**REPORT ON PERSONALIZED VIRTUAL FASHION RECOMMENDATION SYSTEM**

**Business Understanding 1-2**

* Introduction
* Problem statement
* Business Objectives
* Success Criteria

**Data Understanding 2-3**

* Overview
* Data Description
* Data Quality
* Data Preparation

**Preprocessing** **3**

**Modeling 4**

**Deployment 4**

**Next steps 5**

Group 2 Members

David Mwangi : <david.mwangi@student.moringaschool.com>

Lewis Gitari :<lewis.gitari@student.moringaschool.com>

Caroline Njoroge :<caroline.njoroge@student.moringaschool.com>

Suleiman Sheikh :<suleiman.sheikh@student.moringaschool.com>

James Ndiritu: [<james.ndiritu@student.moringaschool.com>](mailto:%3cjames.ndiritu@student.moringaschool.com%3e)

Diannah Wangura: [<diana.wangura@student.moringaschool.com>](mailto:%3cdiana.wangura@student.moringaschool.com%3e)

**Business Overview:**

**Introduction**

In today's digital age, the importance of personalized virtual fashion experiences cannot be overstated. With the exponential growth of online platforms and social media, fashion has become a powerful tool for self-expression. Individuals seek tailored guidance to navigate the vast landscape of digital fashion, where the absence of personalized recommendations often leads to dissatisfaction and confusion.

We aims to revolutionize the fashion industry with the introduction of a Virtual Fashion Recommendation System. In today's fast-paced world, consumers are constantly seeking personalized and convenient shopping experiences. Our solution leverages cutting-edge technology to provide users with tailored fashion recommendations, enhancing their shopping journey and at the same time boosting self-confidence from looking good and less hustle when deciding what to wear.

**Problem Statement**

Traditional fashion shopping experiences often lack personalization and convenience, leading to user frustration and dissatisfaction. Consumers struggle to find clothing items that match their unique style preferences, body type, and budget. Moreover, the overwhelming number of choices available online can make the shopping process daunting and time-consuming.

**Business Objectives**

Our Virtual Fashion Recommendation System addresses these challenges by employing advanced machine learning algorithms and computer vision techniques. By analyzing user preferences, historical behavior, and contextual data, our system generates personalized fashion recommendations tailored to each individual's tastes and preferences. Users can effortlessly discover new clothing items that align with their style, size, and budget, streamlining the shopping experience and

Other objectives are to:

1. Utilize deep learning models, such as convolutional neural networks particularly the Resnet50, to analyze visual features of fashion items.
2. Train models to understand compatibility and create complete, well-balanced outfits.
3. Allow users to upload inspirational images or reference items.
4. Build a search system using deep learning-based image analysis to identify similar or complementary styles.

**Success Criteria**

This project will be deemed successful when the following is achieved:

1. The project is to be completed on time using available resources.

2. Project achieving functional requirements such as:

* Allow users to upload inspirational images or reference items.
* Build a search system using deep learning-based image analysis to identify similar or complementary styles.
* Deployment of the model for easier interaction with the end-user

3. Project achieving other requirements such as:

● enhance target users digital fashion experience

● monetization by advertising partnerships with fashion brands, and commissions from product sales facilitated through the platform.

**Assessing the Situation**

1. Personnel are upcoming data scientists
2. Datasets-We used the FashionAI dataset which is a high-quality fashion dataset, contributed to the academic society. It covers different categories of women’s clothing.
3. Constraints - Large data size

**Data Understanding:**

**Overview**

The dataset used in this project is;

1. Fashion AI tianchi dataset

**Data description**

The total number of images in the FashionAI dataset is 79,573

There are eight mutually exclusive fashion attribute dimensions across the dataset, and each image containing one fashion item is labelled with one specific attribute value under the corresponding attribute dimension.

**Data Quality**

This dataset is a large-scale attribute dataset with manual annotations in high quality that

Addresses common issues like structured noise, occlusion, uncertain problems, and attribute inconsistency that pervasively exist in other public datasets.

This is due to the fact that it was manually put together by students who carefully did the labelling

**Data Preparation**

1. Loading data: the files were downloaded to our local machine. They were then

Loaded to our environment by reading them using open CV and interfacing with numpy to load them as array this will assist in knowing the shape of the image.

1. Data cleaning: We added an extra dimension at the beginning of the array, This expansion is often useful when working with neural network models that expect batch inputs. Adding this extra dimension allows you to process a single image as if it were part of a batch of images, which can simplify the code when handling both single and multiple images

**Preprocessing:**

Below were the steps followed :

We used the the preprocess input () function to preprocess images before feeding them into a neural network model according to the requirements of the ResNet50 model.

Secondly we used the model. Predict () function is used to obtain predictions from a neural network model.

Thirdly we got the results shape where the output was (2048,) suggesting that the model is producing a feature vector with 2048 elements.

Lastly we used the The norm () function to normalize the result array for better predictions. This operation effectively scales down the values in result so that they all fall within the range [0, 1] while preserving their relative dimension.

Since we have a huge dataset and time consuming we used tqdm module to track the progress of our preprocessing

**Modeling:**

**KNN Model**

We performed a grid search on the KNN to get the best hyper parameters and the results were as follows;

Number of neighbors: 3

Algorithm: auto

Metric: Euclidean

Explanation:

Number of neighbors: 3:

With a smaller number of neighbors, the model considers only a few neighboring data points when making predictions. This can lead to a more flexible and sensitive model that may capture local patterns well.

Algorithm: auto:

The algorithm "auto" automatically selects the most appropriate algorithm based on the input data and the specified distance metric. This allows the model to adapt to the characteristics of the dataset without manual intervention.

Metric: Euclidean:

Euclidean distance is a common metric used in KNN to measure the distance between data points in a multidimensional space. It calculates the straight-line distance between two points in space, which is suitable for continuous data.

**Deployment:**

We used Streamlit application for our web application which allows users to upload an image of a clothing item and receive recommendations for similar items from a dataset of fashion images.

User Interaction:

Users can upload an image using the file uploader widget.

The uploaded image is displayed, and recommendations are shown below it.

Functionality:

The application provides an intuitive interface for users to interact with the fashion recommendation system.

It handles image uploads, feature extraction, recommendation generation, and result visualization seamlessly.

Deployment Considerations:

The application is deployed using Streamlit, making it accessible via a web browser without the need for complex setup or installation.

It leverages pre-trained deep learning models and efficient recommendation algorithms to provide real-time recommendations to users.

The deployment process ensures that the application is scalable, reliable, and user-friendly.

**Next Steps:**

Further improvements could include enhancing the recommendation algorithm, incorporating user feedback for personalized recommendations, and optimizing the application's performance.

User interface enhancements, such as adding additional features or refining the layout, could enhance the overall user experience.

Continuous monitoring and updates to the application will ensure its effectiveness and relevance over time.

**Conclusion**

The integration of feature extraction tools, recommendation algorithms, and the Virtual Web App script offers a comprehensive solution for personalized fashion recommendations. Leveraging deep learning models like ResNet50, the Virtual Closet notebook extracts essential features from fashion images, enabling the creation of a robust feature vector dataset. The KNN grid search further refines the recommendation process by optimizing parameters for the K-Nearest Neighbors algorithm. These extracted features and optimized parameters are then utilized by the Virtual Web App script to provide users with personalized fashion recommendations in real-time. Together, these components form a data pipeline that transforms raw fashion images into actionable insights, enhancing the digital fashion experience for users.

**Recommendations**

1. Implement user feedback mechanisms to enhance recommendation accuracy, ensuring that user preferences are continuously updated and reflected in recommendations.

2. Improve the scalability of the virtual closet to accommodate a growing user base by optimizing backend infrastructure and database management.

3. Explore additional deep learning models, such as VGG16 or InceptionV3, to improve fashion feature extraction and further enhance recommendation accuracy.

4. Integrate user engagement analytics to track user interactions and preferences, allowing for a deeper understanding of user behavior and enabling more personalized recommendations over time.