

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Бойко Елена Григорьевна
Должность: Ректор
Дата подписания: 29.02.2024 11:01:42
Уникальный программный ключ:
e69eb689122030af7d22cc354bf0eb9d453ecf8f

Министерство сельского хозяйства РФ
ФГБОУ ВО Государственный аграрный университет Северного Зауралья
Инженерно-технологический институт
Кафедра математики и информатики

«Утверждаю»
И.о. заведующего кафедрой

 М.В. Виноградова

«01» июня 2021 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Введение в анализ данных

для направления подготовки 35.03.06 Агроинженерия

профили: *Технические системы в агробизнесе, Технический сервис в агропромышленном комплексе*

Уровень высшего образования – бакалавриат

Форма обучения: очная, заочная

Тюмень, 2021


При разработке рабочей программы учебной дисциплины в основу положены:

- 1) ФГОС ВО по направлению подготовки (специальности) 35.03.06 «Агроинженерия» утвержденный Министерством образования и науки РФ 23 августа 2017 г., приказ № 813.
- 2) Учебные планы основных образовательных программ профилей «Технический сервис в агропромышленном комплексе», «Технические системы в агробизнесе» одобрены Ученым советом ФГБОУ ВО ГАУ Северного Зауралья от 27 мая 2021 г. Протокол № 11.

Рабочая программа учебной дисциплины (модуля) одобрена на заседании кафедры математики и информатики от 1 июня 2021 г. Протокол № 10.1.

И.о. заведующего кафедрой _____  М.В. Виноградова

Рабочая программа учебной дисциплины (модуля) одобрена методической комиссией института от 8 июня 2021 г. Протокол № 7а.

Председатель методической комиссии института _____  О.А. Мелякова

Разработчик:

Ерёмина Д.В., к.с.-х.н., доцент кафедры математики и информатики

Директор института: _____  Г.А. Дорн

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

Код компетенции	Результаты освоения	Индикатор достижения компетенции	Перечень планируемых результатов обучения по дисциплине
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	ИД-2 _{УК-1} Осуществляет анализ и оценку информации, с использованием различных информационных ресурсов для решения поставленных задач	<p><i>знать:</i></p> <ul style="list-style-type: none"> - теоретические основы и методы решения задач анализа данных; иметь представление об основных тенденциях развития теории и практики данных и методах работы с ними; <p><i>уметь:</i></p> <ul style="list-style-type: none"> - решать типовые задачи в области анализа данных, применять соответствующие методы и знания в профессиональной деятельности; <p><i>владеть:</i></p> <ul style="list-style-type: none"> - методикой анализа данных для решения типовых задач в области профессиональной деятельности.

2. Место дисциплины в структуре образовательной программы

Данная дисциплина относится к *Блоку 1* части, формируемой участниками образовательных отношений.

Для изучения дисциплины необходимы знания в области: математики, информатики и цифровых технологий.

Введение в анализ данных является предшествующей дисциплиной для дисциплин: *Алгоритмы обработки и анализа данных, Машинное обучение и аналитика big data для бизнеса.*

Дисциплина изучается на 4 курсе в 7 семестре по очной форме обучения, на 4 курсе в 8 семестре по заочной форме.

3. Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины составляет 108 часов (3 зачетных единицы).

Вид учебной работы	Форма обучения	
	очная	заочная
Аудиторные занятия (всего)	48	12
<i>В том числе:</i>	-	-
Лекционного типа	24	6
Семинарского типа	24	6
Самостоятельная работа (всего)	60	96
<i>В том числе:</i>	-	-
Проработка материала лекций, подготовка к занятиям	30	72
Самостоятельное изучение тем	6	
Курсовой проект (работа)	-	-
Расчетно-графические работы	-	-
Контрольная работа	-	24
Реферат	-	-
Индивидуальные задания	24	-
Вид промежуточной аттестации:	зачет	зачет
Общая трудоемкость:		
часов	108	108
зачетных единиц	3	3

4. Содержание дисциплины

4.1. Содержание разделов дисциплины

№ п/п	Наименование раздела дисциплины	Содержание раздела
1.	Классификация данных	Данные, подходы и определения, жизненный цикл данных. Понятие метаданных, жизненный цикл метаданных. Большие данные, системы управления Большими данными. Процесс аналитики.
2.	Введение в Data Mining	Введение в когнитивный анализ данных. Классификация задач. Функция конкурентного сходства. Разработка алгоритмов на базе FRIS-функции. Информативность и выбор признаков. Обнаружение ошибок и заполнение пробелов.
3	Google таблицы для анализа данных	Первичная обработка данных. Сводные таблицы и диаграммы. Формулы для анализа данных (готовые формулы статистики; текстовые сложные формулы Lookup, Vlookup; формулы условия IF, ссылки и массивы (ВПР, ГПР). Макросы. Точечное оценивание параметров. Регрессионный и корреляционный анализ.
4.	Основы языка Python	Знакомство с Python. Встроенные типы и операции с ними. Функции. Полезные инструменты. Работа с файлами. Объектно-ориентированное программирование. Библиотеки Python для Data Science.

4.2. Разделы дисциплины и виды занятий

очная форма обучения

№ п/п	Наименование раздела дисциплины	Лекционного типа	Семинарского типа	СР	Всего, часов
1	2	3	4	5	6
1.	Классификация данных	8	-	10	18
2.	Введение в Data Mining	8	-	13	21
3.	Google таблицы для анализа данных	2	12	5	19
4.	Основы языка Python	6	12	32	50
	Итого:	24	24	60	108

заочная форма обучения

№ п/п	Наименование раздела дисциплины	Лекционного типа	Семинарского типа	СР	Всего, часов
1	2	3	4	5	6
1.	Классификация данных	2	-	24	26
2.	Введение в Data Mining	2	-	24	26
3.	Google таблицы для анализа данных	-	6	24	30
4.	Основы языка Python	2	-	24	26
	Итого:	6	6	96	108

4.3. Занятия семинарского типа

№ п/п	№ раздела дисциплины	Тема	Трудоемкость (час)	
			очная	заочная
1	2	3	4	5
1.	3	Google таблицы. Загрузка и первичная обработка данных.	2	2
2.	3	Google таблицы. Сводные таблицы и диаграммы.	2	2
3.	3	Google таблицы. Формулы для анализа данных.	2	2
4.	3	Google таблицы. Макросы.	2	-
5.	3	Google таблицы. Точечное оценивание параметров.	2	-
6.	4	Google таблицы. Регрессионный и корреляционный анализ.	2	-
7.	4	Python. Знакомство с интерпретатором.	2	-
8.	4	Python. Встроенные структуры данных, функции и файлы.	2	-
9.	4	Python. Основы работы с Pandas.	2	-
10.	4	Python. Чтение и запись данных, форматы файлов.	2	-
11.	4	Python. Построение графиков и визуализация.	2	-
12.	4	Python. Примеры анализа данных.	2	-
		Итого:	24	6

4.4. Примерная тематика курсовых проектов (работ) - Не предусмотрено ОПОП

5. Организация самостоятельной работы обучающихся по дисциплине

5.1. Типы самостоятельной работы и её контроль

Тип самостоятельной работы	Форма обучения		Текущий контроль
	очная	заочная	
Проработка материала лекций, подготовка к занятиям	30	- 72	Тестирование
Самостоятельное изучение тем	6		Тестирование
Контрольная работа	-	24	Защита контрольной работы
Индивидуальное задание	24	-	Защита индивидуального задания
всего часов:	60	96	

5.2. Учебно-методические материалы для самостоятельной работы:

1. Введение в анализ данных [Электронный ресурс]: Методические указания к выполнению самостоятельной работы для направления подготовки 35.03.06 Агроинженерия / Автор-сост. Д.В. Ерёмина. - Тюмень: ГАУ Северного Зауралья, 2020. – 12 с.

2. Введение в анализ данных [Электронный ресурс]: Методические указания к выполнению контрольной работы обучающимися заочной формы по направлению подготовки 35.03.06 Агроинженерия / Автор-сост. Д.В. Ерёмина. – Тюмень: ГАУ Северного Зауралья, 2020. – 12 с.

3. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу: учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва: Московский городской педагогический университет, 2012. — 204 с. — ISBN 2227-8397. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/26444.html>. — Режим доступа: для авторизир. пользователей.

4. Буйначев, С. К. Основы программирования на языке Python: учебное пособие / С. К. Буйначев, Н. Ю. Боклаг; под редакцией Ю. В. Песин. — Екатеринбург: Уральский федеральный университет, ЭБС АСВ, 2014. — 92 с. — ISBN 978-5-7996-1198-9. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/66183.html>. — Режим доступа: для авторизир. пользователей.

5.3. Темы, выносимые на самостоятельное изучение:

Раздел 2. Введение в Data Mining.

1. Технологии анализа данных.

1.1. Аналитический и информационный подходы к моделированию.

1.2. Технология KDD (Knowledge Discovery in Database).

1.3. Классы задач, решаемые методами Data Mining.

Раздел 4. Основы языка Python.

1. Встроенные типы данных.

1.1. Тип int и long.

1.2. Тип float.

1.3. Тип bool и tuple.

1.4. Тип string и тип Unicode.

1.5. Тип list. Последовательности.

1.6. Тип dict, тип file.

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

6.1 Перечень компетенций и оценочные средства индикатора достижения компетенций

Код компетенции	Индикатор достижения компетенции	Перечень планируемых результатов обучения по дисциплине	Наименование оценочного средства
УК-1	ИД-3 _{УК-1} Осуществляет анализ и оценку информации, с использованием различных информационных ресурсов для решения поставленных задач	<i>знать:</i> - теоретические основы и методы решения задач анализа данных; иметь представление об основных тенденциях развития теории и практики данных и методах работы с ними; <i>уметь:</i> - решать типовые задачи в области анализа данных, применять соответствующие методы и знания в профессиональной деятельности; <i>владеть:</i> - методикой анализа данных для решения типовых задач в области профессиональной деятельности.	Тест

6.2. Шкалы оценивания

Шкала оценивания тестирования на зачете

% выполнения задания	Результат
50 – 100	зачтено
менее 50	не зачтено

6.4. Типовые контрольные задания или иные материалы:

Указаны в приложении 1.

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

а) основная литература

1. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу: учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва: Московский городской педагогический университет, 2012. — 204 с. — ISBN 2227-8397. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/26444.html>. — Режим доступа: для авторизир. пользователей.

2. Федин, Ф. О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва: Московский городской педагогический университет, 2012. — 308 с. — ISBN 2227-8397. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/26445.html>. — Режим доступа: для авторизир. пользователей.

3. Буйначев, С. К. Основы программирования на языке Python: учебное пособие / С. К. Буйначев, Н. Ю. Боклаг; под редакцией Ю. В. Песин. — Екатеринбург: Уральский федеральный университет, ЭБС АСВ, 2014. — 92 с. — ISBN 978-5-7996-1198-9. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/66183.html>. — Режим доступа: для авторизир. пользователей.

б) дополнительная литература

4. Маккинли, Уэс Python и анализ данных / Уэс Маккинли; перевод А. Слинкина. — 2-е изд. — Саратов: Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/88752.html>. — Режим доступа: для авторизир. пользователей.

5. Сузи, Р. А. Язык программирования Python: учебное пособие / Р. А. Сузи. — 3-е изд. — Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 350 с. — ISBN 978-5-4497-0705-5. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/97589.html>. — Режим доступа: для авторизир. пользователей

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

<http://www.intuit.ru/> - Национальный открытый университет «Интуит».

<http://www.planetaexcel.ru> – сайт о возможностях Excel.

<http://office.microsoft.com/ru-ru> - сайт фирмы Microsoft.

9. Методические указания для обучающихся по освоению дисциплины

Введение в анализ данных [Электронный ресурс]: Методические указания к практическим занятиям для направления подготовки 35.03.06 Агроинженерия / Автор-сост. Д.В. Ерёмкина. – Тюмень: ГАУ Северного Зауралья, 2020. - 48 с.

10. Перечень информационных технологий

1. Microsoft Windows 8 Professional; OpenSUSE Linux;
2. Система электронного обучения Moodle;
3. Сервисы Google Suite for Education;
4. Python 3.9.5.

11. Материально-техническое обеспечение дисциплины

Для проведения практических занятий по дисциплине «Введение в анализ данных» используется компьютерный класс (14-15 компьютеров) с установленным программным обеспечением.

Лекционные занятия проводятся в учебных аудиториях ФГБОУ ВО ГАУ Северного Зауралья с мультимедийным оборудованием (проектор и/или интерактивная доска).

Для выполнения самостоятельной работы обучающиеся могут пользоваться читальными залами библиотеки ФГБОУ ВО ГАУ Северного Зауралья, в том числе оснащёнными компьютерами с локальной сетью и выходом в интернет.

12. Особенности освоения дисциплины для инвалидов и лиц с ограниченными возможностями здоровья

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

В целях освоения учебной программы дисциплины инвалидами и лицами с ограниченными возможностями здоровья обеспечивается:

- для инвалидов и лиц с ограниченными возможностями здоровья по зрению: размещение в доступных для обучающихся, являющихся слепыми или слабовидящими, местах и в адаптированной форме справочной информации о расписании учебных занятий; присутствие ассистента, оказывающего обучающемуся необходимую помощь; выпуск альтернативных форматов методических материалов (крупный шрифт или аудиофайлы), использование версии сайта для слабовидящих ЭБС IPR BOOKS и специального мобильного приложения IPR BOOKS WV-Reader (программы не визуального доступа к информации, предназначенной для мобильных устройств, работающих на операционной системе Android и iOS, которая не требует специально обученного ассистента, т.к. люди с ОВЗ по зрению работают со своим устройством привычным способом, используя специальные штатные программы для незрячих людей, с которыми IPR BOOKS WV-Reader имеет полную совместимость);

- для инвалидов и лиц с ограниченными возможностями здоровья по слуху: надлежащими звуковыми средствами воспроизведение информации;

- для инвалидов и лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата: возможность беспрепятственного доступа обучающихся в учебные помещения, туалетные комнаты и другие помещения кафедры, а также пребывание в указанных помещениях.

Образование обучающихся с ограниченными возможностями здоровья может быть организовано как совместно с другими обучающимися, так и в отдельных группах или в отдельных организациях.

Министерство сельского хозяйства Российской Федерации
ФГБОУ ВО Государственный аграрный университет Северного Зауралья
Инженерно-технологический институт
Кафедра математики и информатики

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по учебной дисциплине «Введение в анализ данных»

для направления подготовки 35.03.06 Агроинженерия

профили: *Технические системы в агробизнесе, Технический сервис в агропромышленном комплексе*

Уровень высшего образования – бакалавриат

Разработчик:

Ерёмина Д.В., к.с.-х.н., доцент кафедры математики и информатики

Утверждено на заседании кафедры

протокол № 10.1 от 1 июня 2021 г.

И.о. заведующего кафедрой  М.В. Виноградова

Тюмень, 2021

**КОНТРОЛЬНЫЕ ЗАДАНИЯ И ИНЫЕ МАТЕРИАЛЫ ОЦЕНКИ
знаний, умений, навыков и (или) опыта деятельности, характеризующие
этапы формирования компетенций в процессе освоения дисциплины
творческих заданий / проектов
*ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ***

1. Вопросы к зачёту

Компетенция	Вопросы
<p>УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</p>	<ol style="list-style-type: none"> 1. Определение «данные». ГОСТы с определениями данных. 2. Определения из ФЗ-149. 3. Жизненный цикл данных. 4. Этап «Синтез данных». 5. Этап «Использование данных». 6. Этап «Публикация данных». 7. Этап «Архивация данных». 8. Метаданные. ГОСТы для метаданных. 9. Определение «Онтология». 10. Жизненный цикл метаданных. 11. Этап «Оценка требований и анализ контента». 12. Этап «Спецификация системных требований». 13. Этап «Система метаданных». 14. Этап «Сервис и оценка». 15. Определение «Большие данные». 16. Пять характеристик, присущих Большим данным. 17. Базовые принципы обработки Больших данных. 18. Распределённые файловые системы и распределённые фреймворки. 20. Бенчмаркинг. 21. Типы баз данных: Column databases (столбцовые). 22. Типы баз данных: Document stores (хранилища документов). 23. Типы баз данных: Streaming data (поточковые данные). 24. Типы баз данных: Key-value stores (хранилища для ключей). 25. Типы баз данных: SQL на Hadoop. 26. Типы баз данных: Новый SQL (New SQL). 27. Типы баз данных: Graph databases (графовые). 28. Определение «Аналитика Больших данных». 29. Основные возможности аналитики Больших данных. 30. Определение «когнитивный анализ». 31. Методика когнитивного анализа ситуаций. 32. Аналитический подход к моделированию. 33. Информационный подход к моделированию. 34. Технология KDD. 35. Классификация задач Data Mining. 36. Функция конкурентного сходства. 37. Интерфейс google таблиц. 38. Настройки доступа к google таблицам. 39. Сводные таблицы и диаграммы. 40. Применение статистических формул для анализа данных. 41. Текстовые сложные формулы Lookup, Vlookup. 42. Формулы условия IF.

43.	Ссылки и массивы (ВПР, ГПР).
44.	Макросы.
45.	Точечное оценивание параметров.
46.	Регрессионный и корреляционный анализ средствами google таблиц.
47.	Плюсы и минусы интерпретатора Python.
48.	Установка и обновление Python-пакетов.
49.	Структуры данных и последовательности.
50.	Встроенные типы данных.
51.	Функции в Python.
52.	Объектно-ориентированное программирование.
53.	Программирование с применением массивов.
54.	Библиотеки Python для Data Science.
55.	Структуры данных pandas.
56.	Построение графиков и визуализация.

Обучающийся допускается к промежуточной аттестации по дисциплине при условии выполнения всех практических работ, индивидуальных заданий (очная форма), контрольной работы (заочная форма), выполнения на положительные оценки тестов по лекционному материалу и темам, вынесенным на самостоятельное обучение.

Тестовое задание в системе электронного обучения Moodle включает 30 вопросов, в случайном порядке выбранных из банка вопросов. Обучающемуся предоставляется 2 попытки, по 45 минут каждая.

Критерии оценки:

- «зачтено» выставляется обучающемуся, если итоговое тестирование в системе электронного обучения Moodle выполнено с результатом 50% и выше;
- «не зачтено» выставляется обучающемуся, если итоговое тестирование в системе электронного обучения Moodle выполнено с результатом 49% и ниже.

2. Варианты заданий для контрольной работы

Вариант № 1

1. Роль анализа данных в современном мире.
2. Базовые принципы обработки Больших данных.
3. Технология KDD.
4. Интерфейс google таблиц.
5. Python: структуры данных и последовательности.
6. Практическое задание: найдите в сети Интернет два сайта, на которых используются системы прогнозирования. Опишите, добавьте ссылки на сайты.

Вариант №2

1. Научные исследования в сфере анализа данных.
2. Распределенные файловые системы.
3. Основные возможности аналитики Больших данных.
4. Регрессионный анализ данных средствами google таблиц.
5. Плюсы и минусы интерпретатора Python.
6. Практическое задание: найдите в сети Интернет два сайта, на которых используются рекомендательные системы. Опишите, добавьте ссылки на сайты.

Вариант №3

1. Примеры применения методов анализа данных.
2. Распределенные фреймворки.
3. Методика когнитивного анализа ситуаций.

4. Точечное оценивание параметров.
5. Установка и обновление Python-пакетов.
6. Практическое задание: пользуясь системами SCOPUS, Web of Science, E-library (РИНЦ), выявите нескольких ведущих ученых в сфере анализа данных.

Вариант №4

1. Алгоритм построения системы анализа данных.
2. Бенчмаркинг.
3. Аналитический подход к моделированию.
4. Google таблицы: ссылки и массивы (ВПР, ГПР).
5. Python: встроенные типы данных.
6. Практическое задание: пользуясь одной из перечисленных систем SCOPUS, Web of Science, E-library (РИНЦ), проанализируйте динамику количества публикаций за пять лет по направлению Deep Learning.

Вариант №5

1. Определение «данные». ГОСТы с определениями данных.
2. Графовые базы данных (Graph databases).
3. Информационный подход к моделированию.
4. Корреляционный анализ данных средствами google таблиц.
5. Функции в Python.
6. Практическое задание: пользуясь одной из перечисленных систем SCOPUS, Web of Science, E-library (РИНЦ), проанализируйте динамику количества публикаций за пять лет по направлению Big Data.

Вариант №6

1. Данные. Определения из ФЗ-149.
2. Столбцовые базы данных (Column databases).
3. Классификация задач Data Mining.
4. Применение макросов в google таблицах.
5. Объектно-ориентированное программирование.
6. Практическое задание: пользуясь одной из перечисленных систем SCOPUS, Web of Science, E-library (РИНЦ), найдите пять публикаций с наибольшей цитируемостью за последние десять лет по направлению Big Data.

Вариант №7

1. Жизненный цикл данных. Основные этапы.
2. Хранилища документов (Document stores).
3. Функция конкурентного схождения.
4. Формулы условия IF в google таблицах.
5. Библиотеки Python для Data Science.
6. Практическое задание: пользуясь одной из перечисленных систем SCOPUS, Web of Science, E-library (РИНЦ), найдите пять публикаций с наибольшей цитируемостью за последние десять лет по направлению Social Network Analysis.

Вариант №8

1. Метаданные. ГОСТы для метаданных.
2. Поточковые данные (Streaming data).
3. Этапы решения задачи анализа данных и их взаимосвязи.
4. Применение статистических формул для анализа данных в google таблицах.
5. Python: программирование с применением массивов.
6. Практическое задание: пользуясь одной из перечисленных систем SCOPUS, Web of Science, E-library (РИНЦ), найдите пять публикаций с наибольшей цитируемостью за последние десять лет по направлению Deep Learning.

Вариант №9

1. Жизненный цикл метаданных. Основные этапы.
2. Хранилища для ключей (Key-value stores).

3. Качественный и количественный этапы анализа данных.
4. Сводные таблицы и диаграммы в google таблицах.
5. Структуры данных pandas.
6. Практическое задание: пользуясь научной-электронной библиотекой E-library (РИНЦ), найдите пять актуальных публикаций по анализу данных.

Вариант №10

1. Большие данные. Пять характеристик, присущих большим данным.
2. SQL на Hadoop.
3. Методики обучения моделей (с учителем, без учителя).
4. Настройки доступа к google таблицам.
5. Python: построение графиков и визуализация.
6. Практическое задание: пользуясь научной-электронной библиотекой E-library (РИНЦ), найдите публикации по теме «Big Data: перспективы развития в России».

Контрольная работа выполняется обучающимися заочной формы. Сдается на проверку преподавателю в период сессии.

При оценке контрольной работы определяются выдержанность требований к оформлению контрольной работы, соответствие содержания заданному варианту, полнота и последовательность изложения, наличие достаточных пояснений, выполнение практического задания, число и характер ошибок (существенные или несущественные), а также ответы на уточняющие вопросы преподавателя по содержанию контрольной.

Критерии оценки:

- оценка «зачтено» выставляется обучающемуся, если контрольная работа выполнена по своему варианту, допущено по каждому заданию по одной несущественной ошибке и в одном задании допущена одна существенная ошибка, приведены схемы, таблицы и рисунки, требующие эти пояснения по работе.
- оценка «не зачтено» выставляется обучающемуся, если контрольная работа выполнена не по своему варианту, допущено по половине заданий по одной существенной ошибке, не приведены схемы, таблицы и рисунки, требующие эти пояснения по работе.

3. Индивидуальные творческие задания (проекты):

Индивидуальное задание «Python: работа с графикой»

Разработайте приложение, которое строит график одной из указанных ниже функций. Построение производится в симметричной системе координат (область рисования делится на четыре равных квадрата), маркировка осей производится в зависимости от масштабирования X и Y. График должен полностью помещаться в области рисования. Пользователю можно изменить цвет фона и цвет графика. Коэффициенты и граничные значения x вводятся с клавиатуры.

Программа должна предусматривать возможность задания значений параметров, диапазона значений по оси OX, при этом диапазон значений по оси OY должен определяться автоматически так, чтобы график полностью помещался в область рисования, также у пользователя должна быть возможность выбора цвета графика, варианта рисования графика: точками или линией (с выбором толщины линии), количества точек, на основании которых должен быть построен график функции (считается, что точки распределяются равномерно по оси OX). Подсчет значений функции должен быть реализован в виде подпрограммы.

Пояснения. При построении графика необходимо учитывать масштаб, заданный пользователем при построении системы координат. При построении графиков функций на

экране монитора необходимо преобразовывать расчетные координаты в графические с учетом дискретности растровой сетки монитора.

Для этого желательно создать процедуры, обеспечивающие универсальность при выводе графических изображений. Ниже приводится алгоритм построения графика функции $Y=F(X)$ в заданной области экрана с возможностью автоматического масштабирования.

Пусть задана непрерывная функция $F(X)$ в диапазоне изменения аргумента $X=[A..B]$. Требуется построить по N точкам график функции $Y=F(X)$ в прямоугольной области экрана left, up, right, down.

1. Определяем массивы значений аргумента и функции.
2. Определяем наибольшее Y_{max} и наименьшее Y_{min} значения функции в заданном интервале изменения аргумента. Эти значения необходимо определить для полного размещения графика в расчетной области.
3. Строим систему координат, предварительно задав начало координат X_0, Y_0 .
4. Определяем коэффициент масштабирования при построении графика в заданной области экрана. Данный коэффициент используется для масштабирования по обеим осям, так как при введении разных масштабов по осям, возможно искажение естественной формы кривой (растяжение или сжатие по одной из осей).
5. Определяем координаты точек для построения графика в системе координат экрана. При этом учитываем необходимость смещения координат точек по соответствующим осям относительно границ области left и down, а также необходимость "переворота" оси Y , которая в координатах монитора направлена сверху вниз.
6. Строим график в виде последовательных отрезков, соединяющих две соседние точки (первый отрезок соединяет 1 и 2 точки, второй отрезок соединяет 2 и 3 точки и т.д., в результате чего получается непрерывная линия).

Варианты к индивидуальному заданию

№ варианта	Функция
1	$y=asin(bx+c)-d\cos(gx^2+fx+h)$
2	$y=asin(bx)$
3	$y=acos(bx^2+cx+d)$
4	$y=ax^3+bx^2+cx+d$
5	$y=ax^5+bx^4+cx^3+dx^2+fx+g$
6	$y=ax^4+bx^3+cx^2+dx+f$
7	$y=(ax^2+bx+c)\sin(dx)$
8	$y=(ax^3+bx^2+cx+d)\sin(x^2)$
9	$y=(ax^4+bx^3+cx^2+dx+f)\cos(ax^2+bx+c)$
10	$y=esin(ax^2+bx+c)+d\cos(fx+g)$
11	$y=dsin(ax^2+bx+c)\exp(-fx)$
12	$y=a^{bx+c}+f$
13	$y=a\ln(bx+c)+d*\lg(fx)$
14	$y=a\exp(bx)+c$
15	$y=acos(bx)\exp(-cx)$

Для защиты индивидуального задания обучающийся показывает преподавателю выполненное задание в электронном виде и, при необходимости, поясняет, как выполнял.

Требования к индивидуальному заданию:

1. Программный код должен быть написан на языке программирования Python.
2. Именованые переменных должно быть осмысленным.

3. В коде программы должны присутствовать комментарии.
4. Код должен производить форматированный вывод результатов.
5. Код должен запрашивать у пользователя входные данные.
6. Код должен иметь базовый уровень проверки входных данных.

Критерии оценки:

- оценка «зачтено» выставляется обучающемуся, если выполнены все требования задания и даны ответы на вопросы преподавателя;
- оценка «не зачтено» выставляется обучающемуся, если задание выполнено менее чем на половину, обучающийся затруднялся с ответами на вопросы преподавателя.

Индивидуальное задание «Оформление отчёта по созданию программы на языке программирования Python»

Оформите отчёт по итогам выполнения индивидуального задания «Python: работа с графикой».

Структура отчета приведена ниже:

1. Титульный лист. На титульном листе указываются фамилия и инициалы обучающегося, номер группы, ФИО преподавателя, № варианта задания.
2. Постановка задачи. Задание копируется из методических указаний согласно варианту.
3. Ход выполнения. В данной части подробно описывается процесс выполнения заданий, приводятся листинг, результаты работы программы и их интерпретация.
4. Заключение. В заключении описывается основной результат, полученный в ходе выполнения задания, формируются выводы по проделанной работе.
5. Список литературы. Список литературы должен содержать минимум две ссылки.

Требования к отчёту:

1. Структура отчета должна быть строго соблюдена.
2. Отчёт должен быть отформатирован.
3. Каждая структурная часть должна начинаться с новой страницы.
4. Все страницы в отчете, кроме титульной, должны быть пронумерованы.
5. Отчёт должен содержать описание всех особенностей реализации кода.

Критерии оценки:

- оценка «зачтено» выставляется обучающемуся, если выполнены все требования задания и даны ответы на вопросы преподавателя;
- оценка «не зачтено» выставляется обучающемуся, если задание выполнено менее чем на половину, обучающийся затруднялся с ответами на вопросы преподавателя.