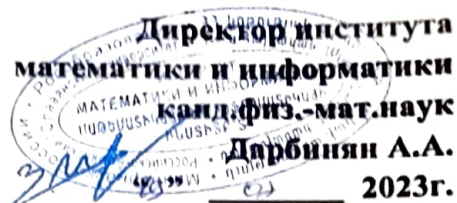


**ГОУ ВПО РОССИЙСКО-АРМЯНСКИЙ (СЛАВЯНСКИЙ)
УНИВЕРСИТЕТ**

Составлен в соответствии с государственными требованиями к минимуму содержания и уровню подготовки выпускников по направлению **01.03.02 Прикладная математика и информатика** и Положением «Об УМКД РАУ».

УТВЕРЖДАЮ:

**Директор института
математики и информатики
Канд. физ.-мат. наук
Дарбинян А.А.
2023г.**

Институт: Математики и информатики
Название института

Кафедра: Системное программирование
Название кафедры

Автор(ы): канд.тех.наук Гукасян Цолак

Ученое звание, ученая степень, Ф.И.О

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС

Дисциплина: **Б1.В.08 Специализация по компьютерной науке 3**

Код и название дисциплины согласно учебному плану

Специальность: **01.03.02 Прикладная математика и информатика**

Код и название специальности

Направление: **Прикладная математика и информатика**
Название направления

ЕРЕВАН

Структура и содержание УМКД

1. Аннотация

1.1. Данный курс является вводным и предназначен для предоставления студентам базовых знаний и практических навыков, необходимых для разработки моделей машинного обучения. В рамках курса основной акцент ставится на методы обучения с учителя и без учителя, самых распространенных видов задач в области машинного обучения. С целью развития практических навыков, рекомендуется регулярное проведение практических занятий, включающих знакомство с библиотеками разработки моделей машинного обучения, демонстрация и изучение применения таких моделей на актуальных примерах. Учитывая популярность языка программирования Python в области машинного обучения и тот факт, что используемые библиотеки написаны для этого языка, курс должен содержать вводное занятие по синтаксису языка и основным структурам данных. Требования к исходным уровням знаний, умений и навыков студентов для прохождения дисциплины (что должен знать, уметь и владеть студент для прохождения данной дисциплины)

1.2. Предварительное условие для прохождения : знание основ программирования

2. Содержание

2.1. Цели и задачи дисциплины

2.2. Требования к уровню освоения содержания дисциплины (какие компетенции (знания, умения и навыки) должны быть сформированы у студента ПОСЛЕ прохождения данной дисциплины)

Трудоёмкость дисциплины и виды учебной работы по учебному плану.

Виды учебной работы	Всего, в акад. часах	Распределение по семестрам						
		I сем	II сем	III сем	IV сем.	V сем	VI сем.	VII сем.
1	3	4	5	6	7	10	11	
1. Общая трудоёмкость изучения дисциплины по семестрам , в т. ч.:	72							
1.1. Аудиторные занятия, в т. ч.:	69						36	
1.1.1. Лекции								
1.1.2. Практические занятия, в т. ч.	36						36	
1.1.2.1. Обсуждение прикладных проектов								
1.1.2.2. Кейсы								
1.1.2.3. Деловые игры, тренинги								
1.1.2.4. Контрольные работы								
1.1.3. Семинары								
1.1.4. Лабораторные работы								
1.1.5. Другие виды аудиторных занятий								
1.2. Самостоятельная работа, в т. ч.:	36						36	
1.2.1. Подготовка к экзаменам								
1.2.2. Другие виды самостоятельной работы, в т.ч. (можно указать)								
1.2.2.1. Письменные домашние задания								
1.2.2.2. Курсовые работы								
1.2.2.3. Эссе и рефераты								
1.3. Консультации								
1.4. Другие методы и формы занятий (контроль)								
Итоговый контроль (Экзамен, Зачет, диф. зачет/указать)							Зачет	

2.3.2. Распределение объема дисциплины по темам и видам учебной работы

Таблица 1. Распределение объема дисциплины по темам и видам учебной работы.

Разделы и темы дисциплины	Всего (ак. часов)	Лекции (ак. часов)	Практ. занятия (ак. часов)	Семинары (ак. часов)	Лабор. (ак. часов)	Другие виды занятий (ак. часов)
Раздел 1: Введение	4		4			
Введение в машинное обучение	2		2			
Введение в Python	2		2			

Раздел 2: Обучение с учителем	22		22			
Линейная регрессия. Функция потерь. Градиентный спуск	4		4			
Библиотека scikit-learn	3		3			
Линейные методы классификации. Метрики качества классификации	2		2			
Переобучение. Регуляризация. Перекрестная проверка	3		3			
Извлечение и отбор признаков. Признаковое описание текстов	2		2			
Kaggle InClass соревнование по обработке текстов. Описание и базовое решение	3		3			
Метод k-ближайших соседей. Наивный байесовский классификатор	3		3			
Дерево решений. Bagging. Random Forest. Boosting	2		2			
Раздел 4: Нейронные сети	10		10			
Введение в нейронные сети	2		2			
Нейронные сети в библиотеке Keras	2		2			
Сверточные нейронные сети и обработка изображений	2		2			
Раздел 5: Обучение без учителя	2		2			
Обучение без учителя: кластеризация, сокращение размерности, фильтрация выбросов	2		2			
Итого	36		36			

2.3.3 Содержание разделов и тем дисциплины

Модуль 1

1. Введение в машинное обучение.

Определение основных понятий. Изучение основных видов задач: обучение с учителем, обучение без учителя, коллаборативная фильтрация, обучение с подкреплением. Обоснование актуальности и демонстрация вариантов применения.

2. Введение в Python.

Введение в программирование на языке Python: синтаксис, структуры данных. Знакомство с библиотекой numpy для быстрых вычислений. Знакомство с библиотеками обработки и визуализации данных pandas и matplotlib. Интерактивная разработка в Jupyter notebook и Google Colaboratory.

3. Линейная регрессия. Функция потерь. Градиентный спуск.

Определение линейных моделей регрессии с одной переменной, со многими переменными. Определение функции потерь. Алгоритм градиентного спуска. Стохастический градиентный спуск. Преимущества градиентного спуска над аналитическим решением. Метрики качества регрессии.

4. Библиотека scikit-learn.

Обучение моделей линейной регрессии с помощью библиотеки scikit-learn. Формат обучающих и тестовых данных. Изучение обученных параметров модели для анализа важности признаков. Изучение значений разных метрик качества для обученной модели.

5. Линейные методы классификации.

Метрики качества классификации. Определение линейных моделей классификации. Логистическая регрессия. Пересептрон. Метод опорных векторов. Метрики качества классификации: ассигасу, точность, полнота, f-1. Обучение линейных классификаторов с помощью библиотеки scikit-learn.

6. Переобучение. Регуляризация.

Перекрестная проверка. Переобучение, недообучение. Методы регуляризации l_1 и l_2 . Влияние l_1 регуляризации на веса признаков. Обучающая, валидационная, тестовая выборки. Перекрестная проверка для подбора гиперпараметров. Использование готовых методов из scikit-learn.

7. Извлечение и отбор признаков. Признаковое описание текстов.

Виды признаков. Кодирование категориальных признаков. Обработка признаков: нормализация, стандартизация. Мультиколлинеарность. Отбор признаков: фильтры, встроенные методы, методы обертки.

8. Kaggle InClass соревнование по обработке текстов. Описание и базовое решение.

Описание соревнования по решению актуальной задачи обработки текстов. Представление базового решения и возможных направлений для исследования.

9. Метод k-ближайших соседей. Наивный байесовский классификатор.

Гипотеза компактности. Алгоритм k-ближайших соседей для задач классификации и регрессии. Метрики расстояния. Взвешенный вариант алгоритма. Наивный байесовский классификатор. Мультиномиальный вариант алгоритма. Гауссовский вариант алгоритма. Реализации алгоритмов в scikit-learn.

10. Дерево решений. Bagging. Random Forest. Boosting.

промежуточного контроля в результирующей оценке промежуточных контролей т.д.								
Вес результирующей оценки промежуточных контролей в результирующей оценке итогового контроля								0,5
Экзамен/зачет (оценка итогового контроля)								0,5
	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$

2.3. Материально-техническое обеспечение дисциплины

Python

1. Руководство по установке: <https://realpython.com/installing-python/>
2. Официальная документация языка: <https://www.python.org/doc/>
3. Краткий обзор основ языка:
<https://cs231n.github.io/python-numpy-tutorial/#python-basic>
4. Платформа Anaconda: <https://www.anaconda.com/distribution/#download-section>
5. Jupyter Notebook для интерактивных вычислений:
http://math-hse.info/f/2018-19/py-polit/instruction_JN.pdf
6. Русскоязычные онлайн-курсы: <https://pythonworld.ru/kursy/free.html>
7. Англоязычный интерактивный туториал: <https://www.learnpython.org/>
8. NumPy туториал: <https://docs.scipy.org/doc/numpy-dev/user/quickstart.html>
9. Matplotlib туториал: http://matplotlib.org/users/pyplot_tutorial.html
10. Pandas туториал 1: <https://yadi.sk/i/pWwVPxvL3N9mX3>
туториал 2: <http://pandas.pydata.org/pandas-docs/stable/tutorials.html>

Машинное обучение

1. Онлайн-курсы:
2. Машинное обучение - онлайн курс Эндрю Бна из Стэнфордского университета:
<https://ru.coursera.org/learn/machine-learning>

3. Машинное обучение - курс Константина Воронцова для Школы анализа данных Яндекса: <https://ru.coursera.org/learn/machine-learning>
4. Учебники:
5. Bishop C. M. Pattern Recognition and Machine Learning. Springer, 2006.
6. Ian Goodfellow , Yoshua Bengio , Aaron Courville, Deep Learning, The MIT Press, 2016: <https://www.deeplearningbook.org/>
7. Другие материалы:
8. [machinelearning.ru](http://www.machinelearning.ru/) - информационный ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных: <http://www.machinelearning.ru/>
9. Руководство по глубокому обучению от YerevaNN: <https://yerevann.com/a-guide-to-deep-learning/>