

Design and Development of Large Language Model Applied to Fashion Analysis

Yiming Fan¹ and Yinan Wang²

¹University of Southampton, Southampton, U.K.

²The University of Hong Kong, HK, China

ABSTRACT

Consumer preferences and intense market competition are characterized by dynamic due to the rapidly evolving fashion industry. However, existing trend analysis tools do not fully exploit large datasets, resulting in a lack of precision and timeliness in trend prediction, and often fail to grasp the complex nuances specific to the fashion sector. To address this research gap, this study aims to develop a cutting-edge, AI-driven model for fashion trend analysis and prediction, leveraging the power of large language models. A mixed method of qualitative and quantitative methods was conducted in this study. To be specific, the latest fashion trend data from WGSN including terncolor trends, fabrics and materials, and market analysis were used to collect data. These data were undergone thorough cleaning and pre-processing. After that, the LLAMA2-13B model was selected for pre-training, which was then fine-tuned using LoRA technology. Finally, the performance evaluation of the model was rigorously evaluated using 5-point Likert scale. In addition, a manual evaluation was conducted, focusing on the specific analytical requirements of the fashion domain to ensure the model's validity and applicability. This study not only enhances the accuracy and efficiency of trend analysis and prediction but also offers valuable market insights and decision-making support for fashion designers, brand managers, and retailers. This advancement marks a substantial step forward in data analysis and prediction methodologies within the fashion industry, paving the way for new perspectives and approaches in future trend analysis and market strategy formulation.

Keywords: Large language model, Fashion trend analysis, Consumer preferences

INTRODUCTION

Fashion trend forecasting, which refers to the profession of envisioning future trends in style and foreseeing consumers desires, is crucial to fashion companies' business success (Furukawa, Miura, Mori, Uchida, & Hasegawa, 2019; Rousso, 2012). With the development of big data and AI technology, Natural language processing (NLP) is widely used in fashion fields, such as fashion trend prediction. NLP can help fashion companies better understand customer needs and optimize business processes. It can also be used as a fashion analyst to provide users with professional fashion analysis services. For example, Xiu Li used NLP to ascertain changes in the fast-fashion industry based

on UN sustainable development goals (Chadha et al., 2023). In addition, Patrick John Chia built on recent developments in contrastive learning to train Fashion CLIP (Chia et al., 2022). Recently, because of the development of large language models (LLMs), especially the recent emergence of Llama, Llama 2, has caused widespread concern in the field of natural language processing (NLP).

Specifically, LLaMA is a series of models developed by Meta which can construct a lot of characters that are not in vocab through byte, and LLaMA has gained better multilingual ability. In order to improve the stability of training, the input of each LLaMA's transformer layer is normalized instead of the output. The LLaMA tokenizer uses the bytewise encoding algorithm (Sennrich et al., 2016) with the implementation from SentencePiece (Kudo & Richardson, 2018). Llama vocabulary has a size of 32k tokens (Touvron et al., 2023). LLaMA uses SwiGLU instead of RELU as the activation function. The advantage is that the activation function can get the optimal value of log-perplexity (Shazeer, 2020). LLaMA incorporates RoPE, which stands for "Rotary Positional Encoding." This approach merges absolute and relative position encoding, preserving simplicity while enhancing understanding of token relationships and sequence context. LLaMA2 is the iterative upgrade large language model of LLaMA. Its parameter scale increases from 7B to 70B. The model architecture and pretraining process involve the use of the RMSNorm pre-normalization (Xu & Zhang, 2014). In addition, the rotary positional embeddings is also involved (Su et al., 2024). In terms of fashion industry, the colour is essential to fashion trend forecasting since it is a building block for creating trend-right, and top-selling designs (Błaszczuk & Wubs, 2018). However, colour analysis involves a wide range of analytical dimensions, scarce training data, and high timeliness, which results in that academia has achieved little in colour analysis. Benefiting from the powerful analytical capabilities of large language models, AI-assisted colour analysis has also ushered in a new dawn.

To address the research gaps mentioned above, this study aims to design and develop of large language model applied to fashion analysis using LLaMA. This development not only improves the precision and effectiveness of trend analysis and forecasting, but also provides important market insights and decision-making assistance for fashion designers, brand managers, and retailers. This progress represents a significant advancement in data analysis and prediction techniques in the fashion industry, especially in terms of colour, opening up new possibilities for future trend analysis and market strategy development. Therefore, the objectives of this study are: firstly, fine-tune a new Llama 2-based large language model specifically for analysis and prediction in the fashion field; In addition, we have constructed a proprietary dataset based on the professional fashion content of WGSN, trained the model, and ensured that it has accurate and specific knowledge and analytical capabilities in the fashion field; Finally, we demonstrate the potential of this model in fashion trend analysis and prediction, providing new perspectives and decision support tools for the fashion industry.

METHODOLOGY

Base Model

In the project, we chose LLaMA2-13B as the pre-trained model. Llama 2 is a series of pre-trained and fine-tuned text generation models developed and publicly released by Meta AI, with parameter scales ranging from 7 billion to 70 billion. These models have performed well in multiple benchmark tests. However, Llama 2 and its variants are mainly oriented to general scenarios, which limits their applicability in handling professional scenarios such as fashion analysis. In the fashion field, the application of Llama 2 faces specific challenges. Fashion data is difficult to quantify, updates quickly. In addition, fashion trends also have a high degree of randomness, and the factors involved in fashion are also extremely extensive. Since Llama 2 is mainly pre-trained based on publicly available online data sources, these data may not contain sufficient fashion domain expertise and details. Therefore, Llama 2 has problems of lack of knowledge and insufficient analytical capabilities when directly applied to the fashion field, especially when dealing with highly professional fashion content. Although Llama 2 performs well in dialogue and other NLP tasks, its ability to understand and analyze fashion trends, styles and designs may be limited.

In order to overcome these limitations and improve the application effect of Llama 2 in the fashion field, we have constructed a dataset based on WGSN, consisting of structured text and unstructured text (Raw text). The former is used to train the model's multidimensional fashion analysis capabilities, while the latter supplements the model's expertise. During fine-tuning, we use LoRA (Low-Rank Adaptation) to fine-tune Llama 2. The LoRA approach was inspired by the structure-aware intrinsic dimension method (Aghajanyan et al., 2021). LoRA is an efficient fine-tuning method that reduces the number of parameters that need to be modified by introducing lower-dimensional, task-specific representations in the weight matrices of the pre-trained model.

On the dataset, we used WGSN fashion content to complete dataset training. WGSN is a globally renowned fashion trend forecasting company that provides in-depth market analysis and design strategies widely used in the fashion and consumer goods industries. By combining Llama 2's powerful language processing capabilities and WGSN's rich fashion data, our model is able to perform in-depth fashion analysis and trend prediction to support decision-making in the fashion industry.

Dataset Construction

Structured Dataset

The difficulty in building large language models for fashion analysis such as color analysis is that the available data in this field is relatively small, and is mostly embedded in texts, highly fragmented, and difficult to construct as structured datasets. The various elements involved in fashion analysis have extensive weak connections between each other. If the structure of the dataset is not set reasonably, it will cause rather serious loss of information.

In addition, fashion fields often rely on rich backgrounds, which means that a lot of information is difficult to extract in a structured way. Therefore, we cannot rely solely on structured data when training large language models. Finally, the industry is known to be facing highly variable demand and fashion trend, and the use of information is an important issue (Chen et al., 2010). The data between different years tend to have a certain independence, which means that we cannot easily use methods based on continuous time sequences to construct datasets, but need to emphasize the relationship between attributes.

This dataset is constructed based on WGSN (Worth Global Style Network) reports, aimed at analyzing future clothing color trends. Color analysis has the characteristics of involving multiple attributes and weak connections between attributes. We want the large language model to obtain the ability to capture multi-dimensional information and analyze the correlation between different dimensions. Therefore, we took this into consideration when constructing the dataset. In the construction process, the dataset tries to capture from the text materials as many attributes related to color trends as possible, so that the large language model has the ability to analyze fashion information from multiple angles, while enabling the large language model to obtain the ability to discover the correlations between these attributes. Specifically, the dataset extracts factors such as regions, seasons, materials, etc. that predict popular colors, providing basic data for fine-tuning large language models.

The dataset contains the following fields:

Color Name: Indicates the name of the popular color, such as “apricot yellow”, “force blue” and so on.

Pantone Code: Corresponds to the Pantone color card encoding for each color, such as “15-1247 TCX”, “18-3922 TCX” and so on.

Region: Indicates the geographical area where the color is predicted to be popular, such as “North America”.

Season: Specifies the season when the color is predicted to be popular, such as “Spring/Summer 2024”.

Application: Describes possible application scenarios or uses for the color.

Evolution: Describes the changes or evolution of the color over time.

Usage: Provides suggestions on how to use or match the color.

Description: Detailed description of the color and its background or meaning.

Data Samples

The dataset includes a variety of colors, each with its unique Pantone encoding and predicted popular regions and seasons. For example, the first record in the dataset is the color “apricot yellow”, with a Pantone code of “15-1247 TCX”, predicted to be popular in North America in spring/summer 2024. We eventually extracted 508 data records from all the WGSN 2022–2024 reports released to the public for model training.

Dataset Features

Multi-dimensional attributes: The dataset contains not only colour names and Pantone codes, but also covers regions, seasons, applications, evolutions,

usage, and descriptions, providing multi-dimensional information. This allows the large language model to obtain the impact information of multiple attributes on colour changes, and to analyze background information from more angles, providing a more profound and comprehensive analysis for colour analysis.

Cross-industry application capabilities: The analysis provided in the dataset based on WGSN involves multiple fields such as fashion, consumer goods, interior decoration, and technology products, in addition to fashion trends, giving the large language model the ability to conduct cross-domain fashion analysis.

In constructing the dataset, the connections between attributes are considered. Specifically: The relationship between colour and region, season: Different regions and seasons have different colour preferences, and these attributes in the dataset help analyze popular trends in specific regions or seasons. The evolution and application of colour: The evolution of colour reflects the change of fashion trends, while the application points out the specific use of colour in different products and scenes. Description and usage: The description of colour provides an in-depth understanding of its characteristics and imagery, which can provide inspiration for design and marketing when combined with usage suggestions.

Raw Text Extraction

The available versions of WGSN are all in PDF format, but because the text in PDF files is embedded in a special way, conventional text extraction methods cannot correctly parse WGSN. Using PdfFileReader to process WGSN directly will make the output text completely garbled. Here we choose to use PyMuPDF combined with OCR for text extraction. PyMuPDF supports PDF text extraction more extensively. At the same time, for texts located at edges, corners or irregular positions, we use OCR to supplement the text extracted by PyMuPDF, eventually obtaining relatively complete text.

Raw Text Cleaning

The text extracted directly from WGSN contains a lot of redundant information, which is difficult to apply directly to the training of large language models. Especially in the task of training large language models for fashion analysis, the amount of data is relatively small, so the number of training rounds and weights are relatively high, and the impact of redundant information is greater. Through textual analysis of the extracted text, the redundant information in the text mainly includes: 1. Headers, footers 2. Social media usernames or tags 3. Color reference information, warnings, and duplicate characters 4. Personal names or organization names. We have established corresponding regular expressions to clean the text for these four types of problems, eventually obtaining text that can be used for training. The number of characters finally generated in this part of the text is 204097.

Specifically, for type 2, use the regular expression `@[A-Za-z0-9_]+\b` to match social media identifiers. For type 3, we statistically established regular expressions to match color reference information, warnings and duplicate characters, and constructed the `remove_color_references_and_warnings`

function. Using list comprehension, call the `remove_color_references_and_warnings` function on each part of the text to delete the content related to color references and warnings in each text. For type 4, we used a processing method similar to type 3. The difference is that we constructed regular expressions to match names and organization names according to the case and general naming principles of letters for personal names and organization names.

LoRA Training

The LoRA uses the following parameters for training structured data: LoRA Rank:96, LoRA Alpha:64, Batch Size:128, Cutoff Length:2048, Epochs:4, Learning Rate:6e-4.

Since our structured data has the characteristics of small quantity, high accuracy and quality, we appropriately increased the Epoch and Learning Rate to 6e-4 during training, and adopted a relatively high LoRA Rank of 96.

The LoRA uses the following parameters for training unstructured data: LoRA Rank:64, LoRA Alpha:64, Batch Size:128, Cutoff Length:2048, Epochs:3, Learning Rate:4e-4.

Unstructured data has the characteristics of larger quantity, lower accuracy and quality than structured datasets. During training, we appropriately reduced the Epochs and Learning Rate to 4e-4, and adopted a relatively low LoRA Rank of 64. Considering the vast data amount discrepancy between the fine-tuning dataset and the data used to train LLaMA2, and the task of enhancing fashion analysis capabilities, we did not excessively reduce the training weight of unstructured data. In training unstructured data, using a Learning Rate over 8e-4 will cause the model to show incoherent statements and confused logic in generation. After multiple generation experiments, we determined the above training parameters.

We compared our large language model with LLaMA2-13B and Llava-13B, and asked each model color matching questions in ten standard scenarios respectively. Table 1 shows Comparison of the results of each model. Due to space constraints, we only show the first color system given in the answers of each model in Table 1.

Survey

Seven questionnaires were disseminated to evaluate the influence of the platform on consumers and their purchasing motivation. The participants included 3 males and 4 females, aged 18-35, all possessing at least a bachelor's degree. Their occupations varied and included teachers, programmers, students, AI computer engineers, and designers. The 5-point Likert scale was used in this questionnaire.

RESULT

The outcomes produced by each model exhibit noticeable disparities. Our model, Fanny, which has been trained on fashion data, presents relatively well-organized responses in terms of content and format. Additionally, the

responses from our model accurately corresponded to Pantone color codes, such as Modern Mocha (Mocha brown-19-1220 TCX, emerald green-17-5641 TCX, sapphire blue-19-4052 TCX, ruby-18-1664 TCX), while the alternative models provided more ambiguous color descriptions, such as neutral colors. In professional fashion analysis applications, such responses are evidently insufficiently precise. Furthermore, in terms of the depth of logical reasoning and the richness of color analysis, Fanny's responses encompass a more comprehensive understanding of color psychology and matching strategies. For instance, in the aforementioned table, Fanny's response highlights the color psychology that darker tones convey a mature, elegant, and mysterious temperament, suitable for expressing individuality and confidence. Fanny's capacity to comprehend scenarios also surpasses that of comparative models, offering more accurate color recommendations and more relevant color psychology and scenarios. In terms of matching the fashion trend, Fanny's color recommendations are more in line with the current fashion trends, having been trained on the latest WGSN dataset. The latest training data addressed the issue of outdatedness caused by outdated training data.

Table 1. Comparison of the results of each model(Intercept the first color system given in the answer).

Questions\Model	Fanny	LLaMA2-13B	Llava-13B
I'm going to attend my boyfriend's graduation ceremony in the summer. Can you recommend any colors on my clothes?	Power pink (Cyclamen powder-17-1937 TCX, light pink-13-2804 TCX, light blue-16-4120 TCX, peppermint green-12-5406 TCX): Color psychology: the tone of this combination gives people a feeling of tenderness, freshness and calmness, reflecting the lightness and vitality of spring and summer. Collocation strategy: light pink and peppermint green can be used as the tone of the overall shape, through light blue embellishment to increase the sense of hierarchy. This combination is suitable for soft fabrics such as chiffon and cotton to create a lightweight appearance. Light pink exudes tenderness and feminine charm, which is a popular color in spring and summer, giving people a fresh and elegant feeling. Light blue brings a feeling of serenity and serenity, giving people a refreshing and comfortable visual enjoyment. Peppermint green is a vibrant color, it is both fresh and fashionable, with a trace of vitality and natural flavor. Such a color combination is suitable for creating a woman who is both gentle and fashionable.	Here are some options for summer graduation: 1. Light blue: when graduating in summer, a light blue dress or suit is a refreshing and versatile choice. This is a classic color that can be worn or worn, and it will complement your boyfriend's bright and celebratory mood.	Congratulations on your boyfriend's graduation! Here are some color palette suggestions for graduation: 1. Classic elegance: * navy blue and white. * black and white. * charcoal gray and white. * dark navy blue and light blue. * Green and rustic tones of the forest. * bright pink (for example, hot pink, bubblegum pink or coral pink).

The findings from the questionnaire survey indicated that the platform recommendations for colours, styles and materials are approved. To be specific, the platform received the highest level of satisfaction for color recommendations (mean: 4.4, median: 4), surpassing other assessment categories (shown in Figure 1). However, it was found to be less satisfactory for style (mean: 3.7, median: 4) and material (mean: 4, median: 4) compared to color recommendations (shown in Figures 2 & 3).



Figure 1: Results of satisfaction with colour recommendations document.

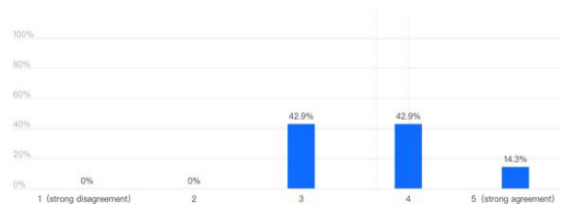


Figure 2: Results of satisfaction with style recommendations document.

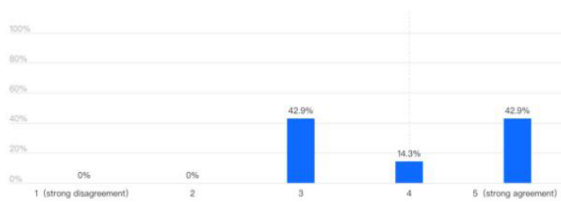


Figure 3: Results of satisfaction with material recommendations document.

Furthermore, the survey demonstrates that platforms have a positive influence on consumers' purchasing intent across all three indicators (shown in Figure 4), as evidenced by their perception of platforms as reliable recommenders and their willingness to consult platforms prior to making purchases (shown in Figures 5 & 6).

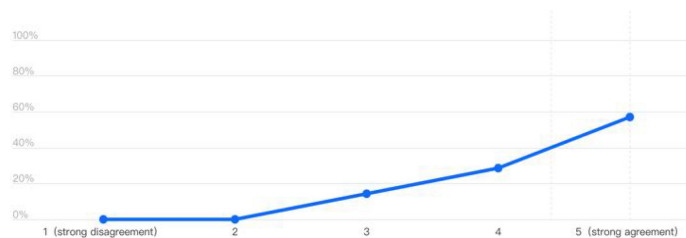


Figure 4: Results of platform influence on willingness to buy.

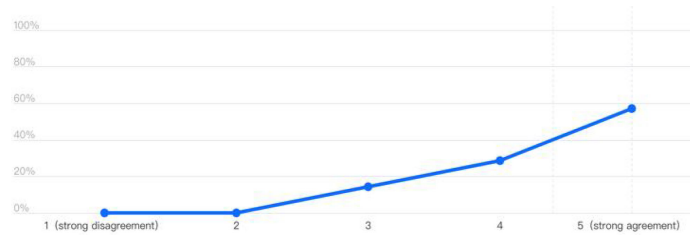


Figure 5: Results of the platform recognition as a qualified fashion recommender.

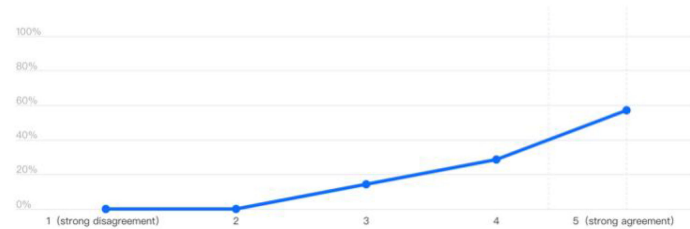


Figure 6: Results of willingness to consult the platform before buying.

CONCLUSION

This paper has demonstrated the innovative application of the Llama 2 Large Language Model in the domain of fashion analysis. Our model, Fanny, fine-tuned using a specialized dataset from WGSN, has shown remarkable capabilities in understanding fashion trends, particularly in color analysis. By integrating structured and unstructured text from WGSN, Fanny has achieved a higher degree of precision in predicting color trends, reflecting the dynamic and complex nature of fashion.

Our research highlights the potential of LLMs, like Llama 2, in specialized fields beyond general language processing. The adaptability of Fanny to the unique demands of fashion analysis, characterized by rapid trend changes and nuanced interpretations, signifies a leap forward in applying AI in niche sectors.

Future research could expand Fanny's capabilities to encompass wider aspects of fashion, such as texture, material, and style analysis. Integrating visual data processing could further enhance its predictive power, offering a more holistic approach to fashion trend forecasting. Additionally, exploring the application of Llama 2 models with greater parameter scales could unveil more profound insights into complex fashion datasets.

In conclusion, our findings open new avenues for AI-assisted fashion analysis, offering a more nuanced and data-driven approach to understanding an ever-evolving industry. Fanny stands as a testament to the potential of large language models in transforming sector-specific research and analysis.

REFERENCES

- Aghajanyan, A., Gupta, S. and Zettlemoyer, L. (2021) 'Intrinsic dimensionality explains the effectiveness of language model fine-tuning', Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) [Preprint]. doi: 10.18653/v1/2021.acl-long.568.

- Chen, H. et al. (2010) ‘Coordination mechanism for the supply chain with leadtime consideration and price-dependent demand’, *European Journal of Operational Research*, 203(1), pp. 70–80. doi: 10.1016/j.ejor.2009.07.002.
- Chen, M. X. et al. (2018) ‘The best of both worlds: Combining recent advances in neural machine translation’, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* [Preprint]. doi: 10.18653/v1/p18-1008.
- Chia, P. J. et al. (2022) ‘Contrastive Language and vision learning of General Fashion Concepts’, *Scientific Reports*, 12(1). doi: 10.1038/s41598-022-23052-9.
- Furukawa, T. et al. (2019) ‘Visualisation for analysing evolutionary dynamics of fashion trends’, *International Journal of Fashion Design, Technology and Education*, 12(2), pp. 247–259. doi: 10.1080/17543266.2019.1587789.
- Kudo, T. and Richardson, J. (2018) ‘Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing’, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* [Preprint]. doi: 10.18653/v1/d18-2012.
- Sennrich, R., Haddow, B. and Birch, A. (2016) ‘Neural machine translation of rare words with Subword units’, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* [Preprint]. doi: 10.18653/v1/p16-1162.
- Shazeer, N. (2020) Glu variants improve transformer, arXiv.org. Available at: <https://arxiv.org/abs/2002.05202> (Accessed: 1 February 2024).
- Su, J. et al. (2024) ‘Roformer: Enhanced Transformer with rotary position embedding’, *Neurocomputing*, 568, p. 127063. doi: 10.1016/j.neucom.2023.127063.
- Touvron, H. et al. (2023) Llama 2: Open Foundation and fine-tuned chat models, arXiv.org. Available at: <https://arxiv.org/abs/2307.09288> (Accessed: 1 February 2024).
- Xu, J. and Zhang, H. (2014) ‘Mean-square stabilization for stochastic systems with multiple delays’, *Proceedings of the 33rd Chinese Control Conference* [Preprint]. doi: 10.1109/chicc.2014.6895865. T. M. Choi and S. Sethi, “Innovative quick response programs: A review,” *Int. J. Prod. Econ.*, vol. 127, no. 1, pp. 1–12, 2010. U.