

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**
**Федеральное государственное автономное образовательное учреждение
высшего образования
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ
ПРАКТИЧЕСКИХ РАБОТ ПО ДИСЦИПЛИНЕ**

Типологический анализ в социологии

Направление подготовки 39.03.01 Социология
Направленность (профиль) – «Аналитика и консалтинг»
Форма обучения – Очная, заочная
Год начала обучения – 2026 г.

Ставрополь, 2026

СОДЕРЖАНИЕ

Введение

Практическая работа 1. Основные понятия типологического анализа

Практическая работа 2. Методологические основы типологического анализа

Практическая работа 3. Математико-статистические понятия и термины, используемые

Практическая работа 4. Кластерный анализ в социологии

Практическая работа 5. Иерархический кластерный анализ

Практическая работа 6. Кластерный анализ методом к-средних

Практическая работа 7. Визуализация данных кластерного анализа, дендрограммы, ствол-лист

Практическая работа 8. Методы снижения размерности

Практическая работа 9. Классификация методом ближайших соседей

Практическая работа 10. Двухэтапный кластерный анализ

Практическая работа 11. Использование факторного анализа в социологии

Практическая работа 12. Методы вращения факторов

Практическая работа 13. Интерпретация результатов факторного анализа

Практическая работа 14. Модели классификации на основе деревьев решений

Практическая работа 15. Методы построения деревьев классификации CHAID

Практическая работа 16. Методы построения деревьев решений SQRT

Практическая работа 17. Построение деревьев решений методом QUEST

Практическая работа 18. Дискриминантный анализ

ВВЕДЕНИЕ

Цель освоения дисциплины: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Задачи освоения дисциплины:

- познакомить с основными методами анализа типологического анализа данных, применяемыми в социологии;
- научиться вести практические расчеты при решении типовых для социологии задач типологизации данных;
- выработать умения проводить интерпретацию полученных при расчетах данных.

Компетенции обучающегося, формируемые в результате изучения дисциплины

Код, формулировка компетенции	Код, формулировка индикатора	Планируемые результаты обучения по дисциплине (модулю), характеризующие этапы формирования компетенций, индикаторов
ПК-1 Способен к подготовке и планированию проектного предложения по реализации фундаментального или прикладного социологического исследования	ИД-2 ПК-1 Разрабатывает методологическую и методическую части программы социологического исследования	Используя компетенции в области типологического анализа данных, готовит аналитическую информацию об исследуемых социальных группах, процессах и явлениях
ПК-2 Способен к разработке программных и методических документов фундаментального или прикладного социологического исследования	ИД-2 ПК-2 Выбирает и описывает методы сбора и обработки социологической информации; анализирует возможность применения социологических методов, методик, технологий	Используя возможности аппарата типологического анализа, демонстрирует возможности использования теоретических знаний и результатов социологических исследований для выявления социально значимых проблем
ПК-3 Способен описать, объяснить, прогнозировать социальные явления и процессы на основе результатов фундаментальных и прикладных социологических исследований	ИД-4 ПК-3 Применяет группировку и типологизацию социальных явлений и процессов при проведении социологического исследования	На основе компетенций в области группировки и типологизации анализирует социальные явления, структуры и процессы

ПЛАНЫ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

ПРАКТИЧЕСКАЯ РАБОТА 1

Тема: «Основные понятия типологического анализа»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Основной целью типологического анализа является проверка гипотезы о существовании типов, интерпретируемые как объекты социального управления. Таковыми можно считать типы электорального поведения, типы времяпрепровождения, типы потребительского поведения, типы общественного мнения и т. д.

Типологический анализ рассматривается как методологическая процедура эмпирической социологии и трактуется как метаметодика — структурный элемент методологии анализа социологических данных. Вводится система языковых конструкторов метаметодики, достаточно полная, непротиворечивая и открытая для дополнений. Предлагается логика проведения типологического анализа в ситуации жестко структурированных и слабо структурированных данных.

Перечень используемого оборудования:

Мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

Персональные компьютеры из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше) Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности.

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. по периметру оставляется свободное место 2-5 сантиметров.

Вопросы и задания для обсуждения:

1. Приведите примеры целесообразности и необходимости использования типологического анализа данных.
2. Назовите цели использования типологического анализа при проведении социологических исследований.

Примерные темы для лабораторных работ:

1. Характеристика специализированных программ для осуществления типологического анализа
2. Характеристика возможностей программы Orange
3. Характеристика возможностей модулей программы Knime

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА2

Тема: «Методологические основы типологического анализа»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате

освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Типология — это метод научного познания, направленный на разделение некоторой совокупности объектов и их объединение в некоторую упорядоченную группу с помощью типа, то есть обобщённой, идеализированной модели. Типология используется в целях сравнительного изучения существенных признаков, связей, функций, отношений, уровней организации объектов, как сосуществующих, так и разделённых во времени. В системе научно-теоретических и практико-ориентированных дискурсов термин «типология» применяется для обозначения результатов типологического описания и сопоставления.

Проблема типологии возникает во всех науках, которые имеют дело с разнородными по-своему составу множествами объектов (как правило, дискретных) и решают задачу упорядоченного описания (систематизации) и объяснения этих множеств. Будучи одной из наиболее универсальных процедур научного мышления, типология опирается на выявление сходства и различия изучаемых объектов, на поиск надёжных способов их идентификации. В своей теоретически развитой форме типология стремится отобразить строение исследуемой системы, выявить её закономерности, позволяющие предсказать существование неизвестных пока ещё объектов.

Перечень используемого оборудования:

Мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

Персональные компьютеры из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше) Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности.

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).

- расположить монитор на уровне глаз и перпендикулярно углу зрения.

- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. по периметру оставляется свободное место 2-5 сантиметров.

Вопросы и задания для обсуждения:

Задание 1. Охарактеризуйте общенаучную методологию, на которую опирается типологический анализ.

Задание 2. Назовите основные теоретические дефиниции типологического анализа.

Контрольные вопросы:

1. Какова схема организации данных при типологическом анализе.

2. Каковы основы методические подходы проведения типологического анализа.

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА 3

Тема: «Математико-статистические понятия и термины, используемые в типологическом анализе данных»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ

эмпирических данных.

Теоретическая часть

Логика решения задач типологического анализа одинакова. Этапы решения. Этапы, при реализации которых возможно использование математических методов разного класса. Комплексное применение математических методов для целей типологизации. Последовательная и параллельная стратегии использования математических конструктов в процессе типологического анализа.

Основания для выделения пяти различных типов социологической информации. Сравнительный анализ пяти типов эмпирических данных по различным основаниям. Формы существования информации: динамические ряды, матрицы типа объект-признак, объект-объект, признак-признак, объект-бюджетная линейка. Специфика типов социологической информации в контексте различия задач, для решения которых необходимы эмпирические данные.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.

- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).

- расположить монитор на уровне глаз и перпендикулярно углу

зрения.

– мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

– клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

– продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.

– длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.

– каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

1. Составные части системы типобразующих признаков.
2. Признаки – характеристики проявления социального явления.
3. Признаки – характеристики факторов, объясняющих существования различных типов изучаемого социального явления в определенных заданных условиях.

4. Измерение на этапе сбора информации и измерение на этапе обработки эмпирических данных

Контрольные вопросы:

1. Примеры всевозможных постановок задач типологического анализа.
2. Различные значения понятия «тип» (типовой, типологический, типический).
3. Реальные и идеальные типы
4. Типологический анализ в исследованиях с жестко формализованными методами сбора

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА 4

Тема: «Кластерный анализ в социологии»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Первое применение кластерный анализ нашел в социологии. Название кластерный анализ происходит от английского слова cluster – гроздь, скопление. Впервые в 1939 был определен предмет кластерного анализа и сделано его описание исследователем Трионом (Tryon). Главное назначение кластерного анализа – разбиение множества исследуемых объектов и признаков на однородные в соответствующем понимании группы или кластеры. Это означает, что решается задача классификации данных и выявления соответствующей структуры в ней. Методы кластерного анализа можно применять в самых различных случаях, даже в тех случаях, когда речь идет о простой группировке, в которой все сводится к образованию групп по количественному сходству.

Большое достоинство кластерного анализа в том, что он позволяет производить разбиение объектов не по одному параметру, а по целому набору признаков. Кроме того, кластерный анализ в отличие от большинства математико-статистических методов не накладывает никаких ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных практически произвольной природы. Это имеет большое значение, например, для прогнозирования конъюнктуры, когда показатели имеют разнообразный вид, затрудняющий применение традиционных эконометрических подходов. Кластерный анализ позволяет рассматривать достаточно большой объем информации и резко сокращать, сжимать большие массивы социально-экономической информации, делать их компактными и наглядными.

Важное значение кластерный анализ имеет применительно к совокупностям временных рядов, характеризующих экономическое развитие (например, общехозяйственной и товарной конъюнктуры). Здесь можно выделять периоды, когда значения соответствующих показателей были достаточно близкими, а также определять группы временных рядов, динамика которых наиболее схожа.

Кластерный анализ можно использовать циклически. В этом случае исследование производится до тех пор, пока не будут достигнуты необходимые результаты. При этом каждый цикл здесь может давать информацию, которая способна сильно изменить направленность и подходы дальнейшего применения кластерного анализа. Этот процесс можно представить системой с обратной связью. В задачах социально-экономического прогнозирования весьма перспективно сочетание кластерного анализа с другими количественными методами (например, с

регрессионным анализом).

Как и любой другой метод, кластерный анализ имеет определенные недостатки и ограничения: в частности, состав и количество кластеров зависит от выбираемых критериев разбиения. При сведении исходного массива данных к более компактному виду могут возникать определенные искажения, а также могут теряться индивидуальные черты отдельных объектов за счет замены их характеристиками обобщенных значений параметров кластера. При проведении классификации объектов игнорируется очень часто возможность отсутствия в рассматриваемой совокупности каких-либо значений кластеров.

В кластерном анализе считается, что:

- выбранные характеристики допускают в принципе желательное разбиение на кластеры;
- единицы измерения (масштаб) выбраны правильно.

Выбор масштаба играет большую роль. Как правило, данные нормализуют вычитанием среднего и делением на стандартное отклонение, так что дисперсия оказывается равной единице.

Задача кластерного анализа заключается в том, чтобы на основании данных, содержащихся во множестве X , разбить множество объектов G на m (m – целое) кластеров (подмножеств) $Q_1, Q_2 \dots, Q_m$, так, чтобы каждый объект G_j принадлежал одному и только одному подмножеству разбиения. А объекты, принадлежащие одному и тому же кластеру, были сходными, в то время как объекты, принадлежащие разным кластерам, были разнородными.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).
- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).

- расположить монитор на уровне глаз и перпендикулярно углу зрения.

- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.

- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.

- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. Назовите основные меры сходства между кластерами

Задание 2. Назовите основные меры различия между кластерами

Задание 3. Охарактеризуйте метрики расчета близостей объектов в кластерном анализе

Контрольные вопросы:

1. Основные понятия и методы кластерного анализа

2. Задачи кластерного анализа

3. Преимущества и недостатки кластерного анализа

4. Этапы проведения кластерного анализа

5. Особенности исходных данных для кластерного анализа

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА5

Тема: «Иерархический кластерный анализ»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических

данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Иерархический кластерный анализ предназначен для выявления относительно однородных групп наблюдений (или переменных) по заданным характеристикам при помощи алгоритма, который вначале рассматривает каждое наблюдение (переменную) как отдельный кластер, а затем последовательно объединяет кластеры, пока не останется только один. Можно анализировать исходные переменные или воспользоваться набором стандартизирующих преобразований. Расстояния или меры сходства формируются процедурой Расстояния (Proximities). Чтобы помочь в выборе наилучшего решения, на каждом шаге выводятся разнообразные статистики.

Переменные могут быть количественными, бинарными или частотами. Масштаб измерения переменных важен - различия в масштабах могут повлиять на полученные кластерные решения. Если масштаб переменных сильно различается (например, одна переменная измерена в долларах, а другая - в годах), то следует подумать об их стандартизации (она может быть проведена автоматически с помощью процедуры Иерархическая кластерный анализ).

Если во входных данных существуют совпадающие расстояния или сходства или они появляются в обновленных кластерах в процессе объединения, то результирующее кластерное решение может зависеть от порядка наблюдений в файле. Возможно, что вы захотите получить несколько различных решений с наблюдениями, упорядоченными случайным образом, чтобы проверить стабильность данного решения.

Используемые расстояния или меры сходства должны соответствовать анализируемым данным (более полную информацию относительно выбора расстояний и мер сходства можно найти в описании процедуры Proximities (Расстояния)). Кроме того, в анализ необходимо включать все переменные, имеющие отношение к проблеме. Игнорирование важных переменных может привести к решению, вводящему в заблуждение. Поскольку иерархический кластерный анализ является разведочным методом, его результаты следует считать предварительными, пока они не будут подтверждены на независимой выборке.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. Осуществите настройку статистик процедуры Иерархический кластерный анализ в программе SPSS.

Задание 2. Проинтерпретируйте результаты процедуры иерархический кластерный анализ.

Контрольные вопросы:

1. Понятие иерархического кластерного анализа
2. Целесообразность применения иерархического кластерного анализа
3. Недостатки иерархического кластерного анализа
4. Интерпретация результатов иерархического кластерного анализа
5. Этапы иерархического кластерного анализа

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА 6

Тема: «Кластерный анализ методом к-средних»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Эта процедура пытается выявить относительно однородные группы наблюдений на основе выбранных характеристик, используя алгоритм, позволяющий обработать большое число наблюдений. Однако этот алгоритм требует указания числа кластеров. Вы можете задать начальные центры кластеров, если такая информация вам доступна. Вы можете выбрать один из двух методов классификации наблюдений, либо итеративно обновляя центры кластеров, либо ограничиваясь только классификацией. Вы можете сохранить принадлежность к кластерам, информацию о расстояниях и окончательные центры кластеров. Дополнительно вы можете задать переменную, значения которой будут использоваться в качестве меток наблюдений при выводе результатов. Вы можете также запросить вывод F -статистик дисперсионного анализа. Относительные величины этих статистик дают информацию о вкладе каждой переменной в разделение групп.

Алгоритм, используемый по умолчанию для выбора начальных центров кластеров, не является инвариантным относительно порядка наблюдений. Для

вычисления расстояний используется простое евклидово расстояние. Если необходимо задать другой тип расстояния или меры сходства, обратитесь к процедуре Иерархический кластерный анализ. Масштабирование переменных играет важную роль. Если ваши переменные имеют различный масштаб измерений (например, одна переменная измерена в долларах, а вторая - в годах), то результаты могут быть некорректными. В этой ситуации необходимо подумать о стандартизации ваших переменных до выполнения кластерного анализа методом k-средних (это можно сделать при помощи процедуры Описательные статистики). Предполагается, что выбрано подходящее число кластеров, а в анализ включены все существенные переменные. Если вы неправильно выбрали число кластеров или не включили важные переменные, то полученные результаты также могут ввести вас в заблуждение.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.

- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).

- расположить монитор на уровне глаз и перпендикулярно углу зрения.

- мышь располагается так, чтобы было удобно работать с ней.

провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

– клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

– продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.

– длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.

– каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью программы SPSS проанализируйте полученную в окне Вывода дендрограмму и сделайте вывод об оптимальном числе кластеров.

Задание 2. Охарактеризуйте алгоритм проведения классификации методом k-средних

Контрольные вопросы:

1. Когда необходимо задать количество кластеров в процедуре кластерного анализа методом k-средних?

2. Какое расстояние используется для нахождения расстояния между кластерам в методе k-средних?

3. По каким данным можно проводить классификацию методом k-средних?

4. Необходимо ли предварительно стандартизировать значения переменных при классификации методом k-средних?

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА7

Тема: «Визуализация данных кластерного анализа, дендрограммы, ствол-лист»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате

освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Первая функция (иллюстрация построения модели), по сути, является визуализацией Data Mining модели. Существует много различных способов представления моделей, но графическое ее представление дает пользователю максимальную "ценность". Пользователь, в большинстве случаев, не является специалистом в моделировании, чаще всего он эксперт в своей предметной области. Поэтому модель Data Mining должна быть представлена на наиболее естественном для него языке или, хотя бы, содержать минимальное количество различных математических и технических элементов.

Таким образом, доступность является одной из основных характеристик модели Data Mining. Несмотря на это, существует и такой распространенный и наиболее простой способ представления модели, как "черный ящик". В этом случае пользователь не понимает поведения той модели, которой пользуется. Однако, несмотря на непонимание, он получает результат - выявленные закономерности. Классическим примером такой модели является модель нейронной сети.

Другой способ представления модели - представление ее в интуитивном, понятном виде. В этом случае пользователь действительно может понимать то, что происходит "внутри" модели. Таким образом, можно обеспечить его непосредственное участие в процессе.

Такие модели обеспечивают пользователю возможность обсуждать ее логику с коллегами, клиентами и другими пользователями, или объяснять ее.

Понимание модели ведет к пониманию ее содержания. В результате понимания возрастает доверие к модели. Классическим примером является дерево решений. Построенное дерево решений действительно улучшает понимание модели, т.е. используемого инструмента Data Mining.

Кроме понимания, такие модели обеспечивают пользователя возможностью взаимодействовать с моделью, задавать ей вопросы и получать ответы. Примером такого взаимодействия является средство "что, если". При помощи диалога "системапользователь" пользователь может получить понимание модели.

Теперь перейдем к функциям, которые помогают интерпретировать и оценить результаты построения Data Mining моделей. Это всевозможные графики, диаграммы, таблицы, списки и т.д.

Примерами средств визуализации, при помощи которых можно оценить качество модели, являются диаграмма рассеивания, таблица сопряженности, график изменения величины ошибки.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для

подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью специализированной компьютерной программы

SPSS осуществите визуализацию процедуры иерархический кластерный анализ.

Контрольные вопросы:

1. Характеристика методов визуализации кластерного анализа
2. Интерпретация результатов построения дендрограммы
3. Интерпретация результатов построения графика «ствол-лист»

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА 8

Тема: «Методы снижения размерности»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

В многомерном статистическом анализе каждый объект описывается вектором, размерность которого произвольна (но одна и та же для всех объектов). Однако человек может непосредственно воспринимать лишь числовые данные или точки на плоскости. Анализировать скопления точек в трехмерном пространстве уже гораздо труднее. Непосредственное восприятие данных более высокой размерности невозможно. Поэтому вполне естественным является желание перейти от многомерной выборки к данным небольшой размерности, чтобы «на них можно было посмотреть».

Кроме стремления к наглядности, есть и другие мотивы для снижения размерности. Те факторы, от которых интересующая исследователя переменная не зависит, лишь мешают статистическому анализу. Во-первых, на сбор информации о них расходуются ресурсы. Во-вторых, как можно доказать, их включение в анализ ухудшает свойства статистических процедур (в частности, увеличивает дисперсию оценок параметров и характеристик распределений). Поэтому желательно избавиться от таких факторов.

Обсудим с точки зрения снижения размерности пример использования регрессионного анализа для прогнозирования объема продаж, рассмотренный

в подразделе 3.2.3. Во-первых, в этом примере удалось сократить число независимых переменных с 17 до 12. Во-вторых, удалось сконструировать новый фактор – линейную функцию от 12 упомянутых факторов, которая лучше всех иных линейных комбинаций факторов прогнозирует объем продаж. Поэтому можно сказать, что в результате размерность задачи уменьшилась с 18 до 2. А именно, остался один независимый фактор (приведенная в подразделе 3.2.3 линейная комбинация) и один зависимый – объем продаж.

При анализе многомерных данных обычно рассматривают не одну, а множество задач, в частности, по-разному выбирая независимые и зависимые переменные. Поэтому рассмотрим задачу снижения размерности в следующей формулировке. Дана многомерная выборка. Требуется перейти от нее к совокупности векторов меньшей размерности, максимально сохранив структуру исходных данных, по возможности не теряя информации, содержащихся в данных. Задача конкретизируется в рамках каждого конкретного метода снижения размерности.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- ☐ не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- ☐ не включать без разрешения оборудование.

- ☐ при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- ☐ знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- ☐ не трогать провода и разъемы (возможно поражение электрическим током).

- ☐ не допускать порчи оборудования.

- ☐ не работать в верхней одежде.

До начала работы:

- ☐ осмотреть рабочее место (расположение блоков и их состояние).

- ☐ расположить монитор на уровне глаз и перпендикулярно углу

зрения.

☐ мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

☐ клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. по периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

☐ продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.

☐ длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.

☐ каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью специализированной компьютерной программы SPSS определите и интерпретируйте основные методы снижения размерности, используемые в типологическом анализе.

Контрольные вопросы:

1. Значение методов снижения размерности в типологическом анализе
2. Целесообразность использования методов снижения размерности в типологическом анализе
3. Этапы реализации методов снижения размерности

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА 9

Тема: «Классификация методом ближайших соседей»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

В метрических методах мерой сходства является расстояние в векторном пространстве, где каждый объект представлен своим вектором признаков. Логика здесь проста: новый объект скорее всего принадлежит к тому же классу, что и большинство соседних с ним объектов. Метрические методы как правило используются для построения классификаторов на основе машинного обучения.

Преимуществом статистических методов является их хорошая математическая обоснованность, недостатком — низкая объясняющая способность. Использование вероятностных оценок позволяет с высокой точностью предсказать к какому классу относится объект, но не позволяет сказать почему. Поэтому результаты статистических методов классификации могут оказаться сложными для понимания и интерпретации результатов.

Недостатком метрических методов является их эвристический характер — они могут дать неточное и неоднозначное решение, но считающееся приемлемым в большинстве практически значимых случаев. Однако при этом они имеют высокую объясняющую способность и поэтому их результаты проще интерпретировать. В простейшем случае можно использовать правило: объект относится к тому же классу, что и большинство его ближайших соседей.

Типичным представителем методов классификации, использующих эту логику, является метод k-ближайших соседей — k-nearest neighbors algorithm (KNN). Метод был впервые разработан Эвелином Фиксом и Джозефом Лоусоном Ходжесом в 1951 году, и позднее развит Томасом Ковером.

Метод относится к классу непараметрических, т.е. не требует предположений о том, из какого статистического распределения была сформирована обучающее множество. Следовательно, классификационные модели, построенные с помощью метода KNN также будут непараметрическими. Это означает, что структура модели не задаётся жёстко изначально, а определяется данными.

Поскольку признаки, на основе которых производится классификация могут иметь различную физическую природу, и, соответственно, диапазоны значений, для улучшения результатов классификации будет полезно выполнить нормализацию обучающих данных.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).
- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью специализированной компьютерной программы SPSS осуществите процедуру кластеризации методом ближайших соседей.

Задание 2. Проинтерпретируйте полученные результаты процедуры кластеризации методом ближайших соседей.

Контрольные вопросы:

1. Сущность метода ближайших соседей
2. Этапы классификации данных методом ближайших соседей
3. Интерпретация данных классификации методом ближайших соседей

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА10

Тема: «Двухэтапный кластерный анализ»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Процедура Двухэтапный кластерный анализ представляет собой средство разведочного анализа для выявления естественного разбиения набора данных на группы (или кластеры), которое без ее применения трудно обнаружить. Алгоритм, используемый этой процедурой, имеет несколько привлекательных особенностей, которые отличают его от традиционных методов кластерного анализа:

Работа с категориальными и непрерывными переменными. Предполагая независимость переменных, можно считать, что категориальные и непрерывные переменные имеют совместное полиномиально-нормальное распределение.

Автоматический выбор числа кластеров. Сравнивая значения критерия отбора модели для различных кластерных решений, процедура может автоматически определить оптимальное число кластеров.

Масштабируемость. Формируя дерево свойств кластеров (СК), которое является компактным представлением информации о наблюдениях, двухэтапный алгоритм позволяет анализировать большие файлы данных.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональный компьютер.

Офисные пакеты: MS Office (версия 7-10 и выше)

Указания по технике безопасности:

– не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

Задания:

1. Демонстрация пройденного материала по темам 1-9.

Контрольные вопросы:

1. Сущность двухэтапного кластерного анализа
2. Этапы классификации данных с помощью двухэтапного кластерного анализа
3. Интерпретация данных классификации с помощью двухэтапного кластерного анализа

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА11

Тема: «Использование факторного анализа в социологии»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Целью факторного анализа является выявление скрытых переменных или факторов, объясняющих структуру корреляций внутри набора наблюдаемых переменных. Факторный анализ часто используется для снижения размерности данных, чтобы найти небольшое число факторов, которые объясняют большую часть дисперсии, наблюдаемой для значительно

большого числа явных переменных. Факторный анализ может также использоваться для формирования гипотез относительно механизмов причинных связей или с целью проверки переменных перед дальнейшим анализом (например, чтобы выявить коллинеарность перед проведением линейного регрессионного анализа).

Рассматриваемая процедура факторного анализа обеспечивает большую гибкость. Доступны семь методов выделения факторов. Доступны пять методов вращения, в том числе прямой облимин и промакс для не ортогональных вращений. Доступны три метода вычисления значений факторов, которые можно сохранить в виде переменных для дальнейшего анализа.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.

- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).

- расположить монитор на уровне глаз и перпендикулярно углу зрения.

- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

– клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. по периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

– продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.

– длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.

– каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. Выберите несколько вопросов, определите и интерпретируйте факторный анализ с помощью метода главных компонент.

Контрольные вопросы:

1. Сущность факторного анализа в социологии
2. Этапы осуществления факторного анализа
3. Проблемы интерпретации факторов

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА12

Тема: «Методы вращения факторов»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Позволяет выбрать метод вращения факторов. Доступные методы: варимакс, прямой облимин, квартимакс, эквимакс и промакс.

Метод варимакс. Ортогональный метод вращения, минимизирующий число переменных с высокими нагрузками на каждый фактор. Этот метод

упрощает интерпретацию факторов.

Метод Прямой облимин. Метод косоугольного (неортогонального) вращения. Самое косоугольное решение соответствует дельте, равной 0 (по умолчанию). По мере того, как дельта отклоняется в отрицательную сторону, факторы становятся более ортогональными. Чтобы изменить задаваемое по умолчанию дельта (равное 0), введите число, меньшее или равное 0,8.

Метод квартимакс. Метод вращения, который минимизирует число факторов, необходимых для объяснения каждой переменной. Этот метод упрощает интерпретацию наблюдаемых переменных.

Метод эквимакс. Метод вращения, объединяющий методы варимакс, упрощающий факторы, и квартимакс, упрощающий переменные. Минимизируется число переменных с большими факторными нагрузками и число факторов, требуемых для объяснения переменной.

Вращение типа промакс. Косоугольное вращение в предположении, что факторы могут коррелировать между собой. Оно производится быстрее, чем вращение типа прямой облимин, поэтому оно полезно для больших наборов данных.

Вывод позволяет запросить вывод повернутого решения, а также графиков нагрузок для первых двух или трех факторов.

Повернутое решение. Чтобы получить повернутое решение, необходимо выбрать метод вращения. Для ортогонального вращения выдаются матрица факторных нагрузок после вращения и матрица преобразования факторов. Для косоугольного вращения выводятся следующие матрицы: факторных нагрузок после вращения, структурная и корреляций факторов.

График факторных нагрузок. Трехмерный график факторных нагрузок для трех первых факторов. Для двухфакторного решения выдается двумерный график. Если выделен только один фактор, график не выдается. Если задано вращение, график выдается для повернутого решения.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. Выберите несколько вопросов, выберите метод вращения факторов для факторного анализа и обоснуйте свой выбор.

Контрольные вопросы:

1. Понятие вращения факторов
2. Выбор процедуры вращения факторов
3. Проблемы вращения факторов
4. Настройки процедуры вращения факторов в программе SPSS

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА13

Тема: «Интерпретация результатов факторного анализа»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Заключительным этапом факторного анализа является процедура содержательной интерпретации выделенных факторов.

Первый вопрос, который здесь возникает, касается количества факторов, которые будут далее содержательно интерпретированы. Как мы уже отметили в предыдущем параграфе, для определения числа факторов используется несколько формальных критериев. Так, при использовании наиболее распространенной в современных исследованиях процедуры главного фактора, как правило, используется критерий Кайзера. Он основан на оценках собственных значений (характеристических корней) матрицы при выделении очередного фактора. Факторы, для которых собственное значение матрицы оказывается меньше единицы, отбрасываются как тривиальные. Важно, однако, иметь в виду, что выбранные для содержательной интерпретации факторы должны в сумме описывать не менее 70—80% общей дисперсии анализируемой корреляционной матрицы. Также отметим, что в конечном итоге выбор факторов зависит не только от этих формальных критериев, но и в еще большей мере — от способности исследователя дать факторам содержательную интерпретацию.

Интерпретация осуществляется на основе исследования матрицы факторных нагрузок или графического представления результатов факторного анализа.

Факторные нагрузки представляют собой стандартизированные весовые коэффициенты для каждой переменной. Они изменяются в диапазоне от -1 до 1. Чем выше по модулю значение факторной нагрузки данной переменной для интерпретируемого фактора, тем в большей степени эта переменная связана с найденным фактором.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим

интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. по периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. Осуществите процедуру факторного и компонентного анализа в программе SPSS, охарактеризуйте и интерпретируйте полученные результаты.

Контрольные вопросы:

1. Проблема выбора количества факторов
2. Какова роль факторного анализа при изучении типологии потребительского поведения?
3. Как используются результаты сохранения факторного анализа в дальнейшем

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА14

Тема: «Модели классификации на основе деревьев решений»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Модели деревьев решений служат для создания систем классификации, которые предсказывают или классифицируют будущие наблюдения на основе набора решающих правил. Если данные разделяются на интересующие вас классы (например, ссуды высокого и низкого риска, подписчики и не-подписчики, голосующие и неголосующие или типы бактерий), можно использовать существующие данные для построения правил, которые можно применять для классификации старых и новых наблюдений с максимальной точностью. Например, можно построить дерево, классифицирующее кредитные риски или намерение покупки на основании возраста и других факторов.

Такой подход, который иногда называют выводом правила, обладает несколькими преимуществами. Во-первых, процесс рассуждения, стоящий в основе модели, виден вам при просмотре дерева. Это важное отличие от способов типа чёрного ящика, в которой внутреннюю логику бывает трудно понять.

Во-вторых, процесс автоматически включает в свои правила только те атрибуты, которые на самом деле имеют значение при принятии решения. Атрибуты, которые не вносят вклада в точность дерева, игнорируются. Это

может дать очень полезную информацию о данных и позволяет сократить данные до релевантных полей перед тренировкой другого метода обучения, такого, как нейросеть.

Слепки модели дерева решений можно преобразовать в собрание правил IF-THEN (набор правил), которое во многих случаях дает информацию в более понятной форме. Презентация дерева решений особенно полезна, когда вы хотите посмотреть, как атрибуты в данных помогут выполнить разбиение или расщепление всей совокупности на подмножества, релевантные для проблемы. Вывод узла Дерево-AS отличается от вывода других узлов дерева решений, так как он включает список правил непосредственно в слепок, без необходимости создания набора правил. Презентация набора правил полезна, если вы хотите посмотреть, как конкретные группы элементов соотносятся с конкретным заключением. Например, следующее правило даёт профиль для группы автомобилей, которых стоит покупать

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.

- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).

- расположить монитор на уровне глаз и перпендикулярно углу зрения.

– мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

– клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

– продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.

– длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.

– каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. Выберите несколько необходимых вопросов и обоснуйте актуальность применения одной или нескольких моделей классификации.

Контрольные вопросы:

1. Охарактеризуйте области применения классификационных моделей в типологическом анализе.

2. Опишите пошаговый алгоритм классификационных моделей в SPSS.

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА15

Тема: «Методы построения деревьев классификации CHAID»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Для классификационного и сегментационного анализа доступно

несколько алгоритмов. Все эти алгоритмы выполняют в основном одно и то же - просматривают все поля из набора данных, чтобы найти то поле, которое даёт наилучшую классификацию или наилучшее предсказание путем разбиения данных на подгруппы. Этот процесс применяется рекурсивно, подгруппы разбиваются на всё меньшие и меньшие блоки до тех пор, пока дерево не будет завершено (что задается определенными критериями остановки). Поля назначения и входные поля, используемые при построении дерева, могут быть непрерывными (числовой диапазон) или категориальными, в зависимости от используемого алгоритма. Если поле назначения непрерывно, генерируется дерево регрессии; если же поле назначения категориальное, генерируется дерево классификации.

Узел CHAID генерирует деревья решений, используя статистику хи-квадрат для определения оптимальных расщеплений. В отличие от узлов дерева C&R и QUEST, CHAID может генерировать не только бинарные деревья, то есть у некоторых расщеплений может быть больше двух ветвей. Входные поля и поле назначения могут быть количественными (числовой диапазон) или категориальными. Исчерпывающий CHAID — это модификация метода CHAID, при котором продельвается более тщательная работа по изучению всех возможных расщеплений для каждого предиктора, но это требует больше времени для вычислений.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.

- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью компьютерной программы SPSS осуществите процедуру построения дерева решений методом CHAID.

Контрольные вопросы:

1. Области применения деревьев решений методом CHAID
2. Преимущества и недостатки метода CHAID
3. Задачи построения деревьев классификации с помощью метода CHAID

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА16

Тема: «Методы построения деревьев решений SQRT»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы , в рамках формируемых компетенций или их части: знает

методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Для классификационного и сегментационного анализа доступно несколько алгоритмов. Все эти алгоритмы выполняют в основном одно и то же - просматривают все поля из набора данных, чтобы найти то поле, которое даёт наилучшую классификацию или наилучшее предсказание путем разбиения данных на подгруппы. Этот процесс применяется рекурсивно, подгруппы разбиваются на всё меньшие и меньшие блоки до тех пор, пока дерево не будет завершено (что задается определенными критериями остановки). Поля назначения и входные поля, используемые при построении дерева, могут быть непрерывными (числовой диапазон) или категориальными, в зависимости от используемого алгоритма. Если поле назначения непрерывно, генерируется дерево регрессии; если же поле назначения категориальное, генерируется дерево классификации.

Узел случайных деревьев аналогичен существующему узлу C&RT; однако узел случайных деревьев сконструирован для обработки большого объема данных с целью создания одного дерева и выводит конечную модель в средстве просмотра вывода, добавленном в SPSS Modeler версии 17. Узел случайных деревьев генерирует дерево решений, которое используется для предсказания или классификации будущих наблюдений. Этот метод использует рекурсивное разделение, чтобы расщепить обучающие записи на сегменты, на каждом шаге минимизируя неоднородность, причем узел дерева считается *чистым*, если все 100% наблюдений в узле попадают в конкретную категорию поля назначения. Входные поля и поля назначения могут быть из числового диапазона или категориальными (номинальными, порядковыми или флагами); все расщепления бинарны (только две подгруппы).

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).
- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. Провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.

– клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. По периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью компьютерной программы SPSS осуществите процедуру построения дерева решений методом SQRT.

Задание 2. Интерпретируйте полученные данные.

Контрольные вопросы:

1. Области применения деревьев решений методом SQRT
2. Преимущества и недостатки метода SQRT
3. Задачи построения деревьев классификации с помощью метода SQRT

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

Тема: «Построение деревьев решений методом QUEST»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

Для классификационного и сегментационного анализа доступно несколько алгоритмов. Все эти алгоритмы выполняют в основном одно и то же - просматривают все поля из набора данных, чтобы найти то поле, которое даёт наилучшую классификацию или наилучшее предсказание путем разбиения данных на подгруппы. Этот процесс применяется рекурсивно, подгруппы разбиваются на всё меньшие и меньшие блоки до тех пор, пока дерево не будет завершено (что задается определенными критериями остановки). Поля назначения и входные поля, используемые при построении дерева, могут быть непрерывными (числовой диапазон) или категориальными, в зависимости от используемого алгоритма. Если поле назначения непрерывно, генерируется дерево регрессии; если же поле назначения категориальное, генерируется дерево классификации.

Узел QUEST предоставляет метод бинарной классификации для построения деревьев решений, разработанный для уменьшения времени обработки, требуемого для анализа больших деревьев C&R, при одновременном подавлении обнаруженного в способах деревьев классификации предпочтения входных полей, допускающих больше расщеплений. Входные поля могут быть в числовом диапазоне (количественными), но поле назначения должно быть категориальным. Все расщепления бинарные.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).
- персональные компьютеры (14-15 шт.) из расчёта 1 ПК на 1 студента.

Основные технические требования: операционная система с графическим интерфейсом, возможность подключения к локальной сети и выхода в Internet; в комплекте: клавиатура, манипулятор «мышь».

Операционные системы и утилиты: о Windows XP/Vista/7/8 или аналогичные.

Офисные пакеты: MS Office (версия 7-10 и выше)

Обязательные приложения: MS Word, MS Excel.

Программные средства для анализа социологических данных: IBM SPSS Statistics (версии 19 и выше) - компьютерная программа для статистической обработки данных.

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.
- не включать без разрешения оборудование.
- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.
- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).
- не трогать провода и разъемы (возможно поражение электрическим током).
- не допускать порчи оборудования.
- не работать в верхней одежде.

До начала работы:

- осмотреть рабочее место (расположение блоков и их состояние).
- расположить монитор на уровне глаз и перпендикулярно углу зрения.
- мышь располагается так, чтобы было удобно работать с ней. провод должен лежать свободно. при работе с мышью по периметру коврика должно оставаться пространство не менее 2-5 сантиметров.
- клавиатуру следует располагать прямо перед пользователем, работающим на компьютере. по периметру оставляется свободное место 2-5 сантиметров.

Рекомендации по организации рабочего времени:

- продолжительность непрерывной работы без регламентированного перерыва не должна превышать 2 часов.
- длительность перерыва должна быть не менее 10 минут, во время которого следует проводить сквозное проветривание с обязательным выходом учащихся из класса.
- каждые 20 минут работы на ПЭВМ следует проводить гимнастику для глаз.

Задания:

Задание 1. С помощью компьютерной программы SPSS осуществите процедуру построения дерева решений методом QUEST.

Задание 2. Интерпретируйте полученные данные.

Контрольные вопросы:

1. Области применения деревьев решений методом QUEST
2. Преимущества и недостатки метода QUEST
3. Задачи построения деревьев классификации с помощью метода QUEST

Список литературы, рекомендуемый к использованию по данной

теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

ПРАКТИЧЕСКАЯ РАБОТА 18

Тема: «Дискриминантный анализ»

Цель: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Формируемые индикаторы компетенции: ИД-2 ПК-1, ИД-2 ПК-2, ИД-4 ПК-3.

Знания и умения, приобретаемые обучающимся в результате освоения темы, в рамках формируемых компетенций или их части: знает методы типологического анализа, применяемые в специализированных компьютерных программах, используемых социологами, применяет методы классификации данных; знает о принципах работы современных информационных технологий, осуществляет сбор, подготовку и анализ эмпирических данных.

Теоретическая часть

При дискриминантном анализе происходит создание прогностической модели для принадлежности к группе. Данная модель строит дискриминантную функцию (или, когда групп больше двух, набор дискриминантных функций) в виде линейной комбинации предикторных переменных, обеспечивающую наилучшее разделение групп. Эти функции строятся по набору наблюдений, для которых их принадлежность к группам известна, и могут в дальнейшем применяться к новым наблюдениям с известными значениями предикторных переменных, но неизвестной групповой принадлежностью.

Если данные переменные являются существенными для разделения двух климатических зон, значения D будут различными для стран с умеренным и тропическим климатом. При использовании метода пошагового отбора переменных может оказаться, что нет необходимости включать в функцию все четыре переменные.

Для каждой переменной: средние значения, стандартные отклонения, однофакторный дисперсионный анализ. Для каждого анализа: M - статистика Бокса, внутригрупповая корреляционная матрица, внутригрупповая ковариационная матрица, ковариационные матрицы для отдельных групп, общая ковариационная матрица. Для каждой канонической дискриминантной функции: собственное значение, процент дисперсии, каноническая корреляция, лямбда Уилкса, хи-квадрат. Для каждого шага: априорные вероятности, коэффициенты функции Фишера, нестандартизованные коэффициенты функции, лямбда Уилкса для каждой канонической функции.

Данные для дискриминантного анализа

Группирующая переменная должна иметь ограниченное число различных категорий, кодированных целыми числами. Независимые переменные, являющиеся номинальными, должны быть перекодированы в фиктивные переменные или переменные контрастов.

Наблюдения должны быть независимыми. Предикторные переменные должны подчиняться многомерному нормальному распределению, а внутригрупповые ковариационные матрицы должны совпадать для всех групп. Групповая принадлежность предполагается взаимоисключающей (т.е. ни одно наблюдение не принадлежит более чем одной группе) и совместно исчерпывающей (т.е. каждое наблюдение принадлежит какой-либо группе). Процедура наиболее эффективна в ситуации, когда группирующая переменная является истинно категориальной; если принадлежность к группе определяется значениями непрерывной переменной (например, высокий IQ (коэффициент интеллекта) низкий IQ), то имеет смысл обратиться к линейной регрессии, чтобы воспользоваться преимуществом большей информативности непрерывной переменной.

Оборудование и материалы:

- мультимедиа проектор (в комплекте: кабель питания, кабели для подключения к компьютеру, видео и аудио источникам).

- персональный компьютер.

Офисные пакеты: MS Office (версия 7-10 и выше)

Указания по технике безопасности:

- не входить в помещение, где находится вычислительная техника без разрешения преподавателя или лаборанта кабинета.

- не включать без разрешения оборудование.

- при несчастном случае или поломке оборудования обратиться к преподавателю или лаборанту кабинета.

- знать, где находится пульт выключения оборудования (выключатель, красная кнопка, рубильник).

- не трогать провода и разъемы (возможно поражение электрическим током).

- не допускать порчи оборудования.

- не работать в верхней одежде.

Задания:

1. Демонстрация пройденного материала по темам 1-17.

Контрольные вопросы:

1. Вопросы задаются после презентации собственных отчетов.

Список литературы, рекомендуемый к использованию по данной теме:

основная: 1-4

дополнительная: 1-3

интернет-ресурсы: 1-6

Перечень литературы необходимый для освоения дисциплины

Перечень основной литературы:

1. Воронова, Л. И. Machine Learning: регрессионные методы интеллектуального анализа данных Электронный ресурс: Учебное пособие / Л. И. Воронова, В. И. Воронов. - Machine Learning: регрессионные методы интеллектуального анализа данных, 2024-02-26. - Москва: Московский технический университет связи и информатики, 2018. - 82 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 2227-8397, экземпляров неограниченно

2. Храмов, А. Г. Методы и алгоритмы интеллектуального анализа данных Электронный ресурс / Храмов А. Г.: учебное пособие. - Самара: СамГУ, 2019. - 176 с. - Рекомендовано редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» в качестве учебного пособия для обучающихся по основной образовательной программе высшего образования по направлению подготовки 01.04.02 Прикладная математика и информатика. - ISBN 978-5-7883-1414-3, экземпляров неограниченно

3. Агалаков, С. А. Статистические методы анализа данных: учебное пособие / С.А. Агалаков; Министерство образования и науки РФ; Омский государственный университет им. Ф. М. Достоевского. - Омск: ОмГУ им. Ф.М. Достоевского, 2017. - 92 с.: табл., граф., схем., ил. - <http://biblioclub.ru/>. - Библиогр. в кн. - ISBN 978-5-7779-2187-1, экземпляров неограниченно

4. Тарасов, И. Е. Статистический анализ данных в информационных системах Электронный ресурс / Тарасов И. Е.: учебно-методическое пособие. - Москва: РТУ МИРЭА, 2020. - 96 с., экземпляров неограниченно

Перечень дополнительной литературы:

1. Глебов, В. И. Практикум по математической статистике. Проверка гипотез с использованием Excel, MatCalc, R и Python: учебное пособие / В. И. Глебов, С. Я. Криволапов. - Практикум по математической статистике. Проверка гипотез с использованием Excel, MatCalc, R и Python, 2026-04-02. - Электрон. дан. (1 файл). - Москва: Прометей, 2019. - 86 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-907100-66-4, экземпляров неограниченно

2. Лихтенштейн, В. Е. Информационные технологии в бизнесе. Том 2. Применение системы Decision в решении прикладных экономических задач Электронный ресурс: Учебное пособие / В. Е. Лихтенштейн, Г. В. Росс. - Информационные технологии в бизнесе. Том 2. Применение системы Decision в решении прикладных экономических задач, 2028-01-26. - Саратов: Ай Пи Эр Медиа, 2018. - 420 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4486-0283-2, экземпляров неограниченно

3. Мартышенко, С. Н. Автоматизация анализа данных в исследовании социально-экономических процессов. Монография Электронный ресурс / Мартышенко С. Н. - Владивосток: ВГУЭС, 2019. - 164 с. - ISBN 978-5-9736-0583-4, экземпляров неограниченно

Перечень ресурсов информационно-телекоммуникационной сети

«Интернет», необходимых для освоения дисциплины:

1 Единый архив экономических и социологических данных // [Электрон. ресурс]. – Режим доступа: <http://sophist.hse.ru>

2. Наследов А. Д. SPSS: Компьютерный анализ данных в психологии и социальных науках. – СПб.: Питер, 2005. —416 с.: ил. [Электронный ресурс]. – Режим доступа: <http://bookre.org/reader?file=720696&pg=1>

3. НАФИ – аналитический центр [Электрон. ресурс]. - Режим доступа: <https://nafir.ru/upload/spss/>

4. Осваиваем синтаксис. Raynald's SPSS Tools [Электрон. ресурс] - Режим доступа: <http://spsstools.net/ru/syntax/learning-syntax/>

5. Пациорковский В.В., Пациорковская В.В. SPSS для социологов. Учебное пособие. - М.: ИСЭПН РАН. 2005.- 433 с. [Электрон. ресурс] - Режим доступа: <http://csl.isc.irk.ru/BD/Books/spss%20для%20социологов.pdf>

6. Университетская библиотека ONLINE (ЭБС)

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**
**Федеральное государственное автономное образовательное учреждение
высшего образования
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО
ОРГАНИЗАЦИИ И ПРОВЕДЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ
ПО ДИСЦИПЛИНЕ**

Типологический анализ в социологии

Направление подготовки 39.03.01 Социология
Направленность (профиль) – «Аналитика и консалтинг»
Форма обучения – Очная, заочная
Год начала обучения – 2026 г.

Ставрополь, 2026

СОДЕРЖАНИЕ

Введение

1. Общая характеристика самостоятельной работы
2. Методические рекомендации по изучению теоретического материала

Список рекомендуемой литературы

Методические указания для самостоятельной работы разработаны на основе программы дисциплины «Типологический анализ в социологии»

Современные стандарты подготовки выпускников вузов предусматривают значительный объем внеаудиторной научной работы студентов. Освоение программы дисциплины предполагает получение как теоретических знаний, так и формирование навыков практической работы. Поэтому самостоятельная работа в рамках курса ориентирована как на теоретический, так и на практический аспект.

Самостоятельная работа студентов – важнейшая составная часть занятий, необходимая для полного усвоения программы курса.

Цель освоения дисциплины: формирование общепрофессиональных и профессиональных компетенций будущего бакалавра по направлению подготовки 39.03.01 Социология в области типологизации и классификации социологических данных.

Задачи освоения дисциплины:

- познакомить с основными методами анализа типологического анализа данных, применяемыми в социологии;
- научиться вести практические расчеты при решении типовых для социологии задач типологизации данных;
- выработать умения проводить интерпретацию полученных при расчетах данных.

Задачами СРС являются:

- развитие и привитие навыков самостоятельной учебной работы и формирование потребностей в самообразовании и социализации;
- освоение, углубление содержания и основных положений курса, выносимых на самостоятельное изучение студента; в ходе составления опорного конспекта, в ходе подготовки к практическим занятиям
- использование материала, собранного и полученного в ходе самостоятельных занятий при подготовке и написании докладов, рефератов, для эффективной подготовки к итоговому контролю.

Самостоятельная работа студента – это особым образом организованная деятельность, включающая в свою структуру такие компоненты, как:

- уяснение цели и поставленной учебной задачи;
- четкое и системное планирование самостоятельной работы;
- поиск необходимой учебной и научной информации;
- освоение собственной информации и ее логическая переработка;
- использование методов исследовательской, научно-исследовательской работы для решения поставленных задач;
- выработка собственной позиции по поводу полученной задачи;
- представление, обоснование и защита полученного решения;
- проведение самоанализа и самоконтроля.

Основными формами реализации самостоятельной работы

являются:

- работа в специализированной компьютерной программе SPSS,
- выполнение исследовательского проекта по заданной проблематике,
- подготовка презентации.

Важным видом самостоятельной работы является подготовка к дискуссии, проведение микроисследований по теме курса. Подготовка к дискуссии должна содержать письменный доклад, который должен состоять из следующих частей: введение, первая часть (реферативная), вторая часть (исследовательская), заключение. Введение - это постановка проблемы, обоснование темы, целеполагание, формулировка задач исследования. Первая часть - состояние исследуемой области, обзор литературы, прочитанной по данной проблеме, выводы. Вторая часть - собственное исследование. Заключение - выводы, к которым пришел студент в результате изучения состояния вопроса и собственного исследования. В конце доклада обязательно должен быть представлен список литературы (не менее пяти-шести названий). Сообщение длится примерно 5-10 минут. При оценке учитывается объем проделанной работы, компетентность студента в избранной области, самостоятельность в подходах, суждениях, выводах; оригинальность решений, культура оформления работы, грамотность.

Их содержание должно показать умение анализировать философско-методологические и научные тексты, обосновывать своё отношение к изложенным в них оценкам, составлять краткие конспекты, тезисы прочитанного, умение применять полученные знания для постановки и решения исследовательских задач, связанных с изучением студентами той или иной области природы и культуры.

В предлагаемом пособии по самостоятельной контролируемой работе студентов представлены задания для самостоятельной работы по отдельным темам курса. Разнообразие заданий позволяет учитывать индивидуальные особенности студентов при организации самостоятельной работы.

Текущий контроль осуществляется в процессе индивидуальных и групповых консультаций, собеседований, в процессе анализа и самоанализа самостоятельной работы. Образовательные продукты (рефераты, словарь научных понятий, опорные конспекты, эссе, мультимедийные доклады, аннотированные перечни статей по темам курса и др.) составляют портфолио («портфель достижений»), который представляется в оформленном виде к зачету.

Общая характеристика самостоятельной работы студента при изучении дисциплины «Типологический анализ в социологии»

Самостоятельная работа студентов наряду с аудиторной представляет одну из форм организации учебного процесса и является его существенной частью. Для ее успешной организации традиционно осуществляется планирование целей, содержания, объема самостоятельной работы, происходит систематический контроль качества, своевременности ее выполнения и получаемых при этом результатов каждым студентом.

Самостоятельная работа понимается как целенаправленная, планируемая деятельность студентов, выполняемая по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

В процессе освоения учебной дисциплины «Типологический анализ в социологии» предполагается аудиторная и неаудиторная самостоятельная работа.

Аудиторная самостоятельная работа по дисциплине выполняется на учебных занятиях (лекциях, лабораторных занятиях) под руководством преподавателя и по его заданию.

Внеаудиторная самостоятельная работа студентов – планируемая учебная, учебно-исследовательская, научно-исследовательская работа студентов, выполняется во внеаудиторное время по заданию и при методическом руководстве и контроле преподавателя, но без его непосредственного участия.

Самостоятельная работа студента по дисциплине «Типологический анализ в социологии» включает в себя:

- подготовку к аудиторным занятиям (лекциям, лабораторным) и выполнение соответствующих заданий;

- работу над отдельными темами учебных дисциплин (модулей) в соответствии с учебно-тематическими планами;

- подготовку к практикам, лабораторным и выполнение заданий, предусмотренных практиками;

- выполнение письменных проверочных работ;

- подготовку ко всем видам промежуточных и итоговых контрольных испытаний;

- проведение статистически-математического анализа эмпирических данных по выбранной теме в специализированной компьютерной программе SPSS;

- подготовка и написание проекта (отчета) по выбранной тематике после проведения анализа эмпирических данных;

- подготовка презентации и доклада по выполнению проекта (отчета);

- другие виды деятельности, организуемой и осуществляемой Университетом, институтами, кафедрами и органами студенческого самоуправления.

Отчеты о результатах самостоятельной работы по дисциплине «Типологический анализ в социологии» могут быть представлены следующими формами:

- текущий контроль усвоения знаний на основе оценки устного ответа на вопрос, сообщения, доклада и т.п. (на лабораторных занятиях);

- решение задач;

- конспект, выполненный по теме, изучаемой самостоятельно;

- отчет, дневник наблюдения, таблицы, схемы и т.п.;

- рейтинговая система оценки знаний студентов по разделам изучаемой дисциплины;

- отчет о научно-исследовательской работе (ее этапе, части работы и т.п.);

статья об учебно-исследовательской работе, тезисы выступления, публикации в научном, научно-популярном, учебном издании по итогам самостоятельной учебной и научно-исследовательской работы, опубликованные по решению кафедры и др.

Методические рекомендации по изучению теоретического материала

При изучении дисциплины у студентов должен вырабатываться рационально-критический подход к изучаемым проблемам и явлениям. Это включает понимание того, что со временем ряд информационных и теоретических материалов устаревает, требуя критического отношения. С другой стороны, каждый текущий вопрос имеет свою историю, которую тоже полезно знать. Каждое событие может иметь разные интерпретации, поэтому слова, сказанные много лет назад, могут иметь важное значение.

Чтобы понять содержание материала, нужно уметь его прочитывать. Начинать следует с предварительного просмотра, в ходе которого ознакомиться с названием работы, с аннотацией, оглавлением, предисловием. Часто замысел работы ясен уже при ознакомлении с ее названием. Но особенно интересен просмотр оглавления, в результате которого становится ясным развитие мысли автора. Неплохо было бы появившиеся при этом мысли зафиксировать на бумаге.

Просматривая текст оглавления, нужно остановиться на тех главах, которые представляют для вас особый интерес, бегло ознакомиться с ними, составляя в общих чертах свое представление о них. Цель этого действия – найти места, относящиеся к искомой теме, определив при этом, что ценного в каждом из них.

Следующий этап – прочтение выделенных мест с фиксацией самых главных сведений. При этом надо четко и ясно осознавать цель чтения, постоянно держа ее перед собой: по какому вопросу нужна информация, для чего нужна, ее характер и т.д. необходимо менять режим чтения – от беглого вдумчивого – в зависимости от ценности информации, останавливаясь там, где это требуется для глубокого понимания текста.

Следует научиться определять структуру текста по соподчиненности его частей, учитывая взаимосвязь текста с рисунками, сносками, примечаниями и таблицами. Все это поможет пониманию текста при беглом ознакомлении с ним. Так вырабатывается способность при прочтении сразу понимать смысл и значение новой информации.

Многие книги и статьи имеют в своем аппарате списки литературы, которые дают возможность пополнить информационную осведомленность о дополнительной литературе по данному вопросу.

Отдельный этап прочтения – ведение записей прочитанного. Существует несколько видов записей: план, выписки, тезисы, аннотация, резюме, конспект.

Планом удобно пользоваться при подготовке к устному выступлению по выбранной теме. Каждый пункт плана должен раскрывать одну из сторон

избранной темы, а весь план должен охватывать ее целиком.

Тезисы предполагают сжатое изложение основных положений текста в форме утверждения или отрицания. Они являются более совершенной формой записей и представляют основу для дискуссии. К тому же их легко запомнить.

Аннотация – краткое изложение содержания – дает общее представление о работе.

Резюме кратко характеризует выводы, главные итоги произведения.

Конспект является наиболее распространенной формой ведения записей. Основную ткань конспекта составляют тезисы, дополненные доказательствами и рассуждениями. Конспект может быть текстуальным, свободным или тематическим. Текстуальный представляет собой цитатник с сохранением логики работы и структуры текста. Свободный конспект основан на изложении материала в том порядке, который более удобен автору. В этом смысле конспект представляет собирание воедино мыслей, разбросанных по всей книге. Тематический конспект может быть составлен по нескольким источникам, где за основу берется тема, интерпретируемая по-разному.

Экономии времени дает использование при записях различного рода сокращений, аббревиатуры и т.д. многие используют для регистрации исследуемых тем систему карточек. Преимущество карточек в том, что тема там излагается очень сжато, и они очень удобны в использовании, т.к. их можно разложить на столе, перегруппировать и без труда найти искомую тему.

Лекционные занятия, проводимые в интерактивной форме, предполагают совместное обсуждение методологических и теоретических положений учебного курса, отработку умения работать с научной литературой, анализируя и обобщая социологические факты, личные примеры и предложенные ситуации. Цель таких занятий – усвоение системы теоретических знаний по основным разделам науки. Для этого по каждому занятию указана литература, непосредственно относящаяся к содержанию темы. С помощью нее студент заранее готовится к занятию по предложенному плану с тем, чтобы в ходе лекции участвовать в коллективном анализе различных подходов к категориям и проблемам социологии, в формулировании дефиниций и в научном обосновании собственных позиций по проблемам. Подготовка к лекционным занятиям предусматривает ознакомление с основными положениями по теме, усвоение нормативной лексики, предложенной к разделу через критическую работу с литературой и научными социологическими текстами. Данная работа необходима студенту для ведения научной дискуссии на занятии.

Работа с конспектом лекций

Просмотрите конспект сразу после занятий. Отметьте материал конспекта лекций, который вызывает затруднения для понимания. Попытайтесь найти ответы на затруднительные вопросы, используя предлагаемую литературу. Если самостоятельно не удалось разобраться в материале, сформулируйте вопросы и обратитесь на текущей консультации или на ближайшей лекции за помощью к преподавателю.

Каждую неделю отводите время для повторения пройденного

материала, проверяя свои знания, умения и навыки по контрольным вопросам и тестам.

Литературу необходимо конспектировать и давать сжатые письменные ответы на вопросы. В них следует отразить содержание главных понятий и связь между ними, дать собственную оценку тем или иным концепциям, выявить связь изучаемого вопроса с современностью, с будущей профессиональной деятельностью. По неясным вопросам необходимо обращаться к преподавателю.

Текст выступления на занятии должен быть написан самим студентом (уже переработанный). Выступления, на которых студент ограничивается только чтением уже опубликованного текста, преподавателем не учитываются.

По пропущенным занятиям студент отчитывается в устной или письменной форме. Преподаватель оценивает подготовку студента к занятию по следующим критериям:

- Наличие конспектов, лекций;
- Понимание главных понятий, их связей;
- Умение мыслить самостоятельно, логично, последовательно, доказательно;
- Умение оценивать идеи, концепции, обосновывать свою теоретическую позицию, связывать теорию с практикой, приводить примеры, приводить аналогии.

На занятия выносятся узловые темы, запланированные в программе учебной дисциплины, освоение которых определяет качество профессиональной подготовки будущих бакалавров.

Цель занятия – формирование знаний по теме, развитие самостоятельности мышления и творческой активности обучающихся путем равноправного и активного их участия в обсуждении рассматриваемых вопросов.

Методические указания по выполнению лабораторных работ

Методические указания по выполнению лабораторных работ по дисциплине «Типологический анализ в социологии» являются самостоятельным элементом учебно-методического комплекса дисциплины и представлены в виде отдельного учебно-методического материала.

Вопросы для коллоквиума

Тема 1.

1. Сущность и смысл типологического анализа данных
2. Целесообразность использования методов типологического анализа
3. Функции типологического анализа данных

Тема 5.

1. Понятие иерархического кластерного анализа
2. Целесообразность применения иерархического кластерного анализа
3. Недостатки иерархического кластерного анализа

4. Интерпретация результатов иерархического кластерного анализа
5. Этапы иерархического кластерного анализа

Тема 7.

1. Характеристика методов визуализации кластерного анализа
2. Интерпретация результатов построения дендрограммы
3. Интерпретация результатов построения графика «ствол-лист»

Тема 8.

1. Значение методов снижения размерности в типологическом анализе
2. Целесообразность использования методов снижения размерности в типологическом анализе
3. Этапы реализации методов снижения размерности

Тема 12.

1. Понятие вращения факторов
2. Выбор процедуры вращения факторов
3. Проблемы вращения факторов
4. Настройки процедуры вращения факторов в программе SPSS

Перечень тем для дискуссии

Тема 2.

1. Предпосылки появления типологического анализа
2. Этапы становления и развития типологического анализа в социологии

Тема 4.

1. Основные понятия и методы кластерного анализа
2. Задачи кластерного анализа
3. Преимущества и недостатки кластерного анализа
4. Этапы проведения кластерного анализа
5. Особенности исходных данных для кластерного анализа

Тема 9.

1. Сущность метода ближайших соседей
2. Этапы классификации данных методом ближайших соседей
3. Интерпретация данных классификации методом ближайших соседей

Тема 10.

1. Сущность двухэтапного кластерного анализа
2. Этапы классификации данных с помощью двухэтапного кластерного анализа
3. Интерпретация данных классификации с помощью двухэтапного кластерного анализа

Тема 11.

1. Сущность факторного анализа в социологии
2. Этапы осуществления факторного анализа
3. Проблемы интерпретации факторов

Тема 13.

1. Проблема выбора количества факторов
2. Какова роль факторного анализа при изучении типологии потребительского поведения?
3. Как используются результаты сохранения факторного анализа в дальнейшем

Тема 15.

1. Области применения деревьев решений методом CHAID
2. Преимущества и недостатки метода CHAID
3. Задачи построения деревьев классификации с помощью метода CHAID

Тема 16.

1. Области применения деревьев решений методом SQRT
2. Преимущества и недостатки метода SQRT
3. Задачи построения деревьев классификации с помощью метода SQRT

Тема 17.

1. Области применения деревьев решений методом QUEST
2. Преимущества и недостатки метода QUEST
3. Задачи построения деревьев классификации с помощью метода QUEST