Model Descriptions:

Baseline: Logistic Regression
A logistic regression model is fit on the noisy dataset (no changes have been made to the labels). RGB histogram features are pulled from each image and work as the basis for the logistic regression model. The logistic regression model classifies images on a binary, meaning, a given image does or does not meet the criteria of a specified class (there is no other outcome possibility). The overall accuracy was found to be 0.24. Though it has a larger rate of accuracy than randomly guessing (0.1), the noisy data is quite disruptive to prediction in a logistic setting given that the model is poorly informed during training.

Model 1: CNN
Model 1 uses a convolutional neural network model to train the noisy dataset (again with no changes to the labels, allowing incorrect labels to exist as it). The images in the training dataset are converted into tensors of dimensions height, width, RGB. Neural networks are comprised of an input layer, hidden middle layers, and an output layer. This model uses 2 convolutional layers and 3 fully connected layers. With a given training input, the image passes through the layers and informs the model of characteristics in this dataset. From this feature extraction, the output is then informed. Padding, stride, and other nuances are set to minimize loss.
With this model, the prediction accuracy on noisy data was found to be 0.263. This demonstrated that moving to a CNN model from a logistic model was enough to allow for greater accuracy in prediction since both models were trained on the same noisy data. In some cases, the CNN model predicted the image correctly, even if it originally was incorrectly labeled as a part of the noisy data.

Model 2: Inception-v3
Model 2 extends the idea of a convolutional neural network and uses Inception-v3. Inception-v3 is a CNN that has 48 layers. This model uses factorized convolutions to simplify the parameters involved in a network, with the aim to make computation more efficient. The overall number of parameters are decreased by involving smaller and asymmetric convolutions. This model features smaller CNNs between layers to refine feature extraction. This model also aims to more efficiently reduce grid size with the goal of smoother computation. Like the preceding models, this model is trained on the clean data to allow for a more accurately informed model. Before use, the clean training images' resolution was changed to 75x75x3 to allow for greater clarity in features, improving spatial resolution. As a result, the inception model accuracy was found to be 0.8037. The allowance for more detailed feature extraction and a model trained with clean labels contributed largely to the great increase in prediction accuracy.

Sources:
https://blog.paperspace.com/popular-deep-learning-architectures-resnet-inceptionv3-squeezenet/