

МИНИСТЕРСТВО ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«РОССИЙСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
МЕДИЦИНСКИЙ УНИВЕРСИТЕТ ИМЕНИ Н.И. ПИРОГОВА»**
МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
(ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России)

СОГЛАСОВАНО

Декан факультета подготовки
кадров высшей квалификации
ФГАОУ ВО РНИМУ
им. Н.И. Пирогова Минздрава России

_____ М.В. Хорева

«23» июня 2022 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)
«МАШИННОЕ ОБУЧЕНИЕ»**

Научная специальность

1.5.8 Математическая биология, биоинформатика

Москва, 2022 г.

Рабочая программа дисциплины «Машинное обучение» разработана в соответствии с Федеральными государственными требованиями, утверждёнными приказом Министерства науки и высшего образования Российской Федерации от 20.10.2021 № 951, педагогическими работниками кафедры биоинформатики МБФ.

№	Фамилия, имя, отчество	Ученая степень, звание	Занимаемая должность в Университете, кафедра
1	Лагунин Алексей Александрович	д-р биол. наук, проф. РАН	Заведующий кафедрой биоинформатики МБФ
2	Иванов Сергей Михайлович	канд. биол. наук	Доцент кафедры биоинформатики МБФ
3	Муравьева Елена Степановна	канд. биол. наук	Доцент кафедры биоинформатики МБФ

Рабочая программа дисциплины «Машинное обучение» рассмотрена и одобрена на заседании кафедры биоинформатики МБФ.

протокол № 7 от «28» апреля 2022 г.

Заведующий кафедрой _____ /А.А. Лагунин/

ОГЛАВЛЕНИЕ

1. Цель и задачи изучения дисциплины (модуля)	4
2. Объем дисциплины (модуля) по видам учебной работы.....	4
3. Содержание дисциплины (модуля).....	4
4. Учебно-тематический план дисциплины (модуля)	4
5. Учебно-методическое обеспечение самостоятельной работы обучающихся	5
6. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся	5
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)	8
8. Материально-техническое обеспечение дисциплины (модуля)	9
9. Методические указания для обучающихся по изучению дисциплины (модуля)	9
10. Методические рекомендации преподавателю по организации учебного процесса по дисциплине (модулю).....	10

1. Цель и задачи изучения дисциплины (модуля)

Цель изучения дисциплины (модуля)

Изучение дисциплины (модуля) «Машинное обучение» предоставляет получение основополагающих знаний и практических навыков программирования на языке R, знаний о методах машинного обучения, опыта применения полученных знаний для анализа молекулярно-биологических и клинико-диагностических данных для подготовки научных и научно-педагогических кадров высшей квалификации по научной специальности 1.5.8 Математическая биология, биоинформатика.

Задачи дисциплины (модуля)

1. Обучить программированию на языке R, включая синтаксис языка, операторы, базовые функции, импорт и сохранение информации, построение графиков и диаграмм;
2. Обучить методам машинного обучения (метод опорных векторов, искусственные нейронные сети, деревья решений, методы кластеризации данных и др.), реализованных в R, и их применению для анализа молекулярно-биологических и клинико-диагностических данных.

2. Объем дисциплины (модуля) по видам учебной работы

Таблица 1

Виды учебной работы	Всего, час.	Объем по полугодиям							
		1	2	3	4	5	6	7	8
Контактная работа обучающегося с преподавателем по видам учебных занятий (Контакт. раб.):	36	-	-	-	36	-	-	-	-
Лекционное занятие (Л)	18	-	-	-	18	-	-	-	-
Семинарское/практическое занятие (СПЗ)	18	-	-	-	18	-	-	-	-
Самостоятельная работа обучающегося, в том числе подготовка к промежуточной аттестации (СР)	36	-	-	-	40	-	-	-	-
Вид промежуточной аттестации: Зачет (З), Зачет с оценкой (ЗО), Экзамен (Э), Кандидатский экзамен (КЭ)	<i>Зачет</i>	-	-	-	3	-	-	-	-
Общий объем	в часах	72	-	-	-	72	-	-	-
	в зачетных единицах	2	-	-	-	2	-	-	-

3. Содержание дисциплины (модуля)

Раздел 1. Методы машинного обучения.

- 1.1. Введение в машинное обучение. Методы классификации и регрессии.
- 1.2. Нейронные сети.
- 1.3. Методы кластеризации и снижения размерности пространства.

4. Учебно-тематический план дисциплины (модуля)

Таблица 2

Номер раздела, темы	Наименование разделов, тем	Количество часов					Форма контроля
		Всего	Конт. акт. раб.	Л	СПЗ	СР	

	Полугодие 4	72	36	18	18	36	Зачет
Раздел 1	Методы машинного обучения	72	36	18	18	36	Доклад с защитой презентации
Тема 1.1	Введение в машинное обучение. Методы классификации и регрессии.	32	16	8	8	16	
Тема 1.2	Нейронные сети.	24	12	6	6	12	
Тема 1.3	Методы кластеризации и снижения размерности пространства.	16	8	4	4	8	
	Общий объем	72	36	18	18	36	

5. Учебно-методическое обеспечение самостоятельной работы обучающихся

Цель самостоятельной работы обучающихся заключается в глубоком, полном усвоении учебного материала и в развитии навыков самообразования. Самостоятельная работа может включать: работу с текстами, литературой, учебно-методическими пособиями, нормативными материалами, в том числе материалами сети интернет, а также проработку конспектов лекций, написание докладов, рефератов, участие в работе семинаров, научных конференциях и пр.

Задания для самостоятельной работы

Таблица 3

Номер раздела	Наименование раздела	Вопросы для самостоятельной работы
1	Методы машинного обучения	Дерево решений. Random Forest. Regression Forest. Метод опорных векторов, метод К ближайших соседей. Отбор признаков. Наивный байесовский классификатор. ROC-кривая. Обобщенные линейные модели. Классическая линейная регрессия. Логистическая, мультиномиальная, порядковая регрессия. Понятие о регуляризации. Регуляризованная регрессия. Регуляризованный метод опорных векторов. Улучшенные методы на основе деревьев решений. Регуляризованный и «наклонный» Random Forest. Градиентный бустинг. Принципы работы с несбалансированными выборками. Нейронные сети. Многослойный персептрон. Бинарная, многоклассовая классификация. Регрессия. Регуляризация и прореживание. Подбор гиперпараметров. Глубокое обучение. Сверточные нейронные сети. Рекуррентные нейронные сети. Иерархическая кластеризация. Метод К-средних. Метод главных компонент. EM алгоритм. Самоорганизующиеся карты Кохонена. Многомерное шкалирование.

Контроль самостоятельной работы осуществляется на семинарских (практических) занятиях.

6. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся

Примерные варианты оценочных заданий для текущего контроля успеваемости

Таблица 4

Раздел, тема	Наименование разделов, тем	Форма контроля	Оценочное задание
	Полугодие 4		
Раздел 1	Методы машинного	Доклад с	Выполнить индивидуальный проект по созданию

	обучения	защитой презентац ии	моделей с использованием методов машинного обучения.
Тема 1.1	Введение в машинное обучение. Методы классификации и регрессии.		
Тема 1.2	Нейронные сети.		
Тема 1.3	Методы кластеризации и снижения размерности пространства.		

Вопросы для подготовки к промежуточной аттестации зачету

1. Методы классификации.
2. Методы регрессии.
3. Нейронные сети.
4. Многослойный перцептрон.
5. Глубокое обучение.
6. Сверточные нейронные сети.
7. Анализ изображений.
8. Глубокое обучение.
9. Рекуррентные нейронные сети.
10. Анализ текста.
11. Методы кластеризации.
12. Методы снижения размерности пространства.
13. Метод дерева решений.
14. Метод Random Forest.
15. Метод Regression Forest.
16. Метод опорных векторов.
17. Метод К ближайших соседей.
18. Отбор признаков.
19. Наивный байесовский классификатор.
20. ROC-кривая.
21. Обобщенные линейные модели.
22. Классическая линейная регрессия.
23. Логистическая регрессия.
24. Мультиномиальная регрессия.
25. Порядковая регрессия.
26. Понятие о регуляризации. Регуляризованная регрессия.
27. Регуляризованный метод опорных векторов.
28. Улучшенные методы на основе деревьев решений. Регуляризованный и «наклонный» Random Forest.
29. Градиентный бустинг.
30. Принципы работы с несбалансированными выборками.
31. Бинарная, многоклассовая классификация. Регрессия.
32. Регуляризация и прореживание. Подбор гиперпараметров.
33. Глубокое обучение. Рекуррентные нейронные сети.
34. Иерархическая кластеризация. Метод К-средних.
35. Метод главных компонент.

36. EM алгоритм.
37. Самоорганизующиеся карты Кохонена.
38. Многомерное шкалирование.

Описание критериев и шкал оценивания

В ходе текущего контроля успеваемости (устный или письменный опрос, подготовка и защита реферата, доклад, презентация, тестирование и пр.) при ответах на учебных занятиях, а также промежуточной аттестации в форме кандидатского экзамена обучающиеся оцениваются по четырёхбалльной шкале: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценка «отлично» – выставляется аспиранту, если он глубоко усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет связывать теорию с практикой, свободно справляется с задачами и вопросами, не затрудняется с ответами при видоизменении заданий, умеет принять правильное решение и грамотно его обосновывать, владеет разносторонними навыками и приемами выполнения практических задач, комплексной оценкой предложенной ситуации.

Оценка «хорошо» – выставляется аспиранту, если он твердо знает программный материал, грамотно и по существу излагает его, не допуская существенных неточностей при ответе на вопрос, но недостаточно полно раскрывает междисциплинарные связи, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения, комплексной оценкой предложенной ситуации.

Оценка «удовлетворительно» – выставляется аспиранту, если он имеет поверхностные знания программного материала, не усвоил его деталей, допускает неточности, оперирует недостаточно правильными формулировками, нарушает логическую последовательность в изложении программного материала, испытывает затруднения при выполнении практических задач, испытывает затруднения с комплексной оценкой предложенной ситуации, не полностью отвечает на вопросы, в том числе при помощи наводящих вопросов преподавателя.

Оценка «неудовлетворительно» – выставляется аспиранту, который не знает значительной части программного материала, допускает грубые ошибки, неуверенно, с большими затруднениями решает практические задачи или не справляется с ними самостоятельно, не владеет комплексной оценкой ситуации, неверно выбирает тактику действий.

В ходе текущего контроля успеваемости (устный или письменный опрос, подготовка и защита реферата, доклад, презентация, тестирование и пр.) при ответах на учебных занятиях, а также промежуточной аттестации в форме зачета обучающиеся оцениваются по двухбалльной шкале:

Оценка «зачтено» – выставляется аспиранту, если он продемонстрировал знания программного материала, подробно ответил на теоретические вопросы, справился с выполнением заданий и (или) ситуационных задач, предусмотренных рабочей программой дисциплины (модуля).

Оценка «не зачтено» – выставляется аспиранту, если он имеет пробелы в знаниях программного материала, не владеет теоретическим материалом и допускает грубые, принципиальные ошибки в выполнении заданий и (или) ситуационных задач, предусмотренных рабочей программой дисциплины (модуля).

Шкала оценивания (четырёхбалльная или двухбалльная), используемая в рамках текущего контроля успеваемости определяется преподавателем, исходя из целесообразности применения той или иной шкалы.

Если текущий контроль успеваемости и (или) промежуточная аттестация, предусматривает тестовые задания, то перевод результатов тестирования в четырёхбалльную шкалу осуществляется по схеме:

Оценка «Отлично» – 90-100% правильных ответов;

Оценка «Хорошо» – 80-89% правильных ответов;

Оценка «Удовлетворительно» – 71-79% правильных ответов;

Оценка «Неудовлетворительно» – 70% и менее правильных ответов.

Перевод результатов тестирования в двухбалльную шкалу:

Оценка «Зачтено» – 71-100% правильных ответов;

Оценка «Не зачтено» – 70% и менее правильных ответов.

7. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Таблица 5

№ п/п	Автор, наименование, место издания, издательство, год издания	Количество экземпляров
1	Буховец, А. Г. Алгоритмы вычислительной статистики в системе R. Санкт-Петербург : Лань, 2015. https://e.lanbook.com/book/68459	Удаленный доступ
2	Балдин К.В., Башлыков В.Н., Рукосуев А.В. Теория вероятностей и математическая статистика, 2-е изд. – Москва: Дашков и К, 2014. https://ibooks.ru/reading.php?productid=34258	Удаленный доступ

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Официальный сайт РНИМУ: адрес ресурса – <https://rsmu.ru.ru/>, на котором содержатся сведения об образовательной организации и ее подразделениях, локальные нормативные акты, сведения о реализуемых образовательных программах, их учебно-методическом и материально-техническом обеспечении, а также справочная, оперативная и иная информация. Через официальный сайт обеспечивается доступ всех участников образовательного процесса к различным сервисам и ссылкам, в том числе к Автоматизированной системе подготовки кадров высшей квалификации (далее – АСПКВК);
2. ЭБС РНИМУ им. Н.И. Пирогова – Электронная библиотечная система;
3. ЭБС IPRbooks – Электронно-библиотечная система;
4. ЭБС Айбукс – Электронно-библиотечная система;
5. ЭБС Букап – Электронно-библиотечная система;
6. ЭБС Лань – Электронно-библиотечная система;
7. ЭБС Юрайт – Электронно-библиотечная система.

Перечень профессиональных баз данных и информационных справочных систем

1. <http://www.consultant.ru> - Консультант студента, компьютерная справочная правовая система в РФ;
2. <https://www.garant.ru> - Гарант.ру, справочно-правовая система по

законодательству Российской Федерации;

3. <https://www.r-project.org/> - основной сайт R;
4. <https://cran.r-project.org/> - архив пакетов для статистического анализа данных в R;
5. <http://www.bioconductor.org/> - Bioconductor – архив пакетов R, предназначенных для анализа молекулярно-биологических данных.

8. Материально-техническое обеспечение дисциплины (модуля)

Таблица 6

№ п/п	Наименование оборудованных учебных аудиторий	Перечень специализированной мебели, технических средств обучения
1	Учебные аудитории для проведения занятий лекционного и семинарского типов, групповых и индивидуальных консультаций, текущего контроля успеваемости и промежуточной аттестации	Персональные компьютеры с доступом в Интернет.
2	Помещения для самостоятельной работы (Библиотека, в том числе читальный зал)	Компьютерная техника с возможностью подключения к сети "Интернет" и обеспечением доступа к электронной информационно-образовательной среде РНИМУ.

Программное обеспечение

- MICROSOFT WINDOWS 7, 10;
- OFFICE 2010, 2013;
- Антивирус Касперского (Kaspersky Endpoint Security);
- ADOBE CC;
- Photoshop;
- Консультант плюс (справочно-правовая система);
- iSpring;
- Adobe Reader;
- Adobe Flash Player;
- Google Chrom, Mozilla Firefox, Mozilla Public License;
- 7-Zip;
- FastStone Image Viewer.

9. Методические указания для обучающихся по изучению дисциплины (модуля)

Преподавание дисциплины (модуля) осуществляется в соответствии с Федеральными государственными требованиями.

Основными формами получения и закрепления знаний по данной дисциплине (модулю) являются занятия лекционного и семинарского типа, самостоятельная работа обучающегося, в том числе под руководством преподавателя, прохождение контроля.

Изучение дисциплины (модуля) согласно учебному плану предполагает самостоятельную работу обучающихся. Самостоятельная работа включает в себя изучение литературы, её конспектирование, подготовку к семинарским (практическим) занятиям, текущему контролю успеваемости и промежуточной аттестации.

Текущий контроль успеваемости по дисциплине (модулю) и промежуточная аттестация осуществляются в соответствии с Порядком организации и проведения

текущего контроля успеваемости и Порядком проведения промежуточной аттестации обучающихся, устанавливающим формы проведения промежуточной аттестации, ее периодичность и систему оценок.

Наличие в Университете электронной информационно-образовательной среды, а также электронных образовательных ресурсов позволяет изучать дисциплину (модуль) инвалидам и лицам с ОВЗ.

Особенности изучения дисциплины (модуля) инвалидами и лицами с ОВЗ определены в Положении об организации получения образования для инвалидов и лиц с ограниченными возможностями здоровья.

10. Методические рекомендации преподавателю по организации учебного процесса по дисциплине (модулю)

Преподавание дисциплины (модуля) осуществляется в соответствии с Федеральными государственными требованиями.

При изучении дисциплины (модуля) рекомендуется использовать следующий набор средств и способов обучения:

- рекомендуемую литературу;
- задания для подготовки к семинарам (практическим занятиям) – вопросы для обсуждения и др.;
- задания для текущего контроля успеваемости (задания для самостоятельной работы обучающихся);
- вопросы и задания для подготовки к промежуточной аттестации по итогам изучения дисциплины (модуля).

При проведении занятий лекционного и семинарского типа, в том числе в форме вебинаров и on-line курсов необходимо строго придерживаться учебно-тематического плана дисциплины (модуля), приведенного в разделе 4 данного документа. Необходимо уделить внимание рассмотрению вопросов и заданий, включенных в оценочные задания, при необходимости, решить аналогичные задачи с объяснением алгоритма решения.

Следует обратить внимание обучающихся на то, что для успешной подготовки к текущему контролю успеваемости и промежуточной аттестации нужно изучить литературу, список которой приведен в разделе 7 данной рабочей программы дисциплины (модуля) и иные источники, рекомендованные в подразделах «Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и «Перечень профессиональных баз данных и информационных справочных систем», необходимых для изучения дисциплины (модуля).

Текущий контроль успеваемости и промежуточная аттестация осуществляются в соответствии с Порядком организации и проведения текущего контроля успеваемости и Порядком проведения промежуточной аттестации обучающихся, устанавливающим формы проведения промежуточной аттестации, ее периодичность и систему оценок, с которыми необходимо ознакомить обучающихся на первом занятии.