АННОТАЦИЯ рабочей программы дисциплины

Б1.В.02 Математические модели защиты информации

Объем трудоемкости: 2 з.е.

Цели 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.

Задачи дисциплины: в соответствии с поставленной целью состоят в следующем:

- Сформировать у студентов понимание фундаментальных математических моделей, используемых в криптографии.
- Научить формализовывать угрозы информационной безопасности для данных и алгоритмов в системах аналитики и ИИ, используя математический аппарат.
- 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.
- Освоить на практическом уровне методы реализации базовых криптографических примитивов и протоколов.
- 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.

Место дисциплины в структуре ООП ВО

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.

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:

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.
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	**(П) Знать** Виды угроз на этапах ЖЦ ИИ-систем, методы моделирования угроз с помощью ИИ, способы оценки и приоритизации рисков. **Уметь** Моделировать сценарии угроз ИИ-систем, оценивать и ранжировать риски, интегрировать риск-метрики в бизнес-показатели. **Владеть** Проектированием процессов моделирования угроз средствами ИИ, разработкой новых сценариев анализа угроз
AI S - 1.2 Обеспечивает соответствие нормативным требованиям и принципам доверенного/этичного ИИ.	**Уровень Б (базовый)** **Знать:** Кодекс этики в сфере ИИ РФ (2021); базовые принципы Responsible AI; закон 152-ФЗ «О персональных данных»; основы GDPR. **Уметь:** описать процесс Data Impact Assessment (оценки воздействия на данные). **Владеть:** знанием базовых этических и нормативных принципов в области ИИ и персональных данных. **Уровень С (средний)** **Знать:** требования NIST AI RMF; основы подготовки документов для Роскомнадзора по 152-ФЗ и для ФСТЭК по классификации ИСПДн. **Уметь:** проводить DPIA / AI Impact Assessment по NIST AI RMF; подготавливать пакеты документов для Роскомнадзора (152-ФЗ) и ФСТЭК (классификация ИСПДн); участвовать в решении конфликтов между этикой, безопасностью и бизнесом. **Владеть:** проведением DPIA/AI Impact Assessment; подготовкой пакетов документов для надзорных органов; навыками участия в конфликтных ситуациях на стыке этики, безопасности и бизнеса. **Уровень П (продвинутый)** **Знать:** EU AI Act; NIST AI RMF; отечественные акты ПНСТ 836-2023 и ГОСТ Р 71476-2024; процедуры сертификации ФСТЭК и ЦБ РФ.

	Уметь: формировать корпоративную политику «Responsible & Secure AI» на стыке международных (EU AI Act, NIST AI RMF) и российских нормативных актов (ПНСТ 836-2023, ГОСТ Р 71476-2024); консультировать при сертификации ФСТЭК и ЦБ РФ. **Владеть:** формированием корпоративной политики в области ответственного и безопасного ИИ; консультированием по вопросам сертификации; участием в работе ТК 164 по разработке новых ГОСТов.
ML-6 Способен применять алгоритмы обучения с подкреплением	
ML-6.1 Обосновывает способы и варианты применения алгоритмов обучения с подкреплением в задачах ИИ, включая их преобразование и адаптацию к специфике задачи	**Уровень Б (базовый)** **Знать:** основные принципы обучения с подкреплением: агент, среда, действие, награда, политика; простейшие алгоритмы Q-обучение и SARSA. **Уметь:** обосновывать выбор Q-обучения или SARSA для решения типовых задач ИИ, включая задачи защиты информации. **Владеть:** описанием базовых принципов RL и способностью применять простейшие алгоритмы для типовых сценариев. **Уровень С (средний)** **Знать:** методы временных разностей, методы Монте-Карло, λ-доход; аппроксимацию функции ценности агента, в том числе с помощью стратегии. **Уметь:** разрабатывать адаптивного агента; задавать цель агента с помощью полного вознаграждения, вознаграждения с обесценением или λ-дохода; применять методы временных разностей и методы Монте-Карло для обучения агента. **Владеть:** применением методов временных разностей и Монте-Карло; проведением аппроксимации функции ценности агента. **Уровень П (продвинутый)** **Знать:** архитектуры DQN, DDPG, A3C, TRPO, PPO, SAC; принципы построения AlphaGo, AlphaZero, MuZero. **Уметь:** применять алгоритмы обучения на основе временных разностей и Q-обучения; разрабатывать процедуру планирования поведения агента с помощью поиска по дереву Монте-Карло (MCTS); адаптировать архитектуры DQN, DDPG, PPO, SAC под конкретную задачу защиты информации. **Владеть:** применением алгоритмов TD и Q-обучения; поиском по дереву Монте-Карло (MCTS); адаптацией продвинутых архитектур RL под модель нарушителя и среду защиты.

Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

№	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.	15,8	2		2	11,8
2.	Математические модели симметричного шифрования.	10	2		2	6
3.	Математические модели хеширования и аутентификации.	8	2		2	4

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
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**				

Курсовые работы: не предусмотрены.

Форма проведения аттестации по дисциплине: зачет.

Автор: В.О. Осипян, профессор кафедры, д-р. ф.-м. н., доцент