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.
подпись
« 29 » августа 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1. В.02 Математические модели защиты информации

Направление подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем

Профиль Искусственный интеллект и аналитика данных

Форма обучения очная

Квалификация бакалавр

Краснодар 2025

Рабочая программа дисциплины «Математические модели защиты информации» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем.

Программу составил(и):
В.О. Осипян, профессор кафедры, д-р.ф.-м.н., доцент
И.О. Фамилия, должность, ученая степень, ученое звание


подпись

Рабочая программа дисциплины «Математические модели защиты информации» утверждена на заседании кафедры анализа данных и искусственного интеллекта
протокол № 01 «28» августа 2025 г.
Заведующий кафедрой Коваленко А.В.


подпись

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


подпись

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

Целью освоения дисциплины «Математические модели защиты информации» является формирование у будущих специалистов в области прикладной информатики и искусственного интеллекта систематизированных знаний о математическом аппарате, лежащем в основе современных методов защиты информации, и выработка навыков его применения для обеспечения конфиденциальности, целостности и доступности данных и моделей на всех этапах жизненного цикла AI-систем.

1.2 Задачи дисциплины

1. Сформировать у студентов понимание фундаментальных математических моделей, используемых в криптографии.
2. Научить формализовывать угрозы информационной безопасности для данных и алгоритмов в системах аналитики и ИИ, используя математический аппарат.
3. Дать представление о принципах работы и области применения основных криптографических алгоритмов (симметричных, асимметричных, хеш-функций) с точки зрения лежащих в их основе сложных математических задач.
4. Освоить на практическом уровне методы реализации базовых криптографических примитивов и протоколов.
5. Показать взаимосвязь между защитой информации и технологиями ИИ, в частности, рассмотреть математические основы методов обеспечения конфиденциальности при обучении моделей (дифференциальная приватность) и противодействия атакам на ML-системы.

1.3 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Математические модели защиты информации» относится к «Части, формируемая участниками образовательных отношений» Блока 1 «Дисциплины (модули)» учебного плана.

Данная дисциплина тесно связана с дисциплинами: «Математический анализ», «Основы программирования», «Дифференциальные уравнения», «Математические модели нейронных сетей».

Материал курса является связующим звеном между математикой, программированием и прикладными задачами обеспечения безопасности данных. Знания, полученные в данной дисциплине, являются основой для преддипломной практики и ВКР.

1.4 Профессиональные роли в структуре 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:

Индикаторы	**Уровни освоения индикаторов компетенции**
ПК-1 Способен анализировать предметную область для выявления угроз информационной безопасности и формулировать требования к защите данных.	
ПК-1.1 Анализирует компоненты информационной системы (данные, модели, каналы связи) с точки зрения потенциальных угроз конфиденциальности, целостности и доступности.	Проводит анализ конвейера данных и AI-модели для выявления уязвимых мест.
ПК-1.2 Выбирает и обосновывает применение конкретных криптографических алгоритмов и методов для парирования выявленных угроз в рамках проектируемой системы.	Проводит сравнительный анализ cryptographic примитивов для решения конкретной прикладной задачи.
ML-6 Способен применять алгоритмы обучения с подкреплением.	
ML-6.1 Применяет методы повышения устойчивости, надежности, безопасности алгоритмов обучения с подкреплением для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.	Применяет готовые RL-библиотеки и базовые методы контроля надежности (ограничение наград, настройка гиперпараметров) для обучения и тестирования моделей в простых симулированных средах.
AI S -1 Способен управлять рисками в разработке систем ИИ, выстраивать управление безопасностью ИИ в компании с учетом этики ИИ	
AI S -1.1 Выявляет и моделирует угрозы на всём жизненном цикле ИИ-систем, оценивает и приоритизирует риски	Понимает основные категории рисков и атак на ИИ. Применяет типовые методики по готовым шаблонам. Знает международные фреймворки и стандарты.
AI S - 1.2 Обеспечивает соответствие нормативным требованиям и принципам доверенного/этичного ИИ.	Знаком с Кодексом этики в сфере ИИ РФ (2021), базовых принципах Responsible AI, законом 152-ФЗ «О перс. данных» и основами GDPR. Может описать процесс Data Impact Assessment.

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 2 зач. ед. (72 часов), их распределение по видам работ представлено в таблице

Вид учебной работы	**Всего часов**	**Семестры (часы)**					
		8					
Контактная работа, в том числе:	**30,2**	**30,2**					
Аудиторные занятия (всего):	**28**	**28**					
Занятия лекционного типа	**14**	14					
Лабораторные занятия	**14**	14					
Занятия семинарского типа (семинары, практические занятия)							
Иная контактная работа:	**0,2**	**0,2**					
Контроль самостоятельной работы (КСР)	**2**	2					

Промежуточная аттестация (ИКР)								
Самостоятельная работа, в том числе:		**41,8**	**41,8**					
Курсовая работа								
Проработка учебного (теоретического) материала								
Выполнение индивидуальных заданий (подготовка сообщений, презентаций)								
Реферат								
Подготовка к текущему контролю								
Контроль:								
Подготовка к экзамену								
Общая трудоемкость	**час.**	**72**	**72**					
	в том числе контактная работа	**30.2**	**30.2**					
	зач. ед	**2**	**2**					

2.2 Структура дисциплины

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

Разделы (темы) дисциплины, изучаемые в 8 семестре

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
1.	Введение. Математические основы и модели угроз.	16	2		2	11.8
2.	Математические модели симметричного шифрования.	10	2		2	6
3.	Математические модели хеширования и аутентификации.	8	2		2	4
4.	Математические модели асимметричного шифрования.	10	2		2	6
5.	Криптографические протоколы: математические модели доверия.	8	2		2	4
6.	Математические модели защиты конфиденциальности в AI: дифференциальная приватность.	12	2		2	8
7.	Математические модели атак на AI-системы и защиты от них.	8	2		2	4
ИТОГО по разделам дисциплины		**69,8**	**14**		**14**	**41,8**
Контроль самостоятельной работы (КСР)		2				
Промежуточная аттестация (ИКР)		0.2				
Подготовка к текущему контролю						
Общая трудоемкость по дисциплине		**72**				

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	2	3	4
1.	Введение. Математические основы и модели угроз.	Цели и задачи дисциплины в контексте разработки безопасных и этичных AI-систем. Основные понятия ИБ: Угрозы Конфиденциальности, Целостности, Доступности (КЦД) применительно к данным, ML-моделям и алгоритмам. Формальные модели нарушителя: классификация (Пассивный/Активный, Внутренний/Внешний), модель Dolev-Yao. Математический аппарат, необходимый для криптографии: модульная арифметика, алгоритм Евклида, функция Эйлера, тест Ферма на простоту. Введение в теорию сложности вычислений (классы P, NP).	*Вопросы к зачету 1-5*
2.	Математические модели симметричного шифрования.	Потоковые шифры: математическая модель гаммирования, требования к ГПСЧ (криптографически стойкие ГПСЧ), уязвимости. Блочные шифры: модель Фейстеля, структура AES (обзор). Математические основы: операции в поле Галуа (GF(2^8)). Режимы шифрования (ECB, CBC, CTR, GCM): математические схемы, преимущества и недостатки, применение для шифрования больших объемов данных (датасеты).	*Вопросы к зачету 6-10*
3.	Математические модели хеширования и аутентификации.	Определение и свойства криптографической хеш-функции (стойкость к прообразам, коллизиям). Модели построения хеш-функций (итеративная схема Меркла-Дамгарда). Обзор алгоритмов (SHA-256). Математические модели атак на хеш-функции ("парадокс дней рождений"). Коды аутентичности сообщений (MAC): модель HMAC. Цифровые подписи на основе симметричных примитивов.	*Вопросы к зачету 11-15*
4.	Математические модели асимметричного шифрования.	Концепция асимметричного шифрования. Математическая задача факторизации и схема RSA: генерация ключей, шифрование, расшифрование. Математическая задача дискретного логарифмирования и схема Эль-Гамаля. Введение в эллиптические кривые (ECC): базовые понятия, группа точек эллиптической кривой, схема ECDH. Сравнительный анализ стойкости и производительности RSA vs ECC.	*Вопросы к зачету 16-20*
5.	Криптографические протоколы: математические модели доверия.	Понятие криптографического протокола. Формальные модели безопасности. Протокол обмена ключами Диффи-Хеллмана (DH и ECDH): математическая модель, защита от атаки "man-in-the-middle". Модели протоколов электронной подписи (RSA, ECDSA). Введение в безопасные многосторонние вычисления (SMPC) и гомоморфное шифрование как модели для конфиденциальных вычислений на данных.	*Вопросы к зачету 21-25*

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
6.	Математические модели защиты конфиденциальности в AI: дифференциальная приватность.	Понятие приватности данных. Неформальное и строгое математическое определение (ε, δ)-дифференциальной приватности (DP). Чувствительность запроса. Механизмы добавления шума: Лапласа (для ε-DP) и Гаусса (для (ε, δ)-DP). Композиционные теоремы (последовательная, параллельная). Применение DP на этапах жизненного цикла AI: обучение с дифференциальной приватностью (DP-SGD), выпуск агрегированной статистики.	*Вопросы к зачету 25-30*
7.	Математические модели атак на AI-системы и защиты от них.	Формальные модели атак на конфиденциальность моделей: Membership Inference Attacks (MIA) и Model Inversion Attacks. Математическая основа Adversarial Attacks: атаки "белого ящика" (Fast Gradient Sign Method - FGSM) и "черного ящика". Модели защиты: adversarial training, регуляризация. Связь с обеспечением устойчивости и надежности алгоритмов.	*Вопросы к зачету 30-40*

2.3.2 Лабораторные занятия

№	Наименование раздела (темы)	Наименование лабораторных работ	Форма текущего контроля
1	2	3	4
1.	Введение. Математические основы и модели угроз.	Практикум по модульной арифметике в Python. Реализация алгоритма Евклида для нахождения НОД и расширенного алгоритма Евклида. Генерация больших простых чисел с использованием вероятностных тестов. Анализ кейса: формальное описание угроз КЦД для конкретного AI-конвейера (например, система рекомендаций).	ЛР
2.	Математические модели симметричного шифрования.	Реализация шифра XOR (гаммирование) на Python. Использование библиотеки cryptography для шифрования/расшифрования файла с данными с использованием AES в режимах ECB и CBC. Визуализация и анализ различий в результатах.	ЛР
3.	Математические модели хеширования и аутентификации.	Практическая работа с хеш-функциями в Python: вычисление хешей для файлов моделей (.pkl, .h5) и проверка их целостности. Реализация простой схемы HMAC. Сравнение скорости работы различных хеш-алгоритмов.	ЛР

№	Наименование раздела (темы)	Наименование лабораторных работ	Форма текущего контроля
1	2	3	4
4.	Математические модели асимметричного шифрования.	Реализация алгоритма RSA на Python (без оптимизаций, для учебных целей) для шифрования коротких сообщений. Использование библиотеки cryptography для генерации ключей ECC и выполнения обмена по ECDH. Демонстрация уязвимости textbook RSA.	ЛР
5.	Криптографические протоколы: математические модели доверия.	Моделирование протокола Диффи-Хеллмана на Python. Реализация простейшей схемы разделения секрета (Шамира). Использование библиотеки для демонстрации работы протокола электронной подписи. **Кейс 3: Защита связи между беспилотником и центром управления** **Описание:** Беспилотник передает телеметрию в центр управления. Нужно защитить канал связи. **Цель:** Реализовать **безопасный обмен ключами** между машиной и сервером. **Технологии:** – Протокол Диффи-Хеллмана – Python библиотеки: cryptography, requests **Реализация:** a) Генерация ключей на стороне автомобиля и сервера b) Обмен открытыми ключами c) Шифрование данных общим ключом **Результат:** – Защита телеметрии от перехвата – Конфиденциальность передаваемых данных	ЛР
6.	Математические модели защиты конфиденциальности в AI: дифференциальная приватность.	Практическое применение дифференциальной приватности в Python с использованием библиотек (например, diffprivlib или TensorFlow Privacy). Вычисление агрегированной статистики (среднее, гистограмма) с добавлением шума Лапласа/Гаусса. Сравнение точности и уровня приватности.	ЛР
7.	Математические модели атак на AI-системы и защиты от них.	Демонстрация Membership Inference Attack на простой классификационной модели. Реализация атаки FGSM на сверточную нейронную сеть с использованием TensorFlow/PyTorch и визуализация adversarial примеров.	ЛР

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.3 Примерная тематика курсовых работ (проектов)

Не предусмотрены.

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	2	3
1	Проработка и повторение лекционного материала, материала учебной и научной литературы, подготовка к семинарским занятиям	Методические указания для подготовки к лекционным и семинарским занятиям, утвержденные на заседании кафедры анализа данных и искусственного интеллекта факультета компьютерных технологий и прикладной математики ФГБОУ ВО «КубГУ», протокол №7 от 22.03.2023 г
2	Подготовка к текущему контролю	Методические указания для подготовки к лекционным и семинарским занятиям, утвержденные на заседании кафедры анализа данных и искусственного интеллекта факультета компьютерных технологий и прикладной математики ФГБОУ ВО «КубГУ», протокол №7 от 22.03.2023 г

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 программа дисциплины предусматривает использование в учебном процессе следующих образовательные технологии: чтение лекций с использованием мультимедийных технологий; лабораторные занятия.

С точки зрения применяемых методов используются как традиционные информационно-объяснительные лекции, так и интерактивная подача материала с мультимедийной системой. Компьютерные технологии в данном случае обеспечивают

возможность разнопланового отображения алгоритмов и демонстрационного материала. Такое сочетание позволяет оптимально использовать отведенное время и раскрывать логику и содержание дисциплины.

Лабораторное занятие позволяет научить студента применять теоретические знания при решении и исследовании конкретных задач. Лабораторные занятия проводятся в компьютерных классах. Подход разбора конкретных ситуаций широко используется как преподавателем, так и студентами при проведении анализа результатов самостоятельной работы. Это обусловлено тем, что в процессе исследования часто встречаются задачи, для которых единых подходов не существует. Каждая конкретная задача при своем исследовании имеет множество подходов, а это требует разбора и оценки целой совокупности конкретных ситуаций.

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

– 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.

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

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 «Математические модели защиты информации».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме тестовых заданий, кейсов и **промежуточной аттестации** в форме вопросов и заданий к **зачету**.

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 зачете;

– 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	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или ее части)	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация
1	Введение. Математические основы и модели угроз.	ПК-1.1, AI S-1.1	Лабораторная работа №1	Вопросы к зачету 1-5
2	Математические модели симметричного шифрования.	ПК-1.2	Лабораторная работа №2	Вопросы к зачету 6-10
3	Математические модели хеширования и аутентификации.	ПК-1.2, AI S-1.2	Лабораторная работа №3	Вопросы к зачету 11-15
4	Математические модели асимметричного шифрования.	ПК-1.2	Лабораторная работа №4	Вопросы к зачету 16-20
5	Криптографические протоколы: математические модели доверия.	ПК-1.2, ML-6.1	Лабораторная работа №5	Вопросы к зачету 21-25
6	Математические модели защиты конфиденциальности в AI: дифференциальная приватность.	AI S-1.1, AI S-1.2	Лабораторная работа №6	Вопросы к зачету 25-30
7	Математические модели атак на AI-системы и защиты от них.	ML-6.1, AI S-1.1	Лабораторная работа №7	Вопросы к зачету 30-40

Показатели, критерии и шкала оценки сформированных компетенций

№ п/п	Код и наименование индикатора	Результаты обучения	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация
Соответствие освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **удовлетворительно /зачтено**)				
на пороговом уровне:				
1.	ПК-1.1 Анализирует компоненты информационной системы (данные, модели, каналы связи) с точки зрения потенциальных угроз конфиденциальности, целостности и доступности.	Понимает базовые принципы КЦД. Может по шаблону идентифицировать 1-2 очевидные угрозы (например, "нешифрованное хранение данных", "отсутствие проверки целостности модели") в упрощенном AI-конвейере.	ЛР	Вопросы к зачету 1- 3, 31-34, 38-39
2.	ПК-1.2 Выбирает и обосновывает применение конкретных криптографических алгоритмов и методов для парирования выявленных угроз в рамках проектируемой системы.	Знает назначение основных типов алгоритмов (шифрование, хеширование, ЭЦП). Может выбрать подходящий тип алгоритма для стандартной задачи по предоставленной инструкции.	ЛР	Вопросы к зачету 4-25, 36
3.	ML-6.1 Применяет методы повышения устойчивости, надежности, безопасности алгоритмов обучения с подкреплением для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.	Понимает на примерах, что данные и среда в RL могут быть скомпрометированы, и это повлияет на работу агента. Знает о существовании adversarial-атак.	ЛР	Вопросы к зачету 14, 25, 33-35, 38
4.	AI S -1.1 Выявляет и моделирует угрозы на всём жизненном цикле ИИ-систем, оценивает и приоритизирует риски	Знает основные типы атак на ИИ (MIA, Data Poisoning). Может назвать соответствующий метод защиты (дифф. приватность, проверка данных).	ЛР	Вопросы к зачету 3, 26- 29, 31-34, 39
5.	AI S -1.2 Обеспечивает соответствие	Знает о существовании 152-ФЗ, GDPR и принципов	ЛР	Вопросы к зачету 23,

	нормативным требованиям и принципам доверенного/этичного ИИ.	этики ИИ. Понимает, что технические меры (обезличивание, шифрование) помогают выполнять эти требования.		30, 35, 37, 39-40
Соответствие освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **хорошо /зачтено**)				
<u>на базовом уровне</u>:				
1.	ПК-1.1 Анализирует компоненты информационной системы (данные, модели, каналы связи) с точки зрения потенциальных угроз конфиденциальности, целостности и доступности.	Может самостоятельно провести анализ простого конвейера, выявить несколько угроз разных типов (К, Ц, Д) и предложить к ним общие классы мер защиты.	ЛР	Вопросы к зачету 1- 3, 31-34, 38-39
2.	ПК-1.2 Выбирает и обосновывает применение конкретных криптографических алгоритмов и методов для парирования выявленных угроз в рамках проектируемой системы.	Может обосновать выбор алгоритма, сравнив базовые свойства (скорость, симметричность). Уверенно использует готовые библиотеки для типовых задач.	ЛР	Вопросы к зачету 4-25, 36
3.	ML-6.1 Применяет методы повышения устойчивости, надежности, безопасности алгоритмов обучения с подкреплением для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.	Может привести конкретные примеры, как искажение данных ломает политику агента. Понимает и может объяснить роль базовых криптографических примитивов в защите RL.	ЛР	Вопросы к зачету 14, 25, 33-35, 38
4.	AI S -1.1 Выявляет и моделирует угрозы на всём жизненном цикле ИИ-систем, оценивает и приоритизирует риски	Может проанализировать конкретный случай утечки данных в ИИ-системе и предложить математическую модель (дифф. приватность) для ее предотвращения.	ЛР	Вопросы к зачету 3, 26- 29, 31-34, 39
5.	AI S -1.2 Обеспечивает соответствие нормативным требованиям и принципам доверенного/этичного ИИ.	Может описать, как конкретная мера (дифф. приватность) обеспечивает соответствие 152-ФЗ. Понимает связь между	ЛР	Вопросы к зачету 23, 30, 35, 37, 39-40

		технических решениями и этическими принципами.		
Соответствие освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **отлично /зачтено**)				
на продвинутом уровне:				
1.	ПК-1.1 Анализирует компоненты информационной системы (данные, модели, каналы связи) с точки зрения потенциальных угроз конфиденциальности, целостности и доступности.	Способен предложить и обосновать комплекс мер для защиты реального AI-конвейера, учитывая взаимосвязь угроз на разных этапах.	ЛР	Вопросы к зачету 1- 3, 31-34, 38-39
2.	ПК-1.2 Выбирает и обосновывает применение конкретных криптографических алгоритмов и методов для парирования выявленных угроз в рамках проектируемой системы.	Может предложить гибридную схему использования алгоритмов для новой задачи и оценить ее безопасность. Способен найти и использовать специализированные криптобиблиотеки.	ЛР	Вопросы к зачету 4-25, 36
3.	ML-6.1 Применяет методы повышения устойчивости, надежности, безопасности алгоритмов обучения с подкреплением для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.	Способен предложить и реализовать простой метод повышения устойчивости RL-агента к определенному типу атак на основе изученных принципов.	ЛР	Вопросы к зачету 14, 25, 33-35, 38
4.	AI S -1.1 Выявляет и моделирует угрозы на всём жизненном цикле ИИ-систем, оценивает и приоритизирует риски	Способен провести сравнительный анализ эффективности разных математических моделей защиты (дифф. приватность, гомоморфное шифрование) для заданного сценария.	ЛР	Вопросы к зачету 3, 26- 29, 31-34, 39
5.	AI S -1.2 Обеспечивает соответствие нормативным требованиям и принципам доверенного/этичного ИИ.	Может критически оценить ИИ-систему на предмет соответствия нормам и этическим принципам и составить список рекомендаций по устранению нарушений.	ЛР	Вопросы к зачету 23, 30, 35, 37, 39-40

Зачетно-экзаменационные материалы для промежуточной аттестации

Примерный перечень вопросов для зачета

Вопросы зачета	Перечень компетенций (части компетенции)
1. Что такое модель нарушителя в информационной безопасности?	ПК-1.1
2. Что понимается под угрозами конфиденциальности, целостности и доступности (КЦД) применительно к AI-системам?	ПК-1.1
3. Какие основные этапы жизненного цикла AI-системы необходимо анализировать на предмет уязвимостей?	ПК-1.1, AI S-1.1
4. Что такое модульная арифметика и почему она важна для криптографии?	ПК-1.2
5. Что такое алгоритм Евклида и как он используется в криптографии?	ПК-1.2
6. Что такое симметричное шифрование? В чем его основные преимущества и недостатки?	ПК-1.2
7. Что такое блочные шифры и чем они отличаются от потоковых?	ПК-1.2
8. Что такое режимы шифрования (ECB, CBC, CTR)? В чем недостаток режима ECB?	ПК-1.2
9. Что такое алгоритм AES? Какие ключевые особенности его математической структуры?	ПК-1.2
10. Как симметричное шифрование может применяться для защиты данных в AI-системах?	ПК-1.2
11. Что такое криптографическая хеш-функция? Какие свойства она должна обладать?	ПК-1.2
12. Что такое HMAC и для каких целей он используется?	ПК-1.2
13. В чем разница между хеш-функцией и электронной подписью?	ПК-1.2
14. Как хеш-функции могут обеспечить целостность данных в системах обучения с подкреплением?	ПК-1.2, ML-6.1
15. Что такое "парадокс дней рождений" в контексте хеш-функций?	ПК-1.2
16. В чем основное различие между симметричным и асимметричным шифрованием?	ПК-1.2
17. Что такое алгоритм RSA и какая математическая задача лежит в его основе?	ПК-1.2
18. Что такое алгоритм Диффи-Хеллмана и для чего он используется?	ПК-1.2
19. Что такое эллиптические кривые в криптографии? Какие преимущества у ECC перед RSA?	ПК-1.2
20. Что такое гибридные криптосистемы и почему они широко используются на практике?	ПК-1.2
21. Что такое криптографический протокол? Приведите примеры протоколов аутентификации.	ПК-1.2
22. Что такое протокол электронной подписи и какие свойства он обеспечивает?	ПК-1.2

23. Что такое инфраструктура открытых ключей (PKI) и как она связана с доверенным ИИ?	ПК-1.2, AI S-1.2
24. Что такое атака "человек посередине" (Man-in-the-Middle) и как от нее защититься?	ПК-1.2
25. Как криптографические протоколы могут обеспечить безопасное взаимодействие в распределенных AI-системах?	ПК-1.2, ML-6.1
26. Что такое дифференциальная приватность? Дайте неформальное и формальное определение.	AI S-1.1
27. Что означают параметры ε и δ в дифференциальной приватности?	AI S-1.1
28. Что такое механизм Лапласа и как он используется в дифференциальной приватности?	AI S-1.1
29. Что такое чувствительность запроса и как она влияет на добавление шума?	AI S-1.1
30. Как дифференциальная приватность связана с выполнением требований GDPR и 152-ФЗ?	AI S-1.2
31. Что такое Membership Inference Attack и как она угрожает конфиденциальности данных?	ПК-1.1, AI S-1.1
32. Что такое Model Inversion Attack? Какая информация может быть восстановлена?	ПК-1.1, AI S-1.1
33. Что такое Adversarial Attacks на модели машинного обучения? Приведите примеры.	ПК-1.1, ML-6.1, AI S-1.1
34. Что такое атака на основе тренировочных данных (Data Poisoning)?	ПК-1.1, ML-6.1, AI S-1.1
35. Какие методы повышения устойчивости AI-моделей к атакам вы знаете?	ML-6.1, AI S-1.2
36. Какие криптографические примитивы вы выберете для защиты конвейера данных в AI-системе и почему?	ПК-1.2
37. Как обеспечить соответствие AI-системы принципам ответственного ИИ с помощью криптографических методов?	AI S-1.2
38. Какие угрозы безопасности наиболее критичны для систем обучения с подкреплением и как их mitigate?	ПК-1.1, ML-6.1
39. Как провести оценку рисков для AI-системы обработки персональных данных?	ПК-1.1, AI S-1.1, AI S-1.2
40. Какие тенденции в развитии математических моделей защиты информации наиболее актуальны для AI-систем будущего?	AI S-1.2

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

Методические рекомендации, определяющие процедуры оценивания ответа на зачете:

Шкала оценивания зачета

«*Зачтено*»**:** Студент демонстрирует понимание фундаментальных математических моделей защиты информации и их связи с практикой AI. Дает точные определения ключевых понятий (КЦД, модель нарушителя, свойства хеш-функций, принципы диф. приватности и т.д.). Объясняет, как криптографические методы применяются в контексте AI (напр., «AES используется для шифрования датасетов на диске», «SHA-256 применяется

для контроля целостности файлов моделей», «Дифференциальная приватность защищает от Membership Inference Attack»). Может сравнить алгоритмы (RSA vs ECC, симметричное vs асимметричное шифрование) и обосновать выбор для простого сценария. бъясняет конкретные угрозы для AI-систем (MIA, Data Poisoning, Adversarial Attacks) и называет соответствующие математические модели защиты.

«*Не зачтено***»:** Отсутствие ключевых понятий или грубые ошибки; изученные математические модели применяются для решения задач защиты информации в AI; не может разобрать коды из лабораторных, проблемы найти связь в темах. Не может объяснить разницу между основными классами алгоритмов или привести пример угрозы и метода защиты.

Методические рекомендации, определяющие процедуры оценивания лабораторных работ:

Оценка зависит от:

- **Качества реализации** (работоспособность кода, эффективность)
- **Анализа результатов** (интерпретация, сравнение методов)
- **Оформления и отчетности** (четкость, соответствие стандартам)

Уровень	Требования
Теоретическая подготовка (20%)	
Пороговый (3/5)	Базовые определения и формулы без глубокого объяснения их роли в защите информации.
Базовый (4/5)	Обоснование выбора алгоритмов и методов защиты, ссылки на теоретические основы.
Продвинутый (5/5)	Сравнение альтернативных криптографических подходов, анализ стойкости и ограничений методов.
Практическая реализация (40%)	
Пороговый (3/5)	Код работает, но без обработки исключений; использование готовых функций по шаблону.
Базовый (4/5)	Чистый код с комментариями, обработкой ошибок, тестированием на различных данных.
Продвинутый (5/5)	Реализация дополнительных функций, оптимизация производительности, сравнение алгоритмов.
Анализ результатов (30%)	
Пороговый (3/5)	Базовые выводы без интерпретации в контексте безопасности AI-систем.

Базовый (4/5)	Сравнение методов с обоснованием выбора, анализ компромиссов.
Продвинутый (5/5)	Анализ влияния параметров на безопасность, оценка рисков, предложения по улучшению защиты.
Оформление отчета (10%)	
Пороговый (3/5)	Есть введение, код и выводы.
Базовый (4/5)	Структурированный отчет с графиками и таблицами.
Продвинутый (5/5)	LaTeX-документ или Jupyter Notebook с интерактивными визуализациями.

Шкала оценивания

Балл	Уровень выполнения	Соответствие критериям
5 (Отлично)	Продвинутый	Все критерии выполнены на высоком уровне + инновации.
4 (Хорошо)	Базовый	Полное соответствие без серьезных недочетов.
3 (Удовл.)	Пороговый	Минимальные требования выполнены, но есть недочеты.
2 (Неуд.)	Ниже порога	Критические ошибки в теории/практике.

Дополнительные критерии по тематикам

Анализ угроз и модели нарушителя (ПК-1.1)

- **Пороговый:** Выявление очевидных угроз КЦД в заданном AI-конвейере по шаблону.
- **Базовый:** Самостоятельный анализ конвейера, классификация угроз с использованием моделей нарушителя.
- **Продвинутый:** Построение комплексной модели угроз для распределенной AI-системы с оценкой рисков.

Криптографические методы защиты (ПК-1.2)

- **Пороговый:** Корректное использование готовых криптографических функций (AES, RSA, SHA-256) по предоставленным примерам.
- **Базовый:** Сравнительный анализ алгоритмов шифрования/хеширования, обоснование выбора для конкретной задачи защиты данных.
- **Продвинутый:** Реализация схемы защиты с оптимизацией параметров для заданных требований безопасности.

Безопасность систем обучения с подкреплением (ML-6.1)

- **Пороговый:** Понимание уязвимостей RL-систем и способов защиты на концептуальном уровне.

- **Базовый:** Реализация механизмов проверки целостности данных среды с использованием HMAC.
- **Продвинутый:** Разработка и тестирование метода защиты от adversarial-атак на функцию вознаграждения.

Математические модели приватности (AI S-1.1)

- **Пороговый:** Применение дифференциальной приватности с заданными параметрами (ε, δ) по шаблону.
- **Базовый:** Анализ компромисса между точностью модели и уровнем приватности, выбор оптимальных параметров.
- **Продвинутый:** Разработка адаптивного механизма приватности для динамического датасета.

Этика и нормативное соответствие (AI S-1.2)

- **Пороговый:** Знание основных положений 152-ФЗ, GDPR и Кодекса этики ИИ.
- **Базовый:** Ан Проведение Data Impact Assessment для проекта с персональными данными.
- **Продвинутый:** Разработка организационных мер и технических стандартов для обеспечения доверенного ИИ.

4.3 Методические указания по организации лабораторных работ по дисциплине " Математические модели защиты информации"

1. Общие сведения

Образовательная программа: «Искусственный интеллект и аналитика данных»
Дисциплина: «Математические модели защиты информации»
Вид обеспечения: Проведение лабораторных работ
Условия применения:
Для успешного выполнения лабораторных работ необходимо обеспечить:

1) **Программное обеспечение:**
   - Python 3.8+ (с библиотеками: cryptography, pycryptodome, hashlib, numpy, pandas, matplotlib, jupyter)
   - TensorFlow/PyTorch (для работ по adversarial attacks)
   - Специализированные библиотеки приватности: diffprivlib, tensorflow-privacy, opacus
   - Дополнительные инструменты: OpenSSL, Wireshark (для анализа сетевого трафика)
2) **Аппаратное обеспечение:**
   - Компьютеры с ОС Windows/Linux, поддерживающие Python и необходимые библиотеки
   - Достаточные вычислительные ресурсы (CPU/GPU) для работы с моделями машинного обучения
3) **Облачная инфраструктура (опционально):**
   - Доступ к Yandex Cloud / AWS для развертывания защищенных решений
   - Google Colab Pro / Kaggle для работы с GPU

2. Цели, задачи и ожидаемые результаты выполнения лабораторных работ

2.1. Обоснование необходимости лабораторных работ

Лабораторные работы в данной дисциплине необходимы, поскольку:

a) **Практическое закрепление теории**: Криптографические методы и модели защиты требуют не только теоретического понимания, но и практических навыков реализации.

b) **Формирование системного мышления:** Студенты учатся анализировать угрозы и выбирать адекватные методы защиты для AI-систем.

c) **Подготовка к реальным проектам:** Работа с современными библиотеками и инструментами формирует компетенции, востребованные в индустрии.

2.2. Задачи лабораторных работ

1. **Освоение криптографических примитивов:**
   - Реализация симметричного и асимметричного шифрования.
   - Работа с хеш-функциями и электронными подписями.
2. **Анализ угроз безопасности:**
   - Идентификация уязвимостей в AI-конвейерах
   - Моделирование атак на конфиденциальность данных
3. **Применение методов защиты:**
   - Реализация дифференциальной приватности
   - Защита от adversarial attacks

2.3. Ожидаемые результаты

После выполнения лабораторных работ студенты смогут:

✓ Анализировать угрозы безопасности в AI-системах

✓ Выбирать и применять криптографические алгоритмы для защиты данных

✓ Реализовывать механизмы дифференциальной приватности

✓ Защищать ML-модели от adversarial attacks

✓ Обеспечивать соответствие требованиям 152-ФЗ и GDPR

2.4. Что необходимо для реализации лабораторных работ?

1) **Методические материалы:**
   - Пошаговые инструкции по каждой теме.
   - Примеры кода (Python).
   - Готовые датасеты для экспериментов.

2) **Техническая поддержка:**
   - Настройка ПО на компьютерах.
   - Доступ к облачным ресурсам (при необходимости).

3) **Контроль выполнения:**
   - Отчеты по каждой работе.

Лабораторные работы в данной дисциплине позволяют студентам **перейти от теории к практике**, освоив ключевые криптографические методы и модели защиты. Это формирует **навыки, необходимые для работы в области криптографии**, включая решения практических задач защиты информации в AI-системах.

Формат:

- Каждая работа включает **теоретическую справку, пошаговое руководство и задания для самостоятельного выполнения**.
- Рекомендуется **сочетание индивидуальной и групповой работы**.

Таким образом, предложенный план лабораторных работ обеспечит **системный подход к обучению** и подготовит студентов к решению реальных задач в математических моделей систем защиты информации.

3. Порядок реализации лабораторных работ

3.1. Задача №1: Провести анализ угроз безопасности для простого AI- конвейера

Цель: Научиться анализировать компоненты AI-конвейера с точки зрения информационной безопасности, классифицировать угрозы по модели КЦД.

Тема 1: Анализ угроз безопасности для простого AI- конвейера

Пошаговая инструкция:

- Дана схема конвейера: "Пользователь загружает данные через веб-форму -> Данные хранятся в базе данных -> Модель ML делает прогноз".
- Для каждого элемента схемы (веб-форма, база данных, модель) предложить по 1-2 возможные угрозы (например, перехват данных при загрузке, утечка из БД, подмена модели).
- Классифицировать каждую угрозу по нарушаемому свойству (Конфиденциальность, Целостность, Доступность).
- Сделать выводы о том, для какого элемента конвейера угрозы кажутся наиболее критичными и почему.

3.2. Задача №2: Освоение симметричного шифрования

Цель: Изучить практическое применение симметричного шифрования для защиты данных.

Тема 2: Математические модели симметричного шифрования

Пошаговая инструкция:

- Взять небольшое черно-белое изображение (например, 100x100 пикселей).
- Используя готовую функцию из библиотеки, зашифровать это изображение с помощью AES в режимах ECB и CBC.
- Отобразить исходное и оба зашифрованных изображения.
- Описать визуальные различия. Объяснить, почему в режиме ECB проступают контуры.
- Сделать выводы о том, почему режим ECB считается небезопасным для структурированных данных, и в каких сценариях предпочтительнее использовать режим CBC или другие.

3.3. Задача №3: Освоение методов хеширования

Цель: Научиться применять хеш-функции для контроля целостности данных.

Тема 3: Создание "цифрового отпечатка" для модели

Пошаговая инструкция:

- Сохранить в файл веса простой модели (например, линейной регрессии из scikit-learn).
- Вычислить хеш-сумму этого файла с помощью алгоритма SHA-256.
- Имитировать инцидент: изменить одну цифру в сохраненном файле весов с помощью текстового редактора.
- Вычислить хеш-сумму измененного файла и сравнить с исходной.
- Сделать выводы о том, как хеш-функции позволяют обнаружить малейшее изменение данных и пригодны для проверки целостности и аутентичности моделей и данных.

3.4. Задача №4: Освоение асимметричного шифрования

Цель: Изучить принципы работы алгоритма Диффи-Хеллмана.

Тема 4: Математические модели асимметричного шифрования (алгоритм Диффи-Хеллмана)

Пошаговая инструкция:

- Используя библиотеку, сгенерировать для двух пользователей (Алисы и Боба) их пары ключей (закрытый и открытый) для протокола Диффи-Хеллмана.
- Продемонстрировать, что Алиса и Боб, обменявшись открытыми ключами, могут независимо вычислить один и тот же общий секрет.
- Показать, что злоумышленник, перехвативший только открытые ключи, не может легко вычислить этот общий секрет.

- Сделать выводы о фундаментальной роли сложности задачи дискретного логарифмирования в безопасности протокола и его значении для установки защищенных соединений.

3.5. Задача №5: Освоение электронной подписи

Цель: Изучить принципы работы электронной подписи.

Тема 5: Криптографические протоколы: математические модели доверия

Пошаговая инструкция:

- Сгенерировать пару ключей (закрытый и открытый) для алгоритма RSA.
- Создать короткое текстовое сообщение (например, "Результат анализа: 5.04").
- Подписать это сообщение своим закрытым ключом.
- Проверить подпись с помощью открытого ключа. Убедиться, что проверка проходит.
- Изменить сообщение и убедиться, что проверка подписи теперь не проходит.
- Сделать выводы о том, как электронная подпись обеспечивает целостность и аутентичность данных, что позволяет, например, проверять подлинность конфигураций моделей или входных данных.

3.6. Задача №6: Освоение дифференциальной приватности

Цель: Изучить применение дифференциальной приватности для защиты данных.

Тема 6: Математические модели защиты конфиденциальности в AI: дифференциальная приватность

Пошаговая инструкция:

- Создать небольшой набор данных с одной колонкой "Зарплата".
- Вычислить и вывести реальное среднее значение зарплат.
- Используя простой механизм Лапласа, добавить шум к вычислению среднего значения, чтобы обеспечить дифференциальную приватность (ε=1.0).
- Повторить пункт 3 несколько раз, чтобы увидеть, как меняется "зашумленный" результат.
- Сделать выводы о том, как дифференциальная приватность "размывает" точный ответ, защищая приватность отдельных людей, но сохраняя полезность данных на групповом уровне.

3.7. Задача №7: Освоение защиты от adversarial-атак

Цель: Изучить методы создания и защиты от adversarial-атак.

Тема 7: Демонстрация уязвимости модели к Adversarial

Пошаговая инструкция:

- Загрузить предобученную модель для классификации изображений и тестовую картинку.
- Используя готовую реализацию атаки FGSM, создать adversarial-пример.
- Убедиться, что модель теперь классифицирует adversarial-пример (который выглядит для человека как "панда") неправильно (например, как "гиббон").
- Визуализировать разницу между исходным и adversarial-изображением.
- Сделать выводы о хрупкости современных моделей ИИ и важности тестирования их на устойчивость перед развертыванием в реальных системах.

Итоговая структура каждой лабораторной работы:

1. **Теоретическая часть** (краткое описание методов).
2. **Пошаговые инструкции** (код + пояснения).
3. **Индивидуальные задания** (задачи на самостоятельное выполнение).
4. **Контрольные вопросы**.

Такой подход обеспечит **постепенное усложнение задач** и **интеграцию знаний** из разных тем.

Соответствие лабораторных работ и индикаторов компетенций

В таблице ниже представлено, как каждая лабораторная работа способствует формированию заявленных компетенций.

Компетенция	Соответствующие лабораторные работы	Обоснование
ПК-1.1 Анализирует компоненты информационной системы (данные, модели, каналы связи) с точки зрения потенциальных угроз конфиденциальности, целостности и доступности.	**Тема 1** (Анализ угроз безопасности для простого AI- конвейера)	Работа формирует навыки анализа уязвимостей и классификации угроз по модели КЦД в контексте AI-систем. Студенты учатся идентифицировать точки уязвимости на всех этапах жизненного цикла данных.
ПК-1.2 Выбирает и обосновывает применение конкретных криптографических алгоритмов и методов для парирования выявленных угроз в рамках проектируемой системы.	**Тема 2** (Освоение симметричного шифрования) **Тема 3** (Создание "цифрового отпечатка" для модели) **Тема 4** (Математические модели асимметричного шифрования (алгоритм Диффи-Хеллмана)) **Тема 5** (Криптографические протоколы: математические модели доверия)	Работы позволяют освоить практическое применение криптографических примитивов и обоснование их выбора для защиты конфиденциальности и целостности данных в AI-системах.
ML-6.1 Применяет методы повышения устойчивости, надежности, безопасности алгоритмов обучения с подкреплением для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.	**Тема 7** (Демонстрация уязвимости модели к Adversarial)	Работа демонстрирует уязвимости ML-моделей к атакам и методы обеспечения устойчивости, что критически важно для надежности RL-алгоритмов в реальных системах.

Компетенция	Соответствующие лабораторные работы	Обоснование
AI S -1.1 Выявляет и моделирует угрозы на всём жизненном цикле ИИ-систем, оценивает и приоритизирует риски	**Тема 1** (Анализ угроз безопасности для простого AI- конвейера) **Тема 7** (Демонстрация уязвимости модели к Adversarial)	Работы формируют системное понимание угроз безопасности throughout жизненного цикла AI-систем - от данных до развернутых моделей.
AI S -1.2 Обеспечивает соответствие нормативным требованиям и принципам доверенного/этичного ИИ.	**Тема 6** (Математические модели защиты конфиденциальности в AI: дифференциальная приватность)	Работа напрямую связана с выполнением требований 152-ФЗ и GDPR через математические методы обеспечения конфиденциальности данных.

Детализация по темам

1. ПК-1.1 (Анализ компонентов информационной системы с точки зрения угроз)

Тема 1 (Анализ угроз безопасности для простого AI- конвейера):

- Идентификация уязвимостей точек передачи данных → необходимо для построения комплексной системы безопасности AI-систем
- Классификация угроз по модели КЦД → помогает приоритизировать меры защиты на разных этапах жизненного цикла данных

2. ПК-1.2 (Выбор и обоснование применения криптографических алгоритмов)

Тема 2 (Освоение симметричного шифрования):

- Сравнение режимов шифрования → обоснование выбора AES-CBC для защиты структурированных данных
- Анализ уязвимостей ECB → понимание требований к защите датасетов и моделей.

Тема 3 (Создание "цифрового отпечатка" для модели):

- Контроль целостности моделей → применение SHA-256 для верификации файлов весов
- Обнаружение изменений данных → основа для систем мониторинга целостности AI-конвейера

Тема 4 (Математические модели асимметричного шифрования (алгоритм Диффи-Хеллмана)):

- Безопасный обмен ключами → основа для защищенного взаимодействия компонентов распределенной AI-системы
- Математические основы стойкости → понимание роли задачи дискретного логарифмирования.

Тема 5 (Криптографические протоколы: математические модели доверия):

- Аутентификация источников данных → защита от подмены моделей и входных данных
- Гарантия целостности конфигураций → обеспечение доверия к настройкам AI-систем.

3. ML-6.1 (Применение методов повышения устойчивости и безопасности алгоритмов)

Тема 7 (Демонстрация уязвимости модели к Adversarial):

- Защита RL-алгоритмов от манипуляций → обеспечение устойчивости политик агента
- Методы обнаружения аномалий в данных среды → основа для безопасного обучения с подкреплением.

4. AI S-1.1 (Выявление и моделирование угроз на жизненном цикле ИИ-систем)

Тема 1 (Анализ угроз безопасности для простого AI- конвейера):

- Моделирование угроз на всех этапах AI-конвейера → комплексный подход к безопасности
- Оценка рисков утечки данных → основа для построения системы управления рисками

Тема 7 (Демонстрация уязвимости модели к Adversarial):

- Моделирование атак на развернутые AI-системы → проактивный подход к безопасности
- Анализ последствий компрометации моделей → оценка бизнес-рисков

5. AI S-1.2 (Обеспечение соответствия нормативным требованиям)

Тема 6 (Математические модели защиты конфиденциальности в AI: дифференциальная приватность):

- Реализация требований 152-ФЗ к обезличиванию данных → правовое обоснование технических мер
- Баланс точности и приватности → соблюдение принципов ответственного ИИ при обработке персональных данных

Предложенные лабораторные работы покрывают **все заявленные индикаторы компетенций**, обеспечивая:

1) Практические навыки анализа угроз и выбора мер защиты (ПК-1.1, ПК-1.2)
2) Понимание уязвимостей AI-систем и методов обеспечения устойчивости (ML-6.1, AI S-1.1)
3) Соответствие нормативным требованиям и этическим принципам (AI S-1.2)

Рекомендация: Включить в отчеты по лабораторным работам раздел "Обоснование выбранных методов защиты", чтобы явно отражать формирование компетенций ПК-1.2 и AI S-1.2.

Чек-лист для проверки выполнения лабораторных работ

Этот чек-лист поможет преподавателю и студентам убедиться, что все этапы работ выполнены корректно.

Общий чек-лист для всех лабораторных работ

1) Подготовка к работе

- Установлено необходимое ПО (Python с библиотеками криптографическими и для машинного обучения, OpenSSL, Git).
- Проверена работоспособность среды (запуск примеров из документации).
- Изучены методические указания и теоретическая часть.

2) Выполнение заданий

- Код написан без синтаксических ошибок.
- Результаты вычислений соответствуют ожидаемым (проверка на тестовых данных).
- Для символьных вычислений проведена аналитическая проверка.
- Для численных методов проведена оценка точности (сравнение с аналитическим решением, если возможно).

3) Отчетность

Отчет содержит:

- Цель работы.
- Листинги кода с комментариями.
- Результаты вычислений (графики, таблицы, выводы).
- Ответы на контрольные вопросы.

Оформление соответствует стандартам (титульный лист, нумерация страниц).

Чек-листы по темам

Тема 1: Анализ угроз для простого AI-конвейера

Критерии проверки:

- Выявлены угрозы для всех 3 компонентов системы.
- Корректная классификация по модели КЦД.
- Обоснование критичности угроз.
- Структурированность отчета.

Пример проверки:

Угрозы для веб-формы:
- Конфиденциальность: перехват данных при передаче
- Целостность: внедрение вредоносного кода
- Доступность: DDoS-атака на форму
Наиболее критичные: утечка данных из БД (нарушение 152-ФЗ)

Тема 2: Математические модели симметричного шифрования

Критерии проверки:

- Корректное шифрование в двух режимах
- Визуализация различий
- Объяснение уязвимостей ECB
- Практические выводы

Пример проверки:

python

```python
# Тестовый код для проверки
from cryptography.hazmat.primitives.ciphers import Cipher, algorithms, modes
import numpy as np

# Проверка корректности шифрования
def test_encryption():
    test_data = b'test_data_for_encryption'
    # Шифрование ECB
    cipher_ecb = Cipher(algorithms.AES(key), modes.ECB())
    encryptor_ecb = cipher_ecb.encryptor()
    encrypted_ecb = encryptor_ecb.update(test_data) + encryptor_ecb.finalize()

    # Должны быть разные результаты для разных режимов
    assert encrypted_ecb != encrypted_cbc
    return True
```

Тема 3: Создание "цифрового отпечатка" для модели

Критерии проверки:

- Создание и сохранение модели
- Корректное вычисление хеш-сумм

- Обнаружение изменений в файле
- Анализ применения в реальных системах

Пример проверки:

```python
# Проверка хеширования
def test_hashing():
    original_hash = 'a1b2c3d4e5f67890123456789012345678901234567890123456789012345678'
    modified_hash = 'f6e5d4c3b2a19876543210987654321098765432109876543210987654321098'

    # Хеши должны различаться при изменении файла
    assert original_hash != modified_hash
    print("✓ Целостность данных проверена успешно")
    return True
```

Тема 4: Математические модели асимметричного шифрования (алгоритм Диффи-Хеллмана)

Критерии проверки:

- Генерация ключевых пар
- Корректный обмен и вычисление секрета
- Анализ стойкости протокола
- Понимание математических основ

Пример проверки:

```python
# Проверка протокола Диффи-Хеллмана
def test_diffie_hellman():
    # Секреты должны совпадать
    assert alice_secret == bob_secret
    # Открытые ключи не должны быть равны
    assert alice_public != bob_public
    # Секрет не должен быть равен открытым ключам
    assert alice_secret not in [alice_public, bob_public]
    return True
```

Тема 5: Криптографические протоколы: математические модели доверия

Критерии проверки:

- Генерация ключевой пары RSA
- Корректное подписание сообщения
- Успешная проверка подписи
- Обнаружение подделки сообщения

Пример проверки:

```python
# Проверка ЭЦП
def test_signature():
    try:
        public_key.verify(signature, original_message)
        print("✓ Подпись верна")

        # Попытка проверки измененного сообщения
```

```
            public_key.verify(signature, modified_message)
            print("✗ Ошибка: подпись прошла проверку для измененного сообщения")
            return False
        except InvalidSignature:
            print("✓ Подпись корректно отклонена для измененного сообщения")
            return True
```

Тема 6: Математические модели защиты конфиденциальности в AI: дифференциальная приватность

Критерии проверки:

- Реализация механизма Лапласа
- Анализ влияния параметра ε
- Понимание компромисса точность-приватность
- Практические выводы для AI-систем

Пример проверки:

python

```
# Проверка дифференциальной приватности
def test_differential_privacy():
    original_data = [50000, 55000, 60000, 45000, 70000]
    true_mean = np.mean(original_data)

    # Несколько запусков с одинаковым ε
    results = [laplace_mechanism(original_data, 1.0) for _ in range(5)]

    # Результаты должны различаться (из-за шума)
    assert len(set(results)) > 1
    # Среднее зашумленных результатов должно быть близко к истинному
    assert abs(np.mean(results) - true_mean) < 10000
    return True
```

Тема 7: Демонстрация уязвимости модели к Adversarial

Критерии проверки:

- Создание adversarial-примера
- Демонстрация изменения предсказания модели
- Визуализация различий
- Анализ методов защиты

Пример проверки:

python

```
# Проверка adversarial-атаки
def test_adversarial_attack():
    # Исходное предсказание
    original_pred = model.predict(original_image)

    # Предсказание для adversarial-примера
    adversarial_pred = model.predict(adversarial_image)

    # Предсказания должны различаться
```

```
    assert np.argmax(original_pred) != np.argmax(adversarial_pred)


    # Визуальное отличие должно быть минимальным
    difference = np.mean(np.abs(original_image - adversarial_image))
    assert difference < 0.1  # Максимальное допустимое отличие
    return True
```

Итоговый контроль

- Все задания выполнены в срок.
- Отчет защищен (студент может объяснить любую часть кода).
- Результаты воспроизводимы (запуск на другом компьютере дает тот же вывод).

Преподаватель может использовать этот чек-лист при проверке работ, а также рекомендовать студентам применять его для самоконтроля. Этот подход минимизирует ошибки и обеспечит системное освоение компетенций.

5. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

5.1 Литература:

1. Зенков, А. В. Информационная безопасность и защита информации : учебник для вузов / А. В. Зенков. - 2-е изд., перераб. и доп. - Москва : Юрайт, 2025. - 107 с. - URL: https://urait.ru/bcode/567915 (дата обращения: 05.03.2025). - Режим доступа: для авториз. пользователей. - ISBN 978-5-534-16388-9. - Текст : электронный. URL:http://megapro.kubsu.ru/MegaPro/UserEntry?Action=Link_FindDoc&id=281310&idb=0
2. Информационный мир XXI века : криптография - основа информационной безопасности / под редакцией Э. А. Белова ; Московский государственный технический университет гражданской авиации. - 7-е изд. - Москва : Издательско-торговая корпорация "Дашков и К°", 2024. - 125 с. : ил. - (Библиотека "Книга будущего инженера"). - Библиогр.: с. 124-125. - ISBN 978-5-394-05556-0 : 162 р. - Текст : непосредственный. URL: http://megapro.kubsu.ru/MegaPro/UserEntry? Action=Link_FindDoc&id=271915&idb=0
3. Чернышев, С. А. Основы программирования на Python : учебное пособие для вузов / С. А. Чернышев. — Москва : Издательство Юрайт, 2022. — 286 с. — (Высшее образование). — ISBN 978-5-534-14350-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/496893. (дата обращения: 19.07.2025).
4. Платонов, А. В. Машинное обучение : учебное пособие для вузов / А. В. Платонов. — Москва : Издательство Юрайт, 2022. — 85 с. — (Высшее образование). — ISBN 978-5-534-15561-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/508804 (дата обращения: 19.07.2025).
5. Рабчевский, А. Н. Синтетические данные и развитие нейросетевых технологий : учебное пособие для вузов / А. Н. Рабчевский. — Москва : Издательство Юрайт, 2024. — 187 с. — (Высшее образование). — ISBN 978-5-534-17716-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/545036 (дата обращения: 19.07.2025).
6. Саломаа А. Криптография с открытым ключом: Пер. с англ. — М.: Мир, 1995. — 318 с.
7. Девянин П.Н. Модели безопасности компьютерных систем. Управление доступом и информационными потоками. / П.Н. Девянин. М.: Горячая линия - Телеком, 2012. 320 с.
8. С.О. Крамаров Криптографическая защита информации : учебное пособие / С.О. Крамаров, О.Ю. Митясова, С.В. Соколов [и др.] ; под ред. С.О. Крамарова. — Москва : РИОР : ИНФРА-М, 2023. — 321 с.

9. Суворова, Г. М. Информационная безопасность : учебное пособие для вузов / Г. М. Суворова. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 277 с.
10. Ерёменко, В.Т. Математические основы защиты информации: учебное пособие для высшего профессионального образования / В.Т. Еременко, М.Н. Орешина, Н.Г. Пеньков – Орёл: ФГБОУ ВПО «Госуниверситет - УНПК», 2015. – 210 с.
11. Росенко А.П. Внутренние угрозы безопасности конфиденциальной информации / А.П. Росенко. Методология и теоретическое исследование. – М: Красанд. – 2010. – 560 с.
12. 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
13. 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
14. 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,

5.2. Периодические издания: издания и конференции (А*):

1. Базы данных компании «Ист Вью» http://dlib.eastview.com
2. 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» http://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

Профессиональные базы данных

1. Scopus http://www.scopus.com/
2. ScienceDirect https://www.sciencedirect.com/
3. Журналы издательства Wiley https://onlinelibrary.wiley.com/
4. Научная электронная библиотека (НЭБ) http://www.elibrary.ru/
5. Полнотекстовые архивы ведущих западных научных журналов на Российской платформе научных журналов НЭИКОН http://archive.neicon.ru
6. Национальная электронная библиотека (доступ к Электронной библиотеке диссертаций Российской государственной библиотеки (РГБ) https://rusneb.ru/
7. Президентская библиотека им. Б.Н. Ельцина https://www.prlib.ru/
8. База данных CSD Кембриджского центра кристаллографических данных (CCDC) https://www.ccdc.cam.ac.uk/structures/
9. Springer Journals: https://link.springer.com/
10. Springer Journals Archive: https://link.springer.com/
11. Nature Journals: https://www.nature.com/
12. Springer Nature Protocols and Methods: https://experiments.springernature.com/sources/springer-protocols
13. Springer Materials: http://materials.springer.com/
14. Nano Database: https://nano.nature.com/
15. Springer eBooks (i.e. 2020 eBook collections): https://link.springer.com/
16. "Лекториум ТВ" http://www.lektorium.tv/

17. 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. КиберЛенинка http://cyberleninka.ru/;
2. Американская патентная база данных http://www.uspto.gov/patft/
3. Министерство науки и высшего образования Российской Федерации https://www.minobrnauki.gov.ru/;
4. Федеральный портал "Российское образование" http://www.edu.ru/;
5. Информационная система "Единое окно доступа к образовательным ресурсам" http://window.edu.ru/;
6. Единая коллекция цифровых образовательных ресурсов http://school-collection.edu.ru/ .
7. Проект Государственного института русского языка имени А.С. Пушкина "Образование на русском" https://pushkininstitute.ru/;
8. Справочно-информационный портал "Русский язык" http://gramota.ru/;
9. Служба тематических толковых словарей http://www.glossary.ru/;
10. Словари и энциклопедии http://dic.academic.ru/;
11. Образовательный портал "Учеба" http://www.ucheba.com/;
12. Законопроект "Об образовании в Российской Федерации". Вопросы и ответы http://xn--273--84d1f.xn--p1ai/voprosy_i_otvety

Собственные электронные образовательные и информационные ресурсы КубГУ

1. Электронный каталог Научной библиотеки КубГУ http://megapro.kubsu.ru/MegaPro/Web
2. Электронная библиотека трудов ученых КубГУ http://megapro.kubsu.ru/MegaPro/UserEntry?Action=ToDb&idb=6
3. Среда модульного динамического обучения http://moodle.kubsu.ru
4. База учебных планов, учебно-методических комплексов, публикаций и конференций http://infoneeds.kubsu.ru/
5. Библиотека информационных ресурсов кафедры информационных образовательных технологий http://mschool.kubsu.ru;
6. Электронный архив документов КубГУ http://docspace.kubsu.ru/
7. 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)

По дисциплине «Математические модели защиты информации» предусмотрено проведение лекционных занятий, на которых формируется систематизированное представление о математических основах современных методов защиты информации и их применении в системах искусственного интеллекта. В ходе лекций студенты знакомятся с фундаментальными концепциями криптографии (симметричное и асимметричное шифрование, хеш-функции, электронные подписи), математическими моделями угроз и нарушителей, а также специализированными методами защиты AI-систем. Особое внимание уделяется анализу криптографических алгоритмов с точки зрения лежащих в их основе сложных математических задач (дискретное логарифмирование, факторизация, эллиптические кривые). Демонстрируются практические примеры применения математических моделей для защиты данных на различных этапах жизненного цикла AI-

систем — от сбора данных до развертывания моделей. Обсуждаются ключевые аспекты дифференциальной приватности, adversarial атак и методов обеспечения конфиденциальности в машинном обучении. После каждой лекции студентам рекомендуется выполнить расчетные задания — проанализировать математические свойства алгоритмов, решить задачи по модульной арифметике, оценить стойкость криптографических схем.

Лабораторные занятия направлены на формирование практических навыков реализации математических моделей защиты информации с использованием современного программного обеспечения. Студенты последовательно осваивают методы работы с криптографическими библиотеками Python (cryptography, hashlib, pycryptodome), выполняют задания по реализации алгоритмов шифрования, хеширования и аутентификации. Рассматриваются практические аспекты интеграции механизмов защиты в AI-конвейеры: шифрование тренировочных данных, обеспечение целостности моделей, реализация безопасного обмена ключами. В ходе лабораторных занятий студенты также изучают методы применения дифференциальной приватности для защиты конфиденциальности данных, проводят эксперименты по генерации adversarial примеров и исследуют методы защиты от них. Выполняются проекты по комплексной защите AI-систем — от анализа угроз до реализации соответствующих математических моделей защиты. После каждого занятия выдаются задания для самостоятельной доработки — оптимизация реализованных алгоритмов, анализ их производительности, исследование компромиссов между безопасностью и эффективностью.

Самостоятельная работа обучающихся направлена на углублённое изучение теоретических основ и прикладных аспектов математических моделей защиты информации. Рекомендуется регулярно обращаться к фундаментальным учебникам по криптографии, научным статьям по безопасности AI-систем и документации используемых библиотек. Студенты должны научиться анализировать архитектуру AI-систем на предмет уязвимостей, формализовывать требования к защите данных, выбирать и математически обосновывать применение конкретных криптографических алгоритмов. Особое внимание уделяется разработке комплексных решений — от построения моделей угроз до реализации прототипов защищенных подсистем. Умения в области Python-программирования и математического моделирования активно применяются при реализации криптографических примитивов, анализе их стойкости и интеграции в существующие AI-приложения. Самостоятельная работа включает также изучение нормативных требований к защите информации (152-ФЗ, GDPR) и принципов ответственного ИИ, что позволяет студентам формировать комплексное понимание вопросов безопасности в современных AI-системах.

Для успешного освоения дисциплины рекомендуется:

- Активно участвовать в обсуждении материала на лекциях и задавать вопросы
- Систематически выполнять лабораторные работы и своевременно сдавать отчеты
- Самостоятельно изучать дополнительные материалы по темам курса
- Применять полученные знания при работе над курсовыми проектами и выпускными квалификационными работами
- Участвовать в научных семинарах и конференциях, связанных с безопасностью AI-систем.

Итоговой формой освоения курса является зачет.

Для студентов с ограниченными возможностями здоровья предусмотрены индивидуальные консультации и адаптированные материалы. Преподаватель помогает осваивать интерфейсы взаимодействия с ИИ, объясняет ключевые понятия в доступной форме, предоставляет инструкции с альтернативным форматированием. При необходимости используются голосовые интерфейсы, увеличенный масштаб экрана, сопровождение при выполнении заданий. Индивидуальный подход обеспечивает равные

условия участия в образовательном процессе и достижения запланированных результатов обучения.

Рассмотрим примеры кейсов.

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.

А. Применение математических моделей защиты информации в кейсах ПАО «Сбербанк»

Кейс 1. Защита транзакционных данных в API перевода средств

Описание: При осуществлении переводов между счетами необходимо гарантировать конфиденциальность и целостность данных.

Цель: Реализовать механизм электронной подписи для транзакций.

Технологии:

- Асимметричное шифрование: RSA-2048
- Электронная подпись: ГОСТ Р 34.10-2012
- Python библиотеки: cryptography

Реализация:

a) Генерация ключевой пары для каждого пользователя
b) Подписание данных транзакции перед отправкой
c) Проверка подписи на стороне банка

Результат:

- Защита от подмены данных транзакций
- Соответствие стандартам Банка России

Кейс 2 Защита данных в мобильном кошельке СберПей**.**

Описание: Модели рекомендательной системы банка должны быть защищены от несанкционированного изменения.

Цель: Реализовать шифрование локальной базы данных приложения.

Технологии:

- Симметричное шифрование: AES-256-GCM
- Ключевая деривация: PBKDF2
- Python библиотеки: cryptography, sqlite3

Реализация:

a) Шифрование локальной БД приложения
b) Защита ключей шифрования на устройстве
c) Реализация безопасного доступа к данным

Результат:

- Защита данных при утере устройства
- Соответствие PCI DSS стандартам

Кейс 3 Защита паролей клиентов в базе данных.

Описание: В базе данных Сбера хранятся пароли клиентов. Нужно защитить их от утечки.

Цель: Хеширование паролей перед сохранением в БД.

Технологии:

- Хеш-функция SHA-256
- Библиотека hashlib в Python

Реализация:

a) Взять пароль пользователя
b) Посчитать его хеш через SHA-256

c) Сохранить только хеш в базу данных

Результат:

- Даже при утечке базы злоумышленник не узнает настоящие пароли
- Соответствие **базовым требованиям безопасности**

Кейс 4 Шифрование номера карты в смс-уведомлениях.

Описание: В смс-уведомлениях нужно **скрыть часть номера карты**.

Цель: Реализовать простое шифрование для маскирования данных.

Технологии:

- Выбранный алгоритм шифрования
- Python, стандартные библиотеки

Реализация:

a) Взять номер карты "1234567812345678"

b) Зашифровать в виде: "5678********5678"

c) Отправить замаскированный номер в смс

Результат:

- Клиенты видят только часть номера своей карты
- Снижение риска перехвата данных

Кейс 5 Создание цифровых пропусков для сотрудников.

Описание: Нужно сделать **электронные пропуска** для доступа в офис Сбера.

Цель: Реализовать систему электронных подписей.

Технологии:

- ЭЦП на базе RSA
- Python библиотека cryptography

Реализация:

a) Сгенерировать ключи для каждого сотрудника

b) Подписывать электронные пропуска

c) Проверять подпись на входе в офис

Результат:

- Защита от подделки пропусков
- Контроль доступа в помещения

Итог:

Каждый кейс сочетает:

Математическую модель (диффуры, графы, генетические алгоритмы).

Инструментальную реализацию (MATLAB, Maple, Python).

Прикладную пользу для Сбера (экономия денег, снижение рисков).

Б. Применение математических моделей защиты информации в кейсах компании AVA Lab

Кейс 1: Защита данных сенсоров беспилотных автомобилей

Описание: Беспилотные автомобили AVA Lab собирают данные с лидаров и камер. Необходимо защитить эти данные от перехвата и подмены.

Цель: Реализовать **систему шифрования данных сенсоров** в реальном времени.

Технологии:

- Симметричное шифрование: AES-128 в режиме CTR
- Python библиотеки: cryptography, numpy

Реализация:

a) Шифрование потоковых данных с камер и лидаров

b) Проверка целостности данных перед обработкой

c) Защита канала передачи между сенсорами и бортовым компьютером

Результат:

- Защита от подмены данных сенсоров
- Предотвращение атак на систему управления

Кейс 2: Аутентификация обновлений прошивки автопилота

Описание: При обновлении ПО беспилотников необходимо убедиться в подлинности прошивки.

Цель: Реализовать **систему проверки электронной подписи обновлений**.

Технологии:

- Электронная подпись: ECDSA на эллиптических кривых
- Хеш-функции: SHA-256

Реализация:

a) Подписание прошивки перед распространением

b) Проверка подписи на устройстве перед установкой

c) Ведение журнала обновлений

Результат:

- Защита от установки вредоносного ПО
- Гарантия подлинности прошивки

Кейс 3: Защита связи между беспилотником и центром управления

Описание: Беспилотник передает телеметрию в центр управления. Нужно защитить канал связи.

Цель: Реализовать **безопасный обмен ключами** между машиной и сервером.

Технологии:

- Протокол Диффи-Хеллмана
- Python библиотеки: cryptography, requests

Реализация:

a) Генерация ключей на стороне автомобиля и сервера

b) Обмен открытыми ключами

c) Шифрование данных общим ключом

Результат:

- Защита телеметрии от перехвата
- Конфиденциальность передаваемых данных

Кейс 4: Защита данных испытаний беспилотников

Описание: Результаты испытаний содержат коммерческую тайну и ноу-хау AVA Lab.

Цель: Реализовать систему **шифрования базы данных испытаний**.

Технологии:

- Шифрование на уровне базы данных
- Python библиотеки: cryptography, sqlite3

Реализация:

a) Шифрование чувствительных полей в БД

b) Разграничение доступа к данным испытаний

c) Ведение аудита доступа

Результат:

- Защита интеллектуальной собственности
- Контроль доступа к критическим данным

Кейс 5: Обнаружение аномалий в работе автопилота

Описание: Необходимо выявлять попытки вмешательства в работу системы управления.

Цель: Реализовать **систему мониторинга аномального поведения**.

Технологии:

- Статистический анализ параметров системы
- Python библиотеки: numpy, scikit-learn

Реализация:

a) Мониторинг критических параметров автопилота

b) Обнаружение отклонений от нормального поведения

c) Автоматическое оповещение о подозрительной активности

Результат:

- Раннее обнаружение кибератак
- Повышение устойчивости системы управления

Итог

Каждый кейс решает задачи информационной безопасности для AVA Lab:

- Защита данных;
- Аутентификация обновлений;
- Контроль целостности.

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

7.1 Перечень информационно-коммуникационных технологий

Инструменты и библиотеки: используются современные криптографические библиотеки и фреймворки для защиты информации. В частности, cryptography, pycryptodome и hashlib для реализации алгоритмов симметричного и асимметричного шифрования, хеширования и электронной подписи. Для обеспечения дифференциальной приватности в машинном обучении применяются библиотеки diffprivlib, tensorflow-privacy и opacus. Для тестирования уязвимостей AI-систем используются adversarial-robustness-toolbox и cleverhans. Инструментарий также включает Jupyter Notebook и Google Colab для интерактивных экспериментов, а для оформления отчетов – MS Word или LibreOffice.

Исходные данные: для экспериментов и заданий используются учебные датасеты и криптографические задачи. Это могут быть стандартные наборы данных MNIST, CIFAR-10 для тестирования защиты моделей компьютерного зрения, синтетические данные для проверки методов дифференциальной приватности, примеры уязвимых AI-систем для анализа угроз безопасности. Данные подбираются преподавателем с учетом изучаемых математических моделей защиты информации.

Программное обеспечение и ИКТ: необходим доступ к интернету и облачным платформам для развертывания защищенных решений. Рекомендуется использовать среды разработки Jupyter Notebook или Google Colab, а также профессиональные код-редакторы (Visual Studio Code, PyCharm). Для хранения и совместной работы со студентами используется учебная информационная система Moodle. Для анализа результатов и визуализации применяются Excel, Google Sheets и специализированные библиотеки Python.

1. Облачные платформы и сервисы

cloud.ru,YandexCloud,AWS/GCP/Azure– облачные вычисления

2. Системы управления версиями и коллаборации

Git/GitHub/GitLab – контроль версий кода и совместная разработка

4. Система управления обучением

Moodle – сдача работ

7.2 Перечень лицензионного и свободно распространяемого программного обеспечения

1. Лицензионное ПО
VSCode – IDE для Python (свободнораспространяемое)
LibreOffice– оформление отчетов (свободнораспространяемое)

2. Свободное ПО (Open Source)
OpenSSL, GnuPG, Libgcrypt - криптографический инструментарии
GitLab, GIT, MLFlow, Docker, Kubernetes, Terraform.Фреймворки для ML:
PyTorch/TensorFlow – разработка нейросетей
scikit-learn – классические алгоритмы ML

Инструменты для визуализации:

Streamlit/Gradio – создание веб-интерфейсов для моделей
Matplotlib/Seaborn – графики и анализ данных

СУБД:

SQLite/PostgreSQL – хранение структурированных данных
FAISS/Annoy – векторный поиск

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	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.