**Model Training with all features ( 13 features )**

Gradient Boosting Machine (GBM)

[1] "Accuracy: 0.890603085553997"

[1] "Precision Large: 0.960552268244576"

[1] "Precision Medium: 0.807387862796834"

[1] "Precision Small: 0.922196796338673"

[1] "Recall Large: 0.855887521968366"

[1] "Recall Medium: 0.9"

[1] "Recall Small: 0.90561797752809"

[1] "F1 Score Large: 0.905204460966543"

[1] "F1 Score Medium: 0.851182197496523"

[1] "F1 Score Small: 0.913832199546485"

[1] "Specificity Large: 0.987261146496815"

[1] "Specificity Medium: 0.899931459904044"

[1] "Specificity Small: 0.945556445156125"

Random Forest (RF)

[1] "Accuracy: 0.977559607293128"

[1] "Precision Large: 0.991087344028521"

[1] "Precision Medium: 0.94413407821229"

[1] "Precision Small: 0.996519721577726"

[1] "Recall Large: 0.977152899824253"

[1] "Recall Medium: 0.994117647058824"

[1] "Recall Small: 0.965168539325843"

[1] "F1 Score Large: 0.984070796460177"

[1] "F1 Score Medium: 0.968481375358166"

[1] "F1 Score Small: 0.980593607305936"

[1] "Specificity Large: 0.996815286624204"

[1] "Specificity Medium: 0.972583961617546"

[1] "Specificity Small: 0.99759807846277"

Multinomial Logistic Regression (MLR)

[1] "Accuracy: 0.809724170172978"

[1] "Precision Large: 0.817184643510055"

[1] "Precision Medium: 0.705722070844687"

[1] "Precision Small: 0.893939393939394"

[1] "Recall Large: 0.785588752196837"

[1] "Recall Medium: 0.761764705882353"

[1] "Recall Small: 0.861797752808989"

[1] "F1 Score Large: 0.801075268817204"

[1] "F1 Score Medium: 0.732673267326733"

[1] "F1 Score Small: 0.877574370709382"

[1] "Specificity Large: 0.936305732484076"

[1] "Specificity Medium: 0.85195339273475"

[1] "Specificity Small: 0.927141713370697"

Naive Bayes (NB)

[1] "Accuracy: 0.711547452080411"

[1] "Precision Large: 0.640416047548291"

[1] "Precision Medium: 0.604060913705584"

[1] "Precision Small: 0.838857142857143"

[1] "Recall Large: 0.757469244288225"

[1] "Recall Medium: 0.525"

[1] "Recall Small: 0.824719101123595"

[1] "F1 Score Large: 0.694041867954911"

[1] "F1 Score Medium: 0.561762391817467"

[1] "F1 Score Small: 0.831728045325779"

[1] "Specificity Large: 0.845859872611465"

[1] "Specificity Medium: 0.839616175462646"

[1] "Specificity Small: 0.8871096877502"

Support Vector Machine (SVM)

[1] "Accuracy: 0.814866760168303"

[1] "Precision Large: 0.857142857142857"

[1] "Precision Medium: 0.700261780104712"

[1] "Precision Small: 0.891203703703704"

[1] "Recall Large: 0.769771528998243"

[1] "Recall Medium: 0.786764705882353"

[1] "Recall Small: 0.865168539325843"

[1] "F1 Score Large: 0.811111111111111"

[1] "F1 Score Medium: 0.740997229916897"

[1] "F1 Score Small: 0.877993158494869"

[1] "Specificity Large: 0.953503184713376"

[1] "Specificity Medium: 0.843043180260452"

[1] "Specificity Small: 0.924739791833467"

Decision Tree (DT)

[1] "Accuracy: 0.770920991117345"

[1] "Precision Large: 0.846827133479212"

[1] "Precision Medium: 0.609409190371991"

[1] "Precision Small: 0.91796875"

[1] "Recall Large: 0.680140597539543"

[1] "Recall Medium: 0.819117647058824"

[1] "Recall Small: 0.792134831460674"

[1] "F1 Score Large: 0.754385964912281"

[1] "F1 Score Medium: 0.698870765370138"

[1] "F1 Score Small: 0.850422195416164"

[1] "Specificity Large: 0.955414012738854"

[1] "Specificity Medium: 0.7553118574366"

[1] "Specificity Small: 0.949559647718175"

Majority Ensembling (Voting Based):

[1] “Accuracy: 0.8503974”

[1] "Precision Large: 0.887429643527205"

[1] "Precision Medium: 0.762845849802372"

[1] "Precision Small: 0.92969696969697"

[1] "Recall Large: 0.83128295254833"

[1] "Recall Medium: 0.851470588235294"

[1] "Recall Small: 0.861797752808989"

[1] "F1 Score Large: 0.858439201451906"

[1] "F1 Score Medium: 0.804725503822099"

[1] "F1 Score Small: 0.894460641399417"

[1] "Specificity Large: 0.961783439490446"

[1] "Specificity Medium: 0.861549006168609"

[1] "Specificity Small: 0.953562850280224"

**Model Training after feature selection ( 10 features )**

Gradient Boosting Machine (GBM)

[1] "Accuracy: 0.886395511921459"

[1] "Precision Large: 0.968253968253968"

[1] "Precision Medium: 0.802110817941952"

[1] "Precision Small: 0.912200684150513"

[1] "Recall Large: 0.857644991212654"

[1] "Recall Medium: 0.894117647058824"

[1] "Recall Small: 0.898876404494382"

[1] "F1 Score Large: 0.909599254426841"

[1] "F1 Score Medium: 0.845618915159944"

[1] "F1 Score Small: 0.905489530277306"

[1] "Specificity Large: 0.989808917197452"

[1] "Specificity Medium: 0.897189856065798"

[1] "Specificity Small: 0.938350680544436"

[1] "AUC Large : 0.984272329374363"

[1] "AUC Medium : 0.962670342297303"

[1] "AUC SMALL : 0.978828456023246"

Random Forest (RF)

[1] "Accuracy: 0.976157082748948"

[1] "Precision Large: 0.991071428571429"

[1] "Precision Medium: 0.939058171745152"

[1] "Precision Small: 0.997666277712952"

[1] "Recall Large: 0.975395430579965"

[1] "Recall Medium: 0.997058823529412"

[1] "Recall Small: 0.960674157303371"

[1] "F1 Score Large: 0.983170947741364"

[1] "F1 Score Medium: 0.967189728958631"

[1] "F1 Score Small: 0.978820835718374"

[1] "Specificity Large: 0.996815286624204"

[1] "Specificity Medium: 0.969842357779301"

[1] "Specificity Small: 0.99839871897518"

[1] "AUC Large : 0.999069212944824"

[1] "AUC Medium : 0.998802060234649"

[1] "AUC SMALL : 0.998666798607425"

Multinomial Logistic Regression (MLR)

[1] "Accuracy: 0.80551659654044"

[1] "Precision Large: 0.813528336380256"

[1] "Precision Medium: 0.701369863013699"

[1] "Precision Small: 0.888631090487239"

[1] "Recall Large: 0.78207381370826"

[1] "Recall Medium: 0.752941176470588"

[1] "Recall Small: 0.860674157303371"

[1] "F1 Score Large: 0.797491039426523"

[1] "F1 Score Medium: 0.726241134751773"

[1] "F1 Score Small: 0.874429223744292"

[1] "Specificity Large: 0.935031847133758"

[1] "Specificity Medium: 0.850582590815627"

[1] "Specificity Small: 0.923138510808647"

[1] "AUC Large : 0.919745222929936"

[1] "AUC Medium : 0.838212111438132"

[1] "AUC SMALL : 0.945409810994863"

Naive Bayes (NB)

[1] "Accuracy: 0.683496961196821"

[1] "Precision Large: 0.558638083228247"

[1] "Precision Medium: 0.605150214592275"

[1] "Precision Small: 0.8375"

[1] "Recall Large: 0.778558875219684"

[1] "Recall Medium: 0.414705882352941"

[1] "Recall Small: 0.828089887640449"

[1] "F1 Score Large: 0.650513950073421"

[1] "F1 Score Medium: 0.492146596858639"

[1] "F1 Score Small: 0.832768361581921"

[1] "Specificity Large: 0.777070063694268"

[1] "Specificity Medium: 0.873886223440713"

[1] "Specificity Small: 0.88550840672538"

[1] "AUC Large : 0.876830510561607"

[1] "AUC Medium : 0.80175684393017"

[1] "AUC SMALL : 0.895982403900649"

Support Vector Machine (SVM)

[1] "Accuracy: 0.808789153810192"

[1] "Precision Large: 0.823747680890538"

[1] "Precision Medium: 0.700542005420054"

[1] "Precision Small: 0.892111368909513"

[1] "Recall Large: 0.780316344463972"

[1] "Recall Medium: 0.760294117647059"

[1] "Recall Small: 0.864044943820225"

[1] "F1 Score Large: 0.8014440433213"

[1] "F1 Score Medium: 0.729196050775741"

[1] "F1 Score Small: 0.877853881278539"

[1] "Specificity Large: 0.939490445859873"

[1] "Specificity Medium: 0.848526387936943"

[1] "Specificity Small: 0.925540432345877"

[1] "AUC Large : 0.912659375594685"

[1] "AUC Medium : 0.811130709994759"

[1] "AUC SMALL : 0.939687480321336"

Decision Tree (DT)

[1] "Accuracy: 0.770920991117345"

[1] "Precision Large: 0.846827133479212"

[1] "Precision Medium: 0.609409190371991"

[1] "Precision Small: 0.91796875"

[1] "Recall Large: 0.680140597539543"

[1] "Recall Medium: 0.819117647058824"

[1] "Recall Small: 0.792134831460674"

[1] "F1 Score Large: 0.754385964912281"

[1] "F1 Score Medium: 0.698870765370138"

[1] "F1 Score Small: 0.850422195416164"

[1] "Specificity Large: 0.955414012738854"

[1] "Specificity Medium: 0.7553118574366"

[1] "Specificity Small: 0.949559647718175"

[1] "AUC Large : 0.895069011451535"

[1] "AUC Medium : 0.796545276781035"

[1] "AUC SMALL : 0.903325356914745"

Majority Based Ensembling (Voting Based):

[1] “Accuracy: 0.8401122"

[1] "Precision Large: 0.840070298769772"

[1] "Precision Medium: 0.741633199464525"

[1] "Precision Small: 0.926150121065375"

[1] "Recall Large: 0.840070298769772"

[1] "Recall Medium: 0.814705882352941"

[1] "Recall Small: 0.859550561797753"

[1] "F1 Score Large: 0.840070298769772"

[1] "F1 Score Medium: 0.77645409950946"

[1] "F1 Score Small: 0.891608391608392"

[1] "Specificity Large: 0.94203821656051"

[1] "Specificity Medium: 0.869773817683345"

[1] "Specificity Small: 0.951160928742994"

**Model Training after feature selection ( 9 features )**

Gradient Boosting Machine (GBM)

[1] "Accuracy: 0.887798036465638"

[1] "Precision Large: 0.966202783300199"

[1] "Precision Medium: 0.805039787798408"

[1] "Precision Small: 0.913832199546485"

[1] "Recall Large: 0.854130052724077"

[1] "Recall Medium: 0.892647058823529"

[1] "Recall Small: 0.90561797752809"

[1] "F1 Score Large: 0.906716417910448"

[1] "F1 Score Medium: 0.846582984658299"

[1] "F1 Score Small: 0.909706546275395"

[1] "Specificity Large: 0.989171974522293"

[1] "Specificity Medium: 0.899246058944483"

[1] "Specificity Small: 0.939151321056846"

[1] "AUC Large : 0.984730166903608"

[1] "AUC Medium : 0.959571422811757"

[1] "AUC SMALL : 0.977896924280998"

Random Forest (RF)

[1] "Accuracy: 0.976624590930341"

[1] "Precision Large: 0.989304812834225"

[1] "Precision Medium: 0.942896935933148"

[1] "Precision Small: 0.996511627906977"

[1] "Recall Large: 0.975395430579965"

[1] "Recall Medium: 0.995588235294118"

[1] "Recall Small: 0.962921348314607"

[1] "F1 Score Large: 0.982300884955752"

[1] "F1 Score Medium: 0.968526466380544"

[1] "F1 Score Small: 0.979428571428571"

[1] "Specificity Large: 0.996178343949045"

[1] "Specificity Medium: 0.971898560657985"

[1] "Specificity Small: 0.99759807846277"

[1] "AUC Large : 0.999257273348035"

[1] "AUC Medium : 0.998682115066726"

[1] "AUC SMALL : 0.998551200510971"

Multinomial Logistic Regression (MLR)

[1] "Accuracy: 0.805984104721833"

[1] "Precision Large: 0.81651376146789"

[1] "Precision Medium: 0.700819672131147"

[1] "Precision Small: 0.888631090487239"

[1] "Recall Large: 0.78207381370826"

[1] "Recall Medium: 0.754411764705882"

[1] "Recall Small: 0.860674157303371"

[1] "F1 Score Large: 0.798922800718133"

[1] "F1 Score Medium: 0.726628895184136"

[1] "F1 Score Small: 0.874429223744292"

[1] "Specificity Large: 0.936305732484076"

[1] "Specificity Medium: 0.849897189856066"

[1] "Specificity Small: 0.923138510808647"

[1] "AUC Large : 0.922375829760559"

[1] "AUC Medium : 0.844540982945611"

[1] "AUC SMALL : 0.946586482669281"

Naive Bayes (NB)

[1] "Accuracy: 0.694249649368864"

[1] "Precision Large: 0.567295597484277"

[1] "Precision Medium: 0.623157894736842"

[1] "Precision Small: 0.849252013808976"

[1] "Recall Large: 0.79261862917399"

[1] "Recall Medium: 0.435294117647059"

[1] "Recall Small: 0.829213483146067"

[1] "F1 Score Large: 0.661290322580645"

[1] "F1 Score Medium: 0.512554112554113"

[1] "F1 Score Small: 0.839113132461626"

[1] "Specificity Large: 0.780891719745223"

[1] "Specificity Medium: 0.87731322823852"

[1] "Specificity Small: 0.895116092874299"

[1] "AUC Large : 0.881997694021246"

[1] "AUC Medium : 0.807180582993993"

[1] "AUC SMALL : 0.900576641088151"

Support Vector Machine (SVM)

[1] "Accuracy: 0.808789153810192"

[1] "Precision Large: 0.823747680890538"

[1] "Precision Medium: 0.700542005420054"

[1] "Precision Small: 0.892111368909513"

[1] "Recall Large: 0.780316344463972"

[1] "Recall Medium: 0.760294117647059"

[1] "Recall Small: 0.864044943820225"

[1] "F1 Score Large: 0.8014440433213"

[1] "F1 Score Medium: 0.729196050775741"

[1] "F1 Score Small: 0.877853881278539"

[1] "Specificity Large: 0.939490445859873"

[1] "Specificity Medium: 0.848526387936943"

[1] "Specificity Small: 0.925540432345877"

[1] "AUC Large : 0.912708629509812"

[1] "AUC Medium : 0.811029915736"

[1] "AUC SMALL : 0.939687480321336"

Decision Tree (DT)

[1] "Accuracy: 0.770920991117345"

[1] "Precision Large: 0.846827133479212"

[1] "Precision Medium: 0.609409190371991"

[1] "Precision Small: 0.91796875"

[1] "Recall Large: 0.680140597539543"

[1] "Recall Medium: 0.819117647058824"

[1] "Recall Small: 0.792134831460674"

[1] "F1 Score Large: 0.754385964912281"

[1] "F1 Score Medium: 0.698870765370138"

[1] "F1 Score Small: 0.850422195416164"

[1] "Specificity Large: 0.955414012738854"

[1] "Specificity Medium: 0.7553118574366"

[1] "Specificity Small: 0.949559647718175"

[1] "AUC Large : 0.895069011451535"

[1] "AUC Medium : 0.796545276781035"

[1] "AUC SMALL : 0.903325356914745"

Majority Based Ensembling (Voting Based):

[1] Accuraccy of Ensembled Data : 0.8410472

[1] "Precision Large: 0.846017699115044"

[1] "Precision Medium: 0.74167776298269"

[1] "Precision Small: 0.928311057108141"

[1] "Recall Large: 0.840070298769772"

[1] "Recall Medium: 0.819117647058824"

[1] "Recall Small: 0.858426966292135"

[1] "F1 Score Large: 0.843033509700176"

[1] "F1 Score Medium: 0.778476589797344"

[1] "F1 Score Small: 0.892002335084647"

[1] "Specificity Large: 0.944585987261147"

[1] "Specificity Medium: 0.867032213845099"

[1] "Specificity Small: 0.952762209767814"

**Model Training after feature selection ( 8 features )**

Gradient Boosting Machine (GBM)

[1] "Accuracy: 0.8835904628331"

[1] "Precision Large: 0.968"

[1] "Precision Medium: 0.800531914893617"

[1] "Precision Small: 0.906426155580609"

[1] "Recall Large: 0.850615114235501"

[1] "Recall Medium: 0.885294117647059"

[1] "Recall Small: 0.903370786516854"

[1] "F1 Score Large: 0.905519176800748"

[1] "F1 Score Medium: 0.840782122905028"

[1] "F1 Score Small: 0.904895891952729"

[1] "Specificity Large: 0.989808917197452"

[1] "Specificity Medium: 0.897189856065798"

[1] "Specificity Small: 0.933546837469976"

[1] "AUC Large : 0.984497330213919"

[1] "AUC Medium : 0.957834737733339"

[1] "AUC SMALL : 0.976801216253902"

Random Forest (RF)

[1] "Accuracy: 0.978027115474521"

[1] "Precision Large: 0.991087344028521"

[1] "Precision Medium: 0.945530726256983"

[1] "Precision Small: 0.996519721577726"

[1] "Recall Large: 0.977152899824253"

[1] "Recall Medium: 0.995588235294118"

[1] "Recall Small: 0.965168539325843"

[1] "F1 Score Large: 0.984070796460177"

[1] "F1 Score Medium: 0.969914040114613"

[1] "F1 Score Small: 0.980593607305936"

[1] "Specificity Large: 0.996815286624204"

[1] "Specificity Medium: 0.973269362577108"

[1] "Specificity Small: 0.99759807846277"

[1] "AUC Large : 0.99846921070601"

[1] "AUC Medium : 0.998927549086804"

[1] "AUC SMALL : 0.998692437095744"

Multinomial Logistic Regression (MLR)

[1] "Accuracy: 0.80551659654044"

[1] "Precision Large: 0.81767955801105"

[1] "Precision Medium: 0.698910081743869"

[1] "Precision Small: 0.888631090487239"

[1] "Recall Large: 0.780316344463972"

[1] "Recall Medium: 0.754411764705882"

[1] "Recall Small: 0.860674157303371"

[1] "F1 Score Large: 0.798561151079137"

[1] "F1 Score Medium: 0.725601131541726"

[1] "F1 Score Small: 0.874429223744292"

[1] "Specificity Large: 0.936942675159236"

[1] "Specificity Medium: 0.848526387936943"

[1] "Specificity Small: 0.923138510808647"

[1] "AUC Large : 0.922432919525819"

[1] "AUC Medium : 0.844618594524856"

[1] "AUC SMALL : 0.9465855830732"

Naive Bayes (NB)

[1] "Accuracy: 0.71575502571295"

[1] "Precision Large: 0.585585585585586"

[1] "Precision Medium: 0.663286004056795"

[1] "Precision Small: 0.861910241657077"

[1] "Recall Large: 0.799648506151142"

[1] "Recall Medium: 0.480882352941176"

[1] "Recall Small: 0.841573033707865"

[1] "F1 Score Large: 0.676077265973254"

[1] "F1 Score Medium: 0.557544757033248"

[1] "F1 Score Small: 0.85162023877203"

[1] "Specificity Large: 0.794904458598726"

[1] "Specificity Medium: 0.886223440712817"

[1] "Specificity Small: 0.903923138510809"

[1] "AUC Large : 0.889541378885742"

[1] "AUC Medium : 0.817807321694956"

[1] "AUC SMALL : 0.905269833844604"

Support Vector Machine (SVM)

[1] "Accuracy: 0.808789153810192"

[1] "Precision Large: 0.822222222222222"

[1] "Precision Medium: 0.701492537313433"

[1] "Precision Small: 0.892111368909513"

[1] "Recall Large: 0.780316344463972"

[1] "Recall Medium: 0.760294117647059"

[1] "Recall Small: 0.864044943820225"

[1] "F1 Score Large: 0.800721370604148"

[1] "F1 Score Medium: 0.729710656316161"

[1] "F1 Score Small: 0.877853881278539"

[1] "Specificity Large: 0.938853503184713"

[1] "Specificity Medium: 0.849211788896505"

[1] "Specificity Small: 0.925540432345877"

[1] "AUC Large : 0.912725420617241"

[1] "AUC Medium : 0.811322219086401"

[1] "AUC SMALL : 0.939694677089987"

Decision Tree (DT)

[1] "Accuracy: 0.770920991117345"

[1] "Precision Large: 0.846827133479212"

[1] "Precision Medium: 0.609409190371991"

[1] "Precision Small: 0.91796875"

[1] "Recall Large: 0.680140597539543"

[1] "Recall Medium: 0.819117647058824"

[1] "Recall Small: 0.792134831460674"

[1] "F1 Score Large: 0.754385964912281"

[1] "F1 Score Medium: 0.698870765370138"

[1] "F1 Score Small: 0.850422195416164"

[1] "Specificity Large: 0.955414012738854"

[1] "Specificity Medium: 0.7553118574366"

[1] "Specificity Small: 0.949559647718175"

[1] "AUC Large : 0.895069011451535"

[1] "AUC Medium : 0.796545276781035"

[1] "AUC SMALL : 0.903325356914745"

Majority Based Ensembling (Voting Based):

[1] Accuraccy of Ensembled Data : 0.8433848

[1] "Precision Large: 0.848056537102474"

[1] "Precision Medium: 0.748993288590604"

[1] "Precision Small: 0.926239419588875"

[1] "Recall Large: 0.843585237258348"

[1] "Recall Medium: 0.820588235294118"

[1] "Recall Small: 0.860674157303371"

[1] "F1 Score Large: 0.845814977973568"

[1] "F1 Score Medium: 0.783157894736842"

[1] "F1 Score Small: 0.892253931275481"

[1] "Specificity Large: 0.945222929936306"

[1] "Specificity Medium: 0.871144619602467"

[1] "Specificity Small: 0.951160928742994"