For Subgroup 1 : ['ENAC10', 'Kmer1', 'DAC7', 'ENAC5', 'CKSNAP9']

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Sensitivity | Specificity | MCC |
| XGBoost | |  |  | | --- | --- | | Train | 0.845069 | | Test | 0.728330 | | |  |  | | --- | --- | | Train | 0.85359 | | Test | 0.733003 | | |  |  | | --- | --- | | Train | 0.836547 | | Test | 0.723806 | | |  |  | | --- | --- | | Train | 0.69024 | | Test | 0.45709 | |
| AdaBoost | |  |  | | --- | --- | | Train | 0.756094 | | Test | 0.703682 | | |  |  | | --- | --- | | Train | 0.757843 | | Test | 0.7079571 | | |  |  | | --- | --- | | Train | 0.754298 | | Test | 0.699160 | | |  |  | | --- | --- | | Train | 0.512238 | | Test | 0.407521 | |
| Ran. Forest | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.692313 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.697405 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.687639 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.385320 | |
| Log. Regression | |  |  | | --- | --- | | Train | 0.798009 | | Test | 0.762187 | | |  |  | | --- | --- | | Train | 0.799270 | | Test | 0.759793 | | |  |  | | --- | --- | | Train | 0.796722 | | Test | 0.764313 | | |  |  | | --- | --- | | Train | 0.596018 | | Test | 0.524334 | |
| SVC | |  |  | | --- | --- | | Train | 0.844190 | | Test | 0.750272 | | |  |  | | --- | --- | | Train | 0.857226 | | Test | 0.758594 | | |  |  | | --- | --- | | Train | 0.831095 | | Test | 0.74239 | | |  |  | | --- | --- | | Train | 0.68860 | | Test | 0.501126 | |

For Subgroup 2: ['ENAC5', 'DAC7', 'CKSNAP9', 'CKSNAP5', 'binary']

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Sensitivity | Specificity | MCC |
| XGBoost | |  |  | | --- | --- | | Train | 0.866941 | | Test | 0.771122 | | |  |  | | --- | --- | | Train | 0.873760 | | Test | 0.771245 | | |  |  | | --- | --- | | Train | 0.860092 | | Test | 0.771215 | | |  |  | | --- | --- | | Train | 0.73392 | | Test | 0.542304 | |
| AdaBoost | |  |  | | --- | --- | | Train | 0.787175 | | Test | 0.744587 | | |  |  | | --- | --- | | Train | 0.781007 | | Test | 0.73113 | | |  |  | | --- | --- | | Train | 0.793336 | | Test | 0.75776 | | |  |  | | --- | --- | | Train | 0.574402 | | Test | 0.48910 | |
| Ran. Forest | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.718581 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.726920 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.709986 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.43699 | |
| Log. Regression | |  |  | | --- | --- | | Train | 0.809994 | | Test | 0.77221 | | |  |  | | --- | --- | | Train | 0.8075766 | | Test | 0.769553 | | |  |  | | --- | --- | | Train | 0.812434 | | Test | 0.775371 | | |  |  | | --- | --- | | Train | 0.620018 | | Test | 0.544844 | |
| SVC | |  |  | | --- | --- | | Train | 0.91176 | | Test | 0.766255 | | |  |  | | --- | --- | | Train | 0.920644 | | Test | 0.767768 | | |  |  | | --- | --- | | Train | 0.902880 | | Test | 0.7647904 | | |  |  | | --- | --- | | Train | 0.823695 | | Test | 0.532638 | |

For Subgroup 3: ['EIIP', 'Kmer1', 'CKSNAP1', 'Kmer5', 'CKSNAP9']

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Sensitivity | Specificity | MCC |
| XGBoost | |  |  | | --- | --- | | Train | 0.854685 | | Test | 0.7667873 | | |  |  | | --- | --- | | Train | 0.862955 | | Test | 0.762372 | | |  |  | | --- | --- | | Train | 0.8464279 | | Test | 0.771319 | | |  |  | | --- | --- | | Train | 0.7095011 | | Test | 0.533851 | |
| AdaBoost | |  |  | | --- | --- | | Train | 0.786294 | | Test | 0.742141 | | |  |  | | --- | --- | | Train | 0.786548 | | Test | 0.74419 | | |  |  | | --- | --- | | Train | 0.786013 | | Test | 0.740584 | | |  |  | | --- | --- | | Train | 0.572587 | | Test | 0.484643 | |
| Ran. Forest | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.69745 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.699872 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.695483 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.395156 | |
| Log. Regression | |  |  | | --- | --- | | Train | 0.7344932 | | Test | 0.709098 | | |  |  | | --- | --- | | Train | 0.7527609 | | Test | 0.728694 | | |  |  | | --- | --- | | Train | 0.716083 | | Test | 0.690606 | | |  |  | | --- | --- | | Train | 0.469228 | | Test | 0.419527 | |
| SVC | |  |  | | --- | --- | | Train | 0.76963 | | Test | 0.721829 | | |  |  | | --- | --- | | Train | 0.788454 | | Test | 0.746419 | | |  |  | | --- | --- | | Train | 0.750774 | | Test | 0.697456 | | |  |  | | --- | --- | | Train | 0.539639 | | Test | 0.444271 | |

For Subgroup 4: ['CKSNAP1', 'ENAC10', 'CKSNAP9', 'CKSNAP5', 'ENAC10']

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Sensitivity | Specificity | MCC |
| XGBoost | |  |  | | --- | --- | | Train | 0.823537 | | Test | 0.702327 | | |  |  | | --- | --- | | Train | 0.82503 | | Test | 0.699473 | | |  |  | | --- | --- | | Train | 0.822037 | | Test | 0.705628 | | |  |  | | --- | --- | | Train | 0.647074 | | Test | 0.405261 | |
| AdaBoost | |  |  | | --- | --- | | Train | 0.736389 | | Test | 0.677410 | | |  |  | | --- | --- | | Train | 0.73822 | | Test | 0.68316 | | |  |  | | --- | --- | | Train | 0.73455 | | Test | 0.671508 | | |  |  | | --- | --- | | Train | 0.472789 | | Test | 0.354770 | |
| Ran. Forest | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.6860747 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.693479 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.678755 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.3725144 | |
| Log. Regression | |  |  | | --- | --- | | Train | 0.7829085 | | Test | 0.752167 | | |  |  | | --- | --- | | Train | 0.785991 | | Test | 0.755532 | | |  |  | | --- | --- | | Train | 0.779772 | | Test | 0.748612 | | |  |  | | --- | --- | | Train | 0.565806 | | Test | 0.504621 | |
| SVC | |  |  | | --- | --- | | Train | 0.797602 | | Test | 0.730495 | | |  |  | | --- | --- | | Train | 0.802009 | | Test | 0.735798 | | |  |  | | --- | --- | | Train | 0.79319 | | Test | 0.725402 | | |  |  | | --- | --- | | Train | 0.595235 | | Test | 0.461292 | |

For Subgroup 5: ['NCP', 'Kmer5', 'CKSNAP9', 'ENAC10', 'Kmer3']

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Sensitivity | Specificity | MCC |
| XGBoost | |  |  | | --- | --- | | Train | 0.869786 | | Test | 0.776266 | | |  |  | | --- | --- | | Train | 0.879160 | | Test | 0.773545 | | |  |  | | --- | --- | | Train | 0.860360 | | Test | 0.778804 | | |  |  | | --- | --- | | Train | 0.739720 | | Test | 0.552451 | |
| AdaBoost | |  |  | | --- | --- | | Train | 0.7910351 | | Test | 0.738348 | | |  |  | | --- | --- | | Train | 0.793142 | | Test | 0.740587 | | |  |  | | --- | --- | | Train | 0.788876 | | Test | 0.736927 | | |  |  | | --- | --- | | Train | 0.582049 | | Test | 0.477255 | |
| Ran. Forest | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.702866 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.703994 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.702384 | | |  |  | | --- | --- | | Train | 1.0 | | Test | 0.406705 | |
| Log. Regression | |  |  | | --- | --- | | Train | 0.80660 | | Test | 0.77193 | | |  |  | | --- | --- | | Train | 0.800792 | | Test | 0.77108 | | |  |  | | --- | --- | | Train | 0.81242 | | Test | 0.77298 | | |  |  | | --- | --- | | Train | 0.613254 | | Test | 0.544357 | |
| SVC | |  |  | | --- | --- | | Train | 0.85400 | | Test | 0.760834 | | |  |  | | --- | --- | | Train | 0.857512 | | Test | 0.761010 | | |  |  | | --- | --- | | Train | 0.850471 | | Test | 0.760596 | | |  |  | | --- | --- | | Train | 0.70805 | | Test | 0.521929 | |

Independent Test:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Accuracy | Sensitivity | Specificity | MCC |
| XGBoost | 0.79166 | 0.803030 | 0.780303 | 0.58348 |
| AdaBoost | 0.753787 | 0.734848 | 0.7727 | 0.50794 |
| Ran. Forest | 0.696969 | 0.7045454 | 0.68939 | 0.39398 |
| Log. Regression | 0.79924 | 0.772727 | 0.82575 | 0.59932 |
| SVC | 0.80303 | 0.8030 | 0.80303 | 0.60606 |