

СИЛЛАБУС
Осенний семестр 2025 -2026 учебного года
Образовательная программа «БВ06105 – Математическое и компьютерное моделирование»

ID и наименование дисциплины	Самостоятельная работа обучающегося (СРО)	Кол-во кредитов			Общее кол-во кредитов	Самостоятельная работа обучающегося под руководством преподавателя (СРОП)
		Лекции (Л)	Практ. занятия (ПЗ)	Лаб. занятия (ЛЗ)		
97726 Машинное обучение	2	1,7	3,3		5	6
АКАДЕМИЧЕСКАЯ ИНФОРМАЦИЯ О ДИСЦИПЛИНЕ						
Формат обучения	Цикл, модуль компонент	Типы лекций	Типы практических занятий		Форма и платформа итогового контроля	
Офлайн	Базовый, Компонент по выбору	Аналитическая лекция	Решение задач		Письменный экзамен	
Лектор - (ы)	Сабирова Юлия Фархатовна					
e-mail:	juliasabirova23@gmail.com					
Телефон:	+77273773330					
Ассистент- (ы)	Сабирова Юлия Фархатовна					
e-mail:	juliasabirova23@gmail.com					
Телефон:	+77273773330					
АКАДЕМИЧЕСКАЯ ПРЕЗЕНТАЦИЯ ДИСЦИПЛИНЫ						
Цель дисциплины	Ожидаемые результаты обучения (РО)*			Индикаторы достижения РО (ИД)		
	РО 1. Демонстрировать знание и понимание основополагающих концепций, задач и методов машинного обучения.			1.1. Знает и умеет объяснять ключевые понятия машинного обучения (обучение с учителем/без учителя, признаки, целевая переменная, классификация, регрессия, кластеризация)		
				1.2 Понимает и интерпретирует принципы работы базовых алгоритмов машинного обучения (линейные модели, метрические методы, деревья решений).		
	РО 2. Применять инструментальные средства языка Python для предварительной обработки, анализа и визуализации данных.			2.1 Умеет использовать библиотеки NumPy и Pandas для загрузки, очистки, фильтрации и агрегации табличных данных.		
				2.2 Проводит исследовательский анализ данных (EDA) с помощью библиотек Matplotlib и Seaborn для выявления закономерностей и подготовки данных к моделированию.		
	РО 3. Выбирать, обучать и оценивать модели машинного обучения для решения практических задач регрессии и классификации.			3.1 Умеет реализовывать полный цикл построения модели (разделение выборки, обучение fit(), получение предсказаний predict()) с использованием библиотеки Scikit-learn.		
				3.2 Проводит оценку качества и надежности моделей с применением релевантных метрик (MAE, R², Accuracy, Precision, Recall) и техники кросс-валидации.		
	РО 4. Проектировать и реализовывать комплексные решения задач анализа данных, сравнивая различные подходы и обосновывая выбор итоговой модели.			4.1 Проводит сравнительный анализ нескольких алгоритмов машинного обучения на одном наборе данных для определения наиболее эффективного в рамках поставленной задачи.		

		4.2 Формулирует выводы по результатам машинного обучения и умеет представлять итоги исследования в рамках учебного проекта.
	РО 5. Критически оценивать применимость и ограничения различных алгоритмов машинного обучения в зависимости от характеристик задачи и данных.	5.1 Владеет навыком определения сильных и слабых сторон алгоритмов (например, чувствительность к масштабу признаков, интерпретируемость, вычислительная сложность) и их соответствия конкретному набору данных.
		5.2 Владеет способностью аргументированно обосновывать выбор или отказ от определенного алгоритма, основываясь на анализе его теоретических свойств и практических требований задачи (например, потребность в интерпретируемости модели, ограничения по времени обучения).
Пререквизиты	Программирование на Python, Линейная алгебра, Теория вероятности и математическая статистика, Математический анализ	
Постреквизиты	Глубокое обучение (Deep Learning), Продвинутые методы машинного обучения	
Учебные ресурсы	<p>Литература:</p> <ol style="list-style-type: none"> 1. Géron, A. Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow : concepts, tools, and techniques to build intelligent systems / A. Géron. – 3rd ed. – Sebastopol : O'Reilly Media, 2022. – 856 p. – ISBN 978-1098125974. 2. McKinney, W. Python for Data Analysis : data wrangling with pandas, NumPy, and IPython / W. McKinney. – 3rd ed. – Sebastopol : O'Reilly Media, 2022. – 634 p. – ISBN 978-1098104030. 3. Raschka, S. Python Machine Learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow 2 / S. Raschka, V. Mirjalili. – 3rd ed. – Birmingham : Packt Publishing, 2019. – 770 p. – ISBN 978-1789955750. 4. Müller, A. C. Introduction to Machine Learning with Python : a guide for data scientists / A. C. Müller, S. Guido. – Sebastopol : O'Reilly Media, 2016. – 398 p. – ISBN 978-1449369415. 5. Flach, P. Machine Learning : the art and science of algorithms that make sense of data / P. Flach. – Cambridge : Cambridge University Press, 2012. – 416 p. – ISBN 978-1107096394. <p>Интернет-ресурсы</p> <ol style="list-style-type: none"> 1. http://elibrary.kaznu.kz/ru 2. Course: Машинное обучение в Python: Machine Learning & Data Science Udemы https://www.udemy.com/course/python-machine-learning-data-science-russian/ 3. https://github.com/yuliya-sabirova/ml-course <p>Программное обеспечение</p> <ol style="list-style-type: none"> 1. www.anaconda.com 	

Академическая политика дисциплины	<p>Академическая политика дисциплины определяется <u>Академической политикой и Политикой академической честности КазНУ имени аль-Фараби</u>. Документы доступны на главной странице ИС Univer.</p> <p>Интеграция науки и образования. Научно-исследовательская работа студентов, магистрантов и докторантов – это углубление учебного процесса. Она организуется непосредственно на кафедрах, в лабораториях, научных и проектных подразделениях университета, в студенческих научно-технических объединениях. Самостоятельная работа обучающихся на всех уровнях образования направлена на развитие исследовательских навыков и компетенций на основе получения нового знания с применением современных научно-исследовательских и информационных технологий. Преподаватель исследовательского университета интегрирует результаты научной деятельности в тематику лекций и семинарских (практических) занятий, лабораторных занятий и в задания СРОП, СРО, которые отражаются в силлабусе и отвечают за актуальность тематик учебных занятий и заданий.</p> <p>Посещаемость. Дедлайн каждого задания указан в календаре (графике) реализации содержания дисциплины. Несоблюдение дедлайнов приводит к потере баллов.</p> <p>Академическая честность. Практические/лабораторные занятия, СРО развивают у обучающегося самостоятельность, критическое мышление, креативность. Недопустимы плагиат, подлог, использование шпаргалок, списывание на всех этапах выполнения заданий. Соблюдение академической честности в период теоретического обучения и на экзаменах помимо основных политик регламентируют <u>«Правила проведения итогового контроля», «Инструкции для проведения итогового контроля осеннего/весеннего семестра текущего учебного года», «Положение о проверке текстовых документов обучающихся на наличие заимствований»</u>. Документы доступны на главной странице ИС Univer.</p> <p>Основные принципы инклюзивного образования. Образовательная среда университета задумана как безопасное место, где всегда присутствуют поддержка и равное отношение со стороны преподавателя ко всем обучающимся и обучающимся друг к другу независимо от гендерной, расовой/ этнической принадлежности, религиозных убеждений, социально-экономического статуса, физического здоровья студента и др. Все люди нуждаются в поддержке и дружбе ровесников и сокурсников. Для всех студентов достижение прогресса скорее в том, что они могут делать, чем в том, что не могут. Разнообразие усиливает все стороны жизни. Все обучающиеся, особенно с ограниченными возможностями, могут получать консультативную помощь по e-mail juliasabirova23@gmail.com либо посредством видеосвязи, предварительно написав преподавателю в MS Teams Sabirova.yuliya@kaznu.kz.</p> <p>Интеграция MOOC (massive open online course). В случае интеграции MOOC в дисциплину, всем обучающимся необходимо зарегистрироваться на MOOC. Сроки прохождения модулей MOOC должны неукоснительно соблюдаться в соответствии с графиком изучения дисциплины.</p> <p>ВНИМАНИЕ! Дедлайн каждого задания указан в календаре (графике) реализации содержания дисциплины, а также в MOOC. Несоблюдение дедлайнов приводит к потере баллов.</p>
--	---

ИНФОРМАЦИЯ О ПРЕПОДАВАНИИ, ОБУЧЕНИИ И ОЦЕНИВАНИИ

Балльно-рейтинговая буквенная система оценки учета учебных достижений				Методы оценивания	
Оценка	Цифровой эквивалент баллов	Баллы, % содержание	Оценка по традиционной системе	<p>Критериальное оценивание – процесс соотнесения реально достигнутых результатов обучения с ожидаемыми результатами обучения на основе четко выработанных критериев. Основано на формативном и суммативном оценивании.</p> <p>Формативное оценивание – вид оценивания, который проводится в ходе повседневной учебной деятельности. Является текущим показателем успеваемости. Обеспечивает оперативную взаимосвязь между обучающимся и преподавателем. Позволяет определить возможности обучающегося, выявить трудности, помочь в достижении наилучших результатов, своевременно корректировать преподавателю образовательный процесс. Оценивается выполнение заданий, активность работы в аудитории во время лекций, семинаров, практических занятий (дискуссии, викторины, дебаты, круглые столы, лабораторные работы и т. д.). Оцениваются приобретенные знания и компетенции.</p> <p>Суммативное оценивание – вид оценивания, который проводится по завершению изучения раздела в соответствии с программой дисциплины. Проводится 3-4 раза за семестр при выполнении СРО. Это оценивание освоения ожидаемых результатов обучения в соотнесенности с дескрипторами. Позволяет определять и фиксировать уровень освоения дисциплины за определенный период. Оцениваются результаты обучения.</p>	
A	4,0	95-100	Отлично	Формативное и суммативное оценивание	
A-	3,67	90-94		Баллы % содержание	
B+	3,33	85-89	Хорошо	Активность на лекциях	5
B	3,0	80-84		Работа на практических занятиях	20
B-	2,67	75-79		Самостоятельная работа	25
C+	2,33	70-74		Проектная и творческая деятельность	10
C	2,0	65-69		Итоговый контроль (экзамен)	40
C-	1,67	60-64		ИТОГО	100
D+	1,33	55-59	Удовлетворительно		
D	1,0	50-54			
FX	0.5	25-49			
			Неудовлетворительно		

F	0	0-24				
Календарь (график) реализации содержания дисциплины. Методы преподавания и обучения.						
Неделя	Название темы				Кол-во часов	Макс. балл
МОДУЛЬ 1. Инструментальные средства анализа данных и машинного обучения						
1	Л 1. Введение в машинное обучение. Основные понятия и задачи	1				
	СЗ 1. Работа в интерактивной среде Jupyter. Основы синтаксиса языка Python.	2				
2	Л 2. Библиотека NumPy. Основы работы с многомерными массивами	1				
	СЗ 2. Реализация векторных вычислений с использованием библиотеки NumPy	2				
3	Л 3. Библиотека Pandas. Структуры данных Series и DataFrame. Загрузка и первичный анализ данных.	1				
	СЗ 3. Выполнение исследовательского анализа данных с применением библиотеки Pandas	2				
	СРОП 1. Консультация по пройденным темам	1				
4	Л 4. Продвинутые методы анализа данных в Pandas. Группировка и агрегация данных.	1				
	СЗ 4. Выполнение исследовательского анализа данных. Часть 2: Агрегация и обработка пропусков.	2			16	
5	Л 5. Визуализация данных. Использование библиотек Matplotlib и Seaborn	1				
	СЗ 5. Построение и анализ графиков для различных типов данных	2			16	
МОДУЛЬ 2. Методы обучения с учителем: регрессия и классификация						
6	Л 6. Жизненный цикл проекта в ML. Метод линейной регрессии	1				
	СЗ 6. Построение и анализ моделей. Часть 1: Реализация линейной регрессии	2			16	
7	Л 7. Оценка качества моделей. Метрики и метод кросс-валидации. Регуляризация.	1				
	СЗ 7. Построение и анализ моделей. Часть 2: Применение метрик и кросс-валидации	2			16	
	СРОП 2. Консультации по выполнению СРО 1	1				
8	Л 8. Задача классификации. Метод логистической регрессии и метрики качества	1				
	СЗ 8. Сравнение моделей классификации. Часть 1: Реализация логистической регрессии	2			16	
	СРОП 3. Консультации по выполнению СРО 1	1				
	СРО 1. Исследовательский анализ данных (EDA).	1			20	
Рубежный контроль 1					100	
9	Л 9. Основы инженерии признаков. Кодирование категориальных и масштабирование численных признаков	1				
	СЗ 9. Сравнение моделей классификации. Часть 2: Предобработка данных.	2				
10	Л 10. Метрические методы. Алгоритм k-ближайших соседей (KNN)	1				
	СЗ 10. Сравнение моделей классификации. Часть 3: Реализация и анализ KNN.	2			16	
11	Л 11. Метод опорных векторов (SVM)	1				
	СЗ 11. Реализация продвинутых методов классификации. Часть 1: SVM	2				
	СРОП 4. Консультация по пройденным темам.	1				
МОДУЛЬ 3. Продвинутые алгоритмы классического машинного обучения						
12	Л12. Логические методы. Алгоритм дерева решений	1				
	СЗ 12. Реализация продвинутых методов классификации. Часть 2: Деревья решений.	2			16	
13	Л 13. Ансамблевые методы. Алгоритм Random Forest и обзор методов бустинга.	1				
	СЗ 13. Применение ансамблевых методов на примере Random Forest	2			16	
14	Л 14. Обучение без учителя. Метод кластеризации K-Means	1				
	СЗ 14. Применение методов кластеризации для сегментации данных. Часть 1.	2			16	
	СРОП 5. Консультация по выполнению СРО 2.					
15	Л 15. Обзор продвинутых тем: метод главных компонент (PCA), основы обработки естественного языка (NLP), глубокое обучение (Deep Learning)	1				
	СЗ 15. Применение методов кластеризации для сегментации данных. Часть 2.	2			16	
	СРОП 6. Консультация по выполнению СРО 2.	1				
	СРО 2. Групповой проект по построению моделей машинного обучения.	1			20	
Рубежный контроль 2					100	
Итоговый контроль (экзамен)					100	
ИТОГО за дисциплину					100	

Декан _____ Досжан Н.С.

Председатель Академического комитета
по качеству преподавания и обучения _____ Ахметова Б.И.

Заведующий кафедрой _____ Бектемесов Ж.М.

Лектор _____ Сабирова Ю.Ф.

**РУБРИКАТОР СУММАТИВНОГО ОЦЕНИВАНИЯ
КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ**

Название задания: СРО 1. Исследовательский анализ данных (EDA). (20 баллов, 20% от РК)

Методы преподавания и обучения: Проектно-ориентированное обучение, исследовательский метод.

Критерий	«Отлично» 20%	«Хорошо» 15%	«Удовлетворительно» 10%	«Неудовлетворительно» 0%
1. Загрузка и первичный анализ данных	Данные корректно загружены. Применены все основные методы первичного анализа (.head, .info, .describe). Сделаны краткие, но содержательные выводы по структуре и статистикам данных.	Данные загружены. Применена большая часть методов первичного анализа. Выводы носят поверхностный характер или отсутствуют.	Данные загружены, но применен только один метод анализа (например, только head()) без каких-либо выводов.	Данные не загружены, или первичный анализ не проведен.
2. Предобработка и очистка данных	Проведен полный анализ пропущенных значений. Выбрана и обоснована стратегия их обработки (удаление/заполнение), которая корректно реализована.	Пропущенные значения найдены и обработаны, но стратегия их обработки не обоснована.	Пропущенные значения найдены, но не обработаны, либо обработаны некорректно.	Пропущенные значения не были проанализированы.
3. Визуальный анализ данных	Использовано не менее 3-х различных и уместных типов графиков. Все графики имеют заголовки и подписанные оси. Каждый график сопровождается текстовой интерпретацией.	Использовано 2-3 типа графиков, но некоторые из них неинформативны или не имеют необходимых подписей. Интерпретация слабая или отсутствует.	Построен только один тип графика, либо использованы неподходящие типы визуализации.	Визуальный анализ не проведен.
4. Глубина анализа и качество выводов	Сформулировано не менее 3-х нетривиальных, подкрепленных данными выводов о взаимосвязях в данных. Выдвинуты обоснованные гипотезы для дальнейшего моделирования.	Выводы в основном сводятся к констатации очевидных фактов (например, "среднее значение признака X равно Y").	Сформулирован только один вывод, либо выводы не подкреплены данными/графиками.	Выводы отсутствуют, либо являются неверными.
5. Структура и оформление отчета (Jupyter Notebook)	Ноутбук имеет четкую структуру с использованием заголовков Markdown. Код чистый и хорошо читаемый. Каждому логическому блоку предшествуют текстовые пояснения.	Структура присутствует, но пояснения к коду минимальны или отсутствуют в некоторых разделах.	Весь код представлен единым блоком без структуры и комментариев ("простыня кода").	Ноутбук неструктурирован и нечитаем.

Формула расчета итоговой оценки:

*Итоговая оценка = (% по критерию 1 + % по критерию 2 + % по критерию 3 + % по критерию 4 + % по критерию 5)*Макс.число баллов*

**РУБРИКАТОР СУММАТИВНОГО ОЦЕНИВАНИЯ
КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ**

Название задания: СРО 2. Групповой проект по построению моделей машинного обучения. (20 баллов, 20% от РК)

Методы преподавания и обучения: Командная работа, проектно-ориентированное обучение, кейс-стади.

Критерий	«Отлично» 20%	«Хорошо» 15%	«Удовлетворительно» 10%	«Неудовлетворительно» 0%
1. Исследовательский анализ данных (EDA) и постановка задачи	Проведен глубокий EDA, выводы которого напрямую используются для дальнейшей предобработки и выбора моделей. Бизнес/исследовательская задача четко сформулирована.	EDA проведен, но его результаты слабо связаны с последующими этапами проекта. Постановка задачи носит общий характер.	Проведен поверхностный EDA (1-2 графика) без глубоких выводов.	EDA или постановка задачи отсутствуют.
2. Инженерия признаков (Feature Engineering)	Все необходимые шаги предобработки (обработка пропусков, кодирование категорий, масштабирование) выполнены корректно и обоснованы с точки зрения их необходимости для конкретных моделей.	Основные шаги предобработки выполнены, но отсутствует обоснование выбора методов, либо допущены незначительные ошибки.	Пропущены важные шаги предобработки (например, не проведено масштабирование для KNN/SVM), либо они реализованы с существенными ошибками.	Данные не были подготовлены для моделирования.
3. Построение и обучение моделей	Обучено не менее 4-х концептуально различных моделей в соответствии с заданием. Код для обучения и получения предсказаний реализован корректно.	Обучено 2-3 модели, либо допущены незначительные ошибки в реализации.	Обучена только одна модель, либо в коде присутствуют серьезные ошибки, влияющие на результат.	Модели не были обучены.
4. Оценка, сравнение моделей и выбор лучшей	Выбраны и корректно рассчитаны адекватные задаче метрики. Проведено четкое сравнение всех моделей (например, в виде итоговой таблицы). Выбор лучшей модели подробно аргументирован .	Метрики рассчитаны верно, но сравнение моделей поверхностное, а выбор лучшей модели не аргументирован или очевиден.	Использованы неподходящие метрики, либо допущены ошибки в их расчете. Сравнение моделей отсутствует.	Оценка моделей не проведена.
5. Качество итогового отчета (Jupyter Notebook) и презентации	Ноутбук представляет собой полноценный, структурированный отчет. Презентация ясно и лаконично отражает все этапы работы. Команда уверенно отвечает на вопросы.	Отчет и презентация содержат всю необходимую информацию, но им не хватает структуры или ясности изложения.	Ноутбук неаккуратно оформлен. Презентация нелогична, либо команда слабо ориентируется в своем же проекте.	Отчетные материалы не предоставлены.

Формула расчета итоговой оценки):

Итоговая оценка = (% по критерию 1 + % по критерию 2 + % по критерию 3 + % по критерию 4 + % по критерию 5)*Макс.число баллов