Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 23.10.2023 17:30:28
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ
Декан факультета информационных технологий
/Д. Г. Демидов/
"30" августа 2021 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Машинное обучение и анализ данных»
Направление подготовки
10.05.03 «Информационная безопасность автоматизированных систем»

Образовательная программа (профиль)
«Безопасность открытых информационных систем»

Квалификация (степень) выпускника
Специалист по защите информации

Форма обучения
Очная
Год приема - 2020

Москва 2020 г.

1. Цели освоения дисциплины

К **основным целям** освоения дисциплины «Машинное обучение и анализ данных» следует отнести:

- овладение студентами моделями и методами интеллектуального анализа данных и машинного обучения в задачах поиска информации, обработки и анализа данных;
- приобретение навыков исследователя данных (data scientist) и разработчика математических моделей, методов и алгоритмов анализа данных.

К **основным задачам** освоения дисциплины «Машинное обучение и анализ данных» следует отнести:

- знать основные модели и методы машинного обучения и разработки данных;
- уметь адекватно применять указанные модели и методы, а также программные средства, в которых они реализованы;
- иметь навыки (приобрести опыт) анализа реальных данных с помощью изученных методов/

Дисциплина «Машинное обучение и анализ данных» относится к числу профессиональных учебных дисциплин базовой части цикла (Б1) основной образовательной программы (Б.1.1.46).

Изучение дисциплины опирается на знания, умения и навыки, приобретенные в предшествующих дисциплинах: «Основы информационной безопасности», «Основы ИКТ», «Криптографические методы защиты информации».

2. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

В результате освоения дисциплины (модуля) у обучающихся формируются следующие компетенции и должны быть достигнуты следующие результаты обучения как этап формирования соответствующих компетенций:

Код компетенции	**В результате освоения образовательной программы обучающийся должен обладать**	**Перечень планируемых результатов обучения по дисциплине**
ОПК—8	Способен применять методы научных исследований при проведении разработок в области защиты информации в автоматизированных системах	**знать:** математические основы теории машинного обучения; основные классы алгоритмов машинного обучения и их представителей и их взаимосвязь, достоинства и недостатки; **уметь:** анализировать, выделять особенности и комбинировать методы машинного обучения; применять методы машинного обучения для решения прикладных задач; **владеть:** программными средствами для разработки алгоритмов машинного обучения, алгоритмами построения искусственных нейронных сетей; навыками разработки моделей машинного обучения.

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, т.е. **144** академических часов (лабораторные занятия – 72 час, самостоятельная работа - 72 часов, форма контроля – экзамен) в 8 семестре.

Структура и содержание дисциплины «Машинное обучение и анализ данных» по срокам и видам работы отражены в приложении.

5. Образовательные технологии

Методика преподавания дисциплины «Машинное обучение и анализ данных» и реализация компетентностного подхода в изложении и восприятии материала предусматривает использование следующих активных и интерактивных форм проведения групповых, индивидуальных, аудиторных занятий в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков у обучающихся:

- выполнение лабораторных работ в лабораториях вуза;
- индивидуальные и групповые консультации студентов преподавателем, в том числе в виде защиты выполненных заданий в рамках самостоятельной работы;
- посещение профильных конференций и работа на мастер-классах экспертов и специалистов индустрии;
- посещение лекций.

Самостоятельная внеаудиторная работа студентов составляет 50% от общего объема дисциплины и состоит из:

- подготовки к выполнению и подготовки к защите лабораторных работ;
- чтения литературы и освоения дополнительного материала в рамках тематики дисциплины;
- подготовки к текущей аттестации;
- подготовки к промежуточной аттестации.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

- экзамен.

Образцы вопросов к экзамену приведены в приложении.

6.1. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине (модулю)

6.1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

В результате освоения дисциплины (модуля) формируются следующие компетенции:

Код компетенции	В результате освоения образовательной программы обучающийся должен обладать
ОПК—8	Способен применять методы научных исследований при проведении разработок в области защиты информации в автоматизированных системах

В процессе освоения образовательной программы данные компетенции, в том числе их отдельные компоненты, формируются поэтапно в ходе освоения обучающимися дисциплин (модулей), практик в соответствии с учебным планом и календарным графиком учебного процесса.

6.1.2. Описание показателей и критериев оценивания компетенций, формируемых по итогам освоения дисциплины (модуля), описание шкал оценивания

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине (модулю):

ОПК—8 Способен применять методы научных исследований при проведении разработок в области защиты информации в автоматизированных системах				
Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
ЗНАТЬ	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие материалу дисциплины знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3).	Обучающийся демонстрирует неполное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые	Обучающийся демонстрирует частичное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие следующих знаний, указанных в индикаторах компетенций дисциплины «Знать» (см. п. 3). Свободно оперирует приобретенными знаниями.

		ситуации.		
УМЕТЬ	Обучающийся не умеет или в недостаточной степени умеет выполнять действия, указанных в индикаторах компетенций дисциплины «Уметь» (см. п. 3).	Обучающийся демонстрирует неполное соответствие умений, указанных в индикаторах компетенций дисциплины «Уметь» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность умений, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании умениями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие умений, указанных в индикаторах компетенций дисциплины «Уметь» (см. п. 3). Умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	Обучающийся демонстрирует полное соответствие умений, указанных в индикаторах компетенций дисциплины «Уметь» (см. п. 3). Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.
ВЛАДЕТЬ	Обучающийся не владеет или в недостаточной степени владеет приемами, методами и иными умениями, указанными в индикаторах компетенций дисциплины «Владеть» (см. п. 3).	Обучающийся в неполном объеме владеет приемами, методами и иными умениями, указанными в индикаторах компетенций дисциплины «Владеть» (см. п. 3). Допускаются значительные ошибки, проявляется недостаточность владения навыками по ряду показателей. Обучающийся испытывает значительные затруднения при применении навыков в новых ситуациях.	Обучающийся частично владеет приемами, методами и иными умениями, указанными в индикаторах компетенций дисциплины «Владеть» (см. п. 3). Навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	Обучающийся в полном объеме владеет приемами, методами и иными умениями, указанными в индикаторах компетенций дисциплины «Владеть» (см. п. 3). Свободно применяет полученные навыки в ситуациях повышенной сложности.

Шкалы оценивания результатов промежуточной аттестации и их описание:

Форма промежуточной аттестации: экзамен.

Промежуточная аттестация обучающихся в форме экзамена проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине (модулю), при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине (модулю) проводится преподавателем, ведущим занятия по дисциплине (модулю) методом экспертной оценки. По итогам промежуточной аттестации по дисциплине (модулю) выставляется оценка «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно».

Шкала оценивания	**Описание**
Отлично	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателей, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
Хорошо	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует неполное, правильное соответствие знаний, умений, навыков приведенным в таблицах показателей, либо если при этом были допущены 2-3 несущественные ошибки.
Удовлетворительно	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, в котором освещена основная, наиболее важная часть материала, но при этом допущена одна значительная ошибка или неточность.
Неудовлетворительно	Не выполнен один или более видов учебной работы, предусмотренных учебным планом. Студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателей, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.

Фонды оценочных средствпредставлены в приложении к рабочей программе.

7. Учебно-методическое и информационное обеспечение дисциплины

1. Основная литература:

- Крутиков, В.Н. Анализ данных : учебное пособие / В.Н. Крутиков, В.В. Мешечкин ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Кемеровский государственный университет». – Кемерово : Кемеровский государственный университет, 2014. – 138 с. : ил. – Режим доступа: по подписке. – URL: http://biblioclub.ru/index.php?page=book&id=278426 (дата обращения: 19.08.2019). – Библиогр. в кн. – ISBN 978-5-8353-1770-7. – Текст : электронный.

- Сергеев, Н.Е. Системы искусственного интеллекта : учебное пособие / Н.Е. Сергеев ; Министерство образования и науки РФ, Южный федеральный университет, Инженерно-технологическая академия. – Таганрог : Издательство Южного федерального университета, 2016. – Ч. 1. – 123 с. : схем., ил., табл. – Режим доступа: по подписке. – URL: http://biblioclub.ru/index.php?page=book&id=493307 (дата обращения: 19.08.2019). – Библиогр. в кн. – ISBN 978-5-9275-2113-5. – Текст : электронный.

2. Дополнительная литература:

- Грин, У. Эконометрический анализ : учебник / У. Грин ; пер. с англ. под науч. ред. С.С. Синельникова, М.Ю. Турунцевой ; Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации. – Москва : Издательский дом «Дело» РАНХиГС, 2016. – Книга 1. – 761 с. : табл. – (Академический учебник). – Режим доступа: по подписке. – URL: http://biblioclub.ru/index.php?page=book&id=563310 (дата обращения: 19.08.2019). – ISBN 978-5-7749-0959-9. - ISBN 978-5-7749-1157-8 (кн. 1). – Текст : электронный.
- Грин, У. Эконометрический анализ : учебник / У. Грин ; пер. с англ. под науч. ред. С.С. Синельникова, М.Ю. Турунцевой ; Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации. – Москва : Издательский дом «Дело» РАНХиГС, 2016. – Книга 2. – 753 с. : табл. – (Академический учебник). – Режим доступа: по подписке. – URL: http://biblioclub.ru/index.php?page=book&id=563313 (дата обращения: 19.08.2019). – Библиогр.: с. 1379-1422. – ISBN 978-5-7749-0959-9. - ISBN 978-5-7749-1158-5 (кн. 2). – Текст : электронный.

8. Материально-техническое обеспечение дисциплины

Для проведения всех видов занятий необходимо презентационное оборудование (мультимедийный проектор, экран) – 1 комплект.

Для проведения лабораторных занятий необходимо наличие компьютерных классов, оборудованных современной вычислительной техникой из расчета одно рабочее место на одного обучаемого.

Оборудование и аппаратура:

1. Компьютер с операционной системой Microsoft Windows.

9. Методические рекомендации для самостоятельной работы студентов

Изучение дисциплины осуществляется в строгом соответствии с целевой установкой в тесной взаимосвязи учебным планом. Основой теоретической подготовки студентов являются лекции.

В процессе самостоятельной работы студенты закрепляют и углубляют знания, полученные во время аудиторных занятий, готовятся к экзамену, а также самостоятельно изучают отдельные темы учебной программы.

10. Методические рекомендации для преподавателя

Программа составлена в соответствии с Федеральным государственным образовательным стандартом высшего образования по направлению подготовки **10.05.03 «Информационная безопасность автоматизированных систем»**.

Программу составил: проф. Федоров Н.В.

Программа утверждена на заседании кафедры “Информационная безопасность” «30» августа 2021 г., протокол № 1

Заведующий кафедрой
«Информационная безопасность»



к.т.н., доцент Н.В. Федоров

Приложение

Структура и содержание дисциплины «Машинное обучение и анализ данных»
по направлению подготовки
10.05.03 «Информационная безопасность автоматизированных систем»
(специалист)

n/n	Раздел	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов, и трудоемкость в часах					Виды самостоятельной работы студентов					Формы аттестации	
				Л	П/С	Лаб	СРС	КСР	К.Р.	К.П.	ДЗ	Реферат	К/р	Э	З
	8 семестр														
1	Введение, основные понятия анализа данных	8	1			4	4								
2	Математические объекты и методы в анализе данных		2			4	4								
3	Линейная регрессия и классификация		3			4	4								
4	Оценивание качества алгоритмов		4			4	4								
5	Логические методы		5			4	4								
6	Композиции алгоритмов		6			4	4								
7	Особенности реальных данных		7			4	4								
8	Анализ частых множеств признаков и ассоциативных правил		8			4	4								
9	Кластеризация данных		9			4	4								
10	Введение в теорию искусственных		10			4	4								

	нейронных сетей. Краткая история														
11	Примеры решения задач на базе искусственных нейронных сетей.		11			4	4								
12	Биологические предпосылки разработки искусственных нейронных сетей и нейрокомпьютеров на их основе.		12			4	4								
13	Место нейрокомпьютеров в современных компьютерных технологиях		13			4	4								
14	Математическая формализация нейронной структуры		14			4	4								
15	Классы задач, решаемых с помощью ИНС.		15			4	4								
16	Примеры построения нейроалгоритмов и нейронных сетей		16			6	6								
17	Классические модели искусственных нейронных сетей.		17-18			6	6								
	Форма аттестации	8	19-21							.				Э	
	Всего часов по дисциплине во восьмом семестре					72	72								
	Всего часов по дисциплине					72	72								

Приложение к
рабочей программе

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ

«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(МОСКОВСКИЙ ПОЛИТЕХ)

Направление подготовки: 10.05.03 «Информационная безопасность автоматизированных систем»

ОП (профиль): «Безопасность открытых информационных систем»

Форма обучения: очная

Вид профессиональной деятельности: эксплуатационная; проектно-технологическая; экспериментально-исследовательская; организационно-управленческая

.

Кафедра: «Информационная безопасность»

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

ПО ДИСЦИПЛИНЕ

«Машинное обучение и анализ данных»

Состав: 1. Паспорт фонда оценочных средств

2. Описание оценочных средств:

список вопросов к экзамену.

Составители: проф. Федоров Н.В.

Москва, 2021 год

Таблица 1

ПОКАЗАТЕЛЬ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ

Машинное обучение и анализ данных					
ФГОС ВО **10.05.03 «Информационная безопасность автоматизированных систем»**					
В процессе освоения данной дисциплины студент формирует и демонстрирует следующие **общепрофессиональные и профессиональные компетенции**:					
КОМПЕТЕНЦИИ		**Перечень компонентов**	**Технология формирования компетен**	**Форма оценоч-ного**	**Степени уровней освоения компетенций**
ИН-ДЕКС	**ФОРМУЛИРОВКА**				

ОПК—8	Способен применять методы научных исследований при проведении разработок в области защиты информации в автоматизированных системах	знать: математические основы теории машинного обучения; основные классы алгоритмов машинного обучения и их представителей и их взаимосвязь, достоинства и недостатки; уметь: анализировать, выделять особенности и комбинировать методы машинного обучения; применять методы машинного обучения для решения прикладных задач; владеть: программными средствами для разработки алгоритмов машинного обучения, алгоритмами построения искусственных нейронных сетей; навыками разработки моделей машинного обучения.	самостоятельная работа, лабораторные занятия, лекции	экзамен	Базовый уровень: знать: математические основы теории машинного обучения; уметь: анализировать, выделять особенности и комбинировать методы машинного обучения; владеть: программными средствами для разработки алгоритмов машинного обучения Повышенный уровень: знать: основные классы алгоритмов машинного обучения и их представителей и их взаимосвязь, достоинства и недостатки; уметь: применять методы машинного обучения для решения прикладных задач; владеть: алгоритмами построения искусственных нейронных сетей; навыками разработки моделей машинного обучения.

Приложение 2
к рабочей программе

Оценочные средства для промежуточной аттестации

Список вопросов для экзамена по дисциплине

1. Записать общую формулу байесовского классификатора (надо помнить формулу).
2. Какие вы знаете три подхода к восстановлению плотности распределения по выборке? 3. Что такое наивный байесовский классификатор?
3. Что такое оценка плотности Парзена-Розенблатта (надо помнить формулу). Выписать формулу алгоритма классификации в методе парзеновского окна.
4. На что влияет ширина окна, а на что вид ядра в методе парзеновского окна?
5. Многомерное нормальное распределение (надо помнить формулу). Вывести формулу квадратичного дискриминанта. При каком условии он становится линейным?
6. На каких предположениях основан линейный дискриминант Фишера?
7. Что такое «проблема мультиколлинеарности», в каких задачах и при использовании каких алгоритмов она возникает? Какие есть подходы к её решению?
8. Что такое «смесь распределений» (надо помнить формулу)?
9. Что такое EM-алгоритм, какова его основная идея? Какая задача решается на E-шаге, на M-шаге? Каков вероятностный смысл скрытых переменных?
10. Последовательное добавление компонент в EM-алгоритме, основная идея алгоритма.
11. Что такое стохастический EM-алгоритм, какова основная идея? В чём его преимущество (какой недостаток стандартного EM-алгоритма он устраняет)?
12. Что такое сеть радиальных базисных функций?
13. Что такое «выбросы»? Как осуществляется фильтрация выбросов?
14. Что такое обобщённый алгоритм классификации (надо помнить формулу)? Какие вы знаете частные случаи?
15. Как определяется понятие отступа в метрических алгоритмах классификации?
16. Что такое окно переменной ширины, в каких случаях его стоит использовать?
17. Основная идея алгоритма СТОЛП.
18. Что такое метод потенциальных функций? Идея алгоритма настройки. Сравните с методом радиальных базисных функций.
19. Зачем нужен отбор опорных объектов в метрических алгоритмах классификации?
20. Что такое функция конкурентного сходства? Основная идея алгоритма FRiS-СТОЛП.
21. Приведите пример метрического алгоритма классификации, который одновременно является байесовским классификатором.
22. Приведите пример метрического алгоритма классификации, который одновременно является линейным классификатором.
23. Что такое логическая закономерность? Приведите примеры закономерностей в задаче распознавания спама.
24. Часто используемые типы логических закономерностей.
25. Дайте определение эпсилон-дельта-логической закономерности (помнить формулы).
26. Сравните области статистических и логических закономерностей в (p,n)-плоскости.
27. С какой целью делается бинаризация?

28. В чём заключается процедура бинаризации признака?
29. Как происходит перебор в жадном алгоритме синтеза информативных конъюнкций
30. Как приспособить жадный алгоритм синтеза конъюнкций для синтеза информативных шаров?
31. Что такое стохастический локальный поиск?
32. В чём отличия редукции и стабилизации? В чём их достоинства и недостатки?
33. Что такое решающий список?
34. Какие критерии информативности используются при синтезе решающего списка и почему?
35. Достоинства и недостатки решающих списков.
36. Что такое решающее дерево? Достоинства и недостатки решающих деревьев.
37. Зачем делается редукция решающих деревьев?
38. Какие есть два основных типа редукции решающих деревьев?
39. Как преобразовать решающее дерево в решающий список, и зачем это делается?
40. Почему возникает проблема предпочтения признаков с меньшими номерами в алгоритме КОРА? Как она решается?
41. Основная идея алгоритма ТЭМП
42. Какие критерии информативности используются в алгоритме ТЭМП и почему?
43. Достоинства и недостатки алгоритма ТЭМП.