**Image Classification: *Fashion Statement***

**Introduction:**

Image classification has emerged into one of the most important areas of machine learning throughout many fields such as biomedical, multimedia marketing, real estate, geospatial intelligence, and even fashion. The classification of images aspires to assists with the solving of problems. These problems can range from detecting disease within human DNA, to the advancement of autonomous vehicles. As the field of data science evolves, the problems become more complex and humans are turning to deep machine learning to create solutions.

The fashion industry is leading not only in the technology used during fashion shows, but also in how merchandise is arranged on the floor within a store. John Cannon, former buyer for Ralph Lauren, mentioned that every store performs in depth analysis based on image classification and the availability of a square feet available by product. He said that this helps to determine placement as well as how pieces of clothing are perceived by the human eye. Humans will look at an article of clothing and make predictions based on color, size, texture, and even the way the article of clothing drapes on the hanger. Machines on the other hand, remove texture and color to help make predictions of articles of clothing being represented based on outlines, grey-scaled pixels, and how the image compares to correctly predicted training images.

As humans we might classify objects differently than another, but by training a machine to use images for classification the accuracy becomes more efficient. When humans classify an image or an object, one might have the tendency to add bias. With deep machine learning we are able to train machines with 2-D flattened images and different algorithms for better classification based on different parameters.

**Business Problem:**

The analysis will discuss several algorithmic approaches, as well as the speed and accuracy in which the images are classified, and the trade-offs with each approach. Each model is scored for the time in which it takes to execute as well as the accuracy of predictions of article of clothing. The analysis will provide a deeper look at the way the fashion industry uses image classification to determine floor placements.

**Obtaining the Data:**

The fashion MNIST data set is first obtained from the repository of zalandoresearch located on GitHub. The data sets can also be obtained from Kaggle if needed. There are two training files and two testing files. Each set contains the images as well as the labels that are needed for each image.

The fashion data set used contains 70,000 images of articles of clothing that are equally distributed among each item of clothing. The data set is loaded into Colab using the Tensorflow package and the keras module within the package. When loading the data using keras, it is split into training and testing sets for prediction analysis. The training set is used to teach the machine about the articles of clothing, and what each piece looks like. The testing set is used for the prediction of the article of clothing and the accuracy in which the model predicted the right piece. The training set consists of 60,000 images, while the testing set contains 10,000.

Within the fashion data set there are 10 classes that correspond to a different article of clothing. These 10 classes can be seen below.

***Classes for each article of clothing:***

|  |  |
| --- | --- |
| Label | Item of Clothing |
| 0 | T-shirt/top |
| 1 | Trouser |
| 2 | Pullover |
| 3 | Dress |
| 4 | Coat |
| 5 | Sandal |
| 6 | Shirt |
| 7 | Sneaker |
| 8 | Bag |
| 9 | Ankle Boot |

***First 25 articles of clothing within the fashion data set:***



**Scrubbing the Data:**

Each image within the fashion data set is 28 x28 pixels, but each pixel has a value between 0 and 255. Normalization of each pixel is completed and changed to value between 0 and 1. This allows for the creation of the models and overfitting. Also, it is imperative that the images be two dimensional and in grey scale. The machine learning techniques used do not evaluate color, but rather the grey scale variation within each image.

**Classification Models:**

***Keras CNN Basic Model:***

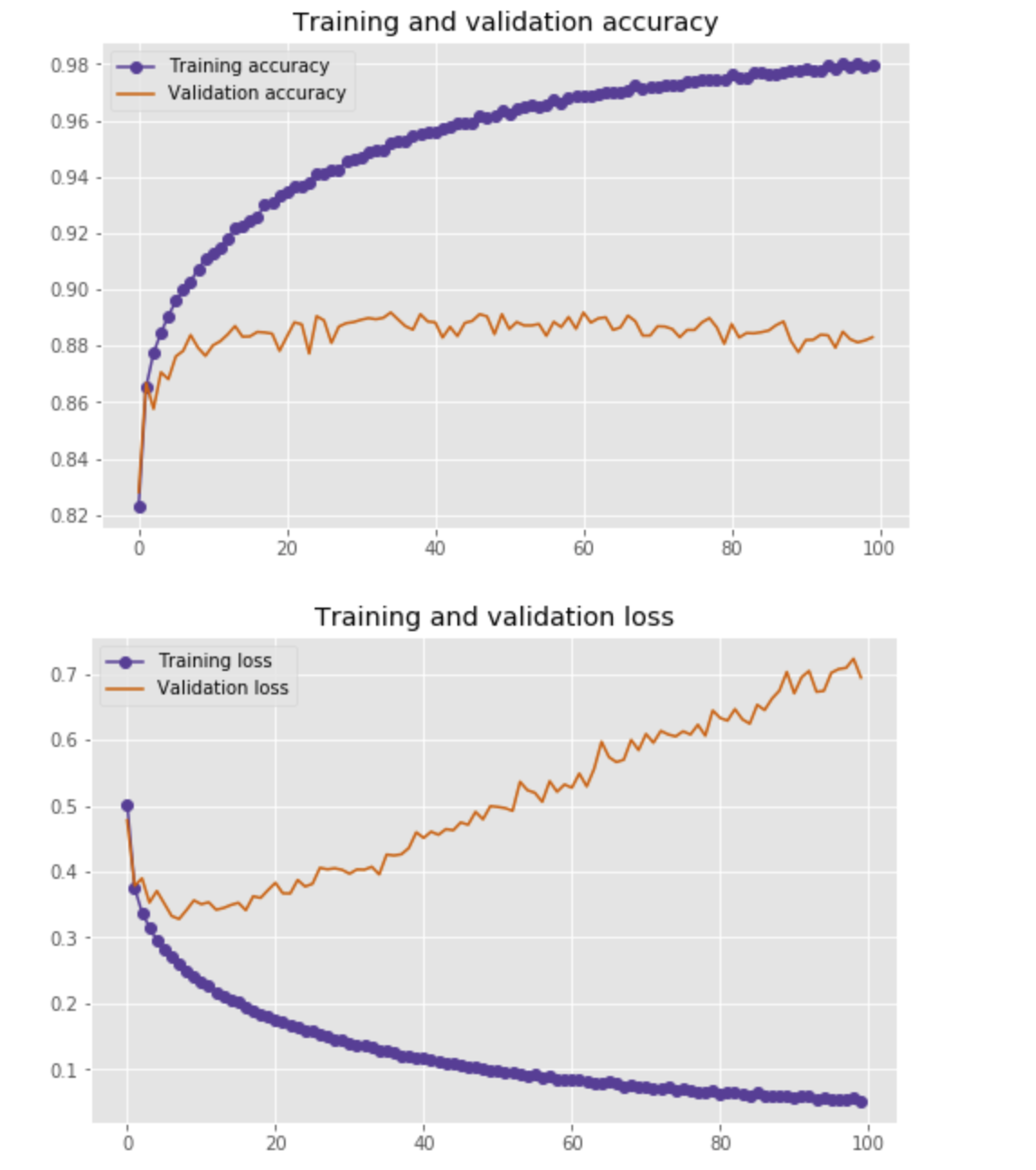
The Keras model is developed first as it is chosen for its user-friendly approach and the ability to add different functions with ease. Keras is a convolutional neural network (CNN) that is pixel dependent. CNNs “can be controlled by varying their depth and breadth, and they also make strong and mostly correct assumptions about the nature of images” (Krizhevsky, A. Sutskever, D. Hinton, G. 1). CNNs have the ability to have fewer connections and parameters so they are easier to train. Even though this is the case one trade-off when using a CNN is that they may be easy to implement, they can be very expensive to apply in large scale to high-resolution images.

The Keras Sequential is set up using layers to extract different representations of the articles of clothing. The model then flattens the images from two dimensional to one dimensional. This model also contains 128 fully connected neural layers with 10 fully connected nodes. Each node has a score that tells the probability that the current image belongs to one of the classes of clothing. The 128 fully connected neural layers are dense layers using the activation function relu. The 10 fully connected nodes use the activation function of softmax.

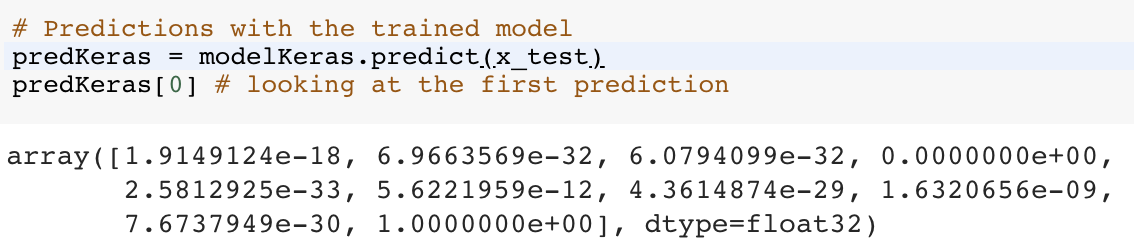
The model is then comprised of more features to determine accuracy during training, loss, metrics of accuracy used during training and testing, and the fraction of the images that are correctly classified.

The Keras model is ran with 100 epochs during the training phase to allow for better accuracy. The Adam optimizer is chosen as the optimizer for the model because it is straightforward to implement, requires little memory, and tuning of parameters. The Adam optimizer computes learning rates that are adaptive for different parameters from estimations. This optimizer is comprised of two popular methods AdaGrad (sparse gradients) and RMSProp(works well in on-line and non-stationary settings).

When the accuracy from the Keras CNN model is plotted we are able to see the training accuracy against the validation accuracy and the same for loss. The following visualizations indicate where the model starts to become stable, which is around 95 epochs. This model should only take a few minutes to run, or around 670 seconds to be more exact. The accuracy of the Keras CNN model is 88%. Overall the model is efficient in the speed it takes to run the model as well as the accuracy, but it could be better.



Looking at the first prediction of the model an array is produced indicating the confidence of the model that the image corresponds to each class of clothing.



The maximum confidence of the first prediction is 9 and the 9th label for the article of clothing is ankle boot.



This indicates that the model predicted correctly.

Looking again at the first 25 articles of clothing and their predictions might be able to tell us where our predictions were not right.



The image in red, indicates that 99% of the time that image is misclassified as a sandal when it is a sneaker. This might be due to the fact that it is facing the other direction from most of the other shoes in the data set.

When designing a store floor, it is imperative that the classification of items are correct. If the store had this dataset and they had the image of the sneaker that was misclassified this could lead to product misplacement on the floor. It would need to be placed with the sneakers not the sandals.

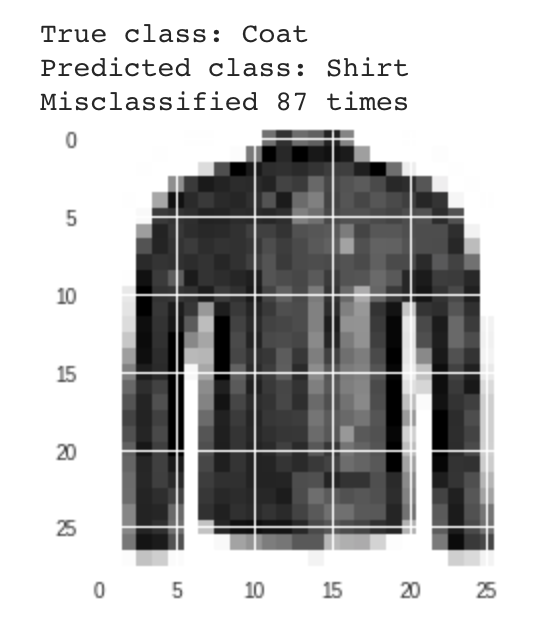
***Logistic Regression:***

Logistic regression is another model used for predictive analysis and is used on the fashion data set to determine the relationship of pixelized arrays. It addresses the probability that the images are classified correctly based on the pixels in the original images.

The logistic regression model did not classify all of the items of clothing correctly, but we are able to look at those that were misclassified for further explanation.

|  |  |  |
| --- | --- | --- |
| True Class | Predicted Class | Misclassified |
| Pullover | Coat | 134 times |
| Shirt | T-shirt/top | 131 times |
| Shirt | Pullover | 122 times |
| Coat | Pullover | 110 times |
| Shirt | Coat | 105 times |
| T-shirt/top | Shirt | 102 times |
| Pullover | Shirt | 91 times |
| Coat | Shirt | 87 times |
| Sandal | Sneaker | 50 times |

There are many shirts, pullovers, coats, and t-shirt/tops that seem to be very misleading when it comes to the image classification. For instance, the following coat, has a strong resemblance to a shirt.



One of the biggest disadvantages of logistic regression is that we cannot solve non-linear problems with this algorithm. It would be better to use random forest to get a better result.

***Gaussian Naïve Bayes:***

Gaussian Naïve Bayes classifier is a conditional probability algorithm also used for classification. This algorithm is based on the probability of object A falling under object B. This classifier is chosen to be ran on the fashion data set because it is considered the simplest and most popular for image classification. However, it may be easy to implement, the accuracy predictions prove otherwise.

When creating the predictions on the fashion data set the Gaussian Naïve Bayes classifier has the lowest accuracy, despite being the fastest. Theoretically if a department store was dealing with these items of clothing and was doing their floor model based on these predictions, they would again have coats with shirts, and tops with t-shits and items would not be very

|  |  |  |
| --- | --- | --- |
| True Class | Predicted Class | Misclassified |
| Sandal | Sneaker | 660 times |
| Pullover | Coat | 545 times |
| Shirt | Coat | 435 times |
| Dress | Trouser | 387 times |
| Ankle Boot | Sneaker | 304 times |
| Shirt | Dress | 200 times |
| T-shirt/top | Dress | 162 times |
| Bag | Coat | 149 times |
| Coat | Dress | 131 times |
| Shirt | T-shirt/top | 117 times |

As discussed previously, giving this prediction to a store would also have detrimental impacts on store layout. If dresses are being classified as trousers that would cause the store to have to move packaging that would be delivered to one department have to be transferred to another. This also results in inefficiencies in planning the floor, delivering the product for placement, and the man power to move products to their intended place.

***Advanced Keras Convolutional Neural Network:***

The last model is an advanced Keras convolutional neural network that consists of six sigmoid activations functions as well as a final softmax activation function. The pooling of this CNN is 2 with the learning rate of 0.1. The accuracy rate of this convolutional neural network is 92% with an hour of training time.

|  |  |  |
| --- | --- | --- |
| True Class | Predicted Class | Misclassified |
| T-shirt/top | Shirt | 109 times |
| Shirt | T-shirt/top | 90 times |
| Coat | Shirt | 71 times |
| Pullover | Shirt | 62 times |
| Shirt | Pullover | 57 times |
| Coat | Pullover | 49 times |
| Dress | Shirt | 45 times |
| Shirt | Coat | 45 times |
| Pullover | Coat | 41 times |
| Ankle Book | Sneaker | 39 times |

This more advanced neural network is able to make better predictions with the parameters that were added.

**Model Comparison:**

The three models have very different computing time as well as accuracy. Tensorflow was used through Colab for faster reporting on a MacBook Pro. Colab uses a GPU for processing the images not on a local server.

Each model’s performance is seen below in the chart.

|  |  |  |
| --- | --- | --- |
| Model: | Compute Performance: | Accuracy: |
| Keras CNN basic | 667.29 seconds / 11.12 min | 88% |
| Logistic Regression | 1070.85 seconds /17.84 min | 85% |
| naïve Bayes | 226.37 seconds / 3.77 min | 58% |
| Keras CNN Advanced | **3622.47 seconds / 60.37 min** | **92%** |

**Recommendation:**

As John Cannon, a buyer from Ralph Lauren, mentioned, “product placement is everything.’ Using object classification for clothing items helps to determine the right product placement on the floor and visualize each classification. Having the right prediction model for this process is vital to a visually appealing floor. As humans we may classify items of clothing differently but by being able to train a machine to do the classification, we are able to have better predictions as it continues to learn each item. It is recommended that using a Convolutional neural network will return the best accuracy with an appropriate learning rate, pooling, and number of fully connected nodes and activation functions.

**Sources:**

A. Vailaya, M.A.T. Figueiredo, A.K. Jain, Hong-Jiang Zhang. Image classification for content-based indexing. *IEEE Transactions on Image Processing.* Vol. 10, Issue 1. January 2001. pg 117-130. DOI 10.1109/83.892448 <https://pdfs.semanticscholar.org/142f/056a365dccd029c0897fcfa7833aecf2212f.pdf>

Kingma, D. Ba J. Adam: A Method for Stochastic Optimization. Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980). <https://arxiv.org/abs/1412.6980v8>

Krizhevsky, A. Sutskever, D. Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>

<https://github.com/zalandoresearch/fashion-mnist>

<https://www.kaggle.com/zalando-research/fashionmnist>

<https://blog.acando.no/thedailypassion/202222/building-a-simple-neural-network-in-1,-2,-3-%20-for-fashion!->