



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

ШКОЛА ЕСТЕСТВЕННЫХ НАУК

СОГЛАСОВАНО:
Руководитель ОП


И.А. Текутьева
«30 сентября» 2018 г.

«УТВЕРЖДАЮ»
Заведующий кафедрой
Информатики, математического и
компьютерного моделирования


Чеботарев А.Ю.
«30 сентября» 2018 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наука о данных и аналитика больших объемов данных

Направление подготовки 38.04.01 Экономика

магистерская программа «Биоэкономика и продовольственная безопасность»
Форма подготовки очная

курс 1 семестр 1
лекции час.
практические занятия 8 час.
лабораторные работы 0 час.
в том числе с использованием МАО /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 1 семестр
экзамен не предусмотрен

Рабочая программа составлена в соответствии с требованиями образовательного стандарта, самостоятельно устанавливаемого ДВФУ по направлению подготовки 38.04.01 Экономика, принят решением Ученого совета ДВФУ, протокол от 04.06.2015 г. № 06-15 и введен в действие приказом ректора ДВФУ от 07.07.2015 № 12-13-1282.

Рабочая программа обсуждена на заседании кафедры Информатики, математического и компьютерного моделирования, протокол № 10 от 13.09 2018 г.

Заведующий кафедрой: Чеботарев А.Ю.
Составитель: канд. техн. наук Смагин С.В.

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)

ABSTRACT

Master's degree in 38.04.01 Economics

Master's Program «Bioeconomy and food security»

Course title: “Data Science and Big Data Analytics”

Basic part of Block, 3 credits.

Instructor: Sergei V. Smagin, Candidate of Technical Sciences

At the beginning of the course a student should be able to:

- the ability to use the foundations of philosophical knowledge for the formation of worldview position;
- the ability to analyze the main stages and the laws of historical development of society for the formation of citizenship position;
- the ability to use the basics of economic knowledge in various fields;
- the ability of self-organization and self-education;
- the ability to collect, analyze and process of data which are necessary for professional tasks;
- the ability to choose the tools to handle economic data, in accordance with the task, analyze the results of the calculations and justify the findings;
- the ability to find organizational and managerial solutions in professional work and willingness to take responsibility for it;
- the ability to collect and analyze the raw data needed to calculate the economic and socio-economic indicators characterizing the activity of economic entities;
- the ability on the basis of the standard methods and operating procedures and regulatory framework to calculate the economic and socio-economic indicators characterizing the activities of economic entities;
- the ability to carry out the necessary preparation for the economic section of the plan calculations, to justify it, and to present the results of work in accordance with accepted standards;
- the ability on the basis of the description of economic processes and

phenomena create standard theoretical and econometric models to analyze and interpret the results meaningfully;

- the ability to analyze and interpret the data of domestic and foreign statistics on socio-economic processes and phenomena, to identify trends in the socio-economic indicators;

- the ability of using the domestic and foreign databases, to collect the necessary data, analyze it and prepare an overview of the information and analytical report;

- the ability to use to solve the problems of modern communication facilities and information technology;

- the ability to summarize and critically evaluate the results obtained by domestic and foreign researchers, to identify promising areas, constitute a research program;

- the ability to forecast the main socio-economic indicators of the enterprise, industry, region, and the economy as a whole;

- the ability to set goals and choose the methods of research, interpret and provide the results of scientific research in the field of tourism;

- the ability to develop training plans, programs and appropriate methodological support for the teaching of economic disciplines in professional educational organizations, educational institutions of high education, additional training programs;

- the ability to make organizational and administrative decisions.

Learning outcomes:

- ability prepare analytical materials to evaluate activities in areas of economic policy and strategic decision-making at the micro and macro level;

- ability to analyze and use various sources of information for economic calculations;

- ability to use modern methods and tools for the study of socio-economic processes, a comparative analysis of national models of the economy.

Course description: The contents of the teaching materials on the course are based on modern science and educational practice and reflects the author's approach to the subject matter.

The content of the course « Data Science and Big Data Analytics » covers the following range of issues: evolution of the development of international economic relations in the conditions of inter-state economic integration and quickly changeable global problems of world economy; different theories of macroeconomic dynamics; specific features and problem issues of the crisis of modern economic theory; the modern trends of world economy.

The main goal of the course is a theoretical and practical part of training bachelors in the use of information and communication technologies (ICT) in their professional activities and is designed to form students' system understanding of the technologies of multidimensional data analysis, data mining (Data Mining), their application and tools, learn the basic methods of applied science. data analysis, develop skills in the study of various processes on a computer, the practical application of multidimensional analysis methods and data mining to solve I have various scientific and technical tasks.

The main objectives of the course: formation of ideas about the goals, methods of implementation and tools for multidimensional data analysis; study of the scope, methods and tools of Data Mining; formation of practical data analysis skills; obtaining theoretical knowledge and practical skills in solving typical economic problems; learning the basics of building decision support systems; consideration of the structure and functions of data warehouses.

Main course literature:

1 Mnogo tsifr. Analiz bol'shikh dannykh pri pomoshchi Excel / Forman D.; Per. s angl. Sokolovoy A. - M.:Al'pina Pabl., 2016. - 461 s., <http://lib.dvfu.ru:8080/lib/item?id=Znanium:Znanium-551044&theme=FEFU>

2 Adler, YU. P. Statisticheskoye upravleniye protsessami. «Bol'shiye dannyye» [Elektronnyy resurs] : uchebnoye posobiye / YU. P. Adler, Ye. A. Chernykh. — Elektron. tekstovyye dannyye. — M. : Izdatel'skiy Dom MISiS,

2016. — 52 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-64199&theme=FEFU>

3 Nauka o dannykh i statistika v obrazovanii budushchego : nauchnoye izdaniye / V.N. Salin. — Moskva : KnoRus, 2018. — 300 s., <http://lib.dvfu.ru:8080/lib/item?id=BookRu:BookRu-931360&theme=FEFU>

4 Bill, Frenks Revolyutsiya v analitike [Elektronnyy resurs] : kak v epokhu Big Data uluchshit' vash biznes s pomoshch'yu operatsionnoy analitiki / Frenks Bill ; per. I. Yevstigneyeva ; pod red. V. Mylov. — Elektron. tekstovyye dannyye. — M. : Al'pina Pabliher, 2017. — 320 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-58563&theme=FEFU>

5 Voronova, L. I. Big Data. Metody i sredstva analiza [Elektronnyy resurs] : uchebnoye posobiye / L. I. Voronova, V. I. Voronov. — Elektron. tekstovyye dannyye. — M. : Moskovskiy tekhnicheskiy universitet svyazi i informatiki, 2016. — 33 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-61463&theme=FEFU>

Form of final control: pass-fail exam.

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

Учебный курс «Наука о данных и аналитика больших данных» предназначен для студентов направления подготовки 38.04.01 Экономика, магистерская программа «Биоэкономика и продовольственная безопасность».

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

Общая трудоемкость дисциплины «Наука о данных и аналитика больших данных» составляет 3 зачетные единицы, 108 часа. Учебным планом предусмотрены лекции (8 часов), самостоятельная работа студентов (108 часов). Дисциплина реализуется на 1 курсе в 1 семестре.

Дисциплина является продолжением подготовки бакалавров и опирается на содержание таких дисциплин бакалаврской программы: «Математика», «Информатика», «Физика», «Информационное обеспечения товароведения и экспертизы товаров».

Содержание дисциплины состоит из самостоятельных тем и охватывает следующий круг вопросов: технологии многомерного анализа данных, интеллектуальный анализ данных (Data Mining), их применении и инструменты, основные методы прикладного анализа данных, методы многомерного анализа и Data Mining для решения различных научных и технических задач.

Цель – призвана сформировать у студентов системное представление о технологиях многомерного анализа данных, интеллектуального анализа данных (Data Mining), их применении и инструментах, изучить основные методы прикладного анализа данных, развить навыки исследования различных процессов на ЭВМ, практического применения методов многомерного анализа и Data Mining для решения различных научных и технических задач.

Задачи:

- формирование представлений о целях, способах реализации и инструментах многомерного анализа данных;
- изучение сфер применения, методов и средств DataMining;
- формирование практических навыков анализа данных;
- получение теоретических знаний и практических навыков при решении типовых экономических задач;
- изучение основ построения систем поддержки принятия решений;
- рассмотрение структуры и функций хранилищ данных.

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

- способность к самоорганизации и самообразованию;
- способность осуществлять сбор, анализ и обработку данных, необходимых для решения профессиональных задач;
- способность выбирать инструментальные средства для обработки экономических данных в соответствии с поставленной задачей, анализировать результаты расчетов и обосновывать полученные выводы;
- способность находить организационно-управленческие решения в профессиональной деятельности и готовностью нести за них ответственность;
- способность собирать и анализировать исходные данные, необходимые для расчета экономических и социально-экономических показателей, характеризующих деятельность хозяйствующих субъектов;
- способность выполнять необходимые для составления экономических разделов планов расчеты, обосновывать их и представлять результаты работы в соответствии с принятыми в организации стандартами;
- способность на основе описания экономических процессов и явлений строить стандартные теоретические и эконометрические модели, анализировать и содержательно интерпретировать полученные результаты;

- способность анализировать и интерпретировать данные отечественной и зарубежной статистики о социально-экономических процессах и явлениях, выявлять тенденции изменения социально-экономических показателей;

- способность, используя отечественные и зарубежные источники информации, собирать необходимые данные, анализировать их и готовить информационный обзор и/или аналитический отчет.

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

Код и формулировка компетенции	Этапы формирования компетенции	
ПК-10 способность готовить аналитические материалы для оценки мероприятий в области экономической политики и принятия стратегических решений на микро- и макроуровне	Знает	теории и методы, позволяющие составлять аналитические материалы по основным социально-экономическим показателям деятельности предприятий, отраслей, региона и экономики в целом
	Умеет	самостоятельно анализировать управленческие ситуации из жизни организаций, позволяющие составить аналитический материал и прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона, экономики страны
	Владеет	навыками проведения самостоятельных научных исследований с подготовкой аналитических материалов для оценки мероприятий в области экономической политики на основе основных социально-экономических показателей деятельности предприятий, отраслей, регионов, экономики в целом
ПК-11 способность анализировать и использовать различные источники информации для проведения экономических расчетов	Знает	теории и методы, позволяющие проводить комплексный анализ и разработку вариантов управленческих решений на основе обоснованного выбора критериев социально-экономической эффективности
	Умеет	анализировать конкретную экономическую ситуацию с помощью разработки вариантов экономических решений и обосновывать их выбор на основе критериев социально-экономической эффективности
	Владеет	навыками использования инструментов по

		разработке вариантов управленческих решений и обоснования их выбора на основе критериев социально-экономической эффективности
ПК-13 способность использовать современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики	Знает	принципы, современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики
	Умеет	применять различные принципы, современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики
	Владеет	навыками использования инструментов, современных методов и инструментов исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики

Для формирования вышеуказанных компетенций в рамках дисциплины «Наука о данных и аналитика больших данных» применяются следующие методы активного/интерактивного обучения: не предусмотрено.

I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА

Тема 1. Определение больших данных. Технологии хранения больших данных.

Большие данные (big data) в информационных технологиях. совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети, сформировавшихся в конце 2000-х годов, альтернативных традиционным системам управления базами данных и решениям класса Business Intelligence. В данную серию включают средства массово-параллельной обработки неопределённо структурированных данных, прежде всего, решениями категории NoSQL, алгоритмами MapReduce, программными каркасами и библиотеками проекта Hadoop.

Тема 2. Процесс анализа больших данных.

Технологии анализа больших данных. Научные проблемы в области больших данных. Методы и техники анализа, применимые к большим данным: методы класса Data Mining: обучение ассоциативным правилам (англ. association rule learning), классификация (методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным), кластерный анализ, регрессионный анализ; краудсорсинг - категоризация и обогащение данных силами широкого, неопределённого круга лиц, привлечённых на основании публичной оферты, без вступления в трудовые отношения; смешение и интеграция данных (англ. data fusion and integration) - набор техник, позволяющих интегрировать разнородные данные из разнообразных источников для возможности глубинного анализа, в качестве примеров таких техник, составляющих этот класс методов приводятся цифровая обработка сигналов и обработка естественного языка (включая тональный анализ); машинное обучение, включая обучение с учителем и без учителя, а также Ensemble learning (англ.) - использование моделей, построенных на базе статистического анализа или машинного обучения для получения комплексных прогнозов на основе базовых моделей (англ. constituent models, ср. со статистическим ансамблем в статистической механике); искусственные нейронные сети, сетевой анализ, оптимизация, в том числе генетические алгоритмы; распознавание образов; прогнозная аналитика; имитационное моделирование; пространственный анализ (англ. Spatial analysis) - класс методов, использующих топологическую, геометрическую и географическую информацию в данных; статистический анализ, в качестве примеров методов приводятся A/B-тестирование и анализ временных рядов; визуализация аналитических данных - представление информации в виде рисунков, диаграмм, с использованием интерактивных возможностей и анимации как для получения результатов, так и для использования в качестве исходных данных для дальнейшего анализа.

Тема 3. Прогнозирование и предвидение в социально-политических и медиа процессах.

Методы прогнозирования. Понятие прогноза и предвидения. Отличие прогнозирования от предвидения. Закон распределения случайной величины. Статистические оценки параметров. Доверительные области. Теория моментов. Корреляционный анализ. Использование модели множественной линейной регрессии для прогнозирования экономических показателей. Доверительные интервалы для зависимой переменной. Сглаживание временных рядов. Динамические модели с распределенными лагами. Стационарные временные ряды. Тестирование стационарности. Коинтеграция. Анализ временных рядов. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание. Авторегрессионные модели. Модели скользящего среднего. Интегрированные процессы. Идентификация авторегрессионной модели скользящего среднего. Прогнозирование с моделями временных рядов. Доверительные интервалы прогноза. Дисперсионный анализ влияния качественных факторов. Ранговые методы. Факторный анализ. Метод главных факторов. Многомерное шкалирование. Классическая модель многомерного шкалирования. Неметрические методы. Кластерный анализ. Дискриминантный анализ. Многомерный статистический анализ.

Тема 4. Программы статистической обработки информации.

Представление возможностей пакета SPSS Statistics для целей анализа социально-политических процессов. SPSS Statistics (аббревиатура англ. "Statistical Package for the Social Sciences", "статистический пакет для социальных наук") - компьютерная программа для статистической обработки данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в социальных науках. Применение программы для решения прикладных задач прогнозирования: ввод и хранение данных; возможность использования переменных разных типов; частотность признаков, таблицы,

графики, таблицы сопряженности, диаграммы; первичная описательная статистика; маркетинговые и медиа исследования; анализ данных маркетинговых и медиа исследований.

II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА

Практические занятия (8 час.), в том числе с использованием МАО – 0 часов.

Занятие 1. Задачи в области больших данных, решаемые методом регрессионного анализа (4 час.)

Занятие 2. Подготовка данных и больших объемов данных, анализ, визуализация, презентация (4 час.)

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	Тема1-4	ПК-10	знает	Конспект (ПР-7)	Вопросы к зачету (УО-1) 1-10
			умеет	Домашнее задание (ПР-11)	Вопросы к зачету (УО-1) 1-10
			владеет	Практическая работа 1-2 (ПР-6)	Вопросы к зачету (УО-1) 1-10
2	Тема1-4	ПК-11	знает	Конспект (ПР-7)	Вопросы к зачету (УО-1) 1-10
			умеет	Домашнее задание (ПР-11)	Вопросы к зачету (УО-1) 1-10
			владеет	Практическая работа 1-2 (ПР-6)	Вопросы к зачету (УО-1) 1-10
3	Тема1-4	ПК-13	знает	Конспект (ПР-7)	Вопросы к зачету (УО-1) 1-10
			умеет	Домашнее задание (ПР-11)	Вопросы к зачету (УО-1) 1-10
			владеет	Практическая работа 1-2 (ПР-6)	Вопросы к зачету (УО-1) 1-10

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)

1Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - М.:Альпина Пабли., 2016. - 461 с., <http://lib.dvfu.ru:8080/lib/item?id=Znanium:Znanium-551044&theme=FEFU>

2Адлер, Ю. П. Статистическое управление процессами. «Большие данные» [Электронный ресурс] : учебное пособие / Ю. П. Адлер, Е. А. Черных. — Электрон. текстовые данные. — М. : Издательский Дом МИСиС, 2016. — 52 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-64199&theme=FEFU>

3Наука о данных и статистика в образовании будущего : научное издание / В.Н. Салин. — Москва : КноРус, 2018. — 300 с., <http://lib.dvfu.ru:8080/lib/item?id=BookRu:BookRu-931360&theme=FEFU>

4Билл, Фрэнк Революция в аналитике [Электронный ресурс] : как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Фрэнк Билл ; пер. И. Евстигнеева ; под ред. В. Мылов. — Электрон. текстовые данные. — М. : Альпина Паблишер, 2017. — 320 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-58563&theme=FEFU>

5Воронова, Л. И. Big Data. Методы и средства анализа [Электронный ресурс] : учебное пособие / Л. И. Воронова, В. И. Воронов. — Электрон. текстовые данные. — М. : Московский технический университет связи и информатики, 2016. — 33 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-61463&theme=FEFU>

Дополнительная литература

(печатные и электронные издания)

1Computational Intelligence for Big Data Analysis [Electronic resource] / D.P. Acharjya, Satchidananda Dehuri, Sugata Sanyal, 2015, <http://lib.dvfu.ru:8080/lib/item?id=chamo:858734&theme=FEFU>

2Sublinear Algorithms for Big Data Applications [Electronic resource] / Dan Wang, Zhu Han, 2015, <http://lib.dvfu.ru:8080/lib/item?id=chamo:852373&theme=FEFU>

3Thinking Ahead - Essays on Big Data, Digital Revolution, and Participatory Market Society [Electronic resource] / Dirk Helbing, 2015, <http://lib.dvfu.ru:8080/lib/item?id=chamo:851898&theme=FEFU>

4Data Mining for Service [Electronic resource] / Katsutoshi Yada, 2014, <http://lib.dvfu.ru:8080/lib/item?id=chamo:857706&theme=FEFU>

5Эрик, Сигель Просчитать будущее [Электронный ресурс] : кто кликнет, купит, совет или умрет / Сигель Эрик ; пер. И. Евстигнеева. — Электрон. текстовые данные. — М. : Альпина Паблишер, 2016. — 374 с., <http://lib.dvfu.ru:8080/lib/item?id=IPRbooks:IPRbooks-48599&theme=FEFU>

6Big Data Benchmarks, Performance Optimization, and Emerging Hardware [Electronic resource] / Jianfeng Zhan, Rui Han, Chuliang Weng, 2014, <http://lib.dvfu.ru:8080/lib/item?id=chamo:852782&theme=FEFU>

7Engaging Customers Using Big Data [Electronic resource] / Arvind Sathi, 2014, <http://lib.dvfu.ru:8080/lib/item?id=chamo:853242&theme=FEFU>

Перечень информационных технологий и программного обеспечения

Аудитория с мультимедиа проектором, 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

Реализация дисциплины « Наука о данных и аналитика больших объёмов 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 объёмов данных» предполагает рейтинговую систему оценки знаний студентов и предусматривает со стороны преподавателя текущий контроль за посещением студентами лекций, подготовкой и выполнением всех практических работ с обязательным предоставлением отчета о работе, выполнением всех видов самостоятельной работы.

Промежуточной аттестацией по дисциплине « Наука о данных и аналитика больших объёмов данных» является экзамен, который проводится в виде собеседования.

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 объемов данных» при условии выполнения всех видов текущего контроля и самостоятельной работы, предусмотренных учебной программой.

Критерии оценки по дисциплине «Наука о данных и аналитика больших объемов данных» для аттестации на экзамене 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.

Рекомендации по планированию и организации времени, отведенного на изучение дисциплины

Каждый студент самостоятельно определяет режим своей самостоятельной работы. В процессе самостоятельной работы студент

приобретает навыки самоорганизации, самоконтроля, самоуправления и становится активным самостоятельным субъектом учебной деятельности. В процессе самостоятельной работы студент должен:

- освоить минимум содержания, выносимый на самостоятельную работу студентов и предложенный преподавателем по данной дисциплине;
- планировать самостоятельную работу в соответствии с графиком самостоятельной работы, предложенным преподавателем;
- осуществлять самостоятельную работу в организационных формах, предусмотренных учебным планом и рабочей программой учебной дисциплины;
- выполнять самостоятельную работу и отчитываться по ее результатам в соответствии с графиком представления результатов, видами и сроками отчетности по самостоятельной работе студентов;
- использовать для самостоятельной работы методические пособия, учебные пособия, разработки сверх предложенного преподавателем перечня.

Алгоритм изучения дисциплины

Изучение курса должно вестись систематически и сопровождаться составлением подробного конспекта. В конспект рекомендуется включать все виды учебной работы: лекции, самостоятельную проработку рекомендуемой основной и дополнительной литературы, отчеты по практическим работам, доклады, ответы на вопросы для самоконтроля и другие задания, предусмотренные для самостоятельной работы студентов.

Основным промежуточным показателем успешности студента в процессе изучения дисциплины является его готовность к выполнению практических и самостоятельных работ.

Приступая к подготовке к практическим работам, прежде всего, необходимо ознакомиться с планом занятия, изучить соответствующую литературу, нормативную и техническую документацию. По каждому вопросу практической работы студент должен определить и усвоить ключевые понятия и представления. В случае возникновения трудностей студент

должен и может обратиться за консультацией к ведущему преподавателю.

Критерием готовности к практическим работам является умение студента ответить на все контрольные вопросы, рекомендованные преподавателем.

Знания, полученные студентами в процессе изучения дисциплины, должны закрепляться не повторением, а применением материала. Этой цели при изучении дисциплины « Наука о данных и аналитика больших объёмов данных» служат активные формы и методы обучения.

Особое значение для освоения теоретического материала и для приобретения и формирования умений и навыков имеет самостоятельная работа студентов. Самостоятельная работа студентов по данной дисциплине предусматривает изучение рекомендуемой основной и дополнительной литературы, подготовка докладов, подготовка к практическим работам и защита отчета, и промежуточной аттестации – экзамену.

Рекомендации по использованию методов активного обучения

Для повышения эффективности образовательного процесса и формирования активной личности студента важную роль играет такой принцип обучения как познавательная активность студентов. Целью такого обучения является не только освоение знаний, умений, навыков, но и формирование основополагающих качеств личности, что обуславливает необходимость использования методов активного обучения, без которых невозможно формирование специалиста, способного решать профессиональные задачи в современных рыночных условиях.

Для развития профессиональных навыков и личности студента в качестве методов активного обучения целесообразно использовать методы ситуационного обучения, представляющие собой описание деловой ситуации, которая реально возникала или возникает в процессе деятельности.

Рекомендации по работе с литературой

Самостоятельная работа с учебниками и книгами, самостоятельное теоретическое исследование проблем, обозначенных преподавателем на

лекциях – важнейшее условие формирования студентом у себя научного способа познания. Сэкономить студенту время и силы помогут рациональные навыки работы с учебной книгой. Для подбора литературы в библиотеке используются алфавитный и систематический каталоги. Изучая материал по учебной книге (учебнику, учебному пособию, монографии, хрестоматии и др.), следует переходить к следующему вопросу только после полного уяснения предыдущего, фиксируя выводы и вычисления, в том числе те, которые в учебнике опущены или на лекции даны для самостоятельного вывода. Особое внимание студент должен обратить на определение основных понятий курса. Надо подробно разбирать примеры, которые поясняют определения, и приводить аналогичные примеры самостоятельно. Полезно составлять опорные конспекты. При изучении материала по учебной книге полезно либо в тетради на специально отведенных полях, либо в документе, созданном на ноутбуке, планшете и др. информационном устройстве, дополнять конспект лекций. Там же следует отмечать вопросы, выделенные студентом для консультации с преподавателем. Выводы, полученные в результате изучения учебной литературы, рекомендуется в конспекте выделять, чтобы при перечитывании материала они лучше запоминались. Студентам рекомендуется составлять лист опорных сигналов, содержащий важнейшие и наиболее часто употребляемые формулы и понятия, основные положения лекции, что может служить постоянным справочником по предмету. Основной смысл подготовки опорных сигналов – это систематизация и оптимизация знаний по данному предмету. Если студент самостоятельно подготовил опорные сигналы, то экзамены он будет сдавать более уверенно, т.к. у него уже сформирована общая ориентировка в сложном материале. Использование сигналов позволяет отвечающему лучше демонстрировать ориентировку в знаниях, что намного важнее знания «тут же забытого» после сдачи экзамена. Следует внимательно и осознанно читать учебную литературу. Различают два вида чтения: первичное, как внимательное, неторопливое чтение, при котором можно остановиться на

трудных местах, и вторичное, после которого у студента не должно остаться ни одного непонятого слова. Содержание учебного или научного материала не всегда может быть понятно после первичного чтения. Задача вторичного чтения полное усвоение смысла целого (по счету это чтение может быть и не вторым, а третьим или четвертым и т.д.).

Правила самостоятельной работы студента с учебной литературой:

1. Составьте перечень книг, с которыми следует познакомиться; не старайтесь запомнить все, что вам в ближайшее время не понадобится, запомните только, где можно отыскать необходимый материал.

2. Перечень должен быть систематизированным (необходимо для семинаров, экзаменов, пригодится для написания курсовых и дипломных работ).

3. Обязательно выписывайте все выходные данные по каждой книге (при написании курсовых и дипломных работ это позволит сэкономить время).

4. Разберитесь для себя, какие книги (или какие главы книг) следует прочитать более внимательно, а какие – просмотреть.

5. При составлении перечня литературы посоветуйтесь с преподавателями и научными руководителями, эрудированными однокурсниками, которые помогут Вам лучше сориентироваться, на что стоит обратить больше внимания.

6. Все прочитанные книги, учебники и статьи конспектируйте – выписывайте кратко основные идеи автора, приводите наиболее яркие и цитаты (с указанием страниц источника).

7. На собственных книгах допускается делать на полях краткие пометки или же в конце книги, на пустых страницах просто сделать свой «предметный указатель», где отмечаются наиболее интересные для Вас мысли и обязательно указываются страницы в тексте - это позволяет экономить время и быстро находить «избранные» места в разных книгах.

8. При малом опыте работы с научной литературой следует выработать в себе способность «воспринимать» сложные тексты посредством приема

«медленного чтения», когда понятно каждое прочитанное слово (если слово незнакомое, то с помощью словаря обязательно его узнать).

9. Эффективный способ оптимизации знакомства с научной литературой – увлечение одной идеей и просматривание всех книг с точки зрения данной идеи. В этом случае Вы будете искать аргументы «за» или «против» интересующей Вас идеи, и одновременно будете мысленно общаться с авторами этих книг по поводу своих идей и размышлений.

Чтение учебной и научной литературы является частью познавательной деятельности студента, цель которой – извлечение из текста необходимой информации. Насколько осознанна Вами собственная внутренняя установка: поиск нужных сведений, усвоение информации полностью или частично, анализ материала и т.п., во многом зависит эффективность осуществляемого Вами действия.

Рекомендации по подготовке к экзамену (зачету)

Подготовка к экзамену способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. В процессе подготовки к экзамену, ликвидируются имеющиеся пробелы в знаниях, углубляются, систематизируются и упорядочиваются знания. На экзамене демонстрируются знания, приобретенные в процессе обучения по конкретной учебной дисциплине.

Перечень экзаменационных вопросов в билете включает два вопроса. Для хорошего ответа на такой вопрос нужно не просто вспомнить материал соответствующего раздела курса, но и мобилизовать относящиеся к вопросу знания из других разделов.

Особенности организации образовательного процесса по образовательной программе для инвалидов и лиц с ограниченными возможностями здоровья

В ДВФУ реализуется организационная модель инклюзивного образования - обеспечение равного доступа к образованию для всех

обучающихся с учетом различных особых образовательных потребностей и индивидуальных возможностей студентов. Модель позволяет лицам, имеющим ограниченные возможности здоровья (ОВЗ), использовать образование как наиболее эффективный механизм развития личности, повышения своего социального статуса.

Содержание высшего образования по образовательным программам и условия организации обучения лиц с ОВЗ определяются адаптированной образовательной программой, а для инвалидов также в соответствии с индивидуальной программой реабилитации, которая разрабатывается Федеральным учреждением медико-социальной экспертизы. Адаптированная образовательная программа разрабатывается при наличии заявления со стороны обучающегося (родителей, законных представителей) и медицинских показаний.

Обучение по образовательным программам инвалидов и обучающихся с ОВЗ осуществляется организацией с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья. Выбор методов обучения в каждом отдельном случае обуславливается целями обучения, содержанием обучения, уровнем профессиональной подготовки педагогов, методического и материально-технического обеспечения, наличием времени на подготовку, с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья обучающихся.

Инвалиды и лица с ограниченными возможностями слуха и речи, с ограниченными возможностями зрения и ограниченными возможностями опорно-двигательной системы могут получить образование в Университете по данной основной образовательной программе по очной форме обучения с использованием элементов дистанционных образовательных технологий.

Университет обеспечивает обучающимся лицам с ОВЗ и инвалидам возможность освоения специализированных адаптационных дисциплин, включаемых в вариативную часть ОПОП. Преподаватели, курсы которых

требуют выполнения определенных специфических действий и представляющих собой проблему или действие, невыполнимое для обучающихся, испытывающих трудности с передвижением или речью, обязаны учитывать эти особенности и предлагать инвалидам и лицам с ОВЗ альтернативные методы закрепления изучаемого материала. Своевременное информирование преподавателей об инвалидах и лицах с ОВЗ в конкретной группе осуществляется ответственным лицом, установленным приказом директора школы.

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

При необходимости для инвалидов и лиц с ОВЗ могут разрабатываться индивидуальные учебные планы и индивидуальные графики обучения. Срок получения высшего образования при обучении по индивидуальному учебному плану для инвалидов и лиц с ОВЗ при желании может быть увеличен, но не более чем на год.

Для осуществления мероприятий текущего контроля успеваемости, промежуточной и итоговой аттестации инвалидов и лиц с ОВЗ применяются фонды оценочных средств, адаптированные для таких обучающихся и позволяющие оценить достижение ими результатов обучения и уровень сформированности всех компетенций, заявленных в образовательной программе. Форма проведения промежуточной и государственной итоговой аттестации для студентов-инвалидов и лиц с ОВЗ устанавливается с учетом индивидуальных психофизических особенностей (устно, письменно на бумажном носителе, письменно на компьютере, в форме тестирования и т.п.).

VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Для осуществления образовательного процесса по дисциплине «Наука о данных и аналитика больших объёмов данных» необходимы:

Наука о данных и аналитика больших объёмов данных	690922, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, корпус G, каб. G423, учебная аудитория для проведения занятий лекционного типа; учебная аудитория для текущего контроля и промежуточной аттестации	90 посадочных мест, автоматизированное рабочее место преподавателя, переносная магнитно-маркерная доска, Wi-Fi Ноутбук Acer ExtensaE2511-30VO Экран с электроприводом 236*147 см Trim Screen Line; Проектор DLP, 3000 ANSI Lm, WXGA 1280x800, 2000:1 EW330U Mitsubishi; Подсистема специализированных креплений оборудования CORSA-2007 Tuarex; Подсистема видеокмутации; Подсистема аудиокмутации и звукоусиления; акустическая система для потолочного монтажа SI 3CT LP Extron; цифровой аудиопроцессор DMP 44 LC Extron.
---	---	---



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ 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 объёмов данных»

**Направление подготовки 38.04.01 Экономика
магистерская программа «Биоэкономика и продовольственная
безопасность »**

Форма подготовки очная

**Владивосток
2018**

План-график выполнения самостоятельной работы по дисциплине

№ п/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
	Еженедельно	Конспект по темам курса	30	Фронтальный просмотр
	еженедельно	Домашнее задание	30	Проверка ДЗ
	12-16 неделя обучения	Подготовка к практическим работам	10	Устный опрос
	18 неделя, зачетная неделя	Подготовка к итоговому контролю	30	зачет
	Итого		100	

В ходе выполнения самостоятельной работы студент должен:

- освоить теоретический материал по изучаемой дисциплине;
- закрепить знания теоретического материала путем выполнения заданий на практических занятиях, выполнения всех видов;
- применить полученные знания и практические навыки для анализа ситуации, выработки правильного решения и формирования собственной позиции при выполнении научно-исследовательской работы и выпускной квалификационной работы.

Написание конспектов по разделам дисциплины

Учебным планом не предусмотрено проведение лекций, поэтому теоретическая часть курса выносится на самостоятельное изучение. По каждой теме теоретической части необходимо написать конспект.

Требования к представлению и оформлению конспекта:

При написании конспекта необходимо придерживаться следующих требований:

- полнота изложения материала;
- логика изложения материала;
- использование соответствующей терминологии и стиля изложения;
- наличие списка использованных источников (не менее 5);
- объем не менее 10 страниц рукописного текста, страница формата А5.

Конспект засчитается при соблюдении вышеперечисленных условий.

Конспект 1. Определение больших данных. Технологии хранения больших данных.

Большие данные (big data) в информационных технологиях. совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети, сформировавшихся в конце 2000-х годов, альтернативных традиционным системам управления базами данных и решениям класса Business Intelligence. В данную серию включают средства массово-параллельной обработки неопределённо структурированных данных, прежде всего, решениями категории NoSQL, алгоритмами MapReduce, программными каркасами и библиотеками проекта Hadoop.

Конспект 2. Процесс анализа больших данных.

Технологии анализа больших данных. Научные проблемы в области больших данных. Методы и техники анализа, применимые к большим данным: методы класса Data Mining: обучение ассоциативным правилам (англ. association rule learning), классификация (методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным), кластерный анализ, регрессионный анализ; краудсорсинг - категоризация и обогащение данных силами широкого, неопределённого

круга лиц, привлечённых на основании публичной оферты, без вступления в трудовые отношения; смешение и интеграция данных (англ. data fusion and integration) - набор техник, позволяющих интегрировать разнородные данные из разнообразных источников для возможности глубинного анализа, в качестве примеров таких техник, составляющих этот класс методов приводятся цифровая обработка сигналов и обработка естественного языка (включая тональный анализ); машинное обучение, включая обучение с учителем и без учителя, а также Ensemble learning (англ.) - использование моделей, построенных на базе статистического анализа или машинного обучения для получения комплексных прогнозов на основе базовых моделей (англ. constituent models, ср. со статистическим ансамблем в статистической механике); искусственные нейронные сети, сетевой анализ, оптимизация, в том числе генетические алгоритмы; распознавание образов; прогнозная аналитика; имитационное моделирование; пространственный анализ (англ. Spatial analysis) - класс методов, использующих топологическую, геометрическую и географическую информацию в данных; статистический анализ, в качестве примеров методов приводятся A/B-тестирование и анализ временных рядов; визуализация аналитических данных - представление информации в виде рисунков, диаграмм, с использованием интерактивных возможностей и анимации как для получения результатов, так и для использования в качестве исходных данных для дальнейшего анализа.

Конспект 3. Прогнозирование и предвидение в социально-политических и медиа процессах.

Методы прогнозирования. Понятие прогноза и предвидения. Отличие прогнозирования от предвидения. Закон распределения случайной величины. Статистические оценки параметров. Доверительные области. Теория моментов. Корреляционный анализ. Использование модели множественной линейной регрессии для прогнозирования экономических показателей. Доверительные интервалы для зависимой переменной. Сглаживание временных рядов. Динамические модели с распределёнными лагами.

Стационарные временные ряды. Тестирование стационарности. Коинтеграция. Анализ временных рядов. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание. Авторегрессионные модели. Модели скользящего среднего. Интегрированные процессы. Идентификация авторегрессионной модели скользящего среднего. Прогнозирование с моделями временных рядов. Доверительные интервалы прогноза. Дисперсионный анализ влияния качественных факторов. Ранговые методы. Факторный анализ. Метод главных факторов. Многомерное шкалирование. Классическая модель многомерного шкалирования. Немеетрические методы. Кластерный анализ. Дискриминантный анализ. Многомерный статистический анализ.

Конспект 4. Программы статистической обработки информации.

Представление возможностей пакета SPSSStatistics для целей анализа социально-политических процессов. SPSS Statistics (аббревиатура англ. "Statistical Package for the Social Sciences", "статистический пакет для социальных наук") - компьютерная программа для статистической обработки данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в социальных науках. Применение программы для решения прикладных задач прогнозирования: ввод и хранение данных; возможность использования переменных разных типов; частотность признаков, таблицы, графики, таблицы сопряженности, диаграммы; первичная описательная статистика; маркетинговые и медиа исследования; анализ данных маркетинговых и медиа исследований.

Критерии оценки конспекта

– 100-86 баллов выставляется студенту, если студент выразил свое 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/или 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.

Домашнее задание (письменно)

При ответе на вопросы необходимо придерживаться следующих требований:

- полнота изложения материала;
- логика изложения материала;
- использование соответствующей терминологии и стиля изложения;
- наличие списка использованных источников (не менее 5);
- объем не менее 5 страниц рукописного текста, страница формата А5.

Темы для домашнего задания

- 1 Линейная регрессия. Метод наименьших квадратов.
- 2 Примеры использования корреляции и регрессионного анализа в области больших данных.
- 3 Логистическая регрессия.
- 4 Наивный классификатор Байеса.
- 5 Алгоритм априори.
- 6 Алгоритм k-means.
- 7 Матрица парных расстояний.
- 8 Виды метрик.
- 9 Анализ стандартных наборов данных с помощью Weka.
- 10 Анализ стандартных наборов данных с помощью Orange.
- 11 Визуализация стандартных наборов данных с помощью Tableau.
- 12 Вычисление индекса TF-IDF.
- 13 Реализация Nadoop.
- 14 Регуляризация для метода наименьших квадратов.
- 15 Машина опорных вектор

Критерии домашнего задания:

100-86 баллов выставляется студенту, если его ответ показывает глубокое и систематическое знание структуры конкретных вопросов по данному разделу дисциплины, а также содержания основной и рекомендованной дополнительной литературы. Студент демонстрирует отчетливое и свободное владение концептуально-понятийным аппаратом, научным языком и терминологией соответствующей научной области.

Полное, логически корректное, аргументированное и убедительное изложение ответа.

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

75-61 баллов - фрагментарные, поверхностные знания раздела дисциплины; затруднения с использованием научно-понятийного аппарата и терминологии учебной дисциплины; неполное знакомство с рекомендованной литературой; частичные затруднения с выполнением предусмотренных программой заданий; стремление логически определенно и последовательно изложить ответ.

60-50 баллов - незнание, либо отрывочное представление о данной проблеме в рамках изучения раздела дисциплины; неумение использовать понятийный аппарат; отсутствие логической связи в ответе.

Подготовка к практическим работам

Занятие 1. Задачи в области больших данных, решаемые методом регрессионного анализа (4 час.)

Вопросы для самоподготовки:

1. В чем принципиальное отличие концепции Big Data от традиционного подхода?
 2. Понятие явной (выраженной) и скрытой (структурной) информации.
 3. Определение контент-анализа.
 4. Каковы основные понятия контент-анализа?
 5. Какие существуют виды контент-анализа?
 6. Какие существуют этапы контент-анализа?
 7. Каковы основные признаки, характеризующие «Большие данные»?
1. Модели развертывания облачных хранилищ.
 2. Модели обслуживания облачных хранилищ.

3. Постановка и описание проблемы «последней мили».
4. Безопасность, производительность и надежность при работе с облачными данными.
5. Экономическая составляющая облачных подходов.
6. Способы машинного обучения.
7. Основные фазы обработки «больших данных».

Занятие 2. Подготовка данных и больших объемов данных, анализ, визуализация, презентация (4 час.)

- 1 Недоопределенные задачи анализа, необходимость визуализации для анализа
- 2 Классификация методов визуализации.
- 3 Преимущества визуализации больших массивов данных в виде сетей. Примеры визуализации

Критерии оценки теоретической подготовки к практическим работам

– 100-86 баллов выставляется студенту, если студент дал правильные ответы на все поставленные вопросы при собеседовании;

– 85-76 баллов – не более 2 неправильных или отсутствующих ответов на поставленные вопросы в собеседовании;

– 75-61 балл – 3-4 неправильных или отсутствующих ответов на поставленные вопросы в собеседовании;

– 60-50 баллов – более 4 неправильных или отсутствующих ответов на поставленные вопросы в собеседовании.

Подготовка к зачету

Подготовка к экзамену (зачету) способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. В процессе подготовки к экзамену, ликвидируются имеющиеся пробелы в знаниях, углубляются, систематизируются и упорядочиваются знания. На экзамене (зачете)

демонстрируются знания, приобретенные в процессе обучения по конкретной учебной дисциплине.

Перечень экзаменационных вопросов в билете включает два вопроса. Для хорошего ответа на такой вопрос нужно не просто вспомнить материал соответствующего раздела курса, но и мобилизовать относящиеся к вопросу знания из других разделов.

Вопросы к зачету

1. Понятие Большие данные. Роль цифровой информации в 21 веке.
2. Виды массивов данных.
3. Базовые принципы обработки больших данных.
4. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop, R.
5. Технологии Business Intelligence и реляционные системы управления базами данных.
6. Прогнозирование и предвидение: общее и особенное.
7. Виды прогнозов
8. Общие методы анализа социально-политических и медиа процессов.
9. Специальные методы анализа социально-политических и медиа процессов.
10. Предварительный анализ данных.
11. Проверка гипотез о законе распределения случайной величины.
12. Статистические оценки параметров. Доверительные области.
13. Теория моментов.
14. Корреляционный анализ.
15. Использование модели множественной линейной регрессии для прогнозирования экономических показателей. Доверительные интервалы для зависимой переменной.
16. Сглаживание временных рядов. Динамические модели с распределенными лагами.
17. Стационарные временные ряды. Тестирование стационарности.

18. Коинтеграция. Анализ временных рядов.
19. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание.
20. Авторегрессионные модели. Модели скользящего среднего.
21. Интегрированные процессы. Идентификация авторегрессионной модели скользящего среднего.
22. Прогнозирование с моделями временных рядов. Доверительные интервалы прогноза.
23. Предсказание и прогнозирование социально-экономических прогнозов.
24. Дисперсионный анализ влияния качественных факторов. Ранговые методы.
25. Факторный анализ. Метод главных факторов.
26. Многомерное шкалирование. Классическая модель многомерного шкалирования.
27. Неметрические методы. Кластерный анализ. Дискриминантный анализ.
28. Многомерный статистический анализ.
29. Статистический анализ в пакете SPSS Statistics.
30. Основные возможности пакета SPSS Statistics

Критерии выставления оценки студенту на зачете

по дисциплине «Наука о данных и аналитика больших объёмов данных»

Баллы (рейтинговой оценки)	Оценка зачета (стандартная)	Требования к сформированным компетенциям
86-100	«зачтено»	Студент 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 и приемами выполнения практических задач.
85-76	«зачтено»	Студент 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.
75-61	«зачтено»	Студент 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.
60-0	«не зачтено»	Оценка «не зачтено» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «не зачтено» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ 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 объемов данных»
Направление подготовки 38.04.01 Экономика
магистерская программа «Биоэкономика и продовольственная
безопасность»
Форма подготовки очная

Владивосток
2018

Паспорт фонда оценочных средств

Код и формулировка компетенции	Этапы формирования компетенции	
ПК-10 способность готовить аналитические материалы для оценки мероприятий в области экономической политики и принятия стратегических решений на микро- и макроуровне	Знает	теории и методы, позволяющие составлять аналитические материалы по основным социально-экономическим показателям деятельности предприятий, отраслей, региона и экономики в целом
	Умеет	самостоятельно анализировать управленческие ситуации из жизни организаций, позволяющие составить аналитический материал и прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона, экономики страны
	Владеет	навыками проведения самостоятельных научных исследований с подготовкой аналитических материалов для оценки мероприятий в области экономической политики на основе основных социально-экономических показателей деятельности предприятий, отраслей, регионов, экономики в целом
ПК-11 способность анализировать и использовать различные источники информации для проведения экономических расчетов	Знает	теории и методы, позволяющие проводить комплексный анализ и разработку вариантов управленческих решений на основе обоснованного выбора критериев социально-экономической эффективности
	Умеет	анализировать конкретную экономическую ситуацию с помощью разработки вариантов экономических решений и обосновывать их выбор на основе критериев социально-экономической эффективности
	Владеет	навыками использования инструментов по разработке вариантов управленческих решений и обоснования их выбора на основе критериев социально-экономической эффективности
ПК-13 способность использовать современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики	Знает	принципы, современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики
	Умеет	применять различные принципы, современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики
	Владеет	навыками использования инструментов, современных методов и инструментов исследования социально-экономических процессов, сравнительного анализа национальных

		моделей экономики
--	--	-------------------

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства	
				текущий контроль	промежуточная аттестация
1	Тема1-4	ПК-10	знает	Конспект (ПР-7)	Вопросы к зачету (УО-1) 1-10
			умеет	Домашнее задание (ПР-11)	Вопросы к зачету (УО-1) 1-10
			владеет	Практическая работа 1-2 (ПР-6)	Вопросы к зачету (УО-1) 1-10
2	Тема1-4	ПК-11	знает	Конспект (ПР-7)	Вопросы к зачету (УО-1) 1-10
			умеет	Домашнее задание (ПР-11)	Вопросы к зачету (УО-1) 1-10
			владеет	Практическая работа 1-2 (ПР-6)	Вопросы к зачету (УО-1) 1-10
3	Тема1-4	ПК-13	знает	Конспект (ПР-7)	Вопросы к зачету (УО-1) 1-10
			умеет	Домашнее задание (ПР-11)	Вопросы к зачету (УО-1) 1-10
			владеет	Практическая работа 1-2 (ПР-6)	Вопросы к зачету (УО-1) 1-10

Шкала оценивания уровня сформированности компетенций

Код и формулировка компетенции	Этапы формирования компетенции		Критерии	Показатели
ПК-10 способность готовить аналитические материалы для оценки мероприятий в области экономической политики и принятия стратегических решений на микро- и макроуровне	знает (пороговый уровень)	основные теории и методы отраслей знаний и особенности видов профессиональной деятельности, методiku организации и проведения научной работы и решения практических задач;	Знание узловых проблем программы и основного содержания лекционного курса	Способность дать определения и сформировать узловые проблемы и основного содержания лекционного материала
	умеет (продвинутый)	самостоятельно анализировать управленческие ситуации из жизни организаций, позволяющие наглядно показать ключевые концепции; анализировать конкретные экономические ситуации, процессы, протекающие в условиях различных социально-экономических систем;		Способность перечислить и раскрыть суть возникающих в процессе научного исследования общих мировоззренческих проблем с помощью научных парадигм; способность осуществлять самостоятельные выводы по теме исследования
	владеет (высокий)	навыками самостоятельной аналитической деятельности; навыками проведения самостоятельных научных исследований и грамотного оформления их результатов;	Владение на высоком уровне понятийным аппаратом, владение навыками научного анализа и методологией научного подхода в научно-исследовательской и практической деятельности	Способность точно применять понятийный аппарат предметной области исследования в устных ответах на вопросы и в

				письменных работах; способность точно применять научный подход в научно- исследовательской деятельности
ПК-11 способность анализировать и использовать различные источники информации для проведения экономических расчетов	знает (пороговый уровень)	методы и подходы для разработки заданий с использованием специальной терминологией на иностранном языке; методы и подходы по подготовке публикации, проведению презентации, ведению дискуссии и защите представленной работы на иностранном языке;	Знание основных теорий и методов отраслей знаний, методику организации и проведения научной работы; знание важнейших работ из списка рекомендованной литературы.	Способность применять основные теории и методы отраслей знаний, способность применять в научно- исследовательской деятельности изученные работы из рекомендованной литературы.
	умеет (продвинутый)	систематизировать и обобщать информацию по актуальным вопросам международных экономических отношений; применять подходы для разработки заданий с использованием специальной терминологией на иностранном языке; применять подходы по подготовке публикации, проведению презентации, ведению дискуссии и защите представленной работы на иностранном язык;	Умение анализировать управленческие ситуации в организациях; умение анализировать конкретные экономические ситуации, процессы в экономических системах.	Способность проводить анализ управленческих ситуаций; способность на должном уровне проводить анализ конкретных экономических ситуаций и процессов в изучаемых экономических системах
	владеет (высокий)	навыками самостоятельной разработки проектных решений в	Владение навыками самостоятельной аналитической	Способность точно применять на

		<p>конкретной ситуации в условиях современных международных экономических отношений;</p> <p>навыками разработки заданий с использованием специальной терминологией на иностранном языке;</p> <p>навыками подготовки публикаций, проведению презентации, ведению дискуссии и защите представленной работы на иностранном языке;</p>	<p>деятельности; владение навыками проведения самостоятельных научных исследований.</p>	<p>практике навыки самостоятельной аналитической деятельности;</p> <p>способность проводить самостоятельные научные исследования по конкретной экономической тематике.</p>
<p>ПК-13 способность использовать современные методы и инструменты исследования социально-экономических процессов, сравнительного анализа национальных моделей экономики</p>	<p>знает (пороговый уровень)</p>	<p>принципы и методы, необходимые для преподавания экономических дисциплин в различных профессиональных образовательных учреждениях;</p>	<p>Знание основной теоретической литературы в области теории экономики и управления; знание принципов и методов макроэкономического анализа.</p>	<p>Способность дать определения основных понятий предметной области исследования;</p> <p>способность перечислить и раскрыть суть методов научного исследования, которые изучил и освоил магистрант;</p> <p>способность самостоятельно сформулировать объект предмет и научного исследования;</p> <p>способность обосновать</p>

				актуальность выполняемого задания или исследования.
	умеет (продвинутый)	применять различные принципы, современные методы и методики при преподавании экономических дисциплин в профессиональных образовательных организациях;	Умение пользоваться концептуально-понятийным аппаратом в процессе анализа основных проблем в рамках тем дисциплины.	Способность работать с данными, каталогов для исследования; способность найти труды учёных и обосновать объективность применения изученных результатов научных исследований в качестве доказательства или опровержения исследовательских аргументов; способность изучить научные определения относительно объекта и предмета исследования; способность применять методы научных исследований для нестандартного решения поставленных задач.
	владеет (высокий)	навыками использования	Владение современным экономическим языком на	Способность бегло и точно применять

		инструментов, современных методик при преподавании экономических дисциплин в профессиональных образовательных организациях;	хорошем уровне.	терминологический аппарат предметной области исследования в устных ответах на вопросы и в письменных работах, способность сформулировать задание по научному исследованию; способность проводить самостоятельные исследования и представлять их результаты на обсуждение на круглых столах, семинарах, научных конференциях.
--	--	---	-----------------	--

Зачетно-экзаменационные материалы

Вопросы к зачету

1. Понятие Большие данные. Роль цифровой информации в 21 веке.
2. Виды массивов данных.
3. Базовые принципы обработки больших данных.
4. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop, R.
5. Технологии Business Intelligence и реляционные системы управления базами данных.
6. Прогнозирование и предвидение: общее и особенное.
7. Виды прогнозов
8. Общие методы анализа социально-политических и медиа процессов.
9. Специальные методы анализа социально-политических и медиа процессов.
10. Предварительный анализ данных.
11. Проверка гипотез о законе распределения случайной величины.
12. Статистические оценки параметров. Доверительные области.
13. Теория моментов.
14. Корреляционный анализ.
15. Использование модели множественной линейной регрессии для прогнозирования экономических показателей. Доверительные интервалы для зависимой переменной.
16. Сглаживание временных рядов. Динамические модели с распределенными лагами.
17. Стационарные временные ряды. Тестирование стационарности.
18. Коинтеграция. Анализ временных рядов.
19. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание.
20. Авторегрессионные модели. Модели скользящего среднего.

21. Интегрированные процессы. Идентификация авторегрессионной модели скользящего среднего.
22. Прогнозирование с моделями временных рядов. Доверительные интервалы прогноза.
23. Предсказание и прогнозирование социально-экономических прогнозов.
24. Дисперсионный анализ влияния качественных факторов. Ранговые методы.
25. Факторный анализ. Метод главных факторов.
26. Многомерное шкалирование. Классическая модель многомерного шкалирования.
27. Неметрические методы. Кластерный анализ. Дискриминантный анализ.
28. Многомерный статистический анализ.
29. Статистический анализ в пакете SPSS Statistics.
30. Основные возможности пакета SPSS Statistics

**Критерии выставления оценки студенту на зачете
по дисциплине «Наука о данных и аналитика больших объёмов данных»**

Баллы (рейтинговой оценки)	Оценка зачета (стандартная)	Требования к сформированным компетенциям
86-100	«зачтено»	Студент 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 и приемами выполнения практических задач.

85-76	«зачтено»	Студент 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.
75-61	«зачтено»	Студент 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.
60-0	«не зачтено»	Оценка «не зачтено» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «не зачтено» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

Оценочные средства для текущей аттестации

Конспект 1. Определение больших данных. Технологии хранения больших данных.

Большие данные (big data) в информационных технологиях. совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети, сформировавшихся в конце 2000-х годов, альтернативных традиционным системам управления базами данных и решениям класса Business Intelligence. В данную серию включают средства массово-параллельной обработки неопределённо структурированных данных, прежде всего, решениями категории NoSQL, алгоритмами MapReduce, программными каркасами и библиотеками проекта Hadoop.

Конспект 2. Процесс анализа больших данных.

Технологии анализа больших данных. Научные проблемы в области больших данных. Методы и техники анализа, применимые к большим данным: методы класса Data Mining: обучение ассоциативным правилам (англ. association rule learning), классификация (методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным), кластерный анализ, регрессионный анализ; краудсорсинг - категоризация и обогащение данных силами широкого, неопределённого круга лиц, привлечённых на основании публичной оферты, без вступления в трудовые отношения; смешение и интеграция данных (англ. data fusion and integration) - набор техник, позволяющих интегрировать разнородные данные из разнообразных источников для возможности глубинного анализа, в качестве примеров таких техник, составляющих этот класс методов приводятся цифровая обработка сигналов и обработка естественного языка (включая тональный анализ); машинное обучение, включая обучение с учителем и без учителя, а также Ensemble learning (англ.) - использование моделей, построенных на базе статистического анализа или машинного обучения для получения комплексных прогнозов на основе базовых моделей (англ. constituent models, ср. со статистическим ансамблем в статистической механике); искусственные нейронные сети, сетевой анализ, оптимизация, в том числе генетические алгоритмы; распознавание образов; прогнозная аналитика; имитационное моделирование; пространственный анализ (англ. Spatial analysis) - класс методов, использующих топологическую, геометрическую и географическую информацию в данных; статистический анализ, в качестве примеров методов приводятся A/B-тестирование и анализ временных рядов; визуализация аналитических данных - представление информации в виде рисунков, диаграмм, с использованием интерактивных возможностей и анимации как для получения результатов, так и для использования в качестве исходных данных для дальнейшего анализа.

Конспект 3. Прогнозирование и предвидение в социально-политических и медиа процессах.

Методы прогнозирования. Понятие прогноза и предвидения. Отличие прогнозирования от предвидения. Закон распределения случайной величины. Статистические оценки параметров. Доверительные области. Теория моментов. Корреляционный анализ. Использование модели множественной линейной регрессии для прогнозирования экономических показателей. Доверительные интервалы для зависимой переменной. Сглаживание временных рядов. Динамические модели с распределенными лагами. Стационарные временные ряды. Тестирование стационарности. Коинтеграция. Анализ временных рядов. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание. Авторегрессионные модели. Модели скользящего среднего. Интегрированные процессы. Идентификация авторегрессионной модели скользящего среднего. Прогнозирование с моделями временных рядов. Доверительные интервалы прогноза. Дисперсионный анализ влияния качественных факторов. Ранговые методы. Факторный анализ. Метод главных факторов. Многомерное шкалирование. Классическая модель многомерного шкалирования. Неметрические методы. Кластерный анализ. Дискриминантный анализ. Многомерный статистический анализ.

Конспект 4. Программы статистической обработки информации.

Представление возможностей пакета SPSSStatistics для целей анализа социально-политических процессов. SPSS Statistics (аббревиатура англ. "Statistical Package for the Social Sciences", "статистический пакет для социальных наук") - компьютерная программа для статистической обработки данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в социальных науках. Применение программы для решения прикладных задач прогнозирования: ввод и хранение данных; возможность использования переменных разных типов; частотность признаков, таблицы, графики, таблицы сопряженности, диаграммы; первичная описательная

статистика; маркетинговые и медиа исследования; анализ данных маркетинговых и медиа исследований.

Критерии оценки конспекта

– 100-86 баллов выставляется студенту, если студент выразил свое 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/или 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.

Темы для домашнего задания

- 1 Линейная регрессия. Метод наименьших квадратов.
- 2 Примеры использования корреляции и регрессионного анализа в области больших данных.
- 3 Логистическая регрессия.
- 4 Наивный классификатор Байеса.
- 5 Алгоритм априори.
- 6 Алгоритм k-means.
- 7 Матрица парных расстояний.
- 8 Виды метрик.
- 9 Анализ стандартных наборов данных с помощью Weka.
- 10 Анализ стандартных наборов данных с помощью Orange.
- 11 Визуализация стандартных наборов данных с помощью Tableau.
- 12 Вычисление индекса TF-IDF.
- 13 Реализация Nadoop.
- 14 Регуляризация для метода наименьших квадратов.
- 15 Машина опорных вектор

Критерии домашнего задания:

100-86 баллов выставляется студенту, если его ответ показывает глубокое и систематическое знание структуры конкретных вопросов по данному разделу дисциплины, а также содержания основной и рекомендованной дополнительной литературы. Студент демонстрирует отчетливое и свободное владение концептуально-понятийным аппаратом, научным языком и терминологией соответствующей научной области. Полное, логически корректное, аргументированное и убедительное изложение ответа.

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

75-61 баллов - фрагментарные, поверхностные знания раздела дисциплины; 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 изучения раздела дисциплины; неумение использовать понятийный аппарат; отсутствие логической связи в ответе.

Подготовка к практическим работам

Занятие 1. Задачи в области больших данных, решаемые методом регрессионного анализа (4 час.)

Вопросы для самоподготовки:

1. В чем принципиальное отличие концепции Big Data от традиционного подхода?
 2. Понятие явной (выраженной) и скрытой (структурной) информации.
 3. Определение контент-анализа.
 4. Каковы основные понятия контент-анализа?
 5. Какие существуют виды контент-анализа?
 6. Какие существуют этапы контент-анализа?
 7. Каковы основные признаки, характеризующие «Большие данные»?
1. Модели развертывания облачных хранилищ.
 2. Модели обслуживания облачных хранилищ.
 3. Постановка и описание проблемы «последней мили».

4.Безопасность, производительность и надежность при работе с облачными данными.

5.Экономическая составляющая облачных подходов.

6.Способы машинного обучения.

7.Основные фазы обработки «больших данных».

Занятие 2. Подготовка данных и больших объемов данных, анализ, визуализация, презентация (4 час.)

4 Недоопределенные задачи анализа, необходимость визуализации для анализа

5 Классификация методов визуализации.

6 Преимущества визуализации больших массивов данных в виде сетей. Примеры визуализации

Критерии оценки теоретической подготовки к практическим работам

– 100-86 баллов выставляется студенту, если студент дал правильные ответы на все поставленные вопросы при собеседовании;

– 85-76 баллов – не более 2 неправильных или отсутствующих ответов на поставленные вопросы в собеседовании;

– 75-61 балл – 3-4 неправильных или отсутствующих ответов на поставленные вопросы в собеседовании;

– 60-50 баллов – более 4 неправильных или отсутствующих ответов на поставленные вопросы в собеседовании.

Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины

Текущая аттестация студентов. Текущая аттестация студентов по дисциплине «Наука о данных и аналитика больших объёмов данных» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация по дисциплине «Наука о данных и аналитика больших объёмов данных» проводится в форме контрольных мероприятий (выполнение практического задания, написание конспекта, выполнение домашнего задания) по оцениванию фактических результатов обучения студентов и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

- учебная дисциплина (активность на занятиях, своевременность выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине);
- степень усвоения теоретических знаний (конспект);
- уровень овладения практическими умениями и навыками по всем видам учебной работы (выполнение практического задания);
- результаты самостоятельной работы (написание домашнего задания).

Промежуточная аттестация студентов. Промежуточная аттестация студентов по дисциплине «Наука о данных и аналитика больших объёмов данных» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

По дисциплине «Наука о данных и аналитика больших объёмов данных» предусмотрен зачет в виде собеседования.

Краткая характеристика процедуры применения используемого оценочного средства. В результате изучения теоретического материала, выполнения и защиты практических занятий и заданий, предусмотренных для самостоятельной работы, студент последовательно осваивает материалы дисциплины и изучает вопросы, предназначенные для подготовки к промежуточной аттестации (зачету). Оценочные средства для текущей и промежуточной аттестации, а также критерии их оценки представлены в соответствующих структурных элементах Фонда оценочных средств.

Критерии выставления оценки студенту на зачете

по дисциплине «Наука о данных и аналитика больших объёмов данных»

Баллы (рейтинговой оценки)	Оценка зачета (стандартная)	Требования к сформированным компетенциям
86-100	«зачтено»	Студент 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 и приемами выполнения практических задач.
85-76	«зачтено»	Студент 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.
75-61	«зачтено»	Студент 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.
60-0	«не зачтено»	Оценка «не зачтено» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «не зачтено» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.