

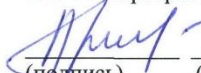


МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Дальневосточный федеральный университет»
(ДВФУ)

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

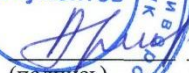
«СОГЛАСОВАНО»

Руководитель ОП Разработка программно-информационных систем по направлению 09.04.04 Программная инженерия


(подпись) Артемяева И.Л.
(Ф.И.О. рук. ОП) « 21 » 07 2018 г.



«УТВЕРЖДАЮ»
Заведующая кафедрой прикладной математики, механики, управления и программного обеспечения


(подпись) Артемяева И.Л.
(Ф.И.О. зав. каф.) « 21 » 07 2018 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Машинное обучение в системах искусственного интеллекта

Направление подготовки – 09.04.04 Программная инженерия

Магистерская программа «Разработка программно-информационных систем»

Форма подготовки (очная)

курс 2 семестр 3

лекции 18 час.

практические занятия 36 час.

лабораторные работы 0 час.

в том числе с использованием МАО лек. ____/пр. ____/лаб. 18 час.

в том числе в электронной форме лек. ____/пр. ____/лаб. ____ час.

всего часов аудиторной нагрузки 54 час.

в том числе с использованием МАО ____ час.

в том числе контролируемая самостоятельная работа 0 час.

в том числе в электронной форме ____ час.

самостоятельная работа 54 час.

в том числе на подготовку к экзамену ____ час.

курсовая работа / курсовой проект _____ семестр

зачет 3 семестр

экзамен не предусмотрен

Рабочая программа составлена в соответствии с требованиями образовательного стандарта, самостоятельно устанавливаемого ДВФУ, утвержденного приказом ректора от 07.07.2015 № 12-13-1282

Рабочая программа обсуждена на заседании кафедры прикладной математики, механики, управления и программного обеспечения, протокол № 7.2 от 21.07.2018 г.

Заведующая кафедрой прикладной математики, механики, управления и программного обеспечения д.т.н., профессор Артемяева И.Л.

Составитель: доцент кафедры прикладной математики, механики, управления и программного обеспечения, к.т.н. Смагин С.В.

Оборотная сторона титульного листа РПУД

I. Рабочая программа пересмотрена на заседании кафедры:

Протокол от «_____» _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) _____ (И.О. Фамилия)

II. Рабочая программа пересмотрена на заседании кафедры:

Протокол от «_____» _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) _____ (И.О. Фамилия)

ABSTRACT

Master's degree in 09.04.04 – Software engineering

Master's Program “Development of software and information systems”

Course title: Machine learning in artificial intelligence systems

Variable part of Block, 3 credits

Instructor: Smagin S.

At the beginning of the course a student should be able to: study independently, be self-organized; know about main concepts, principles, theories and facts related to computer science; carry out the search, the storage, the treatment and the analysis of information from various sources and databases represent this information in a required form with the help of information, computer and network technologies; work with operating systems, network technologies, program interface design tools, languages and methods of formal specifications, database management systems and also various technologies of software design

Learning outcomes: an ability to perceive mathematical, naturally-scientific, social and economic and professional knowledge, an ability to independently get, develop and use it for solving nonstandard problems including the problems from a new or unknown field and an interdisciplinary context; culture of thinking, an ability to form the logic of reasoning and statements based on an interpretation of data integrated from various spheres of science and technology, to make judgments using incomplete data; possession of methods and tools of receiving, storage, processing and broadcasting of information by means of modern computer technologies including global computer networks; possession of the existing methods and algorithms of solving the problems of data recognition and processing; possession of the existing methods and algorithms of solving the problems of digital signal processing

Course description: the main problems of machine learning, its methods and algorithms

Main course literature:

1. Flah P. Mashinnoe obuchenie. Nauka i iskusstvo postroeniya algoritmov, kotorye izvlekayut znaniya iz dannyh [Machine learning. The science and art of constructing algorithms that extract knowledge from data]– M.: DMK Press, 2015. – 400 p. (rus) – access: <http://www.studentlibrary.ru/book/ISBN9785970602737.html>
2. Koehl'o L.P., Richart V. Postroenie sistem mashinnogo obucheniya na yazyke Python. 2-e izdanie [Построение систем машинного обучения на

- языке Python. 2-е издание] / per. s angl. Slinkin A.A. – М.: DMK Press, 2016. – 302 p. (rus) – access: <http://www.studentlibrary.ru/book/ISBN9785970603307.html>
3. V'yugin, V.V. Matematicheskie osnovy mashinnogo obucheniya i prognozirovaniya [EHlektronnyy resurs]: ucheb. posobie [Mathematical foundations of machine learning and forecasting [Electronic resource]: Textbook. allowance]– М.: MCNMO, 2013. – 304 p. (rus) – access: <https://e.lanbook.com/book/56397>
 4. Rashka S. Python i mashinnoe obuchenie: krayne neobhodimoe posobie po noveyshey predskazatel'noy analitike, obyazatel'noe dlya bolee glubokogo ponimaniya metodologii mashinnogo obucheniya [EHlektronnyy resurs]: ruk. [Python and machine learning: an extremely necessary manual on the latest predictive analytics, mandatory for a deeper understanding of the methodology of machine learning [Electronic resource]:] / S. Rashka; per. s angl. Logunova A.V. – М.: DMK Press, 2017. – 418 p. (rus) – access: <https://e.lanbook.com/book/100905>
 5. Kulaichev A.P. Metody i sredstva kompleksnogo analiza dannyh: uchebnoe posobie dlya vuzov. [Methods and means of complex data analysis: a textbook for high schools] – М: [Forum]: INFRA-M, 2014. 511 p. (rus) – access: <https://lib.dvfu.ru:8443/lib/item?id=chamo:795113&theme=FEFU>

Form of final knowledge control: Pass-fail exam.

Аннотация рабочей программы учебной дисциплины «Машинное обучение в системах искусственного интеллекта»

Рабочая программа дисциплины «Машинное обучение в системах искусственного интеллекта» разработана для студентов 2 курса, обучающихся по направлению 09.04.04 Программная инженерия, магистерская программа «Разработка программно-информационных систем». Дисциплина является обязательной дисциплиной вариативной части учебного плана Б1.В.03.02.

Трудоемкость дисциплины 3 зачетных единиц (108 часов). Дисциплина реализуется на 2 курсе, в 3 семестре. Учебным планом предусмотрено: 18 часов лекций, 36 практических занятий (из них 18 часов в интерактивной форме), 54 часа самостоятельной работы.

Дисциплина «Машинное обучение в системах искусственного интеллекта» базируется на дисциплинах «Численные методы», «Теория вероятностей и математическая статистика», «Структуры и алгоритмы компьютерной обработки данных», изучаемых в бакалавриате. Знания, полученные при ее изучении, будут использованы при подготовке выпускных квалификационных работ.

Цель дисциплины – обзор основных задач обучения по прецедентам, изучение методов решения этих задач, а также алгоритмов, реализующих эти методы.

Задачи дисциплины:

1. Изучить основные понятия и примеры прикладных задач обучения по прецедентам в системах искусственного интеллекта.

2. Изучить критерии выбора моделей и методы отбора признаков при решении задач в интеллектуальных системах.

3. Изучить современные методы классификации (метрические методы, логические методы, линейные методы, вероятностные (байесовские) методы), а также методы кластеризации, используемые для решения задач в системах искусственного интеллекта.

Для успешного изучения дисциплины «Машинное обучение в системах искусственного интеллекта» у обучающихся должны быть сформированы следующие предварительные компетенции:

- способность к самоорганизации и самообразованию;
- владение основными концепциями, принципами, теориями и фактами, связанными с информатикой;
- способностью осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных, представлять

ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий;

- владение навыками использования операционных систем, сетевых технологий, средств разработки программного интерфейса, применения языков и методов формальных спецификаций, систем управления базами данных;
- владение навыками использования различных технологий разработки программного обеспечения.

Планируемые результаты обучения по данной дисциплине (знания, умения, владения), соотнесенные с планируемыми результатами освоения образовательной программы, характеризуют этапы формирования следующих компетенций (общекультурные/ общепрофессиональные/ профессиональные компетенции (элементы компетенций)):

Код и формулировка компетенции	Этапы формирования компетенции	
ОПК-1 способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе, в новой или незнакомой среде и в междисциплинарном контексте	Знает	различные модели и методы машинного обучения, используемые при решении задач в системах искусственного интеллекта
	Умеет	использовать и сравнивать различные модели и методы машинного обучения при решении задач в системах искусственного интеллекта
	Владеет	навыками применения программных средств при решении практических задач, связанных с машинным обучением
ОПК-2 культурой мышления, способностью выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных их разных областей науки и техники, выносить суждения на основании неполных данных	Знает	особенности выбора признаков моделей и предварительной обработки данных
	Умеет	формировать набор признаков модели и проводить предварительную обработку данных
	Владеет	технологиями оценивания и выявления информативных признаков модели
ПК-4 владением существующими методами и алгоритмами решения задач распознавания и обработки данных	Знает	основные методы машинного обучения, а также их достоинства и недостатки
	Умеет	применять методы машинного обучения для обработки данных
	Владеет	методами машинного обучения и подходами к их верификации

Для формирования вышеуказанных компетенций в рамках дисциплины «Машинное обучение в системах искусственного интеллекта» применяются следующие методы активного обучения: метод проектов.

I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

Лекционный материал (18 часов)

Тема 1. Введение, цели и задачи курса, основные термины (1.5 час.)

Рассматриваются понятия: модель предметной области, задача обучения по прецедентам, алгоритм обучения, решающее правило, критерий качества обучения, обучающая и контрольная выборки, модельные и реальные данные, объекты, признаки, свойства, классы, кластеры, типы шкал (бинарные, номинальные, порядковые, количественные).

Тема 2. Современная концепция анализа данных, разбор основных типов прикладных задач (1.5 час.)

Рассматриваются типы задач обучения с учителем: классификация, регрессия, ранжирование, прогнозирование.

Рассматриваются типы задач обучения без учителя: кластеризация, поиск ассоциативных правил, фильтрация выбросов, сокращение размерности, заполнение пропущенных значений.

Каждый тип задач сопровождается одним или несколькими детально расписанными примерами.

Тема 3. Формальные определения терминов, введение в проблему переобучение (1.5 час.)

Рассматриваются формальные (математические) определения терминов: модель алгоритмов, алгоритм обучения, этапы обучения, ошибка, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль, проблема переобучения.

Тема 4. Метрические методы классификации (1.5 час.)

Рассматриваются понятия: гипотеза компактности, метрика, метод ближайших соседей и его обобщения, окно Парзена и потенциальные функции, отбор эталонов (понятие отступа, алгоритм отбора эталонных объектов STOLP), отбор признаков и оптимизация метрики (задача выбора метрики, жадный алгоритм отбора признаков, полный скользящий контроль CCV).

Тема 5. Логические методы классификации (1.5 час.)

Рассматриваются понятия: закономерность, информативность, интерпретируемость, решающие деревья (алгоритм ID3, небрежные решающие деревья ODT, бинаризация данных).

Тема 6. Линейные методы классификации. Метод стохастического градиента (1.5 час.)

Рассматриваются понятия: минимизация эмпирического риска для градиентных методов обучения, линейный классификатор, метод стохастического градиента SG.

Тема 7. Линейные методы классификации. Метод опорных векторов (1.5 час.)

Рассматриваются понятия: принцип оптимальной разделяющей гиперплоскости, двойственная задача, понятие опорного вектора, обобщения линейного SVM (ядра и спрямляющие пространства, нейронные сети и SVM, обзор регуляризаторов для SVM).

Тема 8. Байесовские методы классификации (1.5 час.)

Рассматриваются понятия: оптимальный байесовский классификатор, вероятностная постановка задачи классификации, задача восстановления плотности распределения, наивный байесовский классификатор.

Тема 9. Методы кластеризации (6 час.)

Рассматриваются понятия: цели кластеризации, некорректность задачи кластеризации, типы кластерных структур, проблема чувствительности к выбору метрики, графовые методы кластеризации (алгоритм выделения связанных компонент, алгоритм ФОРЭЛ, функционалы качества кластеризации), иерархическая кластеризация (таксономия) (агломеративная иерархическая кластеризация, дендрограмма и свойство монотонности, свойства сжатия, растяжения и редуцируемости).

II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА

Практические занятия (36 час.)

Занятие 1. Метрические методы классификации (4 час.)

Реализация метода ближайшего соседа и метода k ближайших соседей.

Занятие 2. Метрические методы классификации (4 час.)

Реализация метода окна Парзена (фиксированной ширины и переменной ширины).

Занятие 3. Метрические методы классификации (4 час.)

Реализация метода потенциальных функций и алгоритма STOLP.

Занятие 4. Логические методы классификации (4 час.)

Реализация жадного алгоритма построения решающего дерева ID3.

Занятие 5. Логические методы классификации (4 час.)

Реализация алгоритма построения небрежных решающих деревьев ODT.

Занятие 6. Линейные методы классификации (4 час.)

Реализация метода стохастического градиента SG.

Занятие 7. Байесовские методы классификации (4 час.)

Реализация наивного байесовского классификатора.

Занятие 8. Методы кластеризации (4 час.)

Реализация алгоритма ФОРЭЛ (ФОРмальные Элементы).

Занятие 9. Методы кластеризации (4 час.)

Разбор функционалов качества кластеризации и реализация алгоритма Ланса-Уильямса.

Лабораторные работы (0 час.)

Не предусмотрено учебным планом.

Ш. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Трудоемкость самостоятельной работы 54 часов. Учебно-методическое обеспечение самостоятельной работы обучающихся представлено в Приложении 1 и включает в себя:

- план-график выполнения самостоятельной работы по дисциплине, в том числе примерные нормы времени на выполнение по каждому заданию;
- характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению;
- требования к представлению и оформлению результатов самостоятельной работы;
- критерии оценки выполнения самостоятельной работы.

III. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы/темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства – наименование	
				текущий контроль	промежуточная аттестация
1	Темы 1-3	ОПК-1	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 1-11 (теоретическая часть) и 1-3 (практическая)
			Умеет	Практические задания 1-3 ПР-6	
2	Темы 4-9	ОПК-2	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 4-6 (практическая часть)
			Умеет	Практические задания 4-5 ПР-6	
		ПК-4	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 7-10 (практическая часть)
			Умеет	Практические задания 6-7 ПР-6	
		ПК-4	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 11-13 (практическая часть)
			Умеет	Практические задания 8-9 ПР-6	

Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 2.

V. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература

1. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных – М.: ДМК Пресс,

2015. – 400 с. <http://www.studentlibrary.ru/book/ISBN9785970602737.html>
<https://e.lanbook.com/book/69955>
2. Коэльо Л.П., Ричарт В. Построение систем машинного обучения на языке Python. 2-е издание / пер. с англ. Слинкин А.А. – М.: ДМК Пресс, 2016. – 302 с.: ил.
<http://www.studentlibrary.ru/book/ISBN9785970603307.html>
 3. Вьюгин, В.В. Математические основы машинного обучения и прогнозирования [Электронный ресурс]: учеб. пособие – М.: МЦНМО, 2013. – 304 с. <https://e.lanbook.com/book/56397>
 4. Рашка С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения [Электронный ресурс]: рук. / С. Рашка; пер. с англ. Логунова А.В. – М.: ДМК Пресс, 2017. – 418 с. <https://e.lanbook.com/book/100905>
 5. Кулаичев А.П. Методы и средства комплексного анализа данных: учебное пособие для вузов. – М: [Форум]: ИНФРА-М, 2014. 511 с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:795113&theme=FEFU>

Дополнительная литература

1. Загоруйко Н. Г. Прикладные методы анализа данных и знаний. – Новосибирск: ИМ СО РАН, 1999. – 270 с.
<https://lib.dvfu.ru:8443/lib/item?id=chamo:10172&theme=FEFU>
2. Кухаренко Б.Г. Интеллектуальные системы и технологии – М.: Московская государственная академия водного транспорта, 2015. – 116 с. <http://www.iprbookshop.ru/47933.html>
3. Дедукция и обобщение в системах принятия решений / В.Н. Вагин. М: Наука, 1988. – 383 с.
<https://lib.dvfu.ru:8443/lib/item?id=chamo:55526&theme=FEFU>
4. Давнис В.В., Тинякова В.И., Мокшина С.И., Алексеева А.И. Компьютерные решения задач многомерной статистики. Часть 1. Кластерный и дискриминантный анализ: Учебное пособие. – Воронеж: Изд-во ВГУ, 2005. – 37 с. <http://window.edu.ru/resource/417/40417>
5. Симчера В.М. Методы многомерного анализа статистических данных: учеб. пособие. – М.: Финансы и статистика, 2008. – 400 с.: ил.
<http://www.studentlibrary.ru/book/ISBN9785279031849.html>

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. <http://machinelearning.ru/> MachineLearning.ru Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных.
2. http://shad.yandex.ru/lectures/machine_learning.xml Видеолекции курса «Машинное обучение» Школы анализа данных.
3. <http://www.studentlibrary.ru/book/ISBN9785940575061.html> Николенко С.И., Тулупьев А.Л. Самообучающиеся системы. – М.: МЦНМО, 2009. – 288 с.: 24 ил.
4. <http://znanium.com/catalog.php?bookinfo=451186> Информационные аналитические системы: учебник / Т.В. Алексеева, Ю.В. Амириди, В.В. Дик и др.; под ред. В.В. Дика. – М.: МФПУ Синергия, 2013. – 384 с. – (Университетская серия). – ISBN 978-5-4257-0092-6.

Перечень информационных технологий и программного обеспечения

Лекции проводятся с использованием проектора и мультимедийного комплекса для проведения лекций внутренней системы портала ДВФУ. Для составления документации используется текстовый процессор Microsoft Word, для презентаций используется Microsoft PowerPoint.

Для проведения практических занятий требуется следующее программное обеспечение: CPython, RGui, RStudio.

VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Дисциплина «Методы машинного обучения» изучается в следующих организационных формах: лекционное занятие; практическое занятие; самостоятельное изучение теоретического материала; самостоятельное выполнение индивидуального проекта; индивидуальные и групповые консультации.

Основной формой самостоятельной работы студента является изучение конспекта лекций, их дополнение рекомендованной литературой, выполнение проекта, а также активная работа на лабораторных занятиях.

К прослушиванию лекции следует готовиться, для этого необходимо знать программу курса и рекомендованную литературу. Тогда в процессе

лекции легче отделить главное от второстепенного, легче сориентироваться: что записать, что самостоятельно проработать, что является трудным для понимания, а что легко усвоить.

Контроль за выполнением самостоятельной работы студента производится в виде контроля каждого этапа работы, отраженного в документации и защиты проекта.

Студент должен планировать график самостоятельной работы по дисциплине и придерживаться его.

VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Лекции проводятся с использованием проектора и внутренней системы портала ДВФУ. Лабораторные занятия проходят в аудиториях, оборудованных компьютерами типа Lenovo C360G-i34164G500UDK с лицензионными программами Microsoft Office 2013 и аудиовизуальными средствами проектор Panasonic DLPProjectorPT-D2110XE, плазма LG FLATRON M4716CCBAM4716CJ. Для выполнения самостоятельной работы студенты в жилых корпусах ДВФУ обеспечены Wi-Fi.



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Дальневосточный федеральный университет»
(ДВФУ)

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

**УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ
ОБУЧАЮЩИХСЯ**

по дисциплине «Машинное обучение в системах искусственного интеллекта»

Направление подготовки – 09.04.04 Программная инженерия

Магистерская программа «Разработка программно-информационных систем»

Форма подготовки (очная)

**Владивосток
2018**

План-график выполнения самостоятельной работы по дисциплине

№ п/п	Формулировка задачи	Дата/сроки выполнения	Примерные нормы времени на выполнение	Форма контроля
1.	Формулировка задачи, выбор предметной области	7 неделя обучения	2 часа	Собеседование
2.	Разработка модели предметной области	8 неделя обучения	4 часа	Проверка отчетов, собеседование
3.	Исследование свойств модели предметной области	9 неделя обучения	4 часов	Собеседование
4.	Выделение набора информативных признаков	10 неделя обучения	4 часа	Собеседование
5.	Постановка задачи машинного обучения	11 неделя обучения	6 часа	Собеседование
6.	Выбор метода машинного обучения и анализ его свойств	12 неделя обучения	6 часов	Собеседование
7.	Формирование спецификации программного средства, в котором будет реализован выбранный метод машинного обучения	13 неделя обучения	6 часов	Проверка отчетов, собеседование
8.	Выбор/генерация обучающей выборки	14 неделя обучения	6 часов	Проверка отчетов, собеседование
9.	Реализация выбранного метода машинного обучения	15 неделя обучения	12 часов	Проверка отчетов, собеседование
10.	Оценка свойств метода	17 неделя обучения	4 часа	Проверка отчетов, собеседование
		Всего	54 часа	

Рекомендации по самостоятельной работе студентов

Трудоемкость самостоятельной работы 54 часов. Самостоятельная работа обучающихся подразумевает обязательную подготовку к лабораторным занятиям (оформление отчетов), изучение основной и дополнительно литературы по дисциплине, подготовку к текущему контролю и промежуточной аттестации в конце семестра, консультации преподавателя.

Рекомендации по работе с литературой

Для более эффективного освоения и усвоения материала рекомендуется ознакомиться с теоретическим материалом по той или иной теме до проведения лабораторного занятия. Всю учебную литературу желательно изучать «под конспект». Цель написания конспекта по дисциплине – сформировать навыки по поиску, отбору, анализу и формулированию учебного материала. При работе над конспектом обязательно выявляются и отмечаются трудные для самостоятельного изучения вопросы, с которыми уместно обратиться к преподавателю при посещении консультаций, либо в индивидуальном порядке.

Подготовка к лабораторным работам и практическим занятиям

Подготовку к практическому занятию студент должен начать с изучения теоретического материала и ознакомления с планом, который отражает содержание предложенной темы. Все новые понятия по изучаемой теме необходимо выучить наизусть и внести в глоссарий, который целесообразно вести с самого начала изучения курса. Результат такой работы должен проявиться в способности студента свободно ответить на теоретические вопросы по теме задания, и правильном его выполнении.

В процессе выполнения практического задания студент должен создать требуемый документ с помощью предлагаемого программного средства и выполнить требуемые в задании операции. Задание практической работе содержит методические указания по подготовке документа, который должен быть получен в результате выполнения работы. При подготовке следует их внимательно прочесть.



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Дальневосточный федеральный университет»
(ДВФУ)

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине «Машинное обучение в системах искусственного интеллекта»

Направление подготовки – 09.04.04 Программная инженерия

Магистерская программа «Разработка программно-информационных систем»

Форма подготовки (очная)

Владивосток
2018

Паспорт

фонда оценочных средств

по дисциплине «Машинное обучение в системах искусственного интеллекта»

Код и формулировка компетенции	Этапы формирования компетенции	
ОПК-1 способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе, в новой или незнакомой среде и в междисциплинарном контексте	Знает	технологии использования и сравнения различных моделей и методов машинного обучения при решении практических задач
	Умеет	ставить задачи на внедрение методов машинного обучения на предприятии
	Владеет	навыками применения программных средств при решении практических задач, связанных с машинным обучением
ОПК-2 культурой мышления, способностью выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных их разных областей науки и техники, выносить суждения на основании неполных данных	Знает	особенности выбора признаков моделей и предварительной обработки данных
	Умеет	формировать набор признаков модели и проводить предварительную обработку данных
	Владеет	технологиями оценивания и выявления информативных признаков модели
ПК-4 владением существующими методами и алгоритмами решения задач распознавания и обработки данных	Знает	основные методы машинного обучения, а также их достоинства и недостатки
	Умеет	применять методы машинного обучения для обработки данных
	Владеет	методами машинного обучения и подходами к их верификации

№ п/п	Контролируемые разделы/темы дисциплины	Коды и этапы формирования компетенций	Оценочные средства – наименование		
			текущий контроль	промежуточная аттестация	
1	Темы 1-3	ОПК-1	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 1-11 (теоретическая часть) и 1-3 (практическая)
			Умеет	Практические задания 1-3 ПР-6	
2	Темы 4-9	ОПК-2	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 4-6 (практическая)
			Умеет	Практические	

				задания 4-5 ПР-6	часть)
		ПК-4	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 7-10 (практическая часть)
			Умеет	Практические задания 6-7 ПР-6	
		ПК-4	Знает	Собеседование УО-1, круглый стол УО-4	Зачет Вопросы 11-13 (практическая часть)
			Умеет	Практические задания 8-9 ПР-6	

Шкала оценивания уровня сформированности компетенций

Код и формулировка компетенции	Этапы формирования компетенции		критерии	показатели
ОПК-1 способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе, в новой или незнакомой среде и в междисциплинарном контексте	знает (пороговый уровень)	технологии использования и сравнения различных моделей и методов машинного обучения при решении практических задач	Знание способов получения оценок моделей и методов машинного обучения, а также критериев их сравнения.	Способность выбрать один из набора методов машинного обучения для произвольной предметной области на основе оценки и анализа их свойств.
	умеет (продвинутый)	ставить задачи на внедрение методов машинного обучения на предприятии	Умение поставить задачу машинного обучения в заданной предметной области и предложить метод ее решения.	Способность проанализировать произвольную предметную область на предмет необходимости применения в ней методов машинного обучения, способность поставить и обосновать задачу такого применения.
	владеет (высокий)	навыками применения программных средств при решении практических задач, связанных с машинным обучением	Владение навыками работы с программными средствами и средами, реализующими методы машинного обучения в своих библиотеках.	Способность реализовать заданный метод машинного обучения при помощи программного средства, а также получить результат на основе выборки реальных данных.

ОПК-2 культурой мышления, способностью выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных их разных областей науки и техники, выносить суждения на основании неполных данных	знает (пороговый уровень)	особенности выбора признаков моделей и предварительной обработки данных	Знание основных принципов моделирования предметных областей.	Способность сформировать набор признаков упрощенной математической модели для произвольной предметной области.
	умеет (продвинутый)	формировать набор признаков модели и проводить предварительную обработку данных	Умение декомпозировать систему на подсистемы.	Способность выделить в системе подсистемы, способность для каждой подсистемы выделить наиболее характерные ее свойства и характеристики ее объектов.
	владеет (высокий)	технологиями оценивания и выявления информативных признаков модели	Владение навыками выделения и анализа совокупностей признаков, имеющих зависимости в изменении их значений, а также в их влиянии на конечный результат.	Способность выделить из группы признаков подгруппу, в наибольшей степени влияющей на результат решения конкретной задачи машинного обучения.
ПК-4 владением существующими методами и алгоритмами решения задач распознавания и обработки данных	знает (пороговый уровень)	основные методы машинного обучения, а также их достоинства и недостатки	Знание областей применения, преимуществ и недостатков основных методов машинного обучения.	Способность выбрать наиболее подходящий метод машинного обучения для произвольной предметной области.
	умеет (продвинутый)	применять методы машинного обучения для обработки данных	Умение реализовать заданный метод машинного обучения и применить его к заданной обучающей выборке.	Способность подготовить набор данных и обучить на его основе заданный метод машинного обучения.
	владеет (высокий)	методами машинного обучения и подходами к их верификации	Владение навыками формирования решающих правил и способами оценки их свойств (верификации).	Способность оценить решающее правило с точки зрения оценки времени и качества его работы на контрольной выборке.

Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины

Промежуточный контроль

Промежуточный контроль осуществляется в конце семестра и завершает изучение дисциплины. Помогает оценить более крупные совокупности знаний и умений, сформированность определенных профессиональных компетенций по дисциплине. Промежуточный контроль проводится в форме зачета, допуск к экзамену возможен для обучающихся, получивших оценку «зачтено» в результате выполнения самостоятельной работы и успешно выполнившие все лабораторные работы.

Критерии выставления оценки магистранту на зачете

Баллы (рейтинговой оценки)	Оценка (стандартная)	Требования к сформированным компетенциям
86-100	«зачтено»	Оценка «зачтено» выставляется магистранту, если он
76-85		-глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач;
61-75		- твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения;
		- имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ.

0-60	«незачтено»	Оценка «незачтено» выставляется магистранту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «незачтено» ставится магистрантам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.
------	-------------	---

Оценочные средства для промежуточной аттестации

Вопросы к зачету

Теоретическая часть:

1. Определение машинного обучения как направления исследований, его цели и задачи, особенности, основные области применения методов.
2. Типы обучения (дедуктивное, индуктивное, комбинированное). Типичный сценарий индуктивного обучения.
3. Признаковое описание объектов обучающей и контрольной выборок. Основные типы признаков и критерии их отбора.
4. Задача обучения по прецедентам, обобщающая способность, метод и алгоритм обучения, решающее правило, этапы обучения и применения.
5. Критерий качества обучения (функционал качества), ошибка, эмпирический риск, сведение задачи обучения к задаче оптимизации.
6. Классификатор, ошибки 1-ого и 2-ого рода, проблема переобучения, задача чемпионата мира среди алгоритмов классификации.
7. Основные типы задач обучения с учителем (классификация, регрессия, ранжирование, прогнозирование), их постановки и отличия.
8. Основные типы задач обучения без учителя (кластеризация, поиск ассоциативных правил, фильтрация выбросов, сокращение размерности, заполнение пропущенных значений), их постановки и отличия.
9. Другие типы задач: частичное обучение, обучение с подкреплением, динамическое обучение, активное обучение, метаобучение.
10. Примеры задачи классификации: медицинская диагностика, оценивание заемщиков, предсказание ухода клиента. Пример задачи восстановления регрессии: прогнозирование потребительского спроса.
11. Пример задачи ранжирования: ранжирование текстовых документов. Примеры задачи кластеризации: рубрикация текстов, рекомендации.

Практическая часть:

1. Метрические методы классификации: гипотеза компактности, метрика, метрическое пространство, обобщенный метрический классификатор.
2. Метрические методы классификации: метод ближайшего соседа, метод к ближайших соседей, метод окна Парзена, метод потенциальных функций.
3. Метрические методы классификации: понятие отступа, типы объектов (в зависимости от отступа), алгоритм STOLP.
4. Логические методы классификации: логическая закономерность, интерпретируемость, информативность, часто используемые виды закономерностей, часто используемые критерии информативности, обобщенный алгоритм нахождения информативных закономерностей.
5. Логические методы классификации: определение бинарного решающего дерева, пример решающего дерева, жадный алгоритм построения дерева ID3, достоинства и недостатки решающих деревьев ID3.
6. Логические методы классификации: определение бинарного решающего дерева, редукция дерева, небрежные решающие деревья ODT, задача бинаризации вещественного признака.
7. Линейные методы классификации (метод стохастического градиента): задача построения разделяющей поверхности, линейный классификатор, математическая модель нейрона, градиентный метод численной минимизации, алгоритм SG.
8. Линейные методы классификации (метод опорных векторов): кусочно-линейная аппроксимация, задача SVM, оптимальная разделяющая гиперплоскость, достоинства и недостатки метода.
9. Байесовские методы классификации: байесовский классификатор апостериорная вероятность, вероятностная постановка задачи, функционал среднего риска, оптимальный байесовский классификатор.
10. Байесовские методы классификации: разбиение задачи классификации на две подзадачи и алгоритм решение первой из них, три основных подхода к оцениванию плотностей, наивный байесовский классификатор.

11. Методы кластеризации: постановка задачи кластеризации, некорректность задачи кластеризации, цели кластеризации, типы кластерных структур, проблема чувствительности к выбору метрики.
12. Методы кластеризации: графовые методы кластеризации (алгоритм выделения связных компонент, алгоритм кратчайшего незамкнутого пути, алгоритм ФОРЭЛ, функционалы качества кластеризации).
13. Методы кластеризации: агломеративная иерархическая кластеризация (алгоритмы Ланса-Уильямса (обычный и редуцированный), формула Ланса-Уильямса и ее частные случаи, визуализация кластерной структуры).

Критерии оценки проектов

100-86 баллов выставляется, если магистрант/группа точно определили содержание и составляющие части задания, умеют аргументировано отвечать на вопросы, связанные с заданием. Продемонстрировано знание и владение навыками самостоятельной исследовательской работы по теме. Фактических ошибок, связанных с пониманием проблемы, нет.

- 85-76 - баллов - работа магистранта/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

- 75-61 балл – проведен достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимание базовых основ и теоретического обоснования выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы

- 60-50 баллов - если работа представляет собой пересказанный или полностью переписанный исходный текст без каких бы то ни было комментариев, анализа. Не раскрыта структура и теоретическая составляющая темы. Допущено три или более трех ошибок смыслового содержания раскрываемой проблемы

Шкала оценивания проектов

Менее 60 баллов	Не зачтено
От 61 до 75 баллов	зачтено
От 76 до 85 баллов	зачтено
От 86 до 100 баллов	зачтено

Текущий контроль

Текущий контроль предполагает систематическую проверку усвоения учебного материала, сформированности компетенций или их элементов, регулярно осуществляемую на протяжении изучения дисциплины, в соответствии с ее рабочей программой.

Состоит в проверке правильности выполнения заданий по самостоятельной работе. Задание зачтено, если нет ошибок. По текущим ошибкам даются пояснения.

Тесты предназначены для проверки знаний по компетенциям. Проверка достижения умений и навыков по компетенциям проверяется выполнением практических работ и курсовой работы.

Примерные тесты для проверки сформированности компетенций

ОПК-1 способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе, в новой или незнакомой среде и в междисциплинарном контексте	Знает модели и методы машинного обучения, применяемые при решении практических задач.
1. В типологии задач обучения по прецедентам «Обучение без учителя» предполагает, что:	Ответы: а. Для каждого объекта обучающей выборки известен выход (класс). б. Для объектов обучающей выборки выходы (классы) не известны. в. Допустимым ответом является действительное число или числовой вектор. г. Ответы надо получить сразу на множестве объектов.
2. В типологии задач обучения по прецедентам «Частичное обучение» предполагает, что:	Ответы: а. Обучение происходит с частичным привлечением учителя, когда прогноз предполагается делать только для прецедентов из тестовой выборки. б. Обучаемый имеет возможность самостоятельно назначать следующий прецедент, который станет известен. в. Каждый прецедент представляет собой пару «объект, ответ», но ответы известны не для всех прецедентов. г. Прецедентами являются ранее решенные задачи обучения.

<p>ОПК-2 культурой мышления, способностью выстраивать логику рассуждений и высказываний, основанных на интерпретации данных, интегрированных их разных областей науки и техники, выносить суждения на основании неполных данных</p>	<p>Знает особенности выбора признаков моделей и предварительной обработки данных.</p>
<p>1. Объекты могут описываться (можно выбрать один или несколько вариантов):</p>	<p>Ответы: а. Совокупностью числовых показателей, измеряемых у всех объектов. б. Временными рядами, сигналами, изображениями, видеорядами, текстами, попарными отношениями сходства или интенсивности взаимодействия. в. Чем угодно.</p>
<p>2. Какими могут быть цели решения задачи кластеризации (можно выбрать один или несколько вариантов):</p>	<p>Ответы: а. Выделить наиболее типичные объекты. б. Упростить дальнейшую обработку данных. в. Повысить уровень доверия к данным. г. Сократить объем хранимых данных.</p>

<p>ПК-4 владением существующими методами и алгоритмами решения задач распознавания и обработки данных</p>	<p>Знает основные методы машинного обучения, а также их достоинства и недостатки.</p>
<p>1. Укажите метрические алгоритмы классификации (можно выбрать один или несколько вариантов):</p>	<p>Ответы: а. Алгоритм SG. б. Алгоритм ФОРЭЛ. в. Окно Парзена переменной ширины. г. Алгоритм STOLP.</p>
<p>2. Укажите логические алгоритмы классификации (можно выбрать один или несколько вариантов):</p>	<p>Ответы: а. Алгоритм ODT. б. Наивный байесовский классификатор. в. Жадный алгоритм ID3. г. Алгоритм SVM.</p>