

УТВЕРЖДАЮ  
Директор ВШТЭ



## Рабочая программа дисциплины

**Б1.В.ДВ.01.02** Использование больших данных в информационных системах

Учебный план: \_\_\_\_\_ ФГОС3++b090303-1\_23-14.plx

Кафедра:  Информационно-измерительных технологий и систем управления

Направление подготовки:  
(специальность) 09.03.03 Прикладная информатика

Профиль подготовки:  
(специализация) Искусственный интеллект в информационных системах

Уровень образования: бакалавриат

Форма обучения: очная

### План учебного процесса

Семестр (курс для ЗАО)	Контактная работа обучающихся		Сам. работа	Контроль, час.	Трудоё мкость, ЗЕТ	Форма промежуточной аттестации
	Лекции	Практ. занятия				
4	УП	17	17	37,75	0,25	Зачет
	РПД	17	17	37,75	0,25	
Итого	УП	17	17	37,75	0,25	
	РПД	17	17	37,75	0,25	

Рабочая программа дисциплины составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 09.03.03 Прикладная информатика, утверждённым приказом Минобрнауки России от 19.09.2017 г. № 922

Составитель (и):

старший преподаватель

Слюта М.О.

От кафедры составителя:

Заведующий кафедрой информационно-измерительных технологий и систем управления

Сидельников В.И.

От выпускающей кафедры:

Заведующий кафедрой

Сидельников В.И.

Методический отдел:

Смирнова В.Г.

## 1 ВВЕДЕНИЕ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

**1.1 Цель дисциплины:** 1. Изучение математических методов и подходов, используемых в программных системах обработки и анализа больших данных.

2. Развитие профессиональных навыков обучающихся за счет получения практического опыта работы с IT решениями в части обработки и анализа больших данных.

### 1.2 Задачи дисциплины:

усвоение знаний о сущности, структуре и видах математических моделей принятия решений; развитие навыков создания и решения моделей, необходимых в сфере информационных технологий.

### 1.3 Требования к предварительной подготовке обучающегося:

Предварительная подготовка предполагает создание основы для формирования компетенций, указанных в п. 2, при изучении дисциплин:

Программные средства обработки информации

## 2 КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

<b>ПК-1: Способен определять первоначальные требования заказчика к ИС и возможности их реализации в ИС на этапе предконтрактных работ</b>
---

<b>Знать:</b> имеет представление об использовании типовой информационной системы с учетом обработки больших объемов данных.
--

<b>Уметь:</b> формулирует порядок использования программных средств для обработки больших объемов данных.
---

<b>Владеть:</b> демонстрирует навыки модификации информационной системы.
--

### 3 РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Наименование и содержание разделов, тем и учебных занятий	Семестр (курс для ЗАО)	Контактная работа		СР (часы)	Инновац. формы занятий	Форма текущего контроля
		Лек. (часы)	Пр. (часы)			
Раздел 1. Технологии хранения и обработки Больших данных	4					Ко
Тема 1. Введение в Большие данные. Основные направления развития методов обработки и хранения данных. Проблема хранения неструктурированных данных. Проблема преобразования данных.		6	6	10		
Тема 2. Программирование обработки и загрузки Больших данных. Языки для Big Data. Фреймворки. Базы данных. Аналитические платформы.		4	4	10	ИЛ	
Раздел 2. Аналитика в больших данных						Ко
Тема 3. Аналитика Big Data — реалии и перспективы в России и мире. Технологии и методы анализа, которые используются для анализа Big Data.		4	4	10		
Тема 4. Применение и возможности Big data. Решения на основе Big data.		3	3	7,75	ИЛ	
Итого в семестре (на курсе для ЗАО)			17	17	37,75	
Консультации и промежуточная аттестация (Зачет)		0,25				
<b>Всего контактная работа и СР по дисциплине</b>		34,25		37,75		

### 4 КУРСОВОЕ ПРОЕКТИРОВАНИЕ

Курсовое проектирование учебным планом не предусмотрено

### 5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

#### 5.1 Описание показателей, критериев и системы оценивания результатов обучения

##### 5.1.1 Показатели оценивания

Код компетенции	Показатели оценивания результатов обучения	Наименование оценочного средства
ПК-1	1. Имеет представление о методах анализа и хранения больших объемов данных, этапы жизненного цикла обработки больших данных, языки, наиболее приспособленные для обработки и аналитики больших данных, способы организации хранения и доступа к большим данным. 2. Способен выполнять анализ данных и интерпретировать	1. Вопросы устного собеседования. 2. Практико-ориентированные задания.

	результаты, выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. 3. Использует математические методы анализа данных, компьютерные методы обработки данных.	
--	---	--

### 5.1.2 Система и критерии оценивания

Шкала оценивания	Критерии оценивания сформированности компетенций	
	Устное собеседование	Письменная работа
Зачтено	Студент показывает знание фактического материала по программе, в том числе: знание обязательной литературы, современных публикаций по программе курса; студент положительно сдал коллоквиумы; учитываются логика, структура, стиль ответа; культура речи, манера общения; готовность к дискуссии, аргументированность ответа; уровень самостоятельного мышления; умение приложить теорию к практике, решить предложенное практическое задание; отсутствие пропусков занятий по неуважительным причинам	
Не зачтено	Отсутствие знания пройденного материала, плохое знание обязательной литературы; отрицательный результат по прохождению коллоквиумов; студент допускает существенные ошибки при ответе на вопросы преподавателя; невозможность приложить теорию к практике, решить предложенное практическое задание; наличие неуважительных пропусков занятий.	

### 5.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности

#### 5.2.1 Перечень контрольных вопросов

№ п/п	Формулировки вопросов
Семестр 4	
1	Определение больших данных, ключевые характеристики.
2	Примеры задач больших данных. Основные виды данных.
3	Дать краткую сравнительную характеристику инструментария ПО для анализа данных.
4	«Жизненный цикл» проекта по аналитике больших данных
5	Типовая архитектура проекта в области больших данных.
6	Машинное обучение.
7	Искусственные нейронные сети.
8	Распознавание образов.
9	Прогнозная аналитика.
10	Имитационное моделирование.
11	Факторный анализ.
12	Дискриминантный анализ.
13	Big data в банках.
14	Big data в бизнесе.
15	Big data в маркетинге.
16	Проблема преобразования данных.

## 5.2.2 Типовые тестовые задания

Не предусмотрено.

## 5.2.3 Типовые практико-ориентированные задания (задачи, кейсы)

1. Дан набор данных заданной структуры и программа SAS Data step, производящая определенную обработку и вычисления с использованием данного набора. Перепишите эту программу на SAS DS2 с использованием параллельных нитей и созданием пользовательского пакета, чтобы результат обработки сохранился тем же, но код мог выполняться в параллельной среде.

2. Дан набор заданной структуры, постройте модель прогнозирования отклика с использованием процедуры `impstat` с алгоритмом `random forest` с заданным числом деревьев. Примените полученную модель к тестовому набору данных той же структуры, визуализируйте полученный график Lift. Постройте на том же наборе модель с использованием высокопроизводительной версии метода GLM. Примените к тестовому набору. Сравните результаты GLM и Random Forest по AUC.

## 5.3 Методические материалы, определяющие процедуры оценивания знаний, умений, владений (навыков и (или) практического опыта деятельности)

### 5.3.1 Условия допуска обучающегося к промежуточной аттестации и порядок ликвидации академической задолженности

Проведение промежуточной аттестации регламентировано локальным нормативным актом СПбГУПТД «Положение о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся»

### 5.3.2 Форма проведения промежуточной аттестации по дисциплине

Устная  +  Письменная  Компьютерное тестирование  Иная

### 5.3.3 Особенности проведения промежуточной аттестации по дисциплине

При проведении зачета время, отводимое на подготовку к ответу, составляет не более 20 мин.

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### 6.1 Учебная литература

Автор	Заглавие	Издательство	Год издания	Ссылка
<b>6.1.1 Основная учебная литература</b>				
Николаев, Е. И.	Базы данных в высокопроизводительных информационных системах	Ставрополь: Северо-Кавказский федеральный университет	2016	<a href="http://www.iprbookshop.ru/69375.html">http://www.iprbookshop.ru/69375.html</a>
Молдованова, О. В.	Информационные системы и базы данных	Саратов: Профобразование	2021	<a href="http://www.iprbookshop.ru/106617.html">http://www.iprbookshop.ru/106617.html</a>
<b>6.1.2 Дополнительная учебная литература</b>				
Воронова, Л. И., Воронов, В. И.	Big Data. Методы и средства анализа	Москва: Московский технический университет связи и информатики	2016	<a href="http://www.iprbookshop.ru/61463.html">http://www.iprbookshop.ru/61463.html</a>
Танцов, П. Н.	Интеллектуальные информационные системы	Москва: Издательский Дом МИСиС	2015	<a href="http://www.iprbookshop.ru/98854.html">http://www.iprbookshop.ru/98854.html</a>

### 6.2 Перечень профессиональных баз данных и информационно-справочных систем

Электронно-библиотечная система IPRbooks [Электронный ресурс]. URL: <http://www.iprbookshop.ru/>

Электронная библиотека ВШТЭ СПб ГУПТД [Электронный ресурс]. URL: <http://nizrp.narod.ru>

Электронно-библиотечная система «Айбукс» [Электронный ресурс]. URL: <https://www.ibooks.ru/>

### 6.3 Перечень лицензионного и свободно распространяемого программного обеспечения

Microsoft Windows 8

Delphi

**6.4 Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине**

Аудитория	Оснащение
Лекционная аудитория	Мультимедийное оборудование, специализированная мебель, доска
Компьютерный класс	Мультимедийное оборудование, компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду