EXOS Performance Evaluation

Egawati Panjei

2022-05-08

Contents

Case

In this report, we generate 30 datasets for Case 1 (all data attributes have strong correlation) and Case 4 (all data attributes have weak correlation). Dataset of experiment i for each case is generated using the same seed random number. Hence, they have the same number of attributes in each group (stream). Each attribute j in Case 1 and Case 2 has similar mean and range. However, the mean and range of each attribute in each case is different. Cases 1 and 2 also have the same number of inlier and outlier data points. The difference is in their data attribute correlation.

Case 1

```
df_case1 = read.csv('dbpca/pickles/performance/Case1/aggregate_15_10K_Case1_30.csv')
df_case1
```

```
##    experiment precision    recall  f1_score running_time
## 1           1 0.9121222 0.6742460 0.7581076     43.48957
## 2           2 0.8627222 0.6107571 0.7002900     43.71341
## 3           3 0.9391556 0.6709714 0.7651223     43.96031
## 4           4 0.9018079 0.6712333 0.7478692     43.53199
## 5           5 0.9620333 0.6927849 0.7926332     44.23652
## 6           6 0.9174111 0.6836159 0.7611473     44.03168
## 7           7 0.9190889 0.6649190 0.7514966     43.98473
## 8           8 0.8883143 0.6616825 0.7412251     43.62066
## 9           9 0.9294127 0.6860270 0.7742514     44.16683
## 10         10 0.9030111 0.6582762 0.7467393     44.42039
## 11         11 0.9228778 0.6954317 0.7817240     43.90644
## 12         12 0.9018286 0.6648698 0.7456533     43.77420
## 13         13 0.9215627 0.6780095 0.7612376     44.20242
## 14         14 0.8789476 0.6375159 0.7253344     43.83694
## 15         15 0.8948333 0.6548746 0.7411063     43.94390
## 16         16 0.9195794 0.6662611 0.7538053     44.13076
## 17         17 0.8991889 0.6503857 0.7361406     44.05734
## 18         18 0.9053476 0.6897429 0.7638855     43.92518
## 19         19 0.9098931 0.6822778 0.7590703     43.90918
## 20         20 0.9272206 0.6627511 0.7592260     44.11758
## 21         21 0.9217222 0.6662524 0.7589832     43.65052
## 22         22 0.9143063 0.6905286 0.7709330     43.76140
## 23         23 0.9037238 0.7092095 0.7680099     43.99387
## 24         24 0.9449079 0.6812913 0.7786458     43.91784
## 25         25 0.9070770 0.6889778 0.7645008     43.96320
## 26         26 0.9549667 0.6956937 0.7936294     43.69568
## 27         27 0.8746524 0.6806556 0.7406413     43.69984
## 28         28 0.8985841 0.6758754 0.7508982     44.11883
## 29         29 0.8708071 0.6503873 0.7286918     43.93447
## 30         30 0.7988212 0.6627603 0.7036146     44.28376
```

Case 4

```
df_case4 = read.csv('dbpca/pickles/performance/Case4/aggregate_15_10K_Case4_30.csv')
df_case4
```

```
##    experiment precision    recall  f1_score running_time
## 1           1 0.9164778 0.6739762 0.7593747     43.64827
## 2           2 0.8634111 0.6099444 0.7010496     43.80902
## 3           3 0.9410000 0.6702302 0.7641184     43.75594
## 4           4 0.9025063 0.6703429 0.7480583     43.61513
## 5           5 0.9615000 0.6882135 0.7895402     43.91414
## 6           6 0.9195000 0.6831206 0.7626047     43.99893
## 7           7 0.9187444 0.6586095 0.7475022     43.94333
## 8           8 0.8899794 0.6667714 0.7445775     43.70337
## 9           9 0.9365976 0.6864746 0.7784071     44.02096
## 10         10 0.9027389 0.6587937 0.7475211     44.19689
## 11         11 0.9252889 0.7000492 0.7859586     44.05833
## 12         12 0.9029397 0.6679587 0.7476245     43.69912
## 13         13 0.9186333 0.6712206 0.7576184     44.02532
## 14         14 0.8823532 0.6408825 0.7274201     44.00314
```

```
## 15         15 0.8945000 0.6520921 0.7393985     44.16800
## 16         16 0.9188746 0.6652095 0.7536188     44.23551
## 17         17 0.8959333 0.6497294 0.7345150     44.10043
## 18         18 0.8983952 0.6827746 0.7573400     44.12348
## 19         19 0.9078683 0.6791079 0.7571432     44.11483
## 20         20 0.9296429 0.6647786 0.7623860     44.36835
## 21         21 0.9235000 0.6678151 0.7603745     43.79056
## 22         22 0.9164698 0.6909373 0.7718588     44.17115
## 23         23 0.9049587 0.7076984 0.7676054     44.16056
## 24         24 0.9483048 0.6848254 0.7824826     44.18433
## 25         25 0.9022111 0.6908524 0.7613557     44.01492
## 26         26 0.9566000 0.6967937 0.7951088     44.33244
## 27         27 0.8730730 0.6808746 0.7409514     43.70218
## 28         28 0.8968794 0.6750778 0.7495402     43.89961
## 29         29 0.8642238 0.6497476 0.7245399     43.81950
## 30         30 0.7904142 0.6703820 0.7029776     44.45250
```

Precision

Variance Test for Precision

```
var.test(df_case1$precision, df_case4$precision)
```

```
##
##  F test to compare two variances
##
## data:  df_case1$precision and df_case4$precision
## F = 0.88582, num df = 29, denom df = 29, p-value = 0.7463
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4216172 1.8610959
## sample estimates:
## ratio of variances
##          0.8858161
```

The var test shows that the true ratio of precision variances of Case 1 and Case 4 is equal to 1 (accept the null hypothesis).

t Test for Precision

```
t.test(df_case1$precision, df_case4$precision, alternative = "greater", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  df_case1$precision and df_case4$precision
## t = 0.0098005, df = 58, p-value = 0.4961
## alternative hypothesis: true difference in means is greater than 0
```

```
## 95 percent confidence interval:
##  -0.01360973         Inf
## sample estimates:
## mean of x mean of y
## 0.9068643 0.9067840
```

The t test shows that there is no difference between the precision mean of Case 1 and Case 4. We accept the null hypothesis.

Recall

Variance Test for recall

```r
var.test(df_case1$recall, df_case4$recall)
```

```
##
##  F test to compare two variances
##
## data:  df_case1$recall and df_case4$recall
## F = 1.0173, num df = 29, denom df = 29, p-value = 0.9635
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4841843 2.1372783
## sample estimates:
## ratio of variances
##           1.017269
```

The var test shows that the true ratio of recall variances of Case 1 and Case 4 is equal to 1 (accept the null hypothesis).

t Test for recall

```r
t.test(df_case1$recall, df_case4$recall, alternative = "greater", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  df_case1$recall and df_case4$recall
## t = 0.019553, df = 58, p-value = 0.4922
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.008395057          Inf
## sample estimates:
## mean of x mean of y
## 0.6719422 0.6718428
```

The t test shows that there is no difference between the recall mean of Case 1 and Case 4. We accept the null hypothesis.

F1 score

Variance Test for F1 score

```r
var.test(df_case1$f1_score, df_case4$f1_score)
```

```
##
##  F test to compare two variances
##
## data:  df_case1$f1_score and df_case4$f1_score
## F = 0.96131, num df = 29, denom df = 29, p-value = 0.9161
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4575488 2.0197041
## sample estimates:
## ratio of variances
##           0.961308
```

The var test shows that the true ratio of f1 score variances of Case 1 and Case 4 is equal to 1 (accept the null hypothesis).

t Test for F1 score

```r
t.test(df_case1$f1_score, df_case4$f1_score, alternative = "greater", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  df_case1$f1_score and df_case4$f1_score
## t = 0.011983, df = 58, p-value = 0.4952
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.009423786          Inf
## sample estimates:
## mean of x mean of y
## 0.7541538 0.7540857
```

The t test shows that there is no difference between the F1 score mean of Case 1 and Case 4. We accept the null hypothesis since p-value is not small enough.