

**Белорусский государственный университет**

**УТВЕРЖДАЮ**

Проректор по учебной работе и  
образовательным инновациям

\_\_\_\_\_ И. Чуприс  
(И.О.Фамилия)

27.06.2018  
(дата утверждения)

Регистрационный № 1-0239 /уч.



**ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ**

Учебная программа учреждения высшего образования  
по учебной дисциплине для специальности:

1-98 80 03 Аппаратное и программно-техническое обеспечение  
информационной безопасности

2018 г.

Учебная программа составлена на основе образовательного стандарта ОСВО 1-98 80 03-2012 и учебного плана № Р98-286/уч. от 26.05.2017

**СОСТАВИТЕЛИ:**

Елена Ивановна КОЗЛОВА, доцент кафедры интеллектуальных систем Белорусского государственного университета, кандидат физико-математических наук, доцент

Екатерина Александровна ГОЛОВАТАЯ, ассистент кафедры интеллектуальных систем Белорусского государственного университета

Александр Васильевич КУРОЧКИН, ассистент кафедры интеллектуальных систем Белорусского государственного университета

**РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:**

Кафедрой интеллектуальных систем

(протокол № 12 от 06.06.2018 г.);

Учебно-методической комиссией факультета радиофизики и компьютерных технологий Белорусского государственного университета

(протокол № 10 от 19.06.2018 г.)

## ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Учебная программа дисциплины по выбору специальной подготовки «Интеллектуальные технологии обработки данных» разработана в соответствии с требованиями образовательного стандарта по специальности 1-98 80 03 «Аппаратное и программно-техническое обеспечение информационной безопасности». Программа предназначена для магистрантов дневной формы получения высшего образования.

**Целью изучения** данной учебной дисциплины является освоение современных подходов к интеллектуальной обработке и систематизации информации из различных источников с использованием методов машинного обучения и аппарата нечеткой логики.

**Основная задача дисциплины** – сформировать представление о задачах обработки данных, решаемых с помощью интеллектуальных технологий; рассмотреть основные области применения; представить и научить применять подходы и методы для практической реализации систем машинного и глубокого обучения, экспертных систем и систем поддержки принятия решений.

Для успешного усвоения данной учебной дисциплины необходимы знания по дисциплинам «Аналитическая геометрия и линейная алгебра», «Теория вероятностей и математическая статистика», «Численные методы», «Программирование» в объеме программы высшей школы.

В результате изучения дисциплины обучаемый должен **знать:**

- основные области применения методов машинного обучения с учителем и без учителя;
- теоретические основы работы современных подходов и методов машинного обучения;
- архитектурные особенности построения экспертных систем на основе нечеткого вывода;

**уметь:**

- применять подходы построения систем машинного обучения для решения задач прогнозирования, классификации и кластеризации, распознавания образов;
- проектировать экспертные системы и системы поддержки принятия решений с использованием аппарата нечеткого вывода;

**владеть:**

- основными методами прикладного интеллектуального анализа данных, в том числе методами машинного обучения с учителем и без учителя, системами нечеткого вывода.

**Формируемые компетенции:**

- ПК-7. Работать с научно-технической информацией с использованием современных информационных технологий.
- ПК-10. Обосновывать достоверность полученных научных результатов.

- ПК-11. Формулировать выводы и рекомендации по применению результатов научно-исследовательской работы.
- ПК-18. Осуществлять поиск, систематизацию и анализ информации по перспективным направлениям информационной безопасности, инновационным технологиям, проектам и решениям.
- ПК-20. Определять цели инноваций и способы их достижения.

В соответствии с учебным планом на изучение дисциплины в 3 семестре отведено всего 122 часа, в том числе 48 аудиторных часов, из них лекции – 22 часа, лабораторные работы – 20 часов, управляемая самостоятельная работа – 6 часов. Форма текущей аттестации – зачет в 3 семестре.

Зачетные единицы – 3,5.

## СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

**1. Введение.** Основные определения. История развития интеллектуальных технологий обработки данных. Прямые, параллельные и распределенные вычисления. Процессорные архитектуры параллельной обработки, вычисления с использованием видеопроцессоров. Области применения. Обработка больших объемов данных.

**2. Машинное обучение с учителем.** Постановка задачи регрессии и классификации. Требования к обучающей выборке. Линейный и логистический регрессионный анализ. Наивная байесовская классификация. Метод опорных векторов. Метод k ближайших соседей. Методы глобальной, дискретной и комбинаторной оптимизации. Нейронные сети прямого распространения. Функции активации. Методы обучения нейронных сетей. Градиентный спуск и его модификации, метод Ньютона и квазиньютоновские методы, линейный и нелинейный метод сопряженных градиентов, комбинированные методы обучения. Специализированные архитектуры нейронных сетей: рекуррентные нейронные сети, автоэнкодеры, генеративные сети, сети с памятью и т.д. Глубокое обучение на основе нейронных сетей, сверточные нейронные сети. Анализ работы сети, регуляризация параметров.

**3. Машинное обучение без учителя.** Постановка задачи кластеризации. Сингулярное разложение матриц, метод главных компонент, линейное и нелинейное понижение размерности. Иерархическая и неиерархическая кластеризация. Метод k средних и его модификации, метод дендрограмм, алгоритм максимизации ожидания. Задача поиска аномалий и слепого разделения сигнала. Метод независимых компонент. Нейросетевые методы обучения без учителя. Самоорганизующиеся карты, глубокие сети доверия, правило Хебба и сеть Хопфилда. Оценка результатов обучения без учителя.

**4. Нечеткая логика.** Алгебраические структуры, отображения, операции и отношения. Понятия группы, кольца, поля, решетки. Базовые понятия теории категорий. Булевы алгебры как дистрибутивные решетки с дополнениями. Определение нечеткого множества. T-норма и t-конорма. Лингвистические переменные. Фаззификация и дефаззификация. Нечеткие правила и изоморфизм логических операций. Нечеткий вывод. Алгоритмы Мамдани и Такаги-Сугено-Канга. Нечеткие нейронные сети. Построение экспертных систем на основе нечеткого вывода.

## УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Количество часов УСР	Форма контроля знаний
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	2	3	4	5	6	7	8	9
<b>1</b>	<b>Введение</b>	<b>2</b>						
<b>2</b>	<b>Машинное обучение с учителем</b>	<b>10</b>						
	2.1 Постановка задачи регрессии и классификации. Требования к обучающей выборке. Линейный и логистический регрессионный анализ.	2			4			Выборочный опрос на лекции; отчет по лабораторной работе.
	2.2 Наивная байесовская классификация. Метод опорных векторов. Метод k ближайших соседей. Методы глобальной, дискретной и комбинаторной оптимизации.	2						Выборочный опрос на лекции.
	2.3 Нейронные сети прямого распространения. Функции активации. Методы обучения нейронных сетей. Градиентный спуск и его модификации, метод Ньютона и квазиньютоновские методы, линейный и	2					2	Выборочный опрос на лекции; реферат.

	нелинейный метод сопряженных градиентов, комбинированные методы обучения.							
	2.4 Специализированные архитектуры нейронных сетей: рекуррентные нейронные сети, автоэнкодеры, генеративные сети, сети с памятью и т.д. Глубокое обучение на основе нейронных сетей, сверточные нейронные сети.	2			4		2	Выборочный опрос на лекции; отчет по лабораторной работе; реферат.
	2.5 Анализ работы сети, регуляризация параметров.	2						
<b>3</b>	<b>Машинное обучение без учителя</b>	<b>4</b>						
	3.1 Постановка задачи кластеризации. Сингулярное разложение матриц, метод главных компонент, линейное и нелинейное понижение размерности. Иерархическая и неиерархическая кластеризация. Метод k средних и его модификации, метод дендрограмм, алгоритм максимизации ожидания.	2			4			Выборочный опрос на лекции; отчет по лабораторной работе.
	3.2 Задача поиска аномалий и слепого разделения сигнала. Метод независимых компонент. Нейросетевые методы обучения без учителя. Самоорганизующиеся карты, глубокие сети доверия, правило Хебба	2			4		2	Выборочный опрос на лекции; отчет по лабораторной работе; реферат.

	и сеть Хопфилда. Оценка результатов обучения без учителя.							
<b>4</b>	<b>Нечеткая логика</b>	<b>6</b>						
	4.1 Алгебраические структуры, отображения, операции и отношения. Понятия группы, кольца, поля, решетки. Базовые понятия теории категорий. Булевы алгебры как дистрибутивные решетки с дополнениями.	2						Выборочный опрос на лекции.
	4.2 Определение нечеткого множества. Т-норма и t-конорма. Лингвистические переменные. Фаззификация и дефаззификация. Нечеткие правила и изоморфизм логических операций.	2						Выборочный опрос на лекции.
	4.3 Нечеткий вывод. Алгоритмы Мамдани и Такаги-Сугено-Канга. Нечеткие нейронные сети. Построение экспертных систем на основе нечеткого вывода.	2			4			Выборочный опрос на лекции; отчет по лабораторной работе.
	Всего:	22			20		6	

## ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

### Перечень рекомендуемой литературы

#### *Основная*

1. Bishop, C. M., Pattern recognition and machine learning / C. M. Bishop. – New York: Springer, 2016. – 738 p.
2. Николенко, С. Глубокое обучение / С. Николенко, А. Кадурин, Е. Архангельская. – СПб: Питер Мейл, 2018. – 480 с.
3. Goodfellow, I. Deep learning / I. Goodfellow, A. Courville. – Cambridge, MA: MIT Press, 2016. – 800 p.
4. Ross, T. J. Fuzzy Logic With Engineering Applications / T. J. Ross. – Chichester, UK: John Wiley & Sons, Ltd., 3<sup>rd</sup> ed., 2010. – 584 p.
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. Koller, D. Probabilistic Graphical Models / D. Koller, N. Friedman. – The MIT Press, 2009. – 1270 p.

#### *Дополнительная*

1. Rojas, R. Neural Networks. A systematic introduction. / R. Rojas, J. Feldman. – Berlin: Springer Science & Business Media, 1996. – 502 p.
2. Segaran, T. Programming Collective Intelligence: Building Smart Web 2.0 Applications / T. Segaran. – O'Reilly Media, 2008. – 368 p.
3. Grus, J. Data Science from Scratch: First Principles with Python / J. Grus. – O'Reilly Media, 2015. – 330 p.
4. Рашид, Т. Создаём нейронную сеть. / Т. Рашид. – М: Вильямс, 2017. – 272 с.
5. Raschka, S. Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow / S. Raschka. – Packt Publishing – ebooks Account; 2<sup>nd</sup> edition, 2017. – 622 p.

## **ПРИМЕРНЫЙ ПЕРЕЧЕНЬ ЗАДАНИЙ УСР**

1. Библиотеки scіpy для языка программирования Python.
2. Использование библиотеки OpenCV в Python для обработки изображений.
3. Платформа машинного обучения Tensorflow.
4. Параллельные вычисления на видеопроцессоре с использованием NVidia CUDA.

## **ПРИМЕРНЫЙ ПЕРЕЧЕНЬ ТЕМ РЕФЕРАТОВ**

1. Нейронные сети в криптографии и криптоанализе.
2. Детектирование и локализация объектов видеоряда в реальном времени.
3. Уменьшение размерности и сжатие данных на основе автоэнкодеров.
4. Нелинейное уменьшение размерности с использованием диффузионных карт.
5. Семейство алгоритмов DBSCAN/OPTICS.
6. Алгоритм имитации отжига.
7. Графовые методы кластеризации.
8. Обучение по ассоциативным правилам (association rule learning).
9. Алгоритм нечеткого вывода на основе количественного ассоциативного правила (FI-QAR).

## **ПРИМЕРНЫЙ ПЕРЕЧЕНЬ ЛАБОРАТОРНЫХ РАБОТ**

1. Нейросетевой регрессионный анализ временных рядов.
2. Классификация и локализация объектов изображений с использованием сверточных нейронных сетей.
3. Предсказание поведения пользователя веб-сайта с использованием данных о его активности (web click data).
4. Кластерный анализ большого объема данных.
5. Построение экспертной системы на основе нечеткой логики.

## **ПЕРЕЧЕНЬ ИСПОЛЬЗУЕМЫХ СРЕДСТВ ДИАГНОСТИКИ**

С целью текущего контроля знаний и умений студентов по учебной дисциплине используются следующие диагностические средства:

- Выборочный опрос на лекциях;
- Отчеты по лабораторным работам;
- Обсуждение рефератов, презентаций и докладов студента, подготовленных по результатам выполнения лабораторных работ, УСР и самостоятельной работе по индивидуальным заданиям в рамках тематики учебной дисциплины.

Оценивание результатов выполнения лабораторных работ, заданий УСР и выполнения рефератов проводится в соответствии с критериями оценки знаний и компетенций студентов по 10-балльной шкале, изложенными в письме Министерства образования Республики Беларусь №21-04-1/105 от 22.12.2003 г.

Оценка текущей успеваемости определяется как средняя по оценкам лабораторных работ, управляемой самостоятельной работе и рефератам.

## **МЕТОДИКА ФОРМИРОВАНИЯ ИТОГОВОЙ ОЦЕНКИ**

Итоговая оценка формируется на основе:

1. Правил проведения аттестации студентов (Постановление Министерства образования Республики Беларусь № 53 от 29.05.2012 г.);

2. Положения о рейтинговой системе оценки знаний по дисциплине в БГУ (Приказ ректора БГУ от 18.08.2015 г. № 382-ОД);

Критериев оценки знаний студентов (письмо Министерства образования от 22.12.2003 г.)

Итоговый контроль усвоения дисциплины проводится в форме устного собеседования.

Итоговая оценка «зачтено» по дисциплине может быть выставлена студентам, получившим среднюю оценку по результатам итогового собеседования, лабораторным работам, рефератам и управляемой самостоятельной работе не ниже, чем «четыре».

Изложение лекционных материалов рекомендуется сопровождать примерами, иллюстрационным материалом и тестовыми заданиями с контрольными вопросами для закрепления понятий и терминов, устными фронтальными опросами на лекциях. Для успешного выполнения лабораторных работ студентам предлагается предварительно ознакомиться с описанием заданий, соответствующей теоретической частью курса, содержанием рекомендованной литературы. В целях формирования и развития у студентов навыков самоуправления, коммуникативных и

организационно-управленческих умений, а также приобретения опыта командного решения поставленных задач, предлагается организовывать группы студентов численностью до 3 человек для выполнения лабораторных работ. Лабораторные работы выполняются на компьютерах с использованием ресурсов сети Интернет, в средах математических пакетов, отчет подготавливается также на бумажном носителе. Управляемая самостоятельная работа студентов организуется в рамках выполнения лабораторных работ. Формой отчетности по итогам выполнения заданий УСР является реферат.