CS 4361 - Machine Learning

Dr. Aritran Piplai

Group 10: Final Project Submission

Heaven Not Heaven

Ian Flores, Denisse Loya & Christina A. Valtierra

Department of Computer Science, The University of Texas at El Paso

May 9, 2024

Introduction

This project is titled Heaven Not Heaven, inspired by the main goal of accurately identifying which attendees matched the theme for the 2018 MET Gala Exhibition, whose inspiration was Heavenly Bodies. Heavily influenced by catholic fashion and the baroque style, this project used a dataset filled with images directly sourced from that year's exhibition within the museum and elements relevant to the fashion style, as seen after artists walk the red carpet. This exhibition was later released by the MET Museum and Vogue Magazine. These images were mixed alongside other images from other events such as the Kids Choice Awards, which served as our negative control to determine if the attendee was on theme or not after analyzing the image through a Support Vector Machine (SVM) model.

Methodology

This project made use of a custom curated dataset where we had two classes. The first one was *Heaven,* which consisted of images from the exhibition alongside other catholic and heavenly adjacent content. The second dataset, *Not Heaven,* consisted of pictures of people attending other non-related red carpet events as well as casual clothing.

Initially this project was designed utilizing a Convolutional Neural Network, however, after having a check-in session on our implementation's disadvantages, it was advised to not continue the use of this model. This is attributed to the fact that there were too many parameters and very few samples, which could very easily lead to an overfitting model.

These results caused a switch in the model being used. After discussion and through advice given by the professor, the chosen model to continue the project was an SVM. This model was able to

provide satisfactory results as well as being able to work with a higher and limited number of samples in the given dataset, without giving the previous problem of dealing with overfitting.

Setup

This project made use of multiple libraries, described as follows:

- *Pandas:* Used for data manipulation and analysis.
- *Scikit-learn tools:* Used for data mining and analysis, as well as training the SVM model.
- *NumPy:* Used for numerical calculations.
- *Matplotlib:* For result display.
- *Scikit-image (skimage):* For image manipulation and resizing.

Execution

As mentioned before, this project relies on multiple libraries. These were imported at the very beginning of the code execution, as seen in Figure 1, which represents the beginning of the code.

```
# Importing the required libraries for use of the SVM model
import pandas as pd
import os
from skimage.transform import resize
from skimage.io import imread
import numpy as np
import matplotlib.pyplot as plt
from sklearn import svm
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
✓ 5.8s
```

Figure 1 - Imports of the utilized Python libraries

To allow for good use of the images within the training steps in the SVM model, the images were first resized and flattened. A data frame was created in order to split the data observed into input versus output data. Figure 2 showcases how the input and output data was then split into the training and testing sets.

```
# Splitting the data into training and testing sets
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.20,
                                               random_state=77,
                                               stratify=y)
```

✓ 0.0s

Figure 2 - Section of the code dealing with splitting input and output data

The model was then created through the use of the scikit-learn library SVM, as seen in Figure 3.

```
# Building and training the model
# Defining the parameters grid for GridSearchCV
param_grid={'C':[0.1,1,10,100],
            'gamma':[0.0001,0.001,0.1,1],
            'kernel':['rbf','poly']}

# Creating a support vector classifier
svc=svm.SVC(probability=True)

# Creating a model using GridSearchCV with the parameter grid
model=GridSearchCV(svc,param_grid)
```

✓ 0.0s

Figure 3 - Beginning the creation of the SVM model

The model was trained using the fit function by providing the training set. After the training process was complete, a series of metrics were calculated to observe the results of the model, as seen in Figure 4. These included precision, recall, and F-1 score.

```
# Evaluating the model
# Testing the model using the testing data
y_pred = model.predict(x_test)

# Calculating the accuracy of the model
accuracy = accuracy_score(y_pred, y_test)

# Print the accuracy of the model
print(f"The model is {accuracy*100}% accurate")

print(classification_report(y_test, y_pred, target_names=['heaven', 'notheaven']))
```

✓ 1.0s

```
The model is 85.9375% accurate
              precision    recall  f1-score   support

      heaven       0.83      0.91      0.87        32
   notheaven       0.90      0.81      0.85        32

    accuracy                           0.86        64
   macro avg       0.86      0.86      0.86        64
weighted avg       0.86      0.86      0.86        64
```

Figure 4 - Calculation of metrics for measuring the model's accuracy

Lastly, the now trained model was tested by trying out distinct images from both the 2018 MET Gala Red Carpet event, versus other looks from different occasions. Figure 5 illustrates the contrast after running three images, two being on theme and the other not, which gave accurate results as output.

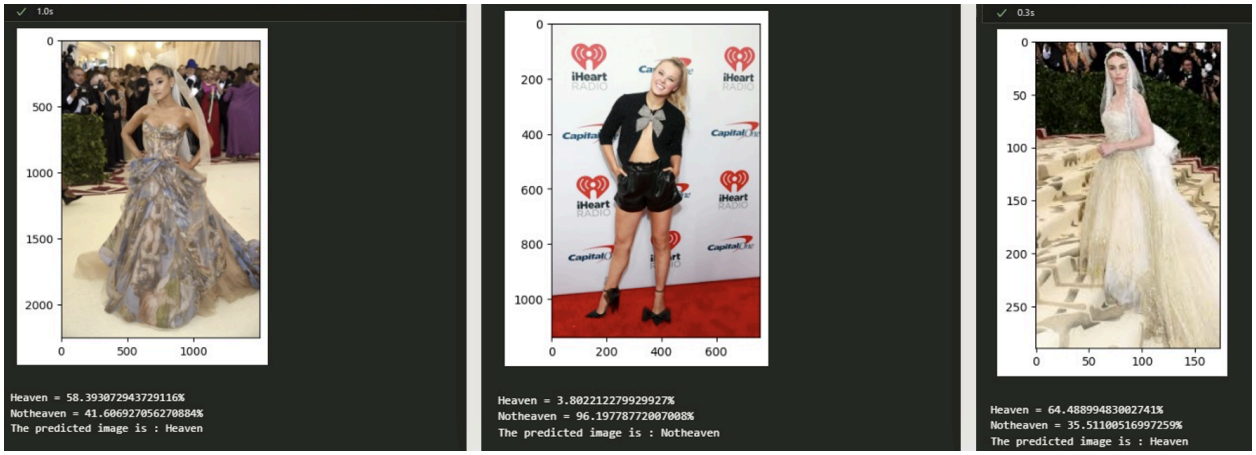

Figure 5 - Prediction results after running the model with three different images

Conclusion

The implementation of the project was both challenging and fulfilling at the same time. By merging two interests, computer science and fashion, this implementation allowed for a deeper understanding on how certain Machine Learning models work in real-life. The inspiration behind this were software engineers at higher fashion companies, which are dedicated to using data analysis on runways and designer showcases. This project was really fun, as well as eye-opening in a sense that it allowed for a different perspective on how these algorithms have a wide, never ending variety of applications. This team looks forward to continuing learning about Machine Learning and its distinct uses within interdisciplinary fields, aspiring to one day getting to apply these in a professional setting.