

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра социальной информатики

**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ПРИКЛАДНОМ
СОЦИОЛОГИЧЕСКОМ ИССЛЕДОВАНИИ «ОТНОШЕНИЕ
МОЛОДЕЖИ ГОРОДА САРАТОВА К ЗДОРОВОМУ ОБРАЗУ ЖИЗНИ»
НА ПРИМЕРЕ SPSS**

(автореферат бакалаврской работы)

студент 5 курса 531 группы
направления 09.03.03 - Прикладная информатика
профиль Прикладная информатика в социологии
Социологического факультета
Орлова Леонида Ильича

Научный руководитель
кандидат физико-математических наук,
доцент

_____ Л.Б. Тяпаев
подпись, дата

Зав. кафедрой
кандидат социологических наук, доцент

_____ И.Г. Малинский
подпись, дата

Саратов 2024

ВВЕДЕНИЕ

Актуальность темы исследования. В эпоху цифровизации данные становятся ключевым активом для многих организаций и исследовательских проектов. Они не только растут по объему, но и становятся все более разнообразными и сложными. В такой среде, где информации множество, но она не всегда проста для понимания, инструменты анализа данных играют критическую роль.

IBM SPSS Statistics - это один из ведущих пакетов программного обеспечения для статистического анализа данных. Его возможности включают в себя широкий спектр методов анализа, начиная от базовых статистических тестов и заканчивая сложными многомерными моделями. Несмотря на его функциональность, принятие решений на основе полученных данных все равно может быть вызовом из-за объема информации и необходимости ее интерпретации.

Параллельно с этим, технологии искусственного интеллекта, особенно модели глубокого обучения, становятся все более распространенными и мощными. Модель ChatGPT (Generative Pre-trained Transformer), разработанная OpenAI, является впечатляющим примером такой технологии. Она обучается на огромном объеме текстовых данных и способна генерировать качественный текст, поддерживать диалоги и даже предсказывать следующие слова или фразы в тексте. Таким образом, ChatGPT представляет собой уникальный инструмент для работы с текстом, который может сэкономить время и усилия при анализе больших объемов информации.

Цель нашего исследования - исследовать, как можно использовать модель ChatGPT в среде IBM SPSS Statistics для улучшения процесса анализа данных. Мы намерены изучить различные способы интеграции ChatGPT с SPSS, чтобы облегчить процесс интерпретации данных, сделать его более эффективным и точным. В этом контексте, мы сфокусируемся на разработке методов автоматизации анализа данных с использованием модели ChatGPT, а также на

оценке эффективности и практической применимости такого подхода в реальных проектах и исследованиях.

В своем исследовании я буду использовать результаты ранее проведенного исследования "Отношение молодежи города Саратова к здоровому образу жизни" при помощи ChatGPT. Данные и выводы из этого исследования станут основой для моего анализа, позволяя мне глубже понять и интерпретировать современные тенденции и взгляды среди молодежи на тему здорового образа жизни. Используя возможности ChatGPT, я смогу проводить более детальный и структурированный анализ, что, в свою очередь, повысит точность и релевантность моих выводов. Результаты ранее проведенного исследования будут продемонстрированы в ПРИЛОЖЕНИИ А.

Объектом исследования является процесс анализа данных с использованием программного обеспечения IBM SPSS Statistics в контексте совместного применения модели искусственного интеллекта ChatGPT.

Предметом исследования является разработка и апробация методов интеграции модели ChatGPT в среду SPSS с целью оптимизации процесса анализа данных и интерпретации результатов.

Возможности интеграции: Анализ технических и методологических аспектов взаимодействия модели ChatGPT с программным обеспечением SPSS для обеспечения совместимости и эффективности их совместного использования.

Методы анализа данных: Разработка и апробация методов автоматизации анализа данных с использованием модели ChatGPT, включая генерацию текстовых описаний результатов, формулировку вопросов для дальнейшего исследования и облегчение взаимодействия с данными на естественном языке.

Эффективность применения: Оценка практической эффективности и применимости разработанных методов в реальных проектах и исследованиях, а также сравнение результатов с традиционными методами анализа данных в SPSS.

Исследование направлено на определение потенциала и перспектив совместного использования модели ChatGPT и программного обеспечения SPSS для улучшения процесса анализа данных и принятия информированных решений на его основе.

Цель исследования:

Провести комплексное сравнительное исследование эффективности анализа данных с использованием программного обеспечения IBM SPSS Statistics и модели искусственного интеллекта ChatGPT, а также оценить их применимость и достоверность результатов в контексте статистического анализа данных.

Для достижения поставленной цели необходимо решить следующие **задачи:**

1. Проверить эффективность использования модели ChatGPT в среде IBM SPSS Statistics для упрощения процесса интерпретации результатов анализа данных путем автоматизации генерации текстовых описаний результатов;
2. Оценить влияние применения методов автоматизации анализа данных с использованием модели ChatGPT в SPSS на сокращение времени, затрачиваемого на выполнение статистических задач и формулировку выводов. Выявить склонность к курению у людей, занимающихся спортом
3. Исследовать воздействие взаимодействия с данными на естественном языке с помощью модели ChatGPT в среде SPSS на уровень понимания данных и эффективность коммуникации результатов между исследователями и заинтересованными сторонами.
4. Определить, как совместное использование модели ChatGPT и SPSS влияет на повышение точности и качества анализа данных за счет обогащения процесса интерпретации результатов текстовыми описаниями и автоматической формулировкой вопросов для дальнейшего исследования.
5. Провести анализ эффективности работы с данными с использованием модели ChatGPT в среде SPSS с целью выявления практической

применимости разработанных методов и их влияния на процесс анализа данных и принятия решений в реальных проектах и исследованиях.

Исходя из изложенных задач, выделим ряд гипотез:

1. Использование модели ChatGPT в среде SPSS упрощает процесс интерпретации результатов анализа данных путем автоматизации генерации текстовых описаний результатов.

2. Применение методов автоматизации анализа данных с использованием модели ChatGPT в SPSS сокращает время, затрачиваемое на выполнение статистических задач и формулировку выводов.

3. Взаимодействие с данными на естественном языке с помощью модели ChatGPT в SPSS повышает понимание данных и упрощает коммуникацию результатов между исследователями и заинтересованными сторонами.

4. Совместное использование модели ChatGPT и SPSS приводит к повышению точности и качества анализа данных за счет обогащения процесса интерпретации результатов текстовыми описаниями и автоматической формулировки вопросов для дальнейшего исследования.

5. Использование модели ChatGPT в среде SPSS способствует повышению эффективности работы с данными путем сокращения времени, затрачиваемого на рутинные задачи анализа данных, и увеличения концентрации на интерпретации результатов и принятии решений.

Практическая значимость исследования, направленного на использование модели ChatGPT в среде IBM SPSS Statistics, может быть определена следующим образом:

Сокращение времени выполнения задач: Внедрение автоматизированных методов анализа данных с использованием ChatGPT может значительно сократить время, затрачиваемое на выполнение статистических задач, что особенно важно в условиях сжатых сроков и высокой нагрузки на исследователей.

Повышение доступности анализа данных: Использование модели ChatGPT позволяет пользователям, не обладающим специализированными знаниями в области статистики или программирования, легко и интуитивно взаимодействовать с данными и проводить статистический анализ на естественном языке, что увеличивает доступность анализа данных для широкого круга пользователей.

Структура выпускной квалификационной работы представлена введением, двумя разделами, заключением, списком использованных источников и одним приложением.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первом разделе «Теоретико-методологические и методические основы SPSS и ChatGPT» представлен подробный обзор программного обеспечения SPSS и ChatGPT, охватывающий их основные функции, историю развития и ключевые возможности.

SPSS (Statistical Package for the Social Sciences):

1. Основные функции и возможности:

- **Анализ данных:** широкий спектр статистических методов (дескриптивная статистика, t-тесты, ANOVA, корреляционный и регрессионный анализ, кластерный и факторный анализ).
- **Визуализация данных:** создание графиков и диаграмм для наглядного представления данных.
- **Импорт и экспорт данных:** поддержка различных форматов данных (текстовые файлы, Excel, базы данных).
- **Манипуляции с данными:** создание новых переменных, преобразование, фильтрация, сортировка данных.
- **Автоматизация и скриптинг:** создание скриптов и макросов для автоматизации анализа данных.

2. История развития SPSS:

- Создан в 1965 году Норманом Наем и Дейлом Вентом для анализа статистической информации.

- Развитие программы от первых версий для IBM до современных версий для Windows.
- Улучшения в интерфейсе и функциональности, адаптация к новым вычислительным платформам и расширение возможностей анализа.

ChatGPT:

1. История развития:

- **GPT-1:** первая версия, способная вести диалоги на естественном языке, но с ограниченными возможностями контекста и точности.
- **GPT-2:** улучшенная версия с большей емкостью для контекста, сниженной вероятностью нежелательного контента и улучшенным качеством генерации текста.
- **GPT-3:** самая мощная версия с 175 миллиардами параметров, превосходным качеством генерации текста и адаптивностью к различным стилям общения.

2. Усовершенствования и возможности:

- **GPT-3:** более высокий уровень точности и разнообразия генерации текста, возможность выполнения различных задач (перевод, ответы на вопросы, генерация кода).
- **GPT-4:** дальнейшие улучшения в контекстуальном понимании и производительности при работе с длинными последовательностями текста, что позволяет более точно и последовательно отвечать на запросы пользователей.

Таким образом, в первый раздел предоставляет всесторонний обзор двух мощных инструментов для анализа данных и ведения диалогов на естественном языке, описывая их эволюцию, функциональность и основные преимущества.

Во втором разделе рассматривается отношение молодежи города Саратова к здоровому образу жизни. Анализ статистических данных при помощи SPSS с использованием модели ChatGPT PLUS

Практическая часть данного исследования посвящена интеграции модели ChatGPT в SPSS для оптимизации анализа данных. Исследование охватывает следующие аспекты:

1. Анализ возможностей интеграции:

- Изучение технических и методологических аспектов взаимодействия ChatGPT с SPSS.
- Оценка совместимости и эффективности совместного использования.
- Анализ различных подходов к интеграции.

2. Методы анализа данных:

- Разработка и апробация методов автоматизации анализа данных с использованием ChatGPT.
- Генерация текстовых описаний результатов, формулировка вопросов и взаимодействие с данными на естественном языке.
- Упрощение и ускорение анализа данных для пользователей без глубоких знаний в статистике.

3. Оценка эффективности применения:

- Тестирование разработанных методов в реальных проектах.
- Сравнение результатов анализа данных, полученных с использованием ChatGPT и традиционных методов SPSS.
- Оценка точности, скорости и удобства использования.

Исследование направлено на выявление потенциала и перспектив совместного использования ChatGPT и SPSS для улучшения анализа данных и принятия решений.

Пример практической работы: Для анализа предоставлены данные на тему "Отношение молодежи города Саратова к здоровому образу жизни". ChatGPT используется для анализа данных и генерации кодов для SPSS. Пример анализа

показывает, что пол не оказывает значимого влияния на здоровый образ жизни (результаты χ^2 -теста).

Процесс интеграции ChatGPT с SPSS:

1. **Подготовка данных:** предоставление и расшифровка данных для ChatGPT.
2. **Анализ данных:** проведение статистических тестов (например, χ^2 -тест) с использованием ChatGPT.
3. **Генерация кода:** получение кода для SPSS от ChatGPT и выполнение его в SPSS.
4. **Интерпретация результатов:** анализ таблиц сопряженности и интерпретация данных с помощью ChatGPT.

Выводы:

- Интеграция ChatGPT с SPSS позволяет автоматизировать и ускорить процесс анализа данных.
- Генерация текстовых отчетов и кодов для SPSS делает анализ более доступным для пользователей.
- Результаты, полученные с помощью ChatGPT, сопоставимы с результатами традиционного анализа в SPSS, что подтверждает эффективность предложенных методов.

ЗАКЛЮЧЕНИЕ

На основе проведенного комплексного сравнительного исследования эффективности анализа данных с использованием программного обеспечения IBM SPSS Statistics и модели искусственного интеллекта ChatGPT, а также оценки их применимости и достоверности результатов в контексте статистического анализа данных, можно сделать следующие выводы:

1. Эффективность использования ChatGPT в среде IBM SPSS Statistics:

Модель ChatGPT продемонстрировала высокую эффективность в упрощении процесса интерпретации результатов анализа данных. Автоматизация генерации текстовых описаний результатов позволила значительно сократить время, необходимое для анализа и интерпретации

данных, что особенно ценно для исследователей с ограниченным опытом в статистике.

2. Влияние автоматизации на время выполнения задач:

Применение методов автоматизации с использованием ChatGPT в SPSS позволило сократить время, затрачиваемое на выполнение статистических задач и формулировку выводов. Это было достигнуто благодаря способности модели быстро обрабатывать большие объемы данных и предоставлять осмысленные текстовые описания результатов, что освобождает исследователей от рутинных операций.

3. Влияние на понимание данных и эффективность коммуникации:

Взаимодействие с данными на естественном языке с помощью ChatGPT улучшило уровень понимания данных среди исследователей. Модель обеспечила более доступное и интуитивное объяснение сложных статистических результатов, что способствовало эффективной коммуникации между исследователями и заинтересованными сторонами, облегчая принятие решений на основе полученных данных.

4. Повышение точности и качества анализа данных:

Совместное использование ChatGPT и SPSS способствовало повышению точности и качества анализа данных. Обогащение процесса интерпретации результатов текстовыми описаниями и автоматическая формулировка вопросов для дальнейшего исследования позволили глубже понять исследуемые явления и выявить скрытые закономерности.

5. Практическая применимость и влияние на процесс принятия решений:

Методы, разработанные для работы с данными с использованием ChatGPT в среде SPSS, продемонстрировали высокую практическую применимость. Они значительно упростили процесс анализа данных и принятия решений в реальных проектах и исследованиях, обеспечивая оперативность, точность и полноту анализа.

Исследование показало, что интеграция моделей искусственного интеллекта, таких как ChatGPT, с традиционными инструментами статистического анализа, такими как IBM SPSS Statistics, может значительно улучшить эффективность, точность и качество анализа данных. Это сочетание технологий открывает новые возможности для автоматизации сложных аналитических задач и улучшения понимания результатов, что в конечном итоге способствует более информированному принятию решений и повышению общей производительности исследовательской работы.