

МИНОБРНАУКИ РОССИИ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ

Рабочая программа дисциплины (модуля)
МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ НА ЯЗЫКЕ R

Направление и направленность (профиль)
38.04.05 Бизнес-информатика. Информационная бизнес-аналитика

Год набора на ОПОП
2022

Форма обучения
очная

Владивосток 2023

Рабочая программа дисциплины (модуля) «Методы машинного обучения на языке R» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 38.04.05 Бизнес-информатика (утв. приказом Минобрнауки России от 12.08.2020г. №990) и Порядком организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры (утв. приказом Минобрнауки России от 06.04.2021 г. N245).

Составитель(и):

Назаров Д.А., кандидат технических наук, доцент, Кафедра информационных технологий и систем

Утверждена на заседании кафедры информационных технологий и систем от 31.05.2023 , протокол № 9

СОГЛАСОВАНО:

Заведующий кафедрой (разработчика)

Свяжина А.С.

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ	
Сертификат	1600852794
Номер транзакции	0000000000BBE21F
Владелец	Свяжина А.С.

1 Цель, планируемые результаты обучения по дисциплине (модулю)

Цели освоения учебной дисциплины:

- изучение теоретических основ и алгоритмов машинного обучения;
- формирование умений и навыков автоматизации решения профессиональных задач в различных областях человеческой деятельности на основе интеллектуального анализа данных.

Задачи освоения дисциплины:

- расширение и углубление теоретических знаний о качественных особенностях различных систем, количественных взаимосвязях и закономерностях их развития;
- овладение методологией построения, анализа и использования регрессионных моделей, как для анализа состояния, так и для прогнозных оценок развития изучаемых систем;
- изучение алгоритмов кластеризации и классификации объектов;
- получение навыков реализации методов машинного обучения с учителем и без учителя на языке R в среде Rstudio.

Планируемыми результатами обучения по дисциплине (модулю), являются знания, умения, навыки. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы, представлен в таблице 1.

Таблица 1 – Компетенции, формируемые в результате изучения дисциплины (модуля)

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине		
			Код результата	Формулировка результата	
38.04.05 «Бизнес-информатика» (М-БИ)	ОПК-3 : Способен принимать решения, осуществлять стратегическое планирование и прогнозирование в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта	ОПК-3.1к : Использует системы искусственного интеллекта в рамках стратегического планирования и прогнозирования в профессиональной деятельности	РД1	Знание	аналитических методов обработки статистических данных
			РД2	Умение	применять методы машинного обучения при решении нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
		ОПК-3.2к : Использует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными для поддержки стратегических управленческих решений	РД3	Знание	методов разработки алгоритмов на языке R для решения профессиональных задач
			РД4	Умение	разрабатывать оригинальные алгоритмы с использованием языка программирования R для решения профессиональных задач

	ПКВ-2 : Способен выбирать и осуществлять рациональные решения по ИС и ИКТ для управления организациями на основе больших данных	ПКВ-2.1к : Использует и совершенствует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными	РД5	Знание	методов исследования современных проблем и методов прикладной информатики
		ПКВ-2.2к : Разрабатывает технические проекты в сфере ИТ, позволяющие анализировать и использовать большие данные в управлении организацией	РД6	Умение	проводить исследования задач различных классов, используя современные методы, средства прикладной информатики и проводить их анализ
			РД7	Навык	использования инструментальных средств при разработке программных средств, основанных на методах машинного обучения

2 Место дисциплины (модуля) в структуре ОПОП

Дисциплина относится к обязательной части блока 1 дисциплин учебного плана направления.

Входными требованиями, необходимыми для освоения дисциплины, является наличие у обучающихся компетенций, сформированных на предыдущем уровне образования.

3. Объем дисциплины (модуля)

Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу с обучающимися (по видам учебных занятий) и на самостоятельную работу, приведен в таблице 2.

Таблица 2 – Общая трудоемкость дисциплины

Название ОПОП ВО	Форма обучения	Часть УП	Семестр (ОФО) или курс (ЗФО, ОЗФО)	Трудо-емкость (З.Е.)	Объем контактной работы (час)					СРС	Форма аттестации	
					Всего	Аудиторная			Внеаудиторная			
						лек.	прак.	лаб.	ПА			КСР
38.04.05 Бизнес-информатика	ОФО	М01.Б	1	4	55	18	36	0	1	0	89	Э

4 Структура и содержание дисциплины (модуля)

4.1 Структура дисциплины (модуля) для ОФО

Тематический план, отражающий содержание дисциплины (перечень разделов и тем), структурированное по видам учебных занятий с указанием их объемов в соответствии с учебным планом, приведен в таблице 3.1

Таблица 3.1 – Разделы дисциплины (модуля), виды учебной деятельности и формы текущего контроля для ОФО

№	Название темы	Код результата обучения	Кол-во часов, отведенное на				Форма текущего контроля
			Лек	Практ	Лаб	СРС	
1	Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	РД1, РД2, РД3, РД4, РД5, РД6, РД7	0	7	0	27	практическая работа № 1
2	Множественная регрессия и корреляция	РД1, РД2, РД3, РД4, РД5, РД6, РД7	0	7	0	30	практическая работа № 2
3	Снижение размерности. Метод главных компонент	РД1, РД2, РД3, РД4, РД5, РД6, РД7	0	7	0	26	практическая работа №3
4	Задачи классификации. Модели с дискретной зависимой переменной	РД1, РД2, РД3, РД4, РД5, РД6, РД7	0	8	0	30	практическая работа № 4
5	Кластерный анализ	РД1, РД2, РД3, РД4, РД5, РД6, РД7	0	7	0	30	практическая работа № 5
Итого по таблице			0	36	0	143	

4.2 Содержание разделов и тем дисциплины (модуля) для ОФО

Тема 1 Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.

Содержание темы: Общая постановка задачи машинного обучения. Способы машинного обучения. Классические задачи машинного обучения. Регрессия. Спецификация модели, связанная с выбором факторов и вида математической функции. Методы выбора математической функции в парной регрессии. Линейная регрессия. Оценка параметров. Оценка значимости уравнения и параметров. Нелинейная регрессия. Ошибки аппроксимации.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: стандартная, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практической работе, подготовка к итоговому тесту.

Тема 2 Множественная регрессия и корреляция.

Содержание темы: Методы построения моделей и спецификация ее систематической части. Мультиколлинеарность. Оценка параметров моделей и их диагностика. Качественные объясняющие переменные. Анализ вклада отдельных предикторов и селекция их оптимальной комбинации. Оценка качества каждой модели-претендента. Предпосылки МНК.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: стандартная, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практической работе, подготовка к итоговому тесту.

Тема 3 Снижение размерности. Метод главных компонент.

Содержание темы: Отбор и выделение признаков. Математическая модель метода главных компонент. Оценка числа главных компонент. Оценка уровня информативности и подбор названий для главных компонент.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: стандартная, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практической работе, подготовка к итоговому тесту.

Тема 4 Задачи классификации. Модели с дискретной зависимой переменной.

Содержание темы: Модели бинарного выбора. Логит-модель, пробит-модель. Оценивание параметров модели. Метод максимального правдоподобия. Модели множественного выбора с неупорядоченными альтернативами. Модели множественного выбора с упорядоченными альтернативами.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: стандартная, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практической работе, подготовка к итоговому тесту.

Тема 5 Кластерный анализ.

Содержание темы: Формальное описание задачи кластеризации. Математические характеристики кластера. Меры близости: меры расстояния, меры сходства, меры предсказуемости вероятностные коэффициенты сходства. Алгоритмы кластерного анализа. Иерархическая кластеризация. Алгоритмы неиерархического разделения. Метод нечетких средних. Оценка качества кластеризации.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: стандартная, метод кооперативного обучения.

Виды самостоятельной подготовки студентов по теме: подготовка к практической работе, подготовка к итоговому тесту.

5 Методические указания для обучающихся по изучению и реализации дисциплины (модуля)

5.1 Методические рекомендации обучающимся по изучению дисциплины и по обеспечению самостоятельной работы

На первом шаге выполнения практических работ необходимо сформировать массив данных из указанных источников, согласно рекомендациям и/или загрузить имеющиеся данные в RStudio.

Далее следует выполнить задания к практической работе, сделать необходимые выводы и сформировать требуемые отчеты.

При подготовке к текущей аудиторной работе рекомендуется использовать материалы лекций, источники из списка учебной литературы, разобрать примеры скриптов, документацию библиотек (пакетов), проанализировать приемы работы с Rstudio.

5.2 Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

При необходимости обучающимся из числа лиц с ограниченными возможностями здоровья и инвалидов (по заявлению обучающегося) предоставляется учебная информация в доступных формах с учетом их индивидуальных психофизических особенностей:

- для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания, консультации и др.

- для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания, консультации и др.

- для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме

электронного документа; индивидуальные задания, консультации и др.

6 Фонд оценочных средств для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине (модулю)

В соответствии с требованиями ФГОС ВО для аттестации обучающихся на соответствие их персональных достижений планируемым результатам обучения по дисциплине (модулю) созданы фонды оценочных средств. Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 1.

7 Учебно-методическое и информационное обеспечение дисциплины (модуля)

7.1 Основная литература

1. Миркин Б. Г. ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ. Учебник и практикум [Электронный ресурс] , 2020 - 174 - Режим доступа: <https://urait.ru/book/vvedenie-v-analiz-dannyh-450262>
2. Под ред. Елисеевой И.И. ЭКОНОМЕТРИКА. Учебник для вузов [Электронный ресурс] , 2020 - 449 - Режим доступа: <https://urait.ru/book/ekonometrika-449677>
3. Энатская Н. Ю. ТЕОРИЯ ВЕРОЯТНОСТЕЙ. Учебное пособие для вузов [Электронный ресурс] , 2020 - 203 - Режим доступа: <https://urait.ru/book/teoriya-veroyatnostey-451176>

7.2 Дополнительная литература

1. Айвазян С. А. Методы эконометрики : Учебник [Электронный ресурс] : Магистр , 2020 - 512 - Режим доступа: <https://znanium.com/catalog/document?id=355480>
2. Балдин К. В., Рукоусев А. В. Общая теория статистики : Учебники [Электронный ресурс] - Москва : Дашков и К^о , 2020 - 312 - Режим доступа: http://biblioclub.ru/index.php?page=book_red&id=573143

7.3 Ресурсы информационно-телекоммуникационной сети "Интернет", включая профессиональные базы данных и информационно-справочные системы (при необходимости):

1. Мастицкий С.Э., Шитиков В.К. Статистический анализ и визуализация данных с помощью R. - Электронная книга, адрес доступа: <http://r-analytics.blogspot.com>
2. Математический форум Math Help Planet (<http://mathhelpplanet.com/static.php>)
3. Официальный сайт RStudio - Режим доступа: <https://rstudio.com/>
4. СПС КонсультантПлюс - Режим доступа: <http://www.consultant.ru/>
5. Шитиков В.К., Мастицкий С.Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. 2017 - 351 с. Электронная книга, адрес доступа: <https://github.com/ranalytics/data-mining>
6. Электронная библиотечная система «Университетская библиотека онлайн» - Режим доступа: <http://biblioclub.ru/>
7. Электронная библиотечная система ZNANIUM.COM - Режим доступа: <https://znanium.com/>
8. Электронно-библиотечная система издательства "Юрайт" - Режим доступа: <https://urait.ru/>
9. Open Academic Journals Index (ОАИ). Профессиональная база данных - Режим

доступа: <http://oaji.net/>

10. Президентская библиотека им. Б.Н.Ельцина (база данных различных профессиональных областей) - Режим доступа: <https://www.prlib.ru/>

8 Материально-техническое обеспечение дисциплины (модуля) и перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения

Основное оборудование:

- Коммутатор SuperStack 3 (16*10/100 19")
- Монитор облачный 23" LG23CAV42K/мышь Geniu
- Облачный монитор 23" LG CAV42K
- Облачный монитор LG Electronics черный +клавиатура+мышь
- П/К DNS Office T300, мышь Genius NetScroll 100, клавиатура Genius KB-06X, монитор AOC919 19"
- Проектор Casio XJ-V1
- Усилитель-распределитель VGA/XGA Kramer VP-200
- Уст-во бесп.питания UPS-3000

Программное обеспечение:

- RStudio

МИНОБРНАУКИ РОССИИ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ

Фонд оценочных средств
для проведения текущего контроля
и промежуточной аттестации по дисциплине (модулю)

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ НА ЯЗЫКЕ R

Направление и направленность (профиль)

38.04.05 Бизнес-информатика. Информационная бизнес-аналитика

Год набора на ОПОП
2022

Форма обучения
очная

Владивосток 2023

1 Перечень формируемых компетенций

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции
38.04.05 «Бизнес-информатика» (М-БИ)	ОПК-3 : Способен принимать решения, осуществлять стратегическое планирование и прогнозирование в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта	ОПК-3.1к : Использует системы искусственного интеллекта в рамках стратегического планирования и прогнозирования в профессиональной деятельности
		ОПК-3.2к : Использует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными для поддержки стратегических управленческих решений
	ПКВ-2 : Способен выбирать и осуществлять рациональные решения по ИС и ИКТ для управления организациями на основе больших данных	ПКВ-2.1к : Использует и совершенствует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными
		ПКВ-2.2к : Разрабатывает технические проекты в сфере ИТ, позволяющие анализировать и использовать большие данные в управлении организацией

Компетенция считается сформированной на данном этапе в случае, если полученные результаты обучения по дисциплине оценены положительно (диапазон критериев оценивания результатов обучения «зачтено», «удовлетворительно», «хорошо», «отлично»). В случае отсутствия положительной оценки компетенция на данном этапе считается несформированной.

2 Показатели оценивания планируемых результатов обучения

Компетенция ПКВ-2 «Способен выбирать и осуществлять рациональные решения по ИС и ИКТ для управления организациями на основе больших данных»

Таблица 2.1 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код результата	Тип результата	Результат	
ПКВ-2.1к : Использует и совершенствует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными	РД5	Знание	методов исследования современных проблем и методов прикладной информатики	правильно выбирает методы и следования при решении прикладных задач
ПКВ-2.2к : Разрабатывает технические проекты в сфере ИТ, позволяющие анализировать и использовать большие данные в управлении организацией	РД6	Умение	проводить исследования задач различных классов, используя современные методы, средства прикладной информатики и проводить их анализ	грамотно излагает и интерпретирует ход и результаты исследований в аналитических отчетах

	Р Д 7	Н а в ы к	использования инструментальных средств при разработке программных средств, основанных на методах машинного обучения	самостоятельно создает программное средство моделирования процессов, используя инструментальные средства
--	-------------	-----------------------	---	--

Компетенция ОПК-3 «Способен принимать решения, осуществлять стратегическое планирование и прогнозирование в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта»

Таблица 2.2 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код результата	Тип результата	Результат	
ОПК-3.1к : Использует системы искусственного интеллекта в рамках стратегического планирования и прогнозирования в профессиональной деятельности	РД1	Знание	аналитических методов обработки статистических данных	правильность постановки задач в новой или незнакомой среде, формулировок выводов, полученных в результате анализа данных
	РД2	Умение	применять методы машинного обучения при решении нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	корректность выбора и правильность использования методов и моделей машинного обучения для проведения исследований в различных областях
ОПК-3.2к : Использует методы, модели, алгоритмы, технологии и инструментальные средства для работы с большими данными для поддержки стратегических управленческих решений	РД3	Знание	методов разработки алгоритмов на языке R для решения профессиональных задач	правильность использования функций языка R для решения поставленных задач
	РД4	Умение	разрабатывать оригинальные алгоритмы с использованием языка программирования R для решения профессиональных задач	корректность составленных алгоритмов для решения профессиональных задач

Таблица заполняется в соответствии с разделом 1 Рабочей программы дисциплины (модуля).

3 Перечень оценочных средств

Таблица 3 – Перечень оценочных средств по дисциплине (модулю)

Контролируемые планируемые результаты обучения	Контролируемые темы дисциплины	Наименование оценочного средства и представление его в ФОС	
		Текущий контроль	Промежуточная аттестация
Очная форма обучения			

РД1	Знание : аналитических методов обработки статистических данных	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест
		1.4. Задачи классификации. Модели с дискретно зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест
РД2	Умение : применять методы машинного обучения при решении нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест
		1.4. Задачи классификации. Модели с дискретно зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест
РД3	Знание : методов разработки алгоритмов на языке R для решения профессиональных задач	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест
		1.4. Задачи классификации. Модели с дискретно зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест
РД4	Умение : разрабатывать оригинальные алгоритмы с использованием языка программирования R для решения профессиональных задач	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест

		1.4. Задачи классификации. Модели с дискретной зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест
РД5	Знание : методов исследования современных проблем и методов прикладной информатики	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест
		1.4. Задачи классификации. Модели с дискретной зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест
РД6	Умение : проводить исследование задач различных классов, используя современные методы, средства прикладной информатики и проводить их анализ	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест
		1.4. Задачи классификации. Модели с дискретной зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест
РД7	Навык : использования инструментальных средств при разработке программных средств, основанных на методах машинного обучения	1.1. Методы машинного обучения. Регрессионный анализ. Парная регрессия. Построение регрессий в RStudio.	Практическая работа	Тест
		1.2. Множественная регрессия и корреляция	Практическая работа	Тест
		1.3. Снижение размерности. Метод главных компонент	Практическая работа	Тест
		1.4. Задачи классификации. Модели с дискретной зависимой переменной	Практическая работа	Тест
		1.5. Кластерный анализ	Практическая работа	Тест

4 Описание процедуры оценивания

Качество сформированности компетенций на данном этапе оценивается по результатам текущих и промежуточных аттестаций при помощи количественной оценки, выраженной в баллах. Максимальная сумма баллов по дисциплине (модулю) равна 100 баллам.

Вид учебной деятельности	Оценочное средство						
	Практическая работа № 1	Практическая работа № 2	Практическая работа № 3	Практическая работа № 4	Практическая работа № 5	Итоговый тест	Итого
Практические занятия	18	18	18	18	18		90
Промежуточная аттестация						10	10
Итого	18	18	18	18	18	10	100

Сумма баллов, набранных студентом по всем видам учебной деятельности в рамках дисциплины, переводится в оценку в соответствии с таблицей.

Сумма баллов по дисциплине	Оценка по промежуточной аттестации	Характеристика качества сформированности компетенции
от 91 до 100	«зачтено» / «отлично»	Студент демонстрирует сформированность дисциплинарных компетенций, обнаруживает всестороннее, систематическое и глубокое знание учебного материала, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой, умеет свободно выполнять практические задания, предусмотренные программой, свободно оперирует приобретенными знаниями и умениями, применяет их в ситуациях повышенной сложности.
от 76 до 90	«зачтено» / «хорошо»	Студент демонстрирует сформированность дисциплинарных компетенций: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
от 61 до 75	«зачтено» / «удовлетворительно»	Студент демонстрирует сформированность дисциплинарных компетенций: в ходе контрольных мероприятий допускаются значительные ошибки, проявляется отсутствие отдельных знаний, умений, навыков по некоторым дисциплинарным компетенциям, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
от 41 до 60	«не зачтено» / «неудовлетворительно»	У студента не сформированы дисциплинарные компетенции, проявляется недостаточность знаний, умений, навыков.
от 0 до 40	«не зачтено» / «неудовлетворительно»	Дисциплинарные компетенции не сформированы. Проявляется полное или практически полное отсутствие знаний, умений, навыков.

5 Примерные оценочные средства

5.1 Примеры заданий для выполнения практических работ

Практическая работа №1

Тема 1.

Используя набор данных, собранных самостоятельно или предоставленных преподавателем, или, например, размещенных на сайте <https://rosstat.gov.ru/>, магистрант выполняет задания лабораторной работы, которая включает:

1. Построение парной линейной регрессии при помощи команды `lm()` в RStudio. Иллюстрация реальных и предсказанных по модели значений.
2. Интерпретация результатов моделирования, полученных в RStudio.
3. Построение точечного прогноза и доверительного интервала для него по результатам расчета модели регрессии.

4. Построение парной нелинейной регрессии с помощью команды `nls()` в RStudio.

Практическая работа №2

Тема 2.

Используя набор данных, собранных самостоятельно или предоставленных преподавателем, или, например, размещенных на сайте <https://rosstat.gov.ru/>, магистрант выполняет задания лабораторной работы, которая включает:

1. Построение модели множественной регрессии при помощи команд `lm()` и `glm()` в RStudio.
2. Исследование данных на мультиколлинеарность.
3. Интерпретация результатов моделирования, полученных в RStudio.
4. Построение моделей с дамми переменными.
5. Выбор оптимального набора предикторов.
6. Проверка предпосылок МНК.

Практическая работа №3

Тема 3.

Используя набор данных, собранных самостоятельно или предоставленных преподавателем, магистрант выполняет задания лабораторной работы, которая включает:

1. Выбор необходимого числа компонент. Анализ собственных значений с использованием функции `fa.parallel()` из пакета `psych`.
2. Выделение главных компонент при помощи функции `principal()`. Анализ полученных результатов.
3. Анализ главных компонент с вращением.
4. Вычисление главных компонент.

Практическая работа №4

Тема 4.

Используя набор данных, собранных самостоятельно или предоставленных преподавателем, магистрант выполняет задания лабораторной работы, которая включает:

1. Построение логит и пробит моделей в RStudio при помощи команды `glm()` с параметром семейства распределений `binomial`.
2. Интерпретация параметров моделей.

Практическая работа №5

Тема 5.

Используя набор данных, собранных самостоятельно или предоставленных преподавателем, магистрант выполняет задания лабораторной работы, которая включает:

1. Метод k средних с использованием функции `kmeans()` из пакета `cluster`. Подбор оптимального числа кластеров при помощи пакета `factoextra`, предназначенного для визуализации результатов кластерного анализа на основе графической системы `ggplot2`.
2. Проведение иерархической кластеризации при помощи функции `hclust()`.
3. Оценка качества кластеризации. Оценка общей предрасположенности имеющихся данных к объединению в кластеры с помощью функции `get_clust_tendency()` из пакета `factoextra`. Поиск оптимальной схемы объединения в кластеры при помощи пакета `NbClust`.

Краткие методические указания

При выполнении задания допускается объединение магистрантов в группы по 2-5 человек. Практические работы позволяют определить уровень усвоения материала. Перед выполнением работы необходимо ознакомиться с теоретическим материалом, проработать методы решения задач, рассмотренных в типовых примерах. После выполнения работы предусмотрен письменный аналитический отчет и публичная защита полученных результатов моделирования, сравнение и обсуждение результатов всех участников

Шкала оценки

Оценка	Баллы	Описание
5	17–18	выставляется студенту, если его группа качественно собрала, обработала информацию и представила результаты по заданной теме, он лично отвечал на вопросы представителей других групп и преподавателя, четко представляя свою позицию и аргументируя точку зрения
4	14–16	выставляется студенту, если его группа качественно собрала, обработала информацию и представила результаты по заданной теме, но он лично, отвечая на вопросы представителей других групп и преподавателя, недостаточно четко представлял свою позицию и аргументировал точку зрения
3	9–13	баллов выставляется студентам групп, недостаточно качественно собравших, обработавших информацию или представивших результаты по заданной теме, он лично, отвечая на вопросы представителей других групп и преподавателя, недостаточно четко представлял свою позицию и аргументировал точку зрения
2	0–8	баллов выставляется студентам групп, недостаточно качественно собравших, обработавших информацию и представивших результаты по заданной теме, или он лично, не участвовал в обсуждении

5.2 Примеры тестовых заданий

1 Выберите методы и модели, которые относятся к методам машинного обучения с учителем:

- 1) регрессия,
- 2) классификация,
- 3) кластеризация,
- 4) снижение размерности.

2 При заданном объеме наблюдений при прочих равных условиях с увеличением числа объясняющих переменных скорректированный коэффициент детерминации:

- 1)увеличивается; 2)уменьшается; 3)не изменяется;4)увеличивается вдвое

3 Число степеней свободы для факторной суммы квадратов в линейной модели множественной регрессии равно:

- 1); 2); 3); 4).

4 Множественный коэффициент детерминации . Определите, какой процент дисперсии зависимой переменной объясняется влиянием факторов и :

- 1)90%; 2)10 %; 3)0,1 %; 4)0,9% .

5 Стандартизованные коэффициенты регрессии :

- 1)позволяют ранжировать факторы по силе их влияния на результат,
- 2)оценивают статистическую значимость факторов,
- 3)являются коэффициентами эластичности,
- 4)являются коэффициентами корреляции.

6 Фиктивные переменные – это:

1) атрибутивные признаки (например, как профессия, пол, образование), которым придали цифровые метки;

2) экономические переменные, которые вводятся в модель, но не используются для прогнозов;

3) переменные, соответствующие статистически не значимым факторам;

4) значения независимой переменной за предшествующий период времени.

7 Если в модели используются два качественных фактора, имеющих по три варианта значений, то необходимое число фиктивных переменных:

- 1)4; 2)6; 3)2; 4)3.

8 Наличие мультиколлинеарности характеризует

- 1) близкий к нулю определитель матрицы межфакторной корреляции,
- 2) близкий к единице определитель матрицы межфакторной корреляции,
- 3) близкий к нулю коэффициент детерминации,
- 4) высокий коэффициент корреляции результата с признак-факторами.

9 Критерий Гольфельда-Квандта для модели с p факторами предполагает

1) вычисление остаточных сумм квадратов для двух парных уравнений регрессии,
2) вычисление остаточных сумм квадратов для двух уравнений регрессии с p параметрами,

3) вычисление коэффициента детерминации,

4) вычисление параметров уравнения регрессии для средних S наблюдений.

10 Какие задачи из нижеследующих относятся к задачам классификации:

1) оценка кредитоспособности заемщика;

2) диагностика наличия заболевания у пациента;

3) определение прибыли по известному набору значений предикторов;

разбиение объектов на группы, которые заранее не predeterminedены?

Краткие методические указания

Тест содержит тестовые задания с выбором одного или нескольких правильных ответов, ввод числа.

Шкала оценки

Оценка	Баллы	Описание
5	10	выставляется студенту, если он ответил на более, чем 90 % поставленных вопросов, учитывая множественный выбор ответов теста
4	8-9	выставляется студенту, если он ответил на 76-90%% поставленных вопросов, учитывая множественный выбор ответов теста
3	6-7	выставляется студенту, если он ответил на 61-75%% поставленных вопросов, учитывая множественный выбор ответов теста
2	0-5	выставляется студенту, если он ответил не более, чем на 60% поставленных вопросов, учитывая множественный выбор ответов теста