Supplementary Material 1.0

Scatterplot of moderate correlated features and GPR:2%/1 mm

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Supplementary Material 1.1

Grid search optimization for each model usinf the reference dataset.

**Random Forest (RF)**

Grid search method (scoring= “roc\_auc”) was performed in RF to find the best hyperparameters. The number of trees (n\_estimators= 100) and the number of nodes (max\_deepth= 3) were calculated three times shuffling the dataset split. The roc curves (cv=5) for the model with their respective learning curve are displayed below.

Different modifications of the number of trees and nodes were performed and measured.It was concluded that 100 and 3 were de best parameters.

A picture containing timeline

Description automatically generated

Diagram

Description automatically generated with medium confidence

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **max\_depth** | **n\_estimators** | **Accuracy** | **ROC\_AUC\_CV\_5** | **Precision** | **Recall** | **F1** |
| 3 | 100 | 0.73 | 0.79+-0.03 | 0.71 | 0.59 | 0.65 |

Graphical user interface, chart

Description automatically generatedChart, line chart

Description automatically generatedChart

Description automatically generated

**XG-Boost**

To achieve up to 95% maximum prediction accuracy, a simple XGBoost identified 72 features as the most important. The hyperparameters resulting from Grid search for XGBoost with all features were learning\_rate=0.01, max\_depth=2, n\_estimators=170.

A picture containing chart

Description automatically generated

Diagram

Description automatically generated

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **max\_depth** | **n\_estimators** | **Learning Rate** | **Accuracy** | **ROC\_AUC\_CV\_5** | **Precision** | **Recall** | **F1** |
| 2 | 170 | 0.01 | 0.80 | 0.87+-0.08 | 0.78 | 0.84 | 0.81 |

Graphical user interface

Description automatically generatedChart, line chart

Description automatically generatedChart, treemap chart

Description automatically generated

**Neural Network**

A simple NN was trained to predict GPRs, changing the number of layers and nodes intuitively keeping the simplest architecture. Using all features the model performance with CV= 5 is shown below in terms of AUC and confusion matrix.

Model: "sequential\_157"

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

layer\_1 (Dense) (None, 210) 12474

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

layer\_2 (Dense) (None, 60) 2211

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

layer\_3 (Dense) (None, 1) 34

=================================================================

Total params: 14,719

Trainable params: 14,719

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Chart, line chart

Description automatically generated Chart, line chart

Description automatically generatedChart

Description automatically generated

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Accuracy** | **ROC\_AUC\_CV\_5** | **Precision** | **Recall** | **F1** |
| 0.72 | 0.81+-0.04 | 0.66 | 0.66 | 0.66 |

Supplementary Material 1.3

Classificatiom performance metrics for each heterogeneity

Random Forest

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Heterogeneity** | **Metric** | **100% - 0%** | | **80% - 20%** | | **60% - 40%** | | **40% - 60%** | | **20% - 80%** | | **0% - 100%** | |
| **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** |
| Dose per fraction | Sensitivity | 0.75 | 0.04 | 0.70 | 0.16 | 0.84 | 0.03 | 0.55 | 0.05 | 0.74 | 0.03 | 0.84 | 0.02 |
| Specificity | 0.75 | 0.01 | 0.73 | 0.03 | 0.79 | 0.06 | 0.78 | 0.00 | 0.67 | 0.04 | 0.58 | 0.04 |
| Precision | 0.77 | 0.01 | 0.65 | 0.05 | 0.61 | 0.03 | 0.68 | 0.02 | 0.69 | 0.03 | 0.73 | 0.02 |
| F1 | 0.76 | 0.03 | 0.70 | 0.09 | 0.70 | 0.02 | 0.61 | 0.04 | 0.71 | 0.02 | 0.78 | 0.01 |
| AUC | 0.82 | 0.04 | 0.70 | 0.03 | 0.62 | 0.12 | 0.71 | 0.11 | 0.77 | 0.09 | 0.88 | 0.07 |
| Number of arcs | Sensitivity | 0.80 | 0.01 | 0.72 | 0.01 | 0.78 | 0.01 | 0.68 | 0.01 | 0.71 | 0.03 | 0.72 | 0.07 |
| Specificity | 0.81 | 0.03 | 0.81 | 0.01 | 0.63 | 0.01 | 0.67 | 0.01 | 0.75 | 0.03 | 0.80 | 0.02 |
| Precision | 0.83 | 0.02 | 0.80 | 0.01 | 0.73 | 0.01 | 0.63 | 0.03 | 0.73 | 0.02 | 0.72 | 0.02 |
| F1 | 0.81 | 0.01 | 0.75 | 0.01 | 0.75 | 0.00 | 0.65 | 0.02 | 0.72 | 0.02 | 0.72 | 0.04 |
| AUC | 0.88 | 0.06 | 0.81 | 0.07 | 0.80 | 0.04 | 0.75 | 0.13 | 0.79 | 0.08 | 0.86 | 0.07 |
| Treatment Unit | Sensitivity | 0.80 | 0.10 | 0.84 | 0.08 | 0.81 | 0.11 | 0.80 | 0.10 | 0.78 | 0.12 | 0.81 | 0.04 |
| Specificity | 0.65 | 0.02 | 0.60 | 0.08 | 0.56 | 0.13 | 0.48 | 0.06 | 0.65 | 0.13 | 0.83 | 0.05 |
| Precision | 0.75 | 0.07 | 0.74 | 0.05 | 0.71 | 0.02 | 0.71 | 0.08 | 0.67 | 0.10 | 0.81 | 0.09 |
| F1 | 0.77 | 0.09 | 0.78 | 0.06 | 0.75 | 0.04 | 0.75 | 0.08 | 0.69 | 0.08 | 0.81 | 0.06 |
| AUC | 0.83 | 0.09 | 0.68 | 0.13 | 0.78 | 0.10 | 0.78 | 0.13 | 0.76 | 0.11 | 0.91 | 0.09 |
| Anatomic Region | Sensitivity | 0.77 | 0.04 | 0.67 | 0.09 | 0.61 | 0.11 | 0.44 | 0.03 | 0.46 | 0.10 | 0.48 | 0.06 |
| Specificity | 0.65 | 0.06 | 0.66 | 0.04 | 0.62 | 0.19 | 0.83 | 0.11 | 0.87 | 0.05 | 0.93 | 0.05 |
| Precision | 0.73 | 0.03 | 0.66 | 0.06 | 0.65 | 0.06 | 0.68 | 0.16 | 0.67 | 0.10 | 0.79 | 0.09 |
| F1 | 0.75 | 0.03 | 0.66 | 0.08 | 0.62 | 0.06 | 0.53 | 0.03 | 0.52 | 0.10 | 0.64 | 0.03 |
| AUC | 0.84 | 0.13 | 0.78 | 0.07 | 0.73 | 0.11 | 0.72 | 0.06 | 0.78 | 0.11 | 0.82 | 0.05 |

XG-Boost

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Heterogeneity** | **Metric** | **100% - 0%** | | **80% - 20%** | | **60% - 40%** | | **40% - 60%** | | **20% - 80%** | | **0% - 100%** | |
| **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** |
| Dose per fraction | Sensitivity | 0.69 | 0.02 | 0.72 | 0.09 | 0.80 | 0.04 | 0.56 | 0.06 | 0.73 | 0.03 | 0.84 | 0.02 |
| Specificity | 0.64 | 0.06 | 0.65 | 0.08 | 0.44 | 0.10 | 0.71 | 0.06 | 0.67 | 0.16 | 0.58 | 0.04 |
| Precision | 0.73 | 0.02 | 0.67 | 0.06 | 0.66 | 0.02 | 0.64 | 0.02 | 0.70 | 0.08 | 0.73 | 0.02 |
| F1 | 0.70 | 0.02 | 0.70 | 0.07 | 0.72 | 0.01 | 0.60 | 0.03 | 0.72 | 0.03 | 0.78 | 0.01 |
| AUC | 0.87 | 0.09 | 0.75 | 0.1 | 0.66 | 0.1 | 0.70 | 0.05 | 0.78 | 0.09 | 0.89 | 0.07 |
| Number of arcs | Sensitivity | 0.82 | 0.02 | 0.71 | 0.04 | 0.77 | 0.02 | 0.56 | 0.07 | 0.61 | 0.13 | 0.75 | 0.15 |
| Specificity | 0.75 | 0.10 | 0.84 | 0.05 | 0.66 | 0.03 | 0.70 | 0.01 | 0.68 | 0.04 | 0.89 | 0.06 |
| Precision | 0.81 | 0.05 | 0.80 | 0.01 | 0.75 | 0.02 | 0.62 | 0.04 | 0.62 | 0.08 | 0.83 | 0.01 |
| F1 | 0.80 | 0.01 | 0.75 | 0.02 | 0.77 | 0.01 | 0.59 | 0.06 | 0.62 | 0.11 | 0.78 | 0.07 |
| AUC | 0.85 | 0.07 | 0.81 | 0.04 | 0.80 | 0.07 | 0.69 | 0.05 | 0.77 | 0.12 | 0.87 | 0.09 |
| Treatment Unit | Sensitivity | 0.80 | 0.03 | 0.74 | 0.06 | 0.76 | 0.09 | 0.78 | 0.12 | 0.68 | 0.13 | 0.77 | 0.07 |
| Specificity | 0.58 | 0.11 | 0.61 | 0.14 | 0.64 | 0.16 | 0.55 | 0.04 | 0.65 | 0.12 | 0.77 | 0.06 |
| Precision | 0.70 | 0.07 | 0.74 | 0.04 | 0.75 | 0.09 | 0.72 | 0.10 | 0.65 | 0.11 | 0.75 | 0.08 |
| F1 | 0.75 | 0.04 | 0.73 | 0.05 | 0.75 | 0.08 | 0.75 | 0.11 | 0.66 | 0.12 | 0.76 | 0.07 |
| AUC | 0.80 | 0.04 | 0.68 | 0.12 | 0.76 | 0.12 | 0.71 | 0.17 | 0.84 | 0.17 | 0.86 | 0.06 |
| Anatomic Region | Sensitivity | 0.81 | 0.02 | 0.65 | 0.11 | 0.60 | 0.05 | 0.46 | 0.14 | 0.63 | 0.24 | 0.65 | 0.10 |
| Specificity | 0.63 | 0.10 | 0.68 | 0.02 | 0.67 | 0.08 | 0.78 | 0.12 | 0.73 | 0.18 | 0.91 | 0.02 |
| Precision | 0.74 | 0.05 | 0.66 | 0.05 | 0.66 | 0.01 | 0.60 | 0.10 | 0.67 | 0.13 | 0.78 | 0.05 |
| F1 | 0.77 | 0.02 | 0.66 | 0.08 | 0.62 | 0.02 | 0.50 | 0.14 | 0.59 | 0.13 | 0.61 | 0.08 |
| AUC | 0.84 | 0.09 | 0.78 | 0.07 | 0.67 | 0.1 | 0.75 | 0.04 | 0.76 | 0.18 | 0.88 | 0.08 |

Neural Network

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Heterogeneity** | **Metric** | **100% - 0%** | | **80% - 20%** | | **60% - 40%** | | **40% - 60%** | | **20% - 80%** | | **0% - 100%** | |
| **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** | **mv** | **sd** |
| Dose per fraction | Sensitivity | 0.93 | 0.06 | 0.91 | 0.03 | 0.91 | 0.06 | 0.80 | 0.05 | 0.78 | 0.11 | 0.92 | 0.01 |
| Specificity | 0.90 | 0.02 | 0.92 | 0.04 | 0.87 | 0.06 | 0.86 | 0.10 | 0.91 | 0.04 | 0.89 | 0.06 |
| Precision | 0.90 | 0.02 | 0.93 | 0.03 | 0.91 | 0.03 | 0.84 | 0.10 | 0.90 | 0.05 | 0.88 | 0.06 |
| F1 | 0.90 | 0.02 | 0.92 | 0.01 | 0.91 | 0.04 | 0.81 | 0.04 | 0.83 | 0.08 | 0.90 | 0.03 |
| AUC | 0.94 | 0.03 | 0.89 | 0.04 | 0.87 | 0.03 | 0.88 | 0.05 | 0.93 | 0.02 | 0.92 | 0.03 |
| Number of arcs | Sensitivity | 0.87 | 0.07 | 0.80 | 0.08 | 0.88 | 0.01 | 0.90 | 0.06 | 0.90 | 0.05 | 0.94 | 0.06 |
| Specificity | 0.80 | 0.03 | 0.84 | 0.05 | 0.67 | 0.04 | 0.68 | 0.01 | 0.70 | 0.04 | 0.84 | 0.02 |
| Precision | 0.83 | 0.04 | 0.82 | 0.03 | 0.89 | 0.05 | 0.87 | 0.02 | 0.89 | 0.02 | 0.92 | 0.03 |
| F1 | 0.84 | 0.01 | 0.80 | 0.01 | 0.90 | 0.03 | 0.88 | 0.01 | 0.91 | 0.01 | 0.91 | 0.01 |
| AUC | 0.90 | 0.05 | 0.84 | 0.01 | 0.90 | 0.02 | 0.83 | 0.11 | 0.88 | 0.04 | 0.96 | 0.04 |
| Treatment Unit | Sensitivity | 0.98 | 0.04 | 0.98 | 0.02 | 0.89 | 0.01 | 0.95 | 0.05 | 0.86 | 0.10 | 0.93 | 0.08 |
| Specificity | 0.92 | 0.08 | 0.90 | 0.10 | 0.91 | 0.02 | 0.86 | 0.07 | 0.89 | 0.02 | 0.95 | 0.08 |
| Precision | 0.94 | 0.06 | 0.92 | 0.07 | 0.93 | 0.02 | 0.91 | 0.06 | 0.88 | 0.01 | 0.96 | 0.07 |
| F1 | 0.96 | 0.02 | 0.95 | 0.05 | 0.90 | 0.01 | 0.92 | 0.04 | 0.87 | 0.07 | 0.95 | 0.05 |
| AUC | 0.97 | 0.03 | 0.96 | 0.02 | 0.89 | 0.07 | 0.92 | 0.03 | 0.94 | 0.01 | 0.98 | 0.01 |
| Anatomic Region | Sensitivity | 0.90 | 0.10 | 0.92 | 0.01 | 0.88 | 0.05 | 0.91 | 0.04 | 0.83 | 0.13 | 0.93 | 0.07 |
| Specificity | 0.79 | 0.07 | 0.87 | 0.05 | 0.93 | 0.06 | 0.92 | 0.07 | 0.91 | 0.04 | 0.98 | 0.03 |
| Precision | 0.83 | 0.06 | 0.88 | 0.05 | 0.92 | 0.08 | 0.88 | 0.10 | 0.86 | 0.06 | 0.97 | 0.04 |
| F1 | 0.86 | 0.03 | 0.89 | 0.03 | 0.90 | 0.06 | 0.89 | 0.07 | 0.85 | 0.09 | 0.95 | 0.03 |
| AUC | 0.94 | 0.04 | 0.90 | 0.02 | 0.90 | 0.05 | 0.93 | 0.03 | 0.95 | 0.04 | 0.95 | 0.04 |