

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Институт бизнеса и делового администрирования

Кафедра менеджмента

Утверждена

на заседании кафедры менеджмента

Протокол от «29» июня 2017 г. № 4

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Индекс ФТД.В.07 «Математика и статистика для Data Science»

по направлению подготовки 38.04.02 Менеджмент

направленность «Прорывные финансовые и цифровые технологии»

квалификация Магистр

очная форма обучения

Год набора - 2018

Москва, 2018 г.

Автор(ы)–составитель(и): Бологов Ярослав Владимирович, к.э.н., Ведущий инженер анализа данных Института прикладного анализа данных Делойта

Заведующий кафедрой менеджмента ИБДА, доктор экономических наук, профессор
Гапоненко Александр Лукич.

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине соотнесенных с планируемыми результатами освоения образовательной программы
2. Объем и место дисциплины в структуре образовательной программы
3. Содержание и структура дисциплины
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине
5. Методические указания для обучающихся по освоению дисциплины
6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине
 - 6.1. Основная литература
 - 6.2. Дополнительная литература
 - 6.3. Учебно-методическое обеспечение самостоятельной работы
 - 6.4. Нормативные правовые документы
 - 6.5. Интернет-ресурсы
 - 6.6. Иные источники
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина ФТД.В.07 «Математика и статистика для Data Science» обеспечивает овладение следующими компетенциями:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК-4	способность управлять организациями, подразделениями, группами (командами) сотрудников, проектными сетями	ПК-4.1	способен управлять организациями, подразделениями, группами (командами) сотрудников, проектными сетями

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

ОТФ/ТФ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
способность управлять организациями, подразделениями, группами (командами) сотрудников, проектными сетями	ПК-4.1	<ul style="list-style-type: none"> - Умеет составлять и решать различные статистические задачи, оценивать неизвестные параметры распределений и проверять гипотезы о согласии с теоретическим распределением. - Умеет представить результаты статистического исследования в виде таблиц, графиков, описания полученных результатов и их интерпретации

2. Объем и место дисциплины в структуре ОП ВО

Объем дисциплины

1 ЗЕ/36 ак. часа/27 астроном. час, на контактную работу с преподавателем 26 ак. часов/19,5 астр. час, на самостоятельную работу обучающихся 6 ак час/4,5 астр.час.

Место дисциплины в структуре ОП ВО

ФТД.В.07 «Математика и статистика для Data Science», 1 курс, 1 семестр.

Для освоения учебной дисциплины студенты должны владеть знаниями и компетенциями, полученными при изучении дисциплин:

☐ «Математика»

☐ «Менеджмент»

– форма промежуточной аттестации – зачет с оценкой.

3. Содержание и структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины (модуля), ак. час./ час.						Форма текущего контроля успеваемости, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Очная форма обучения								
Тема 1	Введение	7	1		4		1	ДЗ

№ п/п	Наименование тем (разделов)	Объем дисциплины (модуля), ак. час./ час.						Форма текущего контроля успеваемости, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Очная форма обучения								
Тема 2	Data Science и его задачи	7	1		4		1	ДЗ
Тема 3	Оценка качества моделей	7	1		4		1	ДЗ
Тема 4	Модели машинного обучения и их построение	7	1		4		1	ДЗ
Тема 5	Управление Data Science проектами	8	2		4		2	ДЗ
Промежуточная аттестация								Зачет с оценкой
Всего:		36/27	6,4,5		20/15		6/4,5	

Примечание* – формы текущего контроля успеваемости: домашнее задание (ДЗ), контрольная работа (КР).

Содержание дисциплины

Темы:

Тема 1. Введение

Лекция No1. Что такое DataScience. Примеры впечатляющих и вдохновляющих проектов. Примеры приземленных проектов из практической деятельности. Описание ролей и задействованных специалистов.

Лекция No2. Этапы проекта. Разбор примера плана проекта с указанием ролей, задач, проблем. Рынок данных. Юридические аспекты работы с данными.

Лекция No3. Инструменты аналитика —python, R, ноутбуки; SAS, SPSS, RapidMiner; библиотеки —какие и зачем; Hadoop, Spark.

Тема 2. Data Science и его задачи

Лекция No4. Типы задач, примеры; выбор метрик -вводная часть; управление ожиданиями по качеству; анализ сценариев использования моделей.

Лекция No5. Валидация задачи; валидация данных; экономический эффект.

Тема 3. Оценка качества моделей

Лекция No6. Метрики классификации; метрики регрессии; метрики ранжирования; стоимостное выражение метрик

Лекция No7. Как и зачем мерять качество; А/Б тестирование; стат.значимость, чувствительность; примеры онлайн и офлайн тестов.

Тема 4. Модели машинного обучения и их построение

Лекция No8. Обзорная лекция по математике, стоящей за DataScience.

Лекция No9. Состав работ в проекте на детальном уровне; предобработка данных; планирование экспериментов; валидация и тестирование в ходе проекта

Тема 5. Управление Data Science проектами и примеры проектов

Лекции No10–12. Разбор примеров проектов с учетом всего ранее изученного.

4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине

4.1. Формы и методы текущего контроля успеваемости.

4.1.1. В ходе реализации дисциплины ФТД.В.07 «Математика и статистика для Data Science» используются следующие методы текущего контроля успеваемости обучающихся:

Тема (раздел)	Методы текущего контроля успеваемости
Тема 1	Домашнее задание 1
Тема 2	Домашнее задание 2
Тема 3	Домашнее задание 3
Тема 4	Домашнее задание 4
Тема 5	Домашнее задание 5

4.1.2. Зачет проводится с применением следующих методов (средств): в виде письменной контрольной работы/аналитического эссе (эсэймент)

4.2. Материалы текущего контроля успеваемости обучающихся

Примерный перечень вопросов:

1. Назовите этапы проекта по машинному обучению
2. Перечислите основные группы инструментов специалистов по машинному обучению, сферы их применения
3. Какие законы регулируют работу с данными? Опишите основные моменты
4. Перечислите шаги валидации задачи машинного обучения
5. Перечислите шаги валидации данных для задач машинного обучения
6. Перечислите основные методы оценки качества моделей при промышленном внедрении.

4.3. Оценочные средства для промежуточной аттестации.

4.3.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Показатели и критерии оценивания компетенций с учетом этапа их формирования.

ОТФ/ТФ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
способность управлять организациями, подразделениями, группами (командами) сотрудников, проектными сетями	ПК-4.1	- Умеет составлять и решать различные статистические задачи, оценивать неизвестные параметры распределений и проверять гипотезы о согласии с теоретическим распределением. - Умеет представить результаты статистического исследования в виде таблиц, графиков, описания полученных результатов и их интерпретации

4.3.2. Типовые оценочные средства.

Шкала оценивания.

10-бальная шкала	Традиционная шкала	«Зачтено»/ «Не зачтено»	Определение
10	Отлично	Зачтено	Полные, глубокие и систематические знания в сфере Data Science , полный и правильный ответ на теоретический вопрос, полное и правильное решение задачи.
9	Отлично	Зачтено	Глубокие и систематические знания в сфере Data Science , полный и правильный ответ на теоретический вопрос, полное и правильное решение задачи.
8	Отлично	Зачтено	Систематические знания в сфере Data Science , полный и правильный ответ на теоретический вопрос, правильное решение задачи.
7	Хорошо	Зачтено	Систематические знания в сфере Data Science , полный и правильный ответ на теоретический вопрос, правильное решение задачи.
6	Хорошо	Зачтено	Систематические знания в сфере Data Science , полный и правильный ответ на теоретический вопрос, правильное решение задачи с незначительными неточностями.
5	Удовлетворительно	Зачтено	Ответ на теоретический вопрос в сфере Data Science , полный и правильный ответ на теоретический вопрос, правильное решение задачи с незначительными неточностями.
4	Удовлетворительно	Зачтено	Ответ на теоретический вопрос на тему банкинг, полный и правильный ответ на теоретический вопрос, решение задачи содержит арифметические ошибки, не влияющие на правильность хода решения задачи.
3	Неудовлетворительно	Не зачтено	Ответ на теоретический вопрос на тему Data Science , решение задачи содержит идеологические ошибки.
2	Неудовлетворительно	Не зачтено	Ответ на теоретический вопрос на тему Data Science неверный и/или решение задачи содержит идеологические ошибки.
1	Неудовлетворительно	Не зачтено	Ответ на теоретический вопрос на тему Data Science неверный и решение задачи отсутствует.
0	Неудовлетворительно	Не зачтено	Ответ на теоретический вопрос на тему Data Science отсутствует и решение задачи отсутствует.

4.4. Методические материалы по проведению промежуточной аттестации

Зачёт проводится в компьютерной аудитории, с установленным пакетом Wolfram Mathematica, с доступом в локальную сеть Академии и глобальную сеть Интернет. Студент выполняет решение задачи на компьютере, затем предоставляет решение преподавателю, по данному решению проводится собеседование со студентом.

5. Методические указания для обучающихся по освоению дисциплины

Все выступления, как преподавателя, так и студентов, могут сопровождаться наглядным материалом в виде презентаций, выполненных на компьютере, отрывками из документальных или художественных фильмов, иллюстрациями, прочим раздаточным материалом, что способствует более полному и глубокому освещению материала и легкости его усвоения слушателями. Развивается способность находить и выделять нужный информационный материал из разных видов источников, проводить его анализ и правильную подачу аудитории.

6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1 Основная литература

1. David Julian, Designing Machine Learning Systems with Python, PACKT, 2016
2. Gene Kim, Kevin Behr, George Spafford, The Phoenix Project: A Novel About IT, DevOps, and Helping Your Business Win, IT Revolution Press, 2014
3. Jennifer Davis, Katherine Daniels, Effective DevOps: Building a Culture of Collaboration, Affinity, and Tooling at Scale, O'Reilly Media, Inc., 2016
4. Mark C. Layton, Agile Project Management For Dummies, John Wiley & Sons, 2012

7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

Для проведения занятий необходим компьютерный класс, оснащённый компьютерами под управлением операционной системой Windows, с доступом в интернет и установленным прикладным программным обеспечением Wolfram Mathematica, Microsoft Office, Adobe Reader. Аудитории с доской, мелом или маркерами – для лекционных и части семинарских занятий.