

# МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учебно-методическое объединение по образованию  
в области информатики и радиоэлектроники

## УТВЕРЖДАЮ

Первый заместитель Министра образования  
Республики Беларусь

\_\_\_\_\_ / тип.  
Регистрационный № ТД-\_\_\_\_\_ / тип.

## ОСНОВЫ МАШИННОГО ОБУЧЕНИЯ

Типовая учебная программа по учебной дисциплине  
для направления образования 28 Электронная экономика;  
группы специальностей:

45 01 Инфокоммуникационные технологии и системы связи;  
специальностей:

1-53 01 02 Автоматизированные системы обработки информации,  
1-98 01 02 Защита информации в телекоммуникациях

### СОГЛАСОВАНО

Председатель Учебно-методического  
объединения по образованию в  
области информатики и  
радиоэлектроники

\_\_\_\_\_ В.А. Богуш  
\_\_\_\_\_

### СОГЛАСОВАНО

Начальник Главного управления  
профессионального образования  
Министерства образования  
Республики Беларусь

\_\_\_\_\_  
\_\_\_\_\_

### СОГЛАСОВАНО

Проректор по научно-методической  
работе Государственного учреждения  
образования «Республиканский  
институт высшей школы»

\_\_\_\_\_ И.В. Титович  
\_\_\_\_\_

Эксперт-нормоконтролер

\_\_\_\_\_  
\_\_\_\_\_

Минск 2023

**СОСТАВИТЕЛЬ:**

Е.А.Баркова, заведующий кафедрой высшей математики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», кандидат физико-математических наук, доцент;

Н.В.Князюк, доцент кафедры высшей математики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», кандидат физико-математических наук;

О.В.Рыкова, доцент кафедры высшей математики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», кандидат физико-математических наук, доцент

**РЕЦЕНЗЕНТЫ:**

Кафедра дифференциальных уравнений и системного анализа Белорусского государственного университета (протокол № 7 от 12.01.2023)

П.В.Лукашевич, старший научный сотрудник государственного научного учреждения «Объединенный институт проблем информатики Национальной академии наук Беларуси», кандидат технических наук

**РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ В КАЧЕСТВЕ ТИПОВОЙ:**

Кафедрой высшей математики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» (протокол № 2 от 14.09.2022);

Научно-методическим советом учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» (протокол № \_\_\_ от \_\_\_\_\_);

Научно-методическим советом по информационной безопасности Учебно-методического объединения по образованию в области информатики и радиоэлектроники (протокол № 3 от 21.11.2022);

Научно-методическим советом по прикладным информационным системам и технологиям Учебно-методического объединения по образованию в области информатики и радиоэлектроники (протокол № 5 от 28.11.2022);

Научно-методическим советом по разработке программного обеспечения и информационно-коммуникационным технологиям Учебно-методического объединения по образованию в области информатики и радиоэлектроники (протокол № 4 от 12.12.2022);

Научно-методическим советом по системам и сетям инфокоммуникаций Учебно-методического объединения по образованию в области информатики и радиоэлектроники (протокол № 8 от 19.12.2022)

## ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

### ХАРАКТЕРИСТИКА УЧЕБНОЙ ДИСЦИПЛИНЫ

Типовая учебная программа по учебной дисциплине «Основы машинного обучения» разработана для студентов учреждений высшего образования, обучающихся по специальностям 1-28 01 01 Экономика электронного бизнеса, 1-28 01 02 Электронный маркетинг, 1-45 01 01 Инфокоммуникационные технологии (по направлениям), 1-45 01 02 Инфокоммуникационные системы (по направлениям), 1-53 01 02 Автоматизированные системы обработки информации, 1-98 01 02 Защита информации в телекоммуникациях в соответствии с требованиями образовательных стандартов высшего образования первой ступени и типовых учебных планов вышеуказанных специальностей.

Теория машинного обучения находится на стыке прикладной статистики, численных методов оптимизации, дискретного анализа и представляет собой самостоятельную математическую дисциплину. Основное содержание которой представлено способами предобработки и визуализации данных, методами машинного обучения (линейными, метрическими, решающими деревьями и их композициями). В результате освоения учебной дисциплины «Основы машинного обучения» студенты смогут освоить применение математического аппарата для решения задач по оценке и разработке моделей, применение методов машинного обучения для решения прикладных задач.

Воспитательное значение учебной дисциплины «Основы машинного обучения» заключается в формировании у обучающихся математической культуры и научного мировоззрения; развитии исследовательских умений, аналитических способностей, креативности, необходимых для решения научных и практических задач; развитии познавательных способностей и активности: творческой инициативы, самостоятельности, ответственности и организованности; формировании способностей к саморазвитию, самосовершенствованию и самореализации.

Изучение данной учебной дисциплины способствует созданию условий для формирования интеллектуально развитой личности обучающегося, которой присущи стремление к профессиональному совершенствованию, активному участию в экономической и социально-культурной жизни страны, гражданская ответственность и патриотизм.

### ЦЕЛЬ, ЗАДАЧИ УЧЕБНОЙ ДИСЦИПЛИНЫ

**Цель учебной дисциплины:** обзорное изучение современных методов автоматического обнаружения и описания скрытых закономерностей в данных.

**Задачи учебной дисциплины:**

освоение навыков самостоятельной разработки подходов к решению прикладных задач;

приобретение навыков практического использования автоматических методов извлечения скрытых закономерностей из данных;

развитие навыков самостоятельного расширения знаний компьютерных алгоритмов обработки данных.

Базовыми учебными дисциплинами по курсу «Основы машинного обучения» являются «Математический анализ», «Теория вероятностей и математическая статистика». В свою очередь учебная дисциплина «Основы машинного обучения» является базой для решения прикладных задач, возникающих в разных разделах науки и техники, в которых не удастся построить решение в замкнутом виде.

### ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины «Основы машинного обучения» формируются следующие компетенции:

*универсальная:* обладать навыками творческого аналитического мышления;

*базовая профессиональная:* использовать современные методологии, программные средства для построения и анализа моделей процессов, данных, объектов.

В результате изучения учебной дисциплины студент должен:

*знать:*

математические основы теории машинного обучения, основные подходы к предобработке данных;

основные алгоритмы классификации объектов;

формулировку задач регрессии, классификации, ранжирования, кластеризации, уменьшения размерности;

*уметь:*

выбирать метод машинного обучения, соответствующий исследуемой задаче;

визуализировать результаты работы алгоритмов машинного обучения;

*владеть:*

навыками применения методов анализа данных и машинного обучения с использованием современных программных средств;

междисциплинарным подходом при решении проблем.

Типовая учебная программа рассчитана на 108 учебных часов, из них – 50 аудиторных. Примерное распределение аудиторных часов по видам занятий: лекции – 26 часов, практические занятия – 24 часа.

## ПРИМЕРНЫЙ ТЕМАТИЧЕСКИЙ ПЛАН

Наименование темы	Всего аудиторных часов	Лекции	Практические занятия
Тема 1. Основные понятия анализа данных и машинного обучения	4	2	2
Тема 2. Классификация моделей по типу обучения	4	2	2
Тема 3. Оценка качества и выбор модели	4	2	2
Тема 4. Линейная регрессия	4	2	2
Тема 5. Классификация	4	2	2
Тема 6. Деревья решений	2	2	-
Тема 7. Метод опорных векторов	2	2	-
Тема 8. Ансамбли моделей. Регуляризация	4	2	2
Тема 9. Прогнозирование временных рядов	4	2	2
Тема 10. Кластеризация данных. Снижение размерности данных	10	4	6
Тема 11. Нейронные сети	8	4	4
<b>Итого:</b>	<b>50</b>	<b>26</b>	<b>24</b>

## СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

### Тема 1. ОСНОВНЫЕ ПОНЯТИЯ АНАЛИЗА ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ

Основные понятия анализа данных и машинного обучения: предикторы и целевые переменные в задачах обучения по прецедентам, типы переменных, модели, методы, обучение модели. Примеры прикладных задач машинного обучения. Основные этапы анализа данных. Стандарт CRISP-DM. Основные библиотеки для работы с данными.

### Тема 2. КЛАССИФИКАЦИЯ МОДЕЛЕЙ ПО ТИПУ ОБУЧЕНИЯ

Обучение с учителем: задачи классификации, регрессии, ранжирования. Обучение без учителя: кластеризация, уменьшение размерности, оценка переобучения. Переобучение модели.

### Тема 3. ОЦЕНКА КАЧЕСТВА И ВЫБОР МОДЕЛИ

Основные метрики для оценки качества модели в задачах классификации, регрессии, ранжирования. Скользящий контроль (кросс-валидация). Стратификация. Аналитические оценки качества: критерий Акаике (AIC), Байесовский информационный критерий (BIC). Методы отбора признаков.

### Тема 4. ЛИНЕЙНАЯ РЕГРЕССИЯ

Обобщенная линейная регрессия. Логистическая регрессия.

### Тема 5. КЛАССИФИКАЦИЯ

Бинарная и множественная классификации. Понятие о нечеткой классификации.

### Тема 6. ДЕРЕВЬЯ РЕШЕНИЙ

Деревья решений в задачах классификации и регрессии. Проблема построения оптимального дерева решений. Алгоритм CART.

### Тема 7. МЕТОД ОПОРНЫХ ВЕКТОРОВ

Метод опорных векторов в задачах классификации и регрессии. Ядра.

### Тема 8. АНСАМБЛИ МОДЕЛЕЙ. РЕГУЛЯРИЗАЦИЯ

Методы создания ансамблей моделей. Случайный лес. Регуляризация. Lasso, Ridge, ElasticNet.

### Тема 9. ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ

Понятие временного ряда. Примеры прикладных задач. Методы прогнозирования.

## Тема 10. КЛАСТЕРИЗАЦИЯ ДАННЫХ. СНИЖЕНИЕ РАЗМЕРНОСТИ ДАННЫХ

Алгоритмы кластеризации данных (k-средних, метод нечеткой кластеризации с-средних, DBSCAN, иерархическая кластеризация). Выбор числа кластеров, оценка качества кластеризации. Визуализация многомерных данных (алгоритм t-SNE). Метод главных компонент.

## Тема 11. НЕЙРОННЫЕ СЕТИ

Многослойный перцептрон. Обучение нейронной сети (метод обратного распространения ошибок). Сверточные нейронные сети (CNN).

## ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

### ЛИТЕРАТУРА

#### Основная

1. Джулли, А. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow / А. Джулли, С. Пал. – Москва : ДМК-Пресс, 2018. – 294 с
2. Горелик, А. Л. Методы распознавания : учебное пособие / А. Л. Горелик, В. А. Скрипкин. – 4-е изд., испр. – Москва : Высшая школа, 2004. – 261 с.
3. Лутц, М. Программирование на Python / М. Лутц ; пер. с англ. – 4-е изд. – Санкт-Петербург : Символ, 2013. – 992 с.
4. Вьюгин, В. Математические основы машинного обучения и прогнозирования / В. Вьюгин. – Москва : МЦНМО, 2014. – 218 с.
5. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. – Москва : ДМК-Пресс, 2012. – 412 с.
6. Силен, Д. Основы Data Science и Big Data. Python и наука о данных / Д. Силен, А. Мейсман, М. Али. – Санкт-Петербург : Питер, 2017. – 336 с.
7. Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль. – 2-е изд., испр. – Москва : ДМК-Пресс, 2018. – 652 с.
8. Флах, П. Машинное обучение / П. Флах ; пер. с англ. – Москва : ДМК-Пресс, 2015. – 400 с.
9. Вентцель, Е. С. Исследование операций: задачи, принципы, методология : учебное пособие / Е. С. Вентцель. – Москва : Высшая школа, 2007. – 208 с.

#### ДОПОЛНИТЕЛЬНАЯ

10. Паттерсон, Дж. Глубокое обучение с точки зрения практика / Дж. Паттерсон, А. Гибсон. – Москва : ДМК-Пресс, 2018. – 418 с.
11. PRIP 2011. Распознавание образов и обработка информации = Pattern Recognition and Information Processing : материалы 11-й международной конференции, Минск, 18 – 20 мая 2011 г. – Минск : БГУИР, 2011. – 472 с.

### МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ И ВЫПОЛНЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

При изучении учебной дисциплины рекомендуется использовать следующие формы самостоятельной работы:

- доклады на студенческих научных конференциях;
- выполнение стандартизированных тестов.



## ПЕРЕЧЕНЬ РЕКОМЕНДУЕМЫХ СРЕДСТВ ДИАГНОСТИКИ КОМПЕТЕНЦИЙ СТУДЕНТА

Типовыми учебными планами специальностей 1-28 01 01 Экономика электронного бизнеса, 1-28 01 02 Электронный маркетинг, 1-45 01 01 Инфокоммуникационные технологии (по направлениям), 1-45 01 02 Инфокоммуникационные системы (по направлениям), 1-53 01 02 Автоматизированные системы обработки информации, 1-98 01 02 Защита информации в телекоммуникациях в качестве формы промежуточной аттестации по учебной дисциплине «Основы машинного обучения» рекомендуется экзамен. Оценка учебных достижений студента осуществляется по десятибалльной шкале.

Для текущего контроля по учебной дисциплине могут использоваться следующие формы:

- контрольные работы;
- доклады на конференциях;
- коллоквиумы.

### РЕКОМЕНДУЕМЫЕ МЕТОДЫ (ТЕХНОЛОГИИ) ОБУЧЕНИЯ

Основные рекомендуемые методы (технологии) обучения, отвечающие целям и задачам учебной дисциплины:

элементы проблемного обучения (проблемное изложение, вариативное изложение, частично-поисковый метод), реализуемые во время чтения лекций и при проведении консультаций;

элементы учебно-исследовательской деятельности, реализация творческого подхода на практических занятиях.

### ПРИМЕРНЫЙ ПЕРЕЧЕНЬ ТЕМ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

1. Знакомство с Python. Основы numpy
2. Знакомство с pandas
3. Основы работы с библиотеками matplotlib, pyplot, seaborn
4. Предобработка данных.
5. Метрики качества и функции ошибок для задач классификации.
6. Метрики качества и функции ошибок для задач регрессии.
7. Ансамбли моделей.
8. Сверточные сети.
9. Нейронные сети.
10. Алгоритмы кластеризации.
11. Анализ и прогнозирование временных рядов.

ПРИМЕРНЫЙ ПЕРЕЧЕНЬ КОМПЬЮТЕРНЫХ ПРОГРАММ  
(необходимого оборудования, наглядных пособий и т.п.)

1. PS с ОП Windows 7, 8, 10.
2. САПР Matrix Laboratory .
3. Среда разработки Android Studio.
4. Пакет для интерактивного исполнения программ Jupyter Notebook.
5. Библиотеки языка программирования Python (Numpy, Scipy, Pandas, Tensorflow, Keras).
6. Microsoft Office 2010 и выше.