

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе и
образовательным инновациям

О.Н. Здрок

« 10 октября 2020 г.

Регистрационный № УД-8607/уч.



НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ В ОБРАБОТКЕ И ЗАЩИТЕ ДАННЫХ

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности:**

1-98 01 01 Компьютерная безопасность (по направлениям)

направление специальности:

1-98 01 01-02 Компьютерная безопасность
(радиофизические методы и программно-технические средства)

2020 г.

Учебная программа составлена на основе ОСВО 1 98 01 01 – 2013 и учебного плана УВО Р98-139/уч. и Р98и-140/уч. от 30.05.2013 г.

СОСТАВИТЕЛИ:

Василий Сергеевич САДОВ, профессор кафедры интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский государственный университет, кандидат технических наук, доцент;

Екатерина Александровна ГОЛОВАТАЯ, старший преподаватель кафедры интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский государственный университет;

Александр Васильевич КУРОЧКИН, старший преподаватель кафедры интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский государственный университет;

Александр Иванович ГОЛОВАТЫЙ, доцент кафедры интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский государственный университет, кандидат физико-математических наук.

РЕЦЕНЗЕНТЫ:

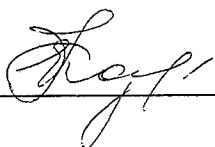
Ирина Петровна МАТВЕЕНКО, доцент кафедры автоматизированных систем управления производством, агроэнергетический факультет, Белорусский государственный аграрный технический университет, кандидат технических наук, доцент.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой интеллектуальных систем факультета радиофизики и компьютерных технологий Белорусского государственного университета (протокол № 1 от 28.08.2020);

Научно-методическим Советом БГУ
(протокол № 1 от 28.09.2020)

Заведующий кафедрой
интеллектуальных систем



Е.И. Козлова

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цели и задачи учебной дисциплины

Цель учебной дисциплины – изучение методов обеспечения безопасности информационных систем при помощи алгоритмов машинного обучения.

Задачи учебной дисциплины:

1. Рассмотреть основные проблемные области и принципы организации данных в системах информационной безопасности;
2. Изучить алгоритмы и методы интеллектуального анализа данных в системах обеспечения информационной безопасности;
3. Приобрести практические навыки проектирования систем интеллектуального анализа и принятия решений в области информационной безопасности.

Место учебной дисциплины в системе подготовки специалиста с высшим образованием.

Учебная дисциплина относится к циклу дисциплин специализации компонента учреждения высшего образования.

Для успешного освоения данной учебной дисциплины необходимы знания по дисциплинам «Теория вероятностей и математическая статистика», «Программирование» и «Основы информационной безопасности».

Требования к компетенциям

Освоение учебной дисциплины «Нейросетевые технологии в обработке и защите данных» должно обеспечить формирование следующих академических, социально-личностных и профессиональных компетенций:

академические компетенции:

- АК-1. Уметь применять базовые научно-теоретические знания для решения теоретических и практических задач.
- АК-2. Владеть системным и сравнительным анализом.
- АК-7. Иметь навыки, связанные с использованием технических устройств, управлением информацией и работой с компьютером.

социально-личностные компетенции:

- СЛК-3. Обладать способностью к межличностным коммуникациям.
- СЛК-6. Уметь работать в команде.

профессиональные компетенции:

- ПК-3. Разрабатывать модели явлений, процессов или систем при организации защиты информации.
- ПК-4. Выбирать необходимые методы исследования, модифицировать существующие, разрабатывать новые методы и применять их для решения поставленных задач при организации защиты информации;
- ПК-5. Выполнять оценку эффективности методов защиты информации;
- ПК-9. Анализировать и оценивать собранные данные;
- ПК-15. Организовывать процесс создания, оценки и эксплуатации средств и систем защиты информации, поддерживать и повышать их безопасность;

осуществлять контроль за их использованием.

В результате освоения учебной дисциплины студент должен:

знать:

- проблемные области информационной безопасности, в которых применимы методы интеллектуального анализа;
- особенности организации данных в системах анализа и обеспечения информационной безопасности;
- основные подходы к статистическому и интеллектуальному анализу набора данных в системах информационной безопасности;

уметь:

- проводить предварительную обработку исходных данных систем сетевого мониторинга, сбора биометрической информации, эвристического анализа исполняемых файлов и интернет-ресурсов с помощью методов извлечения признаков и понижения размерности;
- строить модели принятия решений и анализа пользовательских данных в системах информационной безопасности;
- производить оценку корректности различных моделей классификации, регрессии и кластеризации;

владеть:

- современными инструментами мониторинга, аудита и сбора информации в системах обеспечения безопасности;
- методами и принципами проектирования архитектур систем информационной безопасности, включающих модули интеллектуального анализа данных.

Структура учебной дисциплины

Дисциплина изучается в 6 семестре. Всего на изучение учебной дисциплины «Нейросетевые технологии в обработке и защите данных» отведено:

– для очной формы получения высшего образования – 136 часов, в том числе 62 аудиторных часа, из них: лекции – 34 часов, лабораторные занятия – 28 часов.

Трудоемкость учебной дисциплины составляет 3,5 зачетные единицы.

Форма текущей аттестации–зачет.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Представление данных в системах информационной и компьютерной безопасности

Тема 1.1. Введение. Определение, типы и задачи машинного обучения. Постановка задачи в матричном виде. Представление о датасете.

Тема 1.2. Сбор данных и генерация признаков, формирование датасета. Текстовая информация. Работа с изображениями. Предварительная обработка датасета.

Тема 1.3. Потенциальные угрозы систем информационной безопасности. Уязвимости ресурсов локальных и глобальных сетей. Вредоносное программное обеспечение. Вредоносная текстовая информация. Биометрические системы контроля доступа.

Раздел 2. Модели принятия решений

Тема 2.1. Линейная регрессия. Функция потерь и оптимизация. Алгоритм градиентного спуска. Логистическая регрессия. Задачи классификации.

Тема 2.2. Деревья и леса принятия решений. Метод k ближайших соседей. Метод опорных векторов.

Тема 2.3. Байесовская теория вероятности. Оценка результатов классификации.

Раздел 3. Нейронные сети

Тема 3.1. Формальный нейрон. Функции принадлежности. Многослойный персептрон прямого распространения.

Тема 3.2. Обучение нейронной сети на основе алгоритмов градиентного спуска. Модификации метода градиентного спуска.

Тема 3.3. Метод Ньютона. Квазиньютоновские методы. Метод Бroyдена-Флетчера-Гольдфарба-Шанно. Метод Гаусса-Ньютона. Метод Левенберга-Марквардта. Метод сопряженных градиентов.

Тема 3.4. Анализ работы нейронных сетей.

Тема 3.5. Ограниченные машины Больцмана. Остаточные нейронные сети. Рекуррентные сети. Сети с долгой краткосрочной памятью.

Тема 3.6. Сверточные и рекуррентные сверточные сети. Автоэнкодер. Генеративные состязательные сети. Сиамские нейронные сети.

Раздел 4. Обучение без учителя

Тема 4.1. Методы преобразования/понижения размерности данных. Метод главных компонент. Метод независимых компонент. Многомерное шкалирование. Сингулярное векторное разложение.

Тема 4.2. Кластеризация. Алгоритм k-means. Алгоритм x-means. Самоорганизующиеся карты. Выявление аномалий. Изоляционные леса. Локальный уровень выброса. Одноклассовый метод опорных векторов.

Тема 4.3. Оценка результатов обучения без учителя. Индекс Рэнда и индекс Фолкеса-Мэллоуса. Байесовский информационный критерий. Оценки кластеризации по известной классификации. Кросс-скоринг.

Раздел 5. Нейро-нечеткое моделирование

Тема 5.1. Нечеткие множества как алгебраические структуры. t-норма и t-конорма.

Тема 5.2. Системы нечеткой логики. Лингвистические переменные. Термы и функции принадлежности. Правила. Нечеткий логический вывод. Нейро-нечеткие гибридные системы.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Дневная форма получения образования

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Количество часов УСП	Форма контроля знаний
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	2	3	4	5	6	7	8	9
1	Представление данных в системах информационной и компьютерной безопасности	6			4			
1.1	Введение. Определение, типы и задачи машинного обучения. Постановка задачи в матричном виде. Представление о датасете.	2						выборочный опрос на лекции
1.2	Сбор данных и генерация признаков, формирование датасета. Текстовая информация. Работа с изображениями. Предварительная обработка датасета.	2			2			выборочный опрос на лекции
1.3	Потенциальные угрозы систем информационной безопасности. Уязвимости ресурсов локальных и глобальных сетей. Вредоносное программное обеспечение. Вредоносная текстовая информация. Биометрические системы контроля доступа.	2			2			тест
2	Модели принятия решений	6			6			
2.1	Линейная регрессия. Функция потерь и оптимизация. Алгоритм градиентного спуска. Логистическая регрессия. Задачи классификации.	2			2			выборочный опрос на лекции
2.2	Деревья и леса принятия решений. Метод k ближайших соседей. Метод опорных векторов.	2			2			выборочный опрос на лекции
2.3	Байесовская теория вероятности. Оценка результатов классификации.	2			2			отчёт по лабораторной работе
3	Нейронные сети	12			12			

3.1	Формальный нейрон. Функции принадлежности. Многослойный персептрон прямого распространения.	2			2		отчёт по лабораторной работе
3.2	Обучение нейронной сети на основе алгоритмов градиентного спуска. Модификации метода градиентного спуска.	2			2		выборочный опрос на лекции
3.3	Метод Ньютона. Квазиньютоновские методы. Метод Бройдена-Флетчера-Гольдфарба-Шанно. Метод Гаусса-Ньютона. Метод Левенберга-Марквардта. Метод сопряженных градиентов.	2			2		тест
3.4	Анализ работы нейронных сетей.	2			2		выборочный опрос на лекции
3.5	Ограниченные машины Больцмана. Остаточные нейронные сети. Рекуррентные сети. Сети с долгой краткосрочной памятью.	2			2		отчёт по лабораторной работе
3.6	Сверточные и рекуррентные сверточные сети. Автоэнкодер. Генеративные состязательные сети. Сиамские нейронные сети.	2			2		отчёт по лабораторной работе, тест
4	Обучение без учителя	6			4		
4.1	Методы преобразования/понижения размерности данных. Метод главных компонент. Метод независимых компонент. Многомерное шкалирование. Сингулярное векторное разложение.	2			2		выборочный опрос на лекции
4.2	Кластеризация. Алгоритм k-means. Алгоритм x-means. Самоорганизующиеся карты. Выявление аномалий. Изоляционные леса. Локальный уровень выброса. Одноклассовый метод опорных векторов.	2			2		отчёт по лабораторной работе
4.3	Оценка результатов обучения без учителя. Индекс Рэнда и индекс Фолкеса-Мэллоуса. Байесовский информационный критерий. Оценки кластеризации по известной классификации. Кросс-скоринг.	2					выборочный опрос на лекции
5	Нейро-нечеткое моделирование	4			4		
5.1	Нечеткие множества как алгебраические структуры. t-норма и t-конорма.	2					выборочный опрос на лекции
5.2	Системы нечеткой логики. Лингвистические переменные. Термы и функции принадлежности. Правила. Нечеткий логический вывод. Нейро-нечеткие гибридные системы.	2			4		отчёт по лабораторной работе
	Всего	34			28		

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Чيو, К. Машинное обучение и безопасность / К. Чيو, Д. Фримэн // пер. с англ. А. В. Снастина. – М. : ДМК Пресс, 2020. – 338 с.
2. Рашид, Т. Создаём нейронную сеть. / Т. Рашид. – М: Вильямс, 2017. – 272 с.
3. Николенко, С. Глубокое обучение / С. Николенко, А. Кадури, Е. Архангельская. – СПб: Питер Мейл, 2018. – 480 с.
4. Бирюков, А. А. Информационная безопасность: защита и нападение / А. А. Бирюков. – М. : ДМК Пресс, 2017. – 434 с.
5. Bishop, C. M., Pattern recognition and machine learning / C. M. Bishop. – New York: Springer, 2016. – 738 p.
6. Weaver, K.F. An Introduction to Statistical Analysis in Research: With Applications in the Biological and Life Sciences / K.F. Weaver, V.C. Morales, S.L. Dunn, K. Godde, P.F. Weaver. // Wiley, 2017. – 580 с.

Перечень дополнительной литературы

1. Плас, Дж. В. Python для сложных задач: наука о данных и машинное обучение. – СПб: Питер Мейл, 2018. – 576 с.
2. Грас, Дж. Data science. Наука о данных с нуля. (пер. с англ.) – СПб: БХВ-Петербург, 2018. – 336 с.
3. Стефанова, И.А. Обработка данных и компьютерное моделирование. Учебное пособие / И.А. Стефанова. // Лань : Санкт-Петербург, 2020. – 112 p.
4. Lista L. Statistical Methods for Data Analysis in Particle Physics (Lecture Notes in Physics (941)) / L. Lista. // Springer, 2017. – 273 с.
5. Ghahramani, Z. Unsupervised Learning / Z. Ghahramani. – Gatsby Computational Neuroscience Unit University College London, 2004. – 32 p.
6. Barber, D. Bayesian Reasoning and Machine Learning. – D. Barber, University College London, 2012. – 735 p.
7. Tanenbaum, A. S. Computer Networks / A. S. Tanenbaum, D. J. Wetherall // 5th ed. – Pearson, 2012. – 960 p.

Перечень рекомендуемых средств диагностики и методика формирования итоговой оценки

Опрос на лекциях проводится в форме учебной дискуссии.

Задания к лабораторным работам формируются в соответствии с методом анализа конкретных ситуаций (кейс-методом). При оценке защиты лабораторной работы определяющими факторами являются корректность выполнения заданий, ответы на контрольные вопросы, владение соответствующим теоретическим материалом.

Формой итоговой аттестации по дисциплине «Нейросетевые технологии в обработке и защите данных» учебным планом предусмотрен зачет.

При формировании итоговой оценки используется рейтинговая оценка знаний студента, дающая возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая оценка предусматривает использование весовых коэффициентов для текущего контроля знаний и текущей аттестации студентов по дисциплине.

Формирование оценки за текущую успеваемость:

- тесты – 30 %;
- защита лабораторных работ (отчет) – 70 %;

Рейтинговая оценка по дисциплине рассчитывается на основе оценки текущей успеваемости и зачетной оценки с учетом их весовых коэффициентов. Вес оценки по текущей успеваемости составляет 40 %, зачетной оценки – 60 %.

Примерная тематика лабораторных занятий

Лабораторная работа № 1.

Бинарная классификация в задачах информационной и компьютерной безопасности при обнаружении угроз и атак.

Лабораторная работа № 2.

Обработка естественного языка средствами машинного при обнаружении вредоносных сообщений.

Лабораторная работа № 3.

Прогнозирование количества посещений интернет-ресурса на основе методов анализа временных рядов.

Лабораторная работа № 4.

Обработка биометрической информации в системах идентификации и аутентификации пользователей.

Лабораторная работа № 5.

Анализ данных банковских транзакций на основе методов обучения без учителя и поиска аномалий.

Лабораторная работа № 6.

Балансировка нагрузки на серверах на основе нечеткой логики.

Описание инновационных подходов и методов к преподаванию учебной дисциплины

При организации образовательного процесса используется практико-ориентированный подход, который предполагает:

- освоение содержания образования через решение практических задач;
- ориентацию на генерирование идей, реализацию групповых студенческих проектов;
- использование процедур, способов оценивания, фиксирующих сформированность профессиональных компетенций.

При проведении выборочных опросов на лекциях используется метод учебной дискуссии, который предполагает участие студентов в целенаправленном обмене мнениями, идеями для предъявления и/или согласования существующих позиций по определенной проблеме. Использование метода обеспечивает появление нового уровня понимания изучаемой темы, применение знаний (теорий, концепций) при решении проблем, определение способов их решения.

При проведении лабораторных занятий рекомендуется использовать метод анализа конкретных ситуаций (кейс-метод), который предполагает приобретение студентом знаний и умений для решения практических задач, а также анализ ситуации, используя профессиональные знания, собственный опыт, дополнительную литературу и иные источники.

Методические рекомендации по организации самостоятельной работы обучающихся

При изучении учебной дисциплины рекомендуется использовать следующие формы самостоятельной работы:

- поиск (подбор) и обзор литературы и электронных источников по индивидуально заданной проблеме курса;
- выполнение домашнего задания;
- работы, предусматривающие решение задач и выполнение упражнений, выдаваемых на практических занятиях;
- изучение материала, вынесенного на самостоятельную проработку;
- анализ статистических и фактических материалов по заданной теме, проведение расчетов, составление схем и моделей на основе статистических материалов;
- подготовка к участию в конференциях и конкурсах.

Примерный перечень вопросов к зачету

1. Машинное обучение. Решаемые задачи. Обучение с учителем и без учителя.
2. Машинное обучение с учителем. Понятие модели, параметров и гиперпараметров.
3. Сбор данных и генерация признаков, формирование датасета. Текстовая информация.
4. Сбор данных и генерация признаков, формирование датасета. Работа с изображениями.
5. Сбор данных и генерация признаков, формирование датасета. Предварительная обработка датасета.
6. Потенциальные угрозы систем информационной безопасности. Уязвимости ресурсов локальных сетей.
7. Потенциальные угрозы систем информационной безопасности. Уязвимости ресурсов глобальных сетей.
8. Потенциальные угрозы систем информационной безопасности. Вредоносное программное обеспечение.
9. Потенциальные угрозы систем информационной безопасности. Вредоносная текстовая информация.
10. Потенциальные угрозы систем информационной безопасности. Биометрические системы контроля доступа.
11. Модели принятия решений. Линейная регрессия. Гипотеза и функция стоимости.
Поиск минимума функции стоимости.
Модели принятия решений. Логистическая регрессия. Гипотеза и функция стоимости. Поиск минимума функции стоимости.
12. Модели принятия решений. Деревья и леса принятия решений.
13. Модели принятия решений. Метод k ближайших соседей.
14. Модели принятия решений. Метод опорных векторов.
15. Модели принятия решений. Байесовская теория вероятности. Вероятностная интерпретация машинного обучения. Наивный байесовский классификатор
16. Оценка результатов классификации. Проблема несбалансированной выборки.
Метрики качества модели при несбалансированной выборке.
17. Оценка результатов классификации. Проблема несбалансированной выборки.
ROC-анализ.
18. Нейронные сети. Формальный нейрон и функции активации. Бинарная классификация.
19. Нейронные сети. Множественная классификация. Принцип «один против всех».
20. Нейронные сети. Обучение. Алгоритм градиентного спуска в нейронной сети.
21. Нейронные сети. Модификации алгоритма градиентного спуска.
22. Нейронные сети. Квазиньютоновские методы обучения нейронных сетей.
23. Нейронные сети. Метод сопряженных градиентов.
24. Нейронные сети. Регуляризация параметров.
25. Нейронные сети. Диагностика и анализ работы сети, подбор параметров.

Недообучение и переобучение.

26. Нейронные сети. Архитектуры нейронных сетей. Рекуррентные нейронные сети и сети с долгой краткосрочной памятью.
27. Нейронные сети. Архитектуры нейронных сетей. Генеративные состязательные сети.
28. Нейронные сети. Архитектуры нейронных сетей. Сверточные нейронные сети.
29. Нейронные сети. Архитектуры нейронных сетей. Автоэнкодеры и сверточные автоэнкодеры.
30. Нейронные сети. Архитектуры нейронных сетей. Сиамские нейронные сети.
31. Методы снижения размерности. Метод главных компонент.
32. Методы снижения размерности. Метод независимых компонент.
33. Методы снижения размерности. Сингулярное векторное разложение.
34. Кластерный анализ. Алгоритм k-means. Основные подходы к определению количества кластеров.
35. Кластерный анализ. Алгоритм X-means.
36. Поиск аномалий.
37. Обучение без учителя. Самоорганизующиеся карты.
38. Обучение без учителя. Оценка результатов обучения без учителя.
39. Нечеткая логика. Нечеткие множества как алгебраические структуры.
40. Нечеткая логика. Операции t-норма и t-конорма. Возможности введения операций пересечения, объединения и дополнения на нечетких множествах.
41. Нечеткая логика. Лингвистические переменные и термы.
42. Нечеткая логика. Фаззификация и дефаззификация.
43. Нечеткая логика. Нечеткие правила и изоморфизм логических операций.
44. Нечеткая логика. Нечеткий вывод. Алгоритм Мамдани.
45. Нечеткая логика. Нечеткий вывод. Алгоритм Такаги-Сугено-Канга.
46. Гибридные сети. Нечеткие нейронные сети.

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО
ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
_____ (протокол № ____ от _____ 201_ г.)

Заведующий кафедрой

УТВЕРЖДАЮ

Декан факультета
