Wedding Florist Recommender

Michelle Lee, PhD

/github.com/michellesklee
/in/michelleseulkilee
michellesklee@gmail.com

Background

When planning a wedding, couples have the option of either hiring an expensive wedding planner or going DIY. While many services have moved online, there is currently a gap in online wedding planning. ***Wedfuly****, a startup based in Denver, is the first company to provide online wedding planners who still offer personalized services and are more affordable than traditional wedding planning.*

Objectives

As a new startup, Wedfuly wedding planners do their own research to recommend vendors to clients. The primary objective of this project was to build a recommender that makes the planners' job easier – recommending vendors that fit the clients' style and requirements such as budget and size of the wedding.

The recommender was built by training a model to cluster images and incorporating florist features (e.g., price range) to recommend florists that meet clients' style and needs.

Data

Training set*: 387 images of wedding bouquets collected online from florists in Colorado*
Test set*: 115 images from Wedfuly florists*
Validation set*: 40 unseen images*

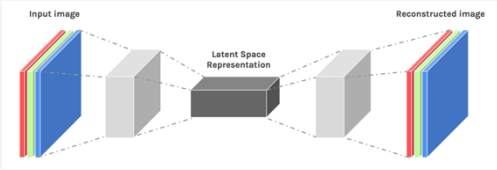
Preprocessing Data:

1. *Crop bouquets from images*
2. *Center bouquet and crop to square*
3. *Resize to 100x100*
4. *Convert to RGB*

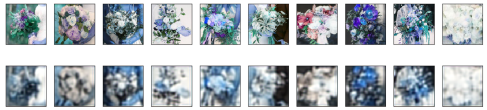
Methods

Step I. Convolutional Autoencoder

The convolutional autoencoder encodes and decodes images, and in the process identifies latent features for clustering similar images.

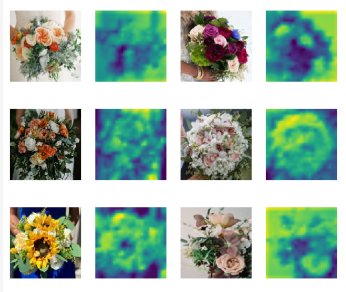


Reconstructed Images



Attention Model

An attention model represents images pooled up to the last convolutional layer and is useful for identifying where the model is paying most attention – bright spots and contrasts.



Methods

Step II. K-Means Clustering

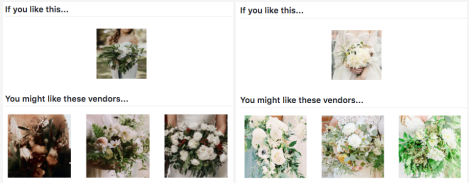
K-means clustering partitions observations into similar clusters, with each observation belonging to the cluster with the nearest mean.
The optimal number of clusters for this data was 7.

Results

Mean squared error (MSE) = .015 for both loss and validation after 30 epochs

Evaluating Model

The validation set of 40 unseen images was also clustered to evaluate model performance.



Step III. Cosine Similarities

Cosine similarity was calculated using florists' assigned clusters and features such as their price range and location. Clients enter their wedding specifications and choose images they like, and the recommender provides 2 vendors that meet their specifications.

Conclusion

A model incorporating a convolutional autoencoder and K-Means clustering was able to create meaningful clusters that became features for the full recommender. As Wedfuly continues to gains vendors and more data is collected, the model may be further fine-tuned and validated.

The current recommender was built for florists; however, the model will extend to other vendors such as photographers as the company grows.

Tech Stack

pandas · scikit-learn machine learning in Python · Keras · Flask · NumPy · OpenCV · matplotlib · TensorFlow · amazon web services