

МИНОБРНАУКИ РОССИИ



Федеральное государственное бюджетное образовательное учреждение высшего образования

**«Российский государственный гуманитарный университет»
(ФГБОУ ВО «РГГУ»)**

*ОТДЕЛЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ В ГУМАНИТАРНОЙ СФЕРЕ
Учебно-научный центр интеллектуальной робототехники*

НАЗВАНИЕ ДИСЦИПЛИНЫ

«Интеллектуальный анализ данных и машинное обучение»

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Интеллектуальный анализ данных и машинное обучение»

Код и наименование направления подготовки (специальности) 45.03.04

«Интеллектуальные системы в гуманитарной сфере»

Наименование направленности (профиля, специализации, программы магистратуры)

«Разработка и программирование интеллектуальных систем в гуманитарной сфере»

Уровень квалификации выпускника: бакалавр

Форма обучения: очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2021



Название дисциплины: **«Интеллектуальный анализ данных и машинное обучение»**
Рабочая программа дисциплины **«Интеллектуальный анализ данных и машинное обучение»**

Составитель:

Д.ф.-м.н., проф. УНЦ ИР Д.В. Виноградов

УТВЕРЖДЕНО

Протокол заседания кафедры

№ 3 от 18.05.2021



ОГЛАВЛЕНИЕ

1. Пояснительная записка

1.1 Цель и задачи дисциплины

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

1.3. Место дисциплины в структуре образовательной программы

2. Структура дисциплины

3. Содержание дисциплины (модуля)

4. Образовательные технологии

5. Оценка планируемых результатов обучения

5.1. Система оценивания

5.2. Критерии выставления оценок

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине (модулю)

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

7. Материально-техническое обеспечение дисциплины (модуля)

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

9. Методические материалы

9.1. Планы семинарских занятий

Приложения

Приложение 1. Аннотация дисциплины

Приложение 2. Лист изменений



1. Пояснительная записка

1.1. Цель и задачи дисциплины «*Интеллектуальный анализ данных и машинное обучение*»

Цель дисциплины — усвоение студентами основных идей интеллектуального анализа данных (ИАД), моделей и методов машинного обучения, основанных на символьном представлении данных.

Задачи дисциплины: изложение основных алгоритмов машинного обучения: голосование для пространств версий, Литтлстоуна для линейно-отделимых понятий, Ривеста для списков решений.

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине:

Коды и содержание компетенций	Индексы компетенций	Перечень планируемых результатов обучения по дисциплине
ОПК-1 способность использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа, логики и моделирования, теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках	ОПК-1.1. Способен использовать основы математического анализа, логики и математического моделирования.	Знать: подход к интеллектуальному анализу данных на основе машинного обучения. Уметь: применять методы машинного обучения для интеллектуального анализа данных.
ОПК-2 способность получать знания в области современных проблем науки, техники и технологии информатики, гуманитарных, лингвистических, и социальных наук	ОПК 2.1. Знает методы доступа к информационным ресурсам.	Знать: основные понятия интеллектуального анализа данных и машинного обучения. Владеть: простейшими навыками применения алгоритмов машинного обучения для переработки больших объемов информации.
ОПК-3 способность разрабатывать алгоритмы и программы, пригодные для	ОПК-3.2. Умеет использовать возможности операционных систем, операционных сред, интегрированных сред	Владеть: простейшими навыками встраивания алгоритмов машинного обучения в новые информационные технологии.



практического применения	программирования и офисных приложений для практической работы на компьютере, подготовки документов, разработки и отладки программного кода.	
--------------------------	---	--

1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Интеллектуальный анализ данных и машинное обучение» относится к базовой части Б1.О.17 блока дисциплин учебного плана.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик: *теория вероятностей и статистика, дискретная математика, теория алгоритмов.*

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин и прохождения практик: *преддипломная практика, методология разработки интеллектуальных систем.*



2. Структура дисциплины

Общая трудоёмкость дисциплины составляет 5 з.е., 190 ч., в том числе контактная работа обучающихся с преподавателем 70 ч., промежуточная аттестация 18 ч., самостоятельная работа обучающихся 102 ч.

№ п/п	Раздел дисциплины/темы	Се ме ст р	Виды учебной работы (в часах)						Формы текущего контроля успеваемости, форма промежуточной аттестации (<i>по семестрам</i>)
			Контактная				Про меж уточ ная атте стац ия	Са мо сто яте ль- на я ра бот а	
			Лек ции	Се ми нар	Прак тичес кие занят ия	Лабор аторны е заняти я			
	<i>Раздел 1</i>	4	2	4				20	
	<i>Раздел 2</i>	4	8	20				48	
	<i>Раздел 3</i>	4	8	24				34	
	<i>зачёт</i>	4		4			18		<i>итоговая контрольная работа</i>
	ИТОГО:		18	52			18	102	

3. Содержание дисциплины¹

№	Наименование раздела дисциплины	Содержание
1	Интеллектуальный анализ данных и машинное обучение	Методы интеллектуального анализа данных. Машинное обучение. Обучение без учителя. Обучение с учителем. Обучение с подкреплением
2	Обучение с ограничением на число ошибок	Обучение с ограничением на число ошибок. Алгоритм пересечения для обучения конъюнктивным понятиям. Пространство версий и алгоритм голосования. Снижение размерности. Алгоритм Winnow1. Персептрон. Алгоритм обучения персептрону. Зазор в обучающей выборке. Теорема о числе ошибок. Отступ центральной функции. Опорные вектора. Машина опорных векторов для персептрона. Двойственная задача. Теорема Каруша-Куна-Таккера и опорные вектора.
3	Вероятно-приближенно корректное обучение	Вероятно-приближенно корректное обучение. Эффективная обучаемость. Алгоритм пересечения для обучения конъюнктивным понятиям. Бритва Оккама. Выборки большого объема для конечного класса гипотез. Списки решений. Алгоритм Ривеста обучения спискам решений. Элементы теории сложности вычислений (напоминание). NP-трудность задачи 3-раскрашивания графа. Вероятностно-полиномиальные языки. Преобразование графа в обучающую выборку. 3-членные ДНФ и 3-раскрашивания. Невозможность эффективного ВПК-обучения 3-членным ДНФ. Расширение пространства версий и эффективная ВПК-обучаемость после этого.

¹ Содержание дисциплины (модуля) должно состоять из разделов, соответствующих структуре дисциплины, подразделов и отдельных тем с той степенью подробности, которая, по мнению автора, оптимально способствуют достижению цели и реализации поставленных задач.

Содержание дисциплины (модуля) может быть представлено в текстовой или в табличной форме



4. Образовательные технологии

Образовательные технологии

№ п/п	Наименование раздела	Виды учебных занятий	Образовательные технологии
1	2	3	4
1.	Интеллектуальный анализ данных и машинное обучение	<p><i>Лекция 1. Интеллектуальный анализ данных и машинное обучение</i></p> <p><i>Семинар 1. Методы интеллектуального анализа данных</i></p> <p><i>Семинар 2. Обучение с учителем и без, обучение с подкреплением</i></p> <p><i>Самостоятельная работа</i></p>	<p><i>Лекция с использованием авторских конспектов</i></p> <p><i>Разбор решения задач</i></p> <p><i>Консультирование посредством электронной почты</i></p>
2.	Обучение с ограничением на число ошибок	<p><i>Лекция 2. Обучение с ограничением на число ошибок. Пространства версий.</i></p> <p><i>Лекция 3. Снижение размерности.</i></p> <p><i>Лекция 4. Обучения перцептрон.</i></p> <p><i>Лекция 5. Машина опорных векторов для перцептрона.</i></p> <p><i>Семинар 3. Алгоритм пересечения для обучения</i></p> <p><i>Семинар 4. Алгоритм голосования для пространства версий</i></p> <p><i>Семинар 5. Алгоритм случайного выбора для пространства версий</i></p> <p><i>Семинар 6. Алгоритм Winnow1.</i></p> <p><i>Семинар 7. Алгоритм обучения перцептрон</i></p> <p><i>Контрольная работа</i></p> <p><i>Самостоятельная работа</i></p>	<p><i>Лекции с использованием авторских конспектов</i></p> <p><i>Разбор решения задач</i></p> <p><i>Контрольная работа</i></p> <p><i>Консультирование посредством электронной почты</i></p>
3.	Вероятно-приближенно корректное обучение	<p><i>Лекция 6. Вероятно-приближенно корректное обучение</i></p> <p><i>Лекция 7. ВПК-обучаемость конечного пространства гипотез</i></p> <p><i>Лекция 8. Невозможность ВПК-обучения.</i></p> <p><i>Лекция 9. Размерность Вапника-Червоненкиса.</i></p>	<p><i>Лекции с использованием авторских конспектов</i></p>



	<p><i>Семинар 8. ВПК-обучаемость алгоритмом пересечений</i></p> <p><i>Семинар 9. Алгоритм Р.Ривеста обучения спискам решений.</i></p> <p><i>Семинар 10. 3-раскрашиваемость графа и обучающая выборка</i></p> <p><i>Семинар 11. Вычисление размерности Вапника-Червоненкиса.</i></p> <p><i>Контрольная работа</i></p> <p><i>Самостоятельная работа</i></p>	<p><i>Разбор решения задач</i></p> <p><i>Контрольная работа</i></p> <p><i>Консультирование посредством электронной почты</i></p>
--	---	--



В период временного приостановления посещения обучающимися помещений и территории РГГУ. для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

- видео-лекции;
- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;
- консультации с использованием телекоммуникационных средств.

5. Оценка планируемых результатов обучения

5.1. Система оценивания²

Система оценивания может быть представлена как в текстовой, так и в табличной форме.

Например:

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль:		
- контрольная работа (темы 1-2)	30 баллов	30 баллов
- контрольная работа (темы 3)	30 баллов	30 баллов
Промежуточная аттестация зачет		40 баллов
Итого за семестр (дисциплину) зачёт		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55		E	
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

²

Система оценивания выстраивается в соответствии с учебным планом, где определены формы промежуточной аттестации (зачёт/зачёт с оценкой/экзамен), и структурой дисциплины, где определены формы текущего контроля. Указывается распределение баллов по формам текущего контроля и промежуточной аттестации, сроки отчётности.



5.2. Критерии выставления оценки по дисциплине³

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ А,В	«отлично»/ «зачтено (отлично)»/ «зачтено»	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	«хорошо»/ «зачтено (хорошо)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	«удовлетвори-тель-но»/ «зачтено (удовлетвори-тель-но)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p>

³ Могут уточняться и дополняться в соответствии со спецификой дисциплины, установленных форм контроля, применяемых технологий обучения и оценивания.



		Демонстрирует достаточный уровень знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	«неудовлетворительно»/ не зачтено	Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине (модулю)⁴

Задачи для промежуточных контрольных по темам 1 и 2-3 и для зачета выдаются (по вариантам) перед проведением текущего контроля (и здесь не будут приведены во избежание их решения до проведения аттестации, возможно, репетиторами).

Контрольная работа 1

1. Для дизъюнкции $x_1 \vee x_5$ выпишите шаги продвижения и устранения в алгоритме Winnow1 (переменные x_1, \dots, x_5).
2. Найдите опорные вектора и проведите разделяющую прямую для обучающей выборки $\{(0,0),-1\}, \{(-3,3),+1\}, \{(3,5),+1\}, \{(0,5),+1\}, \{(1,-2),-1\}, \{(-2,0),+1\}$

Контрольная работа 2

1. Постройте обучающую выборку по графу, заданного матрицей смежности $((01010),(10101),(01010),(10101),(01010))$.
2. Для списка решений $\langle \sim x_3 \ \& \ \sim x_4, 0 \rangle, \langle x_1 \ \& \ \sim x_2, 0 \rangle, \langle \text{true}, 1 \rangle$ оракул выдал примеры с номерами 13, 16, 5, 3, 12, 5, 7, 14, 15, 4, 2, 11. Вычислите список-гипотезу и расхождение (распределение равномерное).

1 Контрольные вопросы к зачету

1. Методы интеллектуального анализа данных.
2. Типы машинного обучения.
3. Обучение с ограничением на число ошибок.

⁴ Приводятся примеры оценочных средств в соответствии со структурой дисциплины и системой контроля: варианты тестов, тематика письменных работ, примеры экзаменационных билетов, типовые задачи, кейсы и т.п. Оценочными средствами должны быть обеспечены все формы текущего контроля и промежуточной аттестации. Они должны быть ориентированы не только на проверку сформированности знаний, но также умений и владений.



4. Алгоритм пересечения для обучения конъюнктивным понятиям.
5. Пространство версий и алгоритм голосования.
6. Снижение размерности. Алгоритм Winnow1.
7. Персептрон. Алгоритм обучения персептрону.
8. Зазор в обучающей выборке.
9. Теорема о числе ошибок.
10. Отступ центральной функции. Опорные вектора.
11. Машина опорных векторов для персептрона.
12. Двойственная задача. Теорема Каруша-Куна-Таккера и опорные вектора.
13. Скользящий контроль для машины опорных векторов.
14. Вероятно-приближенно корректное обучение. Эффективная обучаемость.
15. Алгоритм пересечения для обучения конъюнктивным понятиям.
16. Списки решений. Алгоритм Ривеста обучения спискам решений.
17. Преобразование графа в обучающую выборку. 3-членные ДНФ и 3-раскрашивания.
18. Невозможность эффективного ВПК-обучения 3-членным ДНФ.
19. Расширение пространства версий и эффективная ВПК-обучаемость после этого.

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

Источники

основные

1. конспекты лекций к.ф.-м.н., доц. Д.В.Виноградова (высылаются на электронную почту и раздаются в напечатанном виде студентам перед лекциями)

дополнительные

1. материалы в сети Интернет (находятся студентами самостоятельно)

Литература

основная

1. Вьюгин, В. В. Математические основы машинного обучения и прогнозирования : учебное пособие / В. В. Вьюгин. — Москва : МЦНМО, 2014. — 304 с. — ISBN 978-5-4439-2014-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/56397>.

6.2 Перечень БД и ИСС

№п /п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2021 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2021 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru



4	Компьютерные справочные правовые системы Консультант Плюс, Гарант
---	---





7. Материально-техническое обеспечение дисциплины (модуля)

В качестве материально-технического обеспечения дисциплины указывается необходимое для обучения лицензионное программное обеспечение, оборудование, демонстрационные приборы, мультимедийные средства, учебные фильмы, тренажеры, карты, плакаты, наглядные пособия; требования к аудиториям – компьютерные классы, академические или специально оборудованные аудитории и лаборатории, наличие доски и т.д. необходима аудитория с хорошими досками

1. Перечень ПО

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Adobe Master Collection CS4	Adobe	лицензионное
2	ОС «Альт Образование» 8	ООО «Базальт СПО	лицензионное
3	Windows 10 Pro	Microsoft	лицензионное
4	Kaspersky Endpoint Security	Kaspersky	лицензионное
5	Microsoft Office 2016	Microsoft	лицензионное
6	Zoom	Zoom	лицензионное

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
 - письменные задания выполняются на компьютере в письменной форме;
 - зачёт проводится в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением;
 - зачёт проводится в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.



Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для глухих и слабослышащих:
 - в печатной форме;
 - в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа;
 - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для глухих и слабослышащих:
 - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;
 - акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - передвижными, регулируемыми эргономическими партами СИ-1;
 - компьютерной техникой со специальным программным обеспечением.

9. Методические материалы⁵

9.1. Планы семинарских занятий⁶

Тема 1. (4 ч.) *Методы интеллектуального анализа данных*

Цель занятий: усвоить методологию интеллектуального анализа данных.

Форма проведения – обсуждение, решение задач.

Вопросы для обсуждения:

Какие бывают методы интеллектуального анализа данных?

Какие бывают задачи машинного обучения?

Контрольные вопросы:

Логические методы ИАД
Алгебраические методы ИАД.
Метрические методы ИАД.
Машинное обучение без учителя
Кластерный анализ
Факторный анализ
Машинное обучение с учителем
Обучение с подкреплением

Список источников и литературы:

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

⁵ Методические материалы по дисциплине могут входить в состав рабочей программы, либо разрабатываться отдельным документом.

⁶ План занятий строится в соответствии со структурой дисциплины (п.2). Разделы плана включают: название темы, количество часов, форму проведения занятия, его содержание (вопросы для обсуждения, задания, контрольные вопросы, кейсы и т.п.), список литературы. При необходимости, планы практических и лабораторных занятий могут содержать указания по выполнению заданий и требования к материально-техническому обеспечению занятия.



<https://wikipedia.org>

Материально-техническое обеспечение занятия: доска, видеопроектор, ноутбук.

Тема 2. (20 ч.) Обучение с ограничением на число ошибок

Цель занятий: изучить модель онлайн обучения с ограничением на число ошибок.

Форма проведения – обсуждение, решение задач.

Вопросы для обсуждения:

Как устроено обучение с ограничением на число ошибок?

Как достигается снижение размерности в бинарных данных?

Каково максимальное число ошибок при обучении перцептрон?

Что такое машина опорных векторов?

Как провести разделяющую гиперплоскость по опорным векторам?

Контрольные вопросы:

Алгоритм пересечения для обучения конъюнктивным понятиям.

Пространство версий и алгоритм голосования.

Алгоритм Winnow1 снижения размерности.

Алгоритм обучения перцептрон.

Отступ центральной функции. Опорные вектора.

Машина опорных векторов для перцептрона.

Двойственная задача. Теорема Каруша-Куна-Таккера и опорные вектора.

Скользкий контроль для машины опорных векторов.

Список источников и литературы:

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

Виноградов, Д.В. Статистическое машинное обучение, 2017

<https://vk.com/club190684626>

<https://wikipedia.org>

Материально-техническое обеспечение занятия: доска, видеопроектор, ноутбук.

Тема 3. (24 ч.) Вероятно-приближенно корректное обучение

Цель занятий: усвоить методологию ВПК-обучения.

Форма проведения – обсуждение, решение задач.

Вопросы для обсуждения:

Как обучаться из вероятностно порождаемых данных?

Как устроено ВПК-обучение?

Как обучаться спискам решений?

Почему иногда машинное обучение невозможно?



Контрольные вопросы:

Вероятно-приближенно корректное обучение.
Алгоритм пересечения для обучения конъюнктивным понятиям.
Алгоритм Ривеста обучения спискам решений.
Невозможность эффективного ВПК-обучения
Расширение пространства гипотез

Список источников и литературы:

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

Виноградов, Д.В. Статистическое машинное обучение, 2017

<https://vk.com/club190684626>

<https://wikipedia.org>

Материально-техническое обеспечение занятия: доска, видеопроектор, ноутбук.



АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина «*Интеллектуальный анализ данных и машинное обучение*» входит в состав базовой части дисциплин Б1.О.17 для подготовки бакалавров по направлению подготовки 45.03.04 «Интеллектуальные системы в гуманитарной сфере». Дисциплина реализуется на Отделении интеллектуальных систем в гуманитарной сфере кафедрой Математики, логики и интеллектуальных систем в гуманитарной сфере в 4 семестре.

Содержание дисциплины охватывает круг вопросов, связанных с применением методов машинного обучения в интеллектуальном анализе данных с помощью интеллектуальных систем.

Цель дисциплины — усвоение студентами основных идей интеллектуального анализа данных (ИАД), моделей и методов машинного обучения, основанных на символьном представлении данных.

Задачи дисциплины: изложение основных алгоритмов машинного обучения: голосование для пространств версий, Литтлстоуна для линейно-отделимых понятий, Ривеста для списков решений.

Дисциплина направлена на формирование следующих компетенций:

УК-1 - способность осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач

ОПК-1 - способность использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа, логики и моделирования, теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках

ОПК-2 - способность получать знания в области современных проблем науки, техники и технологии информатики, гуманитарных, лингвистических, и социальных наук

ОПК-3 - способность разрабатывать алгоритмы и программы, пригодные для практического применения

В результате изучения дисциплины студент должен:

Знать:

- основные понятия интеллектуального анализа данных и машинного обучения;
- подход к интеллектуальному анализу данных на основе машинного обучения;
- теоретические ограничения для алгоритмов машинного обучения.

Уметь:

- применять методы машинного обучения для интеллектуального анализа данных.

Владеть:

- простейшими навыками применения алгоритмов машинного обучения для переработки больших объемов информации;
- простейшими навыками встраивания алгоритмов машинного обучения в новые информационные технологии.

Рабочей программой предусмотрены следующие виды контроля: текущий контроль успеваемости в форме устных ответов у доски и написания контрольных работ, промежуточная аттестация в форме зачета.