа данных, необходимые в процессе их профессиональной деятельности.

### 2. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП ВО

Для изучения данной дисциплины студент должен иметь компетенции, полученные при изучении предшествующих дисциплин: информатики, языков программирования, математической логики и баз данных. В результате освоения дисциплины студенты смогут применить полученные теоретические и практические знания при прохождении научно-исследовательской практики и подготовке отчета по практике, а также при написании магистерской диссертации.

### 3. ФОРМИРУЕМЫЕ КОМПЕТЕНЦИИ И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Универсальные и(или) общепрофессиональные компетенции:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-3 [1] – Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями	З-ОПК-3 [1] – Знать: подходы к анализу профессиональной информации и подготовки аналитических обзоров У-ОПК-3 [1] – Уметь: структурировать профессиональную информацию, выделять в ней главное, и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями В-ОПК-3 [1] – Владеть: навыками анализа профессиональной информации и подготовки аналитических обзоров.

ОПК-6 [1] – Способен использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий	З-ОПК-6 [1] – Знать: современные информационные технологии в области системной инженерии. У-ОПК-6 [1] – Уметь: использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации. В-ОПК-6 [1] – Владеть: навыками использования информационных технологий в системной инженерии
УКЦ-2 [1] – Способен к самообучению, самоактуализации и саморазвитию с использованием различных цифровых технологий в условиях их непрерывного совершенствования	З-УКЦ-2 [1] – Знать основные цифровые платформы, технологи и интернет ресурсы используемые при онлайн обучении У-УКЦ-2 [1] – Уметь использовать различные цифровые технологии для организации обучения В-УКЦ-2 [1] – Владеть навыками самообучения, самоактуализации и саморазвития с использованием различных цифровых технологий

Профессиональные компетенции в соответствии с задачами и объектами (областями знаний) профессиональной деятельности:

<b>Задача профессиональной деятельности (ЗПД)</b>	<b>Объект или область знания</b>	<b>Код и наименование профессиональной компетенции; Основание (профессиональный стандарт-ПС, анализ опыта)</b>	<b>Код и наименование индикатора достижения профессиональной компетенции</b>
научно-исследовательский			
Выбор инструментов научного исследования изучаемых областей и объектов, включая элементы системного и цифрового моделирования объектов, экспериментальные методы и методики, методы обработки полученных результатов и выявления закономерностей, в том числе опирающиеся на сквозные цифровые технологии.	Набор инструментов научного исследования и применяемых цифровых продуктов	ПК-1 [1] - Способен осуществлять сбор, обработку и анализ научно-технической информации отечественного и зарубежного опыта по тематике исследования  <i>Основание:</i> Профессиональный стандарт: 06.015, 40.011	З-ПК-1[1] - Знать: современное состояние отечественных и зарубежных исследований и разработок по заданной тематике. ; У-ПК-1[1] - Уметь: осуществлять сбор, обработку и анализ научно-технической информации по тематике исследований.; В-ПК-1[1] - Владеть: современными методами сбора, обработки и анализа научно-технической информации
Планирование и организация исследований и	Календарный план научного исследования, смета	ПК-2 [1] - Способен планировать, организовывать и	З-ПК-2[1] - Знать: подходы и методы планирования,

<p>разработок на базе сквозных цифровых технологий в области информационных систем, систем связи, информационных и коммуникационных технологий, а также цифровых технологий сложных инженерных объектов.</p> <p>Проведение отдельных элементов и этапов, а также полных циклов исследовательских работ с применением выбранного инструментария применительно к объектам исследования в области информационных систем, систем связи, информационных и коммуникационных технологий, а также цифровых технологий сложных инженерных объектов.</p>	<p>научного исследования, бизнес-план внедрения результатов научного исследования</p>	<p>проводить прикладные и фундаментальные научные исследования в области информационных систем и технологий</p> <p><i>Основание:</i> Профессиональный стандарт: 06.015, 40.011</p>	<p>организации и проведения фундаментальных и прикладных исследований в области ИСТ с использованием методов математического моделирования, статистического анализа и языков программирования. ; У-ПК-2[1] - Уметь: применять полученные знания для планирования и проведения фундаментальных и прикладных исследований в области ИСТ.; В-ПК-2[1] - Владеть: фундаментальными знаниями и прикладными навыками планирования, организации и проведения прикладных и фундаментальные научные исследований в области ИСТ</p>
<p>Обработка результатов выполненных исследований, анализ результатов исследования, формирование выводов и заключений, подтверждение или опровержение сформулированных гипотез в сферах исследования информационных систем, систем связи, информационных и</p>	<p>Научные гипотезы, научные обзоры, отчеты и публикации, результаты интеллектуальной деятельности.</p>	<p>ПК-5 [1] - Способен проводить анализ результатов исследования, осуществлять выбор оптимальных решений, подготавливать и составлять обзоры, отчеты и научные публикации по результатам исследований</p> <p><i>Основание:</i> Профессиональный стандарт: 06.016, 40.011</p>	<p>З-ПК-5[1] - Знать: методы системного анализа для комплексной оценки результатов исследований и поиска оптимальных решений. ; У-ПК-5[1] - Уметь: осуществлять выбор оптимальных решений на основе методов системного анализа результатов исследований. ; В-ПК-5[1] - Владеть: навыками подготовки</p>

<p>коммуникационных технологий, а также цифровых технологий сложных инженерных объектов.</p> <p>Формирование отчетов о проведенных научно-исследовательских работах и подготовка публикаций, а также подготовка заявок на изобретения и регистрацию программного обеспечения по результатам выполненных исследований и разработок в сферах исследования информационных систем, систем связи, информационных и коммуникационных технологий, а также цифровых технологий сложных инженерных объектов.</p>			<p>и оформления обзоров, отчетов и научных публикаций по результатам исследований.</p>
---	--	--	--

#### 4. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Разделы учебной дисциплины, их объем, сроки изучения и формы контроля:

№ п.п	Наименование раздела учебной дисциплины	Недели	Лекции/ Практи. (семинары) / Лабораторные работы, час.	Обязат. текущий контроль (форма*, неделя)	Максимальный балл за раздел**	Аттестация раздела (форма*, неделя)	Индикаторы освоения компетенции
	<i>4 Семестр</i>						
1	Введение интеллектуальный анализ данных <span style="float: right;">в</span>	1-8	5/5/0		25	КИ-8	З-ПК-1, У-ПК-1, У-ПК-2, В-ПК-2
2	Анализ представление данных <span style="float: right;">и</span>	9-13	5/5/0		25	КИ-13	З-ОПК-

							6, У- ОПК- 6, В- ОПК- 6, 3-ПК- 1, У- ПК-1, В- ПК-1
	<i>Итого за 4 Семестр</i>		10/10/0		50		
	<b>Контрольные мероприятия за 4 Семестр</b>				50	3	3- ОПК- 3, У- ОПК- 3, В- ОПК- 3, 3- ОПК- 6, У- ОПК- 6, В- ОПК- 6, 3-ПК- 1, У- ПК-1, В- ПК-1, 3-ПК- 2, У- ПК-2, В- ПК-2, 3-ПК- 5, У- ПК-5, В- ПК-5, 3- УКЦ-

							2, У- УКЦ- 2, В- УКЦ- 2
--	--	--	--	--	--	--	---

\* – сокращенное наименование формы контроля

\*\* – сумма максимальных баллов должна быть равна 100 за семестр, включая зачет и (или) экзамен

Сокращение наименований форм текущего контроля и аттестации разделов:

Обозначение	Полное наименование
КИ	Контроль по итогам
З	Зачет

### КАЛЕНДАРНЫЙ ПЛАН

Недели	Темы занятий / Содержание	Лек., час.	Пр./сем., час.	Лаб., час.
	<i>4 Семестр</i>	10	10	0
<b>1-8</b>	<b>Введение в интеллектуальный анализ данных</b>	5	5	0
1 - 3	<b>Введение в интеллектуальный анализ данных. Данные.</b> Набор данных и их атрибутов. Переменная. Значение. Генеральная совокупность. Измерения. Шкалы. Порядковая шкала. Интервальная шкала. Относительная шкала. Типы наборов данных. Форматы хранения данных. Базы данных. Основные положения. Системы управления базами данных, СУБД. Метаданные	Всего аудиторных часов		
		1	1	0
		Онлайн		
		0	0	0
4 - 6	<b>Методы и стадии Data Mining.</b> Классификация стадий Data Mining. Классификация методов Data Mining. Статистические методы Data mining. Кибернетические методы Data Mining. Свойства методов Data Mining.	Всего аудиторных часов		
		2	2	0
		Онлайн		
		0	0	0
7 - 8	<b>Задачи Data Mining. Сферы применения.</b> Классификация. Кластеризация. Ассоциация. Прогнозирование. Оценивание. Визуализация. Классификация задач Data Mining: обучение с учителем; обучение без учителя;	Всего аудиторных часов		
		2	2	0
		Онлайн		
		0	0	0
<b>9-13</b>	<b>Анализ и представление данных</b>	5	5	0
9 - 10	<b>Методы классификации и прогнозирования.</b> Деревья решений. Преимущества деревьев решений. Процесс конструирования дерева решений. Алгоритмы, реализующие деревья решений: CART, C4.5, CHAID, CN2 и др. Метод опорных векторов. Задача классификации новых объектов. Решение задачи прогнозирования. Байесовская классификация. Нейронные сети.	Всего аудиторных часов		
		1	1	0
		Онлайн		
		0	0	0

10 - 11	<b>Методы кластерного анализа</b> Иерархические методы. Иерархические агломеративные методы . Иерархические дивизимные методы . Меры сходства. Методы объединения. Иерархический кластерный анализ. Определение количества кластеров. Итеративные методы. Алгоритм k-средних . Алгоритм	Всего аудиторных часов		
		1	1	0
		Онлайн		
12 - 13	<b>Способы визуального представления данных.</b> Способы визуального представления данных. Методы визуализации. Визуализация инструментов Data Mining. Визуализация Data Mining моделей. Диаграмма рассеивания. Таблица сопряженности. График изменения величины ошибки. Представление данных в одном, двух и трех измерениях.	Всего аудиторных часов		
		1	1	0
		Онлайн		
14 - 15	<b>Инструменты Data Mining.</b> SAS Enterprise Miner. Система PolyAnalyst. Архитектура системы. PolyAnalyst Workplace. Эволюционное программирование. Программные продукты Cognos и система STATISTICA Data Miner. Инструменты Oracle Data Mining и Deductor.	Всего аудиторных часов		
		2	2	0
		Онлайн		
		0	0	0

Сокращенные наименования онлайн опций:

Обозначение	Полное наименование
ЭК	Электронный курс
ПМ	Полнотекстовый материал
ПЛ	Полнотекстовые лекции
ВМ	Видео-материалы
АМ	Аудио-материалы
Прз	Презентации
Т	Тесты
ЭСМ	Электронные справочные материалы
ИС	Интерактивный сайт

## 5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

С целью формирования и развития профессиональных навыков студентов в дисциплине используются как традиционные - лекции, семинары, так и интерактивные формы проведения занятий: доклады и презентации с их обсуждением, дискуссии с разбором конкретных ситуаций в сочетании с внеаудиторной работой. В соответствии со спецификой университета в процессе преподавания дисциплины методически целесообразно в каждом разделе выделить наиболее важные темы и акцентировать на них внимание обучающихся.

## 6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Фонд оценочных средств по дисциплине обеспечивает проверку освоения планируемых результатов обучения (компетенций и их индикаторов) посредством мероприятий текущего, рубежного и промежуточного контроля по дисциплине.

Связь между формируемыми компетенциями и формами контроля их освоения представлена в следующей таблице:

Компетенция	Индикаторы освоения	Аттестационное мероприятие (КП 1)
ОПК-3	З-ОПК-3	З
	У-ОПК-3	З
	В-ОПК-3	З
ОПК-6	З-ОПК-6	З, КИ-13
	У-ОПК-6	З, КИ-13
	В-ОПК-6	З, КИ-13
ПК-1	З-ПК-1	З, КИ-8, КИ-13
	У-ПК-1	З, КИ-8, КИ-13
	В-ПК-1	З, КИ-13
ПК-2	З-ПК-2	З
	В-ПК-2	З, КИ-8
	У-ПК-2	З, КИ-8
ПК-5	З-ПК-5	З
	У-ПК-5	З
	В-ПК-5	З
УКЦ-2	З-УКЦ-2	З
	У-УКЦ-2	З
	В-УКЦ-2	З

### Шкалы оценки образовательных достижений

Шкала каждого контрольного мероприятия лежит в пределах от 0 до установленного максимального балла включительно. Итоговая аттестация по дисциплине оценивается по 100-балльной шкале и представляет собой сумму баллов, заработанных студентом при выполнении заданий в рамках текущего и промежуточного контроля.

Итоговая оценка выставляется в соответствии со следующей шкалой:

Сумма баллов	Оценка по 4-ех балльной шкале	Оценка ECTS	Требования к уровню освоению учебной дисциплины
90-100	5 – «отлично»	A	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, использует в ответе материал монографической литературы.
85-89	4 – «хорошо»	B	Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос.
75-84		C	
70-74		D	
65-69		E	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала,
60-64			

			но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала.
Ниже 60	2 – <i>«неудовлетворительно»</i>	F	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

## **7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ**

### ОСНОВНАЯ ЛИТЕРАТУРА:

1. ЭИ Д 95 Логический анализ данных : учебное пособие, Санкт-Петербург: Лань, 2020
2. ЭИ М 17 Технологии интеллектуального анализа данных : учебное пособие, Санкт-Петербург: Лань, 2019

### ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА:

### ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ:

Специальное программное обеспечение не требуется

### LMS И ИНТЕРНЕТ-РЕСУРСЫ:

<https://online.mephi.ru/>

<http://library.mephi.ru/>

## **8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ**

Специальное материально-техническое обеспечение не требуется

## **9. УЧЕБНО-МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ СТУДЕНТОВ**

Обучение носит прикладной характер. Помимо теоретического материала, используются игровые технологии, требующие присутствия на занятиях.

Успешное освоение дисциплины требует от студентов активной работы во время практических занятий, выполнения всех учебных заданий преподавателя, ознакомления с

базовыми учебниками, основной и дополнительной литературой, а также предполагает творческое участие студента путем планомерной, повседневной работы.

Практические задания являются необходимым элементом данного модуля. Значимость успешного выполнения практических заданий определяется тем, что во время прохождения студенты получают необходимые практические навыки и умения работы с современным цифровым инструментарием. Основная цель практического обучения состоит в формировании и закреплении первичных теоретических знаний и профессиональных навыков. В ходе практических занятий обычно формируется теоретическая и практическая база будущей профессиональной деятельности.

## **10. УЧЕБНО-МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРЕПОДАВАТЕЛЕЙ**

Целью работы преподавателя должно быть эффективное восприятие материала слушателями.

Со стороны преподавателя должен быть установлен контакт со студентами, и они должны быть информированы о порядке прохождения курса, его особенностях, учебно-методическом обеспечении по дисциплине.

При подготовке к практическому занятию преподаватель готовит план его проведения, знакомится с новыми публикациями по теме.

Преподаватель предоставляет учащимся обратную связь о выполненных практических заданиях, ставит перед учащимися четкие цели и представляет новый материал с той степенью подробности изложения, чтобы материал был усвоен, но учащиеся не чувствовали себя перегруженными. Учащимся предоставляется инструкции и стратегии для выполнения практического задания. Когда учащиеся работают индивидуально, преподаватель контролирует их деятельность.

Автор(ы):

Дюмин Александр Александрович