

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский ядерный университет «МИФИ»

ФАКУЛЬТЕТ БИЗНЕС–ИНФОРМАТИКИ И УПРАВЛЕНИЯ КОМПЛЕКСНЫМИ
СИСТЕМАМИ

КАФЕДРА ЭКОНОМИКИ И МЕНЕДЖМЕНТА В ПРОМЫШЛЕННОСТИ

ОДОБРЕНО УМС ФБИУКС

Протокол № 24/08

от 22.08.2024 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

МАТЕМАТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Направление подготовки
(специальность)

[1] 38.04.05 Бизнес-информатика

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, экз./зач./КР/КП
2	2	72	15	0	15		42	0	3
Итого	2	72	15	0	15	15	42	0	

АННОТАЦИЯ

В рамках данной дисциплины рассматриваются основные понятия, задачи и современные методы и технологии в области интеллектуального анализа данных, с использованием методов машинного обучения, и современных инструментальных средств.

Рассматриваются современные подходы к анализу данных, изучаются различия между фактами, информацией и знаниями, основные стандарты анализа данных, такие как CRISP-DM. Даются основные сведения по методам подготовки данных, включая очистку от выбросов и пропущенных значений, по методам и способам визуализации данных, описательного анализа и анализа связей. Также рассматриваются основы кластерных, классификационных, ассоциативных методов, методов обработки текстовой информации, критерии эффективности применения этих методов.

В качестве основного инструмента выполнения заданий в рамках курса, используется аналитическая платформа Rapid Miner.

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Целью изучения дисциплины являются формирование у студентов теоретико-методологических знаний и закрепление профессиональных навыков в области решения прикладных задач в различных сферах государственной, корпоративной и общественной деятельности на основе методов и технологий обработки текстов на естественном языке.

2. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП ВО

Учебная дисциплина относится к дисциплинам базовой части обще-профессионального модуля.

Содержание программы «Математические и инструментальные методы машинного обучения» представляет собой современную теорию и практику использования методов машинного обучения для прикладного анализа данных. Предполагается наличие у слушателей базовых знаний по математической статистике, методам линейного программирования, теории информации, информатике.

3. ФОРМИРУЕМЫЕ КОМПЕТЕНЦИИ И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Универсальные и(или) общепрофессиональные компетенции:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
--------------------------------	--

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

Задача профессиональной деятельности (ЗПД)	Объект или область знания	Код и наименование профессиональной компетенции; Основание	Код и наименование индикатора достижения профессиональной
--	---------------------------	---	---

		(профессиональный стандарт-ПС, анализ опыта)	компетенции
аналитический			
Исследование и разработка методов и инструментов создания и развития цифровых предприятий, цифровых двойников и сквозных цифровых технологий	Методы и инструменты создания и развития цифровых предприятий и цифровых двойников и сквозных цифровых технологий	ПК-2.3 [1] - Способен применять методы машинного обучения и интеллектуального анализа информации для развития цифровых предприятий, цифровых двойников и сквозных цифровых технологий <i>Основание:</i> Профессиональный стандарт: 24.078, Анализ опыта: По согласованию с Заказчиками образовательной программы Трудовая функция: «Выполнение деятельности по использованию методов машинного обучения и интеллектуального анализа информации для развития цифровых предприятий, цифровых двойников и сквозных цифровых технологий»	3-ПК-2.3[1] - Знать: Технологии, методы и инструментальные средства обработки больших данных; Основные методы интеллектуальной обработки данных (Data Mining), в том числе и больших данных; Математические и инструментальные методы машинного обучения; Методы интеллектуального анализа текстов; Основные положения цифровой экономики; Основные направления Стратегии внутренней цифровизации Госкорпорации «Росатом»; Основные положения дорожных карт сквозных цифровых технологий.; У-ПК-2.3[1] - Уметь: Использовать модели данных, адаптированных к технологиям больших данных; Пользоваться методами и инструментами получения, хранения, передачи, обработки больших данных; Применять методы интеллектуальной обработки данных (Data Mining), в том числе и больших данных, в зависимости от предметной области и специфики

			деятельности организации; Применять математические и инструментальные методы машинного обучения в зависимости от поставленной задачи; Применять интеллектуального анализа текстов в зависимости от поставленной задачи.; В-ПК-2.3[1] - Владеть: Разрабатывать модели данных, адаптированных к технологиям больших данных; Разрабатывать предложения по развитию и совершенствованию системы получения, хранения, передачи, обработки больших данных; Использовать математические и инструментальные методы машинного обучения и интеллектуального анализа информации в целях развития цифровых предприятий, цифровых двойников и сквозных цифровых технологий.
научно-исследовательский			
Исследование, разработка и внедрение новых моделей, методов и средств в области экономики, управления и ИКТ для совершенствования архитектуры предприятия	Архитектура предприятия (бизнес-архитектура, архитектура информации, архитектура приложений, инфраструктура)	ПК-1 [1] - Способен проводить исследования и поиск новых моделей и методов в области экономики, управления и ИКТ для совершенствования архитектуры предприятия и выявления инновации;	З-ПК-1[1] - Знать: современные ИТ, широкий кругозор в области ИТ, понимание соотношения целей и путей реализации стратегии развития ИТ; предметная функциональная область применения

		<i>Основание:</i> Профессиональный стандарт: 06.012	ИТ; принципы инновационной деятельности ; У-ПК-1[1] - Уметь: определять возможности использования инноваций ИТ в стратегическом управлении; интегрировать ИТ в деятельность организации; В-ПК-1[1] - Владеть навыками: формирование целей, приоритетов и ограничений формирования ИТ в создание и реализацию инновационной стратегии и изменение их по мере изменения внешних условий и внутренних потребностей; организация работы персонала и выделение ресурсов для формирования вклада ИТ в создание и реализацию инновационной стратегии; контроль формирования вклада ИТ в создание и реализацию инновационной стратегии; анализ формирования вклада ИТ в создание и реализацию инновационной стратегии, целей, приоритетов и ограничений процесса и выполнение управленческих действий по результатам анализа
--	--	---	---

4. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

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

№ п.п	Наименование раздела учебной дисциплины	Недели	Лекции/ Практ. (семинары) / Лабораторные работы, час.	Обязат. текущий контроль (форма*, неделя)	Максимальный балл за раздел**	Аттестация раздела (форма*, неделя)	Индикаторы освоения компетенции
	<i>2 Семестр</i>						
1	Первый раздел	1-8	8/0/8	Т-7 (5), ЛР-8 (20)	25	КИ-8	3-ПК-1, 3-ПК-2.3, 3-ПК-2.3
2	Второй раздел	9-15	7/0/7	Т-14 (5), ЛР-15 (20)	25	КИ-15	3-ПК-1, У-ПК-1, В-ПК-1, 3-ПК-2.3, У-ПК-2.3, В-ПК-2.3, 3-ПК-2.3, У-ПК-2.3, В-ПК-2.3
	<i>Итого за 2 Семестр</i>		15/0/15		50		
	Контрольные мероприятия за 2 Семестр				50	3	3-ПК-1, У-ПК-1, В-ПК-1, 3-ПК-2.3, У-ПК-2.3, В-ПК-2.3, 3-ПК-2.3, У-ПК-2.3, В-ПК-2.3

* – 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
Т	Тестирование
ЛР	Лабораторная работа
КИ	Контроль по итогам
З	Зачет

КАЛЕНДАРНЫЙ ПЛАН

Недели	Темы занятий / Содержание	Лек., час.	Пр./сем., час.	Лаб., час.
--------	---------------------------	------------	----------------	------------

	2 Семестр	15	0	15
1-8	Первый раздел	8	0	8
1 - 8	Базовые методы машинного обучения Базовые методы машинного обучения Отраслевой стандарт CRISP-DM. Методы подготовки данных. Методы описательного анализа. Статистические характеристики выборки. Понятие количественной и качественной шкал. Методы поиска аномалий и выбросов. Методы восстановления пропущенных значений. Методы факторного анализа. Метод главных компонент. Метод TSNE. Визуализация данных. Методы анализа связей. Таблицы сопряженности. Статистика Хи-квадрат. Методы корреляционно-регрессионного анализа. Коэффициенты корреляции. Множественная корреляция. Множественная линейная регрессия. Ошибки линейной регрессии. Коэффициент детерминации.	Всего аудиторных часов		
		8	0	8
		Онлайн		
		0	0	0
9-15	Второй раздел	7	0	7
9 - 15	Продвинутое методы машинного обучения Методы кластерного анализа. Метрики расстояния и меры сходства между объектами. Метод k-средних. Метод иерархической агломеративной кластеризации. Метод DBSCAN. Метрики качества кластеризации. Меж-и внутрикластерные расстояния. Индекс Девиса-Булдена. Коэффициент силуэта. Методы классификации. Логистическая регрессия. Наивный байесовский подход. Деревья решений. Случайный лес. Матрица запутанности. Точность и полнота, F-мера. Ансамбли алгоритмов. Многослойный перцептрон. Методы ассоциативного анализа. Задача анализа рыночной корзины (Market Basket Analysis). Метрики поддержки, достоверности, lift, leverage. Метод Apriori. Метод FP-growth. Методы обработки текстовых данных. Токенизация и векторизация текстов. Стемминг. Лемматизация. Регулярные выражения. Метрика Tf-Idf	Всего аудиторных часов		
		7	0	7
		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

ТЕМЫ ЛАБОРАТОРНЫХ РАБОТ

Недели	Темы занятий / Содержание
	2 Семестр

1 - 2	Лабораторная работа №1. Изучение методов подготовки данных.
3 - 4	Лабораторная работа №2. Изучение методов факторного анализа.
5 - 6	Лабораторная работа №3. Изучение методов визуализации данных.
7 - 8	Лабораторная работа №4. Изучение методов корреляционно-регрессионного анализа.
9 - 10	Лабораторная работа №5. Изучение методов кластерного анализа
11 - 12	Лабораторная работа №6. Изучение методов классификации.
13 - 14	Лабораторная работа №7. Изучение методов ассоциативного анализа
15	Лабораторная работа №8. Изучение методов анализа текстовых данных

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При реализации программы дисциплины используются различные интерактивные образовательные технологии – занятия проводятся в виде анализа деловых ситуаций с визуализацией с использованием технических средств обучения.

Самостоятельная работа студентов подразумевает под собой проработку теоретического материала с использованием рекомендуемой литературы для выполнения практических заданий и подготовке к итоговой форме контроля, а также интерактивные формы обучения в виде выполнения тестов с помощью электронных учебных элементов для системы электронного обучения ИНФОМИФИСТ.

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)
ПК-1	З-ПК-1	З, КИ-8, КИ-15, Т-7, ЛР-8, Т-14, ЛР-15
	У-ПК-1	З, КИ-15, Т-14, ЛР-15
	В-ПК-1	З, КИ-15, Т-14, ЛР-15
ПК-2.3	З-ПК-2.3	З, КИ-8, КИ-15, Т-7, ЛР-8, Т-14, ЛР-15
	У-ПК-2.3	З, КИ-15, Т-14, ЛР-15
	В-ПК-2.3	З, КИ-15, Т-14, ЛР-15

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 – <i>«отлично»</i>	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 – <i>«хорошо»</i>	B	Оценка «хорошо» выставляется студенту, если он твёрдо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос.
75-84		C	
70-74		D	
65-69	3 – <i>«удовлетворительно»</i>	E	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала.
60-64			
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 П 37 Машинное обучение : учебное пособие для вузов, Платонов А. В., Москва: Юрайт, 2022
2. ЭИ М 38 Машинное обучение и большие данные : учеб. пособие, Марухина О.В. [и др.], Санкт-Петербург: ГУАП, 2021
3. ЭИ Б 53 Системы искусственного интеллекта : учебное пособие для вузов, Бессмертный И. А., Москва: Юрайт, 2022

4. ЭИ Е 51 Эконометрика : учебник для вузов, Елисеева И. И., Москва: Юрайт, 2022

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА:

1. ЭИ Н61 Анализ данных : учебное пособие для вузов, Румянцев В.П., Низаметдинов Ш.У., Москва: НИЯУ МИФИ, 2012
2. 004 Т98 Анализ данных на компьютере : , Тюрин Ю.Н., Макаров А.А., М.: ИНФРА-М, 2003

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

При реализации программы дисциплины используются различные интерактивные образовательные технологии с использованием технических средств обучения.

Во время лекционных занятий по дисциплине студент должен уметь сконцентрировать внимание на рассматриваемых проблемах и включить в работу все виды памяти: словесную, образную и моторно-двигательную. Для этого ему необходимо конспектировать материал, излагаемый преподавателем. Во время конспектирования в работу включается моторно-двигательная память, позволяющая эффективно усвоить лекционный материал. Весь иллюстративный материал, представляемый на лекции (на слайдах, на доске, в раздаточном материале) также должен быть зафиксирован в конспекте лекций. Каждому студенту необходимо помнить о том, что конспектирование лекции – это не диктант. Студент должен выделять главное и фиксировать основные моменты.

Важное место в учебном процессе занимают лабораторные работы, призванные закреплять полученные студентами теоретические знания. Перед лабораторной работой студенту необходимо восстановить в памяти теоретический материал по теме практического занятия. Для этого следует обратиться к первоисточникам, конспекту лекций, настоящим методическим указаниям. Каждое занятие начинается с повторения теоретического материала по соответствующей теме. Студенты должны уметь чётко ответить на вопросы, поставленные преподавателем. По характеру ответов преподаватель делает вывод о том, насколько тот или иной студент готов к выполнению различных заданий.

Самостоятельная работа студентов предусматривает: ознакомление с рекомендованной литературой и презентациями лекций, в том числе с использованием Интернет; повторение

пройденного на лекциях материала; работу над электронными тестами; а также интерактивные формы обучения в виде выполнения тестов с помощью электронных учебных элементов для системы электронного обучения ИНФОМИФИСТ.

Оценочными средствами рубежного контроля является Контроль Итогов (КИ) - означающий выставление баллов на основании результатов Текущего контроля отдельно для первого раздела и второго раздела. Текущий контроль проводится на основе проверки теста, выполнения лабораторных работ, подготовки и защиты курсовой работы.

В конце освоения дисциплины студент сдает экзамен.

10. УЧЕБНО-МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРЕПОДАВАТЕЛЕЙ

Целью изучения дисциплины являются формирование у будущих магистров теоретико-методологических знаний и закрепление профессиональных навыков в области решения прикладных задач в различных сферах государственной, корпоративной и общественной деятельности на основе учета закономерностей становления и развития цифровой экономики, общих свойств информации и особенностей информационных процессов.

Интерактивные занятия со студентами проводятся с использованием презентаций, которые публикуются в системе ИНФОМИФИСТ.

Видеозапись каждого проведенного занятия выкладывается на портал home.merphi.ru.

Для наиболее эффективного изучения предусмотрена самостоятельная проработка студентами отдельных тем, освоение которых проверяется при защите курсовой работы. Целесообразно для увеличения времени проработки важных тем предусмотреть рассмотрение отдельных вопросов в форме дискуссий и диспутов. Кроме того, необходимо предусмотреть дополнительные консультации по сложным темам. Самостоятельная работа студентов включает в себя так же интерактивные формы обучения в виде выполнения тестов с помощью электронных учебных элементов для системы электронного обучения ИНФОМИФИСТ.

Оценочными средствами рубежного контроля является Контроль Итогов (КИ) - означающий выставление баллов на основании результатов Текущего контроля отдельно для первого раздела и второго раздела. Текущий контроль проводится с устного опроса и подготовки реферата. Текущий контроль проводится на основе проверки тестов, выполнения лабораторных работ.

Итоговая аттестация проводится в форме зачёта.

Автор(ы):

Киреев Василий Сергеевич, к.т.н.

