Федеральное государственное бюджетное образовательное учреждение
высшего образования
Московский государственный университет имени М.В. Ломоносова
Факультет вычислительной математики и кибернетики

УТВЕРЖДАЮ
декан факультета вычислительной
математики и кибернетики

/И.А. Соколов /
«27» сентября 2023г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине

Безопасность систем искусственного интеллекта

Уровень высшего образования:

бакалавриат

Направление подготовки / специальность:

02.03.02 "Фундаментальная информатика и информационные технологии" (3++)

Направленность (профиль) ОПОП:

Искусственный интеллект и анализ данных

Форма обучения:

очная

Рассмотрен и утвержден
на заседании Ученого совета факультета ВМК
(протокол №7, от 27 сентября 2023 года)

Москва 2023

1. ФОРМЫ И ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

В процессе и по завершении изучения дисциплины оценивается формирование у студентов следующих компетенций:

Планируемые результаты обучения по дисциплине (модулю)		
Содержание и код компетенции.	**Индикатор (показатель) достижения компетенции**	**Планируемые результаты обучения по дисциплине, сопряженные с индикаторами достижения компетенций**
ПК-3. Способен осуществлять концептуальное моделирование проблемной области и проводить формализацию представления знаний в системах искусственного интеллекта	ПК-3.1. Разрабатывает концептуальную модель проблемной области системы искусственного интеллекта ПК-3.2. Выбирает методы представления знаний и проектирует базу знаний системы искусственного интеллекта	ПК-3.1. 3-1. Знает методы концептуального моделирования в аспектах построения объектных, функциональных и поведенческих моделей проблемной области ПК-3.1. 3-2. Знает методы построения онтологий в виде таксономий объектов, установления семантических отношений и определения аксиоматики формирования классов объектов ПК-3.1. У-1. Умеет применять методы концептуального моделирования проблемной области в аспектах построения объектных, функциональных и поведенческих моделей проблемной области ПК-3.1. У-2. Умеет отображать концептуальные модели проблемной области с помощью инструментальных средств построения онтологий и выполнять запросы и навигацию по структуре онтологии ПК-3.2. 3-1. Знает методы представления знаний, основанные на отображении объектного, функционального (процедурного) и поведенческого видов знаний, и критерии их выбора ПК-3.2. 3-2. Знает методы проектирования базы знаний с использованием различных классов методов представления знаний ПК-3.2. У-1. Умеет выбирать

		методы представления знаний в зависимости от класса решаемых задач ПК-3.2. У-2. Умеет проектировать базу знаний с использованием различных классов методов представления знаний

1.1. Текущий контроль успеваемости

Текущий контроль успеваемости осуществляется путем оценки результатов выполнения заданий практических (семинарских) занятий, самостоятельной работы, предусмотренных учебным планом и посещения занятий/активность на занятиях.

В качестве оценочных средств текущего контроля успеваемости предусмотрены:

Примерные практические задания

- **Задание 1.** Разработка порождающей модели МО для генерации изображения лица целевой персоны, позволяющий нарушить работу биометрического классификатора пользователей по лицу. Биометрический классификатор будет предоставлен.
- **Задание 2.** Разработка порождающей модели МО для генерации голоса целевого диктора, позволяющий нарушить работу биометрического классификатора дикторов по голосу. Биометрический классификатор будет предоставлен.
- **Задание 3**. Реализация пула состязательных атак, позволяющих нарушить работу биометрического классификатора пользователей по лицу. Биометрический классификатор будет предоставлен.
- **Задание 4.** Реализация бинарного классификатора синтетических данных, позволяющий идентифицировать такого сорта данные и тем самым защитить модель биометрической классификации лиц. Защита должна эффективно работать от атак, разработанных командой в рамках задания 1. Биометрический классификатор будет предоставлен.
- **Задание 5.** Реализация механизма состязательного обучения, позволяющего защитить модель биометрической классификации лиц. Защита должна эффективно работать от атак, разработанных командой в рамках задания 3. Биометрический классификатор будет предоставлен.

1.2. Промежуточная аттестация

Промежуточная аттестация осуществляется в форме зачет
В качестве средств, используемых на промежуточной аттестации предусматривается:
Билеты

1.3. Типовые задания для проведения промежуточной аттестации

Список вопросов

1. Методологические основы комплексной системы защиты информации систем искусственного интеллекта.
2. Определение состава защищаемой информации.
3. Источники, способы и результаты дестабилизирующего воздействия на информацию.
4. Каналы и методы несанкционированного доступа к информации.
5. Моделирование процессов комплексной системы защиты информации.
6. Нормативно-методическое обеспечение систем защиты информации.
7. Управление комплексной системой защиты информации.
8. Подходы к созданию состязательных примеров.
9. Атаки отравление.
10. Атаки уклонением.
11. Атаки извлечением.
12. Атаки с применением порождающих моделей.

2. КРИТЕРИИ ОЦЕНКИ ПО ДИСЦИПЛИНЕ

ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ результатов обучения (РО) по дисциплине				
Оценка виды оценочных средств	**2 (не зачтено)**	**3 (зачтено)**	**4 (зачтено)**	**5 (зачтено)**
Знания **(виды оценочных средств: приведены в п. 1.2.)**	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Умения **(виды оценочных средств: приведены в п. 1.2.)**	Отсутствие умений	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности непринципиального характера)	Успешное и систематическое умение
Навыки **(владения, опыт деятельности)** **(виды оценочных средств: приведены в п. 1.2..)**	Отсутствие навыков (владений, опыта)	Наличие отдельных навыков (наличие фрагментарного опыта)	В целом, сформированные навыки (владения), но используемые не в активной форме	Сформированные навыки (владения), применяемые при решении задач