

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет компьютерных технологий и прикладной математики

УТВЕРЖДАЮ:

Проректор по учебной работе,
качеству образования – первый
проректор

Хагуров Т.А.

« 29 » августа 2025 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)
Б1. В.03 Искусственный интеллект в оценке рисков и разработке страховых
продуктов

Направление подготовки 02.03.01 Математика и компьютерные науки

Профиль Искусственный интеллект и аналитика данных

Форма обучения очная

Квалификация бакалавр

Краснодар 2025

Рабочая программа дисциплины «Искусственный интеллект в оценке рисков и разработке страховых продуктов» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.03.01 Математика и компьютерные науки

Программу составил(и):
Е.Н. Калайдин, доктор физ.-мат. наук


_____ подписать

Рабочая программа дисциплины утверждена на заседании центра
искусственного интеллекта
протокол № 01 «28» августа 2025 г.
Руководитель центра ИИ Коваленко А.В.


_____ подписать

Утверждена на заседании учебно-методической комиссии факультета
компьютерных технологий и прикладной математики
протокол № 01 «28» августа 2025 г.
Председатель УМК факультета Коваленко А.В.


_____ подписать

Рецензенты:
Мостовой Евгений Викторович, генеральный директор ООО «Портал-Юг»,
e-mail: mostovoy@portal-yug.ru

Луценко Евгений Вениаминович, доктор экономических наук, кандидат
технических наук, профессор кафедры компьютерных технологий и систем
Федерального государственного бюджетное образовательное учреждение
высшего образования «Кубанский государственный аграрный университет
имени И.Т. Трубилина», e-mail: prof.lutsenko@gmail.com

1 Цели и задачи изучения дисциплины (модуля)

1.1 Цель освоения дисциплины

Сформировать у студентов системное понимание современных методов искусственного интеллекта и машинного обучения для решения прикладных задач в страховании: от точной оценки рисков до создания новых, продуктов и оптимизации бизнес-процессов.

1.2 Задачи дисциплины

– Изучить математический аппарат ключевых ML-моделей, применимых в страховании.

– Научить студентов критически оценивать применимость и ограничения различных моделей для конкретных страховых задач (life, non-life, health).

– Сформировать навыки построения, валидации и интерпретации моделей на реальных (или приближенных к реальным) страховых данных.

– Рассмотреть этические и регуляторные аспекты использования AI в страховании (дискриминация, "черный ящик", Solvency II/IFRS 17).

1.3 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Искусственный интеллект в оценке рисков и разработке страховых продуктов» относится к части, формируемой участниками образовательных отношений Блока 1 "Дисциплины (модули)" учебного плана.

Пререквизиты:

– Вероятность и математическая статистика (продвинутый уровень)

– Основы актуарной математики (моделирование продолжительности жизни, теория рисков)

– Линейная алгебра и машинное обучение (базовые модели: линейная регрессия, логистическая регрессия, деревья решений)

– Владение языком программирования Python (NumPy, Pandas, Scikit-learn).

1.4 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Роль 1: Data Analyst (Аналитик данных)

Задачи:

1. Статистический анализ, визуализация данных, предварительная обработка.

2. Создание прогнозных моделей

3. Построение аналитических моделей для поддержки бизнес-решений.

Роль 2: MLOps (Специалист по эксплуатации ИИ)

Задачи:

1. DevOps для ML.

2. Автоматизация, мониторинг ML-систем.

3. Операционное управление жизненным циклом ML-моделей.

Роль 3: AI PM (Менеджер проектов ИИ)

Задачи:

1. Управление ИИ-проектами от идеи до внедрения

2. Анализ бизнес-требований и постановка задач

3. Оценка эффективности и ROI ИИ-решений

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код и наименование индикатора	Результаты обучения по дисциплине
ПК-2	Способен участвовать в исследовании новых математических моделей в прикладных областях

Код и наименование индикатора	Результаты обучения по дисциплине
П-2.1 Умеет анализировать и адаптировать существующие математические модели для решения прикладных задач в конкретной предметной области	Умеет проводить сравнительный анализ классических актуарных моделей (GLM, Chain-Ladder) и современных ML-моделей (Бустинг, НС), выявляя их применимость, ограничения и адаптируя их для задач тарификации и резервирования
П-2.2 Способен предлагать и обосновывать новые математические подходы для моделирования процессов в прикладных исследованиях	Знает и математически обосновывает использование продвинутых подходов (копулы для моделирования зависимостей, survival analysis для цензурированных данных, генеративные сети для синтеза данных) в рамках страховых исследований
	Умеет разрабатывать и обосновать новые или гибридные модели для решения поставленной задачи
Е-1 Способен осуществлять трудовые функции, обусловленные профессиональной ролью, в ОПД «Экономика, финансы и управление»	
Е-1.1. Применяет методы и технологии организации и управления данными и знаниями в финансовой сфере	Умеет применять техники Feature Engineering и ETL для подготовки специфических страховых данных (работа с категориальными переменными, пропусками, временными рядами) в рамках построения ML-пайплайна.
Е-1.2. Применяет современные методы и технологии ИИ для решения задач прогнозирования финансовой сфере	Умеет применять модели Gradient Boosting и Neural Networks для прогнозирования ключевых страховых метрик: частоты и severity убытков, итоговой стоимости убытка, вероятности наступления события.
	Владеет реализацией и валидацией моделей прогнозирования
Е-1.3. Применяет современные методы и технологии ИИ для решения задач оценки рисков и управления рисками в финансовой сфере	Владеет навыками использования и построения ML-моделей для количественной оценки страховых рисков (расчет pure premium, прогнозирование резервов, оценка долгосрочных обязательств) и анализом их влияния на управление капиталом.
SS-1 Способен осуществлять свою трудовую деятельность с учетом определения корректной роли ИИ в различных процессах, критического анализа последствий применения ИИ-технологий, этических принципов	
SS-1.1 Определяет ценностные предпосылки, когнитивные искажения, культурно-обусловленные предвзятости в данных, алгоритмах, постановке задач для ИИ	Знает потенциальные источники bias в страховых данных (например, дискриминация по полу, возрасту, почтовому индексу) и в постановке задач (например, оптимизация только под прибыль, а не под справедливость)
	Умеет анализировать справедливость модели с помощью SHAP и fairness-метрик
SS-1.2 Применяет методики работы с этическими и социальными рисками, возникающими на разных стадиях жизненного цикла ИИ	Знает методики смягчения выявленных рисков (подбор справедливых признаков, регуляризация, post-processing) и рассматривает жизненный цикл модели (MLOps) с точки зрения контроля этических рисков
	Умеет включать в итоговый отчет раздела об этических аспектах и ограничениях модели

Результаты обучения по дисциплине достигаются в рамках осуществления всех видов контактной и самостоятельной работы обучающихся в соответствии с утвержденным учебным планом.

Индикаторы достижения компетенций считаются сформированными при достижении соответствующих им результатов обучения.

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 3 зачетных единиц (108 часов), их распределение по видам работ представлено в таблице

Виды работ	Всего часов	Форма обучения
		очная
		7 семестр (часы)
Контактная работа, в том числе:	52,3	52,3
Аудиторные занятия (всего):	50	50
занятия лекционного типа	16	16
лабораторные занятия	34	34
практические занятия		
семинарские занятия		
Иная контактная работа:	2,3	2,3
Контроль самостоятельной работы (КСР)	2	2
Промежуточная аттестация (ИКР)	0,3	0,3
Самостоятельная работа, в том числе:	20	20
Самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам и т.д.)	15	15
Подготовка к текущему контролю	5	5
Контроль:	35,7	35,7
Подготовка к экзамену	35,7	35,7
Общая трудоемкость	108	108
час.		
в том числе контактная работа	52,3	52,3
зач. ед	3	3

2.2 Содержание дисциплины

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

Разделы (темы) дисциплины, изучаемые в 7 семестре (очная форма обучения)

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	
1.	Введение. Data Foundation и особенности страховых данных.	6	2		4	
2.	Продвинутые модели для тарификации (Pricing) в Non-Life страховании	14	4		6	4
3.	Резервирование убытков (Claim Reserving) и прогнозирование.	12	2		6	4
4.	AI в Life и Health страховании	12	2		6	4
5.	Генеративные модели и создание новых продуктов.	14	4		6	4
6.	Оптимизация, интерпретация и регуляторика.	12	2		6	4
	ИТОГО по разделам дисциплины	70	16		34	20
	Контроль самостоятельной работы (КСР)	2				
	Промежуточная аттестация (ИКР)	0,3				
	Подготовка к текущему контролю	35,7				
	Общая трудоемкость по дисциплине	108				

Примечание: Л – лекции, ПЗ – практические занятия / семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

2.3 Содержание разделов (тем) дисциплины

2.3.1 Занятия лекционного типа

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1.	Введение. Data Foundation и особенности страховых данных.	<ul style="list-style-type: none"> – Эволюция актуарных моделей: от таблиц смертности к AI. Обзор современных вызовов (Big Data, новые риски). – Особенности страховых данных: цензурирование (censoring), усечение (truncation), правые хвосты (heavy tails), разреженность (sparsity). Подготовка данных (feature engineering) для страховых тарифов. – Проблема "информационной асимметрии" (adverse selection, moral hazard) и как AI может помочь в ее решении. 	<i>Вопросы экзамена 1-6</i>
2.	Продвинутое моделирование для тарификации (Pricing) в Non-Life страховании	<ul style="list-style-type: none"> – Обобщенные линейные модели (GLM) как бенчмарк. Интерпретируемость vs. точность. – Градиентный бустинг (Gradient Boosting Machines: XGBoost, LightGBM, CatBoost) для прогнозирования частоты и severity убытков. Анализ важности признаков. – Нейронные сети (Feed-Forward NN) для тарификации. Архитектуры, функции активации, борьба с переобучением. – Модели для учета зависимости (dependency modeling): копулы в сочетании с ML-моделями. 	<i>Вопросы экзамена 7-13</i>
3.	Резервирование убытков (Claim Reserving) и прогнозирование.	<ul style="list-style-type: none"> – Ограничения классических методов (Chain-Ladder). Введение в ML-альтернативы. – Модели на основе градиентного бустинга и случайного леса для прогнозирования индивидуальных убытков и их развития. – Нейронные сети с временными рядами (LSTM, GRU) для анализа т-реугольника убытков. 	<i>Вопросы экзамена 14-16</i>
4.	AI в Life и Health страховании	<ul style="list-style-type: none"> – Модели прогнозирования продолжительности жизни и заболеваемости с использованием методов цензурированных данных (модели Кокса, survival analysis с помощью Random Survival Forests, Gradient Boosting for Survival Analysis) – Анализ медицинских изображений и геномных данных (краткий обзор): возможности и этические ограничения. 	<i>Вопросы экзамена 17-19</i>
5.	Генеративные модели и создание новых продуктов.	<ul style="list-style-type: none"> – Модели прогнозирования продолжительности жизни и заболеваемости с использованием методов цензурированных данных (модели Кокса, survival analysis с помощью Random Survival Forests, Gradient Boosting for Survival Analysis) – Использование AI для проектирования параметрического страхования (parametric insurance) на основе данных с датчиков (IoT) 	<i>Вопросы экзамена 20-22</i>
6.	Оптимизация, интерпретация и регуляторика.	<ul style="list-style-type: none"> – Методы интерпретации сложных моделей (SHAP, LIME). Борьба с дискриминацией (fairness-aware ML). – Интеграция AI-моделей в стресс-тестирование и расчет капитала (Solvency II). Влияние на IFRS 17 – Производственная реализация ML-моделей (MLOps) в страховых компаниях 	<i>Вопросы экзамена 23-27</i>

2.3.2 Занятия семинарского типа (практические / семинарские занятия/ лабораторные работы)

№	Наименование раздела (темы)	Тематика занятий/работ	Форма текущего контроля
1.	Основы обработки страховых данных в Python	<ul style="list-style-type: none"> – Загрузка и первичный анализ данных (например, портфель автострахования). – Обработка пропущенных значений и категориальных переменных (One-Hot Encoding, Target Encoding). – Анализ распределений ключевых переменных: частота убытков, стоимость убытка (severity). 	Очищенный и готовый к моделированию датасет

2.	Бенчмарк-модели: Обобщенные Линейные Модели (GLM)	<ul style="list-style-type: none"> – Разделение данных на обучающую и тестовую выборки с учетом временного фактора. – Построение GLM для частоты убытков (Пуассон/Негативное Биномиальное распределение). – Построение GLM для severity убытков (Гамма/Логнормальное распределение). 	Расчет тарифов по GLM. Интерпретация коэффициентов
3.	Деревья решений и ансамбли (Случайный Лес)	<ul style="list-style-type: none"> – Построение и визуализация одного дерева решений для предсказания вероятности убытка. – Построение модели Случайного Леса для регрессии и классификации. – Сравнение точности с GLM. 	Модель Random Forest. Анализ важности признаков (Feature Importance)
4.	Градиентный бустинг (LightGBM/XGBoost)	<ul style="list-style-type: none"> – Построение моделей LightGBM для прогноза частоты и severity убытков. – Подбор гиперпараметров с помощью кросс-валидации. – Сравнение производительности с Random Forest и GLM. 	Настроенная модель бустинга. Сравнительная таблица метрик (Gini, MAE, RMSE)
5.	Нейронные сети для табличных данных	<ul style="list-style-type: none"> – Создание архитектуры FNN с использованием PyTorch/TensorFlow или scikit-learn. – Обучение модели, мониторинг процесса обучения (loss curves). – Сравнение производительности с бустингом. 	Обученная нейросетевая модель. Понимание ее сильных и слабых сторон
6.	Survival Analysis для Life-страхования	<ul style="list-style-type: none"> – Построение модели Кокса (Cox Proportional Hazards model). – Использование Random Survival Forests для прогнозирования кривых выживаемости. – Расчет вероятности дожития для конкретного клиента. 	Модель для прогнозирования продолжительности и жизни/оттока клиентов.
7.	Прогнозирование резервов убытков с помощью ML	<ul style="list-style-type: none"> – Подготовка треугольника убытков в формате, пригодном для ML. – Построение модели (например, LightGBM) для прогнозирования итоговой стоимости каждого заявленного убытка. – Сравнение итогового резерва с методом Chain-Ladder. 	Оценка общего резерва, полученная ML-методом.
8.	Интерпретация моделей (XAI) и Fairness	<ul style="list-style-type: none"> – Применение SHAP (SHapley Additive exPlanations) для анализа вклада признаков в лучшую модель (например, бустинг). – Анализ индивидуальных предсказаний. – Проверка модели на наличие непреднамеренной дискриминации по защищенным признакам (пол, возраст). 	SHAP-графики. Отчет о "справедливости" модели.
9.	Интеграция: Мини-проект "Сквозной пайплайн"	<ul style="list-style-type: none"> – Студентам дается новый, более сложный набор данных (например, с пропусками, новыми признаками). – Самостоятельно выполнить полный цикл: EDA, feature engineering, построение 2-3 моделей, выбор лучшей, ее интерпретация. – Подготовка краткого отчета с рекомендациями для бизнеса. 	Notebook (Jupyter или Colab) с полным решением
10.	Защита мини-проектов	Студенты представляют свои работы из ЛР №9 в формате коротких (5-7 минут) презентаций с ответами на вопросы от преподавателя и одногруппников.	Умение аргументировать свой выбор и доносить результаты до аудитории.

Защита лабораторной работы (ЛР), выполнение курсового проекта (КП), курсовой работы (КР), расчетно-графического задания (РГЗ), написание реферата (Р), эссе (Э), коллоквиум (К), тестирование (Т) и т.д.

2.3.3 Примерная тематика курсовых работ (проектов)

Не предусмотрены учебным планом

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	Изучение теоретического материала	Методические указания по организации самостоятельной работы студентов, утвержденные кафедрой анализа данных и искусственного интеллекта, протокол №1 от 30.08.2023
2	Решение задач	Методические указания по организации самостоятельной работы студентов, утвержденные кафедрой информационных технологий, протокол №1 от 30.08.2019

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии, применяемые при освоении дисциплины (модуля)

В соответствии с требованиями ФГОС в программа дисциплины предусматривает использование в учебном процессе следующих образовательные технологии: чтение лекций с использованием мультимедийных технологий; метод малых групп, разбор практических задач и кейсов.

При обучении используются следующие образовательные технологии:

- Технология коммуникативного обучения – направлена на формирование коммуникативной компетентности студентов, которая является базовой, необходимой для адаптации к современным условиям межкультурной коммуникации.

- Технология разноуровневого (дифференцированного) обучения – предполагает осуществление познавательной деятельности студентов с учётом их индивидуальных способностей, возможностей и интересов, поощряя их реализовывать свой творческий потенциал. Создание и использование диагностических тестов является неотъемлемой частью данной технологии.

- Технология модульного обучения – предусматривает деление содержания дисциплины на достаточно автономные разделы (модули), интегрированные в общий курс.

- Информационно-коммуникационные технологии (ИКТ) - расширяют рамки образовательного процесса, повышая его практическую направленность, способствуют интенсификации самостоятельной работы учащихся и повышению познавательной активности. В рамках ИКТ выделяются 2 вида технологий:

- Технология использования компьютерных программ – позволяет эффективно дополнить процесс обучения языку на всех уровнях.

- Интернет-технологии – предоставляют широкие возможности для поиска информации, разработки научных проектов, ведения научных исследований.

- Технология индивидуализации обучения – помогает реализовывать личностно-ориентированный подход, учитывая индивидуальные особенности и потребности учащихся.

- Проектная технология – ориентирована на моделирование социального взаимодействия учащихся с целью решения задачи, которая определяется в рамках профессиональной подготовки, выделяя ту или иную предметную область.

- Технология обучения в сотрудничестве – реализует идею взаимного обучения, осуществляя как индивидуальную, так и коллективную ответственность за решение учебных задач.

- Игровая технология – позволяет развивать навыки рассмотрения ряда возможных способов решения проблем, активизируя мышление студентов и раскрывая личностный потенциал каждого учащегося.

- Технология развития критического мышления – способствует формированию разносторонней личности, способной критически относиться к информации, умению отбирать информацию для решения поставленной задачи.

Комплексное использование в учебном процессе всех вышеназванных технологий стимулируют личностную, интеллектуальную активность, развивают познавательные процессы, способствуют формированию компетенций, которыми должен обладать будущий специалист.

Основные виды интерактивных образовательных технологий включают в себя:

- работа в малых группах (команде) - совместная деятельность студентов в группе под руководством лидера, направленная на решение общей задачи путём творческого сложения результатов индивидуальной работы членов команды с делением полномочий и ответственности;

- проектная технология - индивидуальная или коллективная деятельность по отбору, распределению и систематизации материала по определенной теме, в результате которой составляется проект;

- анализ конкретных ситуаций - анализ реальных проблемных ситуаций, имевших место в соответствующей области профессиональной деятельности, и поиск вариантов лучших решений;

- развитие критического мышления – образовательная деятельность, направленная на развитие у студентов разумного, рефлексивного мышления, способного выдвинуть новые идеи и увидеть новые возможности.

Подход разбора конкретных задач и ситуаций широко используется как преподавателем, так и студентами во время лекций, лабораторных занятий и анализа результатов самостоятельной работы. Это обусловлено тем, что при исследовании и

решении каждой конкретной задачи имеется, как правило, несколько методов, а это требует разбора и оценки целой совокупности конкретных ситуаций.

Темы, задания и вопросы для самостоятельной работы призваны сформировать навыки поиска информации, умения самостоятельно расширять и углублять знания, полученные в ходе лекционных и практических занятий.

Подход разбора конкретных ситуаций широко используется как преподавателем, так и студентами при проведении анализа результатов самостоятельной работы.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

4. Оценочные и методические материалы

4.1 Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «название дисциплины».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме отчетов по лабораторным работам и **промежуточной аттестации** в форме вопросов и заданий к экзамену.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Структура оценочных средств для текущей и промежуточной аттестации

№ п/п	Код и наименование индикатора	Темы лекций	Наименование оценочного средства	
			Текущий контроль (Задания лабораторных работ)	Промежуточная аттестация (Экзаменационные вопросы)
1	ПК-2.1: Умеет анализировать и адаптировать существующие математические модели для решения прикладных задач в конкретной предметной области	2,3	2, 3, 4, 7	№7, №8, №14
2	ПК-2.2: Способен предлагать и обосновывать новые математические подходы для моделирования процессов в прикладных исследованиях	2,4,5	6,7	№13, №17, №20
3	Е1.1: Применяет методы и технологии организации и управления данными и знаниями в финансовой сфере	2	1, 6	№7, №8
4	Е1.2: Применяет современные методы и технологии ИИ для решения задач прогнозирования в финансовой сфере	3	4,5,7	№14, №15, №16
5	Е1.3: Применяет современные методы и технологии ИИ для решения задач оценки рисков и управления рисками в финансовой сфере	4, 6,	2, 4, 6, 7	№17, №25, №26
6	SS-1.1: Определяет ценностные предпосылки, когнитивные искажения, культурно-обусловленные предвзятости в данных, алгоритмах, постановке задач для ИИ.	1	8	№19, №24, №29
7	SS-1.2: Применяет методики работы с этическими и социальными рисками, возникающими на разных стадиях жизненного цикла ИИ.	1	8	№24, №27, №30

Показатели, критерии и шкала оценки сформированных компетенций

Для каждого индикатора (ПК 2.1, ПК 2.2, Е 1.1, Е 1.2, Е 1.3, SS-1.1, SS-1.2) определены три уровня освоения компетенции:

ПК-2.1: Умеет анализировать и адаптировать существующие математические модели для решения прикладных задач в конкретной предметной области

Уровень освоения	Критерии оценки
Пороговый (2)	Знает теоретические основы классических (GLM, Chain-Ladder) и современных ML-моделей (Бустинг, НС) Может воспроизвести стандартную модель по заданному алгоритму Понимает базовые ограничения методов на концептуальном уровне
Базовый (3)	Проводит сравнительный анализ моделей по ключевым метрикам (Gini, RMSE, log-loss) Аргументированно выбирает модель для типовой страховой задачи

Уровень освоения	Критерии оценки
	Адаптирует гиперпараметры моделей для улучшения результатов Корректно интерпретирует результаты разных моделей
Продвинутый (4)	Проводит комплексный анализ, учитывая бизнес-контекст и вычислительную сложность Предлагает адаптацию моделей для специфических страховых продуктов (киберстрахование) Критически оценивает адекватность моделей для нетривиальных случаев Синтезирует подходы, создавая гибридные модели

ПК-2.2: Способен предлагать и обосновывать новые математические подходы для моделирования процессов в прикладных исследованиях

Уровень освоения	Критерии оценки
Пороговый (2)	Знает математический аппарат продвинутых методов (копулы, survival analysis, GAN) Понимает области их потенциального применения в страховании Может воспроизвести стандартную реализацию метода
Базовый (3)	Предлагает обоснованное применение advanced-методов для конкретных страховых задач Корректно реализует нетривиальные математические подходы Сравнивает эффективность новых подходов с традиционными
Продвинутый (4)	Творчески адаптирует математические методы из смежных областей для страховых задач Предлагает модификации существующих алгоритмов для специфических потребностей Разрабатывает и математически обосновывает новые подходы к моделированию

E1.1: Применяет методы и технологии организации и управления данными и знаниями в финансовой сфере

Уровень освоения	Критерии оценки
Пороговый (2)	Владеет базовыми техниками ETL для страховых данных Выполняет стандартную предобработку данных (кодирование, масштабирование) Знает основные типы страховых данных и их особенности
Базовый (3)	Грамотно работает со специфическими страховыми данными (цензурирование, усечение) Создает содержательные features на основе доменных знаний Оптимизирует пайплайн обработки данных для конкретной задачи
Продвинутый (4)	Разрабатывает комплексные стратегии работы с данными для всего жизненного цикла модели Создает инновационные признаки, значительно улучшающие качество моделей Проектирует архитектуру данных для промышленного внедрения моделей

E1.2: Применяет современные методы и технологии ИИ для решения задач прогнозирования в финансовой сфере

Уровень освоения	Критерии оценки
Пороговый (2)	Знает архитектуру и принципы работы современных ML-моделей Может реализовать стандартные модели по готовым примерам Понимает базовые метрики качества прогнозирования
Базовый (3)	Грамотно настраивает гиперпараметры моделей для максимизации точности Осознанно выбирает метрики качества в зависимости от бизнес-задачи Интерпретирует результаты прогнозирования в предметной области

Уровень освоения	Критерии оценки
Продвинутый (4)	Создает ансамбли моделей для повышения точности прогнозирования Разрабатывает специализированные архитектуры для конкретных страховых продуктов Критически анализирует и улучшает весь пайплайн прогнозирования

E1.3: Применяет современные методы и технологии ИИ для решения задач оценки рисков и управления рисками в финансовой сфере

Уровень освоения	Критерии оценки
Пороговый (2)	Понимает связь между прогнозами ML-моделей и страховыми рисками Знает основные виды страховых рисков и подходы к их quantification Может рассчитать базовые страховые метрики на основе прогнозов
Базовый (3)	Интегрирует результаты ML-моделей в расчет страховых премий и резервов Оценивает финансовые последствия прогнозов для страховой компании Анализирует влияние моделей на риск-менеджмент и капитал
Продвинутый (4)	Разрабатывает комплексные системы оценки рисков на основе ensemble ML-моделей Моделирует стресс-сценарии и их влияние на финансовую устойчивость Создает системы поддержки принятия решений для риск-менеджеров

SS-1.1: Определяет ценностные предпосылки, когнитивные искажения, культурно-обусловленные предвзятости в данных, алгоритмах, постановке задач для ИИ

Уровень освоения	Критерии оценки
Пороговый (2)	Знает основные типы bias в данных и алгоритмах Понимает этические принципы применения AI в страховании Может идентифицировать явные случаи дискриминации в данных
Базовый (3)	Систематически выявляет скрытые bias в страховых данных и features Анализирует постановку задачи на предмет этических рисков Применяет количественные метрики для оценки fairness моделей
Продвинутый (4)	Проактивно идентифицирует комплексные этические риски на всех этапах проекта Разрабатывает методологии выявления сложных форм дискриминации Критически анализирует ценностные предпосылки в бизнес-требованиях

SS-1.2: Применяет методики работы с этическими и социальными рисками, возникающими на разных стадиях жизненного цикла ИИ

Уровень освоения	Критерии оценки
Пороговый (2)	Знает основные методы mitigation bias (pre-, in-, post-processing) Понимает принципы ответственного AI и регуляторные требования Следует этическим guidelines при реализации моделей
Продвинутый (3)	Осознанно выбирает и применяет методы снижения этических рисков Разрабатывает стратегии мониторинга моделей на предмет drift и bias Учитывает этические аспекты при проектировании ML-пайплайнов
Высокий (4)	Разрабатывает комплексные системы управления этическими рисками Создает методологии балансировки business objectives и ethical constraints Формирует корпоративные стандарты ответственного AI в страховании

Формула итоговой оценки:

Итог = 0.3 × (средний балл по ПК-2) + 0.4 × (средний балл по E1) + 0.3 × (средний балл по SS-1)

Критерии оценивания результатов обучения

Шкала перевода в традиционные оценки:

- **5 (Отлично):** 3.6 - 4.0 (преобладание высоких и продвинутых уровней)
- **4 (Хорошо):** 2.6 - 3.5 (преобладание продвинутых уровней)
- **3 (Удовлетворительно):** 2.0 - 2.5 (преобладание пороговых уровней)
- **2 (Неудовлетворительно):** < 2.0 (не достигнут пороговый уровень)

Типовая лабораторная работа иллюстрирующая материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Типовая лабораторная работа №4: "Градиентный бустинг для тарификации в страховании"

1. Методические указания

Цель работы: Освоить практическое применение алгоритмов градиентного бустинга для решения задач актуарного ценообразования и сформировать профессиональные компетенции в области современных методов оценки рисков.

Задачи:

- Сформировать навыки построения и оптимизации моделей градиентного бустинга
- Освоить методики оценки качества актуарных моделей
- Развить умения интерпретации результатов для бизнес-применений

2. Исходные данные

Набор данных: Portoseguro Safe Driver Prediction (адаптированный)

- **Объем:** 50 000 полисов автострахования
- **Целевые переменные:**
 - o claim_count - количество убытков
 - o claim_amount - сумма убытка

- **Признаки:** демографические данные, характеристики vehicle, история вождения

Программное обеспечение: Python 3.8+, библиотеки: pandas, numpy, matplotlib, seaborn, scikit-learn, lightgbm, shap

3. Задание

Этап 1. Подготовка данных и базовое моделирование

1.1. Разделить данные на обучающую (70%), валидационную (20%) и тестовую (10%) выборки

1.2. Построить baseline-модель на основе GLM с негативным биномиальным распределением

1.3. Рассчитать метрики качества: Gini, MAE, Poisson deviance

Этап 2. Построение модели градиентного бустинга

2.1. Обучить модель LightGBM для прогнозирования frequency убытков

2.2. Подобрать оптимальные гиперпараметры с использованием cross-validation

2.3. Построить модель для severity убытков с учетом zero-inflated data

Этап 3. Анализ и интерпретация результатов

3.1. Провести сравнительный анализ качества моделей

3.2. Проанализировать важность признаков с использованием SHAP

3.3. Рассчитать итоговые страховые премии на основе полученных моделей

4. Содержание отчета

Раздел 1. Теоретическая часть

- Математическое обоснование выбранных методов
- Описание алгоритма градиентного бустинга
- Обоснование выбора метрик качества

Раздел 2. Практическая реализация

- Код подготовки данных и построения моделей
- Визуализация процесса обучения и валидации
- Сравнительные таблицы метрик качества

Раздел 3. Анализ результатов

- Интерпретация важности признаков
- Сравнение с классическими актуарными моделями
- Рекомендации по применению в страховой практике

5. Критерии оценки

Пороговый уровень (2 балла):

- Выполнены этапы 1.1-1.3
- Построена baseline-модель с базовыми метриками
- Понимание теоретических основ методов

Базовый: (3 балла):

- Дополнительно выполнены этапы 2.1-2.3
- Проведена оптимизация гиперпараметров
- Качественный сравнительный анализ моделей

Продвинутый (4 балла):

- Полностью выполнены все этапы работы
- Глубокая интерпретация результатов с использованием SHAP
- Сформулированы практические рекомендации для бизнеса
- Критический анализ ограничений методов

6. Время выполнения: 3 академических часа

7. Форма контроля:

- Проверка программного кода
- Защита отчета с демонстрацией результатов
- Ответы на вопросы по теоретическому обоснованию

Матрица формирования компетенций в лабораторной работе

Компетенция	Этапы формирования	Оценочные материалы
ПК-2.1	Сравнительный анализ GLM и LightGBM, адаптация моделей под страховые данные	Код сравнения моделей, анализ метрик качества
ПК-2.2	Применение advanced методов (SHAP, cross-validation) для улучшения моделей	Отчет с обоснованием выбранных подходов
E1.1	Подготовка и обработка страховых данных, feature engineering	Код предобработки данных, создание новых признаков
E1.2	Построение моделей прогнозирования frequency и severity	Модели LightGBM, расчет прогнозов
E1.3	Расчет страховых премий, оценка рисков портфеля	Тарифные планы, анализ распределения рисков
SS-1.1	Анализ потенциального bias в признаках (демография)	SHAP-анализ, выявление дискриминационных факторов
SS-1.2	Предложения по mitigation bias в финальных моделях	Рекомендации в отчете, этичные feature selection

Формула итоговой оценки:

Итог = $0.3 \times (\text{средний балл по ПК-2}) + 0.4 \times (\text{средний балл по E1}) + 0.3 \times (\text{средний балл по SS-1})$

Критерии оценивания результатов обучения

Шкала перевода в традиционные оценки:

- **5 (Отлично):** 3.6 - 4.0 (преобладание высоких и продвинутых уровней)
- **4 (Хорошо):** 2.6 - 3.5 (преобладание продвинутых уровней)
- **3 (Удовлетворительно):** 2.0 - 2.5 (преобладание пороговых уровней)
- **2 (Неудовлетворительно):** < 2.0 (не достигнут пороговый уровень)

Экзаменационные материалы для промежуточной аттестации (экзамен).

Вопросы для подготовки к экзамену

1. Сравните парадигму классических актуарных моделей (на основе таблиц смертности, Chain-Ladder) и парадигму машинного обучения. Какие принципиальные различия в подходах к моделированию?
2. Объясните, как "большие данные" (Big Data) и новые источники информации (например, телеметрия, носимые устройства) меняют фундаментальные принципы актуарного ценообразования и риск-менеджмента.
3. В чем заключается проблема цензурирования (censoring) и усечения (truncation) данных? Приведите конкретные примеры из life и non-life страхования. Как эти проблемы влияют на выбор математического аппарата для моделирования?
4. Распределения страховых убытков часто обладают "тяжелыми хвостами" (heavy-tailed). Какие математические и вычислительные challenges это создает для стандартных моделей ML, и какие методы можно использовать для решения этих проблем?
5. Что такое "feature engineering" в контексте страховых тарифов? Приведите примеры 3-5 новых признаков, которые можно создать из исходных данных полиса, и обоснуйте их потенциальную прогностическую силу.
6. Опишите, как методы ML могут быть использованы для количественной оценки и снижения риска неблагоприятного отбора (adverse selection). Приведите пример модели или подхода.
7. GLM является бенчмарком в актуарной практике. Сформулируйте, в чем заключаются его ключевые преимущества с точки зрения интерпретируемости и регуляторных требований. Каковы его главные ограничения с точки зрения гибкости?
8. Математически обоснуйте выбор функции связи (link function) и распределения ошибок для моделирования: а) частоты убытков, б) severity убытков.
9. Объясните принцип работы градиентного бустинга (Gradient Boosting). В чем его фундаментальное преимущество перед бэггингом (например, Random Forest) при работе с табличными данными?
10. Как математически определяется и интерпретируется "важность признака" (feature importance) в градиентном бустинге? Какие есть ограничения у этой метрики?
11. Опишите архитектуру полносвязной нейронной сети для задачи тарификации. Какие функции активации вы бы рекомендовали для скрытых слоев и для выходного слоя в задачах регрессии и классификации и почему?
12. Сформулируйте проблему переобучения (overfitting) в контексте нейронных сетей. Какие методы регуляризации, кроме L1/L2, вы знаете и как они работают (например, Dropout)?
13. Что такое копула (copula) и как она используется для моделирования зависимости между случайными величинами? Почему использование копул в сочетании с ML-моделями для маргинальных распределений является мощным подходом в страховании?
14. В чем заключаются основные ограничения детерминистического метода Chain-Ladder? Какие предположения он закладывает и как ML-модели позволяют их нарушить?

15. Предложите архитектуру ML-модели (например, на основе градиентного бустинга) для прогнозирования индивидуальных убытков. Какие признаки вы бы использовали для каждого убытка?

16. Опишите, как рекуррентные нейронные сети (LSTM) могут быть применены к анализу триангуля убытков. В чем состоит преимущество учета временной динамики развития убытка?

17. В чем ключевое математическое отличие моделей survival analysis (например, модели Кокса) от стандартных моделей регрессии? Почему они необходимы для работы с цензурированными данными?

18. Дайте интерпретацию функции выживания (survival function) и hazard function в контексте страхования жизни. Как с их помощью можно рассчитать честную pure premium?

19. Какие этические вызовы и регуляторные ограничения возникают при использовании геномных данных или данных медицинских изображений в актуарном моделировании?

20. Опишите принцип работы Generative Adversarial Networks (GAN). Какую практическую проблему в актуарном моделировании (например, в стресс-тестировании) можно решить с их помощью?

21. В чем разница между GAN и Вариационным автоэнкодером (VAE) с точки зрения математической постановки задачи и качества генерируемых данных?

22. Что такое параметрическое страхование (parametric insurance)? Приведите пример, как данные с IoT-датчиков и простая ML-модель могут автоматизировать процесс урегулирования убытков.

23. Математически объясните, как работает метод SHAP (Shapley Additive exPlanations). Почему значения SHAP обладают свойством "additivity" и почему это важно для интерпретации?

24. Дайте математическое определение "справедливости" (fairness) ML-модели в страховом контексте. Какие метрики fairness вы знаете и как можно "исправить" модель, если она признана несправедливой?

25. Как использование сложных "black-box" ML-моделей влияет на процесс утверждения моделей в рамках Solvency II (требование "Use Test")? Какие дополнительные сложности возникают?

26. Объясните, как прогнозы ML-моделей для убытков должны быть интегрированы в расчеты по IFRS 17 (например, в оценку fulfillment cash flows).

27. Что такое MLOps и почему его внедрение критически важно для промышленного использования AI в страховой компании? Назовите ключевые этапы жизненного цикла ML-модели.

28. Сравнительный анализ: Вам нужно разработать модель для тарификации нового продукта – киберстрахования для малого бизнеса. Опишите пошаговый план действий: от сбора данных до выбора финальной модели, аргументируя каждый шаг. Какие модели вы рассмотрите и почему?

29. Этика и будущее: "Страхование с помощью AI движется к модели динамического, персонализированного ценообразования в реальном времени". Согласны ли вы с этим утверждением? Какие технологические, этические и регуляторные барьеры необходимо преодолеть для реализации этого сценария?

30. Критика модели: Вам представили ML-модель, которая показывает на 15% лучшую точность, чем GLM-бенчмарк, но является сложной для интерпретации. По каким критериям, кроме точности, вы будете оценивать эту модель перед ее внедрением в производство? Составьте checklist из 5-7 ключевых пунктов.

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

Методические рекомендации, определяющие процедуры оценивания лабораторной работы.

Текущая аттестация проводится по лабораторным работам, и может принести в копилку максимум **40 баллов**. В соответствие с критериями оценки выполнения лабораторных работ Пороговый уровень (удовлетворительно) 10 баллов, (хорошо) базовый 20 баллов, продвинутый уровень (отлично) 40 баллов. Текущий бал определяется усреднением баллов по всем (16) лабораторным работам и как результат будет принадлежать отрезку от 0 до 40 баллов.

Промежуточной аттестацией по дисциплине является экзамен. Максимальная оценка, которую можно получить в качестве оценки экзамена 60 баллов: Пороговый уровень 40 баллов, (хорошо) базовый 50 баллов, продвинутый уровень (отлично) 60 баллов.

В стандартной форме экзаменационная оценка определяется следующим соответствием суммарных баллов и оценки:

- 0 – 49 баллов «неудовлетворительно»;
- 50 – 70 баллов «удовлетворительно»;
- 71 – 85 баллов «хорошо»;
- 86 – 100 баллов «отлично».

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

4.3 Методические указания по организации лабораторных работ по дисциплине "Искусственный интеллект в оценке рисков и разработке страховых продуктов"

1. Общие положения

Лабораторные работы предназначены для формирования практических навыков применения методов искусственного интеллекта и машинного обучения в решении

прикладных задач страховой отрасли. Работы построены по принципу "от простого к сложному" и охватывают полный цикл разработки актуарных ML-моделей.

2. Организационная структура

2.1. Временной регламент:

- Общий объем: 34 академических часа
- Формат занятий: 10 лабораторных работ по 3 часа + 4 часа на установочное занятие и защиту проектов
- Режим работы: парные занятия с перерывом 10 минут

2.2. Оборудование и ПО:

- Техническое обеспечение: ПК с доступом в интернет, минимум 8 ГБ ОЗУ
- Программное обеспечение:
 - o Python 3.8+ с библиотеками: pandas, numpy, matplotlib, seaborn
 - o ML-библиотеки: scikit-learn, lightgbm/xgboost, torch/tensorflow
 - o Специализированные библиотеки: lifelines, shap, scipy
 - o Среда разработки: Jupyter Notebook/Lab или VS Code

3. Методика проведения работ

3.1. Подготовительный этап:

- За 1 неделю: выдача задания и теоретических материалов
- Самостоятельное изучение: математического аппарата методов
- Установка и тестирование ПО (при необходимости)

3.2. Структура занятия:

- 0-20 мин: вводный инструктаж, постановка задачи
- 20-120 мин: самостоятельная работа студентов
- 120-140 мин: консультации, разбор типовых ошибок
- 140-170 мин: промежуточный контроль, обсуждение результатов

3.3. Роль преподавателя:

- Консультант и модератор процесса
- Эксперт по предметной области
- Проверяющий и оценивающий

4. Содержание лабораторного практикума

Блок 1. Фундаментальные навыки (Работы 1-3)

- Работа со страховыми данными, включая цензурированные
- Построение и интерпретация классических моделей (GLM)
- Освоение базовых ML-алгоритмов (деревья, случайный лес)

Блок 2. Продвинутое методы (Работы 4-6)

- Градиентный бустинг для табличных данных
- Нейронные сети для страховых задач
- Survival analysis для работы с цензурированными данными

Блок 3. Специализированные приложения (Работы 7-8)

- ML в резервировании убытков
- Интерпретация моделей и анализ справедливости

Блок 4. Интеграция и применение (Работы 9-10)

- Сквозной проект по разработке страхового продукта
- Защита и презентация результатов

5. Система оценивания

5.1. Критерии оценки каждой работы:

- Качество кода и воспроизводимость результатов (25%)
- Глубина анализа и интерпретации (35%)
- Соответствие выводов поставленным задачам (25%)
- Оформление отчета и визуализация (15%)

5.2. Накопительная система:

- Лабораторные работы 1-8: 50% итоговой оценки

- Интеграционный проект (работы 9-10): 40%
 - Активность и участие в обсуждениях: 10%
- 6. Методические рекомендации**
- 6.1. Для студентов:**
- Вести конспект по каждой работе с фиксацией ключевых insights
 - Использовать систему контроля версий (Git) для кода
 - Формировать портфолио из лучших реализованных проектов
- 6.2. Для преподавателей:**
- Дифференцировать сложность заданий по уровню подготовки студентов
 - Использовать реальные страховые кейсы и данные (где возможно)
 - Организовывать peer-review для взаимного обучения
- 7. Требования к отчетности**
- 7.1. Формат отчета:**
- Jupyter Notebook с полным описанием хода работы
 - Текстовый документ с выводами и рекомендациями
 - Презентация для защиты интеграционного проекта
- 7.2. Содержание отчета:**
- Постановка задачи и описание данных
 - Предобработка и feature engineering
 - Построение и валидация моделей
 - Анализ результатов и бизнес-интерпретация
 - Выводы и ограничения подхода
- 8. Особенности организации**
- 8.1. Дистанционный формат:**
- Использование облачных сред (Google Colab, Kaggle Notebooks)
 - Проведение консультаций через видеоконференции
 - Организация совместной работы с помощью Git
- 8.2. Индивидуальные траектории:**
- Дополнительные задания для продвинутых студентов
 - Упрощенные варианты для студентов с трудностями
 - Возможность выбора темы интеграционного проекта
- 9. Ресурсное обеспечение**
- 9.1. Учебные материалы:**
- Шаблоны кода для каждой лабораторной работы
 - Видео-инструкции по сложным темам
 - База данных примеров страховых кейсов
- 9.2. Данные для работы:**
- Анонимизированные реальные страховые данные
 - Сгенерированные датасеты с известными свойствами
 - Публичные страховые датасеты (Kaggle, UCI)
- 10. Контроль качества**
- 10.1. Промежуточный контроль:**
- Еженедельная проверка выполнения работ
 - Взаимопроверка студентами (peer assessment)
 - Автоматизированное тестирование кода
- 10.2. Итоговый контроль:**
- Защита интеграционного проекта
 - Собеседование по теоретическим основам
 - Оценка практических навыков на тестовом задании

Чек-лист для проверки выполнения лабораторных работ

Критерий 1: корректная настройка инфраструктуры

№	Задача	Индикатор выполнения	Статус (+/-)
1.1	Установка и настройка Python-окружения	Установлены Python 3.8+, все требуемые библиотеки (pandas, numpy, scikit-learn, lightgbm, shap), окружение воспроизводимо (requirements.txt)	
1.2	Настройка Jupyter Notebook/Lab	Ноутбук запускается без ошибок, все ячейки выполняются последовательно	
1.3	Организация рабочего пространства	Проектная структура включает папки: /data, /models, /notebooks, /reports	
1.4	Настройка системы контроля версий	Git репозиторий инициализирован, commits отражают этапы работы	

Критерий 2: умение применять spark, nosql, kafka

№	Задача	Индикатор выполнения	Статус (+/-)
2.1	Работа с Spark для больших данных	Создан SparkSession, данные загружены в DataFrame, выполнены операции фильтрации и агрегации	
2.2	Обработка данных в распределенной среде	Применены transformations (map, filter, reduceByKey) и actions (collect, count)	
2.3	Интеграция с NoSQL	Данные сохранены/загружены из MongoDB/Cassandra, выполнены базовые запросы	
2.4	Работа с Kafka для потоковых данных	Настроен producer/consumer, реализована отправка/получение сообщений	

Критерий 3: построение ml-моделей и дашбордов

№	Задача	Индикатор выполнения	Статус (+/-)
3.1	Предобработка данных	Выполнены: обработка пропусков, кодирование категориальных переменных, масштабирование	
3.2	Feature Engineering	Созданы новые признаки, отбор фичей выполнен (feature importance, корреляции)	
3.3	Построение базовых моделей	Реализованы и обучены: Linear Regression, Random Forest, Gradient Boosting	
3.4	Валидация и оценка моделей	Рассчитаны метрики (MAE, RMSE, R ² , ROC-AUC), выполнена кросс-валидация	
3.5	Построение дашбордов	Создан интерактивный дашборд (Plotly Dash/Streamlit) с ключевыми метриками	
3.6	Визуализация результатов	Построены графики: распределения, важности признаков, прогнозы vs факт	

Критерий 4: оптимизация и масштабирование решений

№	Задача	Индикатор выполнения	Статус (+/-)
4.1	Оптимизация гиперпараметров	Выполнен поиск по сетке (GridSearch) или случайный поиск (RandomizedSearch)	
4.2	Ускорение вычислений	Использованы техники: распараллеливание, оптимизация запросов, кэширование	
4.3	Масштабирование решения	Код адаптирован для работы с увеличенными объемами данных	
4.4	Деплой решения	Модель сохранена (pickle/joblib), создан API для предсказаний (Flask/FastAPI)	

Дополнительные критерии качества

№	Критерий	Индикатор выполнения	Статус (+/-)
5.1	Качество кода	Код читаем, содержит комментарии, следует PEP8	
5.2	Документация	Написаны docstrings, README с инструкцией по запуску	
5.3	Воспроизводимость	Результаты идентичны при повторном запуске	
5.4	Анализ ошибок	Проведен анализ неверных предсказаний, выявлены закономерности	

Вывод

Соответствие лабораторных работ индикаторам компетенций подтверждает, что курс покрывает все ключевые аспекты дисциплины.

5. Перечень учебной литературы, информационных ресурсов и технологий

5.1. Учебная литература

1. Новиков, Ф. А. Символический **искусственный интеллект**: математические основы представления знаний : учебник для вузов / Ф. А. Новиков. - Москва : Юрайт, 2025. - 278 с. - URL: <https://www.ura.it.ru/bcode/561410> (дата обращения: 19.09.2025). - Режим доступа: для авториз. пользователей. - ISBN 978-5-534-00734-3 http://megapro.kubsu.ru/MegaPro/UserEntry?Action=Link_FindDoc&id=147743&idb=0

2. Постолиит, А. В. Основы искусственного интеллекта в примерах на Python : самоучитель / Анатолий Постолиит. - 2-е изд., перераб. и доп. - Санкт-Петербург : БХВ-Петербург, 2024. - 446 с. : ил. - (Самоучитель). - Библиогр.: с. 440-443. - ISBN 978-5-9775-1818-5 http://megapro.kubsu.ru/MegaPro/UserEntry?Action=Link_FindDoc&id=276189&idb=0

3. Загорулько, Ю. А. **Искусственный интеллект**. Инженерия знаний : учебное пособие для вузов / Ю. А. Загорулько, Г. Б. Загорулько. - Москва : Юрайт, 2024. - 93 с. - URL: <https://ura.it.ru/bcode/540987> (дата обращения: 07.03.2025). - Режим доступа: для авториз. пользователей. - ISBN 978-5-534-07198-6 http://megapro.kubsu.ru/MegaPro/UserEntry?Action=Link_FindDoc&id=281328&idb=0

4. Жданов, А. А. Автономный **искусственный интеллект** : учебное пособие / А. А. Жданов. - 5-е изд. - Москва : Лаборатория знаний, 2024. - 362 с. - URL: <https://e.lanbook.com/book/387629> (дата обращения: 19.02.2024). - Режим доступа: для авториз. пользователей. - ISBN 978-5-93208-674-2 http://megapro.kubsu.ru/MegaPro/UserEntry?Action=Link_FindDoc&id=144012&idb=0

5.2. Периодическая литература

1. Sun, X., Li, J., Kovalenko, A.V., Feng, W., Ou, Y. Integrating Reinforcement Learning and Learning From Demonstrations to Learn Nonprehensile Manipulation //IEEE Transactions on Automation Science and Engineering, 2023, 20(3), 1735–1744, DOI: 10.1109/TASE.2022.3185071, Q1

2. Petukhova, A.V.; Kovalenko, A.V.; Ovsyannikova, A.V. Algorithm for Optimization of Inverse Problem Modeling in Fuzzy Cognitive Maps. Mathematics 2022, 10, 3452. DOI: 10.3390/math10193452, Q1

3. Kirillova, E.; Kovalenko, A.; Urtenov, M. Study of the Current–Voltage Characteristics of Membrane Systems Using Neural Networks. AppliedMath 2025, 5, 10. <https://doi.org/10.3390/appliedmath5010010>

4. Kadurin, Artur, et al. "The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology." Oncotarget 8.7 (2016): 10883.

5. Kadurin, Artur, et al. "druGAN: an advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico." *Molecular pharmaceutics* 14.9 (2017): 3098-3104.
6. Polykovskiy, Daniil, et al. "Molecular sets (MOSES): a benchmarking platform for molecular generation models." *Frontiers in pharmacology* 11 (2020): 565644.
7. Khrabrov, Kuzma, et al. " ∇^2 DFT: A Universal Quantum Chemistry Dataset of Drug-Like Molecules and a Benchmark for Neural Network Potentials." *Advances in Neural Information Processing Systems* 37 (2024): 36869-36889.
8. Polykovskiy, Daniil, et al. "Entangled conditional adversarial autoencoder for de novo drug discovery." *Molecular pharmaceutics* 15.10 (2018): 4398-4405.
9. Николенко, Сергей, Кадурин, Артур и Архангельская Екатерина. *Глубокое обучение. Издательский дом "Питер"*, 2017.
10. Фёдоров, Е. А., Булыгин, А. В. «Искусственный интеллект в страховании: от теории к практике» // *Вопросы экономики и управления*, 2023, № 2.
11. Сабитов К.Т., Исаев Д.В. ВОЗМОЖНОСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СТРАХОВАНИИ// Научный сетевой журнал «Столыпинский вестник» №9/2024. <https://cyberleninka.ru/article/n/vozmozhnosti-iskusstvennogo-intellekta-v-strahovanii?ysclid=mh5wh3euk7572963311>
12. Ассоциация страховых организаций России (АСО) — *Отчёты по цифровизации страхового рынка РФ* (ежегодно, 2020–2024).
13. Базы данных компании «Ист Вью» <http://dlib.eastview.com>
14. Электронная библиотека GREBENNIKON.RU <https://grebennikon.ru/>

5.3. Интернет-ресурсы, в том числе современные профессиональные базы данных и информационные справочные системы
Электронно-библиотечные системы (ЭБС):

1. ЭБС «ЮРАЙТ» <https://urait.ru/>
2. ЭБС «УНИВЕРСИТЕТСКАЯ БИБЛИОТЕКА ОНЛАЙН» www.biblioclub.ru
3. ЭБС «BOOK.ru» <https://www.book.ru>
4. ЭБС «ZNANIUM.COM» www.znanium.com
5. ЭБС «ЛАНЬ» <https://e.lanbook.com>

Базы данных и аналитические платформы

1. Google BigQuery – облачная аналитика больших данных.
2. Apache Hadoop & Spark – официальная документация и ресурсы.
3. Kaggle – датасеты, соревнования и учебные материалы.
4. Cloudera – платформа для работы с Big Data.
5. Databricks – решения на основе Apache Spark.

Справочные системы и блоги

1. Towards Data Science (Medium) – статьи по Data Science и Big Data.
2. KDnuggets – новости, обучающие материалы и обзоры инструментов.
3. O'Reilly Data & AI – книги и статьи по Big Data и машинному обучению.
4. IBM Big Data Hub – кейсы и руководства по Big Data.

Ресурсы свободного доступа

1. [Apache Spark Documentation](http://www.apache.org/docs/spark/)
2. [Yandex Cloud Big Data](https://cloud.yandex.ru/docs/big-data/)
3. [Kaggle Datasets](https://www.kaggle.com/datasets)

Профессиональные базы данных:

1. Web of Science (WoS) <http://webofscience.com/>
2. Scopus <http://www.scopus.com/>

3. ScienceDirect www.sciencedirect.com
4. Журналы издательства Wiley <https://onlinelibrary.wiley.com/>
5. Научная электронная библиотека (НЭБ) <http://www.elibrary.ru/>
6. Полнотекстовые архивы ведущих западных научных журналов на Российской платформе научных журналов НЭИКОН <http://archive.neicon.ru>
7. Национальная электронная библиотека (доступ к Электронной библиотеке диссертаций Российской государственной библиотеки (РГБ) <https://rusneb.ru/>
8. Президентская библиотека им. Б.Н. Ельцина <https://www.prilib.ru/>
9. Электронная коллекция Оксфордского Российского Фонда <https://ebookcentral.proquest.com/lib/kubanstate/home.action>
10. Springer Journals <https://link.springer.com/>
11. Nature Journals <https://www.nature.com/siteindex/index.html>
12. Springer Nature Protocols and Methods <https://experiments.springernature.com/sources/springer-protocols>
13. Springer Materials <http://materials.springer.com/>
14. zbMath <https://zbmath.org/>
15. Nano Database <https://nano.nature.com/>
16. Springer eBooks: <https://link.springer.com/>
17. "Лекториум ТВ" <http://www.lektorium.tv/>
18. Университетская информационная система РОССИЯ <http://uisrussia.msu.ru>

Информационные справочные системы:

1. Консультант Плюс - справочная правовая система (доступ по локальной сети с компьютеров библиотеки)

Ресурсы свободного доступа:

1. Американская патентная база данных <http://www.uspto.gov/patft/>
2. Полные тексты канадских диссертаций <http://www.nlc-bnc.ca/thesescanada/>
3. КиберЛенинка (<http://cyberleninka.ru/>);
4. Министерство науки и высшего образования Российской Федерации <https://www.minobrnauki.gov.ru/>;
5. Федеральный портал "Российское образование" <http://www.edu.ru/>;
6. Информационная система "Единое окно доступа к образовательным ресурсам" <http://window.edu.ru/>;
7. Единая коллекция цифровых образовательных ресурсов <http://school-collection.edu.ru/> .
8. Федеральный центр информационно-образовательных ресурсов (<http://fcior.edu.ru/>);
9. Проект Государственного института русского языка имени А.С. Пушкина "Образование на русском" <https://pushkininstitute.ru/>;
10. Справочно-информационный портал "Русский язык" <http://gramota.ru/>;
11. Служба тематических толковых словарей <http://www.glossary.ru/>;
12. Словари и энциклопедии <http://dic.academic.ru/>;
13. Образовательный портал "Учеба" <http://www.ucheba.com/>;
14. Законопроект "Об образовании в Российской Федерации". Вопросы и ответы http://xn--273--84d1f.xn--plai/voprosy_i_otvety

Собственные электронные образовательные и информационные ресурсы КубГУ:

1. Среда модульного динамического обучения <http://moodle.kubsu.ru>
2. База учебных планов, учебно-методических комплексов, публикаций и конференций <http://mschool.kubsu.ru/>
3. Библиотека информационных ресурсов кафедры информационных образовательных технологий <http://mschool.kubsu.ru;>
4. Электронный архив документов КубГУ <http://docspace.kubsu.ru/>

5. Электронные образовательные ресурсы кафедры информационных систем и технологий в образовании КубГУ и научно-методического журнала "ШКОЛЬНЫЕ ГОДЫ"
<http://icdau.kubsu.ru/>

6. Методические указания для обучающихся по освоению дисциплины (модуля)

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

По курсу предусмотрено проведение лекционных занятий, на которых дается систематизированный материал по технологиям обработки больших данных. В ходе лекций рассматриваются ключевые концепции.

Лабораторные занятия курса посвящены практическому освоению технологиям обработки больших данных

При самостоятельной работе студентам необходимо изучать рекомендованную литературу в виде официальной документации к используемым открытым программным продуктам, облачным платформам.

Важнейшим компонентом курса является самостоятельная проектная работа, в ходе которой студент разрабатывает законченное решение для решения задач (кейсов) промышленных партнеров. Допускается выполнение проектов в командах.

Подход, определяющий установление соответствия кейсов ИП и УГТ (5-7), позволяет четко соотносить этапы развития технологии с вовлеченностью партнера и снижать риски при переходе от лабораторных испытаний к промышленному внедрению.

Ключевые аспекты взаимодействия с промышленными партнерами:

- Для УГТ 5 – ИП помогает определить реалистичные условия тестирования, но не рискует своей инфраструктурой.
- Для УГТ 6 – ИП предоставляет "песочницу" или изолированную среду, где можно выявить скрытые проблемы.
- Для УГТ 7 – ИП становится соразработчиком, так как технология адаптируется под его конкретные процессы.

А. ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОЦЕНКЕ РИСКОВ И РАЗРАБОТКЕ СТРАХОВЫХ ПРОДУКТОВ В КЕЙСАХ СТРАХОВЫХ СТРУКТУР ПАО «Сбер-банк»

1. ИНТЕЛЛЕКТУАЛЬНАЯ ОЦЕНКА РИСКОВ В РЕАЛЬНОМ ВРЕМЕНИ

Кейс 1: Динамическое ценообразование в автостраховании (Telematics)

• Технологии:

- Gradient Boosting + LSTM-сети
- Обработка потоковых данных с датчиков
- Компьютерное зрение для анализа стиля вождения

• Реализация:

```
# Псевдокод системы оценки риска
risk_score = telematics_model.predict(
    speeding_events,
    braking_intensity,
```

road_type,
time_of_day

)

premium = base_rate * risk_score * customer_segment_factor

Эффект: Снижение убыточности на 15-25% за счет персонализированных тарифов

Кейс 2: Прогнозирование убытков в ипотечном страховании

- **Данные:**
 - Кредитная история + макроэкономические показатели
 - Данные о недвижимости из открытых источников
 - Поведенческие паттерны заемщиков
- **ML-модели:**
 - Survival Analysis для оценки вероятности дефолта
 - Graph Neural Networks для анализа связанных рисков
- **Результат:** Точность прогноза дефолтов на 18% выше традиционных

моделей

2. AI-ДЛЯ РАЗРАБОТКИ СТРАХОВЫХ ПРОДУКТОВ

Кейс 3: Генерация персонифицированных страховых покрытий

- **Технологии:**
 - Generative Adversarial Networks (GAN) для создания продуктовых комбинаций
 - Reinforcement Learning для оптимизации условий страхования
- **Процесс:**
 1. Анализ профиля клиента и его потребностей
 2. Генерация оптимального пакета покрытий
 3. A/B тестирование принятия продукта
- **Пример:**
 - "Страхование путешествий + отмена мероприятий + киберриски"
 - Динамические франшизы на основе поведения

комбинаций

Кейс 4: Параметрическое страхование для SME

- **Решение:**
 - IoT-сенсоры в торговых точках
 - ML-модели для определения страховых случаев
 - Смарт-контракты для автоматических выплат
- **Данные:**
 - Поток клиентов, выручка, погодные условия
 - Социальные сигналы (отзывы, активность)
- **Эффект:** Снижение операционных затрат на 40%

3. ИНТЕГРИРОВАННАЯ AI-ПЛАТФОРМА СТРАХОВОГО БАНКА

Архитектура решения:

Data Sources → Feature Store → ML Models → API Gateway → Business Applications

↓

Банковские
транзакции

↓

Real-time
Features

↓

Risk Scoring
Models

↓

Продуктовые
конструкторы

↓

Дашборды
мониторинга

Ключевые компоненты:

1. **Единый Feature Store**
 - 500+ признаков из банковских и страховых данных
 - Online/offline feature serving
 - Автоматическое обновление фич
2. **MLOps-инфраструктура**
 - Автоматическое переобучение моделей
 - A/B тестирование в production

- Мониторинг дрейфа данных
- 3. **Объясняемые AI (XAI)**
 - SHAP/LIME для регуляторной отчетности
 - Автоматические отчеты для ЦБ РФ
 - Интерпретация моделей для андеррайтеров

4. КОНКРЕТНЫЕ РЕЗУЛЬТАТЫ ВНЕДРЕНИЯ

Направление	Показатель до AI	Показатель после AI	Эффект
Тарификация	Loss Ratio: 78%	Loss Ratio: 62%	↓16% убыточности
Андеррайтинг	Время оценки: 2 дня	Время оценки: 15 мин	↑95% скорость
Фрод-детекция	Обнаружение: 45%	Обнаружение: 89%	↑44% точности
Кросс-селл	Конверсия: 3.2%	Конверсия: 8.7%	↑5.5 п.п.

5. ВЫЗОВЫ И ОГРАНИЧЕНИЯ

Регуляторные требования:

- Объяснимость моделей для ЦБ РФ
- Соответствие 152-ФЗ о персональных данных
- Этическое использование AI (EU AI Act)

Технические сложности:

- Интеграция legacy-систем банка и страховой компании
- Качество и доступность данных
- MLOps в регулируемой отрасли

Организационные барьеры:

- Культурные различия IT-команд банка и страховщика
- Необходимость переобучения андеррайтеров
- Управление change resistance

6. ДОРОЖНАЯ КАРТА ВНЕДРЕНИЯ

Этап 1 (6 месяцев):

- Создание единого Data Lake
- Пилот в автостраховании
- Базовая MLOps-инфраструктура

Этап 2 (12 месяцев):

- AI-андеррайтинг во всех массовых линиях
- Интеграция с банковскими процессами
- Automated claims handling

Этап 3 (18 месяцев):

- Персональные страховые ассистенты
- Predictive capital management
- Полный AI-цикл страхования

Кейс 5: Модель анализа инвестиционной привлекательности малого бизнеса

Описание:

Банк активно развивает кредитование и инвестиционные инструменты для малого и среднего предпринимательства (МСП). Требуется создать модель, которая на основе открытых и банковских данных (выручка, расходы, тип деятельности, отзывы, онлайн-активность) оценивает инвестиционную привлекательность МСП.

Цель:

Разработать систему рейтинговой оценки компаний малого бизнеса с возможностью визуализации факторов и динамики показателей.

Ожидаемый результат:

Модель, присваивающая компании инвестиционный рейтинг (например, А–Е), объясняющая ключевые параметры и дающая рекомендации для инвестора.

Кейс 6: Индивидуальная оценка кредитоспособности клиента на основе поведенческих данных

Описание:

Современный кредитный скоринг выходит за рамки финансовых данных. Необходимо исследовать, как поведенческие и цифровые следы (частота входа в мобильный банк, способы оплаты, география, время отклика) влияют на персональную оценку риска.

Цель:

Разработать ML-модель, оценивающую вероятность дефолта по нестандартным поведенческим признакам (возможно — с explainable AI).

Ожидаемый результат:

Прототип скоринговой модели, которая, помимо стандартных данных, учитывает цифровой профиль клиента и объясняет решения (SHAP, LIME и др.).

Заключение: Внедрение AI в страховые структуры банка позволяет создать уникальное конкурентное преимущество через синергию банковских и страховых данных, сокращение издержек и создание принципиально новых страховых продуктов.

Б. ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОЦЕНКЕ РИСКОВ И РАЗРАБОТКЕ СТРАХОВЫХ ПРОДУКТОВ В КЕЙСАХ КОМПАНИИ AVA GROUP

1. ИИ ДЛЯ УПРАВЛЕНИЯ СТРОИТЕЛЬНЫМИ РИСКАМИ

Кейс 1: Прогнозирование сроков и бюджета строительства

Псевдокод модели прогнозирования задержек

```
delay_risk = model.predict(
    weather_data,
    supply_chain_delays,
    workforce_availability,
    equipment_failure_history
)
```

- **Данные:** Погодные условия, логистика поставок, доступность рабочих, геодезические данные
- **Эффект:** Снижение рисков срыва сроков на 25-30%

Кейс 2: Мониторинг безопасности на стройплощадках

- **Компьютерное зрение:** Анализ видео с камер наблюдения
- **Detection объектов:** Каски, страховочные тросы, опасные зоны
- **Результат:** Снижение страховых случаев на 40%

2. ИИ ДЛЯ ОЦЕНКИ РИСКОВ ДЕВЕЛОПЕРСКИХ ПРОЕКТОВ

Кейс 3: Прогнозирование рыночного спроса

- **ML-модели:** Анализ макроэкономических показателей, демографии, инфраструктурного развития
- **Цель:** Оптимизация портфеля проектов (жилье/офисы/ТРЦ)
- **Данные:** 500+ параметров по каждому объекту

- Динамические тарифы
- Health & Wellness в отелях:**
- Персонализированные страховые пакеты
- Мониторинг состояния гостей
- Проактивные рекомендации
- Цифровые двойники объектов:**
- Визуализация рисков в реальном времени
- Симуляция чрезвычайных ситуаций
- Оптимизация страхового покрытия

8. ДОРОЖНАЯ КАРТА ВНЕДРЕНИЯ

Фаза 1 (0-8 месяцев):

- Внедрение компьютерного зрения на стройплощадках
- Запуск системы прогнозирования логистических рисков
- Пилот в 2-х управляемых объектах

Фаза 2 (8-18 месяцев):

- Интеграция данных по всем направлениям
- Разработка специализированных страховых продуктов
- Внедрение предиктивного maintenance

Фаза 3 (18-36 месяцев):

- Создание единой риск-платформы холдинга
- Запуск коммерческих страховых продуктов для рынка
- Развитие экосистемы "умного" страхования

Итог: Для AVA Group внедрение ИИ в управление рисками позволяет не только снизить операционные потери, но и создать новые страховые продукты, которые могут стать отдельным прибыльным направлением бизнеса. Синергия между различными подразделениями холдинга создает уникальное конкурентное преимущество в области риск-менеджмента.

7. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю)

7.1 Перечень информационно-коммуникационных технологий

Системы управления обучением (LMS):

- **Moodle** – размещение лекций, лабораторных работ, тестирование
- **Google Classroom** – организация групповых проектов, сбор заданий
- **Microsoft Teams for Education** – проведение вебинаров, консультаций

Платформы для онлайн-обучения:

- **Zoom Pro** – интерактивные лекции и мастер-классы
- **Cisco Webex** – видеоконференции с экспертами страховой отрасли
- **Discord** – оперативное общение и техническая поддержка

Облачные сервисы:

- **Google Workspace for Education** – совместная работа над проектами
- **Microsoft 365 A1** – облачное хранение и редактирование материалов
- **Yandex Cloud Educational** – вычислительные ресурсы для ML-моделей

7.2 Перечень лицензионного и свободно распространяемого программного обеспечения

Лицензионное ПО:

Профессиональные среды разработки:

- **PyCharm Professional** – разработка Python-приложений
- **MATLAB Campus Wide License** – математическое моделирование рисков
- **Tableau Academic Program** – визуализация страховых данных

Специализированное ПО:

- **IBM SPSS Statistics** – статистический анализ страховых случаев
- **SAS Analytics Pro** – актуарные расчеты и прогнозирование
- **Oracle Database Enterprise Edition** – управление базами данных

ML-платформы:

- **DataRobot Academic Edition** – автоматизированное машинное обучение
- **H2O.ai Academic License** – построение прогнозных моделей
- **RStudio Pro** – анализ данных и актуарное моделирование

Свободно распространяемое (open-source) ПО:

Языки программирования и библиотеки:

Python 3.8+ с библиотеками:

- o pandas, numpy – обработка данных
- o scikit-learn – машинное обучение
- o lightgbm, xgboost – градиентный бустинг
- o torch, tensorflow – нейронные сети
- o shap, lime – интерпретация моделей
- o lifelines – survival analysis

R с пакетами actuar, insuranceData

Среды разработки:

Jupyter Notebook/Lab – интерактивные вычисления

Visual Studio Code – разработка с поддержкой Python

Apache Spark – распределенная обработка данных

Специализированные инструменты:

Apache Kafka – потоковая обработка данных

MLflow – управление жизненным циклом ML-моделей

Docker – контейнеризация приложений

Образовательные и демонстрационные ресурсы

Датасеты для практических работ:

Kaggle Datasets (Porto Seguro Safe Driver, Health Insurance)

UCI Machine Learning Repository

Synthetic insurance datasets

Онлайн-платформы:

Kaggle Learn – интерактивные курсы по ML

Coursera – специализации по страхованию и AI

Stepik – онлайн-курсы по Python и анализу данных

Технические требования

Аппаратное обеспечение:

- Персональные компьютеры с минимум 8 ГБ ОЗУ
- Видеокарты NVIDIA (для работы с нейросетями)
- Стабильный интернет-канал

Облачная инфраструктура:

- **Google Colab Pro** – доступ к GPU для обучения моделей
- **Yandex DataSphere** – облачная среда для ML-разработки
- **Kaggle Notebooks** – бесплатные вычислительные ресурсы

Специализированные страховые платформы

Демо-версии профессионального ПО:

- **SAS Insurance** – отраслевое решение для страховщиков
- **IBM Watson** – когнитивные технологии для страхования

- **Guidewire** – платформа для страховых продуктов
- Симуляторы страховой деятельности:**
- **AnyLogic** – имитационное моделирование страховых процессов
- **@Risk** – анализ рисков и неопределенностей

8. Материально-техническое обеспечение по дисциплине (модулю)

Виртуальные машины, кластер Managed Kubernetes и ресурсы GPU в облаке предоставляется индустриальным партнером ПАО «Сбербанк»:

№	Продукт	Параметры продукта	Кол-во	Кол-во конфигураций	Ед. изм.
1	2	3	4	5	6
1	Виртуальная машина	Виртуальная машина 10% vCPU 2 vCPU 4 RAM	1	60	Шт
		ОС Ubuntu 22.04	1		Шт
		Системный диск SSD	1		Шт
			10		Гб
		Аренда публичного IP	1		Шт
2	Виртуальная машина с GPU	Виртуальная машина с GPU NVIDIA® Tesla® V100 2 GPU 8 vCPU 128 Гб RAM	1	1	Шт
		ОС Ubuntu 24.04	1		Шт
		Системный диск SSD	1		Шт
			2000		Гб
		Диск SSD	1		Шт
			4096		Гб
		Диск SSD	1		Шт
			4096		Гб
		Аренда публичного IP	1		Шт
3	K8S	Master node 8 vCPU 16 RAM	1	1	Шт
		Worker node 10% доля 4 vCPU 32 RAM	5		Шт
		Worker node SSD-NVME	64		Гб
		Аренда публичного IP	1		Шт
4	ML Inference Instance Type GPU	Время работы в месяц	40	1	ч
		Инстанс 8 x NVIDIA® H100 NVLink PCIe 160 vCPU 1520 GB RAM	1		Шт
		Количество запросов к ML-моделям	1		Млн. Шт
		Кэш ML-моделей	160		Гб
5	LLM	Токены GigaChat 2 Max	50		Млн. Шт
		Токены Embeddings	400		Млн. Шт

Дополнительные облачные ресурсы предоставляются технологическим партнером Yandex Cloud.

№	Вид работ	Наименование учебной аудитории, ее оснащенность оборудованием и техническими средствами обучения
1.	Лекционные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
2.	Лабораторные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, проектором, программным обеспечением

3.	Практические занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
4.	Групповые (индивидуальные) консультации	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
5.	Текущий контроль, промежуточная аттестация	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
6.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.

Примечание: Конкретизация аудиторий и их оснащение определяется ОПОП.