

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Соловьев Дмитрий Александрович
Должность: ректор ФГБОУ ВО «Саратовский университет»
Дата подписания: 18.09.2025 13:58:17
Уникальный программный ключ:
528682d78e671e58a6b07494e1ba2172f735a12

Приложение 1



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Саратовский государственный университет
генетики, биотехнологии и инженерии имени Н.И. Вавилова»

УТВЕРЖДАЮ

И.о. заведующего кафедрой
 / Ключиков А.В./
« 12 » апреля 2024 г.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Дисциплина	МАШИННОЕ ОБУЧЕНИЕ
Специальность	06.05.01 Биоинженерия и биоинформатика
Направленность (профиль)	Генетика и селекция сельскохозяйственных животных
Квалификация выпускника	Биоинженер и биоинформатик
Нормативный срок обучения	5 лет
Форма обучения	Очная
Кафедра разработчик	Цифровое управление процессами в АПК
Ведущий преподаватель	Шибайкин В.А., доцент

Разработчики: доцент, Шибайкин В.А.

ассистент, Гречечук Ю.Н.


(подпись)

(подпись)

Саратов 2024

Содержание

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП.....	3
2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.....	3
3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы.....	6
4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.....	15

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП

В результате изучения дисциплины «Машинное обучение» обучающиеся, в соответствии с ФГОС ВО по специальности 06.05.01 Биоинженерия и биоинформатика, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 12.08.2020 г. № 973, формируют следующие компетенции, указанные в таблице 1.

Таблица 1

Формирование компетенций в процессе изучения дисциплины

Компетенция		Индикаторы достижения компетенций	Этапы формирования компетенции в процессе освоения ОПОП (семестр)	Виды занятий для формирования компетенции	Оценочные средства для оценки уровня сформированности компетенции
Код	Наименование				
1	2	3	4	5	6
ОПК-5	Способен находить и использовать информацию, накопленную в базах данных по биологическим объектам, включая нуклеиновые кислоты и белки, владеть основными биоинформатическими средствами анализа	ОПК-5.2 Владеет основными биоинформатическими средствами анализа геномной, структурной и другой биологической информации, в том числе специализированным программным обеспечением	5	лекции, лабораторные занятия	собеседование, сообщение, типовое задание
ОПК-6	Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения	ОПК-6.2 Внедряет подходы для сбора, хранения и обработки больших данных в биологии, в том числе с использованием технологий искусственного интеллекта (машинного обучения)	5	лекции, лабораторные занятия	собеседование, сообщение, типовое задание

Профиль подготовки «Генетика и селекция сельскохозяйственных животных»

Компетенция ОПК-5 – также формируется в ходе освоения дисциплин:

Молекулярно-генетические методы исследований в животноводстве, Биоинформатика в селекции с.-х. животных, Биоинформационный анализ молекулярно-генетических данных, Базы данных генетической и геномной информации для селекции с.-х. животных, Подготовка к процедуре защиты и защита выпускной квалификационной работы, Молекулярно-генетические методы исследований в животноводстве, Биоинформатика в селекции с.-х. животных, Биоинформационный анализ молекулярно-генетических данных, Базы данных генетической и

геномной информации для селекции с.-х. животных, Подготовка к процедуре защиты и защита выпускной квалификационной работы, Цифровые технологии и программирование, Введение в искусственный интеллект, Глубокое обучение и нейронные сети, Биоинформатика в селекции с.-х. животных, Биоинформационный анализ молекулярно-генетических данных, Базы данных генетической и геномной информации для селекции с.-х. животных, Подготовка к процедуре защиты и защита выпускной квалификационной работы.

Компетенция ОПК-6 – также формируется в ходе освоения дисциплин:

Программирование на языках высокого уровня, Проектирование и управление базами данных, Подготовка к процедуре защиты и защита выпускной квалификационной работы, Введение в искусственный интеллект, Глубокое обучение и нейронные сети, Подготовка к процедуре защиты и защита выпускной квалификационной работы.

2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Перечень оценочных материалов*

Таблица 2

№ п/п	Наименование оценочного материала	Краткая характеристика оценочного материала	Представление оценочного средства в ОМ
1.	собеседование	средство контроля, организованное как специальная беседа педагогического работника с обучающимся на темы, связанные с изучаемой дисциплиной и рассчитанной на выяснение объема знаний обучающегося по определенному разделу	вопросы по темам дисциплины: – перечень вопросов для устного опроса – перечень вопросов для самостоятельной работы
2.	доклад	продукт самостоятельной работы студента, представляющий собой публичное выступление по представлению полученных результатов решения определенной учебно-практической, учебно-исследовательской или научной темы	темы докладов
3.	типовое задание	содержит средство проверки умений и навыков использования цифровых технологий, программного обеспечения для решения аналитических и управленческих задач.	банк типовых заданий
4.	тестирование	метод, который позволяет выявить уровень знаний, умений и навыков, способностей и других качеств личности, а также их соответствие определенным нормам путем анализа способов выполнения обучающимися ряда специальных заданий	банк тестовых заданий

Программа оценивания контролируемой дисциплины

Таблица 3

№ п/п	Контролируемые разделы (темы дисциплины)	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
1	2	3	4
1.	Инструменты и библиотеки для машинного обучения	ОПК-5, ОПК-6	типовое задание
2.	Предобработка данных	ОПК-5, ОПК-6	типовое задание
3.	Разделение данных	ОПК-5, ОПК-6	типовое задание
4.	Стратегия разделения.	ОПК-5, ОПК-6	типовое задание
5.	Построение линейной регрессии	ОПК-5, ОПК-6	типовое задание
6.	Логистическая регрессия	ОПК-5, ОПК-6	собеседование
7.	К-ближайшие соседи	ОПК-5, ОПК-6	тестовое задание, доклад
8.	Метод опорных векторов (SVM)	ОПК-5, ОПК-6	собеседование
9.	Рекуррентные нейронные сети	ОПК-5, ОПК-6	типовое задание
10.	Методы интерпретации	ОПК-5, ОПК-6	собеседование
11.	Глубокое обучение	ОПК-5, ОПК-6	типовое задание
12.	Применение машинного обучения в реальных проектах	ОПК-5, ОПК-6	типовое задание
13.	Оптимизация гиперпараметров	ОПК-5, ОПК-6	собеседование
14.	Градиентный бустинг	ОПК-5, ОПК-6	тестовое задание, доклад, собеседование
15.	Наивный байесовский классификатор	ОПК-5, ОПК-6	типовое задание

Описание показателей и критериев оценивания компетенций по дисциплине «Машинное обучение» на различных этапах их формирования, описание шкал оценивания

Таблица 4

Код компетенции и, этапы освоения компетенции	Индикаторы достижения компетенций	Показатели и критерии оценивания результатов обучения			
		ниже порогового уровня (неудовлетворительно)	пороговый уровень (удовлетворительно)	продвинутый уровень (хорошо)	высокий уровень (отлично)
1	2	3	4	5	6
ОПК-5 5 семестр	ОПК-5.2 Владеет основными биоинформатическими средствами анализа геномной, структурной и другой биологической	обучающийся не знает значительной части программного материала, плохо ориентируется в материале основы теории машинное обучение,	обучающийся демонстрирует знания только основного материала, но не знает деталей, допускает неточности, допускает неточности в	обучающийся демонстрирует знание материала, машинное обучение допускает не существенные неточности	обучающийся демонстрирует знание материала основы теории машинное обучение, для решения профессиональных задач, не затрудняется с

1	2	3	4	5	6
	информации, в том числе специализированным программным обеспечением	теоретические, допускает существенные ошибки	формулировках, нарушает логическую последовательность в изложении программного материала		ответом при видоизменении заданий
ОПК-6 5 семестр	ОПК-6.2 Внедряет подходы для сбора, хранения и обработки больших данных в биологии, в том числе с использованием технологий искусственного интеллекта (машинного обучения)	Обучающийся не знает основы машинного обучения и технологий его применения в биологических исследованиях, не имеет представление об инструментах и технологиях для хранения и обработки больших данных	знает основы машинного обучения и его применение в биологических исследованиях, имеет представление об инструментах и технологиях для хранения и обработки больших данных.	знает методы и подходы к сбору, хранению и обработке больших данных, а также, применяет методы машинного обучения для анализа и интерпретации биологических данных., владеет инструментами и технологиями для хранения и обработки данных	оценивать качество и достоверность результатов, полученных с помощью обучения; глубоко разбираться в этических аспектах работы с большими данными и машинным обучением; понимать и использовать передовые инструменты и технологии для работы с большими данными.

3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

3.1. Входной контроль

Примерный перечень вопросов

1. В каких сферах информационные технологии применяются уже сейчас?
2. В каких сферах информационные технологии применяются уже сейчас?
3. Дизъюнкция, таблица истинности.
4. Комбинаторика. Перестановки.
5. Комбинаторика. Размещения
6. Комбинаторика. Сочетания.
7. Конъюнкция, таблица истинности.
8. Метод математической индукции

9. Не начнется ли деградация людей, если мы научим компьютер думать за нас?
10. Операции над множествами. Объединение.
11. Отрицание, таблица истинности
12. Способы задания графов,
13. Сформулируйте определение Множества, Мощность множества.
14. Что такое информационные технологии?

3.2. Доклады

Рекомендуемая тематика докладов по дисциплине приведена в таблице 5.

Таблица 5

Темы докладов, рекомендуемые к написанию при изучении дисциплины
«Машинное обучение»

№ п/п	Темы докладов
1	2
1	Актуальные вопросы машинного обучения
2	Модели представления знаний
3	Экспертные системы
4	Интеллектуальные информационные системы
5	Интеллектуальные системы поддержки принятия решений
6	История машинного обучения
7	Искусственный интеллект и теория принятия решений: современные тенденции
8	Использование Машинное обучение в современном бизнесе: вопросы теории и практики
9	Машинное обучение: практические примеры
10	Машинное обучение: теория и методы управления роботами
11	Алгоритм принятия управленческих решений с использованием Машинное обучение: теория и практика

3.3. Типовое задание

Тематика типовых расчетов устанавливается в соответствии с разделами дисциплины. Типовой расчет подразделяется на два этапа. На первом этапе проводится теоретически обзор по теме занятия. На втором этапе решается задача по данной теме. Для каждой темы предусмотрен 2 варианта задания.

Пример типового задания

Тема: Введение в теорию Машинное обучение

Как Вы трактуете термин «Искусственный интеллект»?

Чем интеллектуальные системы отличаются от традиционных прикладных программ?

В чем состоит Тест Тьюринга, что он позволяет проверить, какова схема теста

Тьюринга?

В каких конкретных задач компьютерные (интеллектуальные) системы уже сейчас превосходят человека?

Основные этапы исследований в области ИИ.

3.4. Тестовые задания

По дисциплине «Машинное обучение» предусмотрено проведение следующих видов тестирования: письменное, компьютерное и т.п.

Письменное тестирование.

Письменное тестирование рассматривается как рубежный контроль успеваемости и проводится после изучения раздела дисциплины и раздела Интегрированные методы представления знаний.

Результаты тестирования учитываются при проведении промежуточной аттестации

Пример тестового задания занятию. Интегрированные методы представления знаний

Группа _____ ФИО тестируемого

Тест 1

1. Знания о способах решения задач в заданной предметной области, называются
 - a. декларативные знания;
 - b. Процедурные знания;
 - c. **прагматические знания;**
 - d. экспертные знания.
2. К языкам символьной обработки информации относится
 - a. **LISP;**
 - b. PROLOG;
 - c. KRL;
 - d. GURU.
3. Чтобы в рамках продукционной модели знаний подтвердить выбранную цель при помощи имеющихся правил и данных, используется
 - a. прямой вывод;
 - b. **обратный вывод;**
 - c. индуктивный вывод;
 - d. дедуктивный вывод.
4. Отношения в семантических сетях, определяемые глаголами «влияет», «производит», являются
 - a. атрибутивными связями;
 - b. логическими связями;
 - c. **функциональными связями;**
 - d. количественными связями.

3.7. Рубежный контроль

Вопросы рубежного контроля № 1

Вопросы, рассматриваемые на аудиторных занятиях

1. Что представляет собой алгоритм К-ближайших соседей (KNN)?
2. Как выбирается значение параметра K в алгоритме KNN?
3. Какие преимущества и недостатки имеет метод KNN?
4. Как работает классификация в алгоритме KNN?
5. Как изменяется результат при изменении значения K в алгоритме KNN?
6. Что такое метод опорных векторов (Support Vector Machine)?
7. Какая цель оптимизации ставится в SVM?
8. Что такое гиперплоскость и как она определяется в SVM?
9. Как выбрать оптимальное значение параметра C в SVM?
10. Какие типы ядер используются в SVM и как они влияют на результаты?
11. Каковы основные шаги процесса машинного обучения?
12. Какие существуют подходы к решению задач классификации и регрессии?
13. Как определить, какая модель лучше подходит для конкретной задачи?
14. Какие меры принимаются для предотвращения переобучения моделей?
15. Как оценить качество предсказаний модели на новых данных?
16. Как построить простую нейронную сеть с использованием TensorFlow?
17. Как реализовать алгоритм KNN с помощью библиотеки Scikit-learn?
18. Как провести анализ главных компонент (PCA) для уменьшения размерности данных?
19. Как применить регуляризацию $L1$ и $L2$ в логистической регрессии?
20. Как использовать градиентный спуск для минимизации ошибки в линейной регрессии?
21. Как применять кластерный анализ для сегментации клиентов?
22. Как использовать метод деревьев решений для построения классификационной модели?
23. Как прогнозировать временные ряды с помощью рекуррентных нейронных сетей?
24. Как строить ансамблевые модели для улучшения точности предсказания?
25. Как реализовывать глубокое обучение для обработки изображений?

Вопросы для самостоятельного изучения

1. Работа с библиотеками для машинного обучения
2. Какие библиотеки Python наиболее часто используются для машинного обучения?
3. Какую роль играет библиотека NumPy в машинном обучении?
4. Что такое Scikit-learn и какие задачи она решает?
5. Для каких целей используется библиотека Pandas?

6. В чем отличие TensorFlow от PyTorch?
7. Зачем нужна предобработка данных перед построением модели?
8. Какие этапы включает в себя предобработка данных?
9. Что такое нормализация данных и зачем она применяется?
10. Каким образом осуществляется заполнение пропущенных значений в данных?
11. Какие методы применяются для удаления выбросов в данных?
12. Разделение данных на обучающую и тестовую выборки
13. Почему важно разделять данные на обучающую и тестовую выборки?
14. Какой процент данных обычно выделяется на тестовую выборку?
15. Как осуществляется случайная выборка при разделении данных?
16. Что такое кросс-валидация и когда ее применяют?
17. Как влияет несбалансированность классов на результаты разделения данных?
18. Что представляет собой модель линейной регрессии?
19. Как вычисляется коэффициент детерминации R^2 в линейной регрессии?
20. Как интерпретируются коэффициенты в уравнении линейной регрессии?
21. Как выбирается оптимальная сложность модели линейной регрессии?
22. Какие проблемы могут возникнуть при использовании линейной регрессии?
23. Чем отличается логистическая регрессия от линейной?
24. Как определяется функция активации в логистической регрессии?
25. Как оценивается точность модели логистической регрессии?
26. Какие метрики используются для оценки качества классификации в логистической регрессии?
27. Как решается проблема мультиколлинеарности в логистической регрессии?

Вопросы рубежного контроля № 2

Вопросы, рассматриваемые на аудиторных занятиях

1. Что такое гиперпараметры модели?
2. Какие методы используются для оптимизации гиперпараметров?
3. Как проводится поиск по сетке параметров?
4. Как работает метод случайного поиска гиперпараметров?
5. Какие риски связаны с неправильной настройкой гиперпараметров?
6. Что такое обучение с подкреплением?
7. Каковы основные компоненты среды в обучении с подкреплением?
8. Как агент принимает решения в обучении с подкреплением?
9. Какие виды вознаграждений используются в обучении с подкреплением?
10. Какие алгоритмы обучения с подкреплением вы знаете?
11. Что такое трансформеры?
12. Как работают механизмы внимания в трансформерах?
13. Какие архитектуры трансформеров вы знаете?
14. Как трансформеры применяются в задачах обработки естественного языка?

15. Какие преимущества и недостатки имеют трансформеры по сравнению с другими моделями?
16. Какие этические аспекты следует учитывать при разработке моделей машинного обучения?
17. Как избежать предвзятости в данных и моделях?
18. Как обеспечить прозрачность и объяснимость моделей?
19. Какие правовые нормы регулируют использование машинного обучения?
20. Как обеспечивать конфиденциальность и безопасность данных в машинном обучении?
21. Какие метрики используются для оценки качества моделей?
22. Как проводить кросс-валидацию для оценки модели?
23. Как интерпретировать результаты тестирования модели?
24. Какие подходы используются для тестирования моделей в реальном времени?
25. Как убедиться, что модель устойчива к изменениям в данных?

Вопросы для самостоятельного изучения

1. Что означает интерпретация модели в контексте машинного обучения?
2. Какие методы используются для интерпретации моделей?
3. Как измерить важность признаков в моделях машинного обучения?
4. Какие инструменты помогают визуализировать результаты интерпретации моделей?
5. Как интерпретируется модель линейной регрессии?
6. Что такое наивный байесовский классификатор?
7. Какова основная идея метода наивного Байеса?
8. Какие предположения делает наивный байесовский классификатор о данных?
9. Как рассчитывается вероятность класса в наивном байесовском классификаторе?
10. Какие ограничения у наивного байесовского классификатора?
11. Что представляют собой древовидные модели?
12. Как строится дерево решений?
13. Какие критерии используются для разбиения узлов дерева решений?
14. Как предотвратить переобучение в дереве решений?
15. Какие преимущества и недостатки имеют древовидные модели?
16. Что такое градиентный бустинг?
17. Как работает алгоритм градиентного бустинга?
18. Какие параметры настраиваются в градиентном бустинге?
19. Как градиентный бустинг улучшает производительность модели?
20. Какие популярные реализации градиентного бустинга вы знаете?

21. Какие этапы включает процесс разработки проекта с применением машинного обучения?
22. Какие реальные примеры применения машинного обучения вы можете привести?
23. Как происходит сбор и подготовка данных для реального проекта?
24. Какие трудности возникают при внедрении моделей машинного обучения в производство?
25. Как оценивать успешность внедрения модели в реальный проект?

3.8 Промежуточная аттестация

В соответствии с учебным планом по специальности 06.05.01 Биоинженерия и биоинформатика вид промежуточной аттестации – Экзамен

Целью проведения экзамена по дисциплине является определение фактического уровня теоретических знаний и навыков обучающихся. В вопросах для зачета расчетные задания отсутствуют.

Вопросы, выносимые на экзамен

1. Что представляет собой алгоритм K-ближайших соседей (KNN)?
2. Как выбирается значение параметра K в алгоритме KNN?
3. Какие преимущества и недостатки имеет метод KNN?
4. Как работает классификация в алгоритме KNN?
5. Как изменяется результат при изменении значения K в алгоритме KNN?
6. Что такое метод опорных векторов (Support Vector Machine)?
7. Какая цель оптимизации ставится в SVM?
8. Что такое гиперплоскость и как она определяется в SVM?
9. Как выбрать оптимальное значение параметра C в SVM?
10. Какие типы ядер используются в SVM и как они влияют на результаты?
11. Каковы основные шаги процесса машинного обучения?
12. Какие существуют подходы к решению задач классификации и регрессии?
13. Как определить, какая модель лучше подходит для конкретной задачи?
14. Какие меры принимаются для предотвращения переобучения моделей?
15. Как оценить качество предсказаний модели на новых данных?
16. Как построить простую нейронную сеть с использованием TensorFlow?
17. Как реализовать алгоритм KNN с помощью библиотеки Scikit-learn?
18. Как провести анализ главных компонент (PCA) для уменьшения размерности данных?
19. Как применить регуляризацию L1 и L2 в логистической регрессии?
20. Как использовать градиентный спуск для минимизации ошибки в линейной регрессии?
21. Как применять кластерный анализ для сегментации клиентов?
22. Как использовать метод деревьев решений для построения классификационной модели?

23. Как прогнозировать временные ряды с помощью рекуррентных нейронных сетей?
24. Как строить ансамблевые модели для улучшения точности предсказания?
25. Как реализовывать глубокое обучение для обработки изображений?
26. Работа с библиотеками для машинного обучения
27. Какие библиотеки Python наиболее часто используются для машинного обучения?
28. Какую роль играет библиотека NumPy в машинном обучении?
29. Что такое Scikit-learn и какие задачи она решает?
30. Для каких целей используется библиотека Pandas?
31. В чем отличие TensorFlow от PyTorch?
32. Зачем нужна предобработка данных перед построением модели?
33. Какие этапы включает в себя предобработка данных?
34. Что такое нормализация данных и зачем она применяется?
35. Каким образом осуществляется заполнение пропущенных значений в данных?
36. Какие методы применяются для удаления выбросов в данных?
37. Разделение данных на обучающую и тестовую выборки
38. Почему важно разделять данные на обучающую и тестовую выборки?
39. Какой процент данных обычно выделяется на тестовую выборку?
40. Как осуществляется случайная выборка при разделении данных?
41. Что такое кросс-валидация и когда ее применяют?
42. Как влияет несбалансированность классов на результаты разделения данных?
43. Что представляет собой модель линейной регрессии?
44. Как вычисляется коэффициент детерминации R^2 в линейной регрессии?
45. Как интерпретируются коэффициенты в уравнении линейной регрессии?
46. Как выбирается оптимальная сложность модели линейной регрессии?
47. Какие проблемы могут возникнуть при использовании линейной регрессии?
48. Чем отличается логистическая регрессия от линейной?
49. Как определяется функция активации в логистической регрессии?
50. Как оценивается точность модели логистической регрессии?
51. Какие метрики используются для оценки качества классификации в логистической регрессии?
52. Как решается проблема мультиколлинеарности в логистической регрессии?
53. Что такое гиперпараметры модели?
54. Какие методы используются для оптимизации гиперпараметров?
55. Как проводится поиск по сетке параметров?
56. Как работает метод случайного поиска гиперпараметров?
57. Какие риски связаны с неправильной настройкой гиперпараметров?
58. Что такое обучение с подкреплением?
59. Каковы основные компоненты среды в обучении с подкреплением?
60. Как агент принимает решения в обучении с подкреплением?
61. Какие виды вознаграждений используются в обучении с подкреплением?

62. Какие алгоритмы обучения с подкреплением вы знаете?
63. Что такое трансформеры?
64. Как работают механизмы внимания в трансформерах?
65. Какие архитектуры трансформеров вы знаете?
66. Как трансформеры применяются в задачах обработки естественного языка?
67. Какие преимущества и недостатки имеют трансформеры по сравнению с другими моделями?
68. Какие этические аспекты следует учитывать при разработке моделей машинного обучения?
69. Как избежать предвзятости в данных и моделях?
70. Как обеспечить прозрачность и объяснимость моделей?
71. Какие правовые нормы регулируют использование машинного обучения?
72. Как обеспечивать конфиденциальность и безопасность данных в машинном обучении?
73. Какие метрики используются для оценки качества моделей?
74. Как проводить кросс-валидацию для оценки модели?
75. Как интерпретировать результаты тестирования модели?
76. Какие подходы используются для тестирования моделей в реальном времени?
77. Как убедиться, что модель устойчива к изменениям в данных?
78. Что означает интерпретация модели в контексте машинного обучения?
79. Какие методы используются для интерпретации моделей?
80. Как измерить важность признаков в моделях машинного обучения?
81. Какие инструменты помогают визуализировать результаты интерпретации моделей?
82. Как интерпретируется модель линейной регрессии?
83. Что такое наивный байесовский классификатор?
84. Какова основная идея метода наивного Байеса?
85. Какие предположения делает наивный байесовский классификатор о данных?
86. Как рассчитывается вероятность класса в наивном байесовском классификаторе?
87. Какие ограничения у наивного байесовского классификатора?
88. Что представляют собой древовидные модели?
89. Как строится дерево решений?
90. Какие критерии используются для разбиения узлов дерева решений?
91. Как предотвратить переобучение в дереве решений?
92. Какие преимущества и недостатки имеют древовидные модели?
93. Что такое градиентный бустинг?
94. Как работает алгоритм градиентного бустинга?
95. Какие параметры настраиваются в градиентном бустинге?
96. Как градиентный бустинг улучшает производительность модели?
97. Какие популярные реализации градиентного бустинга вы знаете?

98. Какие этапы включает процесс разработки проекта с применением машинного обучения?

99. Какие реальные примеры применения машинного обучения вы можете привести?

100. Как происходит сбор и подготовка данных для реального проекта?

101. Какие трудности возникают при внедрении моделей машинного обучения в производство?

102. Как оценивать успешность внедрения модели в реальный проект?

пример билета

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Саратовский государственный университет генетики, биотехнологии и инженерии имени
Н.И. Вавилова»

Кафедра Цифровое управление процессами в АПК

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1
по дисциплине «Машинное обучение»

1. Как градиентный бустинг улучшает производительность модели?
2. Какие популярные реализации градиентного бустинга вы знаете

30.08.2024

И.о. зав. кафедрой

А.В. Ключиков

4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

4.1 Процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Контроль результатов обучения студентов, этапов и уровня формирования компетенций по дисциплине «Машинное обучение» осуществляется через проведение входного, текущего, рубежных, выходного контролей и контроля самостоятельной работы

Формы текущего, промежуточного и итогового контроля и контрольные задания для текущего контроля разрабатываются кафедрой исходя из специфики дисциплины, и утверждаются на заседании кафедры.

4.2 Критерии оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Описание шкалы оценивания достижения компетенций по дисциплине приведено в таблице 6.

Таблица 6

Уровень освоения компетенции	Отметка по пятибалльной системе (Экзамен)	Описание
высокий	«отлично»	Обучающийся обнаружил всестороннее, систематическое и глубокое знание учебного материала, умеет свободно выполнять задания, предусмотренные программой, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой. Как правило, обучающийся проявляет творческие способности в понимании, изложении и использовании материала
базовый	«хорошо»	Обучающийся обнаружил полное знание учебного материала, успешно выполняет предусмотренные в программе задания, усвоил основную литературу, рекомендованную в программе
пороговый	«удовлетворительно»	Обучающийся обнаружил знания основного учебного материала в объеме, необходимом для дальнейшей учебы и предстоящей работы по профессии, справляется с выполнением практических заданий, предусмотренных программой, знаком с основной литературой, рекомендованной программой, допустил погрешности в ответе на экзамене и при выполнении экзаменационных заданий, но обладает необходимыми знаниями для их устранения под руководством преподавателя
–	«неудовлетворительно»	Обучающийся обнаружил пробелы в знаниях основного учебного материала, допустил принципиальные ошибки в выполнении предусмотренных программой практических заданий, не может продолжить обучение или приступить к профессиональной деятельности по окончании образовательной организации без дополнительных занятий

4.2.1. Критерии оценки устного ответа при промежуточной аттестации

При ответе на вопрос обучающийся демонстрирует:

знания: основных принципов работы с большими данными, основ машинного обучения и его применение в биологических исследованиях. Инструменты и технологии для работы с большими данными, включая базы данных, алгоритмы машинного обучения, библиотеки и фреймворки. Этические аспекты работы с большими данными и машинным обучением в контексте биологических исследований.

умения: собирать и обрабатывать большие объёмы биологических данных с использованием различных методов и подходов. Оценивать качество и достоверность результатов, полученных с помощью машинного обучения.

владение навыками: Использовать инструменты и технологии для хранения и обработки больших данных, такие как базы данных и системы управления данными.

Критерии оценки*

отлично	обучающийся демонстрирует: – знание материала х, практики применения материала, исчерпывающе и последовательно, четко и логично излагает материал, хорошо ориентируется в материале, не затрудняется с ответом при видоизменении заданий; – умение <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – успешное и системное владение навыками чтения и оценки задач машинного обучения
хорошо	обучающийся демонстрирует: – знание материала, не допускает существенных неточностей; – в целом успешное, но содержащие отдельные пробелы, умение <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – в целом успешное, но содержащее отдельные пробелы или сопровождающееся отдельными ошибками владение навыками <i>чтения и оценки задач машинного обучения.</i>
удовлетворительно	обучающийся демонстрирует: – знания только основного материала, но не знает деталей, допускает неточности, допускает неточности в формулировках, нарушает логическую последовательность в изложении программного материала; – в целом успешное, но не системное умение формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки; – в целом успешное, но не системное владение навыками чтения и оценки <i>задач машинного обучения.</i>
неудовлетворительно	обучающийся: – не знает значительной части программного материала, плохо ориентируется в материале по современным технологиям анализа данных; – не умеет <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – обучающийся не владеет навыками чтения и оценки <i>Задач машинного обучения</i>

4.2.2. Критерии оценки доклада

При написании доклада обучающийся демонстрирует:

знания: по вопросам доклада

умения: анализировать и содержательно интерпретировать полученные результаты исследований.

владение навыками: поиска информации в традиционных библиотеках и информационных ресурсах.

Критерии оценки доклада

отлично	обучающийся демонстрирует: - знание исследуемой темы (доклад структурирован; использованы различные точки зрения по рассматриваемому вопросу, сделаны и аргументированы
----------------	--

	основные выводы, прослушивается самостоятельность суждений, основные понятия вопроса изложены подробно) - логичность и структурированность изложения материала; - расширенную электронную презентацию к докладу на 5 слайдов
хорошо	обучающийся демонстрирует: - знание темы доклада (доклад структурирован; использованы различные точки зрения по рассматриваемому вопросу, сделаны и аргументированы основные выводы) - расширенную электронную презентацию к докладу менее 5 слайдов
удовлетворительно	обучающийся демонстрирует: - неполное знание материала (в материале представлена одна точка зрения, отсутствует самостоятельность суждений) - не представлена электронная презентация
неудовлетворительно	обучающийся: не выполнил доклад

4.2.3. Критерии оценки выполнения типового задания

При выполнении типовых заданий обучающийся демонстрирует:

знания: основные концепции и термины машинного обучения, включая модели, алгоритмы и методы., Различия между обучением с учителем, без учителя и с подкреплением, основные этапы процесса машинного обучения: сбор данных, предобработка, обучение модели, оценка и тестирование, принципы работы популярных алгоритмов, таких как линейная регрессия, деревья принятия решений, нейронные сети, методы оценки качества моделей, включая метрики точности.

умения: Применять алгоритмы машинного обучения для решения конкретных задач, Использовать инструменты и библиотеки для разработки моделей, Выполнять предобработку данных: очистка, нормализация и трансформация данных., Разделять данные на обучающие и тестовые выборки., Проводить оценку модели и интерпретировать результаты.

владение навыками: Навыками программирования на Python для реализации алгоритмов машинного обучения, Способностью выбирать подходящие алгоритмы в зависимости от задачи и данных, Навыками оптимизации гиперпараметров моделей для улучшения их производительности, Умением визуализировать данные и результаты работы моделей для лучшего понимания, Способностью критически оценивать результаты экспериментов и вносить необходимые коррективы в процесс обучения моделей.

отлично	обучающийся демонстрирует: - знания современных методов и программных средств выявления тенденций изменения признаков - умения анализировать и правильно интерпретировать, применять инструментарий программы, использовать интегрированную среду разработки, применять язык Python для работы. - владеет навыками использования технологии анализа данных.
----------------	--

хорошо	обучающийся демонстрирует: – современных методов и программных средств выявления тенденций изменения признаков, – умения анализировать и правильно интерпретировать, применять инструментарий программы, использовать интегрированную среду разработки, применять язык Python для работы, – владеет навыками использования технологии анализа данных
удовлетворительно	обучающийся демонстрирует: – знания последовательности решения задания, не всех используемых современных методов и программных средств выявления тенденций изменения признаков. – умения применять инструментарий программы, проводить расчеты, не приводящие к правильному числовому ответу. – владеет навыками современных методов и программных средств выявления тенденций изменения признаков, не может самостоятельно составить выводов по результатам решения задачи.
неудовлетворительно	обучающийся: – не знает современных методов и программных средств выявления тенденций изменения признаков – не умеет применять инструментарий программы, проводить расчеты, – не владеет навыками использования технологии анализа данных.

4.2.4. Критерии оценки выполнения тестовых заданий

При выполнении тестовых заданий обучающийся демонстрирует:

знания: основ функционирования Машинное обучение на основе математических методов.

Критерии оценки выполнения тестовых заданий

отлично	обучающийся демонстрирует: – 85 % правильных ответов
хорошо	обучающийся демонстрирует: – 60 % правильных ответов
удовлетворительно	обучающийся демонстрирует: – 50 % правильных ответов
неудовлетворительно	обучающийся: – Дал менее 45 % правильных ответов

4.2.5. Критерии оценки сообщения

При устном сообщении обучающийся демонстрирует:

знания: о технологиях машинного обучения, о типах данных и базовых алгоритмических конструкциях, языках для работы с массивами и строками.

умения: проводить сравнительный анализ методов и инструментальных средств создания машинного обучения; планировать аналитические работы, использовать современные информационные технологии аналитики для информационно-аналитического сопровождения деятельности.

владение навыками: анализом решений с точки зрения достижения целевых показателей решений.

Критерии оценки сообщения

отлично	обучающийся демонстрирует: – высокий уровень знаний в области информационных технологий в части Машинное обучение, тема при выполнении сообщения раскрыта полностью; – умение проводить сравнительный анализ методов и инструментальных средств анализа данных. – владеет навыками поиска современных средств обработки информации.
хорошо	обучающийся демонстрирует: – обучающийся показывает хороший уровень в области информационных технологий в части Машинное обучение, тема при выполнении сообщения раскрыта полностью, но содержит неточности; – умение представлять информационные технологии и программные средства для анализа данных; – владеет навыками использования специальной терминологии
Удовлетворительно	обучающийся демонстрирует: – обучающийся показывает средний уровень знаний по теме сообщения, тема раскрыта на 50 % – умение представлять информационных технологий и программных средств для анализа данных – владеет навыками малой части использования специальных терминов.
Неудовлетворительно	обучающийся: – обучающийся показывает низкий уровень знаний по теме сообщения, тема при выполнении сообщения не раскрыта, содержит недостоверную информацию, отсутствует специальная терминология

Разработчик: доцент, Шибайкин В.А.



 (подпись)

ассистент, Гречук Ю.Н.



 (подпись)