

Макет учебного комплекса

«Технологии R (часть 2)»
«Data Science with R for economists (part 2)»

Кафедра Экономической информатики. Ауд.548, тел. 7548

Сайт ecinf.econ.msu.ru. e-mail _____

Статус дисциплины: факультатив, читается в 6 семестре

Авторы программы и лекторы:

Пересветов Сергей Борисович, peresvetov@econ.msu.ru

1. Место дисциплины в структуре ООП подготовки бакалавра – Профессиональный цикл

Обеспечиваются базовые знания, умения и навыки в области обработки информации для дисциплин «Информационные системы в управлении организацией», «Экономика информации», «Информационные бухгалтерские системы».

Дисциплина обеспечивает базовые умения и навыки в области обработки информации для дисциплин «Эконометрика», «Статистика», «Теория финансов» и других дисциплин, использующих современные средства обработки данных.

Также обеспечивается овладение умениями и навыками сбора экономической информации из отечественных и зарубежных источников (ПК-9), выбора оптимальных инструментальных средств обработки экономических данных (ПК-5), продвинутый уровень использования для аналитических и исследовательских задач современных технических средств и пакетов прикладных программ (ПК-10), способность участвовать в постановке задач автоматизации своей предметной области (ПК-14).

Входные знания – владение знаниями информационных систем, умениями и навыками работы в прикладных офисных пакетах в объеме дисциплины «Экономическая информатика»

Предшествующие дисциплины – «Экономическая информатика»

2. Объем дисциплины и виды учебной работы

Нагрузка	В часах	В кредитах
Общая трудоемкость	32	
В том числе: Аудиторная	32	
Самостоятельная		
Контактная		
Форма итогового контроля		

3. Цель и задачи дисциплины.

Цель – получение базовых знаний, умений и навыков автоматизированной обработки информации и данных в соответствии с компетенциями ОК-12 и ОК-13, базовые знания основ электронных коммуникаций (ОК-19, ПК-12) и постановки задач автоматизации (ПК-14).

Задачи:

- Получение продвинутых знаний, умений и навыков обработки данных и информации (ПК-10, ПК-5, ПК-12);
- Развитие умений и навыков использования внутренних (корпоративных) и внешних источников данных (ПК-7, ПК-9);
- Выработка знаний, умений и навыков постановки задачи автоматизации своей предметной области (ПК-14).

(Указать соответствие целей и задач дисциплины общекультурным или профессиональным компетенциям и их код в соответствии с приведенной классификацией)

4. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля). В результате освоения дисциплины студент должен:

Знать: Основные принципы построения персональной информационной системы, обеспечивающей производительную работу экономиста

Уметь: Выделить в решаемой задаче необходимые ИТ-сервисы, подобрать пакет прикладных программ, поддерживающий эти сервисы, построить необходимые структуры данных и интерфейсы между пакетами, получить необходимые исходные данные из онлайн-источников через сеть Интернет или корпоративной ИС.

Владеть: Навыками обработки данных в офисных пакетах, статистических пакетах, интеграции данных различных пакетов, поиском данных в электронных библиотеках.

5. Структура и содержание дисциплины (модуля)

5.1. Разделы дисциплин и виды занятий (календарный план) в часах

№ недели	Изучаемые темы	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				
		Аудиторная работа		Контактный час		Самостоятельная работа студента
		Лекция	семинар	консультация	Контактная роль	
	Часть 1. Введение. Элементы технологий машинного обучения в среде R.					
1	Основные задачи анализа данных - классификация и регрессия, кластеризация, задача ассоциации.					
2	Практический бизнес-анализ. Классификация. Скоринг кредитоспособности заемщиков.					
3	Практический бизнес-анализ. Кластеризация и ассоциации. Ассоциативные правила в стимулировании розничных продаж. Сегментация клиентов компании.					
4	Повышение точности предиктивных моделей. Ансамбли моделей.					
	Часть 2. Временные ряды в R.					
5	Работа с финансовыми данными. Эксплораторный анализ активов и оптимизация портфелей инвестиций.					
	Часть 3. Глубокое обучение (Deep Learning) в R.					
6	Пакеты R для доступа к библиотекам глубокого обучения Keras, TensorFlow, Theano, CNTK. Построение нейронных сетей в R. Проблема оверфиттинга (переобучения). Идентификация аномальных данных.					
7	Практический бизнес-анализ. Бинарная классификация - классификация отзывов к фильмам. Классификация на несколько классов - классификация новостных лент.					
	Часть 4. Облачные технологии и					

	технологии обработки больших данных (Big Data) на примере платформы Microsoft Azure.				
8	Технологии распределенных вычислений. Microsoft R Server, Azure HDInsight with R, Azure Machine Learning, Spark, Scala, Hive и т.д.				
9	Практический бизнес-анализ. Анализ оттока клиентов (Churn Prediction) на платформе Microsoft Azure.				
	Часть 5. Подготовка документов и презентаций в среде R/RStudio. Проект Bookdown.				
10	RMarkdown, пакеты bookdown и TinyTex. Создание документов и презентаций. Проблемы русификации.				
Всего			32		

5.2 Форма проведения самостоятельной работы и текущего контроля успеваемости

Указываются обязательные (например, изучение материалов учебника, лекции) и по выбору студента (например, написание реферата, подготовка доклада, презентации, коллективного или индивидуального проекта, участие в конференции, подготовка кейса и прю) формы самостоятельной работы и способы проверки, трудоемкость выполнения самостоятельной работы.

№ нед.	Наименование раздела дисциплины	Форма самостоятельной работы	Форма текущего контроля успеваемости	Трудоемкость (в часах)
Всего				

5.3. Содержание разделов дисциплины

6. Образовательные технологии

Лекции, в т.ч. с презентациями PowerPoint, решение задач в интерактивном режиме в ходе семинара, самостоятельное решение задач вне аудитории, коллоквиумы, электронные тесты. Лекции составляют 33% аудиторного времени, самостоятельная внеаудиторная работа равна аудиторному времени.

(Указываются образовательные технологии, используемые при реализации различных видов учебной работы.)

В соответствии с требованиями ФГОС ВПО по направлению подготовки реализация компетентностного подхода должна предусматривать широкое использование в

учебном процессе активных и интерактивных форм проведения занятий (компьютерных симуляций, деловых и ролевых игр, разбор конкретных ситуаций, психологические и иные тренинги) в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся. В рамках учебных курсов должны быть предусмотрены встречи с представителями российских и зарубежных компаний, государственных и общественных организаций, мастер-классы экспертов и специалистов.

Удельный вес занятий, проводимых в интерактивной форме, определяется главной целью (миссией) программы, особенностью контингента обучающихся и содержанием конкретных дисциплин, и в целом в учебном процессе они должны составлять не менее ___% аудиторных занятий (определяется требованиями ФГОС с учетом специфики ООП). Занятия лекционного типа для соответствующих групп студентов не могут составлять более ___% аудиторных занятий (определяется соответствующим ФГОС)).

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

Самостоятельные работы (темы указаны в разделе 5.2), контрольные работы, зачет. Эссе в рамках дисциплины не предполагаются.

(Указываются темы эссе, курсовых работ и др. Приводятся контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины, а также для контроля самостоятельной работы обучающегося по отдельным разделам дисциплины).

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

8.1 основная литература

1. Даниэла Уиттон, Гарет Джеймс, Роберт Тибширани, Тревор Хасти, Введение в статистическое обучение с примерами на языке R. Издательство: ДМК Пресс, 2017
2. Франсуа Шолле. Глубокое обучение на R. Издательство: Питер, 2018
3. Роберт Кабаков: R в действии. Анализ и визуализация данных на языке R. Издательство: ДМК-Пресс, 2014
4. Мастицкий С.Э., Шитиков В.К. (2014) Статистический анализ и визуализация данных с помощью R – Электронная книга, адрес доступа:<http://r-analytics.blogspot.com>
5. Free online books **bookdown**, адрес доступа <https://bookdown.org/home/archive/>
6. W.N.Venables, D.M.Smith and the R Development Core Team, An Introduction to R, 2010
7. Diethelm Würtz, Yohan Chalabi, Longhow Lam, Andrew Ellis, Basic R for Finance, Finance Online Publishing, Zurich, 2015
8. Diethelm Würtz, Tobias Setz, Yohan Chalabi, William Chen, Andrew Ellis, Portfolio Optimization with R/Rmetrics, Rmetrics Association and Finance Online Publishing, Zurich, Rmetrics eBooks 2009, NEW: Update 2015
9. Eric Zivot and Jiahui Wang, Modelling Financial Time Series with S-PLUS, Second Edition, 2005
10. Paul S.P.Cowpertwait, Andrew V.Metcalfе, Introductory Time Series with R, Springer, 2009

8.2 дополнительная литература

1. Сергей Николенко, А. Кадури́н, Е. В. Архангельская, Глубокое обучение. Погружение в мир нейронных сетей. Издательство: Питер, 2018
2. А.А.Барсе́гян, М.С.Куприянов, И.И.Холод, М.Д.Тесс, С.И. Елизаров, Анализ данных и процессов, СПб: БХВ-Петербург, 2009
3. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: Учебное пособие, 2-е изд., СПб.:Питер, 2010

8.3 Интернет-источники:

1. cran.r-project.org
2. www.bookdown.org

9. Балльная система оценки знаний

1. Самостоятельные работы – 50 баллов
2. Контрольные работы – 50 баллов
3. Зачет – 50 баллов

(условие получения итоговой оценки в соответствии с положением о балльно-рейтинговой системой оценки знаний)

10. Материально-техническое обеспечение дисциплины (модуля)

Для лекций – компьютер и проектор

Для семинаров – компьютерный класс, ПО Microsoft Office 2016 Pro с соответствующими надстройками, пакет R и графическая оболочка RStudio, доступ к институциональной подписке ЭФ МГУ, доступ в Интернет

Указывается материально-техническое обеспечение данной дисциплины (модуля), например, программное обеспечение, специализированные аудитории и пр.