



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

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

«СОГЛАСОВАНО»

Руководитель ОП

Е.И. Бережнова

«УТВЕРЖДАЮ»
Заведующий (ая) кафедрой
компьютерных систем
(название кафедры)

Пустовалов Е.В.

(подпись) * (Ф.И.О. зав. каф.)

«11» июля 2019 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наука о данных и аналитика больших данных

Специальность 38.05.01 Экономическая безопасность

Специализация: Экономико-правовое обеспечение экономической безопасности

Форма подготовки очная

курс 1 семестр 2

лекции 8 час.

практические занятия 0 час.

лабораторные работы 0 час.

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

всего часов аудиторной нагрузки 8 час.

в том числе с использованием МАО час.

самостоятельная работа 100 час.

в том числе на подготовку к экзамену 0 час.

контрольные работы (количество) 0

курсовая работа / курсовой проект семестр

зачет 2 семестр

экзамен не предусмотрен

Рабочая программа составлена в соответствии с требованиями федерального государственного образовательного стандарта высшего образования, утвержденного приказом Министерства образования и науки РФ от 16.01.2017 № 20

Рабочая программа обсуждена на заседании кафедры мировой экономики, протокол № 16 от 18 июля 2019 г.

Заведующий кафедрой: Пустовалов Е.В.

Составитель: к.ф.-м.н., ассистент Шевченко Ю.А.

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)

Аннотация к рабочей программе дисциплины «Наука о данных и аналитика больших данных»

Учебный курс «Наука о данных и аналитика больших объемов данных» предназначен для студентов специальности 38.05.01 Экономическая безопасность, специализация «Экономико-правовое обеспечение экономической безопасности».

Дисциплина «Наука о данных и аналитика больших объемов данных» включена в состав вариативной части блока «Дисциплины (модули)».

Общая трудоемкость дисциплины составляет 3 зачетные единицы, 108 часов. Учебным планом предусмотрены лекционные занятия (8 часов), самостоятельная работа студентов (100 часов). Дисциплина реализуется на 1 курсе во 2 семестре.

Дисциплина «Наука о данных и аналитика больших объемов данных» основывается на знаниях, умениях и навыках, полученных в результате изучения дисциплин «Современные информационные технологии», «Математика в экономике» и позволяет подготовить студента к освоению ряда таких дисциплин, как «Эконометрика», «Финансово-экономические расчеты», подготовить к прохождению учебной и производственной практик.

Содержание дисциплины охватывает следующий круг вопросов: характеристика понятий «данные», «информация», «знания»; проблема «больших данных»; обзор источников информации; основные определения, термины и задачи анализа больших данных; характеристика больших данных – 5V; предпосылки формирования тренда; драйверы рынка больших данных; обзор технологий хранения и обработки больших данных; современные программные средства анализа больших данных; процесс аналитики (стандарт *CRISP-DM*); понятие машинного обучения и его компоненты; основные определения и термины машинного обучения; виды задач в машинном обучении и процесс их решения; обучение без учителя: кластерный анализ;

обучение с учителем: логистическая регрессия, деревья решений, регрессионный анализ.

Цель – формирование у студентов системного представления процессах сбора, хранения и анализа больших объемов данных, а также подготовка обучающихся к работе с большими данными.

Задачи:

- приобретение студентами знаний о технологиях сбора, хранения, обработки и анализа больших данных;
- развитие навыков создания и решения моделей, необходимых в сфере управления;
- изучение методов машинного обучения для анализа данных больших объемов.

Для успешного изучения дисциплины «Наука о данных и аналитика больших объемов данных» у обучающихся должны быть сформированы следующие предварительные компетенции:

- способность осуществлять сбор, анализ и обработку данных, необходимых для решения профессиональных задач;
- способность выбрать инструментальные средства для обработки экономических данных в соответствии с поставленной задачей, проанализировать результаты расчетов и обосновать полученные выводы;
- способность на основе описания экономических процессов и явлений строить стандартные теоретические и эконометрические модели, анализировать и содержательно интерпретировать полученные результаты;
- способность, используя отечественные и зарубежные источники информации, собрать необходимые данные, проанализировать их и подготовить информационный обзор и / или аналитический отчет.

В результате изучения данной дисциплины у обучающихся формируются следующие общекультурные и профессиональные компетенции (элементы компетенций):

Код и формулировка компетенции	Этапы формирования компетенции	
ОК-12 – способность работать с различными информационными ресурсами и технологиями, применять основные методы, способы и средства получения, хранения, поиска, систематизации, обработки и передачи информации	знает	основные направления развития рынка больших данных
	умеет	формулировать задачи в сфере анализа больших объемов данных
	владеет	методологией ведения проектов интеллектуального анализа данных
ПК-5 – способность осуществлять плано-отчетную работу организации, разработку проектных решений, разделов текущих и перспективных планов экономического развития организации, бизнес-планов, смет, учетно-отчетной документации, нормативов затрат и соответствующих предложений по реализации разработанных проектов, планов, программ	знает	технологии хранения и обработки больших данных
	умеет	проводить сравнительный анализ и выбор современных ИКТ для решения прикладных задач
	владеет	навыками использования современных ИКТ для решения прикладных задач

Для формирования вышеуказанных компетенций в рамках дисциплины «Наука о данных и аналитика больших объемов данных» применяются следующие методы активного/ интерактивного обучения: лекция-презентация, лекция с разбором конкретных ситуаций.

I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

Лекционные занятия (8 час.)

МОДУЛЬ 1. ЭНТРОПИЯ И ИНФОРМАЦИЯ

Раздел I. Энтропия и информация в дискретном случае (2 час.)

Тема 1. Энтропия дискретных событий (0.5 час.)

Необходимые сведения из теории вероятности. Энтропия как мера неопределенности опыта. Свойства энтропии. Единицы измерения неопределенности. Теорема: как энтропия одного исхода, так и средняя энтропия всегда не отрицательны. Теорема: Энтропия имеет максимальное

значение, когда исходы опыта равновероятны. Статистический смысл энтропии.

Тема 2. Энтропия сложных событий (0.5 час.)

Энтропия сложных событий. Теорема: Если случайные величины α и β независимы, то полная (совместная) энтропия $H(\alpha\beta)$ распадается на сумму энтропий $H(\alpha) + H(\beta)$. Теорема: Энтропия обладает свойством иерархической аддитивности. Теорема: Условная энтропия не может превосходить безусловную.

Тема 3. Информация и энтропия в дискретном случае (1 час.)

Определение информации. Средняя взаимная информация. Собственная информация. Количественная мера информации. Теорема: пусть $\alpha\beta$ дискретный совместный ансамбль. Для средней взаимной информации между α и β справедливо $I(\alpha,\beta) \geq 0$.

Раздел II. Энтропия и информация в непрерывном случае (1 час.)

Тема 1: Информация и энтропия в непрерывном случае (1 час.)

Эпсилон энтропия. Дифференциальная энтропия. Информация для непрерывного опыта. Дифференциальная энтропия случайного процесса с нормальным распределением.

МОДУЛЬ 2. КАНАЛЫ СВЯЗИ И КОДИРОВАНИЕ

Раздел I. Каналы связи (2 час.)

Тема 1: Каналы связи (0.5 час.)

Передача сообщений по линиям связи. Блок-схема системы связи с кодером и декодером. Определения: Канал связи, емкость канала, Источник информации или сообщения, Сообщение, Алфавит, Код, Кодирование, Основание кода. Характеристики линии связи без помех.

Тема 2: Передача информации по каналам связи без помех (0.5 час.)

Передача сообщений без помех. Пропускная способность линии. Характеристики линии связи с помехами. Теорема Шеннона о кодировании в отсутствии помех.

Тема 3: Передача информации по каналам связи с помехами (1 час.)

Передача сообщений при наличии помех. Пропускная способность линии связи с помехами. Характеристики линии связи с помехами. Двоичная симметричная линия связи. Двоичная несимметричная линия связи. Линия связи со стиранием. Теорема Шеннона о кодировании при наличии помех. Пропускная способность непрерывного канала при наличии аддитивного шума.

Раздел II. Кодирование (3 час.)

Тема 1: Основы кодирования (0.5 час.)

Код Морзе и код Бодо. Коды: Двоичные, троичные и т.п. (по основанию кода); Равномерные, не равномерные; С запятой и без запятой; Префиксные.

Тема 2: Коды Шеннона-Фано и Хаффмана (0.5 час.)

Коды Шеннона — Фано и Хаффмана. Оптимальность кода Хаффмана. Код Гилберта-Мура. Основная теорема о кодировании при отсутствии помех.

Тема 3: Коды обнаруживающие и исправляющие ошибки (1 час.)

Коды обнаруживающие и исправляющие одиночные ошибки. Эффективность кода. Проверки на четность. Коды обнаруживающие и исправляющие двойные ошибки. Линейные (N, M) коды. Определение количества контрольных сигналов.

Тема 4: Неравенство Хемминга, неравенство Варшамова-Гилберта (0,5 час.)

неравенством Варшамова — Гилберта или верхней границей Варшамова — Гилберта. общий (N, M) -код с проверками на четность. систематическими кодами с проверками на четность. линейными кодами или групповыми кодами. Циклические коды. Верхняя граница Хемминга числа кодовых обозначений.

Тема 5: Алгоритмы сжатия информации (0,5 час.)

Словарные методы. Алгоритм RLE. Алгоритмы группы LZ.

II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА

Практические занятия не предусмотрены.

III. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине «Наука о данных и аналитика больших данных» 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.

IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций	Оценочные средства		
			текущий контроль	промежуточная аттестация	
1	Модуль 1 Раздел I. Энтропия и информация в дискретном случае	ПК-10, ПК-11, ПК-13	знает	ПР-7	1-4
			умеет	ПР-13	5-7
			владеет	УО-3	<i>Решение задач</i>
	Модуль 1 Раздел II. Энтропия и информация в непрерывном случае	ПК-10, ПК-11, ПК-13	знает	ПР-7	8,9
			умеет	ПР-13	10-11
			владеет	УО-3	<i>Решение задач</i>
	Модуль 2 Раздел I. Каналы связи	ПК-10, ПК-11, ПК-13	знает	ПР-7	12-14
			умеет	ПР-13	15-16
			владеет	УО-3	<i>Решение задач</i>
	Модуль 2 Раздел II. Кодирование	ПК-10, ПК-11, ПК-13	знает	ПР-7	12, 21-23
			умеет	ПР-13	17- 20, 24-26
			владеет	УО-3	<i>Решение задач</i>

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

(электронные и печатные издания)

1. **Наука о данных и статистика в образовании будущего** : научное издание / В.Н. Салин. — Москва : КноРус, 2018. — 300 с. — ISBN 978-5-4365-2866-3. — <https://www.book.ru/book/931360>

2. Миронов, М. М. Методы и средства исследований [Электронный ресурс] : учебное пособие / М. М. Миронов, Л. Р. Джанбекова. — Электрон. текстовые данные. — Казань : Казанский национальный исследовательский технологический университет, 2009. — 80 с. — 978-5-7882-0654-7. — <http://www.iprbookshop.ru/62490.html>

3. Пальмов, С. В. Интеллектуальный анализ данных [Электронный ресурс] : учебное пособие / С. В. Пальмов. — Электрон. текстовые данные. — Самара : Поволжский государственный университет телекоммуникаций и информатики, 2017. — 127 с. — 2227-8397. — <http://www.iprbookshop.ru/75376.html>

4. Анализ данных качественных исследований [Электронный ресурс] : практикум / сост. А. П. Истомина. — Электрон. текстовые данные. — Ставрополь : Северо-Кавказский федеральный университет, 2016. — 94 с. — 2227-8397.— <http://www.iprbookshop.ru/66014.html>

5. Уэс, Маккинли Python и анализ данных [Электронный ресурс] / Маккинли Уэс ; пер. А. А. Слинкин. — Электрон. текстовые данные. — Саратов : Профобразование, 2017. — 482 с. — 978-5-4488-0046-7. — <http://www.iprbookshop.ru/64058.html>

Дополнительная литература (печатные и электронные издания)

1. Мельниченко, А. С. Математическая статистика и анализ данных [Электронный ресурс] : учебное пособие / А. С. Мельниченко. — Электрон. текстовые данные. — М. : Издательский Дом МИСиС, 2018. — 45 с. — 978-5-906953-62-9. — <http://www.iprbookshop.ru/78563.html>

2. Жуковский, О. И. Информационные технологии и анализ данных [Электронный ресурс] : учебное пособие / О. И. Жуковский. — Электрон.

текстовые данные. — Томск : Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2014. — 130 с. — 978-5-4332-0158-3. — <http://www.iprbookshop.ru/72106.html>

3. Анализ данных : учебник для академического бакалавриата / В. С. Мхитарян [и др.] ; под ред. В. С. Мхитаряна. — Москва : Издательство Юрайт, 2018. — 490 с. — (Серия : Бакалавр. Академический курс). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — <https://www.biblio-online.ru/bcode/412967>

4. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2018. — 174 с. — (Серия : Авторский учебник). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — <https://www.biblio-online.ru/bcode/413060>

Перечень информационных технологий и программного обеспечения

Аудитория с мультимедиапроектором, ПК с MSOfficeили LibreOffice.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Всемирная торговая организация (ВТО) – World Trade Organization (WTO) – www.wto.org
2. Всемирный банк – World Bank Group – www.worldbank.org
3. Европейский союз (ЕС) – The European Union (EU) – europa.eu.int
4. Конференция ООН по торговле и развитию (ЮНКТАД) – United Nations Conference on Trade and Development – www.unctad.org
5. Комиссия ООН по праву в международной торговле (ЮНСИТРАЛ)– United Nations Commission on International Trade Law (UNCITRAL) – www.uncitral.org

6. Международная организация ООН по промышленному развитию (ЮНИДО) – United Nations Industrial Development Organization (UNIDO) – www.unido.org
7. Международный банк реконструкции и развития (МБРР) – International Bank for Reconstruction and Development (IBRD) – www.worldbank.com
8. Международный валютный фонд (МВФ) – International Monetary Fund (IMF) – www.imf.org
9. Организация ООН по вопросам образования, науки и культуры (ЮНЕСКО) – United Nations for Educational Scientific and Cultural Organization (UNESCO) – www.unesco.org
10. Организация экономического сотрудничества и развития (ОЭСР) – Organization for Economic Cooperation and Development – (OECD) – www.oecd.org
11. Организация азиатско-тихоокеанского экономического сотрудничества – Asia-Pacific Economic Cooperation (APEC) – www.apecsec.org.sg
12. Организация Объединенных Наций (ООН) – United Nations (UN) – www.un.org
13. Правительство Российской Федерации – www.government.ru
14. Официальный сайт Президента Российской Федерации – www.kremlin.ru
15. Национальная электронная библиотека [Электронный ресурс].- Режим доступа: www.nns.ru.
16. Российская государственная библиотека [Электронный ресурс]. - Режим доступа: www.rsl.ru.
17. Российская национальная библиотека [Электронный ресурс]. - Режим доступа: www.nlr.ru.
18. Североамериканская ассоциация свободной торговли (НАФТА) – North America Free Trade Association (NAFTA) – www.nafta-sec-alena.org

19. Торгово-промышленная палата РФ – www.tpprf.ru

Перечень информационных технологий и программного обеспечения

1. Справочно-правовая система «КонсультантПлюс». Режим доступа: <http://www.consultant.ru/>
2. Справочно-правовая система «Гарант». Режим доступа: www.garant.ru
3. Справочная система «Кодекс». Режим доступа: <http://www.kodeks.ru/>
4. Программное обеспечение: Microsoft Word, Microsoft Excel, Microsoft PowerPoint.

VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

При изучении дисциплины 8 часов отводится на аудиторные занятия и 100 часов на самостоятельную работу студентов. С целью более эффективного освоения материала рекомендуется практические занятия проводить в конце семестра, когда студенты будут знакомы с теоретическим содержанием курса.

При подготовке к зачету рекомендуется просмотреть материалы лекций и собственные конспекты, разбить вопросы по разделам и темам, затем определить содержание ответов на вопросы.

Выполнение заданий в малых группах и творческого задания предусматривает выбор подраздела курса и согласование его с преподавателем.

VII. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Реализация дисциплины «Наука о данных и аналитика больших данных» предусматривает следующие виды учебной работы: лекции, практические работы, самостоятельную работу студентов, текущий контроль и промежуточную аттестацию.

Освоение курса дисциплины «Наука о данных и аналитика больших данных» предполагает рейтинговую систему оценки знаний студентов и предусматривает со стороны преподавателя текущий контроль за посещением

студентами лекций, подготовкой и выполнением всех домашних работ с обязательным предоставлением отчета о работе, выполнением всех видов самостоятельной работы.

Промежуточной аттестацией по дисциплине «Наука о данных и аналитика больших данных» является зачет, который проводится в виде контрольной работы.

В течение учебного семестра обучающимся нужно:

- освоить теоретический материал (20 баллов);
- успешно выполнить аудиторные и контрольные задания (50 баллов);
- своевременно и успешно выполнить все виды самостоятельной работы (30 баллов).

Студент считается аттестованным по дисциплине «Наука о данных и аналитика больших данных» при условии выполнения всех видов текущего контроля и самостоятельной работы, предусмотренных учебной программой.

Критерии оценки по дисциплине «Наука о данных и аналитика больших данных» для аттестации на зачете следующие: 86-100 баллов – «отлично», 76-85 баллов – «хорошо», 61-75 баллов – «удовлетворительно», 60 и менее баллов – «неудовлетворительно».

Пересчет баллов по текущему контролю и самостоятельной работе производится по формуле:

$$P(n) = \sum_{i=1}^m \left[\frac{O_i}{O_i^{max}} \times \frac{k_i}{W} \right],$$

где: $W = \sum_{i=1}^n k_i^n$ для текущего рейтинга;

$W = \sum_{i=1}^m k_i^n$ для итогового рейтинга;

$P(n)$ – рейтинг студента;

m – общее количество контрольных мероприятий;

n – количество проведенных контрольных мероприятий;

O_i – балл, полученный студентом на i -ом контрольном мероприятии;

O_i^{max} – максимально возможный балл студента по i -му контрольному

мероприятию;

k_i – весовой коэффициент i -го контрольного мероприятия;

k_i^n – весовой коэффициент i -го контрольного мероприятия, если оно является основным, или 0, если оно является дополнительным.

VIII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Сведения о материально-техническом обеспечении и оснащенности образовательного процесса: лекционные и практические занятия по дисциплине проходят в аудиториях, оборудованных компьютерами типа Lenovo C360G-i34164G500UDK с лицензионными программами MicrosoftOffice 2010 и аудио-визуальными средствами проектор Panasonic DLPProjectorPT-D2110XE, плазма LG FLATRON M4716CCBAM4716CJ. Для выполнения самостоятельной работы студенты в жилых корпусах ДВФУ обеспечены Wi-Fi.

В читальных залах Научной библиотеки ДВФУ предусмотрены рабочие места для людей с ограниченными возможностями здоровья, оснащены дисплеями и принтерами Брайля; оборудованные портативными устройствами для чтения плоскочечатных текстов, сканирующими и читающими машинами, видеоувелечителем с возможностью регуляции цветовых спектров; увеличивающими электронными лупами и ультразвуковыми маркировщиками.

В целях обеспечения специальных условий обучения инвалидов и лиц с ограниченными возможностями здоровья ДВФУ все здания оборудованы пандусами, лифтами, подъемниками, специализированными местами, оснащенными туалетными комнатами, табличками информационно-навигационной системы.



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

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

**УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ
РАБОТЫ ОБУЧАЮЩИХСЯ**
по дисциплине «Наука о данных и аналитика больших данных»

Специальность 38.05.01 Экономическая безопасность
Специализация: Экономико-правовое обеспечение экономической безопасности
Форма подготовки очная

Владивосток
2019

План-график выполнения самостоятельной работы по дисциплине

№ п/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
1	еженедельно	Домашнее задание	30	Проверка ДЗ
2	12-16 неделя обучения	Подготовка творческого задания	60	Выступление по результатам
3	18 неделя, зачетная неделя	Подготовка к итоговому контролю	10	зачет
	Итого		100	

Самостоятельная работа по выполнению домашнего задания должна включать в себя повторение лекционного материала, повторение формул по разделу, повторение решенных задач по разделу, решение задач домашнего задания по разделу. Выполненное задание должно быть оформлено в соответствии с требованиями по оформлению решения задач, текст, формулы легко читаемы.

Самостоятельная работа по подготовке творческого задания должна включать в себя поиск информации в сети Интернет и рекомендуемых источниках, обсуждение основных характеристик, подготовка черновиков презентаций и текста выступления. Презентация должна быть не менее 15 слайдов, выступление продолжительностью 15-18 минут. Каждая малая группа должна подготовить вопросы для остальных групп по их темам.

Самостоятельная работа по подготовке к экзамену должна включать повторение теоретического материала, подготовку ответов на вопросы с использованием лекций и рекомендуемых источников.

Оценка результатов самостоятельной работы по подготовке творческого задания выполняется по следующим критериям:

5 баллов выставляется, если студент выразил своё мнение по сформулированной проблеме, аргументировал его, точно определив ее содержание и составляющие. Приведены данные отечественной и зарубежной литературы, статистические сведения, информация нормативно-правового характера. Продемонстрированы знания и владения навыками самостоятельной исследовательской работы по теме исследования; методами и приемами анализа теоретических и/или практических аспектов изучаемой области. Фактических ошибок, связанных с пониманием проблемы, нет; графически работа оформлена правильно

4 балла - работа характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Для аргументации приводятся данные отечественных и зарубежных авторов. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет. Допущены одна-две ошибки в оформлении работы

3 балла - Студент провел достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимает базовые основы и теоретическое обоснование выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы, оформлении работы

2 балла - 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 в смысловом содержании раскрываемой проблемы, в оформлении работы.

Оценка результатов самостоятельной работы в малых группах

выполняется по следующим критериям:

5 баллов выставляется, если студент/группа выразили своё мнение по сформулированной проблеме, аргументировали его, точно определив ее содержание и составляющие. Приведены данные отечественной и зарубежной литературы, статистические сведения, информация нормативно-правового характера. Продемонстрировано знание и владение навыком самостоятельной исследовательской работы по теме исследования; методами и приемами анализа международно-политической практики. Фактических ошибок, связанных с пониманием проблемы, нет

4 балла - работа студента/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Для аргументации приводятся данные отечественных и зарубежных авторов. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

3 балла - 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

2 балла - 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.

Критерии оценки выполнения самостоятельной работы:

Баллы (рейтингов ой оценки)	Оценка (стандартная)	Требования к сформированным компетенциям
100- 85	<i>«отлично»</i>	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал конкретной темы, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач.
84-76	<i>«хорошо»</i>	Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.
75-61	<i>«удовлетворительно»</i>	Оценка «удовлетворительно» выставляется студенту, если он 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.
60-50	<i>«неудовлетворительно»</i>	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по дисциплине.



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

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по дисциплине «Наука о данных и аналитика больших данных»
Специальность 38.05.01 Экономическая безопасность
Специализация: Экономико-правовое обеспечение экономической безопасности
Форма подготовки очная

Владивосток
2019

Паспорт фонда оценочных средств

Код и формулировка компетенции	Этапы формирования компетенции	
ОК-12 – способность работать с различными информационными ресурсами и технологиями, применять основные методы, способы и средства получения, хранения, поиска, систематизации, обработки и передачи информации	знает	основные направления развития рынка больших данных
	умеет	формулировать задачи в сфере анализа больших объемов данных
	владеет	методологией ведения проектов интеллектуального анализа данных
ПК-5 – способность осуществлять планово-отчетную работу организации, разработку проектных решений, разделов текущих и перспективных планов экономического развития организации, бизнес-планов, смет, учетно-отчетной документации, нормативов затрат и соответствующих предложений по реализации разработанных проектов, планов, программ	знает	технологии хранения и обработки больших данных
	умеет	проводить сравнительный анализ и выбор современных ИКТ для решения прикладных задач
	владеет	навыками использования современных ИКТ для решения прикладных задач

Описание критериев и показателей формирования компетенций по дисциплине " Наука о данных и аналитика больших данных "

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства	
				текущий контроль	промежуточная аттестация
1.	Занятия: лекция 1-2	ОК-12 – способность работать с различными информационными ресурсами и технологиями, применять основные	Знает основные направления развития рынка больших данных	Конспект (ПР-7), дискуссия (УО-4)	Контрольная работа 1 (ПР-2), вопросы к экзамену №1-3
			Умеет формулировать задачи в сфере анализа больших объемов данных	Конспект (ПР-7), собеседование (УО-1), дискуссия (УО-4)	Контрольная работа 1 (ПР-2), вопросы к экзамену №4-6
			Владеет методологией ведения проектов интеллектуального анализа данных	Конспект (ПР-7), собеседование (УО-1),	Контрольная работа 1 (ПР-2), вопросы к

		методы, способы и средства получения, хранения, поиска, систематизации, обработки и передачи информации		дискуссия (УО-4)	экзамену №7-8
2.	Занятия: лекция 3-4	ПКПК-5 – способность осуществлять планово-отчетную работу организации, разработку проектных решений, разделов текущих и перспективных планов экономического развития организации, бизнес-планов, смет, учетно-отчетной	Знает технологии хранения и обработки больших данных	Конспект (ПР-7), дискуссия (УО-4)	Контрольная работа 1 (ПР-2), вопросы к экзамену №9-10
			Умеет проводить сравнительный анализ и выбор современных ИКТ для решения прикладных задач	Конспект (ПР-7), собеседование (УО-1), дискуссия (УО-4)	Контрольная работа 1 (ПР-2), вопросы к экзамену №11-12
			Владеет навыками использования современных ИКТ для решения прикладных задач	Конспект (ПР-7), собеседование (УО-1), дискуссия (УО-4)	Контрольная работа 1 (ПР-2), вопросы к экзамену №13-15

		документации, нормативов затрат и соответствующих предложений по реализации и разработанных проектов, планов, программ			
3.	Занятия: лекция 5-8	ПК-5 – способность осуществлять планово-отчетную работу организации, разработку проектных решений, разделов текущих и перспективных планов экономического развития организации, бизнес-планов, смет, учетно-отчетной документации, норматив	Знает технологии хранения и обработки больших данных	Конспект (ПР-7), собеседование (УО-1), дискуссия (УО-4)	Контрольная работа 1 (ПР-2), вопросы к экзамену №16-17
			Умеет проводить сравнительный анализ и выбор современных ИКТ для решения прикладных задач	Конспект (ПР-7), собеседование (УО-1), дискуссия (УО-4)	Контрольная работа 2 (ПР-2), вопросы к экзамену №18-20
			Владеет навыками использования современных ИКТ для решения прикладных задач	Конспект (ПР-7), дискуссия (УО-4)	Контрольная работа 2 (ПР-2), вопросы к экзамену №21-23

		ов затрат и соответствующих предложений по реализации разработанных проектов, планов, программ			
--	--	--	--	--	--

Шкала оценивания уровня сформированности компетенций

Код и формулировка компетенции	Этапы формирования компетенции		Критерии	Показатели
ОК-12 – способность работать с различными информационными ресурсами и технологиями, применять основные методы, способы и средства получения, хранения, поиска, систематизации, обработки и передачи информации	знает (пороговый уровень)	основные направления развития рынка больших данных	знает основные методы, способы и средства получения, хранения, поиска, систематизации, обработки и передачи информации	способность применять основные методы, способы и средства получения, хранения, поиска, систематизации, обработки и передачи информации
	умеет (продвинутый)	формулировать задачи в сфере анализа больших объемов данных	умеет использовать информационные системы для решения профессиональных задач	способность самостоятельно использовать информационные системы для решения профессиональных задач
	владеет (высокий)	методологией ведения проектов интеллектуального анализа данных	владеет навыками обеспечения защиты информации, составляющей государственную, и иную служебную информацию	способность использовать навыки обеспечения защиты информации, составляющей государственную, и иную служебную информацию
ПК-5 – способность осуществлять планово-	знает (пороговый уровень)	технологии хранения и обработки больших данных	знание правовых, нормативных и (законодательные и нормативные акты рф, и др.), необходимых для осуществления	– способность перечислить и охарактеризовать правовые акты, регламентирующие финансовую деятельность; – способность грамотно интерпретировать показатели

<p>отчетную работу организации, разработку проектных решений, разделов текущих и перспективных планов экономического развития организации, бизнес-планов, смет, учетно-отчетной документации, нормативов в затрат и соответствующих предложений по реализации разработанных проектов, планов, программ</p>			финансовой деятельностью	финансовой отчетности предприятия
	умеет (продвинутый)	проводить сравнительный анализ и выбор современных ИКТ для решения прикладных задач	знает основные финансовые показатели для расчета бизнес-планов и критерии деловой и инвестиционной привлекательности	способность проводить расчеты финансовых показателей, составлять финансовую, управленческую отчетность и финансовые отчеты с прогнозированием и построением ожиданий
	владеет (высокий)	навыками использования современных ИКТ для решения прикладных задач	владеет основными критериями эффективности проектами	способность принимать эффективные управленческие решения на основе полученной финансовой информации

Зачетно-экзаменационные материалы

Методические материалы, определяющие процедуры оценивания знаний, умений и навыков

Вопросы к промежуточной аттестации (зачет)

1. Энтропия как мера неопределенности.

2. Введение понятия энтропии.
3. Мера неопределенности по Хартли. Мера неопределенности по Шеннону.
4. Статистический смысл энтропии.
5. Средняя условная энтропия двух опытов. Свойства.
6. Энтропия сложных событий. Правило сложения энтропии зависимые события.
7. Энтропия сложных событий. Правило сложения энтропии независимые события.
8. Дифференциальная энтропия, дифференциальная энтропия нормального распределения.
9. Энтропия источников непрерывных сообщений.
10. Понятие об информации. Соотношение понятий энтропии и информации.
11. Информация в случае непрерывного опыта.
12. Основные понятия передачи информации по линиям связи. Определение кода. Код Морзе. Код Бодо.
13. Линия связи без помех.
14. Линия связи с помехами.
15. Пропускная способность линии связи с помехами.
16. Двоичная симметричная линия. Двоичная симметричная линия со стиранием.
17. Основная теорема о кодировании.
18. Теорема о кодировании при наличии помех.
19. Формула Шеннона для пропускной способности непрерывного канала при наличии аддитивного шума.
20. Мгновенные коды. Равномерные и неравномерные коды. Экономность кода.
21. Коды Шеннона — Фано и Хаффмана, доказательство оптимальности кодов Хаффмана.

22. Коды исправляющие все одиночные ошибки, (N, M) –коды.
23. Коды обнаруживающие и исправляющие ошибки.
24. Коды обнаруживающие ошибки, коды исправляющие ошибки, (N, M) –коды.
25. Неравенство Хемминга, неравенство Варшамова-Гилберта.
26. Верхняя граница Хемминга числа кодовых обозначений.

Типовые задания для текущего контроля

1. Имеется два игральных кубика, определите вероятность выпадения числа семь и энтропию такого опыта.
2. Ключ шифрования содержит 256 бит, сколько времени понадобится для его вскрытия методом перебора, если известно, что за одну секунду компьютер перебирает 10000 вариантов.
3. 1 сентября на первом курсе одного из факультетов запланировано по расписанию три лекции по разным предметам. Всего на первом курсе изучается 7 предметов. Студент, не успевший ознакомиться с расписанием, пытается его угадать. Какова вероятность успеха в данном эксперименте, если считать, что любое расписание из четырех предметов равновероятно? Какова неопределенность такого опыта?
4. В подъезде дома установлен замок с кодом. Дверь автоматически отпирается, если в определенной последовательности набрать три цифры из имеющихся десяти. На набор одной комбинации уходит 20 секунд. Какова вероятность открыть дверь за 40 минут? Какова неопределенность такого опыта?
5. Из набора домино (28 штук) выбирают 7 штук. Какова неопределенность данного события?
6. Из колоды карт (52 карты) извлекают три карты. Какова неопределенность такого события?
7. Бросают две игральные кости. Какова неопределенность данного события?

8. Имеется 12 монет одного достоинства, 11 из них имеют одинаковый вес, а одна — фальшивая — отличается по весу от остальных (причем неизвестно, легче ли она или тяжелее настоящих). Каково наименьшее число взвешиваний на чашечных весах без гирь, которое позволяет обнаружить фальшивую монету и выяснить, легче ли она, чем остальные монеты, или тяжелее. Решить тот же вопрос для случая 13 монет.

9. Какова вероятность того, что взятое наугад целое положительное число до 1000, окажется целой степенью другого целого числа с показателем, больше единицы? Вычислите энтропию такого события.

10. Какую неопределенность содержит сообщение о событие – сдача экзамена студентом, если по опыту предыдущих экзаменов известно, что вероятность того, что студент сдал экзамен $7/8$.

11. Пусть X и Y два случайных опыта; $Z=X+Y$. Чему равна условная энтропия $H(x|z)$, если: а) X и Y независимы; б) X и Y зависимы; в) $X \equiv Y$.

12. Определить среднее количество информации, приходящееся на один символ сообщения 01001000101001, при условии, что источник эргодический, а последовательность типичная.

13. Имеются два дискретных источника с независимыми и равновероятными элементами: двоичный и троичный. На выходе первого зафиксированы два символа, на выходе второго три. Чему равны неопределенности полученных последовательностей букв, образованных парами символов первого источника и тройками символов второго?

14. Символы азбуки Морзе могут появиться в сообщении с вероятностями: для точки - 0.51, для тире - 0.31, для промежутка между буквами - 0.12, между словами - 0.06. Определить среднее количество информации в сообщении из 500 символов данного алфавита, считая, что связь между последовательными символами отсутствует.

15. Известно, что из 100 изготовленных деталей в среднем 10 деталей имеют дефекты. Для выявления брака используется метод, дающий всегда отрицательный эффект, если деталь изготовлена с браком. Если брак

отсутствует, то деталь признается годной лишь в 80% случаев. Какое количество информации о качестве детали можно получить в среднем по результату такого метода отбраковки?

16. Орудие стреляет по удаленной цели. При каждом выстреле она поражается с вероятностью $p = 0.1$. Разведка может только один раз проверить, поражается ли цель хоть один раз. Через некоторое количество выстрелов k следует провести проверку, чтобы она дала максимальное количество информации?

17. Алфавит состоит из двух букв А, Б, В появляющихся в тексте с вероятностями 0.6, 0.3 и 0.1 соответственно. Закодировать отдельные буквы равномерным кодом. Закодировать пары и тройки букв равномерным кодом. Повторить кодирование одной буквы с не равномерным кодом. Сравнить эффективность кодов. Построить кодовое дерево для не равномерного кода

18. По линии связи передаются сообщения из 5-ти равновероятных букв. Закодировать буквы равномерным кодом. Закодировать тройки букв равномерным кодом. Сравнить эффективности кодов.

19. Имеются два дискретных троичных источника с независимыми элементами. На выходе каждого источника появляются сообщения одинаковой длины по 25 элементов. Количество различных элементов в сообщении каждого источника постоянно. Сообщения каждого источника отличаются только порядком элементов. Зафиксированы два типичных сообщения: 0212021202120211201120200 – первого источника и 0121012011012012210200120 – второго. Элемент какого источника несет в среднем большее количество информации?

20. Сообщения с вероятностями 0,5; 0,25; 0,0625; 0,0625; 0,0625; 0,0625 кодируются одним из шести различных кодов: 1) 0-10-110-1110-1011-1101; 2) 1-011-010-001-000-110; 3) 0-10-110-1110-11110-111110; 4) 111-110-101-100-011-010; 5) 0-01-011-0111-01111-01111; 6) 1-01-0011-0010-0001-0000. Определить, какие коды являются разделимыми (мгновенными). Вычислить характеристики кодов.

Критерии оценки (письменного/устного доклада, сообщения, в том числе выполненных в форме презентаций):

✓ 100-86 баллов выставляется студенту, если студент выразил своё мнение по сформулированной проблеме, аргументировал его, точно определив ее содержание и составляющие. Приведены данные отечественной и зарубежной литературы, статистические сведения, информация нормативно-правового характера. Студент знает и владеет навыком самостоятельной исследовательской работы по теме исследования; методами и приемами анализа теоретических и практических аспектов изучаемой области. Фактических ошибок, связанных с пониманием проблемы, нет; графически работа оформлена правильно

✓ 85-76 - баллов - работа характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Для аргументации приводятся данные отечественных и зарубежных авторов. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет. Допущены одна-две ошибки в оформлении работы

✓ 75-61 балл – студент проводит достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимает базовые основы и теоретическое обоснование выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы, оформлении работы

✓ 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 в смысловом содержании раскрываемой проблемы, в оформлении работы.

Критерии оценки (письменный ответ)

✓ 100-86 баллов - если ответ показывает глубокое и систематическое знание всего программного материала и структуры конкретного вопроса, а также основного содержания и новаций лекционного курса по сравнению с учебной литературой. Студент демонстрирует отчетливое и свободное владение концептуально-понятийным аппаратом, научным языком и терминологией соответствующей научной области. Знание основной литературы и знакомство с дополнительно рекомендованной литературой. Логически корректное и убедительное изложение ответа.

✓ 85-76 - баллов - знание узловых проблем программы и основного содержания лекционного курса; умение пользоваться концептуально-понятийным аппаратом в процессе анализа основных проблем в рамках данной темы; знание важнейших работ из списка рекомендованной литературы. В целом логически корректное, но не всегда точное и аргументированное изложение ответа.

✓ 75-61 - балл – фрагментарные, поверхностные знания важнейших разделов программы и содержания лекционного курса; затруднения с использованием научно-понятийного аппарата и терминологии учебной дисциплины; неполное знакомство с рекомендованной литературой; частичные затруднения с выполнением предусмотренных программой заданий; стремление логически определено и последовательно изложить ответ.

✓ 60-50 баллов – незнание, либо отрывочное представление о данной проблеме в рамках учебно-программного материала; неумение использовать понятийный аппарат; отсутствие логической связи в ответе.

Критерии оценки (устный ответ)

✓ 100-85 баллов - если ответ показывает прочные знания основных процессов изучаемой предметной области, отличается глубиной и полнотой раскрытия темы; владение терминологическим аппаратом; умение объяснять

сущность, явлений, процессов, событий, делать выводы и обобщения, давать аргументированные ответы, приводить примеры; свободное владение монологической речью, логичность и последовательность ответа; умение приводить примеры современных проблем изучаемой области.

✓ 85-76 - баллов - ответ, обнаруживающий прочные знания основных процессов изучаемой предметной области, отличается глубиной и полнотой раскрытия темы; владение терминологическим аппаратом; умение объяснять сущность, явлений, процессов, событий, делать выводы и обобщения, давать аргументированные ответы, приводить примеры; свободное владение монологической речью, логичность и последовательность ответа. Однако допускается одна - две неточности в ответе.

✓ 75-61 - балл – оценивается ответ, свидетельствующий в основном о знании процессов изучаемой предметной области, отличающийся недостаточной глубиной и полнотой раскрытия темы; знанием основных вопросов теории; слабо сформированными навыками анализа явлений, процессов, недостаточным умением давать аргументированные ответы и приводить примеры; недостаточно свободным владением монологической речью, логичностью и последовательностью ответа. Допускается несколько ошибок в содержании ответа; неумение привести пример развития ситуации, провести связь с другими аспектами изучаемой области.

✓ 60-50 баллов – ответ, обнаруживающий незнание процессов изучаемой предметной области, отличающийся неглубоким раскрытием темы; незнанием основных вопросов теории, несформированными навыками анализа явлений, процессов; неумением давать аргументированные ответы, слабым владением монологической речью, отсутствием логичности и последовательности. Допускаются серьезные ошибки в содержании ответа; незнание современной проблематики изучаемой области.

Критерии выставления оценки студенту на экзамене по дисциплине

Баллы (рейтингов ой оценки)	Оценка зачета (стандартна я)	Требования к сформированным компетенциям
100- 85	<i>«отлично» /зачтено</i>	Оценка «отлично» выставляется студенту, если он 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, использует в ответе материал монографической литературы, правильно обосновывает принятое решение, владеет разносторонними навыками. Обладает способностью к абстрактному мышлению, анализу, синтезу, а также способностью проводить самостоятельные исследования в соответствии с разработанной программой
84-76	<i>«хорошо»/ зачтено</i>	Оценка «хорошо» выставляется студенту, если он 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. Обладает способностью к абстрактному мышлению, анализу, синтезу.
75-61	<i>«удовлетво рительно» /зачтено</i>	Оценка «удовлетворительно» выставляется студенту, если он 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.
60-50	<i>«неудовлет ворительно » / не зачтено</i>	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.