

**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«УНИВЕРСИТЕТ УПРАВЛЕНИЯ «ТИСБИ»**

Кафедра информационных технологий

Утверждаю
Зав. кафедрой
О.В.Федорова
Протокол заседания
кафедры № 10
от 06.04.2026

Рабочая программа дисциплины

Наименование дисциплины	Введение в машинное обучение
Направление подготовки	09.03.01 Информатика и вычислительная техника
Направленность (профиль) образовательной программы	Инженерия информационных систем
Год набора	2025, 2026

Составители:

Профессор лаборатории машинного обучения и представления данных Хан А.М.
АНО ВО «Университет Иннополис»

Казань

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Разделы рабочей программы:

1. Краткая характеристика дисциплины (модуля).
2. Перечень планируемых результатов обучения.
3. Место дисциплины (модуля) в структуре основной профессиональной образовательной программы высшего образования (ОПОП ВО).
4. Объем, структура и содержание дисциплины (модуля).
5. Учебно-методическое обеспечение.
6. Материально-техническое обеспечение.
7. Оценочные и методические материалы.
8. Дополнительные сведения.

Программа составлена в соответствии с требованиями ОПОП ВО Университета Иннополис по данному направлению подготовки.

Форма обучения	Семестр	Трудоемкость дисциплины в		Контактная работа		Контроль, час.	Самостоятельная работа, час.	Форма промежуточной аттестации (экз./диф.зач./зач.)
		зачетных единиц	академических часах	Лекции, час.	Семинарские (практические) занятия, час.			
очная	5	4	144	30	30	18	66	экзамен

1. КРАТКАЯ ХАРАКТЕРИСТИКА ДИСЦИПЛИНЫ (МОДУЛЯ)

Изучение дисциплины (модуля) «Introduction to Machine Learning / Введение в машинное обучение» обеспечивает формирование и развитие компетенций обучающихся в области машинного обучения, их применение для решения различных прикладных задач в рамках профессиональной деятельности. В ходе освоения дисциплины обучающиеся рассматривают парадигмы машинного обучения, подходы и алгоритмы машинного обучения.

2. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ

Целью освоения дисциплины (модуля) является формирование знаний, умений и навыков в рамках профессиональной деятельности, позволяющих успешно работать в избранной отрасли, а также обладать компетенциями, способствующими социальной мобильности, устойчивости и конкурентоспособности обучающегося в условиях современных рыночных отношений. Задачами дисциплины являются изучение различных парадигм обучения, подходов и алгоритмов обучения, ознакомление с различными настройками обучения, приобретение навыков распознавать и определять показатели производительности, распознавать и определять популярные программные средства машинного обучения, формирование способности применять классификационные подходы к решению задач по машинному обучению с учителем, использовать кластеризацию подходов к решению задач по машинному обучению без учителя, овладение навыками и методами применения обучения ансамбля для повышения производительности модели и применения систематизации для улучшения генерализации модели, а также приобретение навыков использования алгоритмов углубленного обучения для решения реально существующих задач.

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

№ п/п	Код компетенции	Наименование компетенции	Индикаторы компетенций	Методы и технологии обучения, способствующие формированию компетенции
1.	ОПК-2	Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного	Знать: современные информационные технологии, принципы их работы и методы их использования при решении задач профессиональной деятельности. Уметь: выбирать современные информационные технологии и программные средства, в том числе отечественного производства, для решения задач профессио-	Метод дискуссии; Метод контрольных вопросов; Проектная технология; Информационно – коммуникацион-

		производства, и использовать их при решении задач профессиональной деятельности	нальной деятельности. Владеть: способами применения необходимых информационных технологий и программных средств, в том числе отечественного производства, при решении задач профессиональной деятельности.	ная технология; Групповая технология; Технология проблемного обучения; Кейс – технология
2.	ОПК-8	Способен разрабатывать алгоритмы и программы, пригодные для практического применения	Знать: основные языки программирования, операционные системы и оболочки, современные среды разработки программного обеспечения. Уметь: составлять алгоритмы, писать и отлаживать коды на языке программирования, тестировать работоспособность программы, интегрировать программные модули. Владеть: языком программирования, методами отладки и тестирования работоспособности программы.	Метод дискуссии; Метод контрольных вопросов; Проектная технология; Информационно – коммуникационная технология; Групповая технология; Технология проблемного обучения; Кейс – технология
3.	ПК-2	Способен выполнять работы по созданию (модификации) и сопровождению информационных систем	Знать: архитектуру, устройство и функционирование вычислительных и информационных систем; устройство и функционирование современных информационных систем; современные стандарты информационного взаимодействия систем; программные средства и платформы инфраструктуры информационных технологий организаций; современные подходы и стандарты автоматизации организации (например, CRM, MRP, ERP, ITIL, ITSM); методы выявления требований. Уметь: разрабатывать архитектуру, дизайн и прототипы информационной системы; разрабатывать метрики (количественные показатели) работы информационной системы; настраивать и создавать компоненты анализа программного решения информационной системы; исправлять дефекты и несоответствия в архитектуре и дизайне информационной системы; подтверждать исправление дефектов и несоответствий в коде информационной системы и документации к ней; проверять (верифицировать) архитектуру информационной системы, выполнять параметрическую настройку информационной системы; разрабатывать технологии обмена данными; разрабатывать базы данных информационной системы; осуществлять оптимизацию информационной системы; разрабатывать пользовательскую документацию; оценивать временные и материальные затраты на разработку информационной системы при заданных ресурсах.	Метод дискуссии; Метод контрольных вопросов; Проектная технология; Информационно – коммуникационная технология; Групповая технология; Технология проблемного обучения; Кейс – технология

			Владеть: методами оптимизации работы информационной системы и её компонентов; приемами интеграции информационной системы с техническим обеспечением заказчика.	
--	--	--	---	--

Знания: сформированы систематические знания различных парадигм обучения; широкого спектра подходов и алгоритмов обучения; различных настроек обучения; популярных программных средств машинного обучения; методологии определения целей и задач принятия решений и проведения экспериментальных исследований; моделей базовых информационных процессов и технологий, методов и средств их реализации.

Умения: сформированы умения описывать и объяснять различия между различными парадигмами обучения; описывать и объяснять различие между классификацией и регрессией; описывать и объяснять ядерные методы; описывать и объяснять нейронное или глубокое обучение; использовать базовые информационные процессы и технологии для проектирования и реализации информационных систем; формулировать постановку задач принятия решений; вырабатывать, анализировать и принимать решения о наиболее перспективных проектных решениях; корректно выражать и аргументированно обосновывать положения предметной области знания, интерпретировать полученные в практической деятельности результаты.

Навыки (владения): сформировано владение навыками применения классификационных подходов к решению задач по машинному обучению; навыками применения обучения ансамбля для повышения производительности модели; навыками применения систематизации для улучшения генерализации модели; навыками использования алгоритмов углубленного обучения для решения существующих задач; навыками поддержки принятия решений и планирования экспериментов и анализа их результатов; владение стандартными средствами базовых информационных процессов и технологий; навыками работы с электронными каталогами и представлять результат работы в форме, удобной для восприятия.

3. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОСНОВНОЙ ПРОФЕССИОНАЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ВЫСШЕГО ОБРАЗОВАНИЯ (ОПОП ВО)

Данная учебная дисциплина (модуль) включена в «Блок 1. Дисциплины (модули)» образовательной программы по направлению подготовки 09.03.01 Информатика и вычислительная техника, направленность (профиль) образовательной программы «Инженерия информационных систем» и относится к обязательной части программы, обеспечивает формирование общепрофессиональных, профессиональных компетенций. Дисциплина (модуль) «Introduction to Machine Learning / Введение в машинное обучение» позволяет обучающимся получить знания, умения и навыки для успешной профессиональной деятельности в области программного обеспечения компьютерных вычислительных систем и сетей, автоматизиро-

ванных систем обработки информации и управления. Дисциплина осваивается студентами при обучении в очной форме на 2 курсе в 5 семестре.

4. ОБЪЕМ, СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Объем дисциплины (модуля) составляет:

4 зачетные единицы, 144 академических часа, объем контактной работы — 60 академических часов.

Структура дисциплины (модуля):

№ раздела	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы			
		Очная форма			
		Контактная работа		Контроль, час.	Самостоятельная работа, час.
		Лекции, час.	Семинарские (практические) занятия, час.		
1.	Машинное обучение с учителем	7	7	0	16
2.	Дерево решений и обучение ансамбля	7	7	0	16
3.	Машинное обучение без учителя	8	8	0	17
4.	Углубленное изучение	8	8	0	17
	Промежуточная аттестация - экзамен			18	
ИТОГО:		30	30	18	66

Содержание дисциплины (модуля):

№ раздела	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам	Формируемые компетенции
1.	Машинное обучение с учителем	Введение в машинное обучение. Производные инструменты и функция затрат. Предварительная обработка данных. Линейная регрессия. Множественная линейная регрессия. Снижение градиента. Полиномиальная регрессия. Обратное отношение между смещением и дисперсией. Различие между классификацией и регрессией. Логистическая регрессия. Наивный классификатор Байеса. Байесовская сеть. Метод k-ближайших соседей. Метрики путаницы.	

		Показатели производительности. Упорядочение. Классификация на основе гиперплоскости. Алгоритм обучения Perceptron. Классификация максимальной разницы. Метод опорных векторов. Скользкие переменные. Лагранжийский метод опорных векторов. Ядерный способ.	ОПК-2; ОПК-8; ПК-2
2.	Дерево решений и обучение ансамбля	Деревья решений. Упаковка. Ускорение. Случайный лес. Adaboost.	
3.	Машинное обучение без учителя	Кластеризация методом k-средств. K-средства. Иерархическая кластеризация. Плотностный алгоритм кластеризации. пространственных данных с присутствием шума. Сдвиг среднего значения.	
4.	Углубленное изучение	Искусственные нейронные сети. Обратное распространение. Сверточные нейронные сети. Автокодер. Вариантальный автокодер. Генеративно-сопоставительная сеть.	

5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Университет Иннополис обеспечен необходимым комплектом лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства.

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронно-библиотечной системе (электронной библиотеке) и к электронной информационно-образовательной среде Университета (<https://my.university.innopolis.ru>). Электронно-библиотечная система и электронная информационно-образовательная среда обеспечивают возможность доступа, обучающегося из любой точки, в которой имеется доступ к информационно-телекоммуникационной сети «Интернет» (далее — сеть «Интернет»), как на территории Университета, так и вне его.

Также обучающимся обеспечен доступ (удаленный доступ) к современным профессиональным базам данных и информационным справочным системам.

Перечень учебно-методического обеспечения дисциплины (модуля):

Основная литература:

1. Hutter, F., Kotthoff, L. Automated Machine Learning. Methods, Systems, Challenges. — Publishing House: Springer Cham, 2019. — 219 p. — Режим доступа: <https://link.springer.com/book/10.1007/978-3-030-05318-5> — Полнотекстовая англоязычная БД Springer.

Дополнительная литература:

1. Li, H. Machine Learning Methods. — Publishing House: Springer Singapore, 2024. — 532 p. — Режим доступа: <https://link.springer.com/book/10.1007/978-981-99-3917-6> — Полнотекстовая англоязычная БД Springer.

Дистанционная поддержка дисциплины (модуля)

Для дистанционной поддержки дисциплины используется система управления образовательным контентом Moodle, развернутая на портале АНО ВО «Университет Иннополис». Во время обучения все студенты получают индивидуальные пароли для входа в систему, в которой размещаются: программа курса, контрольные вопросы, задания, темы лекций с презентационными материалами и т. д.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

1. Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

2. Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- специализированным оборудованием, включая демонстрационное оборудование.

Перечень материально-технического оснащения учебных аудиторий конкретизируется приказом директора или уполномоченного им лица ежегодно и размещается на официальном сайте Университета в информационно-коммуникационной сети «Интернет» в подразделе «Материально-техническое обеспечение и оснащённость образовательного процесса. Доступная среда» раздела «Сведения об образовательной организации» по ссылке: <https://innopolis.university/sveden/objects>.

3. Для проведения занятий лекционного и семинарского (практического) типа предлагаются наборы демонстрационного оборудования и учебно-наглядных пособий, обеспечивающие тематические иллюстрации, соответствующие программе дисциплины (модуля).

4. Лаборатории Университета, в том числе Кибер-физическая лаборатория информационной безопасности «Иннокиберполигон» (<https://engineerschool.innopolis.university/innocyberpolygon>), Цифровая лаборато-

рия искусственного интеллекта «InnoDataHub» (<https://engineerschool.innopolis.university/innodatahub>), Цифровая образовательная песочница для разработки передовых систем «DevOps Playground» (<https://engineerschool.innopolis.university/devops>), оснащены лабораторным и специализированным оборудованием. Перечень материально-технического оснащения и программного обеспечения лабораторий размещается на официальном сайте Университета по ссылке: <https://innopolis.university/sveden/objects>.

5. Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к в электронную информационно-образовательную среду Университета.

6. Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Ссылка на информационный ресурс	Наименование разработки в электронной форме	Доступность
Электронно-библиотечные системы:			
1.	http://www.iprbookshop.ru	Русскоязычная ЭБС на платформе «IPR Smart». Учебники и учебные пособия для университетов издательства «IPRbooks». Учебники и учебные пособия для лиц с ОВЗ	Индивидуальный неограниченный доступ
2.	http://e.lanbook.com	Русскоязычная ЭБС на платформе «Лань». Учебники и учебные пособия для университетов издательства «Лань». Учебники и учебные пособия для лиц с ОВЗ	Индивидуальный неограниченный доступ
Профессиональные базы данных:			
3.	http://search.ebscohost.com	Платформа EBSCOHost. Доступ к англоязычным полнотекстовым книгам	Индивидуальный неограниченный доступ
4.	https://portal.university.innopolis.ru/reading_hall	Электронный каталог научно-технической библиотеки АНО ВО «Университет Иннополис»	Для студентов, зарегистрированных в системе
5.	https://habr.com	База данных для IT-специалистов	Доступ свободный
6.	https://www.sciencedirect.com	База данных ScienceDirect	Доступ свободный
7.	https://elibrary.ru	Научная электронная библиотека	Доступ свободный
8.	https://link.springer.com	Полнотекстовая англоязычная БД Springer	Доступ из локальной сети УИ. Для удаленного доступа требуется регистрация из локальной сети УИ
9.	https://ar.cnki.net	База Данных Academic Reference. База данных полнотекстовых англоязычных ресурсов по всем академическим дисциплинам, опубликованных в Китае	Доступ из локальной сети УИ
10.	https://www.ams.org/journals	База данных англоязычных полнотекстовых журналов AMS Journals (2017–2022 гг.)	Доступ из локальной сети УИ
11.	https://www.dl.begellhouse.com/collections/6764f0021c05	База данных англоязычных полнотекстовых журналов Begell	Доступ из локальной сети УИ

	bd10.html		
12.	https://saemobilus.sae.org/	База данных англоязычных полнотекстовых журналов SAE International SAE eJournals eBooks	Доступ из локальной сети УИ
13.	https://www.worldscientific.com	База данных англоязычных полнотекстовых журналов WorldScientific 2001–2022 гг.	Доступ из локальной сети УИ
14.	https://www.scitation.org/ebooks	AIPP E-Book Collection I + Collection II. Англоязычная База данных полнотекстовых электронных книг.	Доступ из локальной сети УИ
15.	https://ufn.ru/	Автономная некоммерческая организация Редакция журнала "Успехи физических наук". Электронный журнал на русском языке.	Доступ из локальной сети УИ
16.	https://sk.sagepub.com/books/discipline	SAGE Publications Ltd eBook Collections. Англоязычная База данных полнотекстовых электронных книг	Доступ из локальной сети УИ
17.	https://onlinelibrary.wiley.com/	The Wiley Journals Database Англоязычная База данных полнотекстовых электронных журналов	Доступ из локальной сети УИ
18.	https://www.mathnet.ru/	Полнотекстовая коллекция русскоязычных электронных математических журналов	Доступ из локальной сети УИ
19.	https://quantum-electron.lebedev.ru/arhiv/	Электронная версия журнала «Квантовая электроника»	Доступ из локальной сети УИ
20.	http://journals.rcsi.science/	Полнотекстовая коллекция журналов Российской академии наук включает 140 наименований журналов, охватывающих различные научные специальности.	Доступ из локальной сети УИ
21.	https://www.orbit.com/	База данных патентного поиска, объединяющая информацию о более чем 122 миллионах патентных публикаций, полученную из 120 международных патентных ведомств, включая РосПатент, Всемирную организацию интеллектуальной собственности (ВОИС), Европейскую патентную организацию.	Доступ из локальной сети УИ
Информационные справочные системы:			
22.	https://minobrnauki.gov.ru/	Официальный сайт Министерства науки и высшего образования Российской Федерации	Доступ свободный
23.	https://www.edu.ru/	Федеральный портал «Российское образование»	Доступ свободный
24.	http://window.edu.ru/	Информационная система "Единое окно доступа к образовательным ресурсам"	Доступ свободный
25.	http://school-collection.edu.ru/	Единая коллекция цифровых образовательных ресурсов	Доступ свободный
26.	http://fcior.edu.ru/	Федеральный центр информационно-образовательных ресурсов	Доступ свободный

7. Университет Иннополис обеспечен необходимым комплектом лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, в который входят:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое

Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Системы управления обучением:		
Система LMS Moodle	зарубежное	свободно распространяемое
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. МЕТОДИЧЕСКИЕ И ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Задания для семинарских (практических) занятий:

№ п/п	Наименование раз-дела дисциплины (модуля)	Перечень рассматриваемых тем (вопросов)
1.	Машинное обучение с учителем	Практическая работа: Как применять разнообразные регрессионные модели для решения различных проблем регрессии? Опишите различия между вариативными типами регрессионных моделей, их преимущества и недостатки т. д. Внедрение различных моделей классификации для решения различных задач классификации. Опишите разницу между логистической регрессией и наивным классификатором Байеса. Реализация алгоритма обучения перцептрон, SVM и его вариантов для решения различных задач. Проблемы классификации. Решите заданную задачу оптимизации с помощью метода Лагранжа.

2.	Дерево решений и обучение ансамбля	Практическая работа: Применение различных вариантов деревьев решений для решения различных задач классификации. Решить заданную задачу классификации с помощью ансамблевого классификатора. Использование Adaboost для данной задачи.
3.	Машинное обучение без учителя	Выполнение практических заданий и групповых проектов: Реализация различных алгоритмов кластеризации для решения различных задач кластеризации. Применение сдвига среднего значения для видеотелевизионного сопровождения целей.
4.	Углубленное изучение	Выполнение практических заданий и групповых проектов: Применение различных типов искусственной нейронной сети для решения различных задач классификации. Вычислите расхождение языка ядра между двумя распределениями вероятности. Применить различные генеративные модели для различных задач.

Текущий контроль успеваемости обучающихся по дисциплине (модулю):

№ п/п	Наименование раздела дисциплины (модуля)	Форма текущего контроля	Материалы текущего контроля
1.	Машинное обучение с учителем	Проверка выполнения домашних заданий; Устный / письменный опрос; Коллоквиум	Устный / письменный опрос: Верно ли что простая линейная регрессия R^2 и квадратичная корреляция между X и Y идентичны? Какие два предположения относительно ошибки делает модель линейной регрессии? Подберите модель регрессии для данной задачи с данными и поддержите выбор модели. В списке заданных задач выберите какие из них являются регрессионными, а какие - классификационными. В данной графической модели двоичных случайных переменных сколько параметров должны определить условные распределения вероятности для этой байесовой сети? Напишите математическую форму цели минимизации обучающего алгоритма перцептрона Розенблата для двумерного случая. Что представляет из себя алгоритм обучения перцептрону? Напишите математическую форму его цели минимизации для двумерного случая. Что такое классификатор максимальной разницы? Поясните роль переменной пробег в МОВ (метод опорных векторов).
2.	Дерево решений и обучение ансамбля	Проверка выполнения домашних заданий; Устный / письменный опрос; Коллоквиум	Устный / письменный опрос: Каковы плюсы и минусы деревьев принятия решений над другими классификационными моделями? Объясните, как работает обрезка ветвей. Какова цель обучения ансамбля? Что такое программа начальной загрузки и какова его роль в обучении ансамбля? Поясните роль мнимой переменной в МОВ (метод опорных векторов).
3.	Машинное обучение без учителя	Проверка выполнения домашних	Устный / письменный опрос: Какой явный или неявный критерий выбора реализует K-средство?

		заданий; Устный / письменный опрос; Кол- локвиум	Объясните разницу между k-средствами и k-средствами++ Что такое одиночная связь и каковы ее плюсы и минусы? Поясните как работает плотностный алгоритм кластеризации пространственных данных с присутствием шума.
4.	Углубленное изучение	Проверка выполнения домашних заданий; Устный / письменный опрос; Защита проекта; Коллоквиум	Устный / письменный опрос: Что такое полностью подключенное упреждение искусственной нейронной сети? Что такое генеративная модель и чем она отличается от дискриминационной модели? Задания, в том числе, для групповых проектов: Объясните различные гиперпараметры свёрточной нейронной сети. Вычислите расхождение языка ядра между двумя распределениями вероятности.

Контрольные вопросы для подготовки к промежуточной аттестации:

№ п/п	Наименование раздела дисциплины (модуля)	Наименование вопроса
1.	Машинное обучение с учителем	1. Что означает для оценок стандартного среднеквадратичного коэффициента линейных регрессий быть масштабно эквивариантными? 2. Если для набора данных задана регрессионная модель интерпретируйте ее коэффициенты. 3. Объясните какая регрессионная модель лучше подходит для моделирования взаимосвязи между ответом и прогнозом в исходных данных. 4. Если количество тренировочных примеров уходит в бесконечность каким образом это повлияет на смещение и дисперсию классификационной модели? 5. Если задана двумерная задача классификации, определите может ли проводиться оценка задачи при использовании логической регуляции и линейной регрессии. 6. Объясните какая классификационная модель лучше подходит для данной классификации задачи. 7. Рассмотрите ошибку Leave-one-out-CV стандартного МОВ (метод опорных векторов) двух классов. Решите, является ли заданная математическая формулировка правильной или неправильной при заданном значении мнимой переменной. 8. Как выбор мнимой переменной влияет на смещение и дисперсию классификационной модели в МОВ (метод опорных векторов)? 9. Поясните какое ядро лучше подходит для использования в МОВ(метод опорных векторов) для исходных данных.
2.	Дерево решений и обучение ансамбля	1. Когда дерево решений выращивается на полную глубину, как она влияет на смещение и дисперсию дерева и его ответ на искажённые данные? 2. Решите является ли модель ансамбля лучшим выбором для данной задачи классификации или не является? 3. Учитывая конкретную итерацию ускорения и другую важную информацию рассчитайте вес классификатора Adaboost. 4. Объясните как работает обрезка ветвей. 5. Какова цель обучения ансамбля? 6. Каковы плюсы и минусы деревьев принятия решений над другими классификационными моделями? 7. Что такое программа начальной загрузки и какова его роль в обучении ансамбля?
3.	Машинное обучение	К-средства не используют функцию приспособленности подробно. Каковы

	без учителя	характеристики решений, которые находят K-средства? Какую функцию приспособленности они минимизируют в неявном виде? Предположим, что мы сгруппировали набор из точек ввода данных N используя два различных указанных алгоритма кластеризации. В обоих случаях мы получили 5 кластеров и в обоих случаях центры кластеров точно одинаковы. Могут ли 3 точки, задаваемые различными кластерами одним методом, быть назначены одному кластеру другим методом? Каковы характеристики шумовых точек в плотностном алгоритме кластеризации пространственных данных с присутствием шума? Какой явный или неявный критерий выбора реализует K-средство? Объясните разницу между k-средствами и k-средствами++ Что такое одиночная связь и каковы ее плюсы и минусы? Поясните как работает плотностный алгоритм кластеризации пространственных данных с присутствием шума.
4.	Углубленное изучение	Объясните, что такое блок линейной ректификации, каковы его различные варианты, а также каковы их плюсы и минусы? 2. Вычислите количество параметров, которые необходимо выучить во время обучения свёрточной нейронной сети, учитывая всю важную информацию. 3. Поясните, как вариационный автокодер может использоваться в качестве порождающей модели.

Вопросы/Задания к промежуточной аттестации в устной/письменной форме:

Выберите правильный вариант ответа.

1. Следующие алгоритмы нуждаются в масштабировании / нормализации признаков, за исключением:

- A. логистической регрессии
- B. k ближайших соседей
- C. k средних
- D. дерева решений

2. Следующие алгоритмы НЕ нуждаются в масштабировании / нормализации признаков, за исключением:

- A. Adaboost
- B. бэггинга для деревьев решений
- C. DBSCAN
- D. случайного леса

3. Унитарное кодирование (one-hot encoding) предпочтительнее в случае:

- A. во всех случаях
- B. когда нет порядка между отдельными значениями категориальных признаков
- C. когда категориальные признаки сортируются по их соответствию алгоритму
- D. использования деревьев решений

4. В sklearn по умолчанию пропущенные значения заменяются на:

- A. среднее
- B. медиану
- C. моду
- D. наиболее частое значение

5. "Один против всех" (one-vs-all) адаптирует логистическую регрессию к мульти-классовой классификации при помощи:
- A. построения модели для каждой пары классов
 - B. использования другой функции активации, подходящей для мультиклассовой классификации
 - C. построения модели для каждого класса и класса, содержащего все остальные
 - D. использования кросс-энтропии вместо логистической функции в качестве функции потерь
6. Градиентный спуск гарантирует достижение глобального минимума в случае:
- A. выпуклой функции потерь
 - B. использования среднеквадратической ошибки в качестве функции потерь в проблеме классификации
 - C. использования кросс-энтропии вместо логистической функции в качестве функции потерь
 - D. достижение глобального минимума никак не гарантируется
7. Логистическая регрессия получит ту же точность обучения (accuracy), что и метод опорных векторов, когда:
- A. логистическая регрессия использует высокую лямбду для регуляризации
 - B. метод опорных векторов использует линейное ядро
 - C. параметр регуляризации $C = 0$ в методе опорных векторов
 - D. обучающая выборка линейно разделима
8. Мы можем повысить точность (precision) классификатора:
- A. уменьшая количество признаков
 - B. увеличивая количество признаков
 - C. добавляя регуляризацию
 - D. повышая порог классификации
9. Если ошибки обучения и тестирования высоки, то:
- A. модель обладает низким смещением и высокой дисперсией
 - B. модель обладает высоким смещением и низкой дисперсией.
 - C. модель обладает одинаковыми смещением и дисперсией
 - D. модель хорошо обобщает
10. Метрики precision и recall будут равны, когда:
- A. $FP = FN$
 - B. $FP = 0$
 - C. $TP = TN$
 - D. $FN = 0$
11. Все нижеперечисленные методы хороши для работы с несбалансированными наборами данных, кроме:
- A. использования f1-score вместо точности для достижения высокого качества модели
 - B. избыточной и недостаточной выборки
 - C. применения техник для подбора признаков
 - D. использования стратифицированной выборки
12. Проклятие размерности -:

- A. не проблема для k -ближайших соседей и k -средних
 - B. выполнение метода главных компонент для уменьшения габаритов
 - C. проблема, с которой сталкиваются алгоритмы, не зависящие от расстояния
 - D. эти точки становятся равноудаленными в многомерном пространстве
13. Для уменьшения дисперсии в k -ближайших соседей, мы:
- A. уменьшаем значение k
 - B. увеличиваем значение k
 - C. добавляем полиномиальные признаки
 - D. нет правильного ответа
14. Точность обучения k -ближайших соседей, если $k = 1$:
- A. равна 100%
 - B. равна 50%
 - C. равна 65%
 - D. не может быть определена
15. Проблема применения регрессионной модели в k -ближайших соседей:
- A. не проблема
 - B. практически невозможно реализовать такую модель
 - C. она всегда будет выдавать дискретное значение в пределах выходных значений тренировочной выборки
 - D. нет практического пути разделения узлов в таком случае
16. Опорные вектора в методе опорных векторов — это:
- A. точки, которые правильно классифицированы
 - B. точки, не содержащиеся в отступе
 - C. выбросы
 - D. точки с ненулевым множителем Лагранжа
17. Все перечисленное является гиперпараметрами метода опорных векторов, кроме:
- A. ядра
 - B. альфы
 - C. гаммы
 - D. C
18. Наилучший сценарий, где метод опорных векторов может быть использован:
- A. набор данных с низкой размерностью и относительно маленьким числом выборок
 - B. маленький набор данных с относительно маленькой размерностью
 - C. набор данных с высокой размерностью и относительно маленьким числом выборок
 - D. большой набор данных с относительно маленькой размерностью
19. В методе опорных векторов, если параметр C стремится к бесконечности, то:
- A. он может достичь нелинейной гиперплоскости
 - B. он эквивалентен жесткому отступу метода опорных векторов
 - C. мы должны использовать ядерный метод для улучшения
 - D. ядерный метод бесполезен
20. Нейронная сеть с линейными функциями активации в скрытых слоях и сигмоидой в выходном слое:

- A. эквивалентна дереву решений с одним признаком
 - B. не сможет определить нелинейную границу решения
 - C. сможет определить нелинейную зависимость между признаками и целевым значением
 - D. очень эффективно используется в классификации изображений цифр
21. Лучшая стратегия для инициализации параметров нейронной сети:
- A. поставить их равными нулю
 - B. поставить их равными одной константе
 - C. поставить их случайным образом
 - D. нет лучшей стратегии: у каждой стратегии есть свои случаи применения
22. Все перечисленное является методами регуляризации, кроме:
- A. L1-регуляризация
 - B. отсечение в деревьях решений
 - C. исключение нейронов из сети
 - D. замена пустых значений
23. Область значений L2-регуляризации:
- A. имеет гиперсферическую форму
 - B. более агрессивна, чем регуляризация L1
 - C. не используется в нейронных сетях
 - D. воспроизводит разреженную модель
24. Один из недостатков k-средних:
- A. не гарантирует сходимость
 - B. k - обучаемый параметр
 - C. это детерминированный алгоритм
 - D. плохо масштабируется с наборами данных большой размерности
25. Мы уменьшаем дисперсию в дереве решений следующими методами, кроме:
- A. мы ограничиваем глубину дерева
 - B. мы устанавливаем минимальное количество информации для разделения
 - C. мы устанавливаем минимальное количество выборок на лист
 - D. мы выбираем для разделения признаки, которые имеют больший объем информации
26. В случае одномерного набора данных (1 признак) дерево решений:
- A. переобучится на тренировочных данных
 - B. сможет воспроизвести нелинейную границу решений.
 - C. не сможет выбрать признак для разделения
 - D. будет глубины 1
27. Минимальное и максимальное значение энтропии в случае с m классами:
- A. 0 и $\log_2(m)$, соответственно
 - B. 0 и $\sqrt{m}$, соответственно
 - C. 0 и 1, соответственно
 - D. -1 и 1, соответственно
28. Если мы подправим параметры бэггинга и бустинга (деревьев решений), то:
- A. деревья будут одной длины
 - B. деревья в бустинге будут глубже, чем в бэггинге

С. мы не сможем получить какую-либо информацию относительно длины деревьев

Д. деревья в бэггинге глубже, чем деревья в бустинге

29. В Адабусте мы собираем предсказанные вероятности со всех обученных моделей, используя:

А. сумму

В. взвешенную сумму

С. максимальное значение

Д. среднее значение

30. Какие из следующих высказываний о слабых обученных моделях, используемых в ансамблевых моделях, являются правдой?

1. У них маленькая дисперсия и они обычно не переобучаются

2. У них большое смещение, поэтому они не могут решать сложные проблемы

3. У них большая дисперсия и они обычно не переобучаются

А. 1 и 2

В. 1 и 3

С. 2 и 3

Д. Никакое из них

Критерии оценивания сформированности компетенций

Оценивание уровня учебных достижений, обучающихся по дисциплине (модулю), осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

В рамках внутренней системы оценки качества образовательной деятельности по программе обучающимся предоставляется возможность оценивания условий, содержания, организации и качества образовательного процесса по данной дисциплине (модулю).

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме экзамена, при этом проводится оценка компетенций, сформированных по дисциплине.

Система оценивания результатов обучения по дисциплине (модулю)

Знания, умения и навыки обучающихся при промежуточной аттестации определяются в следующей форме: «отлично» (А), «хорошо» (В), «удовлетворительно» (С), «неудовлетворительно» (D).

Критерии оценивания результатов обучения по дисциплине (модулю)

Результат обучения по дисциплине (модулю)	Общая характеристика результата обучения по дисциплине (модулю)
А «Отлично»	Обучающийся владеет знаниями предмета в полном объеме рабочей программы, достаточно глубоко осмысливает дисциплину; самостоятельно, в логической последовательности и исчерпывающе отвечает на все вопросы, подчеркивая при этом самое существенное, умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделять в нем главное: устанавливать причинно-следственные связи; четко формирует ответы, свободно читает результаты анализов и других исследований и решает

	ситуационные задачи повышенной сложности; хорошо знаком с основной литературой и методами исследования большого объема, необходимым для практической деятельности; увязывает теоретические аспекты предмета с задачами практического значения.
В «Хорошо»	Обучающийся владеет знаниями предмета практически в полном объеме рабочей программы; самостоятельно, в логической последовательности и исчерпывающе отвечает на все вопросы, подчеркивая при этом самое существенное, умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделять в нем главное: устанавливать причинно-следственные связи; четко формирует ответы, свободно читает результаты анализов и других исследований и решает ситуационные задачи повышенной сложности; хорошо знаком с основной литературой и методами исследования большого объема, необходимым для практической деятельности; увязывает теоретические аспекты предмета с задачами практического значения.
С «Удовлетворительно»	Обучающийся владеет основным объемом знаний по дисциплине; проявляет затруднения в самостоятельных ответах, оперирует неточными формулировками; в процессе ответов допускаются ошибки по существу вопросов. Обучающийся способен решать лишь наиболее легкие задачи, владеет только обязательным минимумом методов исследований.
D «Неудовлетворительно»	Обучающийся не освоил обязательного минимума знаний предмета, не способен ответить на вопросы даже при дополнительных наводящих вопросах преподавателя.

Методические указания для обучающихся по освоению дисциплины (модуля)

Вид учебных занятий/деятельности	Деятельность обучающегося
Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения лекции, выводы, формулировки, обобщения; пометить важные мысли, выделять ключевые слова, термины. Обозначить вопросы, термины или другой материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на консультации, во время семинарского (практического) занятия.
Практическое (семинарское) занятие	При подготовке к семинарскому (практическому) занятию необходимо проработать материалы лекций, основной и дополнительной литературы по заданной теме. На основании обработанной информации постараться сформировать собственное мнение по выносимой на обсуждение тематике. Обосновать его аргументами, сформировать список источников, подтверждающих его. Во время семинарского (практического) занятия активно участвовать в обсуждении вопросов, высказывать аргументированную точку зрения на проблемные вопросы. Приводить примеры из источниковой базы и научной и/или исследовательской литературы.
Устный/письменный опрос	Отвечать, максимально полно, логично и структурировано, на поставленный вопрос. Основная цель – показать всю глубину знаний по конкретной теме или ее части.
Подготовка к промежуточной аттестации	При подготовке к промежуточной аттестации необходимо проработать вопросы по темам, которые рекомендуются для самостоятельной подготовки. При возникновении затруднений с ответами следует ориентироваться на конспекты лекций, семинаров, рекомендуемую литературу, материалы электронных и информационных справочных ресурсов, статей. Если тема вызывает затруднение, четко сформулировать проблемный вопрос и задать его преподавателю.
Практические (лабораторные) занятия	Практические занятия предназначены прежде всего для разбора отдельных сложных положений, тренировки аналитических навыков, а также для развития коммуникационных навыков. Поэтому на практических занятиях необходимо участвовать

	в тех формах обсуждения материала, которые предлагает преподаватель: отвечать на вопросы преподавателя, дополнять ответы других студентов, приводить примеры, задавать вопросы другим выступающим, обсуждать вопросы и выполнять задания в группах. Работа на практических занятиях подразумевает домашнюю подготовку и активную умственную работу на самом занятии. Работа на практических занятиях в форме устного опроса заключается прежде всего в тренировке навыков применять теоретические положения к самому разнообразному материалу. В ходе практических занятий студенты работают в группах для обсуждения предлагаемых вопросов.
Самостоятельная работа	Самостоятельная работа состоит из следующих частей: 1) чтение учебной, справочной, научной литературы; 2) повторение материала лекций; 3) составление планов устных выступлений; 4) подготовка видеопрезентации. При чтении учебной литературы нужно разграничивать для себя материал на отдельные проблемы, концепции, идеи. Учебную литературу можно найти в электронных библиотечных системах, на которые подписан АНО Университет Иннополис.
Доклад	Публичное, развернутое сообщение по определенной теме или вопросу, основанное на документальных данных. При подготовке доклада рекомендуется использовать разнообразные источники, позволяющие глубже разобраться в теме. Учебную литературу можно найти в электронных библиотечных системах, на которые подписан АНО Университет Иннополис.
Дискуссия	Публичное обсуждение спорного вопроса, проблемы. Каждая сторона должна оппонировать мнению собеседника, аргументируя свою позицию.
Выполнение домашних заданий и групповых проектов	Для выполнения домашних заданий и групповых проектов необходимо получить формулировку задания от преподавателя и убедиться в понимании задания. При выполнении домашних заданий и групповых проектов необходимо проработать материалы лекций, основной и дополнительной литературы по заданной теме.

Для подготовки к занятиям рекомендуется использовать представленные источники в электронных форматах и дополнительную литературу.

8. ДОПОЛНИТЕЛЬНЫЕ СВЕДЕНИЯ

Данная программа и/или ее фрагменты не могут быть использованы другими образовательными организациями без соответствующего разрешения АНО ВО «Университет Иннополис».