



DATE: 30.01.2024

THE LINGUISTIC DIFFERENCES BETWEEN AI-GENERATED REVIEWS AND REVIEWS WRITTEN BY GERMAN NATIVE SPEAKERS WITH ENGLISH AS THEIR SECOND LANGUAGE

LEONORE SCHÖNAU : 227926 (7TH SEM)

ALVA TERNOW :

AAKASH GOYAL : 229975 (7TH SEM)

AZEEZAT MOSUNMADE MUSTAPHA :

SUPERVISORS:

PROF DR. CLAUD WEIHS

PROF DR. SARAH BUSCHFELD

TABLE OF CONTENTS

Introduction

Previous research, Research question & Hypotheses

Method & Data

Statistical Analysis

Results

Limitations & Conclusion

INTRODUCTION

WHY IS OUR STUDY IMPORTANT?

- **AI is influencing many parts of our daily life online**
- **Reviews are necessary to determine whether a product or a company is worth the money**
- **Fake reviews are tricking customers into buying**
- **Difficult to differentiate between real and fake reviews**

PREVIOUS RESEARCH

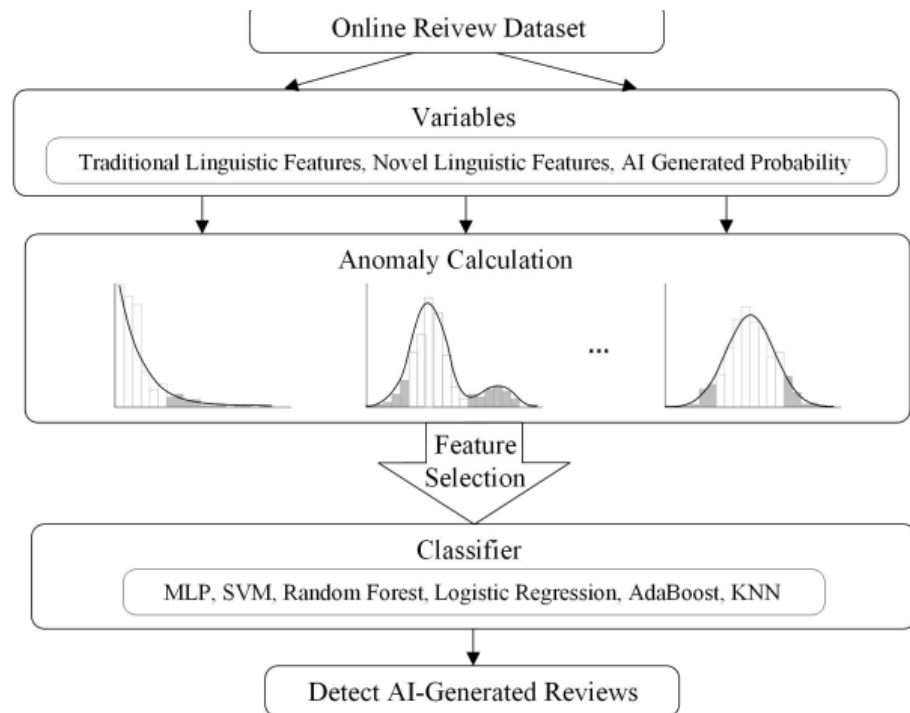


Figure 1. The overall design of AGRDCP.

- Reviews on products have huge impact on purchasing decision
- Rise of Generative Artificial Intelligence (GAI) makes it harder to distinguish
- Two ways to manipulate: merchants generating fake reviews to boost business and attempts to harm business by generating mass negative reviews
- Companies started implementing detection systems
- Different statistical methods applied in previous research: outlier detection

PREVIOUS RESEARCH

Linguistic factors considered in previous study

- Exaggerated emotional expression
- Templated sentence structure
- Inflated use of adjectives
- Reduction of randomness – AI written text based on highest predicted probability, human written texts with higher level of unpredictability and broader lexicon
- "Perplexity" and "burstiness"

RESEARCH QUESTION & HYPOTHESIS

- **Research Questions:**

1. Is the number of adjectives in user reviews a reliable factor to determine the authenticity of a user review?
2. Is the number of personal pronouns in user reviews a reliable factor to determine the authenticity of a user review?
3. Is the number of compound sentences in user reviews a reliable factor to determine the authenticity of a user review?

- **Hypotheses:**

1. The number of adjectives in user reviews is not a reliable factor to determine the authenticity of a user review.
2. The number of personal pronouns in user reviews is not reliable factor to determine the authenticity of a user review.
3. The number of subordinate clause/ compound sentences in user reviews is not a reliable factor to determine the authenticity of a user review.

METHOD – CHAT GPT REVIEWS

- **Our survey question:** "Generate 20 reviews for Netflix"
- **The responses:**
- "Netflix is the ultimate streaming service! With its vast library of movies and TV shows, it's a one-stop-shop for all your entertainment needs. The original content is top-notch, and the convenience of binge-watching your favorite series is unmatched."
- "Netflix continues to impress with its diverse selection of content. From thrilling dramas to hilarious comedies, there's always something for everyone. The user-friendly interface makes navigation a breeze, and the streaming quality is superb."
- "As a long-time Netflix subscriber, I appreciate the constant updates to the content library. The exclusive shows and movies are a game-changer, and the ability to download for offline viewing is a lifesaver for long flights or road trips."

METHOD – CHAT GPT REVIEWS

- **Our survey question:** "Generate 20 reviews for Apple Smartphone"
- **The responses:**
 - "The latest Apple smartphone is a technological marvel! The sleek design, powerful performance, and the stunning display make it a standout device in the market. The integration of new features showcases Apple's commitment to innovation."
 - "Upgrading to the newest Apple smartphone was a game-changer. The camera quality is unparalleled, capturing moments with incredible detail and clarity. The user interface remains intuitive, and the overall performance is smooth and responsive."
 - "Apple continues to set the bar high with its latest smartphone. The build quality is exceptional, and the attention to detail is evident in every aspect. The seamless integration with the Apple ecosystem enhances the overall user experience."

METHOD – OUR SURVEY

- Online questionnaire
- 10 Questions regarding the persons background (Age, nationality, academic level, bilinguality/multilinguality etc)
- 2 Questions asking for reviews
- Methods of reaching participants: Whatsapp groups, Facebook
- Target group:
 - over 18
 - German native speaker
 - German nationality
 - English as a second language

Survey on Online Reviews

Welcome to Our Online Questionnaire!

Thank you for participating in our survey. Your insights are invaluable to our research. Please be assured that your responses are **completely anonymous** and your participation is **voluntary**. Your individual responses **will not be made public**, and your privacy is of utmost importance to us.

The results of this survey will be aggregated and only discussed in the context of a university seminar presentation. Your contribution will play a crucial role in enhancing our understanding of the subject matter.

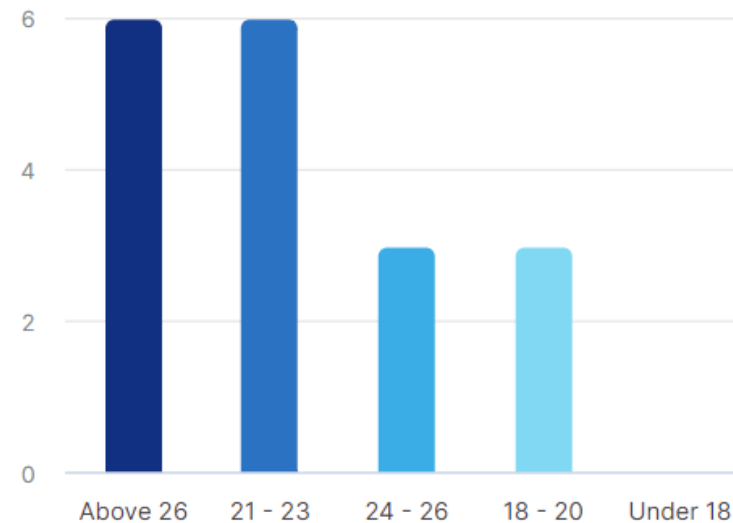
We appreciate your time and honesty. If you have any concerns or questions about the survey, feel free to contact us. Once again, thank you for being a part of our research journey!

[START SURVEY NOW](#)

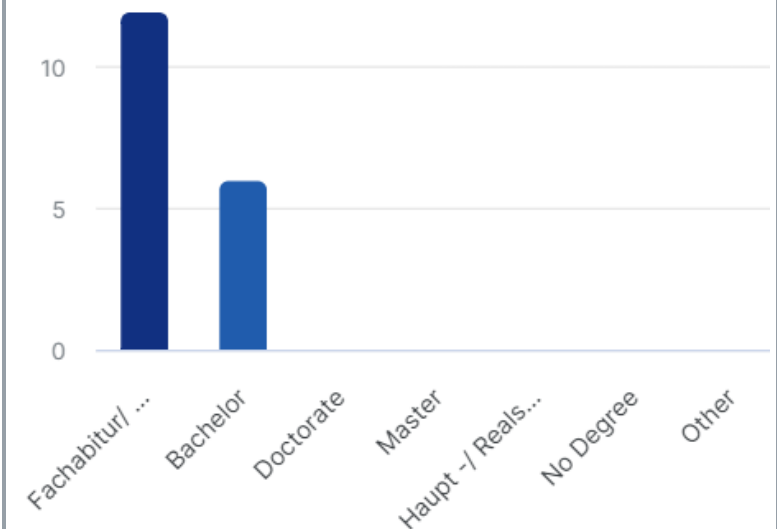
OUR SURVEY BACKGROUND INFO

- Our goal: 20 participants
- Achived: 21 participants
- 3 needed to be removed because they did not write a review or had a different nationality
- Gender: 16 female, 1 male, 1 diverse
- Other native languages (beside German): French, Lingala, Tamil, Rusian
- 3 people grew up bilingual or multilingual

1. How old are you?



3. What is your highest academic level?



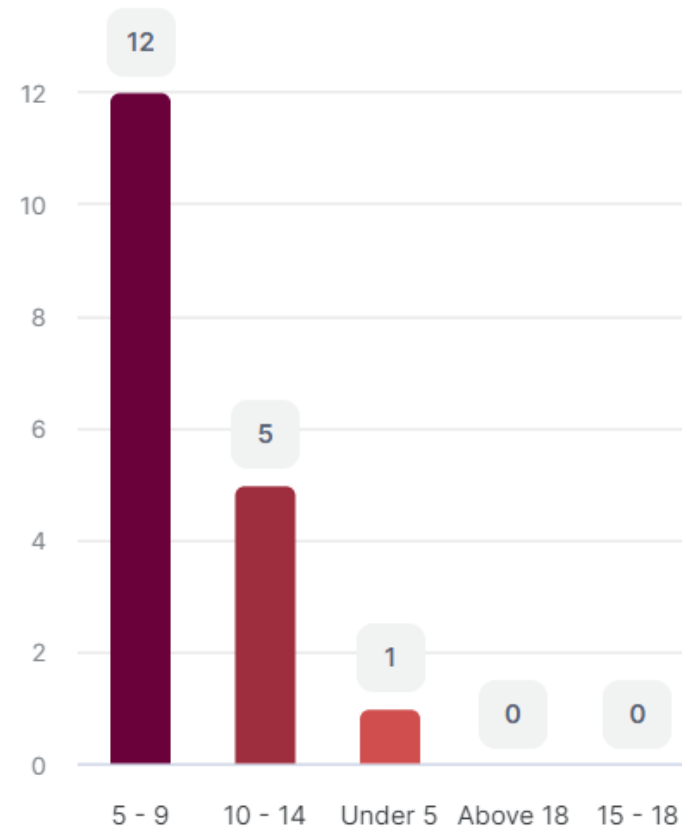
OUR SURVEY BACKGROUND INFO

- All participants study or studied at University

Cultural studies business engineering Applied linguistics
Lehramt BK (Wirtschaftswissenschaften und Englisch)
Master of Science BWL Linguistics
Education **Lehramt** Law
Disability studys medicine Physics GS Lehramt
Chemistry and English - Teacher Lehramt für die Sek 1
English and philosophy for gymnasium GS Lehramt mit Englisch

OUR SURVEY BACKGROUND INFO

9. At what age did you start learning the English language?



METHOD – OUR SURVEY– REVIEWS

- **Our survey question:** Imagine you are on an online review platform for different streaming platforms. You are writing a review in English, based on your experiences with Netflix. Please write an honest review for the streaming platform.
- **The responses:**
 - *„Netflix is a really big place for watching all kinds of shows and movies. It's easy to find stuff you like with the simple menu. They make their own shows too, and those are pretty cool. Sometimes the video quality might not be perfect, and they take away some shows, but overall, Netflix has so much to watch. It's a top choice for sure!“*
 - *„It's expensive but has nice selections and productions“*
 - *„For me Netflix is a 7/10 as there is not as much to select from. A few years back there many more movies and series to choose from“*

METHOD – OUR SURVEY– REVIEWS

- **Our survey question:** Please write an honest review in English, based on your experience with an Apple smartphone.
- **The responses:**
 - *„Using the Apple smartphone has been awesome. It looks cool, works smoothly, and the apps in the App Store add a lot to its features. The camera takes fantastic pictures and videos, making memories look really good. It is a bit expensive, but if you're okay with that, the way it connects with other Apple devices and the strong security features make it a worthwhile choice.“*
 - *„I really love iPhones. My first iPhone was the iPhone 3 and I've been using them ever since. The way they work is really intuitive. I would definitely recommend“*
 - *„It's ok“*

METHOD

LINGUISTIC FACTORS/MARKERS

Use of personal pronouns (I, me, mine)

Use of adjectives

Use of compound sentences – sentences with several main and subordinate clauses

Length (Number of words)

Spelling mistakes

REVIEW DATASET

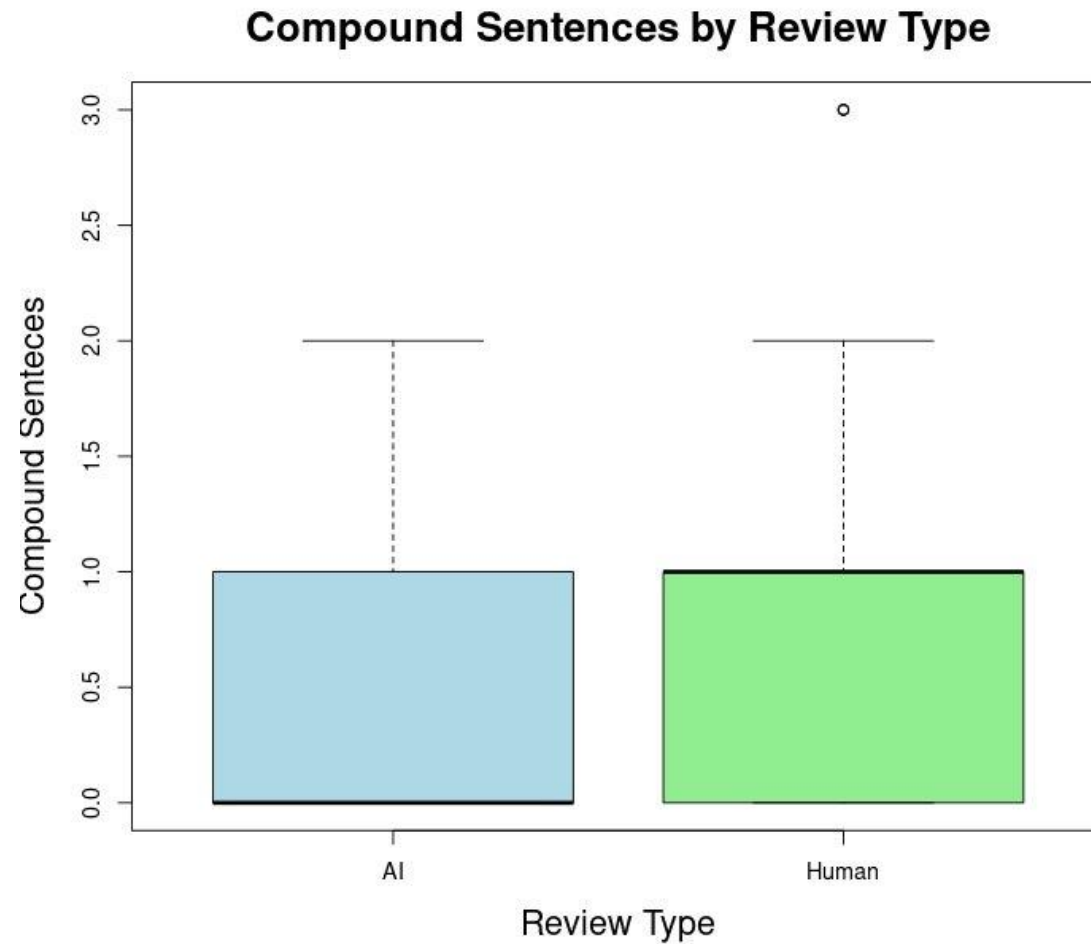
Review	compound_sentences	personal_pronouns	adjectives	word_count	spelling_mistakes	Target
Netflix 1 Chat <u>GPT</u>	1	0	3	29	0	AI
Netflix 2 Chat <u>GPT</u>	1	1	2	30	0	AI
Netflix 3 Chat <u>GPT</u>	0	1	3	34	0	AI
Netflix 4 Chat <u>GPT</u>	0	1	1	32	0	AI
Netflix 5 Chat <u>GPT</u>	0	0	2	36	0	AI
Netflix 6 Chat <u>GPT</u>	1	0	1	32	0	AI
Netflix 7 Chat <u>GPT</u>	0	0	4	36	0	AI
Netflix 8 Chat <u>GPT</u>	0	0	1	32	0	AI
Netflix 9 Chat <u>GPT</u>	0	0	1	31	0	AI
Netflix 10 Chat <u>GPT</u>	0	1	1	33	0	AI
Netflix 11 Chat <u>GPT</u>	0	0	2	32	0	AI
Netflix 12 Chat <u>GPT</u>	1	0	3	27	0	AI
Netflix 13 Chat <u>GPT</u>	0	0	3	29	0	AI
Netflix 14 Chat <u>GPT</u>	1	1	2	30	0	AI
Netflix 15 Chat <u>GPT</u>	0	0	4	27	0	AI
Netflix 16 Chat <u>GPT</u>	1	0	2	37	0	AI
Netflix 17 Chat <u>GPT</u>	0	0	0	29	0	AI
Netflix 18 Chat <u>GPT</u>	1	0	2	36	0	AI
Netflix 19 Chat <u>GPT</u>	1	0	3	38	0	AI
Netflix 20 Chat <u>GPT</u>	1	0	3	35	0	AI
Netflix 1 Survey	0	1	2	16	4	Human
Netflix 2 Survey	1	0	4	51	0	Human
Netflix 3 Survey	0	0	0	4	2	Human
Netflix 4 Survey	2	3	1	45	1	Human
Netflix 5 Survey	2	2	2	95	1	Human
Netflix 6 Survey	1	0	1	22	0	Human
Netflix 7 Survey	0	0	1	29	1	Human
Netflix 8 Survey	1	1	1	21	2	Human
Netflix 9 Survey	0	0	0	8	1	Human
Netflix 10 Survey	3	5	7	93	10	Human
Netflix 11 Survey	1	2	3	33	1	Human

STATISTICAL METHODS

- **BOX PLOTS**

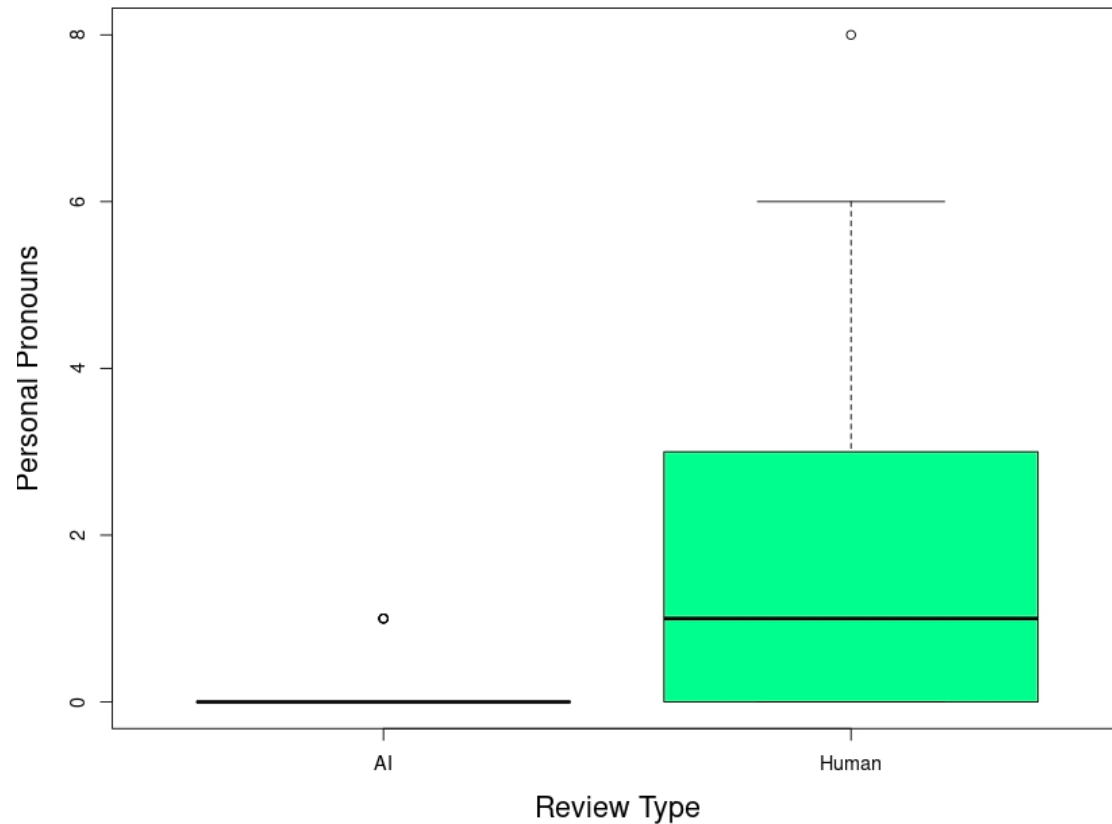
- **DECISION TREES**

UNDERSTANDING COVARIATES

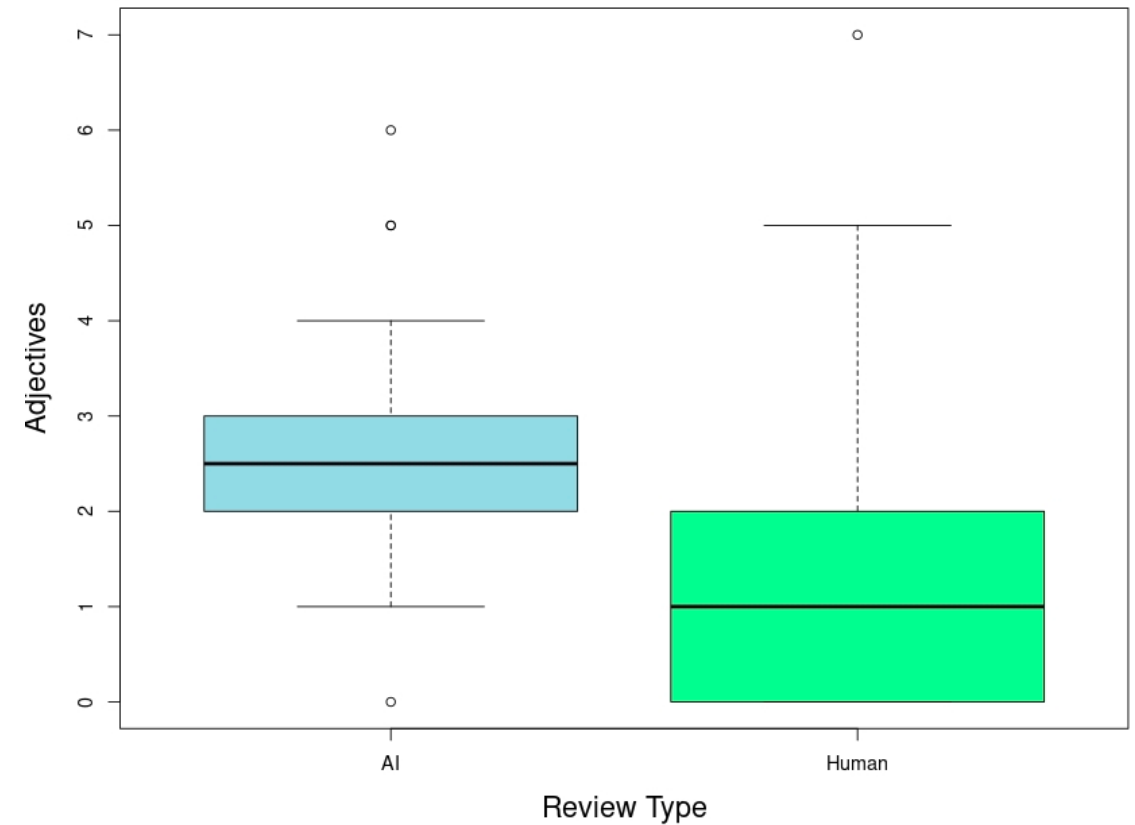


UNDERSTANDING COVARIATES

Personal Pronouns by Review Type

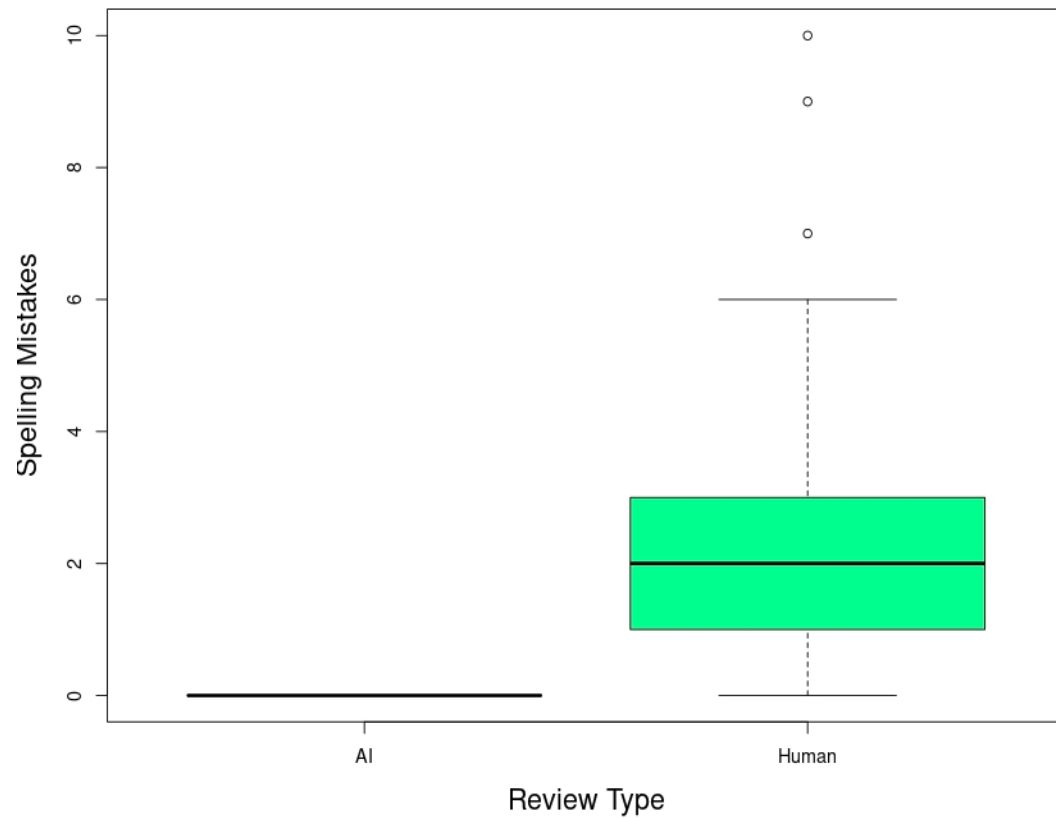


Adjectives by Review Type

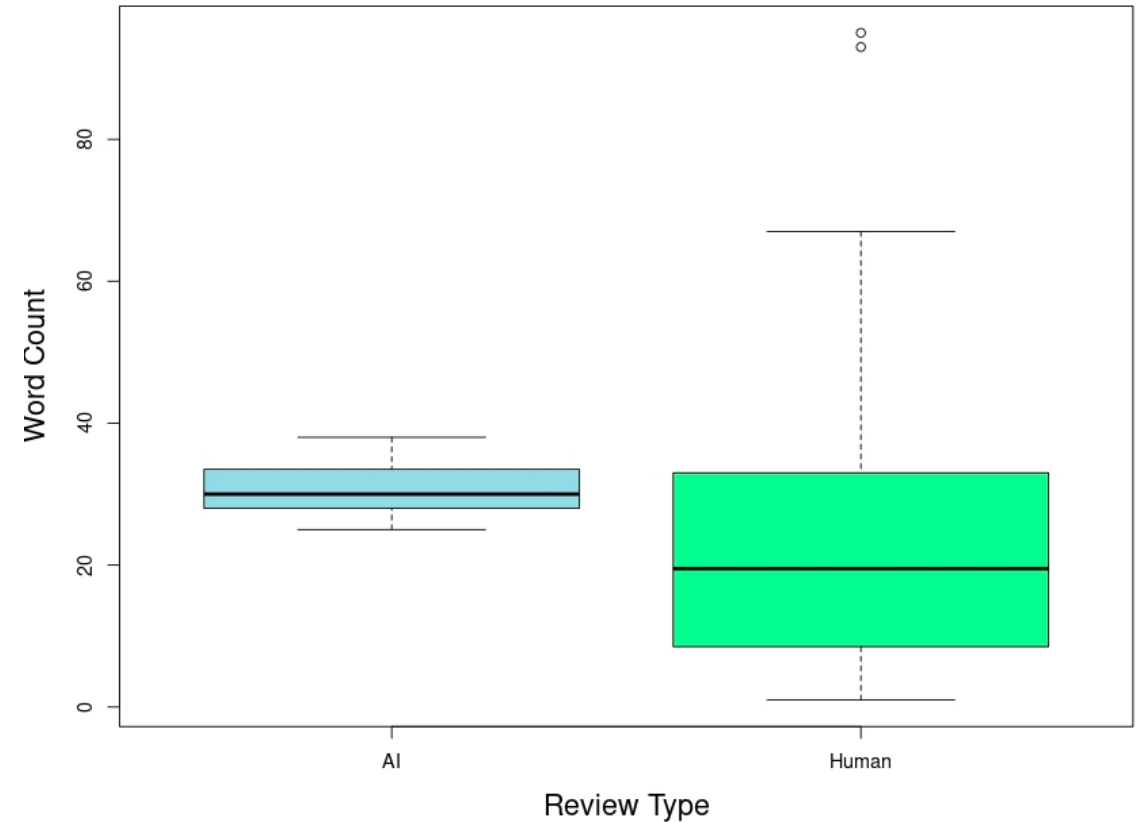


UNDERSTANDING COVARIATES

Spelling Mistakes by Review Type



Word Count by Review Type

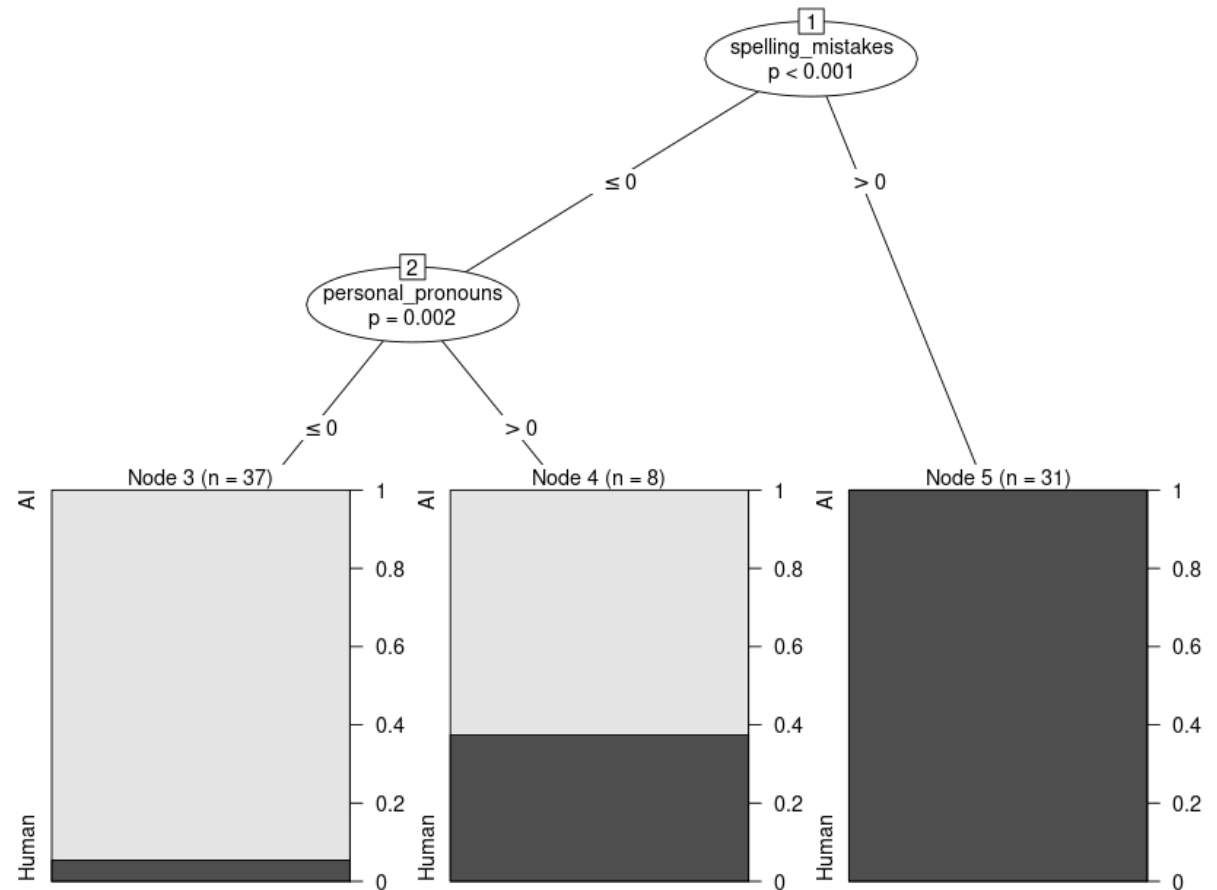


METHOD – DECISION TREES

- Decision trees are used to model the relationship between qualitative dependent variables and independent variables.
- To assess the quality of a classification rule,
 - **Accuracy** is used to assess the quality of a decision rule. It measures the rate of correct results if the rule is applied to the same data set used for rule construction.
 - **Balance accuracy** is use to assess the predictive power of set of classification rule. It is the mean of the accuracies of the two classes.

CONDITIONAL INFERENCE TREE

- Predicts whether the review is written by an AI assistant or Human based on covariates such as number of compound_sentences, personal_pronouns, adjectives, spelling_mistakes and word_count.
- spelling_mistakes and personal_pronouns are the most influential covariates
- High Accuracy of 93.42% and Balanced Accuracy of 93.05%



RESULTS

- **Hypothesis one:**

- **H₀:** The number of adjectives in user reviews is not reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in english.
- **H₁:** The number of adjectives in user reviews is a reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in english.

- **Findings**

From the decision tree, adjective is not significant. **H₀ is not rejected.**

The number of adjectives in user reviews is not reliable factor to determine difference between AI generated reviews and reviews written by German native speakers in English.

RESULTS

- **Hypothesis two:**

- **H₀:** The number of personal pronouns in user reviews is not reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in English
- **H₁:** The number of personal pronouns in user reviews is a reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in English

- **Findings**

From the decision tree, personal pronouns is significant. **H₀ is rejected.**

We conclude that the number of personal pronouns in user reviews is a reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in English.

RESULTS

■ Hypothesis Three:

- **H₀:** The number of subordinate clause/ compound sentences in user reviews is not a reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in English
- **H₁:** The number of subordinate clause/ compound sentences in user reviews is a reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in English

■ Findings

From the decision tree, the subordinate clause/compound feature is not significant. **H₀ is not rejected.**

We conclude that the number of subordinate clause/compound sentences in user reviews is not a reliable factor to determine the difference between AI generated reviews and reviews written by German native speakers in English.

LIMITATIONS & CONCLUSION

Limitations:

- Unequal gender ratio (majority female)
- Small data set
- Decision tree is a greedy algorithm, that finds only locally optimal solutions at each node in the tree.

LIMITATIONS & CONCLUSION

Conclusion

Based on our data set, the most important variables to distinguish between AI generated reviews and reviews written by german native speakers with English are spelling mistakes and personal pronouns.

Based on the classification rule from the data analysis, reviews without spelling mistakes are classified as AI generated while the ones with spelling mistakes are splited by number of personal pronouns.

Reviews with no personal pronouns have approximately 95% chance of being written by human while those with personal pronouns have approximately 65% AI generated.

Future studies

- Increase data set
- Add more explanatory variables
- Explore other classification methods, compare and select the model with best performance.