

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Косенок Сергей Михайлович
Должность: ректор
Дата подписания: 19.06.2024 06:17:03
Уникальный программный ключ:
e3a68f3eaa1e62674b54f4998099d3d6bfdcf836

Бюджетное учреждение высшего образования
Ханты-Мансийского автономного округа-Югры
"Сургутский государственный университет"

УТВЕРЖДАЮ
Проректор по УМР

_____ Е.В. Коновалова

13 июня 2024г., протокол УМС №5

МОДУЛЬ ОБЩЕПРОФЕССИОНАЛЬНЫХ ДИСЦИПЛИН

Машинное обучение

рабочая программа дисциплины (модуля)

Закреплена за кафедрой **Информатики и вычислительной техники**

Учебный план g090402-УпрДан-24-2.plx
09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ
Направленность (профиль): Управление данными

Квалификация **Магистр**

Форма обучения **очная**

Общая трудоемкость **10 ЗЕТ**

Часов по учебному плану 360
в том числе:
аудиторные занятия 64
самостоятельная работа 260
часов на контроль 36

Виды контроля в семестрах:
экзамены 3

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	3 (2.1)		Итого	
	уп	рп		
Неделя	17 1/6			
Вид занятий	уп	рп	уп	рп
Лекции	32	32	32	32
Лабораторные	32	32	32	32
Итого ауд.	64	64	64	64
Контактная работа	64	64	64	64
Сам. работа	260	260	260	260
Часы на контроль	36	36	36	36
Итого	360	360	360	360

Программу составил(и):

к.ф.-м.н., доцент, Лысенкова Светлана Александровна

Рабочая программа дисциплины

Машинное обучение

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 09.04.02 Информационные системы и технологии (приказ Минобрнауки России от 19.09.2017 г. № 917)

составлена на основании учебного плана:

09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Направленность (профиль): Управление данными

утвержденного учебно-методическим советом вуза от 13.06.2024 протокол № 5.

Рабочая программа одобрена на заседании кафедры

Информатики и вычислительной техники

Зав. кафедрой и.о. зав. кафедрой ИВТ Лысенкова С.А.

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	Освоение современных методов машинного обучения. Машинное обучение – это использование алгоритмов анализа данных, получения выводов и решения или предсказания в отношении чего-либо. Т.е. вместо создания программ вручную с помощью специального набора команд для выполнения определенной задачи машину обучают с помощью большого количества данных и алгоритмов, которые дают ей возможность научиться выполнять эту задачу. Дисциплина помогает получить базовые компетенции аналитика данных, среди которых адекватный выбор метода для решения конкретной задачи анализа данных, предобработка данных, настройка параметров метода анализа и интерпретация полученных результатов.
-----	--

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП

Цикл (раздел) ООП:	Б1.О.03
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Основы научных исследований в области анализа данных
2.1.2	Специальные главы математики
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Производственная практика, преддипломная практика
2.2.2	Производственная практика, проектно-технологическая практика

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ОПК-4.1: Демонстрирует знания о новых научных принципах и методах исследования.

ОПК-4.2: Выполняет научные исследования в профессиональной сфере.

ОПК-4.3: Применяет на практике новые научные принципы и методы исследования для решения профессиональных задач

ОПК-2.1: Демонстрирует знания о современных информационно-коммуникационных и интеллектуальных технологиях, инструментальных средах, программно-технических платформах

ОПК-2.2: Делает обоснованный выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывает оригинальные программные средства для решения профессиональных задач

ОПК-2.3: Разрабатывает оригинальные программные средства, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач.

ОПК-1.1: Демонстрирует знания о математических, естественнонаучных и социально-экономических методах для использования в профессиональной деятельности

ОПК-1.2: Решает нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний

ОПК-1.3: Применяет на практике методы теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	основные понятия и парадигмы теории машинного обучения;
3.1.2	основные алгоритмы и принципы интеллектуального анализа данных и машинного обучения;
3.1.3	свободно-распространяемые системы машинного обучения и разработки данных;
3.1.4	возможности и ограничения машинного обучения;
3.1.5	способы правильно формулировать задачи для профессионалов;
3.1.6	международные и отечественные ИТ-стандарты при построении процессов управления в информационных технологиях;
3.2	Уметь:
3.2.1	использовать математические методы в технических приложениях;
3.2.2	строить математические модели и алгоритмы интеллектуального анализа данных в приложении к разным областям знаний;
3.2.3	применять алгоритмы машинного обучения к малым выборкам;
3.2.4	разрабатывать алгоритмы интеллектуального анализа данных для прикладных задач из различных областей знаний;
3.2.5	решать стандартные задачи машинного обучения;
3.2.6	решать стандартные задачи машинного обучения, не требующие подготовки данных и доработки типовых моделей;
3.2.7	проводить разработку и исследование методик анализа, синтеза, оптимизации и прогнозирования качества процессов функционирования информационных систем и технологий;
3.2.8	прогнозировать развитие информационных систем и технологий;
3.2.9	применять международные и отечественные ИТ-стандарты при построении процессов управления в информационных технологиях;

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетен-ции	Литература	Примечание
	Раздел 1. Введение. Задачи и терминология машинного обучения: supervised и unsupervised задачи; регрессия, классификация, кластеризация. Объект, признак, типы признаков, методы работы с ними. Метрики качества.					
1.1	Введение. Задачи и терминология машинного обучения: supervised и unsupervised задачи; регрессия, классификация, кластеризация. Объект, признак, типы признаков, методы работы с ними. Метрики качества. /Лек/	3	8	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
1.2	Введение. Задачи и терминология машинного обучения: supervised и unsupervised задачи; регрессия, классификация, кластеризация. Объект, признак, типы признаков, методы работы с ними. Метрики качества. /Лаб/	3	8	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	

1.3	Введение. Задачи и терминология машинного обучения: supervised и unsupervised задачи; регрессия, классификация, кластеризация. Объект, признак, типы признаков, методы работы с ними. Метрики качества. /Ср/	3	50	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
	Раздел 2. Изучение основ работы с векторными данными и визуализацией. Градиент. Методы оптимизации гладких функций. Реализация градиентного спуска для линейной регрессии.					
2.1	Изучение основ работы с векторными данными и визуализацией. Градиент. Методы оптимизации гладких функций. Реализация градиентного спуска для линейной регрессии. /Лек/	3	8	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
2.2	Изучение основ работы с векторными данными и визуализацией. Градиент. Методы оптимизации гладких функций. Реализация градиентного спуска для линейной регрессии. /Лаб/	3	8	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
2.3	Изучение основ работы с векторными данными и визуализацией. Градиент. Методы оптимизации гладких функций. Реализация градиентного спуска для линейной регрессии. /Ср/	3	50	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
	Раздел 3. Выделение признаков из текста: one-hot encoding, стемминг, лемматизация, tf-idf преобразование. Логистическая регрессия на примере задачи эмоциональной окраски текстов. L1 и L2 регуляризация.					
3.1	Выделение признаков из текста: one-hot encoding, стемминг, лемматизация, tf-idf преобразование. Логистическая регрессия на примере задачи эмоциональной окраски текстов. L1 и L2 регуляризация. /Лек/	3	8	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
3.2	Выделение признаков из текста: one-hot encoding, стемминг, лемматизация, tf-idf преобразование. Логистическая регрессия на примере задачи эмоциональной окраски текстов. L1 и L2 регуляризация. /Лаб/	3	8	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	

3.3	Выделение признаков из текста: one-hot encoding, стемминг, лемматизация, tf-idf преобразование. Логистическая регрессия на примере задачи эмоциональной окраски текстов. L1 и L2 регуляризация. /Ср/	3	50	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
	Раздел 4. Способы оценки качества моделей: holdout и кросс-валидация.Метод ближайших соседей. Деревья решений, случайный лес, градиентный бустинг. Решение соревнования на платформе Kaggle. Способы построения композиций моделей.					
4.1	Способы оценки качества моделей: holdout и кросс-валидация.Метод ближайших соседей. Деревья решений, случайный лес, градиентный бустинг. Решение соревнования на платформе Kaggle. Способы построения композиций моделей. /Лек/	3	4	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
4.2	Способы оценки качества моделей: holdout и кросс-валидация.Метод ближайших соседей. Деревья решений, случайный лес, градиентный бустинг. Решение соревнования на платформе Kaggle. Способы построения композиций моделей. /Лаб/	3	4	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
4.3	Способы оценки качества моделей: holdout и кросс-валидация.Метод ближайших соседей. Деревья решений, случайный лес, градиентный бустинг. Решение соревнования на платформе Kaggle. Способы построения композиций моделей. /Ср/	3	50	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
4.4	Контрольная работа /Контр.раб./	3	0	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
	Раздел 5. Алгоритмы кластеризации: K-means, DBscan, агломеративная кластеризация. Методы понижения размерности на основе матричных разложений (PCA и SVD). T-SNE.					

5.1	Алгоритмы кластеризации: K-means, DBscan, агломеративная кластеризация. Методы понижения размерности на основе матричных разложений (PCA и SVD). T-SNE. /Лек/	3	4	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
5.2	Алгоритмы кластеризации: K-means, DBscan, агломеративная кластеризация. Методы понижения размерности на основе матричных разложений (PCA и SVD). T-SNE. /Лаб/	3	4	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
5.3	Алгоритмы кластеризации: K-means, DBscan, агломеративная кластеризация. Методы понижения размерности на основе матричных разложений (PCA и SVD). T-SNE. /Ср/	3	60	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	
5.4	Машинное обучение /Экзамен/	3	36	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-2.1 ОПК-2.2 ОПК-2.3 ОПК-4.1 ОПК-4.2 ОПК-4.3	Л1.1 Л1.2Л2.1 Л2.2Л3.1 Э1 Э2 Э3 Э4 Э5	

5. ОЦЕНОЧНЫЕ СРЕДСТВА

5.1. Оценочные материалы для текущего контроля и промежуточной аттестации

Представлены отдельным документом

5.2. Оценочные материалы для диагностического тестирования

Представлены отдельным документом

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Топольников А. С., Латыпов Б. М.	Машинное обучение: теория и практическое применение в процессах добычи нефти: учебное пособие	Уфа: УГНТУ, 2020, электронный ресурс	1
Л1.2	Платонов А. В.	Машинное обучение: учебное пособие для вузов	Москва: Юрайт, 2023, электронный ресурс	1

6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
--	---------------------	----------	-------------------	----------

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Маккинли Уэс, Слинкина А.	Python и анализ данных	Саратов: Профобразование, 2019, электронный ресурс	1
Л2.2	Платонов А. В.	Машинное обучение: учебное пособие для вузов	Москва: Юрайт, 2024, электронный ресурс	1

6.1.3. Методические разработки

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л3.1	Лысенкова С. А.	Машинное обучение: методические рекомендации	Сургут: Издательский центр СурГУ, 2019, электронный ресурс	1

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Журнал для ИТ-профессионалов http://www.bytemag.iTi
Э2	Журнал Информационные ресурсы России http://rosenergo.gov.ru/information_and_analytical_support/informatsionnie_resursi_rossii
Э3	Журнал Информационные технологии и вычислительные системы http://www.jitcs.ru
Э4	Российский общеобразовательный портал http://www.school.edu.ru
Э5	Мир Интернет http://www.iworld.ni

6.3.1 Перечень программного обеспечения

6.3.1.1	Операционная система OS Windows XP, W7;
6.3.1.2	Программы браузеры
6.3.1.3	операционные системы Microsoft, пакет прикладных программ Microsoft Office 7
6.3.1.4	неисключительные права (лицензия) на неограниченный период на программное обеспечение MATLAB
6.3.1.5	неисключительные права (лицензия) на неограниченный период на программное обеспечение StatisticaBaseforWindows v.12 English / v.10 Russian) договор № 2014.302750 от 20.10.2014 г. бессрочно
6.3.1.6	Программное обеспечение ГИС MapInfoProfessional для образовательных учреждений, графические пакеты CS5 AdobeDesignPremium 5, CorelDRAWGraphicsSuiteX5, среда разработки EmbarcaderoDelphi, EmbarcaderoC++Builder 2010, договор 123/11-ГК от 12.12.2011 г. бессрочно

6.3.2 Перечень информационных справочных систем

6.3.2.1	Гарант-информационно-правовой портал. http://www.garant.ru/
6.3.2.2	КонсультантПлюс –надежная правовая поддержка. http://www.consultant.ru/

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1	Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа (лабораторных занятий), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации оснащена: комплект специализированной учебной мебели, маркерная (меловая) доска, комплект переносного мультимедийного оборудования - компьютер, проектор, проекционный экран, компьютеры с возможностью выхода в Интернет и доступом в электронную информационно-образовательную среду. Обеспечен доступ к сети Интернет и в электронную информационную среду организации.
-----	---