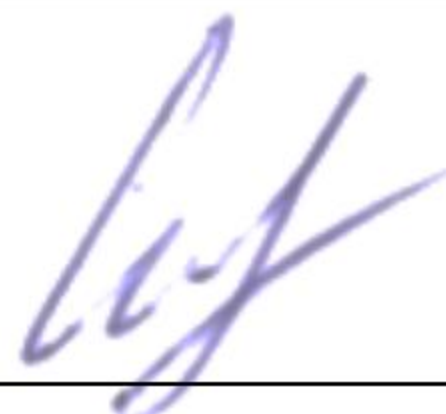


ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ
ПРИЕМ 2020 г.
ФОРМА ОБУЧЕНИЯ очная

Инструментарий для работы с большими данными

Направление подготовки/ специальность	09.04.04 Программная инженерия		
Образовательная программа (направленность (профиль))	Big Data Solutions / Технологии больших данных (язык обучения: английский)		
Специализация	Big Data Solutions / Технологии больших данных		
Уровень образования	высшее образование – магистратура		

Курс	1	семестр	2
Трудоемкость в кредитах (зачетных единицах)	3		

Заведующий кафедрой- руководитель ОИТ на правах кафедры		Шерстнев В.С.
Руководитель ООП		Савельев А.О.
Преподаватель		Губин М.Ю.

2020 г.

1. Роль дисциплины «Аналитика больших данных» в формировании компетенций выпускника:

Код компетенции	Наименование компетенции	Индикаторы достижения компетенций		Составляющие результатов освоения (дескрипторы компетенций)	
		Код индикатора	Наименование индикатора достижения	Код	Наименование
ОПК(У)-3	Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями	И.ОПК (У)-3.1	Анализирует профессиональную информацию оформляет и представляет ее в виде аналитических обзоров.	ОПК(У)-3.1B1	Владеет методами анализа и исследования информационных процессов и технологий
				ОПК(У)-3.1У1	Умеет выполнять анализ и проводить исследования статистически собранных данных
				ОПК(У)-3.131	Знает теоретические основы обработки статистически накопленной информации и методы её анализа
		И.ОПК (У)-3.2	Структурирует профессиональную информацию выделяя в ней основные элементы: цели, гипотезы, результаты, теории, классификации, аргументы и т.п.	ОПК(У)-3.2B1	Владеет навыками анализа современных достижений и методами подготовки научных докладов, публикаций и аналитических обзоров с обоснованными выводами и рекомендациями в области информационных технологий
				ОПК(У)-3.2У1	Умеет структурировать профессиональную информацию, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями

2. Показатели и методы оценивания

Планируемые результаты обучения по дисциплине		Код индикатора достижения контролируемой компетенции (или ее части)	Наименование раздела дисциплины	Методы оценивания (оценочные мероприятия)
Код	Наименование			
РД1	Применять знания подходов к обработке больших данных для определения, какая экосистема и в какой конфигурации может быть использована для решения конкретной задачи обработки больших данных.	И.ОПК (У)-3.1 И.ОПК (У)-3.2	Раздел (модуль) 1. Хранение и обработка больших наборов данных в Apache Hive Раздел (модуль) 2. Анализ больших наборов данных в Apache Hive	Представление и защита результатов практических работ по проектированию структуры таблиц в Apache Hive. Представление и защита результатов лабораторных работ по созданию запросов с использованием языка HiveQL.
РД2	Применять навыки администрирования хранилищ больших данных для планирования конфигурации, развертывания,	И.ОПК (У)-3.1 И.ОПК (У)-3.2	Раздел (модуль) 1. Хранение и обработка больших наборов данных в Apache Hive	Представление и защита результатов практических работ по проектированию структуры таблиц в

	настройки и администрирования хранилищ Hadoop.		Раздел (модуль) 2. Анализ больших наборов данных в Apache Hive	Apache Hive. Представление и защита результатов лабораторных работ по созданию запросов с использованием языка HiveQL.
РД 3	Уметь налаживать обмен данными между кластерами Hadoop и различными источниками и потребителями данных.	И.ОПК (У)-3.1 И.ОПК (У)-3.2	Раздел (модуль) 1. Хранение и обработка больших наборов данных в Apache Hive Раздел (модуль) 2. Анализ больших наборов данных в Apache Hive	Представление и защита результатов практических работ по проектированию структуры таблиц в Apache Hive. Представление и защита результатов лабораторных работ по созданию запросов с использованием языка HiveQL.
РД 4	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, анализ и интерпретацию полученных данных, в области больших данных.	И.ОПК (У)-3.1 И.ОПК (У)-3.2	Раздел (модуль) 1. Хранение и обработка больших наборов данных в Apache Hive Раздел (модуль) 2. Анализ больших наборов данных в Apache Hive	Представление и защита результатов практических работ по проектированию структуры таблиц в Apache Hive. Представление и защита результатов лабораторных работ по созданию запросов с использованием языка HiveQL.
РД 5	Уметь применять для обработки больших данных, находящихся в кластере Hadoop, инструментарий экосистемы Hadoop (Pig, Hive, Spark)	И.ОПК (У)-3.1 И.ОПК (У)-3.2	Раздел (модуль) 1. Хранение и обработка больших наборов данных в Apache Hive Раздел (модуль) 2. Анализ больших наборов данных в Apache Hive	Представление и защита результатов практических работ по проектированию структуры таблиц в Apache Hive. Представление и защита результатов лабораторных работ по созданию запросов с использованием языка HiveQL.

3. Шкала оценивания

Порядок организации оценивания результатов обучения в университете регламентируется отдельным локальным нормативным актом – «Система оценивания результатов обучения в Томском политехническом университете (Система оценивания)» (в действующей редакции). Используется балльно-рейтинговая система оценивания результатов обучения. Итоговая оценка (традиционная и литерная) по видам учебной деятельности (изучение дисциплин, курсовое проектирование) определяется суммой баллов по результатам текущего контроля и промежуточной аттестации (итоговая рейтинговая оценка - максимум 100 баллов).

Распределение основных и дополнительных баллов за оценочные мероприятия текущего контроля и промежуточной аттестации

устанавливается календарным рейтингом-планом дисциплины.

Рекомендуемая шкала для отдельных оценочных мероприятий входного и текущего контроля

% выполнения задания	Соответствие традиционной оценке	Определение оценки
90% – 100%	«Отлично»	Отличное понимание предмета, всесторонние знания, отличные умения и владение опытом практической деятельности, необходимые результаты обучения сформированы, их качество оценено количеством баллов, близким к максимальному.
70% – 89%	«Хорошо»	Достаточно полное понимание предмета, хорошие знания, умения и опыт практической деятельности, необходимые результаты обучения сформированы, качество ни одного из них не оценено минимальным количеством баллов.
55% – 69%	«Удовл.»	Приемлемое понимание предмета, удовлетворительные знания, умения и опыт практической деятельности, необходимые результаты обучения сформированы, качество некоторых из них оценено минимальным количеством баллов.
0% – 54%	«Неудовл.»	Результаты обучения не соответствуют минимально достаточным требованиям.

Шкала для оценочных мероприятий экзамена

% выполнения заданий экзамена	Экзамен, балл	Соответствие традиционной оценке	Определение оценки
90% – 100%	18 – 20	«Отлично»	Отличное понимание предмета, всесторонние знания, отличные умения и владение опытом практической деятельности, необходимые результаты обучения сформированы, их качество оценено количеством баллов, близким к максимальному.
70% – 89%	14 – 17	«Хорошо»	Достаточно полное понимание предмета, хорошие знания, умения и опыт практической деятельности, необходимые результаты обучения сформированы, качество ни одного из них не оценено минимальным количеством баллов.
55% – 69%	11 – 13	«Удовл.»	Приемлемое понимание предмета, удовлетворительные знания, умения и опыт практической деятельности, необходимые результаты обучения сформированы, качество некоторых из них оценено минимальным количеством баллов.
0% – 54%	0 – 10	«Неудовл.»	Результаты обучения не соответствуют минимально достаточным требованиям.

Перечень типовых заданий

	Оценочные мероприятия	Примеры типовых контрольных заданий
--	-----------------------	-------------------------------------

1	Защита лабораторной работы	Вопросы: 1. Создание запросов с использованием строковых функций в Apache Hive. 2. Создание запросов с использованием числовых функций в Apache Hive. 3. Создание запросов с использованием функций преобразования форматов в Apache Hive. 4. Создание запросов с использованием оконных функций в Apache Hive. 5. Создание запросов с использованием аналитических функций в Apache Hive.
2	Защита практической работы	Вопросы: 1. Способы загрузки, хранения и выгрузки данных в Apache Hive. 2. Проектирование и создание разделов (partitions) в таблице Apache Hive. 3. Проектирование и создание кластеров (buckets) в таблице Apache Hive. 4. Проектирование датасета и создание на его основе «внутренних» и «внешних» таблиц. 5. Проектирование сложной структуры таблицы и управление данными.
3	Экзамен	Вопросы на экзамен: 1. Современные подходы к анализу больших данных. 2. Этапы анализа данных согласно стандарту CRISP-DM. 3. Распределенная файловая система Hadoop (HDFS). 4. Особенности модели распределённых вычислений MapReduce. 5. Архитектура хранилища данных Apache Hive. 6. Типы и свойства таблиц в Apache Hive. 7. Применение «разделов» и «кластеров» в таблицах Apache Hive. 8. Применение языка HiveQL для статистического анализа данных. 9. Встроенные аналитические функции в HiveQL. 10. Особенности линейного регрессионного анализа данных.

Методические указания по процедуре оценивания

	Оценочные мероприятия	Процедура проведения оценочного мероприятия и необходимые методические указания
1	Защита лабораторной работы (max 5 баллов)	1. Защита лабораторной работы проводится на занятии, следующем после предыдущей лабораторной работы. 2. Для защиты лабораторной работы студент демонстрирует результаты программирования в соответствии с заданием и отвечает на вопросы, которые включают знание теоретических основ применяемых в работе способов создания запросов (язык HiveQL), используемых функций и т.п. 3. Лабораторная работа считается успешно выполненной при получении более 3 баллов. <u>Оценивание проводит преподаватель по следующим критериям:</u>

		Вид вопроса	Критерии оценки															
		Знание теории	2 б. – знает способы, функции и пр., может уверенно и без ошибок их обсуждать	1 б. – знает способы, функции и пр., может обсуждать их с помощью преподавателя	0,5 б. – затрудняется четко сформулировать способы, функции и пр.													
		Умение применить знания на практике	2 б. – код написан корректно и работает правильно, может продемонстрировать выполнение при изменении исходных данных	1 б. – код написан корректно и работает правильно, затрудняется продемонстрировать выполнение при изменении исходных данных	0,5 б.– код написан не оптимально, возможно некорректное срабатывание при вводе определённых данных													
		Своевременность сдачи работы 0,5–1 б.																
2	Защита практической работы (max 5 баллов)	1. Защита практической работы проводится на занятии, следующем после предыдущего практического занятия. 2. Для защиты практической работы студент демонстрирует результаты проектирования датасета и созданной на его основе таблицы, в соответствии с заданием. Также студент должен ответить на вопросы, которые включают знание способов хранения и обработки данных в Apache HIVE. 3. Практическая работа считается успешно выполненной при получении более 3 баллов. <u>Оценивание проводит преподаватель по следующим критериям:</u> <table><tr><td>Вид вопроса</td><td colspan="3">Критерии оценки</td></tr><tr><td>Знание теории</td><td>2 б. – знает методы, параметры, может уверенно и без ошибок их обсуждать</td><td>1 б. – знает методы, параметры, может обсуждать их с помощью преподавателя</td><td>0,5 б. – затрудняется четко сформулировать методы и параметры</td></tr><tr><td>Умение применить знания на практике</td><td>2 б. – код написан корректно и работает правильно, может продемонстрировать выполнение при изменении исходных данных</td><td>1 б. – код написан корректно и работает правильно, затрудняется продемонстрировать выполнение при изменении исходных данных</td><td>0,5 б.– код написан не оптимально, возможно некорректное срабатывание при вводе определённых данных</td></tr></table> Своевременность сдачи работы 0,5–1 б.					Вид вопроса	Критерии оценки			Знание теории	2 б. – знает методы, параметры, может уверенно и без ошибок их обсуждать	1 б. – знает методы, параметры, может обсуждать их с помощью преподавателя	0,5 б. – затрудняется четко сформулировать методы и параметры	Умение применить знания на практике	2 б. – код написан корректно и работает правильно, может продемонстрировать выполнение при изменении исходных данных	1 б. – код написан корректно и работает правильно, затрудняется продемонстрировать выполнение при изменении исходных данных	0,5 б.– код написан не оптимально, возможно некорректное срабатывание при вводе определённых данных
Вид вопроса	Критерии оценки																	
Знание теории	2 б. – знает методы, параметры, может уверенно и без ошибок их обсуждать	1 б. – знает методы, параметры, может обсуждать их с помощью преподавателя	0,5 б. – затрудняется четко сформулировать методы и параметры															
Умение применить знания на практике	2 б. – код написан корректно и работает правильно, может продемонстрировать выполнение при изменении исходных данных	1 б. – код написан корректно и работает правильно, затрудняется продемонстрировать выполнение при изменении исходных данных	0,5 б.– код написан не оптимально, возможно некорректное срабатывание при вводе определённых данных															