

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Бойко Елена Григорьевна

Должность: Ректор

Дата подписания: 04.10.2024 10:17:27

Уникальный программный ключ:

e69eb689122030af7d22cc354bf0eb9d453ecf8f

Министерство науки и высшего образования РФ

ФГБОУ ВО Государственный аграрный университет Северного Зауралья

Инженерно-технологический институт

Кафедра математики и информатики

«Утверждаю»

И.о. заведующего кафедрой



Н.Н. Мальчукова

«31» мая 2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Введение в анализ данных

для направления подготовки 21.04.02 Землеустройство и кадастры

магистерская программа *Землепользование: организация, оценка и управление*

Уровень высшего образования – магистратура

Форма обучения: очная, заочная

Тюмень, 2024

При разработке рабочей программы учебной дисциплины в основу положены:

- 1) ФГОС ВО по направлению подготовки 21.04.02 Землеустройство и кадастры, утвержденный Министерством образования и науки РФ 11 августа 2020 г., приказ № 945.
- 2) Учебный план основной образовательной программы Землепользование: организация, оценка и управление одобрен Ученым советом ФГБОУ ВО ГАУ Северного Зауралья от 31 мая 2024 г. Протокол № 14.

Рабочая программа учебной дисциплины (модуля) одобрена на заседании кафедры математики и информатики от 31 мая 2024 г. Протокол № 10.

И.о. заведующего кафедрой  Н.Н. Мальчукова

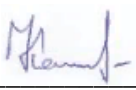
Рабочая программа учебной дисциплины (модуля) одобрена методической комиссией института от 31 мая 2024 г. Протокол № 8.

Председатель методической комиссии института  С.М. Каюгина

Разработчик:

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

Директор института:



М.А. Коноплин

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

<i>Код компетенции</i>	Результаты освоения	Индикатор достижения компетенции	Перечень планируемых результатов обучения по дисциплине
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	ИДЗ ук-1 - Осваивает теоретические основы и методы анализа данных, применяемых при решении прикладных задач в области профессиональной деятельности	<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 *ФТД. Факультативы.*

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

Введение в анализ данных является предшествующей дисциплиной для выпускной квалификационной работы.

Дисциплина изучается на 2 курсе в 3 семестре по очной и заочной формам обучения.

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	8
<i>В том числе:</i>	-	-
Лекционного типа	24	4
Семинарского типа	24	4
Самостоятельная работа (всего)	60	100
<i>В том числе:</i>	-	-
Проработка материала лекций, подготовка к занятиям	30	75
Самостоятельное изучение тем	6	
Курсовой проект (работа)	-	-
Расчетно-графические работы	-	-
Контрольная работа	-	25
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.	Основы анализа данных	Роль анализа данных в современном мире. Научные исследования. Программное обеспечение. Построение системы анализа данных.
2.	Виды данных, классификация методов анализа данных	Данные, подходы и определения, жизненный цикл данных. Понятие метаданных, жизненный цикл метаданных. Большие данные, системы управления Большими данными. Основные задачи и классификация методов анализа данных.
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.
4.	Анализ данных в программе AtteStat excel	Параметрическая и непараметрическая статистика. Дисперсионный, регрессионный, корреляционный, факторный, кластерный и информационный анализ.
5.	Интеллектуальный анализ данных в аналитической платформе Loginom Community Edition	Интерфейс и компоненты обработки данных в Low-code платформе Loginom. Этапы интеллектуального анализа данных. Трансформация данных, визуализация и аналитическая отчетность. Прогнозирование временных рядов. Использование технологии нейронных сетей для задач прогнозирования. Методы кластерного анализа и поиска ассоциативных правил в Loginom режиме Data Mining.

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.	Основы анализа данных.	6	2	8	16
2.	Виды данных, классификация методов анализа данных.	6	2	6	14
3.	Google таблицы для анализа данных.	2	6	6	14
4.	Анализ данных в программе AtteStat excel.	2	6	6	14
5.	Интеллектуальный анализ данных в аналитической платформе Loginom Community Edition	8	8	34	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.	Основы анализа данных.	2	-	20	22
2.	Виды данных, классификация методов анализа данных.	2	-	20	22
3.	Google таблицы для анализа данных.	-	2	16	18
4.	Анализ данных в программе AtteStat excel.	-	2	20	22
5.	Интеллектуальный анализ данных в аналитической платформе Loginom Community Edition	2	2	20	24
	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.	1	MS Excel. Быстродействие систем анализа данных.	2	-
2.	2	MS Excel. Бинарная и множественная классификация данных.	2	-
3.	3	Google таблицы. Формулы для анализа данных.	2	2
4.	3	Google таблицы. Сводные таблицы и диаграммы.	2	-
5.	3	Google таблицы. Автоматизация работы. Макросы.	2	-
6.	4	AtteStat. Параметрическая и	2	-

		непараметрическая статистика.		
7.	4	AtteStat. Дисперсионный и факторный анализ.	2	-
8.	4	AtteStat. Регрессионный и корреляционный анализ.	2	2
9.	5	Loginom Community. Базовые навыки работы. Очистка и предобработка данных.	2	-
10.	5	Loginom Community. Трансформация и визуализация.	2	-
11.	5	Loginom Community. Прогнозирование на основе логистической регрессии.	2	-
12.	5	Loginom Community. Выявление ассоциативных правил.	2	-
		Итого:	24	4

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	75	Тестирование
Самостоятельное изучение тем	6		
Контрольная работа	-	25	Защита контрольной работы
Индивидуальное задание	24	-	Защита индивидуального задания
всего часов:	60	100	

5.2. Учебно-методические материалы для самостоятельной работы:

1. Введение в анализ данных [Электронный ресурс]: Методические указания к выполнению самостоятельной работы для направлений подготовки, реализуемых в ГАУ Северного Зауралья / Автор-сост. Д.В. Ерёмину. - Тюмень: ГАУ Северного Зауралья, 2023. - 12 с.

2. Введение в анализ данных [Электронный ресурс]: Методические указания к выполнению контрольной работы для направлений подготовки, реализуемых в ГАУ Северного Зауралья / Автор-сост. Д.В. Ерёмину. - Тюмень: ГАУ Северного Зауралья, 2023. - 12 с.

3. Шнарева, Г. В. Анализ данных: учебно-методическое пособие / Г. В. Шнарева, Ж. Г. Пономарева. - Симферополь: Университет экономики и управления, 2019. - 129 с. - Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/89482.html> - Режим доступа: для авторизир. пользователей.

4. 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.

5.3. Темы, выносимые на самостоятельное изучение:

Раздел 1. Основы анализа данных

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 6. Методы кластерного анализа и поиска ассоциативных правил в Deductor Studio.

1. Поиск ассоциативных правил.

1.1. Введение в аффинитивный анализ.

1.2. Методика поиска ассоциативных правил.

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	ИДЗ ук-1 - Осваивает теоретические основы и методы анализа данных, применяемых при решении прикладных задач в области профессиональной деятельности	<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. Целых, А. Н. Современные методы прикладной информатики в задачах анализа данных: учебное пособие по курсу «Методы интеллектуального анализа данных» / А. Н. Целых, А. А. Целых, Э. М. Котов. - Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2021. - 130 с. - ISBN 978-5-9275-3783-9. - Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/117165.html> - Режим доступа: для авторизир. пользователей.

2. Котиков, П. Е. Анализ данных: учебно-методическое пособие / П. Е. Котиков. — Санкт-Петербург: СПбГПМУ, 2019. - 48 с. - ISBN 978-5-907184-46-6. - Текст: электронный // Лань: электронно-библиотечная система. - URL: <https://e.lanbook.com/book/174498> - Режим доступа: для авториз. пользователей.

б) дополнительная литература

3. Шнарева, Г. В. Анализ данных: учебно-методическое пособие / Г. В. Шнарева, Ж. Г. Пономарева. - Симферополь: Университет экономики и управления, 2019. - 129 с. - Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/89482.html> - Режим доступа: для авторизир. пользователей.

4. 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.

5. 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.

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

1. Введение в анализ данных [Электронный ресурс]: Методические указания к практическим работам для направлений подготовки, реализуемых в ГАУ Северного Зауралья / Автор-сост. Д.В. Ерёмкина. – Тюмень: ГАУ Северного Зауралья, 2023. - 60 с.

10. Перечень информационных технологий

1. Microsoft Windows 8 Professional;
2. Microsoft Office Standard 2010;
3. Надстройка MS Excel AtteStat;
4. Сервисы Google Suite for Education;
5. Loginom Community Edition CE 6.5.4;
6. Система электронного обучения Moodle.

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 науки и высшего образования РФ
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 21.04.02 Землеустройство и кадастры

магистерская программа *Землепользование: организация, оценка и управление*

Уровень высшего образования – магистратура

Форма обучения: очная, заочная

Разработчик:

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

Утверждено на заседании кафедры

протокол № 10 от 31 мая 2024 г.

И.о. заведующего кафедрой _____  Н.Н. Мальчукова

Тюмень, 2024

**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**

Вопросы к зачёту

Компетенция	Вопросы
<p>УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий</p>	<ol style="list-style-type: none"> 1. Определение «анализ данных». 2. Основные задачи анализа данных. 3. Примеры применения методов анализа данных. 4. Алгоритм построения системы анализа данных. 5. Пример актуального направления в области анализа данных. 6. Технологии анализа данных. 7. Аналитический и информационный подходы к моделированию. 8. Этапы интеллектуального анализа данных. 9. Классы задач, решаемые методами Data Mining. 10. Определение «данные». ГОСТы с определениями данных. 11. Определения из ФЗ-149. 12. Жизненный цикл данных. 13. Этап «Синтез данных». 14. Этап «Использование данных». 15. Этап «Публикация данных». 16. Этап «Архивация данных». 17. Метаданные. ГОСТы для метаданных. 18. Жизненный цикл метаданных. 19. Этап «Оценка требований и анализ контента». 20. Этап «Спецификация системных требований». 21. Этап «Система метаданных». 22. Этап «Сервис и оценка». 23. Определение «Большие данные». 24. Пять характеристик, присущих Большим данным. 25. Базовые принципы обработки Больших данных. 26. Определение «Аналитика Больших данных». 27. Основные возможности аналитики Больших данных. 28. Классификация и прогнозирование как группа задач анализа данных. 29. Кластеризация как группа задач анализа данных. 30. Ассоциации, анализ взаимозависимостей как группа задач анализа данных. 31. Визуализация как группа задач анализа данных. 32. Аналитический подход к моделированию. 33. Информационный подход к моделированию. 34. Технология KDD. 35. Настройки доступа к google таблицам. 36. Сводные таблицы и диаграммы. 37. Применение статистических формул для анализа данных. 38. Текстовые сложные формулы Lookup, Vlookup. 39. Формулы условия IF. 40. Ссылки и массивы (ВПР, ГПР). 41. Макросы. 42. Параметрическая и непараметрическая статистика.

	<p>43. Дисперсионный анализ в Attestat.</p> <p>44. Факторный анализ в Attestat.</p> <p>45. Регрессионный анализ в Attestat.</p> <p>46. Корреляционный анализ в Attestat.</p> <p>47. Определение «нейронная сеть».</p> <p>48. Принципиальное отличие нейронных сетей от обычных программных систем.</p> <p>49. Нейросетевая модель обработки данных.</p> <p>50. Элементы, из которых состоит искусственный нейрон.</p> <p>51. Условия успешного функционирования нейронной сети.</p> <p>52. Основные недостатки нейронных сетей.</p> <p>53. Задача кластеризации, её отличие от классификации.</p> <p>54. Определение «кластер».</p> <p>55. Основные этапы кластерного анализа.</p> <p>56. Процесс обучения без учителя (самоорганизация).</p> <p>57. Отличие процесса обучения с учителем, от обучения без учителя.</p> <p>58. Структура сети Кохонена.</p> <p>59. Принцип конкурентного обучения, его основные шаги.</p> <p>60. Структура карты Кохонена и принцип выбора числа ячеек в ней.</p> <p>61. Оценка результата кластеризации по раскраске карты Кохонена.</p> <p>62. Разведочный анализ.</p> <p>63. Временной ряд.</p> <p>64. Множественная линейная регрессия.</p> <p>65. Прогнозирование с использованием множественной линейной регрессии в платформе Loginom CE 6.5.4.</p> <p>66. Адаптивные модели прогнозирования.</p> <p>67. Компоненты временного ряда.</p> <p>68. Процесс прогнозирования временного ряда с помощью компонентного анализа.</p> <p>69. Методы анализа временных рядов.</p> <p>70. Сезонные колебания.</p> <p>71. Тренд временного ряда.</p> <p>72. Сезонная компонента временного ряда.</p> <p>73. Циклическая компонента временного ряда.</p> <p>74. Случайная компонента временного ряда.</p> <p>75. Цели нахождения ассоциативных правил.</p> <p>76. Область применения деревьев решений.</p> <p>77. Штатный алгоритм построения деревьев решений аналитической платформе Loginom CE 6.5.4.</p> <p>78. Применение деревьев решений для описания данных.</p> <p>79. Применение деревьев решений для классификации.</p> <p>80. Примеры облачных систем.</p>
--	--

Обучающийся допускается к промежуточной аттестации по дисциплине при условии выполнения всех практических работ, контрольной работы, 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 обучение Moodle выполнено с результатом 50% и выше;
 - «не зачтено» выставляется обучающемуся, если итоговое тестирование в системе электронного обучение Moodle выполнено с результатом 49% и ниже.

Индивидуальные творческие задания (проекты):

Индивидуальное задание «Исследование зависимости точности выполнения операций от количества нейронов»

Решение данной задачи произведите с помощью эмулятора нейронных сетей – штатной опцией аналитической системы Loginom SE 6.5.4. Мастер обработки системы позволяет сконструировать нейронную сеть с заданной структурой, определить ее параметры и обучить ее с помощью одного из доступных в системе алгоритмов обучения.

Исследуйте зависимость точности моделирования выполняемых операций от количества нейронов.

В результате выполнения задания должен быть получен график зависимости точности от количества нейронов в нейронной сети и сделан вывод.

Порядок выполнения:

1. Находясь на узле импорта, откройте мастер обработки.
2. Выберите режим обработки «Нейросеть».
3. Установите назначение полей «А», «В» «Операция» как входные, а поле «R» – как выходное.
4. Установите активационную функцию **в зависимости от вашего варианта**. В таблице приведены конкретные значения параметров (тип функции активации: сигмоида, гипертангенс, арктангенс и значение ее крутизны для каждого из вариантов).

Варианты к индивидуальному заданию

Вариант	Тип функции активации	Крутизна
1	Сигмоида	1,2
2	Гипертангенс	0,9
3	Арктангенс	1,0
4	Сигмоида	0,8
5	Гипертангенс	1,0
6	Сигмоида	0,7
7	Сигмоида	1,3
8	Гипертангенс	1,1
9	Гипертангенс	1,4
10	Сигмоида	0,9
11	Гипертангенс	1,2
12	Сигмоида	0,7
13	Сигмоида	1,1
14	Гипертангенс	1,3
15	Сигмоида	0,95

5. Разделите исходное множество данных на обучающее и тестовое, используя параметры «по умолчанию». То есть, установите размер обучающего множества 95%, а тестового 5%.
6. Укажите способ разбиения исходного множества данных «Случайно».
7. Установите параметры остановки обучения нейронной сети – ошибку менее 0,005 или достижение эпохи обучения 35000.
8. Запустите процесс обучения нейронной сети.

9. После окончания обучения выберите визуализаторы Data Mining («Граф нейросети», «Диаграмма рассеяния», «Что-если») и Табличные данные («Обучающий набор», «Таблица» и «Статистика»).
10. В результате обучения нейросети будут получены: граф нейросети, диаграмма рассеяния, таблица результатов, статистические характеристики, инструмент для проведения вычислений «что-если».
11. Последовательно изменяйте количество нейронов внутреннего слоя в диапазоне от 4 до 22 с шагом 2 и проводите построение нейронных сетей.
12. При построении каждой новой нейросети перемещайте курсор в позицию «Текстовый файл».
13. Фиксируйте значения средней ошибки.
14. Получите график зависимости величины средней ошибки от количества нейронов в скрытом слое нейросети.
15. Определите то количество нейронов, которое обеспечивает минимальную среднюю ошибку в данном диапазоне изменения числа нейронов.

Для защиты индивидуального задания обучающийся показывает преподавателю выполненное задание в аналитической системе Loginom CE 6.5.4. при необходимости, поясняет, как выполнял.

Вопросы к индивидуальному заданию:

1. Понятие «нейронная сеть».
2. Принципиальное отличие нейронных сетей от обычных программных систем.
3. Нейросетевая модель обработки данных.
4. Элементы, из которых состоит искусственный нейрон.
5. Примеры наиболее распространенных функций активации.
6. Условия успешного функционирования нейронной сети.
7. «Тренировка» нейронной сети.
8. Основные недостатки нейронных сетей.
9. Эмулятор нейронных сетей.
10. Возможности аналитической платформы Loginom CE 6.5.4.

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

Отчёт по теме «Исследование зависимости точности выполнения операций от количества нейронов»

Оформите отчёт по итогам выполнения индивидуального задания «Исследование зависимости точности выполнения операций от количества нейронов».

Структура отчета приведена ниже:

1. Титульный лист. На титульном листе указываются фамилия и инициалы обучающегося, номер группы, ФИО преподавателя, № варианта задания.
2. Постановка задачи. Берётся из индивидуального задания, указать тип активации, согласно варианту.
3. Ход выполнения. Скриншот каждой полученной нейросети (в программе) с описанием этапов работы.
4. Таблицу зависимости величины средней ошибки от количества нейронов.
5. График зависимости величины средней ошибки от количества нейронов.
6. Заключение. В заключении формулируются выводы по проделанной работе, указывается количество нейронов, обеспечивающих минимальную среднюю.
7. 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.

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.

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. Прогнозирование с использованием множественной линейной регрессии в платформе Loginom CE 6.5.4.
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. Классификация и прогнозирование как группа задач анализа данных.
4. Регрессионный анализ данных средствами google таблиц.
5. Штатный алгоритм построения деревьев решений аналитической платформе Loginom CE 6.5.4.
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. Кластеризация как группа задач анализа данных.
3. Сводные таблицы и диаграммы средствами google таблиц.
4. Принципиальное отличие нейронных сетей от обычных программных систем.
5. Loginom. Применение деревьев решений для классификации.
6. Практическое задание: пользуйтесь 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. Жизненный цикл данных.
3. Аналитический подход к моделированию.
4. Google таблицы: ссылки и массивы (ВПР, ГПР).
5. Loginom. Тренд временного ряда.
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. Задача кластеризации, её отличие от классификации.
3. Информационный подход к моделированию.
4. Корреляционный анализ данных средствами google таблиц.
5. Факторный анализ в Attestat.
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. Пример актуального направления в области анализа данных.
3. Классификация задач Data Mining.
4. Применение макросов в google таблицах.
5. Ассоциации, анализ взаимосвязей как группа задач анализа данных.
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. Визуализация как группа задач анализа данных.
3. Функция конкурентного сходства.
4. Формулы условия IF в google таблицах.
5. Дисперсионный анализ в Attestat.
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. Условия успешного функционирования нейронной сети.
3. Этапы решения задачи анализа данных и их взаимосвязи.
4. Применение статистических формул для анализа данных в google таблицах.
5. Loginom. Оценка результата кластеризации по раскраске карты Кохонена.
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. Определение «Аналитика Больших данных».
3. Качественный и количественный этапы анализа данных.
4. Настройки доступа к google таблицам.
5. Корреляционный анализ в Attestat.

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. Основные недостатки нейронных сетей.
3. Текстовые сложные формулы Lookup, Vlookup в google таблицах.
4. Основные задачи анализа данных.
5. Loginot. Цели нахождения ассоциативных правил.
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.