Федеральное государственное образовательное бюджетное учреждение высшего образования «ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ» (Финансовый университет)

Департамент анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных

25 04 2023 r
Е.А. Каменева
и методической работе
Проректор по учебной
УТВЕРЖДАЮ

Макрушин С.В., Блохин Н.В.

Технологии обработки больших данных

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки 38.03.02 - Менеджмент, ОП «Управление бизнесом», ОП "Управление бизнесом / Bachelor of Business Administration (BBA)", ОП "Управление финансами / Bachelor of Business Administration in Finance", ОП «Финансовый менеджмент»

Рекомендовано Ученым советом Факультета информационных технологий и анализа больших данных (протокол №31 от 18.04.2023г.)

Одобрено Советом учебно-научного Департамента анализа данных и машинного обучения (протокол №2 от 29.03.2023г.)

Москва 2023

Содержание

1. Наименование дисциплины
2.Перечень планируемых результатов освоения образовательной программы
(перечень компетенций) с указанием индикаторов их достижения и
планируемых результатов обучения по дисциплине2
3. Место дисциплины в структуре образовательных программ
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах
с выделением объема аудиторной (лекции, семинары) и самостоятельной
работы обучающихся6
5. Содержание дисциплины, структурированное по темам (разделам)
дисциплины с указанием их объемов (в академических часах) и видов
учебных занятий
5.1. Содержание дисциплины
5.2. Учебно-тематический план
5.3. Содержание семинаров, практических занятий
6. Перечень учебно-методического обеспечения для самостоятельной работы
обучающихся по дисциплине14
6.1. Перечень вопросов, отводимых на самостоятельное освоение
дисциплины, формы внеаудиторной самостоятельной работы14
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю
7. Фонд оценочных средств для проведения промежуточной аттестации
обучающихся по дисциплине
8. Перечень основной и дополнительной учебной литературы, необходимой
для освоения дисциплины
9. Перечень ресурсов информационно-телекоммуникационной сети
«Интернет», необходимых для освоения дисциплины
10. Методические указания для обучающихся по освоению дисциплины. 27
11. Перечень информационных технологий, используемых при
осуществлении образовательного процесса по дисциплине, включая
перечень необходимого программного обеспечения и информационных
справочных систем
12. Описание материально-технической базы, необходимой для
осуществления образовательного процесса по дисциплине

1. Наименование дисциплины

«Технологии обработки больших данных».

2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код	Наименование	Индикаторы	Результаты обучения (уме-						
компе-	компетенции	достижения	ния и знания), соотнесенные						
тен-		компетенции	с индикаторами достижения						
ции			компетенции						
	-		сом / Bachelor of Business Ad-						
ministra			achelor of Business Administra-						
tion in Finance», ОП «Финансовый менеджмент»									
ПКН-5	Владение основами	Применяет результаты	Знать технологии обработки						
	финансового учета	анализа финансовой,	больших данных, применимые						
	и отчетности, а	бухгалтерской, управ-	для анализа отчетности.						
	также принципами	ленческой отчетности							
	управленческого	при составлении страте-	Уметь применять технологии						
	учета в целях ис-	гических финансовых,	обработки больших данных,						
	пользования дан-	инвестиционных пла-	применимые для анализа от-						
	ных учета для при-	нов, отборе проектов и	четности.						
	нятия управленче-	принятии управленче-							
	ских решений	ских решений.							
		Анализирует и оцени-	Знать технологии обработки						
		вает финансово-хозяй-	больших данных, применимые						
		ственное состояние ор-	для комплексной обработки						
		ганизации и результаты	данных.						
		деятельности их внут-							
		ренних подразделений,	Уметь применять технологии						
		формирует ключевые	обработки больших данных,						
		показатели эффективно-	применимые для комплексной						
		сти для сбалансирован-	обработки данных.						
		ного управления дея-							
		тельностью организа-							
		ции.							
		бизнесом», Профиль: Упр							
ПКП-3	Способность рас-	Рассчитывает юнит-эко-	Знать технологии для выпол-						
	считывать юнит-	номику продукта и раз-	нения расчетов на больших						
	экономику про-	рабатывает требования	объемах данных.						
	дукта, разрабаты-	к продукту.							
	вать требования к		Уметь выполнять расчеты на						
	продукту и план		больших объемах данных.						
	развития продукта	Управляет планом раз-	Знать способы расчета слож-						
		вития продукта.	ных показателей и работы с						

			различными форматами дан-
			различными форматами дан-
			Уметь выполнять расчеты
			сложных показателей и рабо-
			тать с различными форматами
			данных.
ОП	«Управление бизнес	ом», Профиль: Менеджме	ент и управление бизнесом
ПКП-3	Способность пла-	Демонстрирует навыки	Знать современные техноло-
	нировать, участво-	использования в своей	гии расчета показателей со
	вать в процессах	работе новых техноло-	сложными зависимостями.
	текущего и страте-	гий для планирования и	
	гического планиро-	контроля в организации.	Уметь выполнять расчета по-
	вания и контроля в		казателей со сложными зави-
	организации		симостями с помощью совре-
	•		менных технологий.
		Использует метрики ре-	Знать современные техноло-
		зультативности компа-	гии расчета экономических по-
		нии на каждом этапе	казателей на большом объеме
		жизненного цикла ком-	данных.
		пании.	
			Уметь рассчитывать экономи-
			ческие показателей на боль-
			шом объеме данных с помо-
			1 ' '
			щью современных технологий.
		иесом / Bachelor of Busines	щью современных технологий. ss Administration (BBA)»
Пр	офиль: «Бизнес и пр	едпринимательство / Bus	щью современных технологий. s Administration (BBA)» iness & Entrepreneurship»
	офиль: «Бизнес и пр Способность опре-	едпринимательство / Bus Применяет теоретиче-	щью современных технологий. ss Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа
Пр	офиль: «Бизнес и пр Способность опре- делять тенденции в	едпринимательство / Bus Применяет теоретиче- ские знания и экономи-	щью современных технологий. s Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые
Пр	офиль: «Бизнес и пр Способность опре- делять тенденции в развитии современ-	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре-	щью современных технологий. s Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических
Пр	офиль: «Бизнес и пр Способность опре- делять тенденции в развитии современ- ной мировой и оте-	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре- деления основных тен-	щью современных технологий. s Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые
Пр	офиль: «Бизнес и пр Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо-	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре- деления основных тен- денций в развитии со-	щью современных технологий. s Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей.
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре- деления основных тен- денций в развитии со- временной мировой эко-	щью современных технологий. ss Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож-	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре- деления основных тен- денций в развитии со-	щью современных технологий. ss Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, при-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре- деления основных тен- денций в развитии со- временной мировой эко-	щью современных технологий. ss Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки эконо-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож-	едпринимательство / Bus Применяет теоретиче- ские знания и экономи- ческие законы для опре- деления основных тен- денций в развитии со- временной мировой эко- номики.	щью современных технологий. S Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей.
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики	щью современных технологий. S Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий	щью современных технологий. ss Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управлен-	щью современных технологий. S Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей.
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сфере	щью современных технологий. SS Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе отно-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сферемеждународного биз-	щью современных технологий. з Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе отно- сящихся к глобальной эконо-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сфере	щью современных технологий. S Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе относящихся к глобальной экономике и международной тор-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сферемеждународного биз-	щью современных технологий. з Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе отно- сящихся к глобальной эконо-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сферемеждународного биз-	щью современных технологий. з Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе отно- сящихся к глобальной экономике и международной тор- говле и финансам.
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сферемеждународного биз-	щью современных технологий. S Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе относящихся к глобальной экономике и международной торговле и финансам. Уметь применять технологии
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сферемеждународного биз-	щью современных технологий. з Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе отно- сящихся к глобальной экономике и международной тор- говле и финансам. Уметь применять технологии анализа больших данных, при-
Пр	офиль: «Бизнес и пре Способность опре- делять тенденции в развитии современ- ной мировой и оте- чественной эконо- мики, искать и находить возмож- ности для создания	едпринимательство / Bus Применяет теоретические знания и экономические законы для определения основных тенденций в развитии современной мировой экономики. Использует методики анализа последствий принимаемых управленческих решений в сферемеждународного биз-	щью современных технологий. S Administration (BBA)» iness & Entrepreneurship» Знать технологии анализа больших данных, применимые для обработки экономических показателей. Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе относящихся к глобальной экономике и международной торговле и финансам. Уметь применять технологии

			числе относящихся к глобаль-						
			ной экономике и международ-						
			ной торговле и финансам.						
Проф	Профиль «Управление маркетингом», «Управление маркетингом / Marketing Management»								
ПКП-2	Способность про-	Использует лучшие	Знать технологии анализа						
	водить маркетинго-	практики при планиро-	больших данных, применимые						
	вые исследования,	вании и проведении	для обработки экономических						
	анализировать	маркетинговых иссле-	показателей и показателей						
	конъюнктуру	дований.	маркетинговых исследований.						
	рынка и интерпре-	дованин.	маркетинговых неследовании.						
	тировать получен-		Уметь применять технологии						
	ные		анализа больших данных, при-						
			менимые для обработки эконо-						
	результаты для								
	принятия управ-		мических показателей и пока-						
	ленческих решений		зателей маркетинговых иссле-						
		Tr.	дований.						
		Демонстрирует навыки	Знать технологии анализа						
		в получении маркетин-	больших данных, применимые						
		говой информации с це-	для обработки маркетинговых						
		лью оценки конъюнк-	показателей и оценки марке-						
		туры рынка.	тинговой конъюнктуры.						
			Уметь применять технологии						
			анализа больших данных, при-						
			менимые для обработки марке-						
			тинговых показателей и						
			оценки маркетинговой конъ-						
			юнктуры.						
ОП «У			dministration in Finance», Про-						
пип э		равление финансами/ВВ							
11K11-3	Спосооность разра-		Знать технологии агрегации и						
	батывать стратегию ценообразова-	тегию ценообразова-	анализа ценовой информации,						
	1	ния.	представленной в больших						
	ния, осуществлять		массивах данных.						
	корректировку		*7						
	цены на основе ис-		Уметь использовать техноло-						
	следования рыноч-		гии агрегации и анализа цено-						
	ной конъюнктуры		вой информации, представлен-						
			ной в больших массивах дан-						
			ных.						
		2.Осуществляет кор-	Знать технологии построения						
		ректировку цены на	прогнозных моделей на основе						
		основе исследований	обработки больших массивов						
		рыночной конъюнк-	числовой информации.						
		туры.							
		1 y pm.	Уметь применять технологии						
			построения прогнозных моде-						
	I .	I .	r - r - r - r - r - r - r - r - r - r -						

			лей на основе обработки боль-		
			ших массивов числовой ин-		
			формации.		
()П «Финансовый ме	неджмент», Профиль: Фи	інансовый менеджмент		
ПКП-1	Способность оце-	Проводит необходимые	Знать технологии анализа		
	нивать тенденции и	для решения финансо-	больших данных, применимые		
	закономерности	вых задач организации	для обработки экономических		
	развития внешней	исследования внешней	показателей, в том числе отно-		
	и внутренней эко-	и внутренней среды с	сящихся к исследованиям		
	номической среды,	использованием совре-	внешней и внутренней среды		
	ее влияние на ре-	менных информацион-	организации.		
	зультаты хозяй-	ных технологий.			
	ственной деятель-		Уметь применять технологии		
	ности организации		анализа больших данных, при-		
	в текущей и долго-		менимые для обработки эконо-		
	срочной перспек-		мических показателей, в том		
	тиве		числе относящихся к исследо-		
			ваниям внешней и внутренней		
			среды организации.		
		Оценивает и прогнози-	Знать технологии анализа		
		рует закономерности	больших данных, применимые		
		развития внешней и	для обработки экономических		
		внутренней среды биз-	показателей, в том числе отно-		
		неса.	сящихся к прогнозированию		
			закономерностей развития		
			внешней и внутренней среды		
			бизнеса.		
			Уметь применять технологии		
			анализа больших данных, при-		
			менимые для обработки эконо-		
			мических показателей, в том		
			числе относящихся к прогно-		
			зированию закономерностей		
			развития внешней и внутрен-		
			ней среды бизнеса.		

3. Место дисциплины в структуре образовательных программ

Дисциплина «Технологии обработки больших данных» является дисциплиной Цикла профиля (элективный) по направлению подготовки 38.03.02 — Менеджмент, ОП «Управление бизнесом», ОП «Управление бизнесом / Bachelor of Business Administration (BBA)», ОП «Управление финансами / Bachelor of Business Administration in Finance», ОП «Финансовый менеджмент».

4. Объем дисциплины(модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

ОП «Управление бизнесом», ОП «Управление бизнесом / Bachelor of Business Administration (BBA)», ОП «Управление финансами / Bachelor of Business Administration in Finance», ОП «Финансовый менеджмент»

Очная форма обучения, 2021 / 2022 г.п. и т.д.

Вид учебной работы по дисциплине	Всего (в з.е. и часах)	Семестр 7 (в часах)
Общая трудоёмкость дисциплины	3/108	108
Контактная работа-	50/36	50/36
Аудиторные занятия		
Лекции	16/2	16/2
Семинары, практические занятия	34/34	34/34
Самостоятельная работа	58/72	58/72
Вид текущего контроля	Зачет	Зачет
Вид промежуточной аттестации	Домашнее	Домашнее
	творческое задание	творческое задание

ОП «Управление бизнесом»

очно-заочная форма обучения, 2021 / 2022 г.п. и т.д.

Вид учебной работы по	Всего	Семестр 8
дисциплине	(в з.е. и часах)	(в часах)
Общая трудоёмкость	3/108	108
дисциплины		
Контактная работа-	32/18	32/18
Аудиторные занятия		
Лекции	16/2	16/2
Семинары, практические занятия	16/16	16/16
Самостоятельная работа	76/90	76/90
Вид текущего контроля	Зачет	Зачет
Вид промежуточной аттестации	Домашнее	Домашнее
	творческое задание	творческое задание

Институт онлайн-образования ОП «Финансовый менеджмент»

Очно-заочная форма обучения, 2021, 2022 г.п. и т.д.

Вид учебной работы по дисциплине	Всего (в з.е. и часах)	Семестр 8 (в часах)
Общая трудоёмкость дисциплины	3/108	108
Контактная работа- Аудиторные занятия	32	32
Лекции	8	8
Семинары, практические занятия	24	24
Самостоятельная работа	76	76
Вид текущего контроля	Зачет	Зачет
Вид промежуточной аттестации	Домашнее	Домашнее
	творческое задание	творческое задание

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Тема 1. Библиотека NumPy и Pandas.

В рамках темы рассматривается технологический стек Python для обработки и анализа данных, возможности Python как glue language, специфика библиотеки NumPy и ее роль в экосистеме Python. Организация массивов в NumPy: хранение данных, создание массивов, принципы реализации операций с едиными исходными данными. Универсальные функции и применение функций по осям в NumPy. Принцип распространения значений при выполнении операций в NumPy: общий алгоритм и примеры Маскирование и прихотливое индексирование в NumPy.

В рамках темы рассматриваются возможности библиотеки Pandas. Организация Pandas DataFrame и организация индексации для DataFrame и Series; применение универсальных функций и работа с пустыми значениями

в Pandas. Объединение данных из нескольких Pandas DataFrame: общая логика и примеры. Рассматривается операция GroupBy в Pandas DataFrame и реализация в ней подхода «разбиение, применение и объединение».

Тема 2. Использование различных форматов файлов в задачах обработки данных.

В рамках темы рассматриваются принципы работы с файлами, файлы и операционные системы. Специфика текстовых и бинарных файлов.

В рамках темы рассматривается задача сериализации и десериализации данных и использование различных форматов файлов для ее решения. Описание формата файла JSON и пример описания данных в этом формате и вза-имодействия с ним в Python.

В рамках темы рассматриваются формат XML и модель DOM: общая характеристика, пример описания данных в XML и DOM, работа с ними с помощью библиотеки BeautifulSoup.

В рамках темы рассматривается проблематика форматов файлов для хранения и обработки больших данных. Форматы файлов NPY и HDF: общая характеристика, пример взаимодействия с данными этих форматов в Python.

Тема 3. Взаимодействие с табличными данными в приложениях обработки данных.

В рамках темы рассматривается формат файлов CSV, представление данных в этом формате и взаимодействие с ним в Python.

В рамках темы рассматриваются возможности использования Excel для внешних приложений обработки данных. Взаимодействие с Excel из Python с помощью библиотеки XLWings: принципы работы и примеры использования.

Тема 4. Визуализация данных.

В рамках темы рассматриваются основы работы с библиотекой matplotlib: организация системы координат, оформление осей, цвета и цветовые карты в matplotlib, стили линий и маркеры. Pyplot и объектно-ориентированный интерфейс matplotlib. Управление фигурами и создание множества графиков на одном рисунке. Различные типы графиков.

В рамках темы рассматривается визуализация данных с помощью библиотеки Pandas: набор методов для построения графиков, реализованный в структурах Series и DataFrame.

В рамках темы проводится введение в разведочный анализ данных: типы признаков, анализ распределений, анализ мер центральной тенденции и поиск выбросов, анализ взаимного распределения и парных корреляций. Проведение разведочного анализа данных с помощью библиотеки Seaborn.

Тема 5. Работа со строками в приложениях обработки данных.

В рамках темы рассматриваются возможности python по форматированию строк: %-форматирование, метод format, f-строки.

В рамках темы рассматриваются основы работы с регулярными выражениями: базовый синтаксис, примеры. Модуль *re* в Python. Примеры использования регулярных выражений.

В рамках темы рассматривается использования хэширования при работе со строками. Строки в библиотеке numpy.

Тема 6. Взаимодействие с базой данных в приложениях обработки данных.

В рамках темы рассматривается взаимодействие из Python с базой данных на примере API SQLite. Базовые возможности работы с транзакциями.

5.2. Учебно-тематический план

ОП «Управление бизнесом», ОП «Управление бизнесом / Bachelor of Business Administration (BBA)», ОП «Управление финансами / Bachelor of Business Administration in Finance», ОП «Финансовый менеджмент»

очная форма обучения, 2021 / 2022 г.п. и т.д.

	Наименование тем (разделов) дисциплины			онтактная		Формы	
No		м (разделов)	Аудиторная работа		Самосто-	текущего	
п/п			Общая, в т.ч.:	Лекции	Семинары, практические занятия	ятельная работа	контроля успеваемости
1	Библиотека NumPy и Pandas	22/22	12/10	4/2	8/8	10/12	
2	Использование различных форматов файлов в задачах обработки данных.	20/16	10/4	4/0	6/4	10/12	Участие в ре- шении задач
3	Взаимодействие с табличными данными в приложениях обработки данных.	14/18	6/6	2/0	4/6	8/12	на практиче- ских занятиях. Обсуждения по результа- там самостоя- тельной ра-
4	Визуализация данных	16/18	8/6	2/0	6/4	8/12	боты
5	Работа со стро- ками в приложе- ниях обработки данных	18/16	6/4	2/0	4/4	12/12	Участие в решении задач на практических занятиях. Обсуждения
6	Взаимодействие с базой данных в приложениях обработки данных.	18/16	8/4	2/0	6/8	10/12	по результатам самостоятельной работы
	В целом по дисциплине	108	50/36	16/2	34/34	58/72	Согласно учебному плану: домашнее творческое задание
	Итого в %		46/33	32/2	68/31	54/67	

ОП «Управление бизнесом»

очно-заочная форма обучения, 2021 / 2022 г.п. и т.д.

№	Наименование тем (разделов) дисциплины		Контактная работа- Аудиторная работа			Самосто-	Формы текущего
п/п		Всего	Общая, в т.ч.:	Лекции	Семинары, практические занятия	ятельная работа	контроля успеваемости
1	Библиотека NumPy и Pandas	22/22	8/6	4/2	4/4	14/16	Участие в решении задач
2	Использование различных форматов файлов в задачах обработки данных.	16/18	4/2	2/0	2/2	12/16	на практиче- ских занятиях. Обсуждения по результа- там самостоя-
3	Взаимодействие с табличными данными в приложениях обработки данных.	16/16	4/2	2/0	2/2	12/14	тельной ра- боты Участие в ре- шении задач
4	Визуализация данных	16/16	4/2	2/0	2/2	12/14	на практиче- ских занятиях. Обсуждения
5	Работа со стро- ками в приложе- ниях обработки данных	16/16	4/2	2/0	2/2	12/14	по результа- там самостоя- тельной ра- боты
6	Взаимодействие с базой данных в приложениях обработки данных.	22/20	8/4	4/0	4/4	14/16	
	В целом по дисциплине	108	32/18	16/2	16/16	76/90	Согласно учебному плану: домаш- нее творче- ское задание
	Итого в %		30/17	50/11	50/89	70/83	

ОП «Финансовый менеджмент»

очно-заочная форма обучения (ИОО), 2021, 2022 г.п. и т.д.

		Трудоемкость в часах						
№	Наименование тем (разделов) дисциплины		Контактная работ Аудиторная работ			Самосто-	Формы текущего	
п/п		Всего	Общая, в т.ч.:	Лекции	Семинары, практические занятия	ятельная работа	контроля успеваемости	
1	Библиотека NumPy и Pandas	20	6	2	4	14	Участие в решении задач	
2	Использование различных форматов файлов в задачах обработки данных	17	5	1	4	12	на практиче- ских занятиях. Обсуждения по результа- там самостоя-	
3	Взаимодействие с табличными данными в приложениях обработки данных	17	5	1	4	12	тельной ра- боты	
4	Визуализация данных	17	5	1	4	12	Участие в ре- шении задач на практиче-	
5	Работа со стро- ками в приложе- ниях обработки данных	17	5	1	4	12	ских занятиях. Обсуждения по результатам самостоятельной ра-	
6	Взаимодействие с базой данных в приложениях обработки данных.	20	6	2	4	14	боты	
	В целом по дисциплине	108	32	8	24	76	Согласно учебному плану: домаш- нее творче- ское задание	
	Итого в %		30	25	75	70		

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Библиотека NumPy и Pandas	 Технологический стек Python для обработки и анализа данных Возможности Python как glue language Организация массивов в NumPy: хранение данных, создание массивов Принципы реализации операций с едиными исходными данными. Универсальные функции и применение функций по осям в NumPy. Организация Pandas DataFrame и организация индексации для DataFrame и Series. Применение универсальных функций и работа с пустыми значениями в Pandas. Объединение данных из нескольких Pandas DataFrame: общая логика и примеры. 8[1], 9[9], 9[10] 	Интерактивная форма, работа на компьютере
Использование различных форматов файлов в задачах обработки данных	 Формат файлов Pickle, представление данных в этом формате и взаимодействие с ним в Python. Формат файлов JSON, представление данных в этом формате и взаимодействие с ним в Python. Формат XML и модель DOM: общая характеристика, пример описания данных в XML и DOM Работа с XML с помощью библиотеки BeautifulSoup. 8[1], 8[2], 9[3], 9[4] 	Интерактивная форма, работа на компьютере
Взаимодействие с табличными данными в приложениях обработки данных.	 Взаимодействие с Excel из Python с помощью библиотеки XLWings. Формат файлов CSV, представление данных в этом формате и взаимодействие с ним в Python 8[1], 8[2] 	Интерактивная форма, работа на компьютере
Визуализация дан- ных	 Построение визуализаций с помощью библиотеки matplotlib Построение визуализаций с помощью библиотеки pandas Построение визуализаций с помощью библиотеки seaborn 8[1], 9[13], 9[15], 9[16] 	Интерактивная форма, работа на компьютере
Работа со строками в приложениях обра- ботки данных	 Основы работы с регулярными выражениями: базовый синтаксис, примеры. Модуль ге в Python. 8[1], 8[2], 9[4] 	Интерактивная форма, работа на компьютере
Взаимодействие с ба-	• Взаимодействие из Python с базой данных с помощью API SQLite.	Интерактивная форма, работа на

зой данных в приложениях обработки данных	8[1], 8[2]	компьютере
Профилирование процессов обработки данных, библиотека Numba.	 профилирование реализации алгоритмов на Python принципы решения задачи оптимизации производительности алгоритма Библиотека Numba: принципы работы, базовые примеры использования. 8[1], 8[2], 9[1], 9[2], 9[3] 	Интерактивная форма, работа на компьютере
Параллельная обработка данных, введение в Dask	 специфика современного аппаратного обеспечения для обработки больших данных и проблема масштабируемости параллельных вычислений. Подходы к декомпозиции крупных вычислительных задач на подзадачи для параллельного исполнения. Проблема Global Interpreter Lock в Python и способы обхода ее ограничений. Модуль Python multiprocessing – назначение и основные возможности, API multiprocessing. Pool. Подход к обработке данных с помощью библиотеки Dask. Структура данных Dask. Array – принцип работы, API, примеры использования. Структура данных Dask. Data Frame – принцип работы, API, примеры использования 8[1], 8[2], 9[5], 9[7], 9[8], 9[10], 9[11] 	Интерактивная форма, работа на компьютере

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисци- плины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
		1
Библиотека NumPy	• Принцип распространения значений при	Обзор литературы и веб-
и Pandas	выполнении операций в NumPy: общий	источников. Самостоя-
	алгоритм и примеры.	тельное освоение инстру-
	• Маскирование и прихотливое индексиро-	ментов аналитической
	вание в NumPy.	обработки. Решение за-
	• Операция GroupBy в Pandas DataFrame и	дач.

	реализация в ней подхода «разбиение,	
Использование различных форматов файлов в задачах обработки данных	применение и объединение». Формат файлов NPY, представление данных в этом формате и взаимодействие с ним в Python. Формат файлов HDF, представление данных в этом формате и взаимодействие с ним в Python.	Обзор литературы и вебисточников. Самостоятельное освоение инструментов аналитической обработки. Решение задач.
Взаимодействие с табличными данными в приложениях обработки данных.	• Продвинутые операции с Excel из Python с помощью библиотеки XLWings.	Обзор литературы и веб- источников. Самостоя- тельное освоение инстру- ментов аналитической обработки. Решение за- дач.
Визуализация дан- ных	• Построение трехмерных графиков Продвинутая работа с цветовыми картами	Обзор литературы и веб- источников. Самостоя- тельное освоение инстру- ментов аналитической обработки. Решение за- дач.
Работа со строками в приложениях обработки данных	Использования хэширования при работе со строками.Строки в библиотеке numpy.	Обзор литературы и вебисточников. Самостоятельное освоение инструментов аналитической обработки. Решение задач.
Взаимодействие с базой данных в приложениях обработки данных	• Базовые возможности работы с транзакциями с помощью API SQLite.	Обзор литературы и вебисточников. Самостоятельное освоение инструментов аналитической обработки. Решение задач.
Профилирование процессов обра- ботки данных, биб- лиотека Numba.	 Векторизация в питру: ключевые параметры функции, примеры применения Использование обобщенной сигнатуры функции в питру и питва. 	Обзор литературы и вебисточников. Самостоятельное освоение инструментов аналитической обработки. Решение задач.
Параллельная обра- ботка данных, вве- дение в Dask	 Модели параллельного программирования и их сочетаемость с архитектурами параллельных вычислительных систем. Специфика различия между потоками и процессами. Организация вычислений с помощью Мар / Filter / Reduce: общий принцип и специфика параллельной реализации обработки данных с помощью Dask.Bag. Многопроцессорные архитектуры с общей и разделяемой памятью — специфика и сравнение. 	Обзор литературы и вебисточников. Самостоятельное освоение инструментов аналитической обработки. Решение задач.

6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю

Примерная тематика домашнего творческого задания

- 1. Прогнозная аналитика и моделирование объемов продаж акций
- 2. Визуализация аналитических данных в области макроэкономики
- 3. Визуализация аналитических данных Московской биржи
- 4. Использование технологии больших данных для анализа портфельных рисков
- 5. Использование параллельных вычислений реализации численных методов решения математических задач
- 6. Анализ и сравнение различных фреймворков для визуализации данных
- 7. Применение распределенных вычислений и экосистемы Hadoop для решения задачи анализа данных
- 8. Анализ больших данных для построения прогнозов на рынке ценных бумаг
- 9. Использование больших данных для оценки кредитоспособности контрагентов на основе анализа текстов новостей
- 10. Проведение анализа собранных из внешних источников данных

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Департамента анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2. «Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».

Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Наименование компетенции ОП «Управлен	Наименование индикаторов достижения компетенции ме бизнесом». ОП «Упр	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
			Business Administration in
	Finance», ОП «Ф	инансовый менеджме	нт»
ПКН-5. Владение основами финансового учета и отчетности, а также принципами управленческого учета в целях использования данных учета для принятия управ-	Применяет результаты анализа финансовой, бухгалтерской, управленческой отчетности при составлении стратегических финансовых, инвестиционных планов, отборе проектов и принятии управленческих решений.	Знать технологии обработки больших данных, применимые для анализа отчетности. Уметь применять технологии обработки больших данных, применимые для анализа отчетности.	Предложите технологический стек для анализа первичной отчетности объемом 100Мб, 1Гб, 10Гб. Реализуйте алгоритм агрегации показателей для набора данных объемом 1Гб.
ленческих решений	Анализирует и оценивает финансовохозяйственное состояние организации и результаты деятельности их внутренних подразделений, формирует ключевые показатели эффективности для сбаланси-	Знать технологии обработки больших данных, применимые для комплексной обработки данных. Уметь применять технологии обработки больших данных, применимые	Предложите технологический стек для многокритериальной группировки данных первичной отчетности объемом 100Мб, 1Гб, 10Гб. Реализуйте алгоритм многокритериальной группировки данных пер-

	<u> </u>		
	рованного управле-	для комплексной об-	вичной отчетности объе-
	ния деятельностью	работки данных.	мом 1Гб.
	организации.	•	
		авление бизнесом»	
HICH 2		правление продуктом	T
ПКП-3. Спо-	1.Рассчитывает	Знать современные	Предложите технологи-
собность рас-	юнит-экономику	технологии расчета	ческий стек для расчета
считывать	продукта и разраба-	показателей со	показателей юнит-эконо-
юнит-эконо-	тывает требования к	сложными зависи-	мики, объем которых со-
мику продукта,	продукту.	мостями.	ставляет 10Мб, 100Мб,
разрабатывать		Vacous printed arrows	1Гб.
требования к		Уметь выполнять	December 2
продукту и		расчета показателей	Реализуйте алгоритм рас-
план развития		со сложными зави-	чета показателей юнит-
продукта		симостями с помо-	экономики, объем данных
		щью современных	для которых составляет
	2 Vypapygar yyayay	технологий.	100Мб.
	2. Управляет планом	Знать современные	Предложите технологический стек для расчета
	развития продукта.	технологии расчета экономических по-	экономических показате-
		казателей на боль-	лей, объем которых со-
		шом объеме данных.	ставляет 10Мб, 100Мб,
		шом ооъсме данных.	1Гб.
		Уметь рассчитывать	11 0.
		экономические по-	Реализуйте алгоритм рас-
		казателей на боль-	чета экономических пока-
		шом объеме данных	зателей (например, при-
		с помощью совре-	были по ЦФО), объем
		менных технологий.	данных для которых со-
			ставляет 100Мб.
	Профиль: Менеджі	мент и управление биз	знесом
ПКП-3. Спо-	1.Демонстрирует	Знать технологии	Предложите технологи-
собность пла-	навыки использова-	анализа больших	ческий стек для расчета и
нировать,	ния в своей работе	данных, примени-	обработки экономиче-
участвовать в	новых технологий	мые для обработки	ских показателей, объем
процессах теку-	для планирования и	экономических по-	которых составляет
щего и страте-	контроля в организа-	казателей.	10Мб, 100Мб, 1Гб.
гического пла-	ции.		
нирования и		Уметь применять	Реализуйте алгоритм рас-
контроля в ор-		технологии анализа	чета экономических пока-
ганизации		больших данных,	зателей и планирования
		применимые для об-	(например, производ-
		работки экономиче-	ственного плана), объем
		ских показателей.	данных для которых со-
	2 77		ставляет 100Мб.
	2. Использует мет-	Знать технологии	Предложите технологи-
	рики результативно-	анализа больших	ческий стек для расчета
	сти компании на каж-	данных, примени-	экономических показате-
	дом этапе жизнен-	мые для обработки	лей международной эко-

	ного цикла компа-	экономических по-	номики, торговли и фи-
	нии.	казателей, в том	нансов объем которых со-
		числе относящихся	ставляет 10Мб, 100Мб,
		к глобальной эконо-	1Гб.
		мике и международ-	
		ной торговле и фи-	Реализуйте алгоритм рас-
		нансам.	чета экономических пока-
			зателей международной
		Уметь применять	экономики, торговли и
		технологии анализа	финансов (например,
		больших данных,	транзакций на крупней-
		применимые для об-	ших глобальных бирже-
		работки экономиче-	вых площадках), объем
		ских показателей, в	данных для которых со-
		том числе относя-	ставляет 100Мб.
		щихся к глобальной	
		экономике и между-	
		народной торговле и	
		финансам.	
ОП «Уп	равление бизнесом / Ва	achelor of Business Adr	ninistration (BBA)»
Профиль:	«Бизнес и предприним	мательство / Business о	& Entrepreneurship»
ПКП-3. Спо-	1.Применяет теоре-	Знать технологии	Предложите технологи-
собность опре-	тические знания и	анализа больших	ческий стек для расчета
делять тенден-	экономические за-	данных, примени-	экономических и марке-
ции в развитии	коны для определе-	мые для обработки	тинговых показателей,
современной	ния основных тен-	экономических по-	объем которых состав-
мировой и оте-	денций в развитии	казателей и показа-	ляет 10Мб, 100Мб, 1Гб.
чественной	современной миро-	телей маркетинго-	
экономики, ис-	вой экономики.	вых исследований.	Реализуйте алгоритм рас-
кать и находить			чета маркетинговых пока-
возможности		Уметь применять	зателей (например, дан-
для создания		технологии анализа	ных о показах баннерной
бизнеса		больших данных,	рекламы), объем данных
		применимые для об-	для которых составляет
		работки экономиче-	100Мб.
		ских показателей и	
		показателей марке-	
		тинговых исследова-	
		ний.	
	2.Использует мето-	Знать технологии	Предложите технологи-
	дики анализа послед-	анализа больших	ческий стек для расчета и
	ствий принимаемых	данных, примени-	анализа экономических и
	управленческих ре-	мые для обработки	маркетинговых показате-
	шений в сфере меж-	маркетинговых по-	лей, объем которых со-
	дународного биз-	казателей и оценки	ставляет 10Мб, 100Мб,
	неса.	маркетинговой	1Γ6.
		конъюнктуры.	
		J 1	Реализуйте алгоритм ана-
		Уметь применять	лиза маркетинговых по-
		технологии анализа	•

			T
		больших данных,	казателей (например, про-
		применимые для об-	гноза эффективности бан-
		работки маркетинго-	нерной рекламы), объем
		вых показателей и	данных для которых со-
		оценки маркетинго-	ставляет 100Мб.
		вой конъюнктуры.	
Профили: «У			ркетингом / Marketing
	M	anagement»	
ПКП-2. Спо-	1.Использует лучшие	Знать технологии	Предложите технологи-
собность про-	практики при плани-	анализа больших	ческий стек для расчета
водить марке-	ровании и проведе-	данных, примени-	экономических и марке-
тинговые ис-	нии маркетинговых	мые для обработки	тинговых показателей,
следования,	исследований.	экономических по-	объем которых состав-
анализировать		казателей и показа-	ляет 10Мб, 100Мб, 1Гб.
конъюнктуру		телей маркетинго-	
рынка и интер-		вых исследований.	Реализуйте алгоритм рас-
претировать		*7	чета маркетинговых пока-
полученные		Уметь применять	зателей (например, обра-
результаты для		технологии анализа	ботки результатов А/В те-
принятия		больших данных,	стирования), объем дан-
управленче-		применимые для об-	ных для которых состав-
ских решений		работки экономиче-	ляет 100Мб.
		ских показателей и	
		показателей марке-	
		тинговых исследова-	
	2 Полеотерия	ний.	Постория
	2. Демонстрирует	Знать технологии анализа больших	Предложите технологи-
	навыки в получении маркетинговой ин-		ческий стек для обра- ботки экономических и
	маркетинговой информации с целью	данных, примени- мые для обработки	маркетинговых показате-
	оценки конъюнктуры	маркетинговых по-	лей, объем которых со-
	рынка.	казателей и оценки	ставляет 10Мб, 100Мб,
	рынка.	маркетинговой	1Гб.
		конъюнктуры.	11 0.
		коныонктуры.	Реализуйте алгоритм об-
		Уметь применять	работки маркетинговых
		технологии анализа	показателей (например,
		больших данных,	прибыли по ЦФО), объем
		применимые для об-	данных для которых со-
		работки маркетинго-	ставляет 100Мб.
		вых показателей и	
		оценки маркетинго-	
		вой конъюнктуры.	
ОП "Управлен	ие финансами / Bache	• • • • • • • • • • • • • • • • • • • •	stration in Finance'', Про -
	филь: "Управление финансами/BBA in Finance"		
ПКП-3. Спо-	1.Разрабатывает	Знать технологии	Для таблицы большого
собность разра-	стратегию ценообра-	агрегации и анализа	объема с ценами для раз-
батывать стра-	зования.	ценовой информа-	ных SCU в разных регио-
тегию ценооб-		ции, представленной	нах и по разным датам
1			рассчитать с помощью

разования, осу-		в больших массивах	технологий обработки
ществлять кор-		данных.	больших данных рассчи-
ректировку			тать средние цены по ре-
цены на основе		Уметь использовать	гиону и укрупненному
исследования		технологии агрега-	периоду (месяцу).
рыночной		ции и анализа цено-	
конъюнктуры		вой информации,	
		представленной в	
		больших массивах	
		данных.	
	2.Осуществляет кор-	Знать технологии	Для таблицы большого
	ректировку цены на	построения прогноз-	объема с ценами для раз-
	основе исследований	ных моделей на ос-	ных SCU в разных регио-
	рыночной конъюнк-	нове обработки больших массивов	нах и по разным датам
	туры		рассчитать с помощью
		числовой информа-	технологий обработки
		ции.	больших данных рассчи-
		W 7	тать на основе экстрапо-
		Уметь применять	ляции прогноз средних
		технологии построе-	цен по региону и укруп-
		ния прогнозных мо-	ненному периоду (ме-
		делей на основе об-	сяцу).
		работки больших	
		массивов числовой	
		информации.	
	нансовый менеджмен		
ПКП-1. Спо-	1. Проводит необхо-	Знать технологии	Предложите технологи-
собность оце-	димые для решения	анализа больших	ческий стек для обра-
нивать тенден-	финансовых задач	данных, примени-	ботки экономических по-
ции и законо-	организации иссле-	мые для обработки	казателей, характеризую-
мерности раз-	дования внешней и	экономических по-	щих внешнюю и внутрен-
вития внешней	внутренней среды с	казателей, в том	нюю среду организации,
и внутренней	использованием со-	числе относящихся	объем которых состав-
экономической	временных информа-	к исследованиям	ляет 10Мб, 100Мб, 1Гб.
среды, ее влия-	ционных техноло-	внешней и внутрен-	
ние на резуль-	гий.	ней среды организа-	Реализуйте алгоритм об-
таты хозяй-		ции.	работки экономических
ственной дея-			показателей, характери-
тельности орга-		Уметь применять	зующих внешнюю и внут-
низации в теку-		технологии анализа	реннюю среду организа-
щей и долго-		больших данных,	ции (например, прибыли
срочной пер-		применимые для об-	по ЦФО организации),
спективе		работки экономиче-	объем данных для кото-
		ских показателей, в	рых составляет 100Мб.
		том числе относя-	pan corrabiner round.
		щихся к исследова-	
		ниям внешней и	
		риутперием спепт	
		внутренней среды организации.	

2. Оценивает и прогнозирует закономерности развития внешней и внутренней среды бизнеса. Знать технологии анализа больших данных, применимые для обработки экономических показателей, в том числе относящихся к прогнозированию закономерностей развития внешней и внутренней среды бизнеса.

Уметь применять технологии анализа больших данных, применимые для обработки экономических показателей, в том числе относящихся к прогнозированию закономерностей развития внешней и внутренней среды бизнеса.

Предложите технологический стек для обработки экономических показателей, в том числе прогнозирования внешней и внутренней среды организации, объем которых составляет 10Мб, 100Мб, 1Гб.

Реализуйте алгоритм прогнозирования экономических показателей, характеризующих внешнюю и внутреннюю среду организации (например, прогноз прибыли по ЦФО организации), объем данных для которых составляет 100Мб.

Примерные вопросы для подготовки к зачету

- 1. Большие данные определение и причины возникновения задач обработки больших данных
- 2. Специфика современного аппаратного обеспечения для обработки больших данных и проблема масштабируемости параллельных вычислений
- 3. Выбор типичных средств обработки данных, адекватных различным объемам данных; принцип обработки данных на базе операций map / filter / reduce
- 4. Многопроцессорные архитектуры с общей и разделяемой памятью специфика и сравнение
- 5. Подходы к декомпозиции крупных вычислительных задач на подзадачи для параллельного исполнения

- 6. Модели параллельного программирования и их сочетаемость с архитектурами параллельных вычислительных систем
- 7. Профилирование реализации алгоритмов на Python, принципы решения задачи оптимизации производительности алгоритма
- 8. Проблема Global Interpreter Lock в Python и способы обхода ее ограничений
- 9. Технологический стек Python для обработки и анализа данных, Python как glue language, специфика библиотеки NumPy и ее роль в экосистеме Python
- 10. Организация массивов в NumPy: хранение данных, создание массивов, принципы реализации операций с едиными исходными данными
- 11. Универсальные функции и применение функций по осям в NumPy
- 12. Принцип распространения значений при выполнении операций в NumPy: общий алгоритм и примеры
- 13. Маскирование и прихотливое индексирование в NumPy
- 14. Векторизация в numpy: ключевые параметры функции, примеры применения, использование обобщенной сигнатуры функции
- 15. Numba: принципы работы, базовые примеры использования
- 16. Организация Pandas DataFrame и организация индексации для Data-Frame и Series
- 17. Применение универсальных функций и работа с пустыми значениями в Pandas
- 18. Объединение данных из нескольких Pandas DataFrame: общая логика и примеры
- 19. Операция GroupBy в Pandas DataFrame и реализация в ней подхода «разбиение, применение и объединение»
- 20. Специфика текстовых и бинарных файлов, форматы файлов CSV и Pickle, представление данных в этих форматах и взаимодействие с ними в Python

- 21.Задача сериализации и десериализации, описание формата файла JSON и пример описания данных в этом формате и взаимодействия с ним в Python
- 22. Формат XML и модель DOM: общая характеристика, пример описания данных в XML и DOM, работа с ними с помощью библиотеки BeautifulSoup
- 23. Форматы файлов NPY и HDF общая характеристика, пример взаимодействие с данными этих форматов в Python
- 24. Взаимодействие из Python с базой данных на примере API SQLite, базовые возможности работы с транзакциями
- 25.Взаимодействие с Excel из Python с помощью XLWings: принципы работы и примеры использования
- 26. Основы работы с регулярными выражениями: базовый синтаксис, примеры использования модуля re в Python
- 27. Сегментация и токенезация текста на естественном языке, стеммминг и лемматизация, примеры на Python
- 28. Различия между потоками и процессами, различие между различными планировщиками в Dask
- 29. Граф зависимостей задач суть структуры данных, ее построение и использование в Dask
- 30. Три ключевых структуры данных Dask: их специфика и принцип выбора структуры данных при решении задач
- 31. Dask. Array структура данных, специфика реализации и применения, процедура создания
- 32.Dask.Array поддерживаемые операции и отличия от NumPy ndarray
- 33. Распараллеливание алгоритмов с помощью dask.delayed принцип и примеры использования
- 34.Дополнительные параметры декоратора dask.delayed назначение и примеры использования

- 35. Использование dask.delayed для объектов и операции над объектами dask.delayed, включая ограничения их использования
- 36.Dask.DataFrame структура данных, специфика реализации и применения, процедура создания Dask.DataFrame
- 37.Ограничения использования Dask.DataFrame и операции мэппинга в Dask.DataFrame
- 38.Поддержка Dask.DataFrame операций работающих со скользящим окном
- 39. Совместное использование промежуточных результатов в Dask: принцип работы и примеры использования
- 40. Dask. Bag структура данных, специфика реализации и применения, процедура создания Dask Bag
- 41. Организация вычислений с помощью Map / Filter / Reduce : общий принцип и специфика параллельной реализации обработки данных в Dask. Bag
- 42. Понятие признака в анализе данных и типы признаков
- 43.Понятие разведочного анализа данных, основные задачи и типовые визуализации для решения этих задач

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Колдаев, В. Д. Структуры и алгоритмы обработки данных : учебное пособие / В. Д. Колдаев. - Москва : РИОР : ИНФРА-М, 2021. - 296 с. - ЭБС ZNANIUM.com. - URL: https://znanium.com/catalog/product/1230215 (дата обращения: 15.06.2023). – Текст : электронный.

Дополнительная литература:

2. Нагаева, И. А. Основы алгоритмизации и программирования: практикум: учебное пособие / И. А. Нагаева, И. А. Кузнецов. – Москва: Берлин: Директ-Медиа, 2021. – 169 с. – ЭБС Университетская библиотека ONLINE. – URL: https://biblioclub.ru/index.php?page=book&id=598404 (дата обращения: 15.06.2023). – Текст: электронный.

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

- 1. Электронная библиотека Финансового университета (ЭБ) http://elib.fa.ru/
 - 2. Электронно-библиотечная система BOOK.RU http://www.book.ru
- 3. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» http://biblioclub.ru/
 - 4. Электронно-библиотечная система Znanium http://www.znanium.com
- 5. Электронно-библиотечная система издательства «ЮРАЙТ» https://urait.ru/
- 6. Электронно-библиотечная система издательства Проспект http://ebs.prospekt.org/books
- 7. Электронно-библиотечная система издательства Лань https://e.lanbook.com/
 - 8. Деловая онлайн-библиотека Alpina Digital http://lib.alpinadigital.ru/
- 9. Электронная библиотека Издательского дома «Гребенников» https://grebennikon.ru/
 - 10. Научная электронная библиотека eLibrary.ru http://elibrary.ru
 - 11. Национальная электронная библиотека http://нэб.рф/
- 12. Pylru 1.0.9 [Электронный ресурс]: сайт. Режим доступа: https://pypi.python.org/pypi/pylru

- 13. Python Data Analysis Library [Электронный ресурс]: сайт. Режим доступа: http://pandas.pydata.org/
- 14Python Documentation [Электронный ресурс]: сайт. Режим доступа: http://python.org/doc/
 - 15. Python Standard Library [Электронный ресурс]: сайт. Режим доступа: https://docs.python.org/2/library/
 - 16. Scikit-learn Machine Learning in Python [Электронный ресурс]: сайт.
 - Режим доступа: http://scikit-learn.org
 - 17. Официальный сайт продукта https://www.python.org/
 - 18. Каталог курсов Интернет Университета Информационных Технологий http://www.intuit.ru/
 - 19. The Python Tutorial // https://docs.python.org/3/tutorial/index.html
 - 20. NumPy User Guide // http://docs.scipy.org/doc/numpy/user/index.html
 - 21. Pandas User Guide http://pandas.pydata.org/pandas-docs/stable/
 - 22. Dask User Guide https://docs.dask.org/en/latest/
 - 23. Dask User Guide https://docs.dask.org/en/latest/
 - 24. Matplotlib User Guide // https://matplotlib.org/stable/users/index.html
 - 25. Seaborn User Guide // https://seaborn.pydata.org/tutorial.html

10. Методические указания для обучающихся по освоению дисциплины

При изучении теоретического материала необходимо опираться на рабочую программу дисциплины, материалы лекций и литературу из основного списка. Кроме этого, необходимо активно работать с Интернет-источниками и пособиями других авторов, помогающими усвоить материал отдельных разделов программы.

Необходимо конспектировать лекции, помечая сложные и непонятные моменты с тем, чтобы задать вопросы лектору в конце лекции или же на консультации.

При подготовке к семинарским занятиям необходимо изучить вопросы, вынесенные на самостоятельное изучение, так как семинарские занятия предполагают их обсуждение и дискуссию по теме; кроме того, задания для самостоятельной работы необходимы для того, чтобы успешно выполнить самостоятельные задания на семинарах.

Индивидуальные задания для работы на компьютере, файлы с выполненными заданиями необходимо хранить в личной сетевой папке в компьютерной сети вуза.

- 11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем
- 11.1. Комплект лицензионного программного обеспечения:
 - 1. Пакет офисных программ
 - 2. Антивирус Kaspersky
- 11.2. Современные профессиональные базы данных и информационные справочные системы:
 - 1. Информационно-правовая система «Гарант»
 - 2. Информационно-правовая система «Консультант Плюс»
 - 3. Электронная энциклопедия: http://ru.wikipedia.org/wiki/Wiki
- 4. Система комплексного раскрытия информации «СКРИН» http://www.skrin.ru/
- 11.3. Сертифицированные программные и аппаратные средства защиты информации
 - не используются

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Для проведения лекций и практических занятий необходима аудитория, оснащенная проектором и компьютерами с постоянным подключением к сети Интернет.