Chicbot: A Hybrid Machine Learning and Deep Learning Fashion stylist

A chatbot that helps the user with fashion using Machine Learning Algorithms

Aditi Ajay Marar
Amrita School of engineering
Bangalore, India
bl.en.u4aie22103

G Krisha
Amrita School of engineering
Bangalore, India
bl.en.u4aie22114

Sanjana R
Amrita School of engineering
Bangalore, India
bl.en.u4aie22149

***Abstract*—Fashion is a highly dynamic aspect of human life. With the growth of Machine Learning, styling ourselves to look crisp is easier. This project aims to develop a Machine Learning-based fashion stylist that can assist the user in putting together outfits that suits the user's style and also stay in the current trend. The system will utilize various Machine Learning algorithms to analyze data like size information, preferred clothing types, and optionally, images of the user's style and suggest suitable outfit combinations. The system will ensure that its user stays fashionable and confident in their attire.**

***Index Terms*—fashion, ML stylist, colour trend**

I. Introduction

In this rapidly evolving world of fashion and technology, integrating the techniques of each landscape can revolutionize the way consumers engage with style. This paper explores the hybrid applications of Machine Learning(ML) and Deep Learning(DL) within the framework of a chatbot-based fashion advisor named Chicbot, which offers personalized outfit recommendations based on the users preferences.

The chatbot technology enables the user to easily interact with the system conversationally. Natural Language Processing(NLP) techniques facilitate meaningful exchanges between the user and the chatbot, this allows the system to understand user preferences, style aspirations, and contextual factors influencing outfit selection. The Machine Learning(ML) model of the system is trained on datasets of fashion images and, color palettes, which allows it to assist the user to be more open and discover new styles.

II. Literature Survey

Han A,Kim J et.al. [1] shows an approach toward color trend analysis in fashion through the application of machine learning. It uses various technologies like computer vision and data analysis. This idea[1] automates the identification of color trends from a huge dataset of fashion images. This approach [1] enables a data-driven understanding of customer preferences and prediction of future trends, which enhances the decision-making processes for designers. The results highlight how machine learning has the potential to transform trend analysis which helps us in learning and adapting features that align with the changing market trends.

Aditya Jain et al[2] talk about how Machine learning(ML) plays a crucial role in Artificial Intelligence(AI). It highlights how Machine Learning has revolutionized various aspects of technology and daily life, particularly in search engines, software products, social media platforms, etc. It talks about different applications of ML in AI, like how ml algorithms are used in NLP.

Zhao et al [3] present an innovative method for fashion trend method for fashion prediction by using data-driven techniques and catwalk analysis. The system uses machine learning and data analytics to examine trends shown in runaway events which offers an understanding of emerging styles and preferences.This system analyzes visual cues and patterns from catwalk presentations which enables the designers to fulfill the customer requirements with a better accuracy. This idea[3] gives a better and effective analysis strategy in predicting fashion trends.

X Gu et al [4] introduce a framework for fashion analyzing fashion and obtaining data by using a unified latent feature representation from various fashion data sources. The framework [4] considers homologous and heterogeneous which captures similarity and also preserving similarity within the same perspective.The approach includes two steps.First step is quintuplet-based ranking loss for multi-domain fashion and the other one is a cross-vies similarity ranking loss for multi-model data.By using the common latent feature representation the frameworks shows the similarity evaluation between vectors from the same or different modalities.This approach[4] shows its effectiveness in enhancing fashion retrieval and analysis tasks showing its potential for improving recommendation systems and trend prediction in the fashion industry.

M A Waheed et al [5] present a computer visions system using convolutional neural networks(CNN) like YOLO v3 for clothing detection and Residual Networks for classification aiming at analyzing fashion trends in e-commerce images.The model reaches a 90% accuracy in detecting clothing after being trained and tested on a portion of the Fashion image dataset and some manually collected dataset. The study[5]

shows the efficiency of CNN algorithms for clothing detection and classification tasks.The system [5] can also classify Indian,Western and Indo-Western cuisines based on state and season, showcasing the diversity of CNN based approaches beyond just fashion trend.

S Kim et al[6] presents a method to analyze fashion styles based on online customer reviews from e-commerce platforms. By using Topic Modelling with Latent Dirichlet Allocation(LDA) on datasets of online reviews from Amazon and Rakuten, the study[6] shows hidden topics within the review texts,throwing light on customer preferences in fashion industry.The obtained fashion style models provide valuable learning for marketing and product design specialists helping them to understand and customize the product based on customer preferences.This approach shows the potential of machine learning techniques in using online customer feedback to enhance business strategies and customer experiences in the fashion industry.

Ying Hong and Xiwen Shao [7] explores the application of machine learning methods for sentimental analysis of e-commerce product reviews. The collect data by web crawling and use 3 machine learning algorithms: Naive Bayes, Logistic Regression and Support Vector Machine for sentiment analysis. They find that Naive Bayes with informative words and two-word collocation gives the best results. They also say that too less features will give low accuracy while too many features will include a lot of noise, hence it is important to find the correct number of features to train the model.

R Verma et al [8] talks about the steps of building a fashion recommendation system. A very detailed literature survey was done and learning from it was used to design a new fashion recommendation system. The system is built in four steps: Information collection phase, learning phase, recommendation phase, and trend evaluation. The system depends highly on user feedback and provides recommendations accordingly. They have also suggested ways to make the system better including, feature engineering, selecting algorithms depending on the specific problem, and using hybrid algorithms.

Dr Ashok Kumar K et al[9] talks about how machine learning(ML) was used in a chatbot system of a college, so as to enhance its services. The Natural Language Processing(NLP) technique is used to input and generate appropriate responses. This [9] gives a general idea about the methods to use in the chatbot of this project, and the services it should provide.

MimiMalika et al[10] talks about the different kind of Machine Learning techniques used to categorize clothing images. The Machine Learning algorithms explored are support vector machine(SVM), KNN, etc. This [10] gives a general idea on how to classify the dataset, and more importantly which algorithms needed to be used to get the desired output.

M Ahmed et al [11] talk about building a Machine learning model which can predict colour trends in fashion. It works by scraping fashion data (images) from several websites using BeautifulSoup. It identifies the images as different types of clothings, later extracts the dominant colours from each image using K-means clustering. The most occuring colours from all the images are used to find the most trending colour in the fashion industry. ARIMA (Autoregressive integrated moving average) model is used to forecast the colour trend. The system has very low root mean squared error of 0.025 and accuracy of 97.5

Zadeh et al [12] talk about colour trend analysis using machine learning. It also introduces how histograms can be used to analyse colour patterns and trends. It talks about the use of histograms and how colour clustering and K-means clustering can be used in the same. It concludes that histograms are a powerful tool to depict clustering (colour trends) while clustering techniques are the best to classify colour attributes.

III. Methodology

A diverse dataset of fashion images and textual description representing various styles, colours and types are collected from Kaggle and other websites. Some more data about user preferences are collected through survey and user interaction through our chatbot. The dataset collected is cleaned and preprocessed. The steps include resizing and normalization of the images. Feature extraction is also be done for better results. A recommendation system is developed using several ML algorithms such as collaborative filtering, content-based filtering and hybrid models. The model is trained using the dataset collected and preprocessed to learn patterns and styles. The user preference data is also included.

In this project, implementation of a chatbot interface using Natural Language Processing(NLP) techniques is being done to enable interactions with the user. After integrating the ML model with the chatbot it will provide output recommendations based on user inputs and preferences. The model will be evaluated using metrics such as accuracy, precision, recall and F1-score. The model will also be evaluated based on customer review and satisfaction. Improvements shall be made according to the user review.

The performance of the metrices play a very important role in telling whether the classifier is good or not. We have implemented 4 metrices: Accuracy, Precision, Recall, F1 score. If accuracy, precision, recall and F1 score are all high, it tells that the classifier implemented is performing well. If the accuracy is high but rest are low, then the classifier could be biased towards the majority class. If accuracy is low, classifier is not performing well.

IV. Results and Discussion

The dataset used for our application is not well separated. It is also a multi-label dataset which makes the use of K-nn classifier not the best choice. Also the dataset is highly unbalanced with the class ”Casual” being more than 77% of the classes.

A high k-value will increase accuracy, but if it is too high

it will result in failing to capture the underlying patterns in data. (Underfitting). When the k is too small, the decision boundary becomes too complex and fits too closely to the training data (Overfitting). This may happen if the dataset contains irrelevant features. The dataset needs to be cleaned more and a better classifier can be applied.

From the assignment done practically, it was found that k value 3 is the most optimal one as per RandomSearchCV(). This was when the dataset was label encoded and only 2 features and 2 classes (gender were taken into consideration.). Overfitting might happen when the training data is learnt too well by the model, or in case of imbalanced dataset.

From the experiments done in lab 6, single layer perceptron can be used when we have a linearly separable data and multilayer perceptron when we non-linearly separable data like XOR gate logic. But it is important to have good hyperparameter tuning of weights and bias for good output and performance. The implementation took a lot of iterations and was slow. Experiments done involving multiple classifier models including, perceptron, multilayer perceptron, Naive Bayes classifier, Random Forest classifier, Decision Tree, AdaBoost Classifier, XGBoost Classifier and CatBoost, revealed that CatBoost worked best on our dataset with an apporximate accuracy of 0.87, precision of 0.94, recall of 0.88 and f1 score of 0.914.

REFERENCES

[1] Han, A., Kim, J., and Ahn, J. (2022). Color Trend Analysis using Machine Learning with Fashion Collection Images. Clothing and Textiles Research Journal, 40(4), 308-324. https://doi.org/10.1177/0887302X21995948

[2] *Use of Machine Learning in application of Artificial Intelligence* - Aditya Jain, Anubhuti Singh, Ishika Jain

[3] Zhao, L., Li, M., and Sun, P. (2024). Neo-Fashion: A Data-Driven Fashion Trend Forecasting System Using Catwalk Analysis. Clothing and Textiles Research Journal, 42(1), 19-34. https://doi.org/10.1177/0887302X211004299

[4] X. Gu, Y. Wong, L. Shou, P. Peng, G. Chen and M. S. Kankanhalli, "Multi-Modal and Multi-Domain Embedding Learning for Fashion Retrieval and Analysis," in IEEE Transactions on Multimedia, vol. 21, no. 6, pp. 1524-1537, June 2019, doi: 10.1109/TMM.2018.2876822.

[5] D. Shubhangi, B. Gadgay, P. J. Nandyal and M. A. Waheed, "Analyzing fashion trends using CNN and Cloth classification with respect to season and state," 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2022, pp. 629-634, doi: 10.1109/ICCCIS56430.2022.10037748.

[6] V. R. Hananto, S. Kim, M. Kovacs, U. Serdült and V. Kryssanov, "A Machine Learning Approach to Analyze Fashion Styles from Large Collections of Online Customer Reviews," 2021 6th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, 2021, pp. 153-158, doi: 10.1109/ICBIR52339.2021.9465830.

[7] Y. Hong and X. Shao, "Emotional Analysis of Clothing Product Reviews Based on Machine Learning," 2021 3rd International Conference on Applied Machine Learning (ICAML), Changsha, China, 2021, pp. 398-401, doi: 10.1109/ICAML54311.2021.00090.

[8] A. Bajpai, N. Dhanda, R. Verma and K. K. Gupta, "A Novel Approach Towards Fashion Trend Analysis System Using Machine Learning," 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2023, pp. 453-457, doi: 10.1109/ICPCSN58827.2023.00080.

[9] *Smart college chatbot using ML and Python*- Dr. Ashok Kumar K et all

[10] *Fashion images classification using Machine Learning, Deep Learning and transfer learning models*- Mimi Malika et all

[11] K. Nadeem, M. Ahmad, Z. Javed and M. A. Habib, "Development of a Machine Learning Model for Prediction of Colour Trends in Fashion Industry," 2022 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, 2022, pp. 296-301, doi: 10.1109/FIT57066.2022.00061.

[12] Zadeh, S.A., Iwendi, C., Uhumuavbi, I., Boulouard, Z. (2023). Colour Trend Analysis Using Machine Learning and Histogram. In: Iwendi, C., Boulouard, Z., Kryvinska, N. (eds) Proceedings of ICACTCE'23 — The International Conference on Advances in Communication Technology and Computer Engineering. ICACTCE 2023. Lecture Notes in Networks and Systems, vol 735. Springer, Cham. https://doi.org/10.1007/978-3-031-37164-6_33