

EXTRACTION OF ATYPICAL ASPECTS FROM CUSTOMER REVIEWS:
DATASETS AND EXPERIMENTS WITH LANGUAGE MODELS

by

Smita Nannaware

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Computer Science

Charlotte

2023

Approved by:

Dr. Razvan Bunescu

Dr. Wlodek Zadrozny

Dr. Xi (Sunshine) Niu

Dr. Tomek Strzalkowski

ABSTRACT

SMITA NANNAWARE. Extraction of Atypical Aspects from Customer Reviews:
Datasets and Experiments with Language Models. (Under the direction of DR.
RAZVAN BUNESCU)

A restaurant dinner can become a memorable experience due to an unexpected aspect that is appreciated by the customer, such as an origami-making station in the waiting area. If aspects that are atypical for a restaurant experience were known in advance, they could be leveraged to make recommendations that have the potential to engender serendipitous experiences, further increasing user satisfaction. Although relatively rare, due to their memorable quality, atypical aspects often end up being mentioned in reviews. Correspondingly, in this thesis, I propose the task of detecting atypical aspects in customer reviews. To facilitate the development of extraction models, I manually annotate benchmark datasets of reviews in three domains: restaurants, hotels, and hair salons. The datasets are then used to evaluate a number of language models, ranging from fine-tuning the instruction-based text-to-text transformer Flan-T5 to zero-shot and few-shot prompting of the much larger ChatGPT.

DEDICATION

To my parents, thank you for teaching me the importance of education.

ACKNOWLEDGEMENTS

I sincerely thank my thesis advisor, Prof. Razvan Bunescu, for giving me the opportunity to work on interesting research in the field of Natural Language Processing. I am grateful for his constant support and guidance with patience. I couldn't have enjoyed the topic without his direction. I would like to thank the committee members, Dr. Wlodek Zadrozny, Dr. Xi (Sunshine) Niu, and Dr. Tomek Strzalkowski, for being part of my thesis committee and providing ideas and valuable feedback.

A special thanks to Erfan Al-Hossami, whose encouragement and assistance have been a guiding light during moments of uncertainty. I am deeply grateful to UNC Charlotte for providing the essential resources for computational tasks, which played a vital role in the successful completion of this research.

My heartfelt thanks extend to all my friends and fellow PhD colleagues, whose constant encouragement and shared passion for research have been a source of motivation and inspiration throughout this journey.

Lastly, this would not have been possible without immense motivation from my husband, Darshan. Thank you for your belief in my abilities and unconditional support.

Thank you all for being an integral part of this remarkable journey.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: RELATED WORK	5
CHAPTER 3: XtrAtA TASK DEFINITION AND ANNOTATION GUIDELINES	7
CHAPTER 4: XtrAtA DATASETS OF CUSTOMER REVIEWS	13
4.1. Manual Annotation Of Atypical Aspects	13
4.2. Artificial Addition Of Atypical Aspects	16
CHAPTER 5: XtrAtA APPROACHES USING LANGUAGE MODELS	23
CHAPTER 6: EXPERIMENTAL EVALUATIONS	26
6.1. Experiments Using Artificially Augmented Reviews	29
CHAPTER 7: CONCLUSION	31
APPENDIX A: ANNOTATED EXAMPLES	40
APPENDIX B: ATYPICAL ASPECTS IN ARTIFICIAL DATASET	46

LIST OF TABLES

TABLE 3.1: Examples of extractive and <i>abstractive</i> annotations of customer reviews in the restaurants domain. Secondary or optional annotations are shown in blue .	10
TABLE 4.1: Statistics for the 3 datasets, split between Train+Test and Development (Dev). The number of atypical reviews and atypical aspects are presented separately for primary atypical aspects (Primary) vs. both primary and secondary atypical aspects (+ Secondary).	15
TABLE 4.2: Statistics for the 3 datasets, split between Train+Test, Artificial Train Type 1 (ATT-1), Type 2 (ATT-2) and Development (Dev). The number of atypical reviews and atypical aspects are presented separately for primary atypical aspects (Primary) vs. both primary and secondary atypical aspects (+ Secondary).	20
TABLE 4.3: Taxonomy of restaurant atypical aspects	21
TABLE 6.1: Extractive (exact vs. partial match) and abstractive results (%), on primary (default) and primary + secondary atypical aspects across the 3 domains. Precision (P), Recall (R), and F1 scores are reported for the extractive setting. The F1 scores of Rouge-1 (R-1), Rouge-2 (R-2), rougeLsum (RLS) and BERTScore (BS) are reported for the abstractive setting. Best results in each domain are in bold .	27
TABLE 6.2: Extractive (exact vs. partial match) and abstractive results (%), on primary atypical aspects across the 3 domains. Precision (P), Recall (R), and F1 scores are reported for the extractive setting. The F1 scores of Rouge-1 (R-1), Rouge-2 (R-2), rougeLsum (RLS) and BERTScore (BS) are reported for the abstractive setting. Best, and second best results in each domain are in bold , <u>underline</u> respectively.	30
TABLE A.1: Examples of extractive and <i>abstractive</i> annotations of customer reviews in the hotels domain. Secondary or optional annotations are shown in blue .	42
TABLE A.2: Examples of extractive and <i>abstractive</i> annotations of customer reviews in the hair salons domain. Secondary or optional annotations are shown in blue .	45
TABLE B.1: Atypical aspects of restaurants for Type 2 artificial data	48
TABLE B.2: Atypical aspects of hotels for Type 2 artificial data	51

TABLE B.3: Atypical aspects of hair salons for Type 2 artificial data	54
---	----

LIST OF FIGURES

- FIGURE 1.1: The user profile on the left shows (in bold) a user interest that is relevant for the atypical aspect (in bold) shown on the right. 3

LIST OF ABBREVIATIONS

BNP Base Noun Phrase

LM Language Model

PEFT Parameter-Efficient Fine-Tuning

XtrAtA Extraction of Atypical Aspects

CHAPTER 1: INTRODUCTION

When looking for a restaurant or a hotel, people are often faced with an overwhelming number of options matching their search constraints. Even when ranked by their average review scores, there may be numerous high quality choices that satisfy the basic search criteria, especially in a metropolitan area. This may lead to choice overload, or overchoice [38, 5], where an individual is presented with a large number of choices that are too difficult to compare, particularly under time constraints [9]. Making a decision in the presence of overchoice becomes mentally exhausting and can lead to subsequent impaired self-regulation [39], decision paralysis, and anxiety [10]. The level of satisfaction that people experience when faced with an increasing number of choices has been observed to follow the well-known Wundt curve [2], an inverted U-shape curve originally relating stimulus intensity with its pleasantness. According to this functional dependency, as the number of choices goes up, satisfaction initially increases and then decreases [35, 14]. In this context, choice overload can be alleviated by reducing the number of consumer choices [32] or by making one option stand out and appear better than the others [31]. To this end, I propose to emphasize options that possess aspects with the potential to surprise the user in a positive way, i.e., serendipity. For example, as shown in Figure 1.1, suppose a user Jane is looking for a ramen restaurant in her locality. The system knows that she has been passionate about creating crafts from paper since childhood. Upon asking the system for recommendations, the system finds a number of highly rated ramen restaurants, however, one of these, Nikita Ramen, stands out because it has an origami making station in its waiting area (an atypical aspect for a restaurant). The system recommends this restaurant to her, **without telling her about the origami station**. Upon entering

the restaurant, she is very pleasantly surprised to see the origami making station in the waiting area, which brings feelings of nostalgia and happy memories from childhood. She takes some time making various origami figures, before being seated at her table. This serendipitous experience was facilitated by the fact that an origami making station is an atypical aspect, hence it would be experienced as surprising in a restaurant. After the dinner event, the system further confounds her expectations by asking her if she enjoyed the origami making station, which surprises her because she did not expect that the system was responsible for the initial surprise.

To enable such recommendations with potential for serendipity, the input would ideally consist of three parts:

1. The user's query, be it a standalone request or a turn in a longer conversation. This would be used to find the initial, often large set of items that satisfy the user's information need.
2. The item's data, including not only information about the typical aspects of items from the same category, but also any atypical aspects that are likely to generate surprise.
3. The user's data, especially in terms of their interests, their likes and dislikes. This would be useful for determining if an atypical aspect would be enjoyed by the user, i.e. serendipitous.

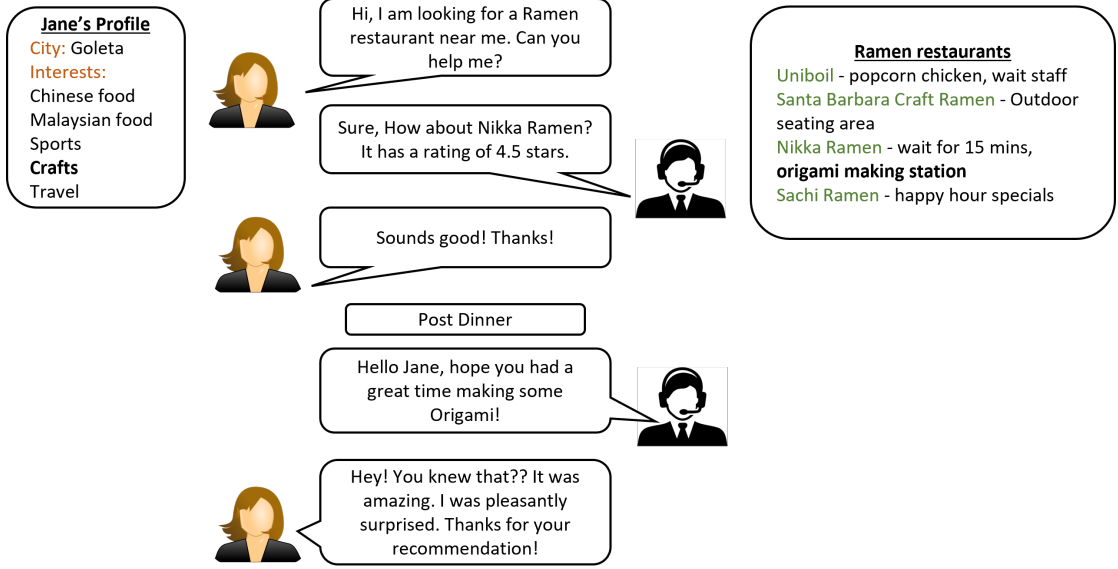


Figure 1.1: The user profile on the left shows (in bold) a user interest that is relevant for the atypical aspect (in bold) shown on the right.

To increase the chance of serendipity, the system would need to (i) have knowledge about the user preferences, and (ii) also ensure that the user notices / takes advantage of the atypical aspect, e.g. estimating that there will be some wait involved in the origami example. In this thesis, I introduce a more focused task where I assume that the category of items requested in the user's query at (1) is known, e.g. restaurants, and the task is to use the item's data to extract aspects that are *atypical* for its category, e.g. origami station for restaurants. Given that the user's expectations are shaped by what they think is common for the category of interest in their query, atypical aspects are likely to confound the user's expectations, and hence be perceived as surprising. Henceforth, we will use the term *surprising aspects* to refer solely to *atypical aspects*. Because atypical aspects tend to be more noticeable and often lead to more memorable experiences, they are likely to be mentioned in customer reviews. Therefore, I use customer reviews as the source of an item's data at (2). At this time, no user data is used as input, which means that the atypical aspects that are extracted, while surprising for the user searching for that particular category of items, cannot be said for sure to lead to serendipity due to unknown user preferences. In

short, the task is that of extracting atypical, surprising aspects from customer reviews, where atypical is defined to be relative to a predefined item category.

The rest of the thesis proceeds as follows. Chapter 2 describes briefly related work in the area. Chapter 3 introduces the task of surprising aspect extraction from customer reviews. Chapter 4 details the development of benchmark datasets of customer reviews that are manually annotated for atypical aspects with respect to three categories: restaurants, hotels, and hair salons. In Chapter 5 I describe a number of extraction approaches that rely on language models (LM), ranging from Flan-T5 [40, 6] and ChatGPT [26] in zero-shot or few-shot setting, to fine-tuning of Flan-T5 with and without the help of artificial data generated using ChatGPT. Experimental evaluations of these models in both extractive and abstractive settings are detailed in Chapter 6. The thesis ends with concluding remarks in Chapter 7. Data is available for review at <https://anonymous.4open.science/r/surprising-aspects>.

CHAPTER 2: RELATED WORK

Mitigating overchoice is an important focus of recommender systems and it is typically addressed by recommending between 5 to 20 attractive and diverse items [3], based on user preferences [11, 44], user ratings, item attributes, or user reviews [23, 33, 17]. Recommending items with potential for serendipity is one way of diversifying an item set. In [28], unexpectedness is defined as the distance of an item from a set of obvious items for that user, relative to the user’s preferred level of unexpectedness. Li et al. [18] recommended unexpected items by modeling user interests as clusters of historical data in a latent space and calculating the weighted distance between a new item and the clusters of interests. Kotkov et al. [16] crowdsourced serendipity labels for a movie dataset using multiple definitions of serendipity. Fan and Niu [7] and Niu et al. [25] develop a number of computational approaches to serendipity, decomposed into surprise and utility components, and evaluate them on the task of recommending news articles in the health domain. In contrast, the task described in this proposal targets a more fine-grained category of surprise, namely that of *atypical aspects*. Atypical aspects inherently have a high potential for surprise, which precludes the need for a computational definition of surprise.

User reviews have been used to learn latent features of users [17], extract sentiment [33], derive user preferences [44], or to perform aspect based sentiment analysis in order to recommend better quality products with aspects relevant to the user [23, 22]. Conversational recommender systems use reviews to provide explanations [24], to maintain fluency in conversation [20], or to understand the users requirements by asking questions about aspects mentioned in reviews [45].

To the best of my knowledge, no prior work has looked into extracting atypi-

cal aspects from item data such as customer reviews, which have high potential for serendipity and for alleviating overchoice.

CHAPTER 3: XtrAtA TASK DEFINITION AND ANNOTATION GUIDELINES

Given a domain category, e.g. restaurants, and a customer review of a particular item in that category, e.g. a restaurant, the task is to extract aspects of that particular item that are atypical of items in its category. I call this task as Extraction of Atypical Aspects (XtrAtA). Throughout most of this document, I will use the category of restaurants as an example. Correspondingly, in Table 3.1 I show samples taken from 4 reviews, illustrating two types of manual annotations, *extractive* and *abstractive*, analogue to the extractive [13, 15] and abstractive [12, 30] annotation schemes widely used in summarization datasets¹. Similar to Table 3.1 examples for hotels and hair salons are given in Table A.1 and Table A.2 in the appendix.

In the extractive annotation, only base noun phrases referring to atypical aspects are annotated. If multiple phrases refer to the same atypical aspects, I only annotate the grounding instance of the coreference chain. For example, the noun phrase "adjoining park" in the second review is not annotated, as it refers to "Goose Creek Park" which is already annotated. However, if a review mentions an atypical category, such as "arcade games" in the first review or "local wildlife" in the third review, any category instance that is mentioned will also be annotated as atypical, such as "nfl blitz 99" in the first review or "anarchists" in the third review, respectively. If the atypical aspect is a more complex noun phrase, I only annotate the base noun phrase that expresses the semantic core (often the syntactic head), as in "the *Goose Creek Park* behind the restaurant". The extractive annotation is meant to be used together with the original review in downstream applications, which makes it acceptable to annotate only the most important part of the phrase.

¹<https://duc.nist.gov>, <https://tac.nist.gov>

► A group of work friends and I stumbled upon Upper Deck a little over a year ago and everyone from our office has turned Upper Deck into our local watering hole ever since. Their happy hour special is unbeatable, they have a good selection of draft beers, and the food is out of this world good. The stand out feature of Upper Deck is the offering of **life size beer pong** at their outside patio. This takes traditional beer pong and substitutes solo cups with garbage cans (painted to look like solo cups) and Ping Pong balls with dodgeballs. They also have a **pool table** and recently added **arcade games** (**nfl blitz 99** beats **madden 15** all day). Get some friends and bring your appetites and some quarters, you won't be disappointed.

The restaurant offers life size beer pong at their outside patio. They have a pool table. They recently added arcade games, such as nfl blitz 99 and madden 15.

► The big draw of this place is the excellent pizza, which you can have with beer on an outside deck with a view of the **parklands**. It's a nice place to hang out on a sunny afternoon. You can even go for a walk in the **Goose Creek Park** behind the restaurant afterwards to burn off the calories you just consumed. The big minus is that if you go for lunch or in early afternoon, the menu is really limited. This place also seems to attract a goodly number of families with kids at lunch times, probably because it serves pizza and there's a **playground** in the adjoining park,

Continued on next page

Table 3.1 – *Continued from previous page*

The restaurant has an outside deck with a view of the parklands. Customers can go for a walk in the Goose Creek Park behind the restaurant. There's a playground in the adjoining park.

► Classic West Philly spot where you can see **local wildlife**. Everyone from **moms** to **anarchists** to **hackers** to **organic gardeners** to **activists** hangs out there. The coffee is excellent, the baked goods are great as well, and if you're working on something, you might run into a possible collaborator there. If you're thinking of moving to West Philly, definitely check out Satellite and the farmer's market. They've replaced the cracked and chipped cups with awesome new cups, which are awesome.

In this restaurant you can see local wildlife. Everyone from moms to anarchists to hackers to organic gardeners to activists hangs out there.

Continued on next page

Table 3.1 – *Continued from previous page*

► This is such a cool place! Three words was all it took to add this gem to my list of places to visit while in St. Louis - "**Good Burger Car!!**" YES! They have the car from the movie Good Burger! A movie I was obsessed with as a child & have since gotten my kids to love just as much! The place is made out of **cool, colorful shipping containers** with many neat decorations, what looks like an **alien spaceship** from Toy Story with a laser on it adorned the top of the place along with a **cow**. Now, on to the food. They have many different options to choose from including create-your-own burgers & many of their own creations, sandwiches, salads, sides, kid's meals, & shakes & floats ... Such a unique place & worth a visit! They also sell **souvenir T-shirts & hats**, & my fiance had to get himself a "**HI AF**" **shirt**. The shirts were heavily influenced by the **movie Good Burger** & there was one in particular I had my heart set on, unfortunately they did not have it in my size.

The restaurant has the car from the movie Good Burger. They sell souvenir T-shirts and hats, and a customer got a "HI AF" shirt. The shirts are heavily influenced by the movie Good Burger.

The place is made out of cool, colorful shipping containers, with many neat decorations, what looks like an alien spaceship from Toy Story with a laser on it adorned the top of the place along with a cow.

Table 3.1: Examples of **extractive** and *abstractive* annotations of customer reviews in the restaurants domain. Secondary or optional annotations are shown in **blue**.

In the abstractive annotation, one or more sentences are generated that enumerate the atypical aspects mentioned in the review. The formulation is kept as close as

possible to the original text while maintaining naturalness. The generated sentences are intended to be concise, usually maintaining details that are expressed in the same sentence in the review, however keeping out unimportant information about the atypical aspect that is mentioned in other sentences, or details that are vague or uncertain. Sentiment words are maintained only if it helps keep the text natural and faithful to the original. The abstractive annotation is meant to be standalone and used without the original review in downstream tasks, as such it may require some minimal rewriting of the original review formulation, e.g. adding the phrase "the restaurant", or removing opinion words such as "beats" in the first example. Sometimes, reviewer use metaphors to refer to atypical aspects, in which case it is important that the abstractive version preserves the metaphorical meaning. This is the case for "local wildlife" in the third review, which refers metaphorically to types of customers that are seen relatively less often in that context.

All aspects that are related to the core business of a restaurant, including but not limited to food, service, price, opening hours, parking, are considered typical aspects and are not annotated. Conversely, an aspect is considered atypical and is annotated if it is not related to the core business of *a restaurant*, yet it belongs to or is a feature of *the restaurant* ("a restaurant" refers generically [4] to the restaurant category, "the restaurant" refers to a specific restaurant). A special case is made of aspects that are related to the ambience or atmosphere of a restaurant: while they might be considered as an important part of, and thus subordinate to, the core restaurant experience, there are cases where they stand out and become an attraction on their own. When that happens, I annotate them on a secondary, optional layer. In my annotations, the distinction between typical vs. atypical aspects is done with respect to common cultural expectations in the US. This is to control for the fact that aspects that are typical in the US, e.g. a bar in a restaurant, may be atypical in other cultures.

In many aspect-based sentiment analysis approaches [34], identifying *typical* aspects that are mentioned in a review is done explicitly as *aspect term extraction*. However, the task of extracting *atypical* aspects, as introduced above, cannot be solved simply by first (a) identifying all typical aspects of restaurants that are mentioned in a review, followed by (b) extracting all other noun phrases, i.e. phrases that do not refer to a typical aspect of a restaurant. In their reviews, people often mention entities or events that are not associated with the restaurant, such as "our office" in the first review, or "the farmer's market" in the third review, and these phrases should not be extracted either. Thus, it is important that the noun phrase refers to an aspect that is *associated with the reviewed restaurant*, and that at the same time is atypical of restaurants (or unexpected for a restaurant). Finally, while Table 3.1 may induce the perception that atypical aspects are common, the opposite is actually true. As will be detailed in Section 4.1 below, it takes going through at least 50 reviews in order to find one review that mentions an atypical aspect. The difficulty of manually finding this "needle in a haystack" further motivates the development of automated approaches for surprising aspect extraction.

CHAPTER 4: XtrAtA DATASETS OF CUSTOMER REVIEWS

In this chapter I introduce a real dataset of reviews that are manually annotated for atypical aspects (Section 4.1), followed by the description of a procedure for automatic augmentation of originally typical reviews (Section 4.2) with sentences describing atypical aspects.

4.1 Manual Annotation Of Atypical Aspects

I used the Yelp dataset [42] as a source of reviews for the 3 target categories: Restaurant (~ 5 M reviews), Hotel (~ 190 K reviews), and Hair Salon (~ 115 K reviews). Because most aspects are expressed as nouns and less frequently as verbs, I use spaCy [36] to collect lemmas of all nouns and verbs and compute their frequencies for each domain. I rank words in ascending order based on their counts and filter out words that appear with very low frequency, e.g., less than 10 times for the Restaurant domain, as these tend to be spelling mistakes or interjections that are purposely misspelled for extra emphasis, e.g., "amaazzing". I then consider the remaining rare words in ascending order of their frequency as candidate atypical words, extract the reviews that mention them, and read these reviews to determine which occurrences truly refer to an atypical aspect. When reading a review, all atypical aspects are annotated, not only the ones corresponding to the search word. Notwithstanding the heuristic selection of reviews based on the occurrence of rare words, overall this was still a very time-consuming process, because rare words very often appear in a review without necessarily referring to atypical aspects. For example, out of the 43 restaurant reviews that contain the lemma "poncho", in only 1 review the word "ponchos" was deemed to refer to an atypical aspect (the restaurant was selling them).

The other 42 reviews contained references to ponchos that were not associated with the restaurants itself, e.g. staff helping customers put their ponchos on a rainy day, or customers describing their arrival at the restaurant on a rainy day. As I went down the list of rare words, their frequency increased, resulting in a larger number of reviews to skim through for each rare word. Overall, for the Restaurant dataset, I used search words as the rare words that appeared with a frequency of up to 187. Upon semi-automatically sifting through the $\sim 97\text{K}$ reviews found to contain these words, I was able to collect 114 reviews that contained atypical aspects. On average, one hour of following this process led to finding between 2 and 3 reviews containing atypical aspects for the restaurant category, whereas for the hair salon category it took on average two hours to find 1 atypical reviews. Henceforth the term *atypical review* will be used to refer to a review that contains one or more atypical aspects; analogously, the term *typical review* will be used to refer to reviews that do not contain any atypical aspect.

As illustrated in the examples from Chapter 3, I organize annotations of atypical aspects on two layers:

- A *primary* layer that contains atypical aspects that are clearly not connected to any core feature of that domain.
- A *secondary* layer that contains atypical aspects that are related to a typical aspect, such as ambiance or location, but that stand out and are interesting on their own, separate from the core features of the domain.

For example, *'I was even encouraged to visit their **petting zoo** in the back'* would be considered a primary atypical aspect in any of the 3 categories, whereas *'There is an **interesting giant stuffed spider** that goes up and down when the door leading to the bathrooms opens and closes'* would be annotated as a secondary atypical aspect.

Table 4.1 shows summary statistics for the 3 datasets, one for each domain (cate-

Table 4.1: Statistics for the 3 datasets, split between Train+Test and Development (Dev). The number of atypical reviews and atypical aspects are presented separately for primary atypical aspects (Primary) vs. both primary and secondary atypical aspects (+ Secondary).

Domain	Dataset split	Primary		+ Secondary		Total
		# reviews	# aspects	# reviews	# aspects	reviews
Restaurant	Train+Test	100	253	107	340	200
	Dev	14	32	16	46	28
Hotel	Train+Test	69	274	85	401	150
	Dev	10	33	11	49	20
Hair Salon	Train+Test	45	147	48	181	90
	Dev	5	24	5	29	10

gory), split between data used for training and testing, and data used for development. Under the Primary column, I show the number of atypical reviews and atypical aspect annotations contained in them. The next column shows the same statistics for when both primary and secondary atypical aspects are considered. The total number of reviews in each dataset, shown in the third column, is about double the number of primary atypical reviews, reflecting a balanced dataset where the number of typical reviews was selected to be about the same as the number of atypical reviews.

Often a reviewer expresses positive or strong sentiment towards the atypical aspect by explicitly writing that they like it or implicitly through exclamation points. In the restaurant domain, 119 out of 253 primary atypical aspects are associated with a positive or a strong sentiment, whereas 9 are associated with a negative sentiment. The remaining atypical aspects are mentioned in neutral, informative sentences such as "There is a gift shop". In the hotel domain, 138 out of 274 primary atypical aspects are associated with a positive or a strong sentiment, whereas 7 are associated with a negative sentiment. And in the hair salon domain, 51 out of 147 primary atypical aspects are associated with positive sentiment, whereas there are no atypical aspects associated with a negative sentiment.

With the help of Erfan Al-Hossami, I computed inter-annotator agreement (ITA)

on both the extractive and abstractive annotations in the development set of Hair Salon. The ITA metrics are shown in the last row of Table 6.1 and are calculated by holding one annotator’s annotations as the ground truth, while the other annotator is considered as the system.

4.2 Artificial Addition Of Atypical Aspects

Finding atypical reviews is a time-consuming process and finetuning a LM often requires a large amount of data. To alleviate this issue, I created a synthetic dataset using state of the art generative models such as ChatGPT. I have used two methods to generate atypical reviews: Type 1 and Type 2. It is important to note here that the synthetic reviews are meant to be used only during training, as such testing will always be conducted solely on real reviews. In Type 1, I instructed ChatGPT to think of any atypical aspect and incorporate it into a given typical review as shown in **Ex.1**. This method helps to identify atypical aspects which ChatGPT has seen in its training data that I had not discovered while creating the manual dataset earlier. Whereas, in Type 2, I provided an atypical aspect and asked ChatGPT to incorporate it into a given typical review as shown in **Ex.2**.

Ex.1 ChatGPT Input prompt: *Below is an example of a restaurant review that contains an atypical aspect.*

Review: I’m addicted to the smoothies here. They make me feel like a glowing goddess, full of light energy and goji berries. The menu is impressive, spanning many different types of cuisine in a vegan rendering. My one problem is the price- The smoothies come out to \$8 plus tax, which make this place more of a special treat for me rather than a morning routine, which it could be if the prices were about 25% lower. This trend spans the rest of the menu, which is unfortunate because while I love their mission and I really want to support this place more often, I balk at the prices and end up shying away from ever trying anything substantial. However, I understand

that the pricing is due to the quality of the ingredients and the costs of running a very niche restaurant (not to be confused with Niche, the restaurant), so I chalk up my 8+ dollars and indulge as often as I can. I love PuraVegan because it's such a refreshing change of pace in a place like St. Louis, where things like "pork steak" rule and ravioli is deep-fried and BBQ is a lunch staple. I really like the people who work here, and they host donation yoga classes, which is so awesome. This place is worth the price if you want to treat yourself to a truly healthy meal and nurture your body. This place is not for people who got lost on their way to The Kitchen Sink.

Atypical aspect: The restaurant hosts donation yoga classes.

Follow these steps to answer the user queries.

Step 1: Read the review given by the user.

Step 2: Think of one or more suitable atypical aspects that would be coherent with the review. Atypical aspects are not related to food, service, beverages, decor, location, price, menu, discounts, parking, policies, staff, customer satisfaction, or other items commonly associated with a restaurant.

Step 3: Incorporate the atypical aspects into the original review in the same writing style as the original review. Add one or two sentences having atypical aspects anywhere in the original review.

Step 4: Check the updated review for naturalness and coherence.

Step 5: Return the updated review.

Review: After living 25 years thinking Greek food was like the kind served in Daphne's Cafes, I was blown away by Zorba's. Each dish has obviously been nurtured to the best possible quality and flavor, seasoned and cooked into perfection. You're probably discarding my opinion since my comparison is like Kraft Singles vs. smoked gouda, but hey, I love Daphne's and fake cheddar slices. They remind me of college :)The service is also great. I've been there twice, the first time as a save haven from the

soreness and cold from visiting the Eastern State Penitentiary in winter. I had no idea I'd walk into Greek heaven! It also happened to be Valentine's Day, so we got free heart-shaped ice cream dessert!

ChatGPT response: *Review: After living 25 years thinking Greek food was like the kind served in Daphne's Cafes, I was blown away by Zorba's. Each dish has obviously been nurtured to the best possible quality and flavor, seasoned and cooked into perfection. You're probably discarding my opinion since my comparison is like Kraft Singles vs. smoked gouda, but hey, I love Daphne's and fake cheddar slices. They remind me of college :)The service is also great. I've been there twice, the first time as a save haven from the soreness and cold from visiting the Eastern State Penitentiary in winter. I had no idea I'd walk into Greek heaven! It also happened to be Valentine's Day, so we got free heart-shaped ice cream dessert! Oh, and did I mention, Zorba's also has **live Greek music performances** every Saturday night, which adds an amazing cultural touch to the dining experience.*

Ex.2 ChatGPT Input prompt: *Follow these steps to answer the user queries.*

Step 1: Read the review and the atypical aspect given by the user.

Step 2: Incorporate the atypical aspect into the original review in the same writing style as the original review. Add one or two sentences having the atypical aspect anywhere in the original review.

Step 3: Check the updated review for naturalness and coherency.

Step 4: Return the updated review.

Atypical Aspect: painting classes

Review: Wow! It happened again! My weekend was just weird! My cousin died, and we found out he was related to frank sinatra! See that's weird! So i go to Firecracker and sure enough Katie Brown is there! She says to me, gee Rick did anybody ever tell you you look like Frank Sinatra. I said no. Before too long she had me singing "I did it my way". I received over \$50 in tips. Then a guy approached me and asked me if

i would sing at the atomic cowboy. After resisting he offered me a \$300 advance. I took it! Will be setting up date soon. Katie then encouraged me to get lady gaga to sing with me. She called her. We are all set. That katie is the best bartender alive!

ChatGPT response: *Wow! It happened again! My weekend was just weird! My cousin died, and we found out he was related to Frank Sinatra! See that's weird! But you know what else is weird? I decided to try something new and took painting classes at Firecracker. So, when I went there, Katie Brown, the bartender, said to me, "Gee Rick, did anybody ever tell you you look like Frank Sinatra?" I said no. Before too long, she had me singing "I did it my way" while painting on the canvas. It was such a unique experience! I even received over \$50 in tips for my singing and painting skills. Then, a guy approached me and asked if I would sing at the Atomic Cowboy. After resisting at first, he offered me a \$300 advance, and I just couldn't resist. I'll be setting up the date soon. Katie, being the amazing bartender that she is, even encouraged me to get Lady Gaga to sing with me. And guess what? She called her, and now we are all set to perform together. That Katie is truly the best bartender alive!*

Table 4.2 shows summary statistics for the 3 datasets, one for each domain (category), split between original data used for training and testing, data created using artificial methods Type 1 and Type 2, and data used for development. Under the Primary column, I show the number of atypical reviews and atypical aspect annotations contained in them. The next column shows the same statistics for when both primary and secondary atypical aspects are considered. The total number of reviews in each dataset, shown in the third column, is about double the number of primary atypical reviews, reflecting a balanced dataset where the number of typical reviews was selected to be about the same as the number of atypical reviews.

To create an artificial dataset I used a temperature between 0.6 to 1.2 and a number of samples equal to 6. I have manually verified the coherency of 6 samples from Chat-

Table 4.2: Statistics for the 3 datasets, split between Train+Test, Artificial Train Type 1 (ATT-1), Type 2 (ATT-2) and Development (Dev). The number of atypical reviews and atypical aspects are presented separately for primary atypical aspects (Primary) vs. both primary and secondary atypical aspects (+ Secondary).

Domain	Dataset	Primary		+ Secondary		Typical reviews
		# reviews	# aspects	# reviews	# aspects	
Restaurant	Train+Test	100	253	107	340	93
	ATT-1	50	75	50	76	50
	ATT-2	50	71	50	71	50
	Dev	14	32	16	46	14
Hotel	Train+Test	69	274	85	401	65
	ATT-1	50	66	51	90	49
	ATT-2	50	77	50	88	50
	Dev	10	33	11	49	10
Hair Salon	Train+Test	45	147	48	181	42
	ATT-1	50	72	50	72	50
	ATT-2	50	68	50	68	50
	Dev	5	24	5	29	5

GPT output and selected one in order to create a lexically diverse dataset. While ChatGPT is able to generate phrases in the style of a review, such as "Oh, did I mention", when the same prompt is used with a different review, it tends to generate the same phrase again. To alleviate this behavior, I change the temperature and use sampling. Minor manual updates are done for some reviews to remove additional/repetitive words e.g. repetitive use of words such as 'unique' and 'resident' in 'unique touch, unique activity, resident pianist, resident painter'.

To come up with atypical aspects for the Type 2 data generation method, I first looked at the atypical aspects generated by ChatGPT, and abstracted a taxonomy of general categories for atypical aspects that make sense to appear in that domain. For example, "free English classes for immigrants" is categorized into an "Educational" category, whereas "massage lounge" goes into a "Personal care" category. The resulting taxonomy of atypical aspects for the Restaurant domain is shown in Table 4.3. After the categories were created, for each category I manually added new

atypical aspects that could be surprising/useful for a user while visiting the domain (restaurants/hotels). These aspects together with the unused Type 1 aspects are then used for Type 2 artificial data generation. The list of atypical aspects used for Type 2 data generation for each domain are shown in Appendix B.1, Appendix B.2 and Appendix B.3.

Table 4.3: Taxonomy of restaurant atypical aspects

Category	Atypical Aspects	
Retail	<ul style="list-style-type: none"> • Books • Hiking gear 	<ul style="list-style-type: none"> • Travel items
Entertainment	<ul style="list-style-type: none"> • Passive <ul style="list-style-type: none"> – Live music – Magic shows 	<ul style="list-style-type: none"> • Active <ul style="list-style-type: none"> – Games – Dancing
Personal care	<ul style="list-style-type: none"> • Counseling 	<ul style="list-style-type: none"> • Spa, massage, hair, etc.
Socializing/dating services	<ul style="list-style-type: none"> • Speed dating, networking 	
Activities	<ul style="list-style-type: none"> • Food(fishing, picking, etc.) 	<ul style="list-style-type: none"> • Children activities • Petting animals
Educational	<ul style="list-style-type: none"> • Free English classes 	<ul style="list-style-type: none"> • Workshops
Cultural	<ul style="list-style-type: none"> • Museum 	<ul style="list-style-type: none"> • Festivals

Continued on next page

Table 4.3 – *Continued from previous page*

Tourism	<ul style="list-style-type: none"> • Tours • Bike rentals
Nature	<ul style="list-style-type: none"> • Garden
Community service	<ul style="list-style-type: none"> • Donation booths
Location	<ul style="list-style-type: none"> • On the water in a boat
Misc	<ul style="list-style-type: none"> • Composting program

CHAPTER 5: XtrAtA APPROACHES USING LANGUAGE MODELS

This chapter describes language model (LM) based approaches to detecting surprising aspects in customer reviews. I experimented with 2 language models: Flan-T5 and ChatGPT. The 3 billion parameter FLAN-T5 is an encoder-decoder transformer based on the T5 model [29] that was further instruction-tuned on the FLAN dataset [40, 6]. I decided to use the FLAN-T5 model due to its exposure to narratives in the style of reviews, e.g., blog posts, during pre-training on the C4 corpus [29], and also due to its instruction-tuning on summarization and sentiment analysis tasks. I also experimented with zero-shot and few-shot prompting of the much larger ChatGPT (`gpt-3.5-turbo`) [26] in order to evaluate the performance of a state-of-the-art language model without any fine-tuning.

With the exception of abstractive generation for Hotels, which did not benefit from a prompt, the FLAN-T5 fine-tuning experiments employed the following prompts:

- **Fine-tuning FLAN-T5 Extractive Prompt:** *question: Based on the following restaurant review, list aspects that are atypical for a restaurant. Separate them using commas. context: {{Review}}*
- **Fine-tuning FLAN-T5 Abstractive Prompt:** *question: Based on the following restaurant review, what are the atypical aspects for a restaurant? context: {{Review}}*

In the 0-shot setup for ChatGPT, I include an instruction to either extract lists of atypical aspects (extractive) or to generate naturally sounding text about the atypical aspects in the review (abstractive):

- **0-shot ChatGPT Extractive Prompt:** *Given the following restaurant review,*

can you list atypical aspects for a restaurant? Atypical aspects are not related to service, food, drinks, location, price, menu, discounts, policies, staff, customer satisfaction, or other items commonly associated with a restaurant. Please be precise in your response; it should contain only atypical aspects associated with the restaurant that is reviewed. Extract base noun phrases in the output format as below: 'Atypical aspects: aspect 1, aspect 2, aspect 3.' Output `<None>` if there are no atypical aspects. Please follow the output format strictly.

Passage: {{Review}}

- **0-shot ChatGPT Abstractive Prompt:** *Which aspects mentioned in the review are atypical for a restaurant? Unlike common aspects such as service, food, drinks, location, price, menu, discounts, policies, staff, or customer satisfaction, atypical aspects are not commonly associated with a restaurant. In the output, formulate each aspect as sentences, e.g., "Atypical aspects:*

- The restaurant has `<aspect 1>`.*
- The restaurant has `<aspect 2>`.*
- The restaurant has `<aspect 3>`."*

If there are no atypical aspects, output "None".

Passage: {{Review}}

In the few-shot setup for ChatGPT, I include in the prompt both the instruction and 5 worked-out examples:

- **Few-shot ChatGPT Extractive and Abstractive Prompt:** *Given the following restaurant review, can you list atypical aspects for a restaurant? Atypical aspects are not related to service, food, drinks, location, price, menu, discounts, policies, staff, customer satisfaction or other types of items that are commonly*

associated with a restaurant. Please be precise in your response, which should contain only atypical aspects that are associated with the restaurant that is reviewed. Output <None> if there are no atypical aspects.

Example 1: {{Example Review 1}}

Atypical aspects: {{comma-separated extractive annotations OR bullet-listed abstractive sentences}}

...

Example 5: {{Example Review 5}}

Atypical aspects: {{comma-separated extractive annotations OR bullet-listed abstractive sentences}}

Can you try for the restaurant review below? {{Review}}

I use the Hugging Face Transformers package [41] for fine-tuning Flan-T5 with the following hyper-parameters: an effective batch size of 32, a number of epochs of 30, a learning rate of 3e-5 for Restaurants and 5e-5 for Hotels and Hair Salons, a weight decay of 0.001, and a generation max length set to 512. Those hyper-parameter values were found through tuning on the development portion of each dataset. I perform the fine-tuning experiments on a high-performance computing cluster using 8 CPU cores, 128 GB RAM, and 2 A100 80 GB GPUs, for around 96 hours. I use the OpenAI API Python package [27] for the ChatGPT experiments, where I do greedy decoding by setting the temperature parameter to 0.

CHAPTER 6: EXPERIMENTAL EVALUATIONS

The LM-based approaches are evaluated in a 10-fold evaluation scenario where the Train+Test review dataset is partitioned into 10 folds, 9 folds are used for training and 1 fold is used for testing. This process is repeated 10 times until each fold in the dataset is used as a test fold. The evaluation metrics computed across the 10 folds are then micro-averaged yielding the final evaluation metric. For the extractive evaluation, I report the Precision (P), Recall (R), and F1 scores using exact [F.1] and partial [F.2] matches of the extracted base noun phrase (BNP) with the ground truth phrase. For the abstractive evaluation, I follow prior work in summarization [37, 1, 21] and compare the generated output with the ground truth using BERT F1 Score [43] instantiated with DeBERTa [8], Rouge-1, Rouge-2, and Rouge-L-Sum [19].

F.1 Extractive Exact Match:

In the exact match method, an entire ground truth (gold) BNP is matched with extracted whole BNP.

$$P = \# \text{ correct extracted BNPs} / \# \text{ total extracted BNPs}$$

$$R = \# \text{ correct extracted BNPs} / \# \text{ gold BNPs}$$

F.2 Extractive Partial Match:

In the partial match method, I use a greedy method [Algorithm.1] to compute a bipartite matching between gold phrases (GP) and extracted phrases (EP).

$$P = TP_e / (TP_e + FP) \quad R = TP_g / (TP_g + FN)$$

where, TP_e = True positives w.r.t extracted BNPs

TP_g = True Positives w.r.t gold BNPs

FP = False Positives, FN = False Negatives

Table 6.1: Extractive (exact vs. partial match) and abstractive results (%), on primary (default) and primary + secondary atypical aspects across the 3 domains. Precision (P), Recall (R), and F1 scores are reported for the extractive setting. The F1 scores of Rouge-1 (R-1), Rouge-2 (R-2), rougeLsum (RLS) and BERTScore (BS) are reported for the abstractive setting. Best results in each domain are in **bold**.

Language Model & Experimental Setup	Exact Match			Partial Match			Abstractive			
	P	R	F1	P	R	F1	R-1	R-2	RLS	BS
Restaurants										
ChatGPT (0-shot)	22.0	37.7	27.7	27.9	56.8	37.4	35.0	24.0	52.0	66.0
+ secondary	25.1	35.2	29.3	32.5	56.0	41.2	42.0	29.0	53.0	68.0
ChatGPT (5-shot)	33.0	43.6	37.6	39.3	61.7	48.0	42.0	31.0	59.0	71.0
+ secondary	26.8	38.9	31.7	33.5	60.7	43.2	46.0	36.0	57.0	70.0
FLAN-T5 (0-shot)	28.9	23.5	25.9	32.4	28.7	30.5	32.0	24.0	44.0	54.0
+ secondary	27.9	18.5	22.3	32.3	23.9	27.4	30.0	22.0	44.0	56.0
FLAN-T5 (fine-tuned)	67.5	60.2	63.4	72.9	65.3	68.6	58.0	50.0	73.0	79.0
+ secondary	57.7	58.6	56.4	64.2	64.4	62.3	56.0	48.0	68.0	75.0
Hotels										
ChatGPT (5-shot)	30.5	35.2	32.7	34.8	46.9	39.9	34.0	26.0	56.0	65.0
FLAN-T5 (fine-tuned)	60.2	54.9	55.9	63.8	57.6	59.0	34.0	26.0	59.0	63.0
Hair Salons										
ChatGPT (5-shot)	34.7	43.2	38.5	42.4	61.1	50.1	48.0	37.0	57.0	69.0
FLAN-T5 (fine-tuned)	66.6	62.9	63.9	75.1	69.1	71.1	38.0	33.0	61.0	65.0
Human ITA (est.)	92.3	82.8	87.3	94.2	85.1	89.4	59.0	41.0	78.0	91.0
+ secondary	74.2	67.6	70.8	79.0	71.8	75.3	65.0	47.0	77.0	91.0

The overall experimental results are shown in Table 6.1. For Restaurants I show results on extracting primary atypical aspects as well as results on extracting both primary and secondary atypical aspects. Since fine-tuned Flan-T5 and ChatGPT (5-shot) obtained the best results on Restaurants, they were selected to be evaluated on the other two domains, using solely primary atypical aspects. Fine-tuning FLAN-T5 yields the best performance in the extractive task across all domains. While I observe a big performance gap between ChatGPT and fine-tuned FLAN-T5 in the extractive setting, that gap shrinks considerably in the abstractive setting for Hotels and Hair Salons, where ChatGPT (5-shot) occasionally outperforms FLAN-T5 on some of the metrics. For both LMs the Hotel domain appears to be more challenging.

Algorithm 1 An algorithm to compute Extractive Partial Match metrics

```

for  $gp$  in GP do
  if EP is empty then
     $FN \leftarrow FN + |gp|/|gp|$ 
  else Find the extracted phrase  $ep$  in EP that best overlaps with  $gp$  (in terms
  of maximizing Jaccard similarity)
     $TPe \leftarrow TPe + |ep \wedge gp|/|ep|$   $\triangleright$  # matching tokens between  $ep$  and  $gp$ 
     $TPg \leftarrow TPg + |ep \wedge gp|/|gp|$ 
     $FP \leftarrow FP + |ep - gp|/|ep|$   $\triangleright$  # tokens of  $ep$  that are not present in  $gp$ 
     $FN \leftarrow FN + |gp - ep|/|gp|$   $\triangleright$  # tokens of  $gp$  that are not present in  $ep$ 
  end if
  Remove  $ep$  from the set EP.
end for
for  $ep$  in EP do
   $FP \leftarrow FP + |ep|/|ep|$ 
end for

```

Compared to the other domains, atypical aspects are more common and more diverse in hotels, likely because hotels try to differentiate themselves from other hotels more than restaurants or hair salons do. Table 4.1 shows that indeed there are more primary and secondary atypical aspects per review in the hotel domain.

To determine how well Flan-T5 generalizes to unseen atypical aspects in the Restaurant domain, I manually created groupings of atypical aspects where semantically similar atypical aspects, e.g. greeting cards and anniversary gifts, are grouped together, such that aspects in different groups are semantically very different. I then partition the set of groups into 10 folds of groups, which ensures that the atypical aspects that the language model sees in the test fold have not been seen during training (either literally or semantically similar). Upon fine-tuning and evaluating FLAN-T5 on this dataset, I observe a similar precision as reported in Table 6.1, however there is a significant drop in recall from 60.2 to 46.1 for primary atypical aspects and from 58.6 to 49.3 when extracting both primary and secondary atypical aspects. Improving generalization to semantically novel atypical aspects is therefore an interesting avenue for future work.

Error analysis reveals that fine-tuned Flan-T5 is more succinct in its answers, leading it to sometimes ignore atypical aspects in its response. Conversely, ChatGPT tends to be more verbose, often generating unnecessary details about the atypical aspects that it extracts, or mistaking typical for atypical aspects.

6.1 Experiments Using Artificially Augmented Reviews

In these experiments, the artificially augmented reviews are added to the original 9 training folds, and the fine-tuned LM is tested on the original test fold. This process is repeated 10 times until each fold in the original dataset is used as a test fold. These experiments are done for primary aspects only for each domain.

The addition of artificial examples during training leads to a 3-4% increase in the performance for extractive tasks on all domains as reported in Table 6.2. For the abstractive task, performance increases up to 5% in the case of the restaurant domain, 3% in the case of the hotel domain but in the case of the hair salon domain performance did not improve as much.

As ChatGPT is very sensitive to prompts and the current artificial dataset is created using only 2 or 3 prompts, the artificial review examples are not as lexically diverse and natural when compared to the original review examples. Also, 32% of the artificial atypical reviews contain more than one atypical aspect however in the original dataset 62% of atypical reviews contain more than one atypical aspect. Therefore to improve generalization performance further using artificial datasets, more prompts and atypical aspects could be used in future work in order to make the artificial data as similar to the real data as possible.

Table 6.2: Extractive (exact vs. partial match) and abstractive results (%), on primary atypical aspects across the 3 domains. Precision (P), Recall (R), and F1 scores are reported for the extractive setting. The F1 scores of Rouge-1 (R-1), Rouge-2 (R-2), rougeLsum (RLS) and BERTScore (BS) are reported for the abstractive setting. Best, and second best results in each domain are in **bold**, underline respectively.

Language Model & Experimental Setup	Exact Match			Partial Match			Abstractive			
	P	R	F1	P	R	F1	R-1	R-2	RLS	BS
Restaurants										
ChatGPT (0-shot)	22.0	37.7	27.7	27.9	56.8	37.4	35.0	24.0	52.0	66.0
ChatGPT (5-shot)	33.0	43.6	37.6	39.3	61.7	48.0	42.0	31.0	59.0	71.0
FLAN-T5 (0-shot)	28.9	23.5	25.9	32.4	28.7	30.5	32.0	24.0	44.0	54.0
FLAN-T5 (fine-tuned)	67.5	60.2	<u>63.4</u>	72.9	65.3	<u>68.6</u>	58.0	50.0	<u>73.0</u>	<u>79.0</u>
+ Artificial Data	69.7	63.7	66.3	75.6	69.3	71.8	62.0	53.0	77.0	84.0
Hotels										
ChatGPT (5-shot)	30.5	35.2	32.7	34.8	46.9	39.9	34.0	26.0	56.0	<u>65.0</u>
FLAN-T5 (fine-tuned)	60.2	54.9	<u>55.9</u>	63.8	57.6	<u>59.0</u>	34.0	26.0	<u>59.0</u>	<u>63.0</u>
+ Artificial Data	61.8	60.3	59.3	65.3	63.0	62.1	48.0	34.0	62.0	66.0
Hair Salons										
ChatGPT (5-shot)	34.7	43.2	38.5	42.4	61.1	<u>50.1</u>	48.0	37.0	57.0	69.0
FLAN-T5 (fine-tuned)	66.6	62.9	<u>63.9</u>	75.1	69.1	71.1	38.0	33.0	<u>61.0</u>	65.0
+ Artificial Data	71.6	66.5	67.7	75.4	69.7	71.1	58.0	52.0	62.0	<u>67.0</u>
Human ITA (est.)	92.3	82.8	87.3	94.2	85.1	89.4	59.0	41.0	78.0	<u>91.0</u>

CHAPTER 7: CONCLUSION

I introduced the new task of extracting atypical aspects from customer reviews. Items with aspects that are atypical for their category can be promoted in search results or recommendation lists in order to alleviate overchoice, on the premise that atypical items are likely to be perceived as surprising and thus to lead to serendipitous and more memorable experiences. To enable the training and evaluation of atypical aspect extraction models, I manually annotated two layers of atypical aspects in customer reviews from three domains. Experimental evaluations show that fine-tuning of Flan T5 obtains better accuracy than few-shot prompting of GPT 3.5. Even though fine-tuning on reviews that are artificially augmented with atypical aspects leads to further gains in performance, the LM performance is still far from human performance, justifying future efforts in this area.

REFERENCES

- [1] Ojas Ahuja, Jiacheng Xu, Akshay Gupta, Kevin Horecka, and Greg Durrett. ASPECTNEWS: Aspect-oriented summarization of news documents. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6494–6506, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-long.449. URL <https://aclanthology.org/2022.acl-long.449>.
- [2] Daniel E Berlyne. Aesthetics and psychobiology. *Journal of Aesthetics and Art Criticism*, 31(4), 1973.
- [3] Dirk Bollen, Bart P. Knijnenburg, Martijn C. Willemsen, and Mark Graus. Understanding choice overload in recommender systems. In *Proceedings of the Fourth ACM Conference on Recommender Systems, RecSys '10*, page 63â70, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605589060. doi: 10.1145/1864708.1864724. URL <https://doi.org/10.1145/1864708.1864724>.
- [4] G.N. Carlson and F.J. Pelletier. *The Generic Book*. University of Chicago Press, 1995. ISBN 978-0-226-09292-8.
- [5] Alexander Chernev, Ulf BÄ¶ckenholt, and Joseph Goodman. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2):333–358, 2015. ISSN 1057-7408. doi: <https://doi.org/10.1016/j.jcps.2014.08.002>. URL <https://www.sciencedirect.com/science/article/pii/S1057740814000916>.
- [6] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen,

- Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. Scaling instruction-finetuned language models, 2022.
- [7] Xiangyu Fan and Xi Niu. Implementing and evaluating serendipity in delivering personalized health information. *ACM Trans. Manage. Inf. Syst.*, 9(2), aug 2018. ISSN 2158-656X. doi: 10.1145/3205849. URL <https://doi.org/10.1145/3205849>.
- [8] Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint arXiv:2006.03654*, 2020.
- [9] Yoel Inbar, Simona Botti, and Karlene Hanko. Decision speed and choice regret: When haste feels like waste. *Journal of Experimental Social Psychology*, 47(3):533–540, 2011. ISSN 0022-1031. doi: <https://doi.org/10.1016/j.jesp.2011.01.011>. URL <https://www.sciencedirect.com/science/article/pii/S002210311100028X>.
- [10] Sheena S. Iyengar and Mark R. Lepper. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79:995–1006, 2000. ISSN 1939-1315. doi: 10.1037/0022-3514.79.6.995.
- [11] Niklas Jakob, Stefan Hagen Weber, Mark Christoph Müller, and Iryna Gurevych. Beyond the stars: Exploiting free-text user reviews to improve the accuracy of movie recommendations. In *Proceedings of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion*, TSA '09, page 57â64, New York, NY, USA, 2009. Association for Computing Machinery. ISBN

9781605588056. doi: 10.1145/1651461.1651473. URL <https://doi.org/10.1145/1651461.1651473>.
- [12] Hongyan Jing. Using hidden Markov modeling to decompose human-written summaries. *Computational Linguistics*, 28(4):527–543, 2002. doi: 10.1162/089120102762671972. URL <https://aclanthology.org/J02-4006>.
- [13] Hongyan Jing and Kathleen R. McKeown. The decomposition of human-written summary sentences. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '99, page 129–136, New York, NY, USA, 1999. Association for Computing Machinery. ISBN 1581130961. doi: 10.1145/312624.312666. URL <https://doi.org/10.1145/312624.312666>.
- [14] Daniel Kaimann, Nadja Stroh-Maraun, and Joe Cox. Variety in the video game industry: An empirical study of the wundt curve. *Managerial and Decision Economics*, 39(3):354–362, 2018. doi: <https://doi.org/10.1002/mde.2909>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mde.2909>.
- [15] Kevin Knight and Daniel Marcu. Summarization beyond sentence extraction: A probabilistic approach to sentence compression. *Artificial Intelligence*, 139(1):91–107, 2002. ISSN 0004-3702. doi: [https://doi.org/10.1016/S0004-3702\(02\)00222-9](https://doi.org/10.1016/S0004-3702(02)00222-9). URL <https://www.sciencedirect.com/science/article/pii/S0004370202002229>.
- [16] Denis Kotkov, Joseph A. Konstan, Qian Zhao, and Jari Veijalainen. Investigating serendipity in recommender systems based on real user feedback. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, SAC '18, page 1341–1350, New York, NY, USA, 2018. Association for Computing Machinery.

ISBN 9781450351911. doi: 10.1145/3167132.3167276. URL <https://doi.org/10.1145/3167132.3167276>.

- [17] Pan Li and Alexander Tuzhilin. Learning latent multi-criteria ratings from user reviews for recommendations. *IEEE Transactions on Knowledge and Data Engineering*, 34(8):3854–3866, 2022. doi: 10.1109/TKDE.2020.3030623.
- [18] Pan Li, Maofei Que, Zhichao Jiang, YAO HU, and Alexander Tuzhilin. Purs: Personalized unexpected recommender system for improving user satisfaction. In *Proceedings of the 14th ACM Conference on Recommender Systems, RecSys '20*, page 279â288, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450375832. doi: 10.1145/3383313.3412238. URL <https://doi.org/10.1145/3383313.3412238>.
- [19] Chin-Yew Lin. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain, July 2004. Association for Computational Linguistics. URL <https://aclanthology.org/W04-1013>.
- [20] Yu Lu, Junwei Bao, Yan Song, Zichen Ma, Shuguang Cui, Youzheng Wu, and Xiaodong He. RevCore: Review-augmented conversational recommendation. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1161–1173, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.findings-acl.99. URL <https://aclanthology.org/2021.findings-acl.99>.
- [21] Khalil Mrini, Can Liu, and Markus Dreyer. Rewards with negative examples for reinforced topic-focused abstractive summarization. In *Proceedings of the Third Workshop on New Frontiers in Summarization*, pages 33–38, Online and in Dominican Republic, November 2021. Association for Computational Linguistics.

- doi: 10.18653/v1/2021.newsum-1.4. URL <https://aclanthology.org/2021.newsum-1.4>.
- [22] Matteo Muffo, Aldo Cocco, Edoardo Negri, Enrico Bertino, Devi Sreekumar, Giorgio Pennesi, and Riccardo Lorenzon. Sbertiment: A new pipeline to solve aspect based sentiment analysis in the zero-shot setting. *The International FLAIRS Conference Proceedings*, 36, 05 2023. doi: 10.32473/flairs.36.133058.
- [23] Cataldo Musto, Marco de Gemmis, Giovanni Semeraro, and Pasquale Lops. A multi-criteria recommender system exploiting aspect-based sentiment analysis of users’ reviews. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, RecSys ’17, page 321â325, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346528. doi: 10.1145/3109859.3109905. URL <https://doi.org/10.1145/3109859.3109905>.
- [24] Cataldo Musto, Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro. Justifying recommendations through aspect-based sentiment analysis of users reviews. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, UMAP ’19, page 4â12, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450360210. doi: 10.1145/3320435.3320457. URL <https://doi.org/10.1145/3320435.3320457>.
- [25] Xi Niu, Fakhri Abbas, Mary Lou Maher, and Kazjon Grace. Surprise me if you can: Serendipity in health information. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, page 1â12, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. doi: 10.1145/3173574.3173597. URL <https://doi.org/10.1145/3173574.3173597>.
- [26] OpenAI. Introducing chatgpt, 2022. URL <https://openai.com/blog/chatgpt>.

- [27] OpenAI. Openai python library, 2022. URL <https://github.com/openai/openai-python>.
- [28] Adamopoulos Panagiotis and T Alexander. On unexpectedness in recommender systems: Or how to expect the unexpected. In *Workshop on Novelty and Diversity in Recommender Systems*, volume 2011, 2011. URL <https://ceur-ws.org/Vol-816/paper2.pdf>.
- [29] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 21(1):5485–5551, 2020.
- [30] Alexander M. Rush, Sumit Chopra, and Jason Weston. A neural attention model for abstractive sentence summarization. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 379–389, Lisbon, Portugal, September 2015. Association for Computational Linguistics. doi: 10.18653/v1/D15-1044. URL <https://aclanthology.org/D15-1044>.
- [31] Benjamin Scheibehenne, Rainer Greifeneder, and Peter M. Todd. Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload. *Journal of Consumer Research*, 37(3):409–425, February 2010. ISSN 0093-5301. doi: 10.1086/651235. URL <https://doi.org/10.1086/651235>. _eprint: <https://academic.oup.com/jcr/article-pdf/37/3/409/5173186/37-3-409.pdf>.
- [32] Barry Schwartz. *The paradox of choice: Why more is less*. Harper Perennial, New York, 2004.
- [33] Safa Selmene and Zahra Kodia. Recommender system based on user’s tweets sentiment analysis. In *Proceedings of the 4th International Conference on E-Commerce, E-Business and E-Government, ICEEG ’20*, page 96â102, New York,

- NY, USA, 2020. Association for Computing Machinery. ISBN 9781450388030. doi: 10.1145/3409929.3414744. URL <https://doi.org/10.1145/3409929.3414744>.
- [34] SemEval. Task description: Aspect based sentiment analysis (absa) < semeval-2014 task 4, 2014. URL <https://alt.qcri.org/semeval2014/task4/>.
- [35] Avni M Shah and George L. Wolford. Buying behavior as a function of parametric variation of number of choices. *Psychological Science*, 18:369 – 370, 2007.
- [36] spaCy. Industrial-strength natural language processing, 2023. URL <https://spacy.io/>.
- [37] Rahul Tangsali, Aditya Jagdish Vyawahare, Aditya Vyankatesh Mandke, Onkar Rupesh Litake, and Dipali Dattatray Kadam. Abstractive approaches to multidocument summarization of medical literature reviews. In *Proceedings of the Third Workshop on Scholarly Document Processing*, pages 199–203, Gyeongju, Republic of Korea, October 2022. Association for Computational Linguistics. URL <https://aclanthology.org/2022.sdp-1.24>.
- [38] Alvin Toffler. *Future shock*. Random House, New York, 1970.
- [39] Kathleen D. Vohs, Roy F. Baumeister, Brandon J. Schmeichel, Jean M. Twenge, Noelle M. Nelson, and Dianne M. Tice. Making choices impairs subsequent self-control: a limited-resource account of decision making, self-regulation, and active initiative. *Journal of Personality and Social Psychology*, 94(5):883–898, May 2008. ISSN 0022-3514. doi: 10.1037/0022-3514.94.5.883.
- [40] Jason Wei, Maarten Bosma, Vincent Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V Le. Finetuned language models are zero-shot learners. In *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=gEZrGCozdqR>.

- [41] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, 2020.
- [42] Yelp. The yelp dataset, 2023. URL <https://www.yelp.com/dataset>.
- [43] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*, 2020. URL <https://openreview.net/forum?id=SkeHuCVFDr>.
- [44] Yongfeng Zhang, Guokun Lai, Min Zhang, Yi Zhang, Yiqun Liu, and Shaoping Ma. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '14, page 83â92, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450322577. doi: 10.1145/2600428.2609579. URL <https://doi.org/10.1145/2600428.2609579>.
- [45] Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W. Bruce Croft. Towards conversational search and recommendation: System ask, user respond. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 177â186, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450360142. doi: 10.1145/3269206.3271776. URL <https://doi.org/10.1145/3269206.3271776>.

APPENDIX A: ANNOTATED EXAMPLES

A.1 Hotels Domain

► This was an awesome stop for me and my dog after a long day of travel. Hotel McCoy has a whimsical, Route 66 vibe with a modern sensibility. After an incredibly exit off the freeway, We were welcomed by super friendly staff. Our room was comfortable and cool upon entry and features a retro fridge, chalk and a chalkboard on back of the door, and a nice book display. An overnight stay included a free drink at the very comfortable bar and a complimentary breakfast delivered to your room at your time of choice! There is a nice wide sidewalk out front to walk your pet as well. The buildings feature whimsical, colorful murals. Finally, this place features a **fabulous small bookstore** called **Barrio Books**! I highly recommend Hotel MCCoy if you are looking for accommodations in Tucson. Cost is SO reasonable and SO many pluses.

The hotel features a fabulous small bookstore called Barrio Books.

► We stayed here for a night and loved it. The rooms and hotel have a charm, nicely decorated boutique style, cozy, and comfortable. They have a nice rooftop bar and pool and pretty much smack in the middle of State St. Everything is within walking distance on State St.

We loved the bed with pillow top mattress and I liked the plush pillows. There is a **bird book** and a pair of **binoculars** in every room to bird (or people) watch from the balcony. The bathroom was nice in the room we had, it had a TV and a nice bath tub/shower, soft towels, and good water pressure...

Continued on next page

Table A.1 – *Continued from previous page*

The hotel offers a bird book and a pair of binoculars in every room to bird watch from the balcony.

► The largest non-casino hotel in the US definitely lives up to it's title. This place is huge and it is quite easy to get lost in it. Having been to the other location in Orlando, I can say that I prefer this location. The staff are extremely helpful and nice, and the food choices, although priced for a resort, are plentiful. Going for a convention is great, as you really don't need to go out for much if you don't want to. It is within walking distance to the [Opry Mills mall](#), the [Grand Ole Opry](#) (they also have a **free shuttle** there), and they have a **shuttle** to downtown Nashville for a nominal fee (much less than taxi).

The hotel has a free shuttle to Grand Ole Opry. They have a shuttle to downtown Nashville for a nominal fee (much less than taxi).

[It is within walking distance to the Opry Mills mall, the Grand Ole Opry.](#)

Continued on next page

Table A.1 – *Continued from previous page*

► My friend M says: "They get everything wrong, but they're so nice about it!"

The Shining comparisons can't be avoided – long windy hallways, dark wood trim. Like many Doubletrees, it's a 1970's hotel, but I didn't see any writing on the mirrors. I had no choice in the hotel – I go where the meetings are – but the three stars come from my highly scientific pro/con meeting hotel evaluation formula(TM).

Pros: Cheap (\$88 - \$100 a night), soft pillows, **shuttle** to airport & downtown, decent gym and outdoor pool, lots of meeting rooms, good tea & coffee, rose garden makes for a fragrant break, location right on the **riverside bike path** offers **jogging/fishing/blackberry picking options**.

Cons: Far from town, depressing architecture, very limited Wifi, the food, bar prices are high, if you get the one cranky lady at the front desk she won't let the shuttle take you downtown even though you had it booked already.

The hotel has a shuttle to downtown.

The hotel's location is right on the riverside bike path that offers jogging/fishing/blackberry picking options.

Table A.1: Examples of **extractive** and *abstractive* annotations of customer reviews in the hotels domain. Secondary or optional annotations are shown in **blue**.

A.2 Hair Salons Domain

► I've been coming to this salon for a long time. The staff is very friendly and the building is nicely decorated. Katie Chavez has been doing my hair for the last 13 years. I have followed her to a couple new salons, but I'm so happy she decided to make her home here. The place is clean and modern. I believe they offer **exercise classes**, but I've never attended on. If I could give this salon 6 stars, I would!

The salon offers exercise classes.

► Such an amazing space. Guaranteed - you will become just a little bit cooler and more stylish from even just walking into the place. Not only is the environment and decor inspiring - from the photography on the walls to the vintage furniture everywhere, but they also sell all kinds of things that you're going to want to buy. They sell **vintage dresses, new designer T shirts**, and best of all - all kinds of amazing, **one of a kind jewelry pieces** and **handmade hats and scarfs**. There were easily 5 things that I wanted to buy immediately upon arrival. I decided on one new scarf and I love it! I will be back to buy some jewelry next time.

Also you will leave with a hairstyle that is flattering, edgy, and stylish. The do cuts, colors, braids, wedding hair - everything. There is a **photography studio** in the back where you can book photo shoots as well. The owners and employees are beyond friendly. You should totally check this place out for so many reasons, it's really unique and inspiring.

The salon sells vintage dresses, new designer T shirts, jewelry pieces and handmade hats and scarfs.

There is a photography studio in the back where you can book photo shoots as well.

Continued on next page

Table A.2 – *Continued from previous page*

► Ok, so I gave Icon another try. I was walking through Commerce Place when I passed a mirror and remembered I was desperately in need of a haircut. I decided to give them a try again - and I'm glad I did. A woman named Jayme cut my hair and I honestly haven't had such a hip, modern cut in long time. This time the service was exceptional, and I have to say that the salon has improved by leaps and bounds. I learned it's under new management. The studio has adopted a more **boutique-style environment**, with **artwork** on the walls and they even have a **tattoo studio** attached. The place was quiet, as it was in the middle of the afternoon during the week, but I wouldn't be surprised if weekends are packed. I was really impressed with Jayme's friendly service and she even gave me tips for styling my hair, without pushing product on me - a pet peeve of mine at salons. I will definitely go back for a haircut, and may even try a colour there next time. Keep up the good work, Icon.

The salon has a tattoo studio attached.

The salon has adopted a more boutique-style environment, with artwork on the walls.

Continued on next page

Table A.2 – *Continued from previous page*

► I decided to go to Plume on a whim since it’s just a block from my house. I was very pleased with the service and how my hair turned out. I was greeted by name upon arriving and offered water, coffee, or a glass of wine. You really can’t beat the **eclectic feel** you get in this intimate space. They sell **adorable potted plants** and **succulents**, as well as **unique jewelry**, **clothes**, and **items** that I assume are made by **local artists**. The esthetic has a **very relaxing gypsy vibe** with **beautifully mismatched furniture** and **exposed brick**. I was a fan of the bowl of candy they had in the waiting area.

I can’t remember my hair dresser’s name but she was very nice and helpful. I left it in her hands to decide which color to highlight hair and I’m so happy I did that. She suggested I go for a more golden blonde tone because my current color was washing me out. All of the products she used smelled SO good and I left a very happy camper. It’s not just a haircut appointment - It’s an experience. I will be coming back to Plume for my next cut!

The salon sells adorable potted plants and succulents, as well as unique jewelry, clothes, and items made by local artists.

You really can’t beat the eclectic feel you get in this salon. The esthetic has a very relaxing gypsy vibe with beautifully mismatched furniture and exposed brick.

Table A.2: Examples of **extractive** and *abstractive* annotations of customer reviews in the hair salons domain. Secondary or optional annotations are shown in **blue**.

APPENDIX B: ATYPICAL ASPECTS IN ARTIFICIAL DATASET

B.1 Restaurants Domain

Category	Atypical Aspects
Retail	<ul style="list-style-type: none"> • selling culinary books • car washing service at the back • small travel store that includes travel size toiletries, water bottles, granola bars, small hiking gear, etc. • anime/comic figurines for sale • guitar store
Entertainment	<ul style="list-style-type: none"> • group of clowns entertaining children • collection of sea creatures • stand-up comedy night • calligrapher to write your name in any language • outdoor movie nights • live snake charmer • collection of travel books of the places of all countries in the world • binocular for bird watching • cool science experiments for the kids
Personal care	<ul style="list-style-type: none"> • massage lounge to relax while waiting • men's grooming service • miniature therapy horse

Continued on next page

Table B.1 – *Continued from previous page*

Socializing/dating	<ul style="list-style-type: none"> • speed dating event
services	<ul style="list-style-type: none"> • find a foodie friend program • single mom day
Activities	<ul style="list-style-type: none"> • escape room • weekly ballroom dance event for locals • their monthly salsa dancing nights • you can fish your own fish in fish tank • pottery man, you can create own pottery and take home • restaurant having pool and water slide activity • small tree house for kids to play • gold panning activity • puzzles on the table to solve • camel ride
Educational	<ul style="list-style-type: none"> • karate classes for womens for self defense • free English classes for immigrants • food photography • sharing their recipe of a dish

Continued on next page

Table B.1 – *Continued from previous page*

Cultural	<ul style="list-style-type: none"> • mini-museum for cultural(Indonesian/Japanese/etc) • small museum-like space, vintage cars and motorcycles • weekly cultural performances, traditional music and dance
Tourism	<ul style="list-style-type: none"> • bicycle rental to roam around town • rent kayaks, snorkel gear • horse carriage ride around the town
Nature	<ul style="list-style-type: none"> • butterfly garden • bonsai garden with bonsai for sale • view of a mountain range
Community service	<ul style="list-style-type: none"> • donation booths where you can donate clothes, shoes, etc to needy
Location	<ul style="list-style-type: none"> • they are on the water in a boat • beside the museum • white water rafting center
Misc	<ul style="list-style-type: none"> • composting program

Table B.1: Atypical aspects of restaurants for Type 2 artificial data

B.2 Hotels Domain

Category	Atypical Aspects
Retail	<ul style="list-style-type: none"> • handmade jewelry and accessories • car washing service at the back • souvenir shop • handcrafted pottery sale • shop selling perfumes, scarfs, hats, sunglasses • toys for sale • booking of flights, train, bus • mobile accessories shop with cellphone chargers, head-phones, cables, etc.
Entertainment	<ul style="list-style-type: none"> • magician • fun karaoke area • giant aquarium • therapeutic treehouse that overlooks a serene lake • rescued farm animal sanctuary on the property • monthly stand-up comedy show with local comedians • fortune-telling booth • ice-sculpting show • live snake charmer • henna tattoo artist

Continued on next page

Table B.2 – *Continued from previous page*

Personal care	<ul style="list-style-type: none"> • men's grooming service • miniature therapy horse • hair salon
Activities	<ul style="list-style-type: none"> • maze made of hedges • water park access • indoor ice skating rink • resident llama named Simon who you can take for walks around the property • complimentary dance class • cherry picking activity • gold panning activity
Educational	<ul style="list-style-type: none"> • on-site pottery studio where guests can participate in pottery classes • complimentary music lessons • workshops on artful jewelry making • archery sessions for kids • hobby classes for kids • embroidery sessions
Cultural	<ul style="list-style-type: none"> • collection of antique clocks • weekly cultural performances, traditional music and dance

Continued on next page

Table B.2 – *Continued from previous page*

Tourism	<ul style="list-style-type: none"> • horse carriage rides around town • daily tours including hiking, kayaking, etc • rent ski • guided boat tour in a lake • taxi service • beach chair rental
Nature	<ul style="list-style-type: none"> • butterfly garden filled with rare species of butterflies • natural hot spring
Community service	<ul style="list-style-type: none"> • community outreach program where they hire and train individuals from underprivileged backgrounds in hospitality management • "Adopt-A-Boat" program, where you can adopt a local fishing boat and help support the preservation of the local fishing industry • donation booths where you can donate clothes, shoes, etc to needy
Location	<ul style="list-style-type: none"> • they are in the lake
Misc	<ul style="list-style-type: none"> • Test-drives of Tesla

Table B.2: Atypical aspects of hotels for Type 2 artificial data

B.3 Hair Salons Domain

Category	Atypical Aspects
Retail	<ul style="list-style-type: none"> • car washing service at the back • handcrafted pottery • shop selling perfumes, scarfs, hats, sunglasses • toys for sale • anime/comic figurines for sale • handmade Moroccan jewelry and accessories • cultural cookbooks for sale • antique store • eclectic antique salt and pepper shakers
Entertainment	<ul style="list-style-type: none"> • collection of travel books of the places of all countries in the world • live snake charmer • outdoor gaming area • audiobooks of short stories • caricature artist onsite • claw machine game • monthly stand-up comedy show with local comedians • miniature golf course • board games and puzzles

Continued on next page

Table B.3 – *Continued from previous page*

Personal care	<ul style="list-style-type: none"> • massage lounge to relax while waiting • dietician on-site for hair wellness
Socializing services	<ul style="list-style-type: none"> • single mom day
Activities	<ul style="list-style-type: none"> • painting studio; paint your own ceramics • coloring book and crayons • themed escape rooms • binocular for bird watching • gold panning activity • zumba sessions • design your own stickers • ice-skating rink
Educational	<ul style="list-style-type: none"> • music lessons • on-site pottery studio where guests can participate in pottery classes • free origami lessons • karate classes for womens for self defense • free English classes for immigrants • photography lessons • embroidery sessions

Continued on next page

Table B.3 – *Continued from previous page*

Cultural	<ul style="list-style-type: none"> • mini-museum of Indonesian culture • collection of antique musical instruments
Tourism	<ul style="list-style-type: none"> • ski rental • guided boat tour in a lake • borrow bicycles for free • camel ride around the town
Nature	<ul style="list-style-type: none"> • butterfly garden filled with rare species of butterflies • bonsai garden with bonsai for sale
Community service	<ul style="list-style-type: none"> • community outreach program where they hire and train individuals from underprivileged backgrounds in hair salon services • donation booths where you can donate clothes, shoes, etc to needy • pet adoption center
Misc	<ul style="list-style-type: none"> • office space for work while waiting • daycare facility

Table B.3: Atypical aspects of hair salons for Type 2 artificial data