3 ML Features and RFE

- GLCM (Gray-Level Co-occurrence Matrix): Analyzes texture by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, useful in image classification.

- LBP (Local Binary Patterns): A texture descriptor that summarizes local structures in images by comparing each pixel with its neighborhood, often used in facial recognition.

- HOG (Histogram of Oriented Gradients): Captures the edge directions and distributions in localized portions of an image, commonly employed in human detection within images.

- RFE (Recursive Feature Elimination): A feature selection method that fits a model and removes the weakest feature (or features) until the specified number of features is reached, enhancing model accuracy.

6 ML Classifiers

- Logistic Regression: A statistical model that in its basic form uses a logistic function to model a binary dependent variable, widely used for binary classification tasks.

- Random Forest: An ensemble learning method for classification and regression that operates by constructing multiple decision trees and outputting the mode of their predictions, providing robustness.

- SVM (Support Vector Machine): A supervised learning model that uses classification algorithms for two-group classification problems, it's effective in high-dimensional spaces.

- XGBoost (eXtreme Gradient Boosting): An optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable, often winning many machine learning competitions.

- Adaboost (Adaptive Boosting): A boosting algorithm that can be used with many types of data and classifiers, it focuses on classification problems and aims to convert a set of weak classifiers into a strong one.

- HistGrading Boosting: A type of Gradient Boosting that uses histograms to summarize the gradient information, which can speed up training and handle large datasets with reduced memory usage.

6 CNN Models

- VGG16: A deep convolutional network for image recognition that proved that depth is a critical component for good performance, its uniform architecture makes it easy to manage.

- VGG19: Similar to VGG16 but with three more convolutional layers, this network has more depth which can lead to better feature learning, at the cost of increased computational complexity.

- ResNet152V2: A deep residual network that uses skip connections or shortcuts to jump over some layers, helps to solve the vanishing gradient problem and allows for very deep networks.

- InceptionV3: An advanced CNN with inception modules that apply convolutions of different sizes simultaneously, improving computational efficiency and model performance on image data.

- MobileNet: Designed for mobile and embedded vision applications, MobileNet is based on depth-wise separable convolutions which significantly reduce the number of parameters without losing performance.

- DenseNet121: A CNN where each layer is directly connected to every other layer in a feed-forward fashion, strengthens feature propagation and encourages feature reuse, which significantly reduces the number of parameters.

Hyperparameters

- Epochs: Refers to one cycle through the full training dataset, more epochs can mean better learning, but also risk overfitting if not managed with techniques like early stopping.

- Loss: The objective that a model's training algorithm seeks to minimize, different loss functions can be used depending on the nature of the problem, like mean squared error for regression.

- Learning rate: A hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated, a critical parameter that can affect model performance and training speed.

- Classifier: A type of machine learning algorithm that makes predictions or classifications based on the input data, examples include logistic regression and neural networks.

- Train-test split: The process of dividing a dataset into subsets for training and testing, which allows for the evaluation of a model on unseen data to assess its predictive performance.

- Augmented images: Enlarged dataset through transformations like rotation, translation, and flipping of images, helps the model generalize better by simulating various perspectives and conditions.