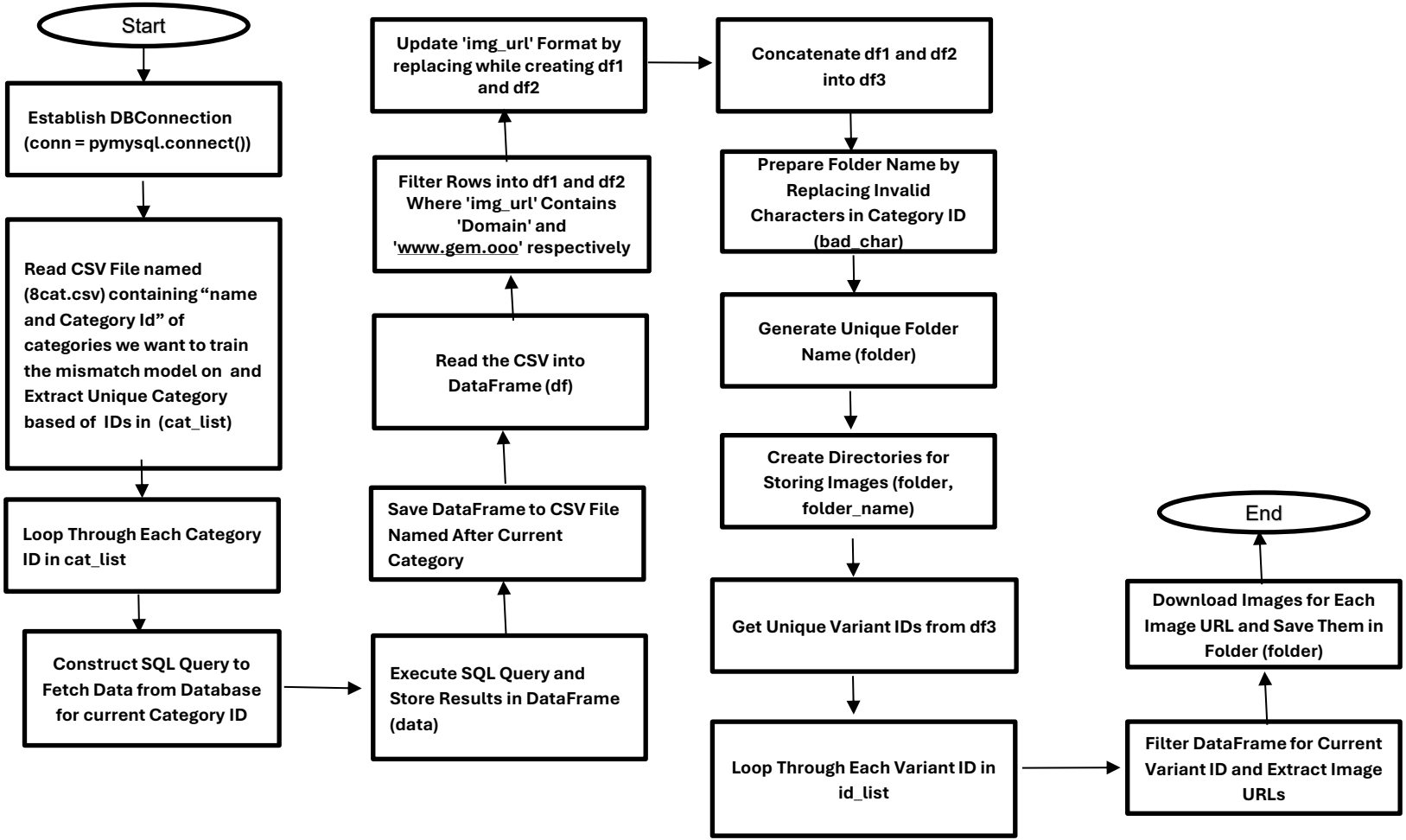
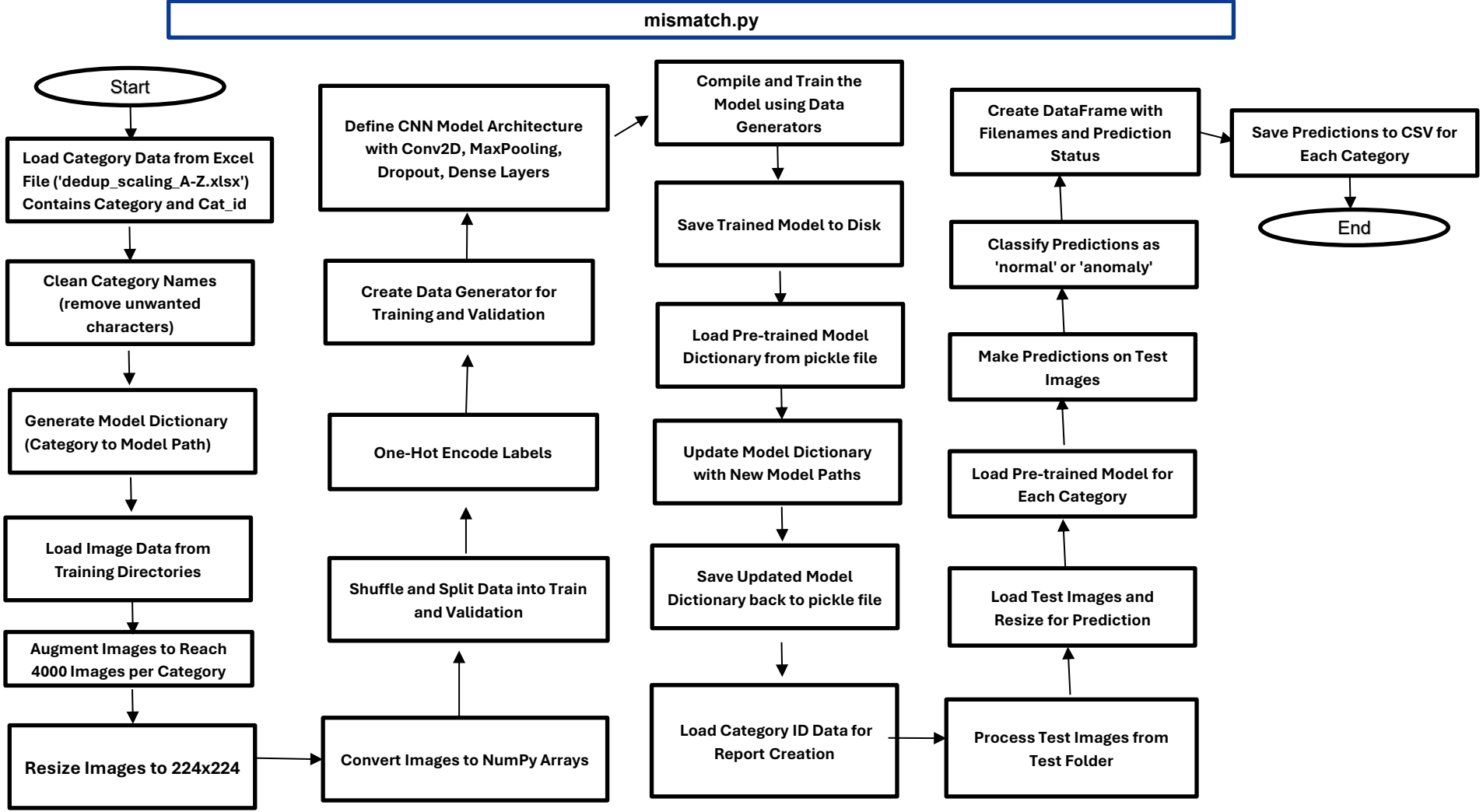
deduplication_market.py




mismatch.py
Start
Load Category Data from Excel File ('dedup_scaling_A-Z.xlsx') Contains Category and Cat_id
Clean Category Names (remove unwanted characters)
Generate Model Dictionary (Category to Model Path)
Load Image Data from Training Directories
Augment Images to Reach 4000 Images per Category
Resize Images to 224x224
Convert Images to NumPy Arrays
Shuffle and Split Data into Train and Validation
One-Hot Encode Labels
Create Data Generator for Training and Validation
Define CNN Model Architecture with Conv2D, MaxPooling, Dropout, Dense Layers
Compile and Train the Model using Data Generators
Save Trained Model to Disk
Load Pre-trained Model Dictionary from pickle file
Update Model Dictionary with New Model Paths
Save Updated Model Dictionary back to pickle file
Load Category ID Data for Report Creation
Process Test Images from Test Folder
Load Test Images and Resize for Prediction
Load Pre-trained Model for Each Category
Make Predictions on Test Images
Classify Predictions as 'normal' or 'anomaly'
Create DataFrame with Filenames and Prediction Status
Save Predictions to CSV for Each Category
End