



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(ДГТУ)**

Кафедра «Организация перевозок и дорожного движения»

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ
ДЛЯ ВЫПОЛНЕНИЯ КОНТРОЛЬНЫХ РАБОТ ПО ДИСЦИПЛИНЕ
«ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ НА ТРАНСПОРТЕ»**

Направление подготовки 23.04.01 «Технология транспортных процессов»

Профиль: «Транспортная логистика»; «Организация перевозок на автомобильном транспорте»; «Интеллектуальные транспортные системы»

Ростов-на-Дону
ДГТУ
2023

УДК 656

Составитель: Роговенко Т.Н.

Методические указания для выполнения контрольных работ по дисциплине «Интеллектуальный анализ данных на транспорте» –
Ростов-на-Дону: Донской гос. техн. ун-т, 2023. – 7 с.

УДК 656

Методические указания предназначены для магистрантов заочной формы обучения по направлению подготовки 23.04.01 «Технология транспортных процессов», программ «Транспортная логистика», «Организация перевозок на автомобильном транспорте», «Интеллектуальные транспортные системы».

Печатается по решению редакционно-издательского совета Донского государственного технического университета

Ответственный за выпуск зав. кафедрой «Организации перевозок и дорожного движения» д-р техн. наук, профессор В.В. Зырянов

В печать ____ . ____ .2023 г.
Формат 60×84/16. Объем ____ усл. п. л.
Тираж ____ экз. Заказ № ____

Издательский центр ДГТУ
Адрес университета и полиграфического предприятия:
344000, г. Ростов-на-Дону, пл. Гагарина, 1

© Донской государственный
технический университет, 2023

Введение

Дисциплина «Интеллектуальный анализ данных на транспорте» изучается 2 семестра. В каждом семестре учебным планом предусмотрено выполнение студентом одной контрольной работы. Контрольная работа №1 предусматривает выполнение заданий по подготовке и статистическому анализу транспортных данных. Контрольная работа №2 – применение к транспортным данным методов глубокого обучения и анализ его результатов.

Теоретические основы

Теоретическими основами интеллектуального анализа применительно к данным транспорта являются: методы системного анализа и широкий спектр методов статистического анализа, методы машинного обучения (классификации и кластеризации, ассоциативный анализ) и теория принятия решений.

Требования по выполнению и оформлению контрольной работы

Контрольная работа выполняется с применением информационных технологий для индивидуального набора данных (dataset), который выдает преподаватель. Все вычисления и графики выполняются в электронных таблицах (например, MS Excel), текст контрольной работы (пояснительная записка) оформляется с помощью текстовых редакторов (например, MS Word) в соответствии с примерами из методических указаний для практических занятий по дисциплине и общими требованиями по оформлению расчетно-пояснительной записки. Следует обратить внимание, что правила оформления методических указаний отличаются от правил оформления пояснительных записок. Оформлять контрольную работу следует в соответствии с ГОСТ 2.105-95 . На проверку следует сдавать (загружать на do.skif.donstu.ru в раздел «Заочникам: контрольная работа №1» и «Заочникам: контрольная работа №2», соответственно) два файла: текст в формате .docx и расчеты в формате .xlsx. В имени файлов должны содержаться фамилия студента, группа и номер зачетки. Например: Иванов_АМЗТЛ21_123456789.docx . Контрольная работа должна быть распечатана и сдана преподавателю.

Если содержание контрольной работы отвечает предъявляемым требованиям, то она допускается к защите. При неудовлетворительном выполнении контрольной работы она возвращается студенту на доработку.

По контрольной работе проводится устный опрос (зачет контрольной работы), после которого студент приступает к сдаче зачета по дисциплине.

Алгоритм выбора варианта контрольной работы

Для задания 1 исходные данные следует получить у преподавателя. Для задания 2 номер теоретического вопроса равен последней цифре в номере зачетной книжки студента.

Контрольная работа №1

Задание 1

Все пункты задания выполняются и оформляются в соответствии с методическими указаниями для практических занятий по дисциплине «Интеллектуальный анализ данных на транспорте», которые можно скачать с сайта do.skif.donstu.ru [1].

Задание выполняется последовательно:

1. Провести очистку данных из файла. Найти противоречия в исходных данных и провести преобразования.
2. Получить характеристики описательной статистики для всех параметров после очистки и преобразований данных.
3. Провести статистическое оценивание одного параметра, например speed. Провести однофакторный дисперсионный анализ средних скоростей движения (выборкой считать данные за сутки). Провести корреляционный анализ параметров counter, speed, k.
4. Получить линейную многофакторную регрессионную модель.
5. Агрегировать данные: определить суммарные часовые интенсивности counter, средние скорости speed и плотности k. Для агрегированных данных выполнить п.6-8.

6. Подобрать однофакторную нелинейную регрессионную модель с максимальным коэффициентом корреляции для каждого из параметров counter, speed, k на точечной диаграмме, добавляя линию тренда. Фактором считать время наблюдения.
7. Построить коррелограмму для параметров counter, speed, k и провести ее анализ.
8. Применить гармонический анализ к исходным данным о транспортном потоке, например к параметру speed.
9. Сравнить модели из п.п. 8 и 6. Сделать вывод.

Задание 2.

Следует дать развернутый ответ на один из вопросов из списка. Номер теоретического вопроса равен последней цифре в номере зачетной книжки (студенческого билета) студента. Объем ответа не менее 2 страниц (, шрифт Times New Roman, 14 пт, интервал 1,5).

0. Основные понятия интеллектуального анализа данных: интеллектуальный анализ данных, набор данных, обнаружение знаний в данных, "сырые" данные, качество данных.
1. Какие методы используются для получения набора данных?
2. Этапы интеллектуального анализа данных. Характеристики данных, определяющие их качество.
3. Этапы очистки данных. ETL-технология. Нерелевантные данные, дубликаты, преобразование типов, синтаксические ошибки.
4. Масштабирование, преобразование, стандартизация, нормализация данных. Методы обработки пропущенных значений.
5. Уровни интеграции данных. Задачи интеграции данных. Преобразование данных с целью интеграции.
6. Описательный анализ эмпирических данных. Статистическое оценивание.
7. Корреляционный анализ эмпирических данных. Дисперсионный анализ эмпирических данных.

8. Регрессионный анализ (линейные и нелинейные модели).
9. Анализ временных рядов применительно к транспортным данным.

Контрольная работа №2

Задание 1

Все пункты задания выполняются и оформляются в соответствии с методическими указаниями для практических занятий по дисциплине «Интеллектуальный анализ данных на транспорте, которые можно скачать с сайта do.skif.donstu.ru или НТБ ДГТУ [2].

Задание выполняется последовательно:

1. Провести предварительную подготовку и анализ исходных данных.
2. Применить алгоритм k-means к транспортным данным. Построить диаграмму, иллюстрирующую разбиение на кластеры.
3. Применить алгоритм kNN к транспортным данным. Построить иллюстрирующую диаграмму.
4. Применить метод «Наивный Байес» к нечисловым характеристикам объектов исследования.
5. Применить метод ассоциативного анализа к исходным данным о наборах исследуемых объектов.
6. Построить дерево решений для исходных данных о погодных условиях и транспортном событии.
7. Создать простую нейронную сеть прямого распространения для распознавания цифры, равной номеру варианта.
8. Сделать вывод.

Задание 2.

Следует дать развернутый ответ на один из вопросов из списка. Номер теоретического вопроса равен последней цифре в номере зачетной книжки (студенческого билета) студента. Объем ответа не менее 2 страниц (, шрифт Times New Roman, 14 пт, интервал 1,5).

0. Методы машинного обучения. Их цели и сфера применения. Применение методов машинного обучения на транспорте.

1. Кластеризация данных для методов глубокого обучения. Алгоритмы кластеризации.
2. Алгоритмы классификации и их применение в транспортных процессах.
3. Деревья решений как метод анализа данных.
4. Генетические алгоритмы как метод интеллектуального анализа данных.
5. Базовые понятия искусственных нейронных сетей.
6. Применение искусственных нейронных сетей в управлении дорожным движением и транспортными средствами.
7. Применение искусственных нейронных сетей для автоматизации логистических процессов.
8. Описательные и предсказательные задачи knowledge discovery in data mining.
9. Задачи нечеткой логики в анализе данных, возможности их применения на транспорте.

Список используемых источников

1. Методическими указаниями для практических занятий по дисциплине «Интеллектуальный анализ данных на транспорте». Часть 1. – Ростов-на-Дону: Донской гос. техн. ун-т, 2023. – 28 с. Режим доступа:
<https://do.skif.donstu.ru/course/view.php?id=4610>
2. Методическими указаниями для практических занятий по дисциплине «Интеллектуальный анализ данных на транспорте». Часть 2. – Ростов-на-Дону: Донской гос. техн. ун-т, 2023. – 18 с. Режим доступа:
<https://do.skif.donstu.ru/course/view.php?id=4610>
2. Learn Data Mining Through Excel: A Step-by-Step Approach for Understanding Machine Learning Methods. Hong Zhou. Режим доступа:
<https://doi.org/10.1007/978-1-4842-5982-5>
3. Data mining : practical machine learning tools and techniques / Ian H. Witten, Eibe Frank. -: Morgan Kaufmann Publishers, 2020. - 525 p.
4. Воскобойников Ю. Е. Эконометрика в Excel. Новосиб. гос. архитектур.-строит. ун-т. – Новосибирск : НГАСУ (Сибстрин), 2008. – 152 с.