МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

МИНИСТЕРСТВО НАУКИ, ВЫСШЕГО ОБРАЗОВАНИЯ И ИННОВАЦИЙ КЫРГЫЗСКОЙ РЕСПУБЛИКИ

МОО ВО Кыргызско-Российский Славянский университет имени первого Президента Российской Федерации Б.Н. Ельцина

Высшая школа Цифровых технологий

Фонд оценочных средств

по дисциплине «**Машинное обучение**»

Уровень высшего образования

МАГИСТРАТУРА

Направление подготовки

09.04.04 – РФ, 710400 – КР Программная инженерия

(код и наименование направления подготовки)

«Основы анализа и разработки приложений с большими объемами распределенных данных» (профиль образовательной программы)

Квалификация

магистр

Фонд оценочных средств предназначен для контроля знаний обучающихся по направлению подготовки 09.04.04 – РФ, 710400 – КР «Программная инженерия» по дисциплине *«Машинное обучение»*.

Фонд оценочных средств рассмотрен и утвержден
на заседании Высшей школы Цифровых технологий, протокол № 1 от 02.09.2026

Директор Высшей школы Цифровых технологий:

к.т.н., доцент Хмелева И.В.

Руководитель образовательной программы:

д.т.н., профессор Лыченко Н.М.

Исполнители (разработчики):

д.т.н., профессор Верзунов С.Н.

д.т.н., профессор Лыченко Н.М.

СОГЛАСОВАНО:
Зам. директора по учебной работе
Политехнического института КРСУ Комарцов Н.М.

Раздел 1 Перечень компетенций и этапы их формирования

Результаты и формулировки компетенций соответствуют рабочей программе дисциплины. Уровни 1, 2 и 3 характеризуют выполнение по образцу, самостоятельное решение типовой задачи и обоснованный анализ решения.

Формируемая компетенция	Планируемые результаты и этапы формирования	Виды оценочных средств
ПК-8: Способен проводить поиск и анализ инноваций в управлении и машинном обучении	Знать Уровень 1. источники научной и технической информации о методах машинного обучения Уровень 2. критерии сравнения новизны, применимости и ограничений методов Уровень 3. подходы к проверке заявлений о качестве и воспроизводимости новых решений	Блок A; вопросы 1–10, диагностический тест; блок D – теория. Формирование: модули 1–3.
	Уметь Уровень 1. найти и описать метод из научного или технического источника Уровень 2. сопоставить метод с базовым подходом на одинаковых данных и метриках Уровень 3. обосновать выбор метода с учетом качества, ресурсов, смещения и рисков	Блок B; практические работы и задания 11–20; блок D – решение задачи. Формирование: модули 1–3.
	Владеть Уровень 1. приемами документированного поиска и отбора источников Уровень 2. методами критического обзора и воспроизводимого сравнения решений Уровень 3. навыками подготовки обоснованного заключения о применимости инновации	Блок C; контрольные работы и задания 21–30; блок D – самостоятельное объяснение и защита. Формирование: модули 1–3.
ПК-13: Способен совершенствовать и разрабатывать новые методы, модели и алгоритмы, в том числе для работы с большими данными	Знать Уровень 1. этапы подготовки данных, обучения и проверки моделей Уровень 2. принципы построения конвейеров, настройки моделей и оценки затрат Уровень 3. ограничения методов при росте объема данных и способы проверки модификаций алгоритмов	Блок A; вопросы 1–10, диагностический тест; блок D – теория. Формирование: модули 1–3.
	Уметь Уровень 1. реализовать базовый воспроизводимый конвейер машинного обучения Уровень 2. внести и проверить ограниченную модификацию признаков, процедуры обучения или модели Уровень 3. сравнить прототип с базовым решением по качеству и ресурсам на подходящем наборе данных	Блок B; практические работы и задания 11–20; блок D – решение задачи. Формирование: модули 1–3.
	Владеть Уровень 1. инструментами Python для анализа данных и обучения моделей Уровень 2. навыками документирования данных, кода, разбиений и результатов Уровень 3. приемами аргументированной защиты экспериментального прототипа и его ограничений	Блок C; контрольные работы и задания 21–30; блок D – самостоятельное объяснение и защита. Формирование: модули 1–3.

Раздел 2 Технологическая карта дисциплины Машинное обучение

Курс 2; семестр 3; трудоемкость 4 ЗЕТ, 144 часов. Промежуточная аттестация – экзамен.

Название модулей дисциплины согласно РПД	**Контроль**	**Форма контроля**	**зачетный минимум**	**зачетный максимум**	**график контроля**
Модуль 1. Подготовка данных и разведочный анализ	Текущий контроль	Практические работы № 1–2	4	8	5
	Рубежный контроль	Контрольное задание № 1 (теория и анализ результатов); самостоятельное задание № 1	6	8	5
Модуль 2. Обучение с учителем: регрессия, классификация и оценка качества моделей	Текущий контроль	Практические работы № 3–4	4	9	9
	Рубежный контроль	Контрольное задание № 2 (теория и анализ результатов); самостоятельное задание № 2	6	9	9
Модуль 3. Ансамблевые методы обучения с учителем, анализ ошибок и интерпретация моделей	Текущий контроль	Практические работы № 5–6	4	9	13
	Рубежный контроль	Контрольное задание № 3 (теория и анализ результатов); самостоятельное задание № 3	6	9	13
Модуль 4. Обучение без учителя и воспроизводимый исследовательский проект	Текущий контроль	Практические работы № 7	4	9	16
	Рубежный контроль	Контрольное задание № 4 (теория и анализ результатов); самостоятельное задание № 4	6	9	16
ВСЕГО за семестр			40	70	
Промежуточный контроль (экзамен)		Письменная работа: вопросы и практическое задание	20	30	
Семестровый рейтинг по дисциплине			60	100	

Раздел 3 Типовые контрольные задания и оценочные материалы

3.1 Блок А Вопросы и задания для проверки планируемых результатов

Вопросы и задания проверяют знания, умения и владение приемами работы по трем тематическим разделам дисциплины. Для конкретного контроля преподаватель выбирает задания в пределах изученного материала.

Знать

1. Опишите этапы анализа данных и постановки задач обучения с учителем и без учителя.
2. Какие ошибки качества, происхождения и смещения данных влияют на выводы?
3. Чем различаются случайное, групповое и временное разбиения?
4. Объясните назначение Pipeline и вложенного обучения преобразований при CV.
5. Сравните линейные модели, деревья, случайный лес и градиентный бустинг.
6. Как выбор метрики и порога связан со стоимостью ошибок и дисбалансом?
7. Чем настройка гиперпараметров отличается от независимой оценки качества?
8. Сравните k-средних, DBSCAN и PCA по задачам, предпосылкам и ограничениям.
9. Какие сведения нужны для оценки новизны и воспроизводимости опубликованного метода?
10. Как рост объема и размерности данных влияет на память, время обучения и выбор метода?

Уметь

11. Составьте паспорт набора данных, включая единицу наблюдения, источник и ограничения.
12. Проведите разведочный анализ и обоснуйте обработку пропусков и выбросов.
13. Выберите схему разбиения для повторных наблюдений по объектам или времени.
14. Соберите конвейер предобработки и модели без утечки данных.
15. Сравните базовую модель с альтернативами на одинаковых фолдах.
16. Настройте ограниченную сетку параметров и сохраните независимый тест.
17. Рассчитайте метрики классификации и обоснуйте выбор с учетом цены ошибок.
18. Проведите кластеризацию или PCA и проверьте устойчивость интерпретации.
19. Найдите описание метода и отделите авторские утверждения от проверенных результатов.
20. Проверьте ограниченную модификацию признаков или алгоритма по качеству и затратам.

Владеть

21. Представьте воспроизводимую структуру проекта с данными, кодом и инструкцией запуска.
22. Сохраните разбиения, параметры, версии среды и журнал экспериментов.
23. Покажите проверки отсутствия утечек и пересечения объектов между выборками.
24. Организуйте честное сравнение базового и измененного метода при общем бюджете.
25. Оцените разброс результатов на нескольких разбиениях и обозначьте пределы обобщения.
26. Документируйте время обучения, размер данных и ограничения аппаратной среды.
27. Разберите ошибки модели и отличите объяснение прогноза от причинного вывода.
28. Продемонстрируйте повторный запуск итогового конвейера в чистом сеансе.
29. Подготовьте отчет, связывающий исследовательский вопрос, эксперимент и выводы.
30. Защитите прототип, обосновав новизну изменения, практический смысл и ограничения.

Диагностический тест один вариант

Выберите один правильный ответ в каждом вопросе. Вопросы 1–4 относятся к модулю 1, 5–8 – к модулю 2, 9–12 – к модулю 3. Тест предназначен для самопроверки и обсуждения: 1 верный ответ – 1 диагностический пункт, максимум 12. Эти пункты не прибавляются к семестровому рейтингу; балльные позиции технологической карты сохраняются.

1. Как избежать попадания записей одного клиента в train и test?

А) Разделять по идентификатору клиента Б) Удалить идентификатор после случайного разбиения В) Переименовать столбцы Г) Перемешать строки дважды

2. Когда обучается заполнение пропусков при CV?

А) Внутри обучающей части каждого фолда Б) Один раз на всей таблице В) После выбора ответа экзаменатора Г) На независимом тесте

3. Что описывает смещение выборки?

А) Только формат файла Б) Только число признаков В) Любую случайную перестановку строк Г) Систематическое отличие выборки от целевой совокупности

4. Какое разбиение соответствует прогнозу будущих событий?

А) Случайное по перекрывающимся окнам Б) По значению будущего отклика В) Хронологическое с исключением утечки будущих наблюдений Г) По алфавиту имен столбцов

5. На чем сравнивают исходный и модифицированный методы?

А) На общих разбиениях, метриках и сопоставимом бюджете Б) На разных задачах без оговорок В) Только по лучшему обучающему результату Г) По числу терминов в описании

6. Чему равен recall при TP = 30 и FN = 10?

А) 0,60 Б) 0,75 В) 0,40 Г) 0,25

7. Что необходимо для обоснованной проверки статистической значимости превосходства?

А) Большее среднее одного запуска Б) Любое положительное отличие в третьем знаке В) Два округленных средних без протокола Г) Обоснованная процедура сравнения с учетом зависимости результатов

8. Где выбирают порог классификации с учетом стоимости ошибок?

А) По фамилии разработчика Б) На независимом тесте после просмотра ответов В) На train/validation или внутри CV до итогового теста Г) На всех данных одновременно

9. Когда после исключения шума допустим silhouette для DBSCAN?

А) Когда все точки являются шумом Б) Когда число кластеров от 2 до $n-1$ В) При одном кластере из всех точек Г) При отсутствии объектов

10. Что проверяется при поиске инновационного метода?

А) Только число страниц Б) Источник, заявленная новизна, базовое сравнение и воспроизводимость В) Только наличие слова AI Г) Только привлекательность названия

11. Как поступить, если ограниченная модификация не улучшила качество?

А) Объявить улучшение без расчета Б) Скрыть исходный метод В) Повторно подбирать по тесту до выигрыша Г) Документировать результат и ограничения по тому же протоколу

12. Что входит в оценку ресурсов экспериментального прототипа?

А) Только название библиотеки Б) Только дата запуска В) Время, память или размер модели, объем данных и условия запуска Г) Только итоговая метрика

Ключ диагностического теста

1 – А; 2 – А; 3 – Г; 4 – В; 5 – А; 6 – Б; 7 – Г; 8 – В; 9 – Б; 10 – Б; 11 – Г; 12 – В.

3.2 Блок В Практические задания

Практические работы выполняются в объеме и по тематике РПД. Проверяются способ решения, самостоятельность, результаты и воспроизводимость. Оценка проводится по разделу 4; контрольные работы и итоговая защита проверяют отдельные действия и не дают повторных баллов за один результат.

Семестр 3. Общий объем практических занятий – 22 часа. Семь работ объединены в три модуля; задания и рубежный контроль выполняются в пределах этих часов. Подготовка данных, отчета и повторный запуск входят в самостоятельную работу соответствующих тем. Компетенции: ПК-8, ПК-13.

Практическая работа № 1. Паспорт данных и разведочный анализ

Модуль 1. Строки тематического плана: 1.4. Объем – 2 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Использовать учебный набор breast_cancer из scikit-learn; для регрессионной ветки допускается diabetes. Записать источник, единицу наблюдения, признаки, отклик, объем и ограничения. Данные рассматриваются только как учебный пример.

2. Проверить размерности, типы, пропуски, дубликаты и распределение отклика. Для breast_cancer явно перекодировать положительный класс как злокачественное образование; медицинских рекомендаций по учебному эксперименту не формулировать.

3. Отложить 20 % наблюдений в тест с seed = 2026 и стратификацией для классификации. Сохранить индексы. Разведочные графики и решения о преобразованиях строить по оставшейся обучающей части.

4. Рассчитать базовый прогноз: частота класса для классификации или среднее train для регрессии. Сформулировать вопрос сравнения и выбрать основную метрику до обучения моделей.

Материалы для сдачи и проверка

Паспорт данных, журнал проверок, 2–3 содержательных графика и обоснование базового ориентира. Проверить размер разбиений, отсутствие пересечений и наличие обоих классов.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

Практическая работа № 2. Воспроизводимая предобработка

Модуль 1. Строки тематического плана: 1.6. Объем – 4 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Продолжить работу с сохраненными разбиениями. Составить список числовых и категориальных полей. Пропуски, кодирование и масштабирование включить в ColumnTransformer и Pipeline.

2. Для проверки обработки пропусков создать копию обучающей таблицы с фиксированным маскированием части значений, сохранив seed и маску. Исходные данные и тест не перезаписывать.

3. Обучить заполнение и масштабирование только на train. На выданной заготовке с базовой моделью выполнить трехкратную CV, при которой преобразования переобучаются внутри каждого фолда. Описать опасность предварительного fit_transform всего набора.

4. Проверить совместимость столбцов при transform и predict, обработку неизвестной категории на отдельном искусственном примере и воспроизводимость повторного запуска.

Материалы для сдачи и проверка

Исполняемый конвейер, схема потока данных и таблица проверок. Передать параметры преобразований, индексы разбиений и объяснение отсутствия утечки.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

Практическая работа № 3. Базовые модели регрессии и классификации

Модуль 2. Строки тематического плана: 2.2. Объем – 2 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Для основной классификационной задачи сравнить DummyClassifier и LogisticRegression в конвейерах. На регрессионной ветке diabetes кратко сравнить DummyRegressor и линейную модель.

2. Использовать общие три фолда: стратифицированные для классификации, обычные для регрессии. Зафиксировать seed, метрики и состав признаков.

3. Для классификации показать precision, recall и F1; для регрессии RMSE и R^2. Одну матрицу ошибок и один расчет MSE проверить вручную на небольшом фрагменте.

4. Сопоставить модели по CV и объяснить, что улучшение над базовым ориентиром не доказывает применимость вне учебной выборки. Тест пока не открывать.

Материалы для сдачи и проверка

Таблица базового сравнения, формулы метрик и объяснение выбора модели. Оба алгоритма каждой задачи оцениваются на одинаковых наблюдениях.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

Практическая работа № 4. Кросс-валидация и ограниченный поиск параметров

Модуль 2. Строки тематического плана: 2.5. Объем – 2 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Для логистической регрессии выбрать три значения C, например 0,1; 1; 10. Выполнить поиск по одинаковым фолдам только на train.

2. Сохранить среднее качество, разброс и время каждого запуска. Объяснить различие параметра C и коэффициента L2-штрафа: увеличение C ослабляет регуляризацию.

3. Рассмотреть учебный кейс повторных наблюдений одного клиента и предложить GroupKFold; для временных данных описать последовательное разбиение. Обучать дополнительные модели для кейсов не требуется.

4. Зафиксировать выбранную конфигурацию и объяснить, почему лучшая CV-оценка после поиска не является независимым подтверждением качества.

Материалы для сдачи и проверка

Таблица трех конфигураций, обоснование выбора и две схемы разбиения. Указать, какие действия привели бы к утечке и как они исключены.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

Практическая работа № 5. Ансамбли и анализ ошибок

Модуль 2. Строки тематического плана: 2.8 и 2.11. Объем – 6 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Сравнить выбранную линейную модель, дерево глубины 3, случайный лес из 100 деревьев и градиентный бустинг с не более чем 100 итерациями. Использовать общий train и три фолда; не расширять сетку.

2. Записать качество, разброс и время обучения. По out-of-fold прогнозам разобрать ложноположительные и ложноотрицательные ошибки и связь метрики с условной стоимостью ошибок.

3. Построить permutation importance на валидационном фолде выбранного решения либо разобрать готовый пример. Указать влияние корреляции признаков и невозможность причинного вывода.

4. Выбрать базовое решение для проекта, сформулировать ограничение по памяти или времени и описать, какое изменение алгоритма или признаков будет проверено в работе № 7.

Материалы для сдачи и проверка

Единая таблица четырех моделей, анализ ошибок, интерпретация признаков и план одной модификации. Результаты могут быть отрицательными при корректном протоколе.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

Практическая работа № 6. Кластеризация и снижение размерности

Модуль 3. Строки тематического плана: 3.2. Объем – 2 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Использовать признаки того же набора без целевой переменной; стандартизацию обучить на train. Для описательного эксперимента обозначить, что известные метки не используются при построении кластеров.

2. Сравнить k-средних при k = 2 и 3 с фиксированным n_init и DBSCAN при двух заранее заданных eps. Указать число кластеров и долю шума; Для DBSCAN исключить шум с меткой −1 из расчета silhouette и отдельно указать его долю. Метрика определена, если на оставшихся наблюдениях число кластеров от 2 до N − 1, где N – число этих наблюдений. При разном составе оставшихся точек оценки не являются прямым сравнением на одной выборке.

3. Выполнить PCA, вывести объясненную дисперсию и проекцию. Проверить повторным запуском устойчивость k-средних и объяснить условность номеров кластеров.

4. Не трактовать совпадение проекции или кластера с известной меткой как доказательство причинной структуры. Для предсказательного конвейера объяснить необходимость PCA внутри CV.

Материалы для сдачи и проверка

Сводная таблица, проекция и вывод об устойчивости. При единственном кластере либо сплошном шуме честно указать неприменимость выбранной метрики.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

Практическая работа № 7. Исследовательский прототип и воспроизводимый отчет

Модуль 3. Строки тематического плана: 3.5. Объем – 4 ч. Максимум – 5 баллов.

Цель и порядок выполнения

1. Найти первичный научный или технический источник по одному изученному методу. Указать задачу, предпосылки, сравниваемый базовый подход и авторское утверждение о результате.

2. Предложить одну ограниченную модификацию: удаление группы признаков, настройку веса класса или проверку другого правила подготовки признаков. Сформулировать ожидаемый эффект до запуска.

3. Сравнить исходный и измененный конвейеры на общих фолдах при одинаковом бюджете. Записать качество, разброс, время и объем данных. Не выдавать средний выигрыш на нескольких фолдах за доказанную статистическую значимость.

4. Зафиксировать итоговый конвейер, обучить на полном train и однократно оценить на тесте. Подготовить код, README, файл зависимостей, разбиения и отчет 4–6 страниц.

5. Повторить запуск в чистом сеансе и объяснить расхождения. Защитить вывод, включая отсутствие улучшения, ограничения масштаба и условия возможного дальнейшего исследования.

Материалы для сдачи и проверка

Прототип, отчет, журнал экспериментов и проверка повторного запуска. Контейнер, веб-сервис и распределенное обучение не являются обязательными требованиями.

Критерии: корректность подготовки и протокола – 1 балл; выполнение основного эксперимента – 2; проверка результата – 1; отчет и самостоятельное объяснение – 1. Полный балл требует предъявления соответствующего результата; половина балла критерия допускается при частичном выполнении. Баллы за контрольную работу начисляются отдельно по подразделу 3.3.

3.3 Блок С Контрольные и исследовательские задания

Три рубежные работы проверяют ПК-8 и ПК-13. Максимумы: 10, 15 и 10 баллов; всего – 35. Практические работы дают еще 35 баллов. Итог за семестр – 70; экзамен – 30. Продолжительность каждой работы – до 45 минут в пределах практических занятий. Выполняется один вариант; вариант Б заменяет исходные данные варианта А.

Контрольная работа № 1. Данные и корректный эксперимент

Модуль 1. Максимум – 10 баллов. Ответ оформляется индивидуально: ход решения, вычисления, обоснование и ограничения.

Вариант А

1. Даны 240 наблюдений по 60 клиентам, по четыре записи на клиента. Разбиение выполнено случайно по строкам, медиана пропусков рассчитана по всей таблице. Найдите два источника утечки и предложите исправленный порядок. Максимум – 4 балла.

2. В обучающей части 180 наблюдений, 18 положительных исходов. Обоснуйте базовый ориентир, метрику и способ стратификации с учетом группировки. Максимум – 3 балла.

3. Составьте перечень материалов, по которым можно проверить источник, разбиение и воспроизводимость предобработки. Максимум – 3 балла.

Вариант Б

В варианте Б используются 300 наблюдений по 75 клиентам, по четыре записи на клиента. В исходной ошибочной схеме разбиение выполнено по строкам, а медиана рассчитана по всей таблице. При исправленном групповом разбиении в train – 224 строки по 56 клиентам, из них 27 положительных исходов; в test – 76 строк по 19 клиентам. Выполните три задания варианта А с этими данными; максимумы баллов сохраняются.

Проверка и критерии

Ориентир проверки: сначала разделение по клиентам, затем обучение преобразований внутри train/CV; баланс классов обеспечивается настолько, насколько допускают группы. Точное равенство долей классов не гарантируется.

Пункт	За что начисляются баллы	Максимум
1	Правильный основной вывод – 2; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	4
2	Правильный основной вывод – 1; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	3
3	Правильный основной вывод – 1; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	3

Внутри каждого критерия полный балл ставится за верное и проверяемое действие; половина максимума – за частичное выполнение без принципиальной ошибки; 0 – при отсутствии результата или неверном методе. Преподаватель отмечает основание частичной оценки. За один и тот же фрагмент проекта баллы повторно не начисляются.

Контрольная работа № 2. Сравнение моделей и анализ ошибок

Модуль 2. Максимум – 15 баллов. Ответ оформляется индивидуально: ход решения, вычисления, обоснование и ограничения.

Вариант А

1. По одинаковой CV три модели имеют F1 (среднее ± стандартное отклонение): A 0,72 ± 0,04; B 0,75 ± 0,06; C 0,74 ± 0,02. Время обучения: 2, 60 и 8 с. При лимите 10 с обоснуйте выбор и укажите, чего недостаточно для утверждения о статистически значимом превосходстве. Максимум – 5 баллов.

2. Для TP = 30, FN = 10, FP = 20, TN = 140 вычислите precision, recall и F1. При условной стоимости FN = 5 и FP = 1 рассчитайте суммарную стоимость ошибок; объясните, на каких данных можно выбирать порог. Максимум – 5 баллов.

3. Предложите план сравнения базового метода с одним изменением признаков. Укажите общие разбиения, бюджет, журнал результатов и момент использования независимого теста. Максимум – 5 баллов.

Вариант Б

A 0,68 ± 0,03; B 0,74 ± 0,08; C 0,71 ± 0,02, время 2, 90 и 7 с, лимит 10 с. Для расчета: TP = 24, FN = 16, FP = 12, TN = 148; стоимости те же.

Проверка и критерии

Ориентир проверки: в A precision = 0,6, recall = 0,75, F1 = 2/3, стоимость = 70; в Б precision = 2/3, recall = 0,6, F1 ≈ 0,6316, стоимость = 92. Выбор модели требует учета ограничения времени; статистическая значимость по одним средним и отклонениям не установлена.

Пункт	За что начисляются баллы	Максимум
1	Правильный основной вывод – 3; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	5
2	Правильный основной вывод – 3; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	5
3	Правильный основной вывод – 3; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	5

Внутри каждого критерия полный балл ставится за верное и проверяемое действие; половина максимума – за частичное выполнение без принципиальной ошибки; 0 – при отсутствии результата или неверном методе. Преподаватель отмечает основание частичной оценки. За один и тот же фрагмент проекта баллы повторно не начисляются.

Контрольная работа № 3. Методы без учителя и исследовательский прототип

Модуль 3. Максимум – 10 баллов. Ответ оформляется индивидуально: ход решения, вычисления, обоснование и ограничения.

Вариант А

1. Для стандартизованных данных k-средних дает silhouette 0,42 и 0,45 при k = 2 и 3; DBSCAN – один кластер и 35 % шума. Обоснуйте, какие сравнения допустимы, и предложите проверку устойчивости. Максимум – 3 балла.

2. Источник утверждает, что модификация улучшает качество на 5 %. Укажите, какие сведения о метрике, базе сравнения, разбиениях и затратах нужны для проверки этого заявления. Максимум – 3 балла.

3. Представьте схему воспроизведения проекта другим исследователем: получение данных, зависимости, запуск, контрольные результаты и ограничения. Объясните, как документируется отрицательный результат модификации. Максимум – 4 балла.

Вариант Б

silhouette 0,38 и 0,40, DBSCAN пометил все наблюдения как шум; источник обещает уменьшение времени обучения на 30 %. Структура задания и баллы сохраняются.

Проверка и критерии

Ориентир проверки: после исключения шума silhouette не определен для единственного кластера или отсутствия кластеров; шум не считается дополнительным содержательным кластером; проверяются параметры, масштабирование и повторные запуски. Процентное улучшение интерпретируется только относительно явно указанной базы.

Пункт	За что начисляются баллы	Максимум
1	Правильный основной вывод – 1; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	3
2	Правильный основной вывод – 1; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	3
3	Правильный основной вывод – 2; обоснованный расчет или порядок действий – 1; проверка и ограничение – 1.	4

Внутри каждого критерия полный балл ставится за верное и проверяемое действие; половина максимума – за частичное выполнение без принципиальной ошибки; 0 – при отсутствии результата или неверном методе. Преподаватель отмечает основание частичной оценки. За один и тот же фрагмент проекта баллы повторно не начисляются.

3.4 Блок D Промежуточная аттестация

Экзамен включает три элемента по 10 баллов: постановку задачи и методы, проверку эксперимента и интерпретацию, защиту исследовательского прототипа. Максимум – 30 баллов. Во всех вариантах студент предъявляет собственный проект.

Образец билета № 1

1. По первичному источнику сопоставьте выбранный метод с базовым подходом и определите ограниченную исследовательскую модификацию. (10 баллов.)

2. При TP = 30, FN = 10, FP = 20, TN = 140 вычислите precision, recall, F1 и стоимость ошибок при цене FN = 5 и FP = 1; объясните выбор порога. (10 баллов.)

3. Защитите свой прототип: общие фолды, бюджет, журнал, фиксация итоговой модели и однократная оценка на тесте. (10 баллов.)

Образец билета № 2

1. Для повторных записей по клиентам обоснуйте групповое разбиение, предобработку и сравнение методов. (10 баллов.)

2. У моделей A и B F1 на CV равны 0,72 ± 0,04 и 0,75 ± 0,06. Объясните, почему этих чисел недостаточно для вывода о значимости, и предложите проверку с учетом времени обучения. (10 баллов.)

3. Представьте модификацию своего прототипа, качество, затраты и повторный запуск; обоснуйте пределы обобщения. (10 баллов.)

Образец билета № 3

1. Сопоставьте k-средних, DBSCAN и PCA по задачам и ограничениям; сформулируйте исследовательский вопрос. (10 баллов.)

2. После исключения шума DBSCAN оставил один кластер. Объясните неприменимость silhouette и предложите проверку устойчивости без вымышленной численной оценки. (10 баллов.)

3. Защитите отчет о прототипе, включая отрицательный результат, ограничения объема данных и условия воспроизведения. (10 баллов.)

Ориентиры проверки итоговых образцов

Вариант 1. precision = 0,6; recall = 0,75; F1 = 2/3; стоимость = 10·5 + 20 = 70. Порог выбирают до итогового теста.

Вариант 2. Нужны результаты на согласованных разбиениях и обоснованная процедура сравнения; зависимые CV-оценки нельзя безусловно считать независимой выборкой.

Вариант 3. Silhouette требует $2 \leq$ число кластеров $\leq n - 1$; шум не считают вторым содержательным кластером. Проверить масштабы, параметры, повторные запуски и устойчивость.

Раздел 4 Процедуры критерии и шкалы оценивания

За семестр – 70 баллов: семь практических работ по 5 баллов (35) и три контрольные работы на 10, 15 и 10 баллов (35). Максимумы модулей – 20, 30 и 20. Экзамен – 30; общий максимум – 100.

Практическая работа	Максимум
Подготовка данных и корректный протокол	1
Основной эксперимент и выполнение задания	2
Проверка результата	1
Отчет и самостоятельное объяснение	1
Итого за одну работу	5

Критерии каждой контрольной работы приведены в подразделе 3.3. Один результат оценивается один раз. Работы № 1–2 и контрольная № 1 относятся к модулю 1; работы № 3–5 и контрольная № 2 – к модулю 2; работы № 6–7 и контрольная № 3 – к модулю 3.

Экзамен	Максимум
Постановка задачи и методы	10
Проверка эксперимента и интерпретация	10
Защита проекта и ответы	10

По каждому элементу экзамена: правильность и полнота – до 5 баллов; аргументация – до 3; самостоятельность и ответы на уточнения – до 2. Полный балл требует подтверждения результата; частичная оценка сопровождается пояснением.

Шкала итоговой оценки: $85 \leq B \leq 100$ – отлично; $70 \leq B < 85$ – хорошо; $60 \leq B < 70$ – удовлетворительно; $0 \leq B < 60$ – неудовлетворительно, где B – сумма баллов за семестр и экзамен. Для положительной оценки необходимо не менее 40 баллов за семестр и 20 за экзамен. Баллы

сохраняются без округления. При невыполнении любого из двух общих минимумов итоговая оценка – «неудовлетворительно», даже если сумма В достигла 60.

Итоговая оценка и общие минимумы

Итоговый балл В	Оценка
$85 \le B \le 100$	отлично
$70 \le B < 85$	хорошо
$60 \le B < 70$	удовлетворительно
$B < 60$ либо не выполнен один из общих минимумов	неудовлетворительно

В – сумма баллов за семестр и промежуточную аттестацию. Положительная оценка требует не менее 40 баллов за семестр и 20 за промежуточную аттестацию. Значения минимумов строк технологической карты раскрывают распределение семестрового минимума; баллы начисляются за подтвержденный результат, от 0 до максимума.

Порядок проведения контроля

Текущий контроль проводится при сдаче и обсуждении практической работы. Студент предъявляет материалы, объясняет выбранный метод, воспроизводит ключевую проверку и отвечает на уточняющие вопросы. Устный опрос помогает проверить личное понимание и учитывается внутри критериев соответствующей работы.

Рубежная проверка выполняется индивидуально в пределах предусмотренных практических занятий. В каждой контрольной работе выполняется один вариант. Вариант Б заменяет вариант А; их баллы не суммируются. Продолжительность и допустимые средства указаны в подразделе 3.3. Результат сопровождается ходом решения и проверкой.

Промежуточная аттестация проводится по материалам подраздела 3.4: преподаватель сообщает выбранный вариант, порядок ответа и время по расписанию. Проверяются знания, самостоятельность выполнения и обоснованность выводов. Критерии доводятся до обучающихся до начала контроля; необязательные расширения программы не требуются для максимума.

Преподаватель фиксирует баллы по отдельным критериям и основание частичной оценки. Полный балл означает выполненное и проверяемое действие; 0 – отсутствие результата либо принципиально неверное решение. Один результат оценивается один раз. Баллы сохраняются без округления. Порядок допуска, повторной сдачи и оформления результатов определяется действующими локальными актами университета.

Раздел 5 Методические указания обучающемуся

5.1 Организация самостоятельной работы

Начните с целей, тематического плана и планируемых результатов РПД. Для каждого модуля сопоставьте вопросы подраздела 3.1 с практическими и контрольными заданиями. Изучайте литературу, указанную в актуальной РПД; номера источников в практических заданиях относятся к этой библиографии.

После занятия восстановите основные определения и алгоритмы, выполните небольшой расчет, отметьте затруднения. Диагностический тест используйте для поиска пробелов; объясните, почему отвергнуты остальные ответы. Непонятный вопрос сформулируйте до консультации с преподавателем.

5.2 Подготовка к практическим и контрольным работам

До занятия изучите цель и исходные данные, проверьте доступность инструментов и ограничений работы. Сохраняйте оригиналы данных, промежуточные результаты и принятые решения. В отчете укажите задачу, метод, исходные условия, ход выполнения, независимую проверку, выводы и ограничения.

На защите будьте готовы воспроизвести ключевой расчет, объяснить результат и диагностировать ошибку. Для контрольной работы подготовьте определения и формулы по

соответствующему модулю; выполните пробный вариант с обоснованием каждого действия, затем проверьте его по критериям.

5.3 Подготовка к промежуточной аттестации

Повторите вопросы трех групп «Знать», «Уметь», «Владеть», разберите образцы билетов и критерии раздела 4. Подготовьте собственные отчеты, исходные данные и материалы проекта или портфолио. Тренируйтесь отличать наблюдаемый результат от предположения и формулировать границы вывода.

Сохраните разбиения, параметры, версии библиотек и инструкцию запуска. Обучаемые преобразования выполняйте только на train или внутри CV; итоговый тест используйте после фиксации модели. Подготовьте честное сравнение с базовым решением и объяснение ошибок без повторного подбора по тесту.