CODS-COMAD Data Challenge Report (Sponsored by Meesho )

1.Introduction

We developed a Custom Neural Network for multi-attribute fashion image prediction, utilizing softmax activation, Adam optimizer (0.0002 learning rate), and sparse categorical cross-entropy loss function. The system processes five clothing categories: Men's T-shirts, Sarees, Kurtis, Women's T-shirts, and Women's Tops & Tunics.

Our model takes in one category and we predict the values of that category.

Our data preprocessing involved analyzing CSV files for each category, where we identified and removed columns with complete null values. The remaining attributes showed varying distributions of values, for example:

* Color attributes: default, multicolor, black, white
* Neckline types: round, polo
* Pattern styles: printed, solid
* Design variations: default, solid, typography
* Sleeve types: short sleeves, long sleeves

After visualizing the remaining columns' distributions, we addressed missing values by selecting optimal replacement values based on accuracy testing. The final steps involved integer encoding of categorical variables before loading and processing the images for prediction.

The system successfully integrates image processing with attribute prediction, providing an efficient solution for fashion item classification.

2. Data Preprocessing

**Data Preprocessing:** The initial dataset was organized into separate CSV files for distinct clothing categories (Men's T-shirts, Sarees, Kurtis, Women's T-shirts, and Women's Tops & Tunics). During preprocessing, we first identified and removed columns containing entirely null values. The remaining attributes showed varying value distributions across different features, such as color (default, multicolor, black, white), neckline (round, polo), pattern (printed, solid), and sleeve types (short sleeves, long sleeves).

**Data Exploration and Learnings:** Through exploratory data analysis, we discovered significant missing values across multiple attributes. After visualizing the remaining columns' distributions, we implemented a strategic approach of selecting and testing different value replacements for missing data, choosing the ones that yielded the best accuracy for our model.

**Feature Engineering:** The categorical variables in our dataset required integer encoding before model implementation. This transformation converted text-based attributes into numerical format suitable for our neural network. The final preprocessing step involved loading and standardizing the images alongside their corresponding encoded attribute values for model training.

3. Modeling Approach

Model Selection:

We implemented a custom Convolutional Neural Network (CNN) with multiple outputs for simultaneous attribute prediction. This architecture was chosen because:

* It enables multi-task learning, allowing the model to predict multiple attributes simultaneously
* CNNs are particularly effective for image processing tasks
* The shared convolutional base reduces computational overhead while maintaining feature learning capabilities

Architecture:

The model employs a hierarchical structure:

* Input Layer: Accepts 128x128x3 RGB images
* Convolutional Base:
  + Three sequential Conv2D layers (32, 64, 128 filters)
  + Each block includes MaxPooling2D, BatchNormalization, and Dropout(0.2)
  + Progressive feature extraction with increasing filter sizes

Final Model and Hyperparameters:

Model Details:

* Shared Dense Layer: 128 units with ReLU activation
* Output Branches:
  + The output of the layers are no of values each attribute gives out.
  + Softmax activation for multi-class classification
  + Output dimensions: 4, 2, 2, 3, and 2 classes respectively for T-shirt for example

Final Hyperparameters:

* Optimizer: Adam with learning rate 0.0002
* Batch Size: 64
* Training Epochs: 10
* Dropout Rates: 0.2 (convolutional layers), 0.5 (dense layer)
* Train-Validation Split: 70-30

Hyperparameter tuning focused on balancing model complexity with training stability. The learning rate was specifically chosen to be small (0.0002) to ensure stable convergence given the multi-task nature of the problem.

4. Novelty and Innovation

Unique Techniques:

1. Multi-Task Learning Architecture:
   * Single model predicting multiple attributes simultaneously
   * Shared convolutional base for efficient feature extraction
   * Separate output heads for specialized attribute prediction
2. Balanced Regularization Strategy:
   * Dual dropout rates (0.2 and 0.5) at different network depths
   * BatchNormalization after each major layer for training stability
   * Combined approach prevents overfitting while maintaining model capacity
3. Efficient Training Approach:
   * Unified loss function incorporating multiple attribute predictions
   * Balanced architecture allowing for simultaneous optimization of all attributes
   * Shared feature learning reducing computational requirements while maintaining prediction accuracy

These innovations allow for efficient processing of fashion images with multiple attribute predictions in a single forward pass, making it particularly suitable for real-world e-commerce applications.