**Supplemental File**

1. **Methods and materials**

**Supplemental table S1.** The semantic definition of pyradiomics features.

|  |  |  |
| --- | --- | --- |
| pyradiomics feature | Abbreviation | semantic feature description |
| original\_firstorder\_10Percentile | FO\_10P | The 10th percentile of the gray level intensity in the region of interest, indicating the intensity value below which 10% of the voxels fall. |
| original\_firstorder\_90Percentile | FO\_90P | The 90th percentile of the gray level intensity in the region of interest, indicating the intensity value below which 90% of the voxels fall. |
| original\_firstorder\_Energy | FO\_E | A uniformly bright tumor = more light energy \_ high energy A patchy or dark tumor = less light energy \_ low energy |
| original\_firstorder\_Entropy | FO\_En | Measure of the randomness or disorder in the distribution of voxel intensities. High entropy indicates greater heterogeneity. |
| original\_firstorder\_InterquartileRange | FO\_IQR | The difference between the 75th and 25th percentile of the gray level intensity in the region of interest, representing the spread of the middle 50% of intensity values. |
| original\_firstorder\_Kurtosis | FO\_Ku | A measure of the 'peakedness' or 'flatness' of the gray level intensity distribution. High kurtosis indicates a sharper peak and fatter tails, while low kurtosis indicates a flatter distribution. |
| original\_firstorder\_Maximum | FO\_MAD | The largest gray level intensity in the region of interest. |
| original\_firstorder\_Mean | FO\_MaxI | The average gray level intensity in the region of interest. |
| original\_firstorder\_MeanAbsoluteDeviation | FO\_MedI | The average of the absolute differences between each gray level intensity and the mean intensity. |
| original\_firstorder\_Median | FO\_MI | The median gray level intensity in the region of interest. |
| original\_firstorder\_Minimum | FO\_Min | The smallest gray level intensity in the region of interest. |
| original\_firstorder\_Range | FO\_R | The difference between the maximum and minimum gray level intensity in the region of interest. |
| original\_firstorder\_RobustMeanAbsoluteDeviation | FO\_rMAD | The mean absolute deviation calculated on the intensities within the 10th and 90th percentile, reducing the effect of outliers. |
| original\_firstorder\_RootMeanSquared | FO\_RMS | The square root of the mean of the squared gray level intensities, providing a measure of the magnitude of the intensities. |
| original\_firstorder\_Skewness | FO\_Sk | A measure of the asymmetry of the gray level intensity distribution. Positive skewness indicates a tail to the right (more high intensity values), negative skewness indicates a tail to the left (more low intensity values). |
| original\_firstorder\_TotalEnergy | FO\_TE | Sum of squared voxel intensities, indicating overall energy/brightness of the tumor. High energy can be found in both uniform and heterogeneous tumors. |
| original\_firstorder\_Uniformity | FO\_Un | Measures the sum of the squares of each intensity value divided by the total number of voxels. A higher value indicates a more uniform distribution of intensities. |
| original\_firstorder\_Variance | FO\_V | The variance of the gray level intensities, indicating the spread of intensity values around the mean. Higher values indicate more variation. |
| original\_glcm\_Autocorrelation | GLCM\_AC | Measures the magnitude of the fineness and coarseness of the texture. High values suggest a coarser texture or more ordered patterns. I will classify it as 'Rest (0)' because coarseness aligns with heterogeneity. |
| original\_glcm\_ClusterProminence | GLCM\_Co | Measures the skewness and uniformity of the GLCM. Higher values indicate greater asymmetry and less uniform distribution of paired intensities, hence more prominent clusters of higher or lower intensity values. Aligns with heterogeneity. |
| original\_glcm\_ClusterShade | GLCM\_Corr | Measures the skewness and uniformity of the GLCM. Higher values indicate greater asymmetry and less uniform distribution of paired intensities, hence more prominent clusters of higher or lower intensity values. Aligns with heterogeneity. |
| original\_glcm\_ClusterTendency | GLCM\_CP | Measures the groupings of voxels with similar gray-level intensities. Higher values indicate more prominent clusters of intensity values. Aligns with heterogeneity. |
| original\_glcm\_Contrast | GLCM\_CS | Measures the local variations in the image. High contrast indicates large differences in intensity between neighboring pixels, suggesting a heterogeneous texture. |
| original\_glcm\_Correlation | GLCM\_CT | Measures the linear dependency of gray levels in the GLCM. Higher values indicate a more linear relationship, suggesting a more uniform and less complex texture. |
| original\_glcm\_DifferenceAverage | GLCM\_DA | Measures the average intensity difference between neighboring pixels. Higher values indicate more heterogeneity. |
| original\_glcm\_DifferenceEntropy | GLCM\_DiEn | Measures the randomness of intensity differences between neighboring pixels. High entropy indicates greater heterogeneity. |
| original\_glcm\_Id | GLCM\_ID | Inverse Difference. Measures the homogeneity of the image. Higher values indicate more homogeneous texture. |
| original\_glcm\_Idm | GLCM\_IDM | Inverse Difference Moment. Measures the homogeneity of the image. Higher values indicate more homogeneous texture. |
| original\_glcm\_Idmn | GLCM\_IDMN | Inverse Difference Moment Normalized. Measures the homogeneity of the image. Higher values indicate more homogeneous texture. |
| original\_glcm\_Idn | GLCM\_IDN | Inverse Difference Normalized. Measures the homogeneity of the image. Higher values indicate more homogeneous texture. |
| original\_glcm\_Imc1 | GLCM\_IMC1 | Informational Measure of Correlation 1. Measures the complexity of the texture. Higher values may indicate more ordered textures, but context is needed. I'll classify as 'Rest (0)'. |
| original\_glcm\_Imc2 | GLCM\_IMC2 | Informational Measure of Correlation 2. Measures the complexity of the texture. Higher values may indicate more ordered textures, but context is needed. I'll classify as 'Rest (0)'. |
| original\_glcm\_InverseVariance | GLCM\_IV | Measures the inverse of the variance of the paired intensities. Higher values indicate less variation and more uniformity. |
| original\_glcm\_JointAverage | GLCM\_JA | Measures the average of the gray level co-occurrence matrix elements. Higher values indicate higher average intensity in paired voxels. |
| original\_glcm\_JointEnergy | GLCM\_JE | Measures the sum of squared elements in the GLCM. Higher values indicate more uniform or less complex texture. |
| original\_glcm\_JointEntropy | GLCM\_JEn | Measures the randomness or disorder of paired voxel intensities. High entropy indicates greater heterogeneity. |
| original\_glcm\_MaximumProbability | GLCM\_MCC | The maximum probability of any pair in the GLCM. A higher value indicates a more uniform texture (dominant pairs). |
| original\_glcm\_MCC | GLCM\_MP | Maximal Correlation Coefficient. Measures the maximum correlation between a gray level and its neighborhood. Higher values indicate more uniform texture. |
| original\_glcm\_SumAverage | GLCM\_Sav | Measures the average of the sum of the gray level intensities in the GLCM. Higher values indicate higher average intensity in paired voxels. |
| original\_glcm\_SumEntropy | GLCM\_SEn | Measures the randomness of the sum of paired voxel intensities. High entropy indicates greater heterogeneity. |
| original\_glcm\_SumSquares | GLCM\_SQ | Measures the variance of the paired intensities. Higher values indicate more heterogeneity. |
| original\_gldm\_DependenceEntropy | GLDM\_DeEn | Measures the randomness of dependence sizes. High entropy indicates greater heterogeneity. |
| original\_gldm\_DependenceNonUniformity | GLDM\_DN | Measures the variability in the size of homogeneous dependencies. Higher values indicate greater heterogeneity. |
| original\_gldm\_DependenceNonUniformityNormalized | GLDM\_DNN | Normalized version of DependenceNonUniformity. Higher values indicate greater heterogeneity. |
| original\_gldm\_DependenceVariance | GLDM\_DV | Measures the variance of the dependence sizes. Higher values indicate greater heterogeneity. |
| original\_gldm\_GrayLevelNonUniformity | GLDM\_GLN | Measures the variability of gray-level intensities in the dependencies. Higher values indicate greater heterogeneity. |
| original\_gldm\_GrayLevelVariance | GLDM\_GLV | Measures the variance of gray-level intensities in the dependencies. Higher values indicate greater heterogeneity. |
| original\_gldm\_HighGrayLevelEmphasis | GLDM\_HGLE | Emphasizes dependencies with high gray-level values. |
| original\_gldm\_LargeDependenceEmphasis | GLDM\_LDE | Measures the proportion of large dependencies. Higher values suggest a texture with large, homogeneous regions. |
| original\_gldm\_LargeDependenceHighGrayLevelEmphasis | GLDM\_LDHGLE | Emphasizes large dependencies with high gray-level values. |
| original\_gldm\_LargeDependenceLowGrayLevelEmphasis | GLDM\_LDLGLE | Emphasizes large dependencies with low gray-level values. |
| original\_gldm\_LowGrayLevelEmphasis | GLDM\_LGLE | Emphasizes dependencies with low gray-level values. |
| original\_gldm\_SmallDependenceEmphasis | GLDM\_SDE | Measures the proportion of small dependencies. Higher values suggest a texture with many small, fragmented regions. |
| original\_gldm\_SmallDependenceHighGrayLevelEmphasis | GLDM\_SDHGLE | Emphasizes small dependencies with high gray-level values. |
| original\_gldm\_SmallDependenceLowGrayLevelEmphasis | GLDM\_SDLGLE | Emphasizes small dependencies with low gray-level values. |
| original\_glrlm\_GrayLevelNonUniformity | GLRLM\_GLN | Measures the variability of gray-level intensities in the runs. Higher values indicate greater heterogeneity. |
| original\_glrlm\_GrayLevelNonUniformityNormalized | GLRLM\_GLNN | Normalized version of GrayLevelNonUniformity. Higher values indicate greater heterogeneity. |
| original\_glrlm\_GrayLevelVariance | GLRLM\_GLV | Measures the variance of gray-level intensities in the runs. Higher values indicate greater heterogeneity. |
| original\_glrlm\_HighGrayLevelRunEmphasis | GLRLM\_HGLRE | Emphasizes short runs with high gray-level values. |
| original\_glrlm\_LongRunEmphasis | GLRLM\_LGLRE | Measures the proportion of long runs. Higher values suggest a texture with continuous regions of similar intensity, implying uniformity/smoothness. |
| original\_glrlm\_LongRunHighGrayLevelEmphasis | GLRLM\_LRE | Emphasizes long runs with high gray-level values. |
| original\_glrlm\_LongRunLowGrayLevelEmphasis | GLRLM\_LRHGLE | Emphasizes long runs with low gray-level values. |
| original\_glrlm\_LowGrayLevelRunEmphasis | GLRLM\_LRLGLE | Emphasizes short runs with low gray-level values. |
| original\_glrlm\_RunEntropy | GLRLM\_REn | Measures the randomness of run lengths. High entropy indicates greater heterogeneity. |
| original\_glrlm\_RunLengthNonUniformity | GLRLM\_RLN | Measures the variability in the length of homogeneous runs. Higher values indicate greater heterogeneity. |
| original\_glrlm\_RunLengthNonUniformityNormalized | GLRLM\_RLNN | Normalized version of RunLengthNonUniformity. Higher values indicate greater heterogeneity. |
| original\_glrlm\_RunPercentage | GLRLM\_RP | Measures the fraction of runs in the image. |
| original\_glrlm\_RunVariance | GLRLM\_RV | Measures the variance of the run lengths. Higher values indicate greater heterogeneity. |
| original\_glrlm\_ShortRunEmphasis | GLRLM\_SRE | Measures the proportion of short runs. Higher values suggest a texture with frequent changes in intensity, implying heterogeneity. |
| original\_glrlm\_ShortRunHighGrayLevelEmphasis | GLRLM\_SRHGLE | Emphasizes short runs with high gray-level values. |
| original\_glrlm\_ShortRunLowGrayLevelEmphasis | GLRLM\_SRLGLE | Emphasizes short runs with low gray-level values. |
| original\_glszm\_GrayLevelNonUniformity | GLSZM\_GLN | Measures the variability of gray-level intensities in the zones. Higher values indicate greater heterogeneity. |
| original\_glszm\_GrayLevelNonUniformityNormalized | GLSZM\_GLNN | Normalized version of GrayLevelNonUniformity. Higher values indicate greater heterogeneity. |
| original\_glszm\_GrayLevelVariance | GLSZM\_GLV | Measures the variance of gray-level intensities in the zones. Higher values indicate greater heterogeneity. |
| original\_glszm\_HighGrayLevelZoneEmphasis | GLSZM\_HGLZE | Emphasizes zones with high gray-level values. |
| original\_glszm\_LargeAreaEmphasis | GLSZM\_LAE | Measures the proportion of large zones in the image. Higher values suggest a texture composed of large, homogeneous regions. |
| original\_glszm\_LargeAreaHighGrayLevelEmphasis | GLSZM\_LAHGLE | Emphasizes large zones with high gray-level values. |
| original\_glszm\_LargeAreaLowGrayLevelEmphasis | GLSZM\_LALGLE | Emphasizes large zones with low gray-level values. |
| original\_glszm\_LowGrayLevelZoneEmphasis | GLSZM\_LGLZE | Emphasizes zones with low gray-level values. |
| original\_glszm\_SizeZoneNonUniformity | GLSZM\_SZN | Measures the variability in the size of homogeneous zones. Higher values indicate greater heterogeneity. |
| original\_glszm\_SizeZoneNonUniformityNormalized | GLSZM\_SZNN | Normalized version of SizeZoneNonUniformity. Higher values indicate greater heterogeneity. |
| original\_glszm\_SmallAreaEmphasis | GLSZM\_SAE | Measures the proportion of small zones in the image. Higher values suggest a texture composed of many small, possibly heterogeneous, regions. |
| original\_glszm\_SmallAreaHighGrayLevelEmphasis | GLSZM\_SAHGLE | Emphasizes small zones with high gray-level values. |
| original\_glszm\_SmallAreaLowGrayLevelEmphasis | GLSZM\_SALGLE | Emphasizes small zones with low gray-level values. |
| original\_glszm\_ZoneEntropy | GLSZM\_ZE | Measures the randomness of zone sizes. High entropy indicates greater heterogeneity. |
| original\_glszm\_ZonePercentage | GLSZM\_ZP | Measures the fraction of the image occupied by homogeneous zones. |
| original\_glszm\_ZoneVariance | GLSZM\_ZV | Measures the variance of the zone sizes. Higher values indicate greater heterogeneity. |
| original\_ngtdm\_Busyness | NGTDM\_B | Measures the variability of gray levels in the neighborhood. Higher busyness implies more heterogeneity or 'busy' texture. |
| original\_ngtdm\_Coarseness | NGTDM\_Coar | Measures the average gray level difference between a voxel and its neighborhood. Higher coarseness implies a less uniform, more granular texture. |
| original\_ngtdm\_Complexity | NGTDM\_Comp | Measures the spatial complexity of the image texture. Higher complexity implies more heterogeneity. |
| original\_ngtdm\_Contrast | NGTDM\_Con | Measures the spatial intensity change. High contrast indicates large differences in intensity between neighboring pixels, suggesting a heterogeneous texture. |
| original\_ngtdm\_Strength | NGTDM\_S | Measures the primitive structures and their strengths. Higher strength implies a more consistent, uniform texture. |
| original\_shape\_Elongation | SF\_E\_3D | Measures the ratio of the minor axis length to the major axis length. Higher values indicate a more elongated shape. |
| original\_shape\_Flatness | SF\_F\_3D | Measures the ratio of the least axis length to the major axis length. Higher values indicate a flatter shape. |
| original\_shape\_LeastAxisLength | SF\_LAL\_3D | The length of the smallest principal axis of the ROI. |
| original\_shape\_MajorAxisLength | SF\_MAL\_3D | The length of the largest principal axis of the ROI. |
| original\_shape\_Maximum2DDiameterColumn | SF\_Max2DD (Column) | The largest pairwise Euclidean distance between any two voxel centers in the 2D plane of the largest column. |
| original\_shape\_Maximum2DDiameterRow | SF\_Max2DD (Row) | The largest pairwise Euclidean distance between any two voxel centers in the 2D plane of the largest row. |
| original\_shape\_Maximum2DDiameterSlice | SF\_Max2DD (Slice) | The largest pairwise Euclidean distance between any two voxel centers in the 2D plane of the largest slice. |
| original\_shape\_Maximum3DDiameter | SF\_Max3DD | The largest pairwise Euclidean distance between any two voxel centers within the ROI. |
| original\_shape\_MeshVolume | SF\_MV\_3D | The volume of the ROI calculated from the mesh. |
| original\_shape\_MinorAxisLength | SF\_MiAL\_2D | The length of the second largest principal axis of the ROI. |
| original\_shape\_Sphericity | SF\_Sp\_2D | A measure of the roundness of the shape relative to a sphere. A value of 1 indicates a perfect sphere. Higher values indicate a more spherical shape. |
| original\_shape\_SurfaceArea | SF\_SA\_3D | The surface area of the ROI. |
| original\_shape\_SurfaceVolumeRatio | SF\_SAVR\_3D | The ratio of the surface area to the volume of the ROI. Lower values indicate a more spherical or compact shape, higher values indicate more irregular or diffuse shapes. |
| original\_shape\_VoxelVolume | SF\_VV\_3D | The volume of the ROI calculated from the number of voxels. |

**Supplemental table S2.**Data-Driven Feature Interpretation Dictionary (DDFID).

|  |  |
| --- | --- |
| Feature name | Semantic meaning |
| original\_firstorder\_10Percentile | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_firstorder\_Median | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_firstorder\_RootMeanSquared | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_firstorder\_TotalEnergy | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glcm\_JointAverage | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_gldm\_DependenceVariance | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_gldm\_GrayLevelNonUniformity | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_gldm\_HighGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_gldm\_LargeDependenceLowGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_LongRunEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_LongRunHighGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_LowGrayLevelRunEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_RunLengthNonUniformityNormalized | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_RunVariance | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_ShortRunEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_ShortRunHighGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glrlm\_ShortRunLowGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_GrayLevelNonUniformityNormalized | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_GrayLevelVariance | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_HighGrayLevelZoneEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_LargeAreaLowGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_LowGrayLevelZoneEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_SmallAreaHighGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_glszm\_SmallAreaLowGrayLevelEmphasis | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_ngtdm\_Strength | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_shape\_Elongation | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_shape\_LeastAxisLength | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_shape\_MajorAxisLength | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_shape\_Maximum2DDiameterSlice | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_shape\_MeshVolume | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_shape\_MinorAxisLength | Heterogenous - Irregular - Spiculated - Diffiuse |
| original\_firstorder\_90Percentile | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_firstorder\_Energy | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_firstorder\_Maximum | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_firstorder\_Mean | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_firstorder\_Minimum | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_glcm\_SumAverage | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_gldm\_DependenceEntropy | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_gldm\_DependenceNonUniformity | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_gldm\_DependenceNonUniformityNormalized | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_gldm\_GrayLevelVariance | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_gldm\_LargeDependenceHighGrayLevelEmphasis | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_glrlm\_HighGrayLevelRunEmphasis | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_glrlm\_RunLengthNonUniformity | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_glrlm\_RunPercentage | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_glszm\_GrayLevelNonUniformity | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_glszm\_LargeAreaHighGrayLevelEmphasis | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_shape\_Maximum2DDiameterColumn | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_shape\_Maximum2DDiameterRow | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_shape\_Maximum3DDiameter | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_shape\_SurfaceArea | Rim Enhancement - Round/Oval - Smooth - Segmental |
| original\_shape\_VoxelVolume | Rim Enhancement - Round/Oval - Smooth - Segmental |

* 1. **Feature selection algorithms (FSAs)**

Feature selection algorithmsemployed in this study spanned statistical, wrapper-based, and embedded strategies to ensure robust dimensionality reduction and effective identification of informative predictors. Statistical techniques included Chi-Square Test, Correlation Coefficient, Mutual Information (MI), Mutual Info Gain Ratio, ANOVA F-Test and P-Value selection, Chi2 P-Value filtering, and Variance Thresholding. Wrapper-based methods involved Recursive Feature Elimination (RFE), Univariate Feature Selection, and Sequential Feature Selection (both forward and backward variants), along with model-driven selection via Logistic Regression (Select From Model LR). Embedded approaches incorporated LASSO, Elastic Net, Embedded Elastic Net, and Stability Selection with LASSO to impose regularization-based constraints, alongside Feature Importance extraction via Random Forest (RF), Extra Trees, and Permutation Importance. Additionally, ReliefF, SelectFDR, SelectFWE, Information Gain, and VIF-based filtering were utilized, supplemented by PCA-inspired techniques leveraging loadings or sparse dictionary representations. This diverse suite allowed comprehensive evaluation of feature relevance under varying statistical assumptions and modeling perspectives.

* 1. **Classification Algorithms (CA)**

The classification pipeline integrated a wide range of algorithms across decision-based, probabilistic, linear, distance-based, ensemble, and neural paradigms to assess generalizability and task adaptability. Core models included Decision Tree, Logistic Regression, Linear Discriminant Analysis (LDA), and Ridge Classifier. Probabilistic classifiers encompassed Gaussian, Bernoulli, and Complement Naive Bayes, as well as variants like GaussianNB with priors. Distance-based techniques included k-Nearest Neighbors (KNN) and Nearest Centroid. Ensemble learners such as Random Forest, Extra Trees (in several configurations), AdaBoost, Gradient Boosting, HistGradientBoosting, and Bagging were applied, along with meta-learners like Stacking and Voting Classifier. Advanced models included MLPClassifier, LightGBM (LGBM), and XGBoost (XGB), supplemented by baseline and experimental benchmarks such as Dummy Classifier, Decision Stump, Gaussian Process Classifier, and SGDClassifier. Hyperparameter optimization was conducted via exhaustive grid search, and classifiers were evaluated through metrics including accuracy, precision, recall, F1-score, AUC, and specificity to maintain consistency across validation and test folds.

1. **Results**