



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

КГЭУ «КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ ЭНЕРГЕТИЧЕСКИЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «КГЭУ»)

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

для проведения текущего контроля успеваемости и промежуточной
аттестации студентов по итогам освоения дисциплины

Машинное обучение

Направление подготовки 09.04.01 Информатика и вычислительная
техника

Направленность (профиль) Инженерия искусственного интеллекта

Квалификация Магистр

Форма обучения Очная

Составлено авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Долганов Антон Юрьевич	Кандидат технических наук, без ученого звания	Доцент	Кафедра радиоэлектроники и телекоммуникаций, ИРИТ- РТФ, УрФУ

1. Цель и задачи текущего контроля и промежуточной аттестации студентов по дисциплине «Машинное обучение»

Цель текущего контроля - систематическая проверка степени освоения программы дисциплины «Машинное обучение», уровня сформированности знаний, умений, навыков, компетенций на текущих занятиях

Задачи текущего контроля:

1. определение индивидуального учебного рейтинга студентов;
2. своевременное выполнение корректирующих действий по содержанию и организации процесса обучения; обнаружение и устранение пробелов в усвоении учебной дисциплины;
3. подготовки к промежуточной аттестации.

В течение семестра при изучении дисциплины реализуется комплексная система поэтапного оценивания уровня освоения – балльно-рейтинговая система. За каждый вид учебных действий студенты получают определенное количество баллов. В течение семестра студент может набрать до 60-ти баллов.

Цель промежуточной аттестации - проверка степени усвоения студентами учебного материала за время изучения дисциплины, уровня сформированности компетенций после завершения изучения дисциплины. Аттестация проходит в форме зачета.

Задачи промежуточной аттестации:

1. определение уровня усвоения учебной дисциплины;
2. определение уровня сформированности компетенций.

2. Основное содержание текущего контроля и промежуточной аттестации студентов

В результате изучения дисциплины «Машинное обучение» формируются следующие компетенции или их составляющие:

2.1. Основное содержание текущего контроля

Коды компетенций	Совокупность ожидаемых результатов образования студентов в форме компетенций по завершении модуля / освоения дисциплины	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3
ПК-1	ПК-1.1. Исследует и разрабатывает архитектуры систем искусственного интеллекта для различных предметных областей ПК-1.2. Выбирает комплексы методов и инструментальных средств искусственного интеллекта для решения задач в зависимости от особенностей предметной области ПК-1.3. Разрабатывает единые стандарты в области безопасности (в том числе отказоустойчивости) и совместимости программного обеспечения,	Контрольная работа; домашняя работа; зачёт; экзамен

	эталонных архитектур вычислительных систем и программного обеспечения, а также определяет критерии сопоставления программного обеспечения и критерии эталонных открытых тестовых сред (условий) в целях улучшения качества и эффективности программного обеспечения технологий и систем искусственного интеллекта	
ПК-3	ПК-3.1. Ставит задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области ПК-3.2. Руководит исследовательской группой по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области ПК-3.3. Разрабатывает унифицированные и обновляемые методологии описания, сбора и разметки данных, а также механизмы контроля за соблюдением указанных методологий	Контрольная работа; домашняя работа; зачёт; экзамен
ПК-5	ПК-5.1. Руководит работами по оценке и выбору моделей искусственных нейронных сетей и инструментальных средств для решения поставленной задачи ПК-5.2. Руководит созданием систем искусственного интеллекта на основе моделей искусственных нейронных сетей и инструментальных средств ПК-5.3. Руководит проектами по разработке систем искусственного интеллекта на основе моделей глубоких нейронных сетей и нечетких моделей и методов	Контрольная работа; домашняя работа; зачёт; экзамен
ОПК-9	ОПК-9.1 Применяет инструментальные среды, программно-технические платформы для решения задач в области создания и применения искусственного интеллекта ОПК-9.2 Разрабатывает оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта	Контрольная работа; домашняя работа; зачёт; экзамен

2.2 Основное содержание промежуточной аттестации студентов

Код и наименование компетенций, формируемые с участием дисциплины	Индикаторы достижения компетенции	Планируемые результаты обучения	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3	4
ПК-1	ПК-1.1. Исследует и разрабатывает архитектуры систем искусственного	<i>Знать:</i> архитектурные принципы построения систем искусственного интеллекта,	Контрольная работа; домашняя работа; зачёт

	<p>интеллекта для различных предметных областей ПК-1.2. Выбирает комплексы методов и инструментальных средств искусственного интеллекта для решения задач в зависимости от особенностей предметной области ПК-1.3. Разрабатывает единые стандарты в области безопасности (в том числе отказоустойчивости) и совместимости программного обеспечения, эталонных архитектур вычислительных систем и программного обеспечения, а также определяет критерии сопоставления программного обеспечения и критерии эталонных открытых тестовых сред (условий) в целях улучшения качества и эффективности программного обеспечения технологий и систем искусственного интеллекта</p>	<p>методы декомпозиции основных подсистем (компонентов) и реализации их взаимодействия на основе методологии предметно-ориентированного проектирования</p> <p><i>Уметь:</i> выстраивать архитектуру системы искусственного интеллекта, осуществлять декомпозицию основных подсистем (компонентов) и реализации их взаимодействия на основе методологии предметно-ориентированного проектирования</p> <p><i>Знать:</i> методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках создания интегрированных гибридных интеллектуальных систем различного назначения</p> <p><i>Уметь:</i> выбирать, применять и интегрировать методы и инструментальные средства систем искусственного интеллекта, критерии их выбора и методы комплексирования в рамках создания интегрированных гибридных интеллектуальных систем различного назначения</p> <p><i>Знать:</i> единые стандарты в области безопасности (в том числе отказоустойчивости) и совместимости программного обеспечения, эталонных архитектур вычислительных систем и программного обеспечения технологий и систем искусственного интеллекта</p> <p><i>Уметь:</i> применять и разрабатывать единые стандарты в области безопасности (в том числе отказоустойчивости) и совместимости программного обеспечения, эталонных архитектур вычислительных</p>	
--	---	--	--

		систем и программного обеспечения технологий и систем искусственного интеллекта	
ПК-3	<p>ПК-3.1. Ставит задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области</p> <p>ПК-3.2. Руководит исследовательской группой по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области</p> <p>ПК-3.3. Разрабатывает унифицированные и обновляемые методологии описания, сбора и разметки данных, а также механизмы контроля за соблюдением указанных методологий</p>	<p>Знать:</p> <p>классы методов и алгоритмов машинного обучения</p> <p>Уметь:</p> <p>ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения</p> <p>Знать:</p> <p>методы и критерии оценки качества моделей машинного обучения</p> <p>Уметь:</p> <p>определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области</p> <p>Знать:</p> <p>унифицированные и обновляемые методологии описания, сбора и разметки данных, а также механизмы контроля за соблюдением указанных методологий</p> <p>Уметь:</p> <p>разрабатывать унифицированные и обновляемые методологии описания, сбора и разметки данных, а также механизмы контроля за соблюдением указанных методологий</p>	Контрольная работа; домашняя работа; экзамен
ПК-5	<p>ПК-5.1. Руководит работами по оценке и выбору моделей искусственных нейронных сетей и инструментальных средств для решения поставленной задачи</p> <p>ПК-5.2. Руководит созданием систем искусственного интеллекта на основе моделей искусственных нейронных сетей и инструментальных средств</p> <p>ПК-5.3. Руководит проектами по разработке систем искусственного</p>	<p>Знать:</p> <p>функциональность современных инструментальных средств и систем программирования в области создания моделей искусственных нейронных сетей</p> <p>Уметь:</p> <p>проводить оценку и выбор моделей искусственных нейронных сетей и инструментальных средств для решения задач машинного обучения</p> <p>Знать:</p> <p>принципы построения систем искусственного интеллекта на основе искусственных нейронных сетей, методы и подходы к планированию и</p>	Контрольная работа; домашняя работа; экзамен

	интеллекта на основе моделей глубоких нейронных сетей и нечетких моделей и методов	реализации проектов по созданию систем искусственного интеллекта Уметь: руководить выполнением коллективной проектной деятельности для создания, поддержки и использования систем искусственного интеллекта на основе искусственных нейронных сетей Знать: принципы построения моделей глубоких нейронных сетей и глубокого машинного обучения (с подкреплением и без) Уметь: руководить выполнением коллективной проектной деятельности для создания, поддержки и использования систем искусственного интеллекта на основе моделей глубоких нейронных сетей и нечетких моделей и методов	
ОПК-9	ОПК-9.1 Применяет инструментальные среды, программно-технические платформы для решения задач в области создания и применения искусственного интеллекта ОПК-9.2 Разрабатывает оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта	Знать: инструментальные среды, программно-технические платформы для решения профессиональных задач Уметь: применять инструментальные среды, программно-технические платформы для решения профессиональных задач Знать: принципы разработки оригинальных программных средств для решения профессиональных задач Уметь: разрабатывать оригинальные программные средства для решения задач в области создания и применения искусственного интеллекта	Контрольная работа; домашняя работа; экзамен

3. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины

1 семестр
1.Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0,4

Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Контрольная работа 1</i>	<i>1 сем., 15 нед.</i>	<i>100</i>
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 1		
Промежуточная аттестация по лекциям – не предусмотрено		
Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0,6		
Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Домашняя работа 1</i>	<i>1 сем., 7 нед.</i>	<i>50</i>
<i>Домашняя работа 2</i>	<i>1 сем., 15 нед.</i>	<i>50</i>
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям– 0,5		
Промежуточная аттестация по практическим/семинарским занятиям– зачёт		
Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям– 0,5		

2 семестр		
1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0,4		
Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Контрольная работа 2</i>	<i>2 сем., 15 нед.</i>	<i>100</i>
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 1		
Промежуточная аттестация по лекциям – не предусмотрено		
Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0,6		
Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Домашняя работа 3</i>	<i>1 сем., 5 нед.</i>	<i>20</i>
<i>Домашняя работа 4</i>	<i>1 сем., 10 нед.</i>	<i>40</i>
<i>Домашняя работа 5</i>	<i>1 сем., 15 нед.</i>	<i>40</i>
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям– 0,5		
Промежуточная аттестация по практическим/семинарским занятиям– зачёт		
Весовой коэффициент значимости результатов промежуточной аттестации по		

практическим/семинарским занятиям– 0,5		
3 семестр		
1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0,4		
Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Домашняя работа 6</i>	<i>3 сем., 15 нед.</i>	<i>100</i>
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 1		
Промежуточная аттестация по лекциям – не предусмотрено		
Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0,6		
Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>Домашняя работа 7</i>	<i>3 сем., 15 нед.</i>	<i>100</i>
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям– 0,5		
Промежуточная аттестация по практическим/семинарским занятиям– экзамен		
Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям– 0,5		

4. Практические занятия

№ п/п	Примерный перечень тем практических работ
1.1	Операции над векторами и матрицами с использованием библиотеки NumPy
1.2	Работа с Наборами Данных с использованием библиотеки Pandas
1.3	Разложение матриц с использованием метода главных компонент и сингулярного разложения матрица
1.4	Предварительная обработка данных
1.5	Кластеризация данных методом k-Средних
1.6	Реализация алгоритма линейной регрессии
1.7	Реализация алгоритма логистической регрессии
2.1	Применение библиотеки sklearn для решения задач регрессии, классификации и кластеризации
2.2	Кластеризация данных методом DBSCAN
2.3	Кластеризация данных методом иерархической кластеризации
2.4	Классификация и Регрессия данных методом опорных векторов
2.5	Классификация и Регрессия данных методом k-Ближайших соседей
2.6	Визуализация данных методом Neighborhood Component Analysis
2.7	Визуализация данных методом t-SNE
2.8	Классификация данных с использованием Наивного Байесовского классификатора

2.9	Классификация данных с использованием Дискриминантного Анализа
2.1 0	Визуализация данных с использованием Линейного дискриминанта Фишера
2.1 1	Классификация и Регрессия данных с использованием Деревьев Решений
2.1 2	Классификация и Регрессия данных с использованием ансамблевых методов
3.1	Продвинутая генерация признаков

Примерная тематика контрольных работ:

Номер работы	Примерная тематика контрольных работ
1.1	История машинного обучения и базовые понятия
1.2	Данные
1.3	Линейная Алгебра
1.4	Методы разложения матриц
1.5	Предварительная обработка данных
1.6	Кластеризация
1.7	Основы математического анализа
1.8	Регрессия
1.9	Классификация
2.1	Библиотеки Машинного Обучения
2.2	Продвинутые алгоритмы кластеризации
2.3	Метод опорных векторов
2.4	Ближайшие соседи.
2.5	Байесовские методы
2.6	Деревья Решений
2.7	Ансамблевые методы
2.8	Лучшие практики применения методов машинного обучения
3.1	Продвинутая генерация признаков

Примерные задания в составе контрольных работ:

- 1.1.1 Опишите разницу между подходом машинного обучения и традиционным программированием
- 1.1.2 Опишите разницу между задачами классификации и задачами регрессии
- 1.2.1 Приведите несколько примеров непрерывных и дискретных данных
- 1.2.2 К какому типу данных можно отнести диагноз, поставленный врачом?
- 1.3.1 У вас есть три матрицы A, B, C: A имеет размеры 5×4 , B имеет размеры 4×6 , C имеет размеры 3×5 . Напишите все возможные матрицы, которые можно перемножить между собой, и укажите размеры результирующих матриц
- 1.4.1 Что означают собственные значения и собственные векторы ковариационной матрицы в методе главных компонент?
- 1.4.2 Что означают матрицы U, S и V в сингулярном разложении матрицы?
- 1.5.1 Опишите разные ситуации, в которых вы будете использовать разные типы предварительной обработки данных
- 1.6.1 Какая связь между евклидовым расстоянием и расстоянием Минковского?
- 1.6.2 Какой гиперпараметр нужно настроить для работы алгоритма кластеризации k-средних?
- 1.6.3 Может ли коэффициент силуэта быть равным отрицательному числу? Если «да» - в каких случаях, если «нет» - почему?
- 1.8.1 В чем заключаются основные различия между методом наименьших квадратов и градиентным спуском для нахождения коэффициентов регрессии?
- 1.8.2 В каком случае среднеквадратичная логарифмическая ошибка более подходящая метрика, чем среднеквадратичная ошибка?
- 1.8.3 Может ли коэффициент детерминации быть отрицательным числом? Если «да» - в каких случаях, если «нет» - почему?
- 1.8.4 Почему L1-регуляризация может привести к отбору значимых признаков (в отличие от L2-регуляризации)?
- 1.9.1 В чем основное различие между задачами классификации и задачами регрессии?
- 1.9.2 Допустим, пришли результаты теста мр. К на коронавирус. Тест дал положительный ответ, хотя на самом деле у мр. К нет коронавируса. Ошибку какого рода допустил тест?
- 1.9.3 Как можно построить поверхность принятия решений для логистической регрессии?
- 2.2.1 Иерархическая кластеризация: в чем разница между разными типами связей?
- 2.2.2 DBSCAN: какие точки считаются шумом, граничной точкой, основной точкой?

- 2.3.1 Какие точки считаются опорными векторами (для задач классификации и регрессии)?
- 2.3.2 В чем разница между Hard Margin SVM (жестким зазором) и Soft Margin SVM (мягким зазором) ?
- 2.3.3 Почему Kernel trick помогает улучшить результаты метода опорных векторов?
- 2.4.1 В чем основное отличие использования метода k-ближайших соседей в при классификации и при регрессии?
- 2.4.2 Как найти оптимальное значение гиперпараметра k для методов ближайших соседей?
- 2.4.3 Как можно уменьшить размерность данных с помощью Nearest Components Analysis?
- 2.4.4 Какой гиперпараметр в реализации t-sne связан с балансом между локальными и глобальными аспектами структуры данных?
- 2.5.1 Какого рода информацию нужно получить, чтобы использовать теорему Байеса?
- 2.5.2 Что означает «наивный» в наивном байесовском классификаторе?
- 2.5.3 Предположим есть некое заболевание. Распространенность заболевания 1%. У Вас есть тест, который имеет чувствительность 90% и специфичность 91%. Сколько испытуемых, у которых тест дал положительный результат на самом имеют заболевание?
- 2.5.4 В чем разница между линейным и квадратичным дискриминантным анализом?
- 2.5.5 Почему линейный дискриминантный анализ может использоваться как метод уменьшения размерности?
- 2.6.1 Перечислите основные элементы дерева решений?
- 2.6.2 Как выбираются наиболее оптимальные узлы решения?
- 2.6.3 В чем разница между использованием деревьев решений при классификации и при регрессии?
- 2.7.1 В чем разница между методами бустинга и методами усреднения?
- 2.7.2 Что такое «слабый предсказатель» в контексте ансамблевых методов?
- 2.7.3 В чем разница в объединении деревьев для алгоритмов Random Forest, Gradient Boosting и AdaBoost?
- 2.8.1 В чем заключается основная идея методов фильтрации для выбора значимых параметров?
- 2.8.2 Каков основной принцип, лежащий в основе методов Wrapper для выбора значимых параметров?

2.8.3 Какие концепции необходимы для успешного применения генетического алгоритма для выбора значимых параметров?

3.1 Каким образом категориальные признаки могут быть корректно использованы в линейных моделях?

3.2 Каким образом методы, основанные на деревьях решений, могут использовать комбинаторные признаки?

Примерная тематика домашних работ:

Основы линейной алгебры в среде Python.

Базовые алгоритмы машинного обучения.

Алгоритмы кластеризации и визуализации данных в библиотеке sklearn.

Алгоритмы регрессии в библиотеке sklearn.

Алгоритмы классификации в библиотеке sklearn.

Методы продвинутой генерации признаков.

Применение алгоритмов Машинного обучения для решения задач уменьшения размерности, кластеризации, регрессии и классификации.

Примерные задания в составе домашних работ:

Тема: Основы линейной алгебры в среде Python

Примерные задания:

а. выполните импорт и визуализацию набора данных с использованием библиотеки Pandas

б. выполните предварительную обработку данных используя векторные и матричные операции с использованием библиотеки NumPy

в. реализуйте алгоритм разложения матриц с использованием библиотеки NumPy

Тема: Базовые алгоритмы машинного обучения

Примерные задания:

а. выполните реализацию алгоритма кластеризации данных методом k-Средних с использованием библиотеки NumPy

б. выполните реализацию алгоритма линейной регрессии с использованием библиотеки NumPy

в. выполните реализацию алгоритма логистической регрессии с использованием библиотеки NumPy

Тема: Алгоритмы кластеризации и визуализации данных в библиотеке sklearn

Примерные задания:

- а. выполните кластеризацию учебного набора данных с использованием метода k-Средних
- б. выполните кластеризацию учебного набора данных с использованием метода DBSCAN
- в. выполните кластеризацию учебного набора данных с использованием метода иерархической кластеризации
- г. выполните визуализацию учебного набора данных методом Neighborhood Component Analysis
- д. выполните визуализацию учебного набора данных методом t-SNE

Тема: Алгоритмы регрессии в библиотеке sklearn

Примерные задания:

- а. выполните регрессию для учебного набора данных с использованием линейной регрессии
- б. выполните регрессию для учебного набора данных с использованием метода ближайших соседей
- в. выполните регрессию для учебного набора данных с использованием метода опорных векторов
- г. выполните регрессию для учебного набора данных с использованием деревьев решений
- д. выполните регрессию для учебного набора данных с использованием ансамблевых методов

Тема: Алгоритмы классификации в библиотеке sklearn

Примерные задания:

- а. выполните классификацию для учебного набора данных с использованием логистической регрессии
- б. выполните классификацию для учебного набора данных с использованием метода ближайших соседей
- в. выполните классификацию для учебного набора данных с использованием метода опорных векторов
- г. выполните классификацию для учебного набора данных с использованием деревьев решений
- д. выполните классификацию для учебного набора данных с использованием ансамблевых методов

Тема: Методы продвинутой генерации признаков

Примерные задания:

- а. выполните генерацию дополнительных категориальных признаков с использованием one-hot encoding для реального набора данных
- б. выполните генерацию дополнительных числовых признаков с использованием mean encoding для реального набора данных

Тема: Применение алгоритмов Машинного обучения для решения задач уменьшения размерности, кластеризации, регрессии и классификации.

Примерные задания:

- а. Приведите описание набора данных. Описание должно как минимум включать название набора данных, краткую аннотацию, количество параметров и экземпляров в наборе данных. Желательно добавлять небольшие таблицы с примерами
- б. Для представления результаты алгоритмов уменьшения размерности необходимо включить представление исходного набора данных в новом пространстве, выделив одни и те же классы отдельными маркерами. Также рекомендуется представить матрицы преобразования, чтобы увидеть какие параметры каждый метод использует. Для t-sne представьте несколько результатов для разных значений perplexity
- с. Представьте результаты наиболее оптимальной кластеризации, указав гиперпараметры, которые использовались для получения этого оптимального разделения
- д. Представьте результаты оптимальной классификации, указав гиперпараметры, которые использовались для получения этой оптимальной классификации. Представьте метрики классификации как для обучающих, так и для тестовых данных. Выполните визуализацию в зависимости от используемого алгоритма.
- е. Представьте результаты оптимальной регрессии, указав гиперпараметры, которые использовались для получения этой оптимальной регрессии. Представьте метрики регрессии как для обучающих, так и для тестовых данных. Выполните визуализацию в зависимости от используемого алгоритма.