Optimizing ECommerce Listing: LLM Based Description and Keyword Generation from Multimodal Data

Article *in* International Journal of Science and Research (IJSR) · October 2023

DOI: 10.21275/SR24304113521

3 authors:

Satish Kathiriya
San Jose State University

Mahidhar Mullapudi
Microsoft

Rajath Karangara
Prin. L. N. Welingkar Institute of Management Development & Research

Optimizing ECommerce Listing: LLM Based Description and Keyword Generation from Multimodal Data

Satish Kathiriya[1], Mahidhar Mullapudi[2], Rajath Karangara[3]

[1]Software Engineer, CA, USA

[2]Senior Software Engineer, WA, USA
American Express, U. S

Abstract: ***The way products are displayed on web pages is a crucial decision for e - commerce websites since it may significantly impact the sales of the products. This research explores the integration of Large Language Models with image recognition and meta - data analysis to auto - generate keywords and descriptions for e - commerce product listings. By analyzing visual content, item size, weight, color, and other meta - data, the LLM creates rich, accurate, and search - optimized product descriptions. This method aims to enhance discoverability and accuracy in e - commerce catalogs, reducing the workload for sellers and improving the shopping experience for buyers. This study proposes an advanced approach for e - commerce platforms using LLMs to generate product descriptions and keywords from multimodal data inputs, including images and meta - data. The research focuses on how the integration of visual and textual analysis by AI can create more detailed, accurate, and appealing product descriptions, ultimately leading to enhanced user engagement and sales conversion rates.***

Keywords: Multimodal Data, E - Commerce, Large Language Model, Keyword, Generation

1. Introduction

1.1 Background Information

In 2023, the estimated value of the worldwide e - commerce industry would reach $6.3 trillion [13, 14]. With that number expected to rise in the next few years, borderless ecommerce is clearly proving to be a lucrative choice for online merchants. More than 21.2% of all retail purchases will take place online by the year 2024 [13, 14]. With the continuous growth of subscriptions, a change in retention measures, the rising significance of AI and automation, and social platforms changing product discovery, 2024 is shaping up to be an interesting year for e - Commerce. Make use of these methods to write effective descriptions for e - Commerce products:

Strategy 1: Use Persuasive Language
Words and phrases used with the intent to influence another person's opinion or behavior are known as persuasive language. That is the act of purchasing in online retail.

Strategy 2: Focus on Benefits, Not Features
Instead, then just outlining the product's qualities, highlight the advantages that the buyer will experience while utilizing it.

Strategy 3: Use Emotional Triggers
Words, phrases, or images that have the power to stir up intense feelings in people are called emotional triggers. By appealing to the customer's emotions, you may establish a more personal connection with them and inspire them to take action.

Strategy 4: Optimizing Product Descriptions for SEO
Improve product's exposure and increase organic traffic to e - Commerce sites by making product descriptions accessible and appealing to search engines.

Strategy 5: Test and Refine
Find out what needs fixing by conducting regular evaluations. Making many variations of the description and seeing which one gets people to buy is part of this process.

1.2 Problem Statement

In the ever - changing world of online retail, search engine optimization (SEO) and consumer engagement rely on the depth and accuracy of product descriptions. Improved click - through rates, trend alignment, and elimination of the "cold start" issue are all possible outcomes of well - written product descriptions. Conventional approaches to writing these descriptions often need a lot of human intervention and could not be consistent or scalable. Instead of providing formal product descriptions, sellers on e - commerce platforms describe their offerings using textual features like titles and descriptions. Extracting attribute/value pairs from textual product attributes is essential for processing such offers in the applications. It takes a lot of effort and time to write an effective product description. There has been a lot of interest from the scholarly and industry worlds in automating the development of product descriptions because of its relevance. In this study, we investigate the possibility of using LLMs to derive attribute values from product names.

1.3 Purpose of the Research

The purpose of this article is to look at GPT - 3.5 tweaking, in - context demonstration selection, and the supply of sample attribute values. Making a system that can automatically produce product descriptions from a given language is the goal of this challenge. In addition to assisting buyers in making educated decisions, a detailed and appealing description increases the possibility that they will make a purchase. Writing an effective product description, however, is a laborious and time - consuming process. Compared to PLM - based attribute/value extraction approaches, large language models (LLMs) have the ability to use training data more efficiently and are more resilient, which is something we investigate in this study along with a novel approach to creating individualized product descriptions.

1.4 Research Questions:

1.4.1 Question

a) How to examine the possibility of using LLMs to the problem of extracting the value of product attributes.
b) To construct TaoDescribe is a new large - volume dataset for creating product descriptions.

1.4.2 Hypothesis

a) In situations with and without task - specific training data, several prompt designs are used to teach LLMs about the goal schema of the attribute value extraction task. These prompts are evaluated based on the dimensions of depth of information and representation format.
b) Concentrate on automating the creation of product descriptions for online product recommendations.

1.5 Significance of the Study

One key component of an exceptional e - commerce experience is product suggestion, the goal of which is to provide consumers with relevant material at the optimal moment. One of the most difficult things is to read consumers' minds quickly, guide them to the products they need, and be there for them every step of the way. In order to assist clients in making an educated purchase, it is essential that textual descriptions provide pertinent knowledge and precise product details. Automating the development of product descriptions for online product recommendations is our primary emphasis. Based on product details and customer choices, the system can autonomously create an appealing and accurate product description.

2. Literature Review

2.1 Previous Wor

2.1.1 Keyword Optimization

Srivastava, Shashank & Kshatriya, Suhani & Rathore, Rajkumar (2017) [1] To maximize the amount of free, targeted traffic your website receives from search engines, you need to use search engine optimization strategies. Search engine optimization (SEO) has become more popular in the digital marketing space in recent years, thanks to the proliferation of mobile devices. There is intense rivalry among e - commerce websites due to the rising popularity of online purchasing. Their company is highly dependent on their financial profits.

Wu, Huanwei (2011) [2] Websites need search engine optimization, or SEO, to increase their visibility in search engine results and the number of visitors to their pages. In this paper, we lay out the groundwork for e - commerce website optimization by discussing the unique challenges posed by dynamic web pages, redirect web pages, and similar or repeated pages. We then offer solutions to these problems, including making dynamic pages static, decreasing the number of redirect pages, and improving similar pages.

2.1.2 LLMs

Shi, et al (2023) [3] In order to increase product sales, e - commerce authors must produce promotional material that is appealing, plentiful, and tailored to the target audience. A new paradigm has emerged with the advent of large language models (LLMs), providing a cohesive answer to the many writing responsibilities at play here. But traditional LLMs trained on common - sense corpora fail miserably when it comes to fitting the intricate and individualized characteristics of e - commerce items and buyers. Additionally, there are worries about the transmission security of large amounts of consumer private data since LLMs like GPT - 3.5 need remote accessibility. The LLaMA - E, a set of unified and customizable instruction - following language models with an emphasis on various e - commerce writing activities, is proposed in this study. Domain experts construct the first set of instructions from the following tasks: advertising creation, product categorization, query - enhanced title rewriting, general Q&A, and buy intent guessing. By integrating characteristics that include common service components of clients, sellers, and platforms, these jobs allow the models to fully comprehend accurate e - commerce writing knowledge. As a teacher model, the GPT - 3.5 is shown; it builds on the seed instructions to train LLaMA - E models of different sizes. The experimental findings demonstrate that the suggested LLaMA - E models outperform the state - of - the - art models in both quantitative and qualitative assessments, and they even perform better in zero - shot scenarios. As far as we are aware, this research is the first one to use LLMs to targeted e - commerce writing situations.

Soni, Vishvesh (2023) [4] A new method that considerably improves the customer experience and company outcomes is a Large Language Model, integrates into the customer lifecycle management process. This article delves at the effects of LLM at many points in the customer lifecycle, including pre - purchase (awareness and acquisition), post - purchase (onboarding and retention), and purchase. While it comes to lead identification and targeting, LLMs show their advantage over standard AI - driven approaches at the Awareness and Acquisition stage. Improved lead quality is the result of LLMs' ability to analyze intricate patterns in consumer and market data, which allows for more precise audience segmentation and targeting. The combination of this automation with the creation of hyper - personalized message and content guarantees that prospective consumers get interesting and relevant information, which in turn increases their involvement with the brand. When it comes to the Purchase stage, LLMs play an important supporting role in the sales process. In order to streamline the sales process and improve productivity, this involves providing real - time negotiating help, predictive insights, and acting as a virtual assistant to sales teams. Benefits of tailored onboarding, ongoing assistance, and engagement with chatbot features and help from sales leadership are highlighted in the post - purchase stage. Customer retention and company development are greatly aided by LLMs' real - time advice, churn modeling, and opportunity identification capabilities.

2.1.3 Multimodal Data Analysis

Chen, Jinyu & Zhong, Ziqi & Feng, Qindi & Liu, Lei (2022) [5] Product promotion describes the ways in which e - commerce encourages customers to engage in consumption, and e - commerce has grown quickly. Improving the efficiency of e - commerce product line dynamic pricing choices requires immediate attention to the demand side and the computational complexity of the decision - making process. Consequently, we investigate the product line's dynamic pricing issue and suggest a neural network - based Q - learning algorithm model based on multimodal emotion information detection and analysis. The findings demonstrate the establishment of a multi - modal fusion model for consumer emotion classification by combining voice and visual emotion detection. Then, you may utilize them as supplementary resources to learn about and analyze the demand in the market. The LSTM classifier does a great job at extracting features from images. Compared to other comparable classifiers, the accuracy rate is 3.92–6.74% higher, and when comparing the picture and voice single - feature optimum models, the former has a 9.32 percent better accuracy rate.

Sinha, et al (2023) [6] For those who want to maximize the usability of their websites, usability testing is a complicated process with numerous components. It doesn't need any advanced knowledge; all that's required are certain fundamental assumptions. The first and most apparent is that testing is essential if you want a great website. There is a correlation between the most pressing usability concerns with online stores and the points of friction that customers encounter while making a purchase. A specific, measurable result (in terms of time, effort, accuracy, etc.) is what you can expect to see when you set and work toward a well - defined goal. This article presents an unproven notion and compares and contrasts usability testing with security testing in a nutshell. Even if it's ridiculous to compare these jobs, the outcome is unexpected. Usability and security testing may not be that dissimilar after all. If usability testing is further studied, new approaches to security testing could emerge.

2.2 Theoretical Framework

Currently, several corporations have built massive language models that have been trained using datasets and variables totaling in the billions. But we'll be seeing some of the best LLMs available right now:

1) **GPT - 3**– Generative Pre - trained Transformer 3, often known as GPT - 3, is the biggest and most well - known huge language model in the world. It was released in 2020. Microsoft has allowed OpenAI to alter and utilize its GPT - 3 code. Deep learning allows GPT - 3 to generate prompt - specific text output that is remarkably accurate in its resemblance to human writing. One well - known AI chatbot that uses the GPT - 3.5 model is ChatGPT. Another perk is the public API that allows you to connect with ChatGPT and get text results.
2) **BERT**– An artificial intelligence language model created by Google; Bidirectional Encoder Representations from Transformers (BERT) was introduced in 2018. It is the first natural language processing model to consider the left and right sides of a word simultaneously, setting it apart from its competitors. In order to get a better grasp of the prompt, BERT consults pre - trained plain text data sources, such as Wikipedia.
3) **LaMDA**– Google's Language Model for Dialogue Applications (LaMDA), released in 2022, is a big conversational language model. It is pre - trained on a 1.56 trillion - word text corpus that include both documents and conversations, and it employs a decoder - only transformer language model. Google's conversational AI chatbot Bard is powered by LaMDA, which also offers a Generative Language API for third - party app integration.
4) **PaLM**– Google AI's Pathways Language Model (PALM) is a vast, proprietary language model that was built in 2022. It has been pre - trained using high - quality datasets such as curated web pages, books, Wikipedia articles, news stories, open - source code from GitHub repositories, and talks from social media.
5) **LLaMA**– Vast Language Encoding In 2023, Facebook develops what is known as Meta AI (LLaMA). In the same way that other large language models create text endlessly, LLaMA takes a string of words as input and uses word prediction to do the same. Developers prioritized languages employing the Latin and Cyrillic alphabets while selecting texts to train the LLaMA model. The top 20 languages by speaker count were considered.

2.3 Research Gap

In our literature we are finding ways to improve e - commerce catalogs' discoverability and accuracy—which would make life easier for vendors and more enjoyable for consumers—has received very little academic attention. As far as the researcher is aware, a model has been created that uses a combination of picture recognition, meta - data analysis, and Large Language Models to automatically create product listings for online stores. Despite claims in the literature to the contrary, the exact processes by which social media may assist e - commerce platforms with LLMs are still not understood.

3. Methodology

3.1 Data Collection

Our data comes from actual Taobao transactions, and we're sharing it with the community so it can fuel their growth. A massive Chinese e - commerce platform, Taobao, is the source of the data. the datasets, LLMs, baseline information extraction techniques, and assessment criteria that make up our experimental setup. [7]

3.1.1 LLM Integration

This study compares and contrasts private LLMs hosted by OpenAI, such as GPT - 3.5 and GPT - 4, with open - source LLMs built on Llama2. All LLMs that are evaluated in this article are listed in Table 1. The OpenAI API, via which we access GPT - 3.5 and GPT - 4, levies use costs that are depending on both the model and the token count. You can run open - source models like Solar, Beluga, and Beluga - 7B on local GPUs; they are publicly accessible on the Hugging face hub.

Table 1: Detailed list of LLMs, with each model's name, parameter count, and API or GPU access for local execution

LLM	Exact Name	Model Size	API/ GPUs
GPT - 3.5	gpt - 3.5 - turbo - 0613	175B	API
GPT - 4	gpt - 4 - 0613	~1.8T	API
Beluga2	StableBeluga2	70B	4
Beluga - 7B	StableBeluga - 7B	7B	1
Solar	SOLAR - 0 - 70b - 16bit	70B	3

The outputs' degree of unpredictability is controlled by the temperature parameter of the LLMs. For the sake of reproducibility, we have set it to 0. The langchain Python module integrates with the OpenAI API and allows for the local execution of open - source models. Scientific investigations were carried out on a communal server that has 96 × 3.6 GHz CPU cores, 1024 GB of RAM, and 8 NVIDIA RTX A6000 graphics processing units.

3.2 Keyword and Description Generation:

The goal of this paper is to develop an automated system that can take text input and produce product descriptions. Using the product name as an input, we can simplify the issue. The goal of the system is to provide the product description y = (y1, y2,. . ., ym) ∈ Y, which is a sequence of words that may be used to describe the product, given the product title x = (x1, x2,. . ., xn) ∈ X. In order to have a detailed and tailored description, we add qualities and information and then provide a better definition. [8]

Definition 1.
What we mean by "knowledge" is a large database of unstructured text entries W that are indexed by named entities V. This database may include everything from Wikipedia to CN - DBpedia. The key for each entry is a named item v ∈V, and the value, which is a string of words w = (w1, w2,. . ., wu), is knowledge w ∈W.

We take hosted LLMs like GPT - 3.5 and GPT - 4 from OpenAI and open - source LLMs like StableBeluga and SOLAR into account, which may be executed locally. We test many LLMs on the product attribute value extraction job, including hosted ones like GPT - 3.5 and GPT - 4 and open - source ones based on Llama2.

3.3 Evaluation Methods

There are five different ways in which the model can predict an attribute's value given its value: NN (no predicted value), NV (incorrect predicted value), VN (no predicted value when the ground truth contains an attribute value), VC (correct predicted value that exactly matches the attribute value in the ground truth), and VW (incorrect predicted value that does not match the attribute value in the ground truth). P = VC/ (NV + VC + VW) is the formula for precision, a = VC/ (VN + VC + VW) is the formula for recall, and F1 - score is F1 = 2PR/ (P+R). We provide the F1 - score in our trials. Generation quality, variety, and attribute capture ability are the metrics we use to assess our model.

BLEU To make sure this isn't all a sham, we examine the BLEU score to make sure the qualities really improve performance. Using the same data, we compare the conditioned model's test BLEU score to that of the unconditioned baseline.

Lexical Diversity The tendency for the system to provide safe replies with insufficient variability is a prevalent issue in automated text production. In order to see whether our model can produce diverse texts, we also assess the lexical variety. Due to the varied nature of most human - generated writings, this is also an indication of their informativeness. If the variety score is low, then the produced material is probably boring and generic, but if it's high, then it's instructive and fascinating. One way to quantify the variety of descriptions created is by counting the number of unique n - grams that are formed on the test set. [9]

Attribute Capturing To evaluate the degree to which the produced results match the given characteristics, however, BLEU and lexical diversity alone are insufficient. The difficulty of recovering the input qualities from the model's descriptions is of special relevance to us. The accuracy of the prediction is used to calculate the attribute capturing score for attribute k:

$$\frac{1}{|\hat{Y}_T|} \sum_{(\hat{\boldsymbol{y}},\boldsymbol{y})\in(\hat{Y}_T, Y_T)} \mathbf{1}_{M(\hat{\boldsymbol{y}})=M(\boldsymbol{y})},$$

the user category classifier is denoted by M: Y → A2, and the indicator function is 1M (yˆ) =M (y), with a value of 1 when M (yˆ) =M (y) and 0 otherwise.

KOBE
An innovative model known as KOBE. Attribute Fusion and Knowledge Incorporation are the other two components that make up KOBE, in addition to the Transformer framework. Knowledge Incorporation is in charge of incorporating the pertinent information acquired from the knowledge base, while Attribute Fusion is in charge of combining the title representation with the attributes, which include product characteristics and associated user categories.

Baseline
Without taking qualities or knowledge into account, our baseline is a transformer that takes the product title as its source sequence and the description as its target sequence. The first step in training conditioned models is to train a specialized model on each data subset. Using the Attr - S (Source token) property, we may annotate the source text with characteristics, such as the product title. Using an attribute - specific start token for the target sequence, as opposed to a common token, is another approach to attribute fusion. This token is known as Attr - T (start of Target sequence). All we do at each timestep using Attr - A (Add) is add the attribute embeddings to the title embeddings. [10]

4. Results

4.1 Findings

Table 2 displays the models' BLEU scores. Attribute fusion improves BLEU score, as seen in Table 2. With an advantage of +0.7 BLEU (compared to 9.7%), Attr - A (Both) surpasses the baseline. Considering the qualities are categorical and lack information that is highly relevant to the product, this improvement in BLEU score is noteworthy.

Table 2: Evaluate BLEU score using various Attribute Fusion techniques

Model	Aspect	User Category	Both
Baseline	7.2	7.2	7.2
Attr - D	7.5	-	-
Attr - S	7.6	7.3	7.9
Attr - T	7.3	7.1	7.7
Attr - A	7.6	7.5	8.2

Table 3: Acquiring attributes score

Model	Aspect (%)	User Category (%)
Baseline	63.0	71.9
KOBE	74.6	86.0

Table 4: Evaluate the n - gram diversity score using various Attribte Fusion techniques

Model	n=3 ($\times 10^5$)	n=4 ($\times 10^5$)	n=5 ($\times 10^6$)
Baseline	2.4	7.8	2.0
Attr - S (Aspect)	2.8	9.7	2.5
Attr - T (Aspect)	2.8	9.6	2.5
Attr - A (Aspect)	2.8	9.6	2.5
Attr - S (User)	2.8	9.6	2.5
Attr - T (User)	2.6	9.1	2.4
Attr - A (User)	2.7	9.4	2.4
Attr - S (Both)	3.1	11.1	2.9
Attr - T (Both)	2.8	9.7	2.5
Attr - A (Both)	3.3	11.1	3.0

Table 5: Test n - gram diversity score was enhanced with the use of knowledge

Model	n=3 ($\times 10^5$)	n=4 ($\times 10^5$)	n=5 ($\times 10^6$)
Baseline	2.4	7.8	2.0
Know - BiAttn	3.1	12.1	3.3

All models, including those with knowledge and attribute constraints, have their diversity scores shown in Tables 4 and 5. Aspect and user category both considerably increase variety by 46.8 percent. The diversity score is raised by 56.4% when knowledge is included. The increased variety in our model's output shows that it generates more detailed descriptions compared to the baseline.

4.1.1 Through the Use of Training Data

(i). Extracting attribute values for the attribute representations, (ii) providing example demonstrations for in - context learning, and (iii) fine - tuning the GPT3.5 are all made possible by this training data.

Example Values

What follows is an examination of the effects of representing the target attributes using values taken from the training set as examples. There are two stages to this assessment. We begin by contrasting the compact attribute representation—which does not provide sample values—with the textual, compact, and json representations of the same attributes. Figure 2 shows the integration of the sample values into the attribute representations. The second step is to compare the optimal attribute representation to the zero - shot prompt list and then test out various quantities of sampled attribute values. Each attribute's short training set is often used to randomly choose up to ten distinct values. If an attribute in the training set has less than ten distinct values, all of those values are obtained.

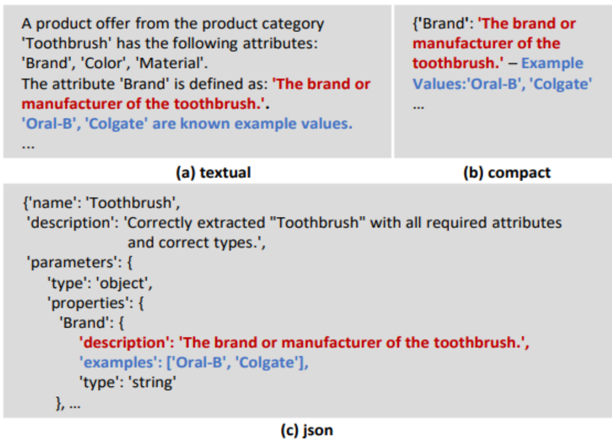

Figure 2: Formats for representing target attributes: (a) text, (b) compact, and (c) json

Table 6: For textual, compact, and json with 10 example values (10 - val), as well as for compact without example values, F1 - scores are available. Bold is the highest F1 - score for each LLM and dataset.

	Dataset	list	Textual 10 - val	Compact 10 - val	Json 10 - val
GPT - 3.5	OA - Mine	65.1	60.7	61.4	69.8
GPT - 4		68.8	64.8	66.4	75
Beluga - 7B		37.6	25.6	40.3	4.6
Beluga2		36.2	55.1	41.4	0
Solar		56.6	12.2	52	52.9
GPT - 3.5	AE - 110K	63.6	55	40.5	74.4
GPT - 4		53.7	53.5	46.5	70.1
Beluga - 7B		42.7	35.8	37.3	0.2
Beluga2		38.2	50.8	36.2	0
Solar		47.3	6	47.7	46.8

4.1.2 Fine - Tuning

When it comes to fine - tuning the GPT3.5 LLM, we assess its impact on performance using the training data. We do this by conducting a two - pronged analysis of how well a fine - tuned LLM performs. We begin by contrasting the adjusted GPT3.5 model's output with that of GPT4, which makes use of the same training data for its example values and demonstrations. Second, we look at the possibility that the model learns anything during tweaking that it might use for future prompts and datasets. Figure 3 shows the four fine - tuned GPT - 3.5 models that were produced by applying the two prompts list and json to the massive OA - Mine and AE - 110K training data. [11]

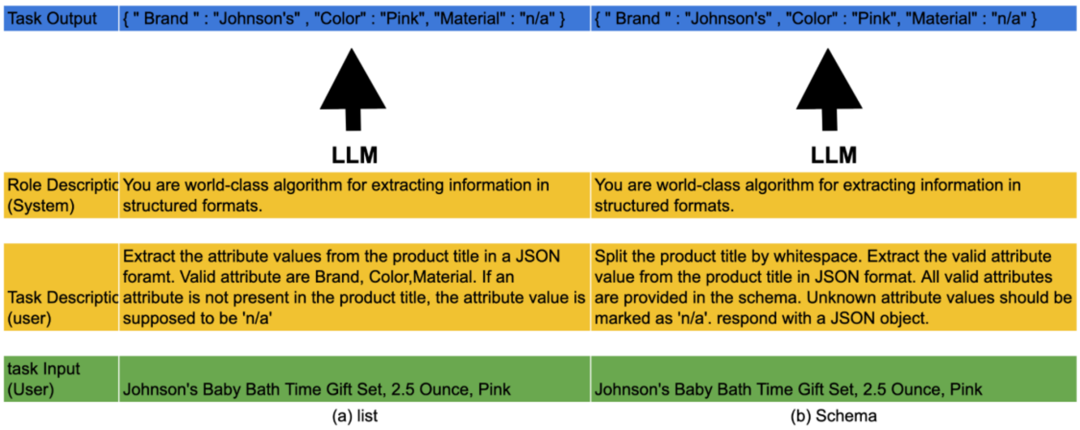

Figure 3: Design schema and list of zero - shot prompts

4.2 Data Interpretation

Table 7 shows that both GPT - 3.5 and GPT - 4 get F1 - scores greater than 60%. The compact form is optimal for GPT - 3.5, however the json representation with attribute descriptions is often more beneficial for GPT - 4. For GPT - 3.5 and GPT - 4, there is a little discrepancy in the attribute representations that do not include sample values. The json representation is particularly useful for GPT - 3.5 and GPT - 4. Without the advantage of very detailed attribute representations, the open - source models outperform OpenAI's LLMs on all challenges. The open - source models are the most effective with the list prompt. Compared to the best - performing baseline AVEQA that was fine - tuned on the huge training set, which achieves 79% F1 - score on OA - Mine and 81% F1 - score on AE - 110K, the zero - shot results of the LLMs are worse.

Table 7: List, textual, condensed, and JSON scores F1 - scores devoid of example values. Bolded is the highest F1 - score for each LLM and dataset.

	Dataset	list	textual	compact	json
GPT - 3.5	OA - Mine	63.3	62.7	65.1	64.8
GPT - 4		63.1	68.9	68.8	68.1
Beluga - 7B		36.9	39.6	37.6	6.8
Beluga2		58.3	50.8	36.2	0.0
SOLAR		60.8	55.2	56.6	11.4
GPT - 3.5	AE - 110K	61.4	61.3	63.6	61.5
GPT - 4		56.0	55.5	53.7	62.1
Beluga - 7B		46.4	47.4	42.7	0.3
Beluga2		52.5	52.1	38.2	0.0
SOLAR		52.1	49.2	47.3	11.1

Table 8 displays the outcomes of the KOBE ablation trial. To learn how various parts affect overall performance, this ablation research is useful. Knowledge Incorporation, User Category Fusion, and Aspect Fusion are the three parts. We dismantle KOBE component by component until all of them have been eliminated. Knowledge removal raises the BLEU score to 8.2 from 7.6 but lowers the diversity to 11.1 from 15.1. The model's BLEU performance may be negatively affected by irrelevant external information. Still, it has the potential to introduce a plethora of new things, expanding the variety.

Table 8: Ablation of three model components

Model	BLEU	Diversity (n=4) ($\times 10^6$)
KOBE	7.6	15.1
- knowledge	8.2	11.1
- user category	7.6	9.6
- aspect	7.2	7.8

A portion of the test set is selected at random for human review. The instance's three components are the input product title, the baseline description, and the KOBE description. Human annotators who are not familiar with the description systems are given the examples that have been chosen for annotation. The fluency score indicates the level of fluency in the description. The description's diversity and the amount of competitive material are reflected in its diversity score. The reasonableness of the description, or its congruence with what is known about the world, is reflected in the total quality score. A score of 1–5 is possible. The greater the score, the closer the description is to the requirement. Table 9 displays the outcomes. In human evaluations, KOBE achieves a diversity score that is 0.06 higher than the baseline, which is in accordance with the findings of the automated assessment. Nevertheless, fluency is not sacrificed in the enhancement of variety. On the contrary, as compared to the baseline, KOBE's fluency performance is +0.17 times better. When compared to the baseline, KOBE maintains a significant quality advantage of +0.13, indicating that our model can provide more plausible descriptions after including external information. [12]

Table 9: The advantage of the proposed KOBE model over the baseline, as determined by pairwise human assessments

Model	Fluency	Diversity	Overall Quality
Baseline	3.78	3.73	3.67
KOBE	3.95	3.79	3.80

Paper ID: SR24304113521 DOI: https://dx.doi.org/10.21275/SR24304113521

4.3 Optimizing E - Commerce Listings LLM - Based Description and Keyword Generation from Multimodal Data

Optimizing E - Commerce listings using Large Language Models (LLMs) for description and keyword generation from multimodal data involves several critical steps, each contributing to the overall effectiveness and attractiveness of product listings.

1) Data Collection
The first step involves collecting various forms of data, including images, characteristics, and technical details of the product. Image recognition techniques can be employed to extract vital visual data from product images, which forms an integral part of the multimodal data approach.

2) Data Preprocessing
Next, the collected data needs preprocessing. Textual data is processed and normalized to ensure consistency and accuracy. Similarly, image data is normalized and converted into a format compatible with LLM input, preparing it for the subsequent stages of description generation and keyword extraction.

3) LLM - Based Description Generation
In this phase, LLMs like GPT - 3 or BERT play a crucial role. These models are capable of generating comprehensive and engaging product descriptions by integrating both textual and visual data. This results in thorough product narratives that accurately reflect the brand's tone and style, tailored to the intended audience. The key is to engage with LLMs during the ideation phase for generating innovative ideas and refining concepts. LLMs can also be used as effective editors in the revision process to correct errors and improve readability [15].

4) Keyword Extraction
Utilizing Natural Language Processing (NLP) techniques, relevant keywords are extracted from the generated descriptions. This includes product specs, unique attributes, and other pertinent phrases. These keywords are then optimized for search engine visibility, prioritizing terms with high search volumes.

5) Multimodal Integration
The generated textual content is integrated with the original data and keywords. It's crucial to ensure that the written descriptions are properly matched with their corresponding product photos to maintain consistency.

6) SEO Optimization
Incorporating the extracted keywords into product names, descriptions, and meta tags enhances the content's search engine ranking, boosting organic traffic. The use of LLMs in crafting unique and compelling product descriptions and optimizing title tags and meta descriptions can significantly improve search engine visibility and click - through rates [16].

Benefits
- **Improved Search Visibility**: Enhanced product descriptions and targeted keywords increase the likelihood of products appearing in relevant search results.
- **Enhanced User Engagement:** LLM - generated descriptions provide detailed information, helping potential buyers make informed decisions.
- **Consistent Branding:** Uniform brand tone across product listings strengthens brand recognition.
- **Higher Conversion Rates:** Engaging descriptions and relevant keywords increase the chances of converting visitors into customers.

Challenges and Considerations
However, working with LLMs also brings certain challenges and considerations:
- LLMs are most effective with familiar themes and styles, and their results can vary for more specialized requirements.
- LLMs are not experts in specific fields, so it's essential to bring in domain expertise and carefully word the prompts [17].
- Bias and ethical considerations are important when implementing LLMs in e - commerce. These models can inadvertently include societal biases in their outputs, so active monitoring and evaluation are necessary [17].

In summary, the use of LLMs in optimizing e - commerce listings is a multifaceted process that involves careful data collection and preprocessing, creative and technical use of LLMs for content generation, meticulous keyword extraction, and strategic SEO optimization. While offering significant benefits in terms of search visibility, user engagement, branding consistency, and conversion rates, it also requires mindful handling of challenges related to LLM functionality and ethical considerations.

5. Discussion

5.1 Implications

Our innovative approach KOBE makes use of product qualities and external information to enhance the quality of generation, with an emphasis on individualized and comprehensive product descriptions, in order to achieve results in recommendation. Our technique is able to provide more accurate and detailed product descriptions than the baseline model, as shown by the assessment and experimental studies. The results of the ablation investigation highlight the need of fine - tuning the parameters. It highlights how subtle changes to hyperparameters may significantly impact performance, even in complex models driven by massive volumes of data.

5.2 Limitations

Phrasal expression and discourse structure were both severely constrained in this research, making it unable to produce descriptive writing that is both unique and useful to the reader. On top of that, they disregard user data and external knowledge completely. One of the world's foremost e - commerce platforms provided the data used to build the report. This is a solid starting point, but it would be helpful to see how well our concept works on other e - commerce platforms to get a more complete picture. Our method is naturally extensible, but there is always room for improvement when it comes to computing performance, particularly in real - time applications.

5.3 Future Research

We have created and annotated a large - volume dataset called TaoDescribe for product description creation. This dataset might serve as a standard for future study in this area. Our hope is that future research will build upon this study's strengths and address the study's limitations, which we do not see as flaws but as openings for further discovery and improvement.

6. Conclusion

6.1 Summary

A big language model is an algorithm that uses a huge dataset for training and learning. Incorporating the patterns and information gleaned from such training, it is designed to comprehend and produce text that resembles human writing. Online marketplaces with a wide variety of items are becoming more common, and e - commerce websites are a key component of this trend. Various things in a catalog have different brands, models, and primary characteristics, yet all of them are distinct from one another (clothing, gadgets, books). MLLMs show signs of being able to generate visual narratives and answer questions based on images, which might pave the way for AI and real - world human - computer interactions. We go into a fascinating issue with the automated generation of customized product descriptions inside an e - commerce platform. This research investigates the possibility of using LLMs instead of PLMs for attribute/value extraction, since they are more resilient and need less training data.

6.2 Conclusions

We can say that our classification and analysis of current modality alignment approaches draws attention to their unique characteristics and future research objectives in the realm of MLLMs. In comparison to state - of - the - art PLM - based models, our findings further demonstrate that LLMs need less domain specific data for training. Using the identical training data, GPT4 with in - context learning surpasses the best fine tuned PLM - based baseline by a large margin. Optimized GPT - 3.5 may cut extraction costs in half compared to GPT - 4 without sacrificing speed, making it ideal for use cases that need a high number of product attribute data. Taobao, the biggest online e - commerce site in the world, has the recommended framework that we have successfully applied. KOBE outperforms state - of - the - art by 9.7 percentage points in terms of BLEU. To back up our claims that the suggested method works, we provide a number of case examples. Taobao now has the structure put into action. The existing literature on the topic of massive language models and their practical use in e - commerce is expanded upon by this study.

References

[1] S. Srivastava, S. Kshatriya, and R. Rathore, "Search Engine Optimization in E - Commerce Sites, " 2017.

[2] H. Wu, "Search Engine Optimization of E - Commerce Websites, " in Proc.2011 IEEE International Conference on Management and Service Science (ICMSS), 2011, doi: 10.1109/ICMSS.2011.5999008.

[3] K. Shi, X. Sun, D. Wang, Y. Fu, G. Xu, and Q. Li, "LLaMA - E: Empowering E - commerce Authoring with Multi - Aspect Instruction Following, " 2023.

[4] V. Soni, "Large Language Models for Enhancing Customer Lifecycle Management, " 2023.

[5] J. Chen, Z. Zhong, Q. Feng, L. Liu, "The Multimodal Emotion Information Analysis of E - Commerce Online Pricing in Electronic Word of Mouth, " Journal of Global Information Management, vol.30, pp.1 - 17, 2022, doi: 10.4018/JGIM.315322.

[6] A. Sinha, B. Kumar, S. Roy, H. Mahmood, N. Garg, and M. Hashmi, "ANALYSIS OF E - COMMERCE USABILITY AND SECURITY FACTOR BASED ON MULTIMODAL BUSINESS DIMENSION IN PHARMACEUTICAL SUPPLY CHAIN AND CBIR BASED TEXTILE RECOMMENDATION SYSTEM, " 2022, doi: 10.47750/pnr.2022.13. S07.336.

[7] C. R. Kothari, "Research Methodology: Methods and Techniques, " 2004.

[8] M. Booth, S. P. Reinhardt, and A. Roy, "Partitioning optimization problems for hybrid classical/quantum execution, " 2017. [Online]. Available: https: //www.dwavesys. com/resources/publications.

[9] D - Wave Systems Inc., "D - wave solver properties and parameters reference, " 2017. [Online]. Available: https: //www.dwavesys. com/resources/publications.

[10] M. Henderson, J. Novak, and T. Cook, "Leveraging adiabatic quantum computation for election forecasting, " arXiv reprint arXiv: 1802.00069, 2018.

[11] M. Hernandez and M. Aramon, "Enhancing quantum annealing performance for the molecular similarity problem, " Quantum Information Processing, vol.16, 2017.

[12] Z. Yang et al., "Mmreact: Prompting chatgpt for multimodal reasoning and action, " arXiv preprint arXiv: 2303.11381, 2023.

[13] M. Fokina, "Online Shopping Statistics: Ecommerce Trends for 2023, " 2023. [Online]. Available: https: //www.tidio. com/blog/online - shopping - statistics/. Accessed Jun.21, 2023.

[14] J. Howarth, "27 New Online Shopping Statistics for 2023, " 2023. [Online]. Available: https: //explodingtopics. com/blog/online - shopping - stats. Accessed Jun.21, 2023.

[15] Cohere, "How to Build a Product Description Generator with LLMS, " 2022. [Online]. Available: https: //txt. cohere. com/how - to - build - a - product - description - generator - with - llms/. Accessed Oct.10, 2023.

[16] Marin Software, "How to Optimize Your E - Commerce Product Descriptions for SEO, " 2023. [Online]. Available: https: //www.marinsoftware. com/blog/how - to - optimize - your - e - commerce - product - descriptions - for - seo. Accessed Oct.10, 2023.

[17] Analytics Vidhya, "Customized Marketing Copywriting Using LLMS for E - Commerce, " 2023. [Online]. Available: https: //www.analyticsvidhya. com/blog/2023/09/customized - marketing - copywriting - using - llms - for - e - commerce/. Accessed Oct.10, 2023.