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

УТВЕРЖДАЮ
директор Политехнического института
_____ Г.В. Гоцев
3 сентября 2026 г.

Машинное обучение

рабочая программа дисциплины (модуля)

Закреплена за **Высшая школа Цифровых технологий**

Учебный план g090404_26_1 пи_оаирп.plx
Направление подготовки 09.04.04 - РФ, 710400 - КР Программная инженерия
Магистерская программа "Основы анализа и разработки приложений с большими объемами распределенных данных"

Квалификация **магистр**

Форма обучения **очная**

Общая трудоемкость **4 ЗЕТ**

Часов по учебному 144

в том числе:

аудиторные занятия 38

самостоятельная 105

Виды контроля в семестрах:
экзамен 3

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	**3 (2.1)**		Итого	
Недель	19			
Вид занятий	УП	РП	УП	РП
Лекции	16	16	16	16
Практические	22	22	22	22
Контактная работа в период теоретического	1	1	1	1
В том числе инт.	8	8	8	8
В том числе в форме практ.подготовки	22	22	22	22
Итого ауд.	38	38	38	38
Контактная работа	39	39	39	39
Сам. работа	105	105	105	105
Итого	144	144	144	144

Программу составил(и):
д.т.н., профессор, Лыченко Н.М.;д.т.н., профессор, Верзунов С.Н. _____

Рецензент(ы):
к.т.н., доцент, Хмелева И.В. _____

Согласовано РОП: Лыченко Н.М.
_____

Рабочая программа дисциплины
Машинное обучение

разработана в соответствии с ФГОС 3++:
Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 09.04.04 Программная инженерия (приказ Минобрнауки России от 19.09.2017 г. № 932)

составлена на основании учебного плана:

Направление подготовки 09.04.04 - РФ, 710400 - КР Программная инженерия
Магистерская программа "Основы анализа и разработки приложений с большими объемами распределенных данных"
утвержденного учёным советом вуза от 26.06.2026 протокол № 11.

Рабочая программа одобрена на заседании ВШ Цифровые технологии

Протокол от 02.09.2026 г. № 1
Срок действия программы: 2026-2028 уч.г.
Руководитель подразделения: Хмелева Ирина Владимировна

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ
к рабочей программе дисциплины (модуля)

Машинное обучение

_____ *наименование дисциплины*

№ п/п	Основание внесения изменения	Раздел/ пункт	Содержание изменения
1.			
2.			
3.			
4.			
5.			

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	
1.1	Цель дисциплины – сформировать у магистрантов способность находить и критически оценивать новые подходы к машинному обучению, строить воспроизводимые модели и исследовательские прототипы алгоритмов для данных различного объема с учетом качества, вычислительных затрат и ограничений применения.
1.2	Задачи: освоить подготовку данных и построение регрессионных, классификационных и кластеризационных моделей; находить и оценивать новые методы; проводить корректное сравнение алгоритмов по качеству и вычислительным затратам; обосновывать и проверять ограниченную модификацию модели в исследовательском проекте.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП	
Цикл (раздел) ООП:	Б1.В
2.1	**Требования к предварительной подготовке обучающегося:**
2.1.1	Знание линейной алгебры, теории вероятностей и математической статистики; навыки программирования на Python и работы с табличными данными. Достаточно базового опыта с NumPy, pandas и визуализацией.
2.2	**Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:**
2.2.1	Результаты применяются в исследовательской, технологической и преддипломной практике, разработке приложений с большими объемами распределенных данных и подготовке выпускной квалификационной работы.

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
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
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
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

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
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
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

В результате освоения дисциплины обучающийся должен

3.1	**Знать:**
3.1.1	этапы анализа данных, основные задачи обучения с учителем и без учителя.
3.1.2	метрики, схемы разбиения, кросс-валидации и поиска новых методов.
3.1.3	ограничения моделей при росте объема данных и требования к воспроизводимому сравнению.

3.2	**Уметь:**
3.2.1	подготовить данные, выявить пропуски, ошибки и ограничения источника.
3.2.2	найти метод, построить конвейер и сравнить базовую модель с альтернативой.
3.2.3	проверить ограниченную модификацию метода по качеству и затратам, сформулировать ограничения.
3.3	**Владеть:**
3.3.1	приемами документирования данных, кода, источников и результатов эксперимента.
3.3.2	инструментами Python для анализа данных, визуализации и обучения моделей.
3.3.3	навыками подготовки воспроизводимого отчета и защиты исследовательского прототипа.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетен-ции	Литература	Инте ракт.	Пр. подг.	Примечание
	Раздел 1. Подготовка данных и разведочный анализ							
1.1	Постановка задачи анализа данных; качество, происхождение и структура наборов данных /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2			
1.2	Постановка задачи анализа данных; качество, происхождение и структура наборов данных /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
1.3	Разведочный анализ, визуализация, пропуски и выбросы /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2	1		Интерактивный разбор кейса: обсуждение альтернатив и совместная проверка решения.
1.4	Разведочный анализ, визуализация, пропуски и выбросы /Пр/	3	2	ПК-13 ПК-8	Л1.1 Л1.2		2	
1.5	Разведочный анализ, визуализация, пропуски и выбросы /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
1.6	Воспроизводимая предобработка: разбиение, кодирование, масштабирование и конвейеры /Пр/	3	4	ПК-13 ПК-8	Л1.1 Л1.2		4	
1.7	Воспроизводимая предобработка: разбиение, кодирование, масштабирование и конвейеры /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
	Раздел 2. Обучение с учителем: регрессия, классификация и оценка качества моделей							
2.1	Регрессия и классификация: базовые модели и корректное сравнение /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2	1		Интерактивный разбор кейса: обсуждение альтернатив и совместная проверка решения.
2.2	Регрессия и классификация: базовые модели и корректное сравнение /Пр/	3	2	ПК-13 ПК-8	Л1.1 Л1.2		2	

2.3	Регрессия и классификация: базовые модели и корректное сравнение /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
2.4	Метрики, кросс-валидация, утечка данных и выбор гиперпараметров /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2	1		Интерактивный разбор кейса: обсуждение альтернатив и совместная проверка решения.
2.5	Метрики, кросс-валидация, утечка данных и выбор гиперпараметров /Пр/	3	2	ПК-13 ПК-8	Л1.1 Л1.2		2	
2.6	Метрики, кросс-валидация, утечка данных и выбор гиперпараметров /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
	Раздел 3. Ансамблевые методы обучения с учителем, анализ ошибок и интерпретация моделей							
3.1	Регуляризация, деревья решений, случайный лес и градиентный бустинг /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2			
3.2	Регуляризация, деревья решений, случайный лес и градиентный бустинг /Пр/	3	4	ПК-13 ПК-8	Л1.1 Л1.2	2	4	Работа в парах: взаимная проверка протокола эксперимента и обсуждение результатов.
3.3	Регуляризация, деревья решений, случайный лес и градиентный бустинг /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
3.4	Анализ ошибок, дисбаланс классов и интерпретация результатов /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2			
3.5	Анализ ошибок, дисбаланс классов и интерпретация результатов /Пр/	3	2	ПК-13 ПК-8	Л1.1 Л1.2		2	
3.6	Анализ ошибок, дисбаланс классов и интерпретация результатов /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
	Раздел 4. Обучение без учителя и воспроизводимый исследовательский проект							
4.1	Кластеризация и снижение размерности: k-средних, DBSCAN, PCA /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2	1		Интерактивный разбор кейса: обсуждение альтернатив и совместная проверка решения.
4.2	Кластеризация и снижение размерности: k-средних, DBSCAN, PCA /Пр/	3	2	ПК-13 ПК-8	Л1.1 Л1.2		2	
4.3	Кластеризация и снижение размерности: k-средних, DBSCAN, PCA /Ср/	3	12	ПК-13 ПК-8	Л1.1 Л1.2			
4.4	Масштабирование данных и исследовательский прототип /Лек/	3	2	ПК-13 ПК-8	Л1.1 Л1.2			

4.5	Масштабирование данных и исследовательский прототип /Пр/	3	4	ПК-13 ПК-8	Л1.1 Л1.2	2	4	Работа в парах: взаимная проверка протокола эксперимента и обсуждение результатов.
4.6	Масштабирование данных и исследовательский прототип /Ср/	3	13	ПК-13 ПК-8	Л1.1 Л1.2			
4.7	Подготовка воспроизводимого проекта и защиты /Ср/	3	10	ПК-13 ПК-8	Л1.1 Л1.2			
4.8	Подготовка воспроизводимого проекта и защиты /КрТО/	3	1	ПК-13 ПК-8	Л1.1 Л1.2			

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

5.1. Контрольные вопросы и задания

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.

5.2. Темы курсовых работ (проектов)

Курсовая работа учебным планом не предусмотрена. В рамках практических занятий выполняется небольшой аналитический проект.

5.3. Фонд оценочных средств

Задания, критерии и технологическая карта приведены в документе ФОС.

5.4. Перечень видов оценочных средств
Семь практических работ – 35 баллов; три рубежные контрольные работы – 35 баллов. Итого текущий и рубежный контроль – 70. Промежуточная аттестация – экзамен, 30 баллов. Формы заданий и критерии приведены в документе ФОС.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы,	Заглавие	Издательство, год
Л1.1	Рашка С., Мирджалили В	. Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2	. 3-е изд. М.: Вильямс, 2024. 848 с. ISBN 978-5-907203-57-0.
Л1.2	Джеймс Г., Уиттен Д., Хасти Т., Тибширани Р.,	Введение в статистическое обучение с примерами на языке Python М	.: ДМК Пресс, 2024
Л1.3	Бишоп К. М.	Распознавание образов и машинное обучение / пер. с англ. Д. А. Клюшина.	М.; СПб.: Диалектика, 2020. 960 с. ISBN 978-5-907144-55-2.

6.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство, год
Л2.1	Шолле Ф	. Глубокое обучение на Python (дополнительное чтение для продолжения)	СПб.: Питер, 2023
Л2.2	Hastie T., Tibshirani R., Friedman J.	The Elements of Statistical Learning.	2nd ed. Springer, 2009. URL: https://hastie.su.domains/ElemStatLearn/
Л2.3	Pedregosa F. et al.	Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. 2011. Vol. 12. P. 2825–2830. URL: https://jmlr.org/papers/v12/pedregosa11a.html	Journal of Machine Learning Research. 2011. Vol. 12
Л2.4	Virtanen P. et al. SciPy 1.0:	Fundamental Algorithms for Scientific Computing in Python // Nature Methods. 2020. Vol. 17. P. 261–272. DOI: 10.1038/s41592-019-0686-2. URL:	Nature Methods. 2020. Vol. 17.
Л2.5	Wilson G. et al.	Good enough practices in scientific computing // PLOS Computational Biology. 2017. Vol. 13(6). e1005510. DOI:	PLOS Computational Biology. 2017. Vol. 13(6).

6.1.3. Методические разработки

	Авторы,	Заглавие	Издательство, год
Л3.1	Верзунов С. Н.	Практические работы и требования к отчетам по дисциплине «Машинное обучение». Учебные материалы дисциплины, 2026.	Электронная версия
Л3.2	Верзунов С. Н.	Контрольные работы и критерии оценивания по дисциплине «Машинное обучение». Учебные материалы дисциплины, 2026.	Электронная версия
Л3.3	Верзунов С. Н.	Шкалы оценивания и организация самостоятельной работы по дисциплине «Машинное обучение». Учебные материалы дисциплины, 2026.	Электронная версия

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Официальное руководство scikit-learn. URL	https://scikit-learn.org/stable/user_guide.
Э2	Учебные материалы авторов ISLP. URL	https://www.statlearning.com/

6.3. Перечень информационных и образовательных технологий	
6.3.1 Компетентностно-ориентированные образовательные технологии	
6.3.1.1	Разбор кейсов, работа с открытыми и учебными наборами данных, взаимная проверка аналитических отчетов, проектная защита.
6.3.2 Перечень информационных справочных систем и программного обеспечения	
6.3.2.1	Python, Jupyter, NumPy, pandas, Matplotlib, scikit-learn; Git для хранения версий кода и отчета. Допускаются функционально равнозначные свободные инструменты.

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)	
7.1	Компьютерный класс с рабочими местами для Python/Jupyter, доступом к учебным наборам данных, проектором и возможностью представления результатов. Облачные вычислительные мощности для обязательных заданий не требуются.

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)
Вести проект на одном наборе данных: зафиксировать источник и ограничения, отделить тестовую выборку до подбора преобразований, сохранить воспроизводимый код и параметры эксперимента. Найти и критически оценить выбранный метод; предложить небольшую модификацию, сравнить ее с базовым подходом по качеству и затратам. Отчет и код представляются по заданиям приложения 3.