<TO DO>

10-re run Average runtime for traintime & testtime of all ML methods

Independent Sample T test

Student’s T test

Two-sample location test : such that means of two populations are equal

(variance of the two populations are assumed to be equal)

Welch’s T test

Two-sample location test : such that means of two populations are equal

Assumption dropped

**<Testing Time T-test comparison>**

Comparing lr\_traintime and dtree\_traintime

New degree of freedom: 9.024333841492712

Test T-Score: 41.15836389300091

Comparable T-score: 3.2476188262299925

Significantly, lr\_traintime is better than dtree\_traintime / lr\_traintime mean accuracy: 0.38272788519999995 / dtree\_traintime mean accuracy: 0.008533930700000001

Comparing lr\_traintime and adaboost\_traintime

New degree of freedom: 9.831453763894432

Test T-Score: 33.28794258692248

Comparable T-score: 3.1814849392671953

Significantly, lr\_traintime is better than adaboost\_traintime / lr\_traintime mean accuracy: 0.38272788519999995 / adaboost\_traintime mean accuracy: 0.0733719349

Comparing lr\_traintime and rforest\_traintime

New degree of freedom: 9.429140710038748

Test T-Score: 37.23254176828583

Comparable T-score: 3.2127530042271824

Significantly, lr\_traintime is better than rforest\_traintime / lr\_traintime mean accuracy: 0.38272788519999995 / rforest\_traintime mean accuracy: 0.0404434443

Comparing lr\_traintime and kernelsvm\_traintime

New degree of freedom: 9.68841184920536

Test T-Score: 32.420233167629064

Comparable T-score: 3.1922472292295643

Significantly, lr\_traintime is better than kernelsvm\_traintime / lr\_traintime mean accuracy: 0.38272788519999995 / kernelsvm\_traintime mean accuracy: 0.0825884103

Comparing lr\_traintime and gp\_traintime

New degree of freedom: 9.062823082587759

Test T-Score: -11.948609640990709

Comparable T-score: -3.244141945766534

Statistically, no difference detected. But in this sample, gp\_traintimeis slightly better. gp\_traintime mean accuracy: 2.2234712123

Comparing lr\_traintime and nn\_traintime

New degree of freedom: 9.070111347899156

Test T-Score: -67.48362708259448

Comparable T-score: -3.2434875796024705

Statistically, no difference detected. But in this sample, nn\_traintimeis slightly better. nn\_traintime mean accuracy: 10.22571

Comparing dtree\_traintime and adaboost\_traintime

New degree of freedom: 9.525225735237756

Test T-Score: -32.69534819858653

Comparable T-score: -3.2049978037215383

Statistically, no difference detected. But in this sample, adaboost\_traintimeis slightly better. adaboost\_traintime mean accuracy: 0.0733719349

Comparing dtree\_traintime and rforest\_traintime

New degree of freedom: 10.016821273271653

Test T-Score: -22.121661662975235

Comparable T-score: -3.1680806229037968

Statistically, no difference detected. But in this sample, rforest\_traintimeis slightly better. rforest\_traintime mean accuracy: 0.0404434443

Comparing dtree\_traintime and kernelsvm\_traintime

New degree of freedom: 9.634538953422856

Test T-Score: -40.93242838624701

Comparable T-score: -3.196399542462392

Statistically, no difference detected. But in this sample, kernelsvm\_traintimeis slightly better. kernelsvm\_traintime mean accuracy: 0.0825884103

Comparing dtree\_traintime and gp\_traintime

New degree of freedom: 9.000084930463888

Test T-Score: -14.402607240847145

Comparable T-score: -3.2498277819849837

Statistically, no difference detected. But in this sample, gp\_traintimeis slightly better. gp\_traintime mean accuracy: 2.2234712123

Comparing dtree\_traintime and nn\_traintime

New degree of freedom: 9.000094783746077

Test T-Score: -70.18521416887616

Comparable T-score: -3.2498268814776896

Statistically, no difference detected. But in this sample, nn\_traintimeis slightly better. nn\_traintime mean accuracy: 10.22571

Comparing adaboost\_traintime and rforest\_traintime

New degree of freedom: 16.329422717836668

Test T-Score: 13.684419492486557

Comparable T-score: 2.913013167806354

Significantly, adaboost\_traintime is better than rforest\_traintime / adaboost\_traintime mean accuracy: 0.0733719349 / rforest\_traintime mean accuracy: 0.0404434443

Comparing adaboost\_traintime and kernelsvm\_traintime

New degree of freedom: 17.840848124976812

Test T-Score: -3.4878255042847277

Comparable T-score: -2.881426326509641

Statistically, no difference detected. But in this sample, kernelsvm\_traintimeis slightly better. kernelsvm\_traintime mean accuracy: 0.0825884103

Comparing adaboost\_traintime and gp\_traintime

New degree of freedom: 9.002908169728219

Test T-Score: -13.979902473075683

Comparable T-score: -3.2495698590672815

Statistically, no difference detected. But in this sample, gp\_traintimeis slightly better. gp\_traintime mean accuracy: 2.2234712123

Comparing adaboost\_traintime and nn\_traintime

New degree of freedom: 9.003245563566725

Test T-Score: -69.7337172283953

Comparable T-score: -3.2495390488272413

Statistically, no difference detected. But in this sample, nn\_traintimeis slightly better. nn\_traintime mean accuracy: 10.22571

Comparing rforest\_traintime and kernelsvm\_traintime

New degree of freedom: 17.07746462282735

Test T-Score: -18.6062621763615

Comparable T-score: -2.8966057130195977

Statistically, no difference detected. But in this sample, kernelsvm\_traintimeis slightly better. kernelsvm\_traintime mean accuracy: 0.0825884103

Comparing rforest\_traintime and gp\_traintime

New degree of freedom: 9.001498645120677

Test T-Score: -14.19455854720382

Comparable T-score: -3.2496986046861727

Statistically, no difference detected. But in this sample, gp\_traintimeis slightly better. gp\_traintime mean accuracy: 2.2234712123

Comparing rforest\_traintime and nn\_traintime

New degree of freedom: 9.001672511744696

Test T-Score: -69.96295090474513

Comparable T-score: -3.249682721118045

Statistically, no difference detected. But in this sample, nn\_traintimeis slightly better. nn\_traintime mean accuracy: 10.22571

Comparing kernelsvm\_traintime and gp\_traintime

New degree of freedom: 9.002406229198515

Test T-Score: -13.920171190539653

Comparable T-score: -3.24961570060083

Statistically, no difference detected. But in this sample, gp\_traintimeis slightly better. gp\_traintime mean accuracy: 2.2234712123

Comparing kernelsvm\_traintime and nn\_traintime

New degree of freedom: 9.00268538997566

Test T-Score: -69.67149563142766

Comparable T-score: -3.249590204471039

Statistically, no difference detected. But in this sample, nn\_traintimeis slightly better. nn\_traintime mean accuracy: 10.22571

Comparing gp\_traintime and nn\_traintime

New degree of freedom: 17.946053166832296

Test T-Score: -37.78922115921443

Comparable T-score: -2.8794460283593137

Statistically, no difference detected. But in this sample, nn\_traintimeis slightly better. nn\_traintime mean accuracy: 10.22571

**<Testing Time T-test comparison>**

Comparing lr\_testtime and dtree\_testtime

New degree of freedom: 9.067273311447666

Test T-Score: 82.92676472579211

Comparable T-score: 3.243742237558301

Significantly, lr\_testtime is better than dtree\_testtime / lr\_testtime mean accuracy: 0.45868430149999995 / dtree\_testtime mean accuracy: 0.0012292863000000001

Comparing lr\_testtime and adaboost\_testtime

New degree of freedom: 9.078355301152882

Test T-Score: 81.77891768778332

Comparable T-score: 3.2427489367197455

Significantly, lr\_testtime is better than adaboost\_testtime / lr\_testtime mean accuracy: 0.45868430149999995 / adaboost\_testtime mean accuracy: 0.0074229002999999984

Comparing lr\_testtime and rforest\_testtime

New degree of freedom: 9.006319239688464

Test T-Score: 82.85264579499179

Comparable T-score: 3.2492584943537914

Significantly, lr\_testtime is better than rforest\_testtime / lr\_testtime mean accuracy: 0.45868430149999995 / rforest\_testtime mean accuracy: 0.002409792

Comparing lr\_testtime and kernelsvm\_testtime

New degree of freedom: 9.09211446182014

Test T-Score: 76.77552604992763

Comparable T-score: 3.2415197410124335

Significantly, lr\_testtime is better than kernelsvm\_testtime / lr\_testtime mean accuracy: 0.45868430149999995 / kernelsvm\_testtime mean accuracy: 0.0348707438

Comparing lr\_testtime and gp\_testtime

New degree of freedom: 9.000775218140717

Test T-Score: -21.37284689157829

Comparable T-score: -3.2497647012495086

Statistically, no difference detected. But in this sample, gp\_testtime is slightly better. gp\_testtime mean accuracy: 18.391133951

Comparing lr\_testtime and nn\_testtime

New degree of freedom: 10.34965642257543

Test T-Score: -14.06682264649891

Comparable T-score: -3.145434548345073

Statistically, no difference detected. But in this sample, nn\_testtime is slightly better. nn\_testtime mean accuracy: 0.75118

Comparing dtree\_testtime and adaboost\_testtime

New degree of freedom: 17.89635819031156

Test T-Score: -12.505751999477427

Comparable T-score: -2.8803782476976183

Statistically, no difference detected. But in this sample, adaboost\_testtime is slightly better. adaboost\_testtime mean accuracy: 0.0074229002999999984

Comparing dtree\_testtime and rforest\_testtime

New degree of freedom: 10.67599763726382

Test T-Score: -3.3530577371014583

Comparable T-score: -3.124840151527637

Statistically, no difference detected. But in this sample, rforest\_testtime is slightly better. rforest\_testtime mean accuracy: 0.002409792

Comparing dtree\_testtime and kernelsvm\_testtime

New degree of freedom: 17.573110697046104

Test T-Score: -64.92850184301071

Comparable T-score: -2.8865841999692687

Statistically, no difference detected. But in this sample, kernelsvm\_testtime is slightly better. kernelsvm\_testtime mean accuracy: 0.0348707438

Comparing dtree\_testtime and gp\_testtime

New degree of freedom: 9.000002897345556

Test T-Score: -21.918536310759137

Comparable T-score: -3.2498352792160152

Statistically, no difference detected. But in this sample, gp\_testtime is slightly better. gp\_testtime mean accuracy: 18.391133951

Comparing dtree\_testtime and nn\_testtime

New degree of freedom: 9.005072967426358

Test T-Score: -37.39680892764113

Comparable T-score: -3.2493722218539567

Statistically, no difference detected. But in this sample, nn\_testtime is slightly better. nn\_testtime mean accuracy: 0.75118

Comparing adaboost\_testtime and rforest\_testtime

New degree of freedom: 10.442265761678023

Test T-Score: 13.274540581839481

Comparable T-score: 3.139435999535531

Significantly, adaboost\_testtime is better than rforest\_testtime / adaboost\_testtime mean accuracy: 0.0074229002999999984 / rforest\_testtime mean accuracy: 0.002409792

Comparing adaboost\_testtime and kernelsvm\_testtime

New degree of freedom: 17.883485483780614

Test T-Score: -51.223851919936166

Comparable T-score: -2.880620656495208

Statistically, no difference detected. But in this sample, kernelsvm\_testtime is slightly better. kernelsvm\_testtime mean accuracy: 0.0348707438

Comparing adaboost\_testtime and gp\_testtime

New degree of freedom: 9.00000337464456

Test T-Score: -21.91115398343426

Comparable T-score: -3.249835235593877

Statistically, no difference detected. But in this sample, gp\_testtime is slightly better. gp\_testtime mean accuracy: 18.391133951

Comparing adaboost\_testtime and nn\_testtime

New degree of freedom: 9.005908670922405

Test T-Score: -37.08709941209162

Comparable T-score: -3.2492959562504495

Statistically, no difference detected. But in this sample, nn\_testtime is slightly better. nn\_testtime mean accuracy: 0.75118

Comparing rforest\_testtime and kernelsvm\_testtime

New degree of freedom: 10.229020708848193

Test T-Score: -79.72172098022362

Comparable T-score: -3.153441183354744

Statistically, no difference detected. But in this sample, kernelsvm\_testtime is slightly better. kernelsvm\_testtime mean accuracy: 0.0348707438

Comparing rforest\_testtime and gp\_testtime

New degree of freedom: 9.000000272154994

Test T-Score: -21.91713088935856

Comparable T-score: -3.249835519142088

Statistically, no difference detected. But in this sample, gp\_testtime is slightly better. gp\_testtime mean accuracy: 18.391133951

Comparing rforest\_testtime and nn\_testtime

New degree of freedom: 9.00047651669131

Test T-Score: -37.34270904577926

Comparable T-score: -3.249791996125047

Statistically, no difference detected. But in this sample, nn\_testtime is slightly better. nn\_testtime mean accuracy: 0.75118

Comparing kernelsvm\_testtime and gp\_testtime

New degree of freedom: 9.000003967259566

Test T-Score: -21.878439122722952

Comparable T-score: -3.2498351814325845

Statistically, no difference detected. But in this sample, gp\_testtime is slightly better. gp\_testtime mean accuracy: 18.391133951

Comparing kernelsvm\_testtime and nn\_testtime

New degree of freedom: 9.0069462812611

Test T-Score: -35.71739643937128

Comparable T-score: -3.2492012885991937

Statistically, no difference detected. But in this sample, nn\_testtime is slightly better. nn\_testtime mean accuracy: 0.75118

Comparing gp\_testtime and nn\_testtime

New degree of freedom: 9.010280412717599

Test T-Score: 21.018686306024303

Comparable T-score: 3.2488972733929042

Significantly, gp\_testtime is better than nn\_testtime / gp\_testtime mean accuracy: 18.391133951 / nn\_testtime mean accuracy: 0.75118