**Machine Learning Workflow Summary**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pre-Preprocessing &  Feature Engineering | Supervised Learning | Unsupervised Learning | Model Evaluation | Model Tuning |
| * Encoding and Binning * Label Encoding * One Hot Encoding * Ordinal Encoding * Binary Encoding * Frequency Encoding * Scaling and Transformation * Standard Scalar * MinMax Scalar * Robust Scalar * MaxAbs Scalar * Log Transformation * Power Transformation * Text Processing * Regex * Stemming/Lemmatization * Tokenization * POS Tagging * Stopwords Removal * Vectorization * Count * Tf-Idf * Word2Vec * BERT Embeddings * Dim Reduction * PCA * Filter Methods * Embedded Methods * Wrapper Methods * Feature Creation * Mathematical Combinations * Integration Features * Aggregations * Imbalanced Dataset * Under Sampling * Over Sampling * SMOT Tomek * Python Libraries * Pandas * NumPy * NLTK/SpaCy for Text * Imblearn * Statsmodels * Scikit-learn | * Regression * Linear Regression * Polynomial * Lasso/Ridge * Elastic Net * SVR * KNN Regressor * Booting Regressor * Random Forest Regressor * Classification * Tree Based Models * Decision Tree * Random Forest * Extra Trees * Gradient Boosting Machine (GBM) * XGBoost * AdaBoost * LightGBM * CatBoost * KNN Classifier * Logistic Regression * SVC * Naive Bays * Ensemble Techniques | * Clustering * PCA | * Bias Variance Tradeoff * Cross Validation   + K-Fold   + Stratified k-Fold   + LOOCV   + Time Series Split * Confusion Matrix for Classification   + Accuracy   + Precision   + Recall   + F1 Score   + AUC/ROC * Model Evaluation for Regression   + R-Squared (Accuracy)   + MSE   + MAE   + RMSE | * Grid Search * Random Search * Genetic Algorithm * Bayesian Optimization |

**Deep Learning Workflow Summary**

|  |  |  |  |
| --- | --- | --- | --- |
| **DL Basics** | **Image/Video (CV)** | **Text (NLP)** | **Library/Framework** |
| * ANN Architecture * Neural Networks Architecture * Neural, Input-Output, Hidden Layers * Activation Functions (ReLU, Sigmoid, SoftMax, Swish, Tanh & Linear) * Loss Functions (Cross-Entropy, MSE/MAE) * Optimizers (SGD, Adam, RMSprop) * Forward and Backward Propagation * Epochs, Batch Size, Learning Rate * Overfitting and Regularization (Dropout, L2) * Vanishing and Exploring Gradient | * Image Processing * Normalization * Augmentation * Color Conversion * Image Classification using CNN * Raw Pixels * CNNs (Backbone) * Pretrained CNNs * VGG * ResNet * ImageNet * Others | * Text Processing (Details in ML) * Sequence Models * RNNs * LSTMs * GRUs * BERT * Core Text Tasks in NLP * Text Classification * Sentiment Analysis | * Keras * TensorFlow |