Here’s a **complete categorized list** of all the essential components in a practical Machine Learning (ML) and Deep Learning (DL) pipeline, based on your request:

**🟡 1. Types of Target Values (Output Types)**

**🔹 Supervised Learning:**

* **Binary Classification** → Target: 0/1 or True/False
* **Multiclass Classification** → Target: categories (e.g., cat/dog/horse)
* **Multilabel Classification** → Multiple classes at once (e.g., "dog" + "brown")
* **Regression** → Continuous value (e.g., price, temperature)
* **Ordinal Regression** → Ordered discrete categories (e.g., low < medium < high)

**🔹 Unsupervised Learning:**

* No target value → Clustering, Dimensionality Reduction

**🔹 Time Series:**

* Target: sequential values (e.g., stock price next day)

**🟡 2. Data Cleaning Techniques**

**🔹 Missing Values**

* Drop rows/columns
* Impute with mean/median/mode
* Use advanced imputation (e.g., KNN, MICE)

**🔹 Outliers**

* Detect with IQR, Z-score, or Isolation Forest
* Remove, cap, or transform them

**🔹 Duplicates**

* Remove duplicated rows

**🔹 Inconsistent Data**

* Fix typos in categories (e.g., "Male", "male", "MALE")
* Standardize units/formats

**🔹 Imbalanced Classes**

* Resampling: SMOTE, ADASYN, RandomOversampler
* Use class weights in model

**🟡 3. Data Preprocessing Techniques**

**🔹 Feature Encoding**

* Label Encoding
* One-Hot Encoding
* Target Encoding / Mean Encoding
* Embedding layers (for DL)

**🔹 Feature Scaling**

* Standardization (Z-score)
* Normalization (MinMax)
* RobustScaler (for outliers)

**🔹 Feature Engineering**

* Polynomial features
* Date/time features (day, month, is\_weekend)
* Text features (TF-IDF, Bag of Words, embeddings)

**🔹 Text Preprocessing**

* Lowercasing, removing stopwords/punctuation
* Tokenization
* Stemming/Lemmatization
* Word embeddings (Word2Vec, GloVe)

**🔹 Image Preprocessing**

* Resize, normalize pixels
* Data augmentation (flip, rotate, zoom)
* Color space transforms (RGB → grayscale)

**🟡 4. Types of Models**

**🔹 Traditional ML Models**

* Linear Regression, Logistic Regression
* Decision Trees
* Random Forest
* XGBoost / LightGBM / CatBoost
* K-Nearest Neighbors (KNN)
* Naive Bayes
* Support Vector Machine (SVM)
* K-Means / DBSCAN (clustering)
* PCA, t-SNE, UMAP (dimensionality reduction)

**🔹 Deep Learning Models**

* **Feedforward Neural Networks (FNN / MLP)**
* **Convolutional Neural Networks (CNN)** – for vision
* **Recurrent Neural Networks (RNN)**, **LSTM**, **GRU** – for sequences
* **Transformers (BERT, GPT)** – for NLP & vision
* **Autoencoders** – for unsupervised learning
* **GANs (Generative Adversarial Networks)** – for generative tasks
* **Diffusion Models** – image generation

**🟡 5. Model Evaluation Techniques**

**🔹 Classification Metrics**

* Accuracy
* Precision, Recall, F1 Score
* ROC-AUC
* Confusion Matrix
* PR AUC (for imbalanced data)
* Log Loss

**🔹 Regression Metrics**

* Mean Absolute Error (MAE)
* Mean Squared Error (MSE)
* Root Mean Squared Error (RMSE)
* R² Score

**🔹 Clustering Metrics**

* Silhouette Score
* Davies-Bouldin Index
* Adjusted Rand Index
* Calinski-Harabasz Index

**🔹 Cross-Validation Methods**

* K-Fold CV
* Stratified K-Fold (for classification)
* Leave-One-Out CV
* Time Series CV (forward chaining)

**🔹 Model Monitoring (Post-deployment)**

* Drift detection (data distribution change)
* Model confidence
* Latency & throughput
* Production metrics (user success, revenue impact)