

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Пекаревский Борис Владимирович
Должность: Проректор по учебной и методической работе
Дата подписания: 02.11.2023 13:15:16
Уникальный программный ключ:
3b89716a1076b80b2c167df0f27c09d01782ba84



МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Санкт-Петербургский государственный технологический институт
(технический университет)»

УТВЕРЖДАЮ
Проректор по учебной
и методической работе
_____ Б.В. Пекаревский
« 24 » мая 2021 г.

**Рабочая программа дисциплины
БОЛЬШИЕ ДАННЫЕ**

Направление подготовки
27.03.03 Системный анализ и управление
Направленности программы бакалавриата
«Системный анализ в информационных технологиях»

Квалификация

Бакалавр

Форма обучения

Заочная

Факультет **информационных технологий и управления**
Кафедра **системного анализа и информационных технологий**

Санкт-Петербург

2021

ЛИСТ СОГЛАСОВАНИЯ

Должность разработчика	Подпись	Ученое звание, фамилия, инициалы
доцент		доцент, Ананченко И.В.

Рабочая программа дисциплины «Большие данные» обсуждена на заседании кафедры системного анализа и информационных технологий
протокол от «28» 04 2021 № 7

Заведующий кафедрой

А.А. Мусаев

Одобрено учебно-методической комиссией факультета информационных технологий и управления
протокол от «19» 05 2021 № 8

Председатель

В.В. Куркина

СОГЛАСОВАНО

Руководитель направления подготовки «Системный анализ и управление»		Д.А. Краснобородько
Директор библиотеки		Т.Н. Старостенко
Начальник методического отдела учебно-методического управления		Т.И. Богданова
Начальник учебно-методического управления		С.Н. Денисенко

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.....	04
2. Место дисциплины (модуля) в структуре образовательной программы.....	05
3. Объем дисциплины	05
4. Содержание дисциплины	
4.1. Разделы дисциплины и виды занятий.....	06
4.2. Занятия лекционного типа.....	07
4.3. Занятия семинарского типа.....	08
4.3.1. Семинары, практические занятия	08
4.4. Самостоятельная работа.....	10
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	12
6. Фонд оценочных средств для проведения промежуточной аттестации.....	12
7. Перечень учебных изданий, необходимых для освоения дисциплины.....	13
8. Перечень электронных образовательных ресурсов, необходимых для освоения дисциплины	13
9. Методические указания для обучающихся по освоению дисциплины.....	13
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине	
10.1. Информационные технологии.....	14
10.2. Программное обеспечение.....	14
10.3. Базы данных и информационно-справочные системы	14
11. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине.....	14
12. Особенности освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья	14

Приложения: 1. Фонд оценочных средств для проведения промежуточной аттестации.

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

В результате освоения образовательной программы обучающийся должен овладеть следующими результатами обучения по дисциплине:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения (дескрипторы)
<p>ПК-5 Способен управлять работами по созданию (модификации) и сопровождению информационных систем, автоматизирующих задачи организационного управления</p>	<p>ПК-5.2 Разработка алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности</p>	<p>Знать: - принципы применения методов анализа больших данных для решения задач профессиональной деятельности (ЗН-1). Уметь: - внедрять методы анализа больших данных для решения задач профессиональной деятельности (У-1). Владеть: - методами анализа больших данных и способен сформировать выводы в соответствии с поставленной задачей. (Н-1).</p>
		<p>Знать: - принципы разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности (ЗН-2). Уметь: - внедрять алгоритмы анализа и обработки большого объема данных для решения задач профессиональной деятельности (У-2). Владеть: - навыками решения задач анализа и обработки большого объема данных для решения задач профессиональной деятельности (Н-2).</p>

2. Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам части, формируемой участниками образовательных отношений (Б1.В.17) и изучается на 3 курсе.

Изучение данной дисциплины базируется на знании студентами основ математики, информатики и основ алгоритмизации, на знаниях, полученных в процессе изучения дисциплин «Операционные системы», «Введение в информационные технологии», «Алгоритмы и структуры данных».

Полученные в процессе изучения дисциплины «Большие данные» знания, умения и навыки могут быть использованы при изучении дисциплин, «Лингвистическое и программное обеспечение автоматизированных информационных систем», «Разработка программного обеспечения для мобильных систем», а также в научно-исследовательской работе бакалавра и при выполнении выпускной квалификационной работы.

3. Объем дисциплины

Вид учебной работы	Всего, академических часов
	Заочная форма обучения
Общая трудоемкость дисциплины (зачетных единиц/ академических часов)	5/ 180
Контактная работа с преподавателем:	14
занятия лекционного типа	4
занятия семинарского типа, в т.ч.	8
семинары, практические занятия (в том числе практическая подготовка)	8(1)
лабораторные работы	-
курсовое проектирование (КР или КП)	2
КСР	
другие виды контактной работы	
Самостоятельная работа	157
Форма текущего контроля (Кр, реферат, РГР, эссе)	Кр(2)
Форма промежуточной аттестации (КР, КП, зачет, экзамен)	Экзамен(9), КР

4. Содержание дисциплины

4.1. Разделы дисциплины и виды занятий

№ п/п	Наименование раздела дисциплины	Занятия лекционного типа, академ. часы	Занятия семинарского типа, академ. часы		Самостоятельная работа, академ. часы	Формируемые компетенции	Формируемые индикаторы
			Семинары и/или практические занятия	Лабораторные работы			
1.	Наука о данных в Azure. Введение в процесс Data Science (наука о данных). Работа с данными с помощью службы машинного обучения Azure. Автоматизация машинного обучения с помощью службы Azure Machine Learning.	0,5	0,5		15	ПК-5	ПК-5.2
2	Управление и мониторинг моделей машинного обучения с помощью службы Azure Machine Learning. Введение в Azure Cosmos DB. Проектирование и реализация приложения баз данных API-интерфейса SQL.	0,3	0,5		15	ПК-5	ПК-5.2
3.	Реализация операций на стороне сервера. Оптимизация и мониторинг производительности. Проектирование и реализация базы данных графов.	0,2	1		15	ПК-5	ПК-5.2
4.	Извлечение и анализ больших данных (Big Data) в Cosmos DB. Введение в решения самостоятельной бизнес-аналитики. Знакомство с Power BI.	0,5	1		15	ПК-5	ПК-5.2
5.	Данные Power BI. Формирование и объединение данных. Моделирование данных.	0,5	1		15	ПК-5	ПК-5.2
6.	Интерактивная визуализация данных. Прямое подключение. API для разработчиков.	0,5	1		15	ПК-5	ПК-5.2
7.	Power BI Mobile. Анализ данных в Excel. Модель данных Excel.	0,5	1		25	ПК-5	ПК-5.2

8.	Импорт данных из файлов. Импорт данных из базы данных. Импорт данных из отчетов Excel.	0,5	1		25	ПК-5	ПК-5.2
9.	Создание и форматирование меры. Визуализация данных в Excel. Использование Excel с Power BI.	0,5	1		17	ПК-5	ПК-5.2

4.2. Занятия лекционного типа

№ раздела дисципли-	Наименование темы и краткое содержание занятия	Объем, акад. часы	Инновационная форма
1.	Наука о данных в Azure. Введение в процесс Data Science (наука о данных). Работа с данными с помощью службы машинного обучения Azure. Автоматизация машинного обучения с помощью службы Azure Machine Learning. Процесс науки о данных и роли специалиста по данным (data scientist), использование служб Azure в поддержке и расширении процесса обработки данных. Использование службы машинного обучения Azure Machine Learning для полной автоматизации процесса обработки данных. Описание конвейера машинного обучения, а также использование служб AutoML и HyperDrive для автоматизации некоторых трудоемких задач.	0,5	ЛВ
2.	Управление и мониторинг моделей машинного обучения с помощью службы Azure Machine Learning. Введение в Azure Cosmos DB. Проектирование и реализация приложения баз данных API-интерфейса SQL. Обзор структуры баз данных NoSQL. Модели документов в Cosmos DB. Извлечение данных в базах API-интерфейса SQL	0,3	ЛВ
3.	Реализация операций на стороне сервера. Оптимизация и мониторинг производительности. Проектирование и реализация базы данных графов. Программные возможности на стороне сервера в Cosmos DB. Оптимизация производительности базы данных. Проектирование моделей базы данных графов для эффективной работы.	0,2	ЛВ
4.	Извлечение и анализ больших данных (Big Data) в Cosmos DB. Введение в решения самостоятельной бизнес-аналитики. Знакомство с Power BI. Интеграция базы данных Cosmos DB с поиском Azure для оптимизации запросов. Анализ данных в базе данных Cosmos DB с использованием Apache Spark. Введение в анализ данных. Введение в визуализацию данных.	0,5	ЛВ
5.	Данные Power BI. Формирование и объединение данных. Моделирование данных. Использование Excel как источник данных Power BI. Модель данных Power BI. Связывание данных. DAX запросы.	0,5	ЛВ

№ раздела дисципли-	Наименование темы и краткое содержание занятия	Объем, акад. часы	Инновационная форма
6.	Интерактивная визуализация данных. Прямое подключение. API для разработчиков. Создания отчетов Power BI. Управление решениями Power BI. Облачные данные. Подключение к службе анализа данных.	0,5	ЛВ
7	Power BI Mobile. Анализ данных в Excel. Модель данных Excel. Приложения Power BI Mobile. Классический анализ данных с помощью Excel. Сводные Таблицы Excel. Использование модели данных в Excel. DAX. Изучение модели данных в Excel	0,5	ЛВ
8	Импорт данных из файлов. Импорт данных из базы данных. Импорт данных из отчетов Excel. Импорт данных в Excel. Обработка и преобразование данных. Доступные источники данных. Просмотр, обработка и преобразование данных. Связи и иерархии таблиц.	0,5	ЛВ
9	Создание и форматирование меры. Визуализация данных в Excel. Использование Excel с Power BI. Расширенные функции DAX. Создание мер с помощью расширенных функций DAX.	0,5	ЛВ

4.3. Занятия семинарского типа

4.3.1. Семинары, практические занятия

№ раздела дисциплины	Наименование темы и краткое содержание занятия	Объем, акад. часы		Инновационная форма
		всего	в том числе на практическую подготовку	
1	Использование службы машинного обучения Azure Machine Learning для полной автоматизации процесса обработки данных. Внедрение службы машинного обучения Azure Machine Learning (AML). Регистрация и развертывание модели ML с помощью службы AML. Автоматизация выбора модели машинного обучения. Автоматизация настройки гиперпараметра с помощью HyperDrive.	0,5	0,25	Слайд-презентация, групповая дискуссия
2	Миграция данных и приложений в Cosmos DB. Управление данными в Cosmos DB. Модели документов в Cosmos DB. Извлечение данных в базах API-интерфейса SQL. Программное извлечение и сопровождение данных. Проектирование и реализация приложения баз данных API-интерфейса SQL.	0,5	0,25	Слайд-презентация, групповая дискуссия

№ раздела дисциплины	Наименование темы и краткое содержание занятия	Объем, акад. часы		Инновационная форма
		всего	в том числе на практическую подготовку	
3	Программные возможности на стороне сервера в Cosmos DB. Создание и использование хранимых процедур. Использование триггеров для реализации целостности данных. Оптимизация производительности базы данных. Мониторинг производительности базы данных. Проектирование моделей базы данных графов для эффективной работы.	1		Слайд-презентация, групповая дискуссия
4	Визуализация данных в базе данных Cosmos DB. Извлечение и анализ больших данных (Big Data) в Cosmos DB. Реализация потоковой обработки с помощью базы данных Cosmos DB. Работа с поддержкой веб-канала изменений в Azure Cosmos DB. Инструменты Microsoft для самостоятельной бизнес-	1		Слайд-презентация, групповая дискуссия
5	Использование баз данных как источник данных. Power BI. Служба Power BI. Импорт данных в Power BI. Объединение данных. Формирование и объединение данных. DAX запросы.	1		Слайд-презентация, групповая дискуссия
6	Создания отчетов Power BI. Управление решениями Power BI. Создания отчетов Power BI. API для разработчиков. Пользовательские визуализации. Использование API для разработчиков.	1	0,25	Слайд-презентация, групповая дискуссия
7	Ограничения классического анализа данных. Построение классической экспресс-панели в Excel. Фильтрация и форматирование данных. Изучение модели данных в Excel. Создание вычисляемых столбцов. Форматирование данных модели данных. Создание меры.	1		Слайд-презентация, групповая дискуссия
8	Обработка и преобразование данных. Загрузка данных. Импорт данных из файла CSV. Импорт и преобразование данных из файла CSV. Добавление данных из папки. Импорт данных из SQL Server. Импорт данных из CSV-файла.	1	0,25	Слайд-презентация, групповая дискуссия

№ раздела дисциплины	Наименование темы и краткое содержание занятия	Объем, акад. часы		Инновационная форма
		всего	в том числе на практическую подготовку	
9	Визуализация данных в Excel. Создание табличного отчета. Создание сводной диаграммы. Добавление срезов в диаграмму. Как использовать Excel с Power BI. Power BI. Выгрузка данных в Excel для Power BI Мобильное приложение Power BI. Создание экспресс-панели Power BI в Excel. Выгрузка данных в Excel. Создание экспресс-панели Power BI.	1		Слайд-презентация, групповая дискуссия

4.4. Самостоятельная работа обучающихся

№ раздела дисциплины	Перечень вопросов для самостоятельного изучения	Объем, акад. часы	Форма контроля
1	Использование службы машинного обучения Azure Machine Learning для полной автоматизации процесса обработки данных. Внедрение службы машинного обучения Azure Machine Learning (AML). Регистрация и развертывание модели ML с помощью службы AML. Автоматизация машинного обучения с помощью службы Azure Machine Learning. Описание конвейера машинного обучения, а также использование служб AutoML и HyperDrive для автоматизации некоторых трудоемких задач.	15	Устный опрос №1
2	Введение в Azure Cosmos DB. Обзор структуры баз данных NoSQL. Миграция данных и приложений в Cosmos DB. Управление данными в Cosmos DB. Создание и использование баз данных SQL API в Cosmos DB. Проектирование и реализация приложения баз данных API-интерфейса SQL. Модели документов в Cosmos DB. Извлечение данных в базах API-интерфейса SQL. Программное извлечение и сопровождение данных. Проектирование и реализация приложения баз данных API-интерфейса SQL.	15	Устный опрос №1

№ раздела дисциплины	Перечень вопросов для самостоятельного изучения	Объем, акад. часы	Форма контроля
3	Создание и использование хранимых процедур. Использование триггеров для реализации целостности данных. Написание пользовательских функций, хранимых процедур и триггеров. Оптимизация и мониторинг производительности. Оптимизация производительности базы данных. Мониторинг производительности базы данных. Настройка базы данных и мониторинг производительности. Проектирование и реализация базы данных графов. Модели базы данных графов и Cosmos DB. Проектирование моделей базы данных графов для эффективной работы.	15	Контрольная работа №1
4	Power BI. Служба Power BI. Создание Power BI dashboard. Введение в решения самостоятельной бизнес-аналитики. Знакомство с бизнес-аналитикой. Введение в анализ данных. Введение в визуализацию данных. Обзор самостоятельной бизнес-аналитики. Инструменты Microsoft для самостоятельной бизнес-аналитики. Обзор корпоративных решений анализа данных.	15	Устный опрос №2
5	Формирование и объединение данных. Запросы Power BI desktop. Формирование данных. Объединение данных. Формирование и объединение данных. Моделирование данных. Связывание данных. DAX запросы. Вычисления и меры. Моделирование данных.	15	Устный опрос №2
6	Управление решениями Power BI. Создания отчетов Power BI. Прямое подключение. Облачные данные. Подключение к службе анализа данных. Прямое подключение. API для разработчиков. API для разработчиков. Пользовательские визуализации. Использование API для разработчиков.	15	Устный опрос №2
7	Сводные Таблицы Excel. Ограничения классического анализа данных. Построение классической экспресс-панели в Excel. Фильтрация и форматирование данных. Построение сводной таблицы. Построение сводной диаграммы. Экспресс-панели Модель данных Excel. Рассматривается модель данных в Excel и способы расширения ее возможностей. Использование модели данных в Excel. DAX. Изучение модели данных в Excel. Создание вычисляемых столбцов. Форматирование данных модели данных. Создание меры. Анализ данных.	25	Контрольная работа №2

№ раздела дисциплины	Перечень вопросов для самостоятельного изучения	Объем, акад. часы	Форма контроля
8	Импорт данных в Excel. Обработка и преобразование данных. Загрузка данных. Импорт данных из файла CSV. Импорт и преобразование данных из файла CSV. Добавление данных из папки. Импорт данных из базы данных. Как импортировать данные в Excel из базы данных SQL Server. Доступные источники данных. Просмотр, обработка и преобразование данных. Связи и иерархии таблиц. Загрузка данных. Импорт данных из нескольких источников. Импорт данных из SQL Server. Импорт данных из CSV-файла. Создание таблицы данных. Импорт данных из отчетов Excel. Как импортировать данные из отчета Excel. Импорт данных из отчетов в формате Excel. Преобразование данных отчета Excel. Импорт данных из отчетов.	25	Устный опрос №3
9	DAX. Меры. Расширенные функции DAX. Создание мер с помощью расширенных функций DAX. Сравнение показателей с прошлым годом. Показатели текущего года. Доли рынка. Сводные диаграммы. Функции куба. Диаграммы для функций куба. Визуализация данных в Excel. Создание табличного отчета. Создание сводной диаграммы. Добавление срезов в диаграмму.	17	Устный опрос №3

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

Методические указания для обучающихся по организации самостоятельной работы по дисциплине, включая перечень тем самостоятельной работы, формы текущего контроля по дисциплине и требования к их выполнению размещены в электронной информационно-образовательной среде СПбГТИ(ТУ) на сайте: <http://media.technolog.edu.ru>

6. Фонд оценочных средств для проведения промежуточной аттестации

Промежуточная аттестация по дисциплине проводится в виде экзамена и защиты курсовой работы.

Экзамен предусматривает выборочную проверку освоения предусмотренных элементов компетенций и комплектуется теоретическими вопросами для проверки знаний.

При сдаче экзамена студент получает два вопроса из перечня вопросов, время подготовки студента к ответу - до 30 мин.

Пример варианта вопросов на экзамене:

Вариант № 1
<ol style="list-style-type: none"> 1. Извлечение данных в базах API-интерфейса SQL. 2. Использование Excel с Power BI.

Фонд оценочных средств по дисциплине представлен в Приложении № 1

Результаты освоения дисциплины считаются достигнутыми, если для всех элементов компетенций достигнут пороговый уровень освоения компетенции на данном этапе – оценка «удовлетворительно».

7. Перечень учебных изданий, необходимых для освоения дисциплины

а) печатные издания:

1. Информатика. Базовый курс: Учебное пособие для вузов / под ред. С. В. Симоновича. - 3-е изд. - Москва; Санкт-Петербург; Новгород: Питер, 2016. - 640 с. – ISBN 978-5-496-00217-2.
2. Шапорев, С. Д. Информатика. Теоретический курс и практические занятия: Учебник для вузов по направлениям 230100 "Информатика и вычислительная техника", 230200 "Информационные системы" / С. Д. Шапорев. - Санкт-Петербург : БХВ Петербург, 2009. - 469 с. – ISBN 978-5-9775-0242-9.
3. Норенков, И. П. Автоматизированные информационные системы: учеб. пособие для студ. вузов, обуч. по направлению 230100«Информатика и вычислительная техника» (УМО) / И. П. Норенков. - Москва: МГТУ им. Н. Э. Баумана, 2011. - 343 с. – ISBN 978-5-7038-3446-6.

б) электронные учебные издания:

1. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : Московский инженерно-строительный институт им. В. В. Куйбышева – Московский государственный строительный университет, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/145102> (дата обращения: 23.03.2021). — Режим доступа: по подписке.
2. Методы и модели исследования сложных систем и обработки больших данных : монография / И. Ю. Парамонов, В. А. Смагин, Н. Е. Косых, А. Д. Хомоненко ; под редакцией В. А. Смагина и А. Д. Хомоненко. — Санкт-Петербург : Лань, 2020. — 236 с. — ISBN 978-5-8114-4006-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/126938> (дата обращения: 23.03.2021). — Режим доступа: по подписке.

8. Перечень электронных образовательных ресурсов, необходимых для освоения дисциплины

учебный план, РПД и учебно-методические материалы: <http://media.technolog.edu.ru>

электронно-библиотечные системы:

«Электронный читальный зал – БиблиоТех» <https://technolog.bibliotech.ru/>;

«Лань» <https://e.lanbook.com/books/>

9. Методические указания для обучающихся по освоению дисциплины

Все виды занятий по дисциплине «Большие данные» проводятся в соответствии с требованиями следующих СТП:

СТП СПбГТИ 040-02. КС УКДВ. Виды учебных занятий. Лекция. Общие требования;

СТО СПбГТИ 018-2014. КС УКДВ. Виды учебных занятий. Семинары и практические занятия. Общие требования к организации и проведению.

СТП СПбГТИ 048-2009. КС УКДВ. Виды учебных занятий. Самостоятельная планируемая работа студентов. Общие требования к организации и проведению.

Планирование времени, необходимого на изучение данной дисциплины, лучше всего осуществлять на весь семестр, предусматривая при этом регулярное повторение пройденного материала.

Основными условиями правильной организации учебного процесса для студентов является:

- плановость в организации учебной работы;
- серьезное отношение к изучению материала;
- постоянный самоконтроль.

На занятия студент должен приходить, имея знания по уже изученному материалу.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

10.1. Информационные технологии

В учебном процессе по данной дисциплине предусмотрено использование информационных технологий:

- чтение лекций с использованием слайд-презентаций;
- взаимодействие с обучающимися посредством ЭИОС.

10.2. Программное обеспечение

Программы: ОС Microsoft Windows, ОС Kali Linux, ОС AstraLinux, ОС Ubuntu, MathCAD, Microsoft Office (Microsoft Word, Microsoft Excel, Microsoft Access, Microsoft PowerPoint), интегрированная среда Microsoft Visual Studio Community. VMware Workstation Player. Hyper-V. MS Virtual PC.

10.3. Базы данных и информационно справочные системы

Справочно-поисковая система «Консультант-Плюс»

11. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине.

Для ведения лекционных и практических занятий используется компьютерный класс, оснащенный объединенными в сеть персональными компьютерами, оборудованием и техническими средствами обучения на необходимое количество посадочных мест. При проведении занятий используется аудитория, оборудованная при необходимости проектором для отображения презентаций. Кроме того, при проведении лекций и практических занятий необходим компьютер с установленным на нем браузером и программным обеспечением для демонстрации презентаций (Power Point и др.). Для самостоятельной работы с медиаматериалами каждому студенту требуется персональный компьютер или планшет, широкополосный доступ в сеть Интернет, браузер последней версии, устройство для воспроизведения звука (динамики, колонки, наушники и др.)

12. Особенности освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья.

Для инвалидов и лиц с ограниченными возможностями учебные процесс осуществляется в соответствии с Положением об организации учебного процесса для обучения инвалидов и лиц с ограниченными возможностями здоровья СПбГТИ(ТУ), утвержденным ректором 28.08.2014г.

**Фонд оценочных средств
для проведения промежуточной аттестации по
дисциплине «Большие данные»**

1. Перечень компетенций и этапов их формирования.

Индекс компетенции	Содержание	Этап формирования
ПК-5	Способен управлять работами по созданию (модификации) и сопровождению информационных систем, автоматизирующих задачи организационного управления	промежуточный

2. Показатели и критерии оценивания компетенций на различных этапах их формирования, шкала оценивания

Код и наименование индикатора достижения компетенции	Показатели сформированности (дескрипторы)	Критерий оценивания	УРОВНИ СФОРМИРОВАННОСТИ (описание выраженности дескрипторов)		
			«удовлетворительно» (пороговый)	«хорошо» (средний)	«отлично» (высокий)
ПК-5.2 Разработка алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности	Правильно определяет применение методов анализа больших данных для решения задач профессиональной деятельности (ЗН-1).	Ответы на вопросы №1 - 33 к экзамену	Затрудняется в четком определении основных методов анализа больших данных для решения задач профессиональной деятельности.	Определяет основные методы анализа больших данных для решения задач профессиональной деятельности.	Демонстрирует глубокие знания методов анализа больших данных для решения задач профессиональной деятельности.
	Демонстрирует навыки применения методов анализа больших данных для решения задач профессиональной деятельности (У-1).	Ответы на вопросы №1 - 33 к экзамену	Демонстрирует слабые навыки применения методов анализа больших данных для решения задач профессиональной деятельности.	Демонстрирует с ошибками навыки применения методов анализа больших данных для решения задач профессиональной деятельности.	Демонстрирует хорошие навыки применения методов анализа больших данных для решения задач профессиональной деятельности.
	Перечисляет и приводит примеры решения задач применения методов анализа больших данных для решения задач профессиональной деятельности (Н-1).	Ответы на вопросы №1 - 33 к экзамену	Затрудняется с решением задач применения методов анализа больших данных для решения задач профессиональной деятельности.	Справляется с решением типовых задач применения методов анализа больших данных для решения задач профессиональной деятельности.	Демонстрирует хорошие навыки и умения решения задач применения методов анализа больших данных для решения задач профессиональной деятельности

	Правильно определяет принципы разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности (ЗН-2).	Ответы на вопросы №34 - 72 к экзамену	Затрудняется в четком определении основных принципов разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.	Определяет основные принципы разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.	Демонстрирует глубокие знания принципов разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.
	Демонстрирует навыки разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности (У-2).	Ответы на вопросы №34 - 72 к экзамену	Демонстрирует слабые навыки разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.	Демонстрирует с ошибками навыки разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.	Демонстрирует хорошие навыки разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.
	Перечисляет и приводит примеры решения задач разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности (Н-2).	Ответы на вопросы №34 - 72 к экзамену; защита курсовой работы	Затрудняется с решением задач разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.	Справляется с решением типовых задач разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.	Демонстрирует хорошие навыки и умения решения задач разработки алгоритмов анализа и обработки большого объема данных для решения задач профессиональной деятельности.

Шкала оценивания соответствует СТО СПбГТИ(ТУ):

По дисциплине промежуточная аттестация проводится в форме экзамена, шкала оценивания – балльная («отлично», «хорошо», «удовлетворительно», «неудовлетворительно»).

Шкала оценивания курсовой работы – балльная («отлично», «хорошо», «удовлетворительно», «неудовлетворительно»).

3. Типовые контрольные задания для проведения промежуточной аттестации

а) Вопросы для оценки знаний, умений и навыков, сформированных у студента по компетенции ПК-5:

1. Процесс науки о данных и роли специалиста по данным (data scientist), использование служб Azure в поддержке и расширении процесса обработки данных.
2. Azure Notebooks. Работа с данными с помощью службы машинного обучения Azure.
3. Использование службы машинного обучения Azure Machine Learning для полной автоматизации процесса обработки данных.
4. Внедрение службы машинного обучения Azure Machine Learning (AML)
5. Регистрация и развертывание модели ML с помощью службы AML.
6. Автоматизация машинного обучения с помощью службы Azure Machine Learning
7. Описание конвейера машинного обучения, использование служб AutoML и HyperDrive для автоматизации некоторых трудоемких задач.
8. Автоматизация выбора модели машинного обучения. Автоматизация настройки гиперпараметра с помощью HyperDrive
9. Управление и мониторинг моделей машинного обучения с помощью службы Azure Machine Learning.
10. Введение в Azure Cosmos DB.
11. Проектирование и реализация приложения баз данных API-интерфейса SQL.
12. Управление и мониторинг моделей машинного обучения с помощью службы Azure Machine Learning
13. Введение в Azure Cosmos DB. Обзор структуры баз данных NoSQL. Миграция данных и приложений в Cosmos DB
14. Управление данными в Cosmos DB. Создание и использование баз данных SQL API в Cosmos DB.
15. Проектирование и реализация приложения баз данных API-интерфейса SQL. Модели документов в Cosmos DB.
16. Извлечение данных в базах API-интерфейса SQL. Программное извлечение и сопровождение данных. Проектирование и реализация приложения баз данных API-интерфейса SQL.
17. Реализация операций на стороне сервера. Оптимизация и мониторинг производительности. Проектирование и реализация базы данных графов.
18. Реализация операций на стороне сервера. Программные возможности на стороне сервера в Cosmos DB. Создание и использование хранимых процедур.
19. Использование триггеров для реализации целостности данных. Написание пользовательских функций, хранимых процедур и триггеров.
20. Оптимизация и мониторинг производительности. Оптимизация производительности базы данных. Мониторинг производительности базы данных.
21. Настройка базы данных и мониторинг производительности. Проектирование и реализация базы данных графов.
22. Модели базы данных графов и Cosmos DB. Проектирование моделей базы данных графов для эффективной работы. Проектирование и реализация базы данных графов.
23. Извлечение и анализ больших данных (Big Data) в Cosmos DB. Введение в решения самостоятельной бизнес-аналитики. Знакомство с Power BI.
24. Извлечение и анализ больших данных (Big Data) в Cosmos DB. Интеграция базы данных Cosmos DB с поиском Azure для оптимизации запросов.
25. Анализ данных в базе данных Cosmos DB с использованием Apache Spark. Визуализация данных в базе данных Cosmos DB. Извлечение и анализ больших данных (Big Data) в Cosmos DB.

26. Реализация потоковой обработки с помощью базы данных Cosmos DB. Работа с поддержкой веб-канала изменений в Azure Cosmos DB. Интеграция базы данных Cosmos DB с потоковыми решениями.
27. Использование базы данных Cosmos DB с потоковыми решениями. Введение в решения самостоятельной бизнес-аналитики.
28. Знакомство с бизнес-аналитикой. Введение в анализ данных. Введение в визуализацию данных. Обзор самостоятельной бизнес-аналитики.
29. Инструменты Microsoft для самостоятельной бизнес-аналитики. Обзор корпоративных решений анализа данных.
30. Power BI. Служба Power BI. Создание Power BI dashboard
31. Данные Power BI. Формирование и объединение данных. Моделирование данных.
32. Данные Power BI. Использование Excel как источник данных Power BI.
33. Модель данных Power BI. Использование баз данных как источник данных Power BI.
34. Служба Power BI. Импорт данных в Power BI.
35. Формирование и объединение данных. Запросы Power BI desktop.
36. Формирование данных. Объединение данных. Формирование и объединение данных.
37. Моделирование данных. Связывание данных. DAX запросы. Вычисления и меры. Моделирование данных.
38. Интерактивная визуализация данных. Прямое подключение. API для разработчиков.
39. Интерактивная визуализация данных. Создания отчетов Power BI. Управление решениями Power BI. Создания отчетов Power BI.
40. Прямое подключение. Облачные данные. Подключение к службе анализа данных.
41. Прямое подключение. API для разработчиков.
42. API для разработчиков. Пользовательские визуализации.
43. Использование API для разработчиков.
44. Power BI Mobile. Анализ данных в Excel. Модель данных Excel.
45. Power BI Mobile. Приложения Power BI Mobile. Загрузка и использование приложений Power BI Mobile.
46. Встраиваемые компоненты Power BI. 1: Анализ данных в Excel.
47. Классические возможности Excel и способы их расширения. Классический анализ данных с помощью Excel.
48. Сводные Таблицы Excel. Ограничения классического анализа данных. Построение классической экспресс-панели в Excel.
49. Фильтрация и форматирование данных. Построение сводной таблицы. Построение сводной диаграммы. Экспресс-панели.
50. Модель данных Excel. Рассматривается модель данных в Excel и способы расширения ее возможностей. Использование модели данных в Excel. DAX.
51. Изучение модели данных в Excel. Создание вычисляемых столбцов. Форматирование данных модели данных. Создание меры. Анализ данных.
52. Импорт данных из файлов. Импорт данных из базы данных. Импорт данных из отчетов Excel.
53. Импорт данных из файлов. Как предварительно подготовить и импортировать CSV-файлы.
54. Импорт данных в Excel. Обработка и преобразование данных. Загрузка данных. Импорт данных из файла CSV.
55. Импорт и преобразование данных из файла CSV. Добавление данных из папки.
56. Импорт данных из базы данных. Как импортировать данные в Excel из базы данных SQL Server.

57. Доступные источники данных. Просмотр, обработка и преобразование данных. Связи и иерархии таблиц.
58. Загрузка данных. Импорт данных из нескольких источников. Импорт данных из SQL Server.
59. Импорт данных из CSV-файла. Создание таблицы данных.
60. Импорт данных из отчетов Excel. Как импортировать данные из отчета Excel. Импорт данных из отчетов в формате Excel.
61. Преобразование данных отчета Excel. Импорт данных из отчетов. Импорт данных из Excel. Преобразование данных Excel.
62. Загрузка данных в модель данных Excel.
63. Создание и форматирование меры. Визуализация данных в Excel. Использование Excel с Power BI.
64. Создание и форматирование меры. Как создавать и форматировать меры. DAX. Меры
65. Расширенные функции DAX. Создание мер с помощью расширенных функций DAX. Сравнение показателей с прошлым годом.
66. Показатели текущего года. Доли рынка.
67. Визуализация данных в Excel. Сводные диаграммы. Функции куба.
68. Диаграммы для функций куба. Визуализация данных в Excel. Создание табличного отчета.
69. Создание сводной диаграммы. Добавление срезов в диаграмму.
70. Использование Excel с Power BI. Выгрузка данных в Excel для Power BI.
71. Мобильное приложение Power BI. Создание экспресс-панели Power BI в Excel.
72. Выгрузка данных в Excel. Создание экспресс-панели Power BI.

При сдаче экзамена студент получает два вопроса, сформированных на основе перечня, приведенного выше. Время подготовки студента к устному ответу на вопросы - до 30 мин.

Темы и содержание контрольных работ

Контрольная работа №1

Поиск и определение больших данных.

Требуется осуществить поиск источников информации в сети интернет: открытые и закрытые источники данных. Рассмотреть портал открытых данных РФ. Определить можно ли классифицировать найденные данные как «Big data» .

Контрольная работа №2

Работа с моделью больших данных.

Модель данных Excel. Рассматривается модель данных в Excel и способы расширения ее возможностей. Использование модели данных в Excel. DAX.

4. Темы курсовых работ

1. Создание модели документов в Cosmos DB.
2. Мониторинг производительности базы данных.
3. Проектирование и реализация базы данных графов.
4. Оптимизация производительности базы данных.
5. Проектирование и реализация приложения баз данных.
6. Облачные данные.

5. Методические материалы для определения процедур оценивания знаний, умений и навыков, характеризующих этапы формирования компетенций.

Промежуточная аттестация по дисциплине проводится в соответствии с требованиями СПб СТО СПбГТИ(ТУ) 016-2015. КС УКДВ Порядок проведения зачетов и экзаменов.

СПб СТО СПбГТИ 044-2012. КС УКВД. Виды учебных занятий. Курсовой проект. Курсовая работа. Общие требования.