Visualizations LDA coefficients

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```r
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.0.5
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 4.0.5
```

```r
library("ggpubr")
```

```
## Warning: package 'ggpubr' was built under R version 4.0.5
```

```r
library(LDATS)
```

```
## Warning: package 'LDATS' was built under R version 4.0.5
```

```r
library(ggVennDiagram)
```

```
## Warning: package 'ggVennDiagram' was built under R version 4.0.5
```

```r
library(stringr)
library(abind)
```

```
## Warning: package 'abind' was built under R version 4.0.3
```

```r
library(patchwork)
```

```
## Warning: package 'patchwork' was built under R version 4.0.3
```

```r
source("utils.R")
```

```
## Warning: package 'hash' was built under R version 4.0.5
```

```
## hash-2.2.6.1 provided by Decision Patterns
```

```
## Warning: package 'reticulate' was built under R version 4.0.5
```

```
## Warning: package 'berryFunctions' was built under R version 4.0.5
```

```
##
## Attaching package: 'berryFunctions'
```

```
## The following object is masked from 'package:ggVennDiagram':
##
##     circle
```

```
## The following object is masked from 'package:dplyr':
##
##     between
```

```
## Warning: package 'purrr' was built under R version 4.0.3
```

```
## Warning: package 'reshape2' was built under R version 4.0.3
```

```
##
## Attaching package: 'reshape2'
```

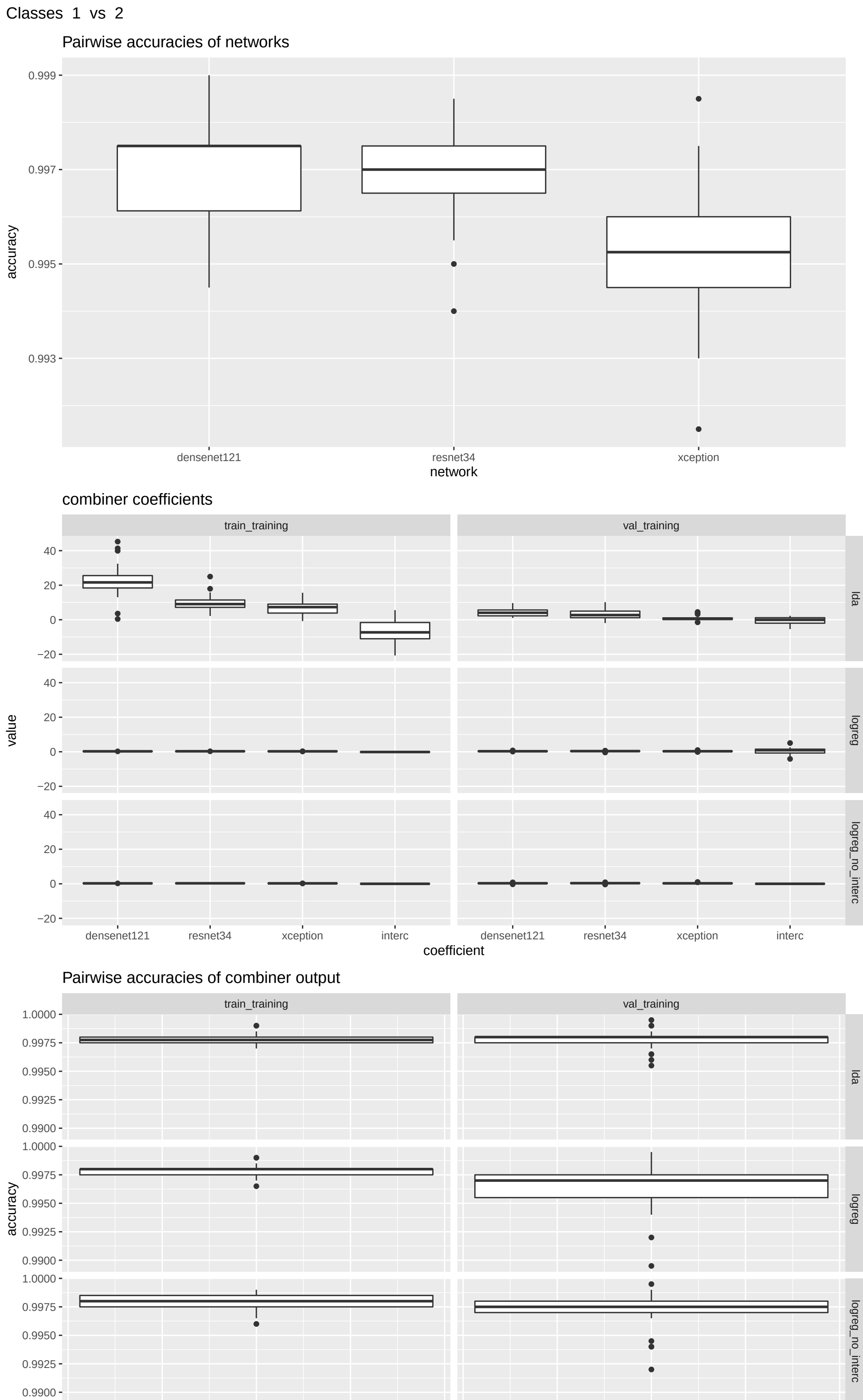
```
## The following object is masked from 'package:tidyr':
##
##     smiths
```

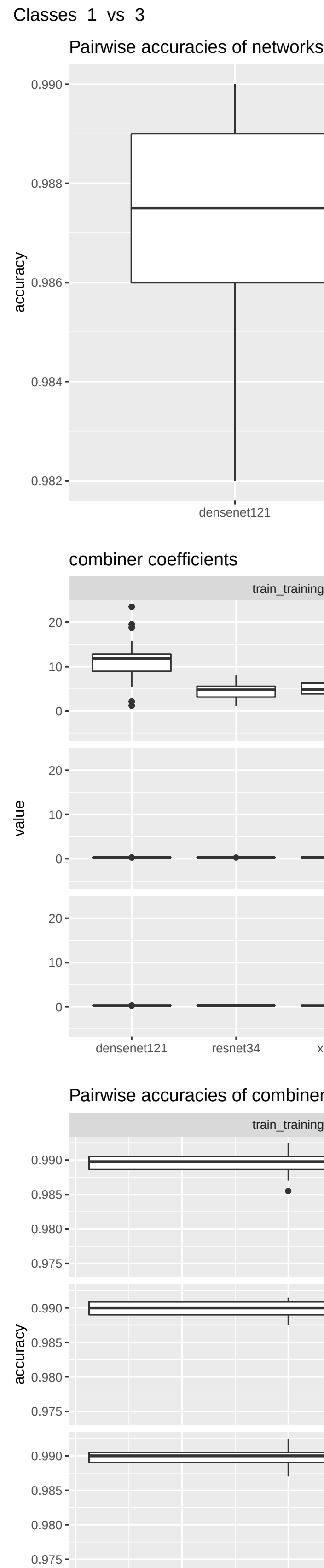
Visualizations on three sets of networks. First set consists of networks trained in 30 replications on CIFAR 10 training set with 500 sample validation set extracted randomly in each replication. Second set consists of networks trained in a single replication on half of CIFAR 10 training set. Third set are networks trained in 10 replications on half of CIFAR 100 training set.

CIFAR 10

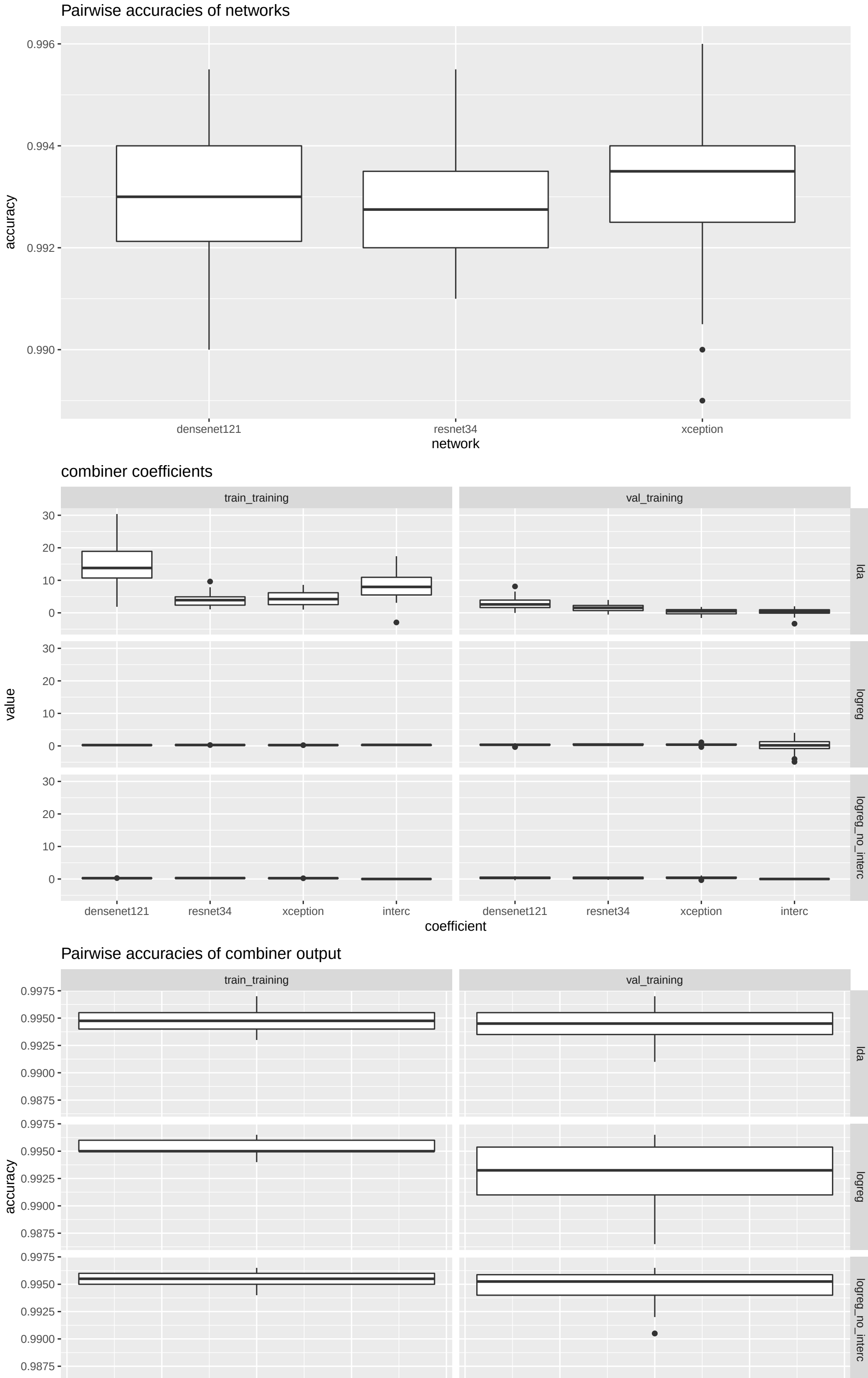
```
base_dir <- "../data/data_train_val_c10"
repls <- 0:29
classes <- 10
combiner_coefs <- load_combiner_coefs(base_dir, repls)
net_pw_results <- read.csv(file.path(base_dir, "net_pw_accuracies.csv"))
ens_pw_results <- read.csv(file.path(base_dir, "ensemble_pw_accuracies.csv"))
ens_pw_cal <- read.csv(file.path(base_dir, "ensemble_pw_calibration.csv"))
ens_pw_irrel <- read.csv(file.path(base_dir, "ensemble_pw_irrelevant.csv"))
net_pw_results[, c("class1", "class2")] <- lapply(net_pw_results[, c("class1", "class2")], as.factor)
ens_pw_results[, c("class1", "class2", "combining_method")] <- lapply(ens_pw_results[, c("class1", "cla
ens_pw_cal$bin_c <- (ens_pw_cal$conf_min + ens_pw_cal$conf_max) / 2
ens_pw_cal[, c("class1", "class2", "bin_c", "combining_method")] <- lapply(ens_pw_cal[, c("class1", "cla
ens_pw_irrel[, c("class1", "class2", "combining_method")] <- lapply(ens_pw_irrel[, c("class1", "class2"
```

```
for (cl1 in 1:(classes - 1))
{
  for (cl2 in (cl1 + 1):classes)
  {
    combiner_plt <- combiner_coefs %>% filter(class1 == cl1 & class2 == cl2) %>% ggplot() + geom_boxplot
      facet_grid(cols=vars(train_type), rows=vars(combining_method)) + ggtitle("combiner coefficients")

    acc_plt_net <- net_pw_results %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot(mappi
      geom_boxplot() + ggtitle("Pairwise accuracies of networks")

    acc_plt_ens <- ens_pw_results %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot(mappi
      geom_boxplot() + facet_grid(cols=vars(train_set), rows=vars(combining_method)) + ggtitle("Pairwise
      theme(axis.ticks.x=element_blank(), axis.text.x=element_blank())


    print((acc_plt_net/combiner_plt/acc_plt_ens) + plot_annotation(title=paste("Classes ", cl1, " vs ",
  }
}
```

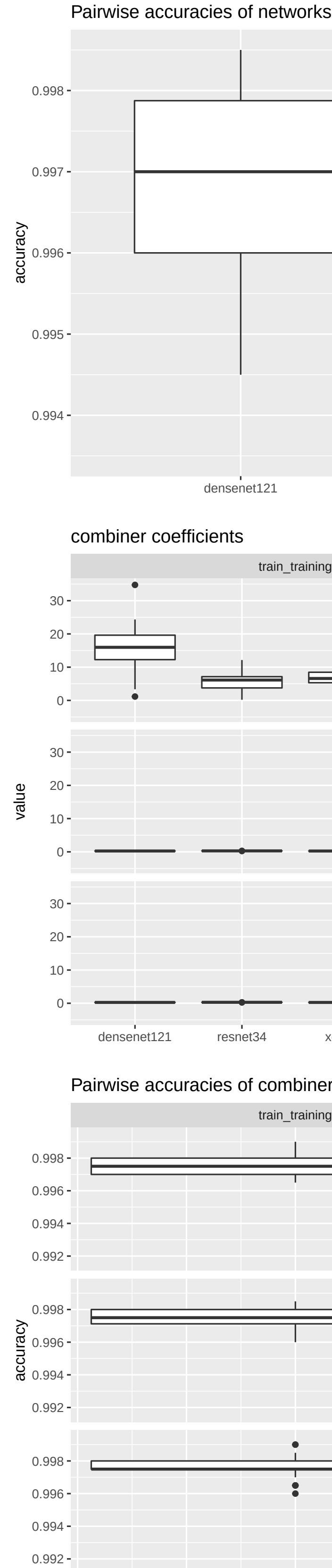

Classes 1 vs 2
Pairwise accuracies of networks
accuracy
0.999
0.997
0.995
0.993
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
40
20
0
−20
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
accuracy
1.0000
0.9975
0.9950
0.9925
0.9900


Classes 1 vs 3
Pairwise accuracies of networks
accuracy
0.990
0.988
0.986
0.984
0.982
densenet121
combiner coefficients
train_training
value
20
10
0
densenet121
resnet34
Pairwise accuracies of combiner
accuracy
0.990
0.985
0.980
0.975

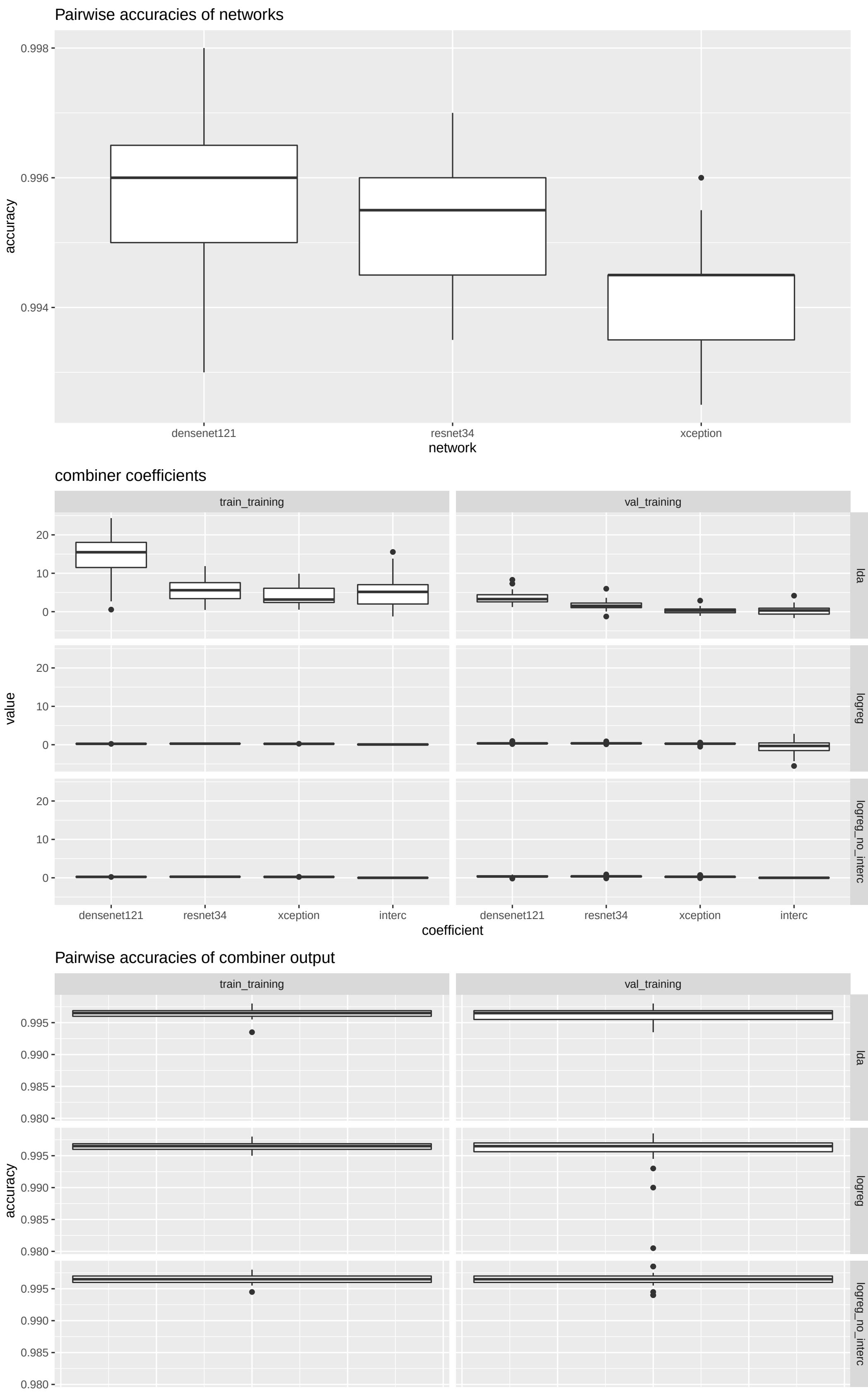
Classes 1 vs 4



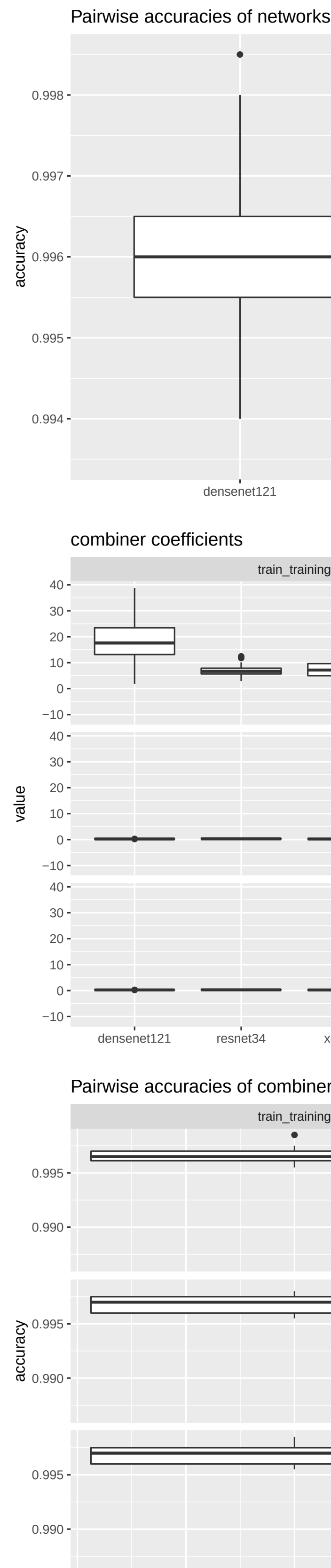
Classes 1 vs 5

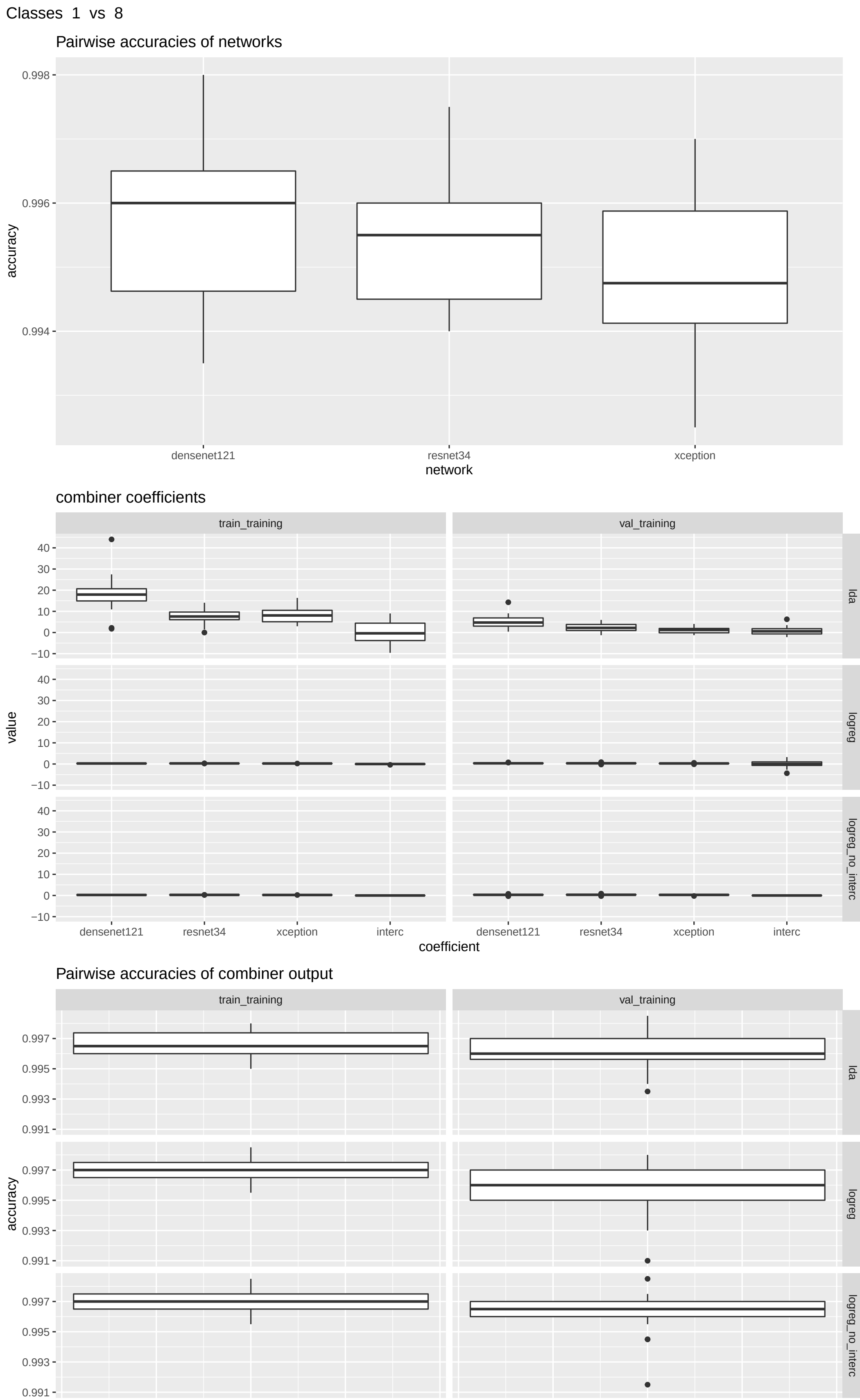


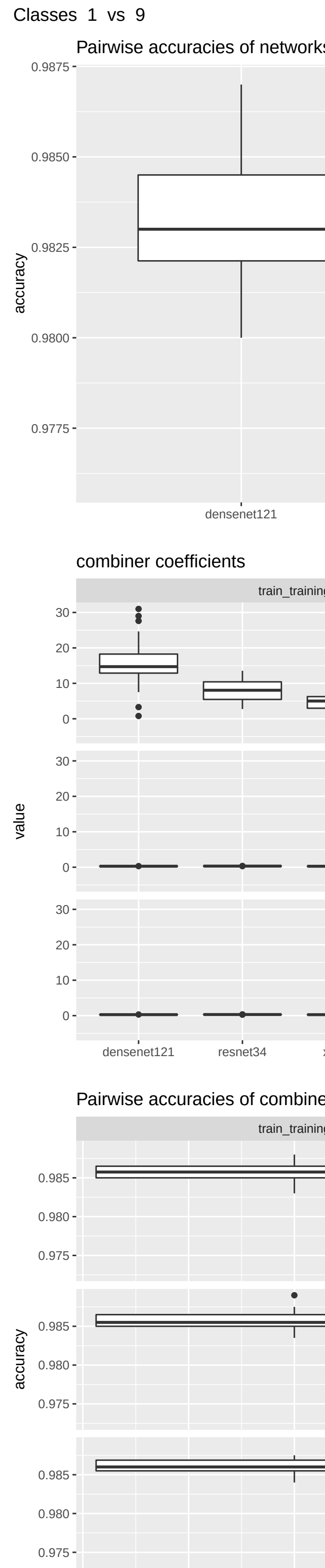
Classes 1 vs 6

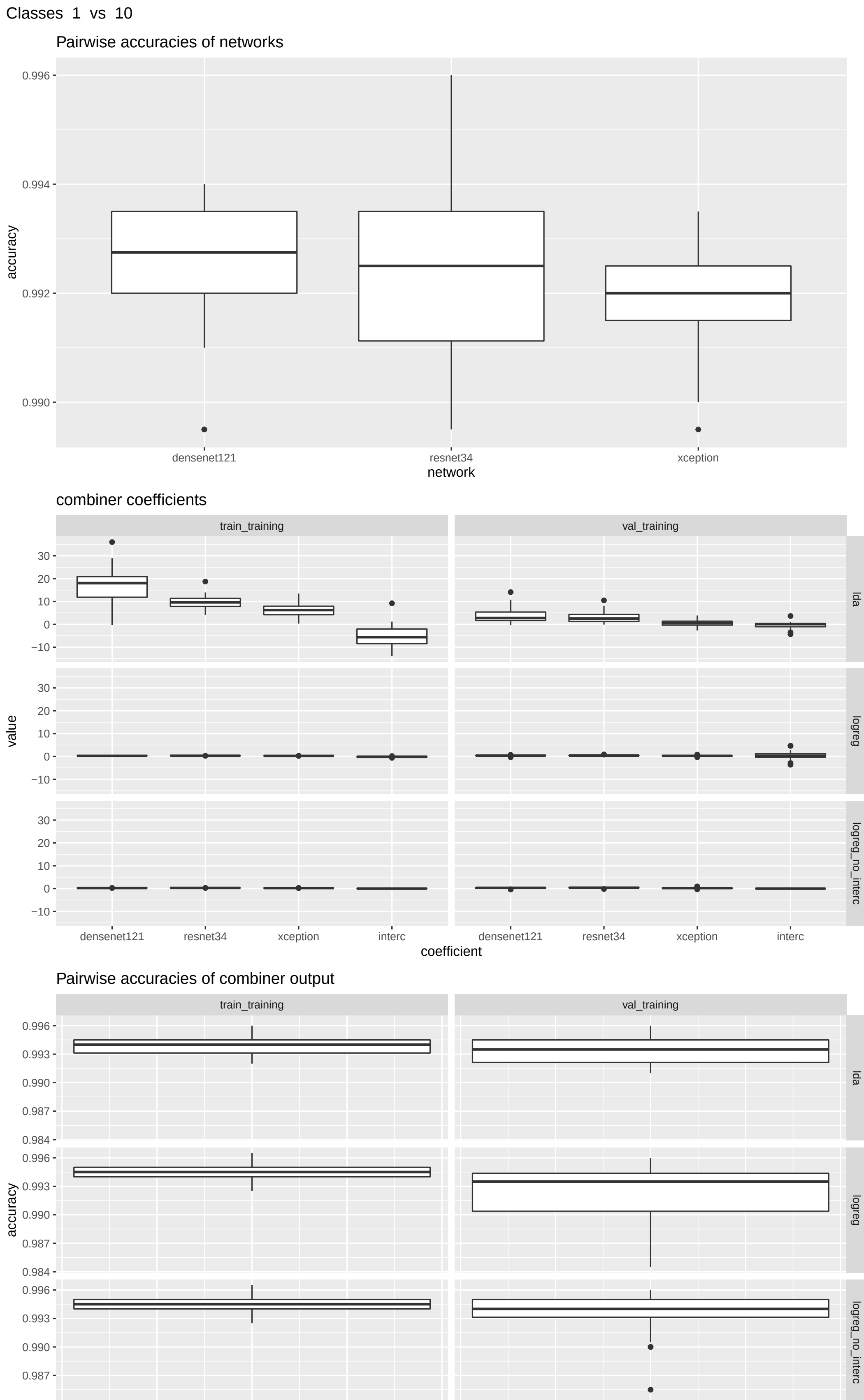


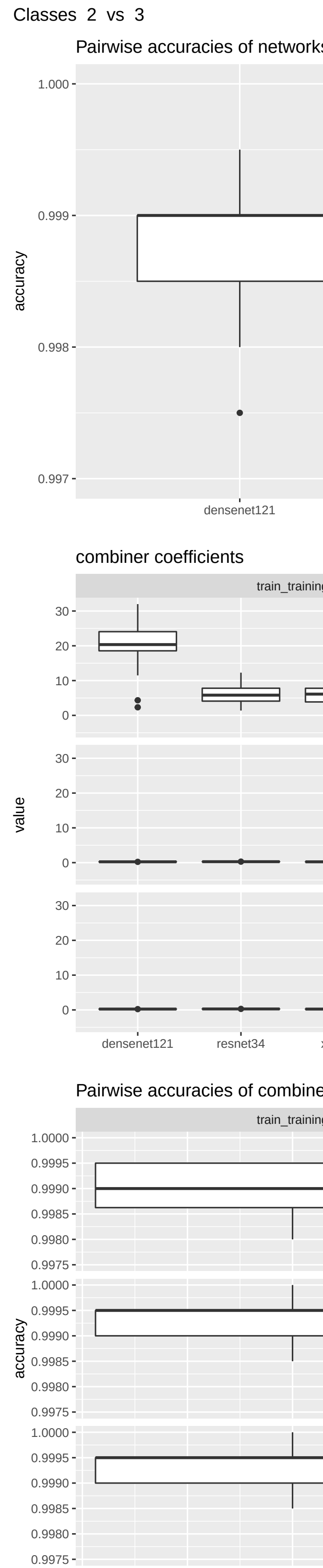
Classes 1 vs 7

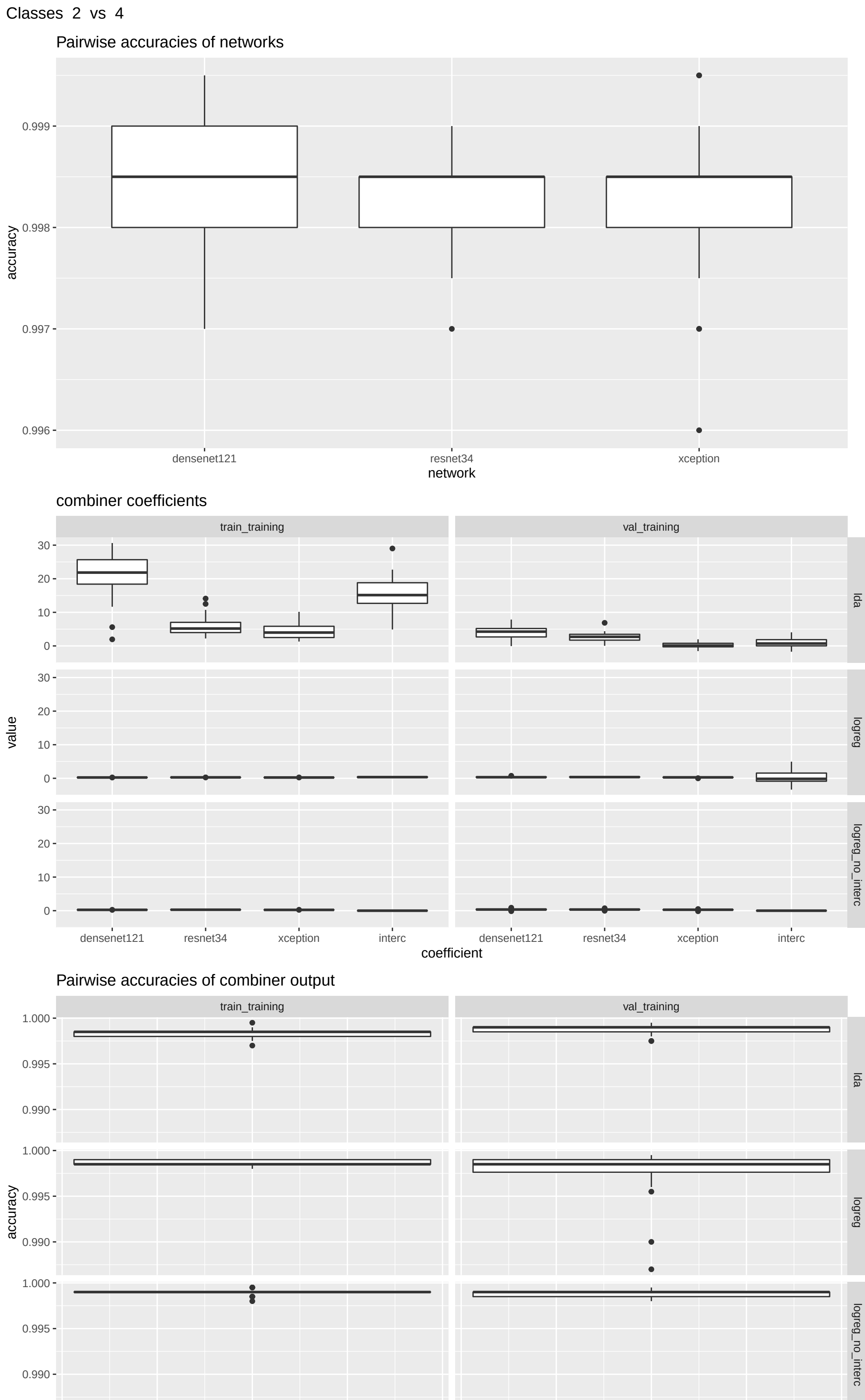


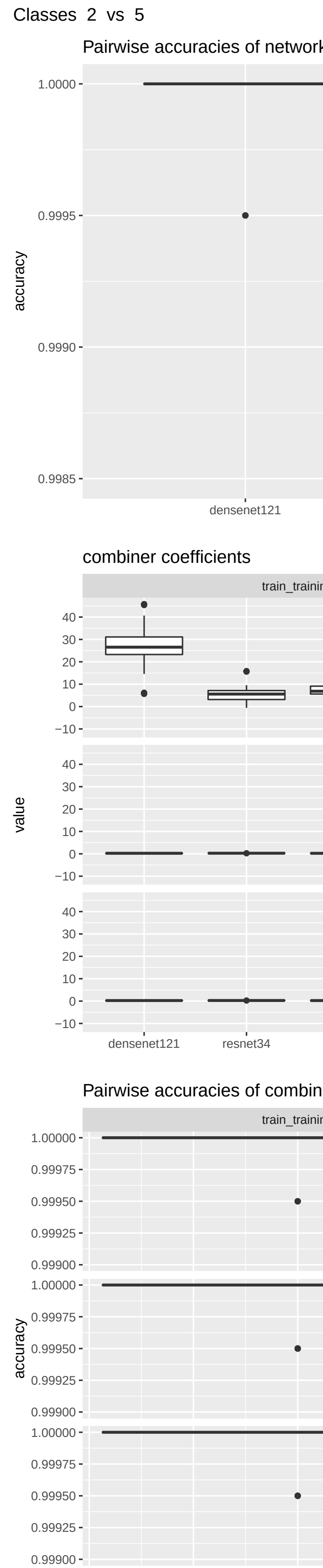

Classes 1 vs 8
Pairwise accuracies of networks
accuracy
0.998
0.996
0.994
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
40
30
20
10
0
−10
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.997
0.995
0.993
0.991

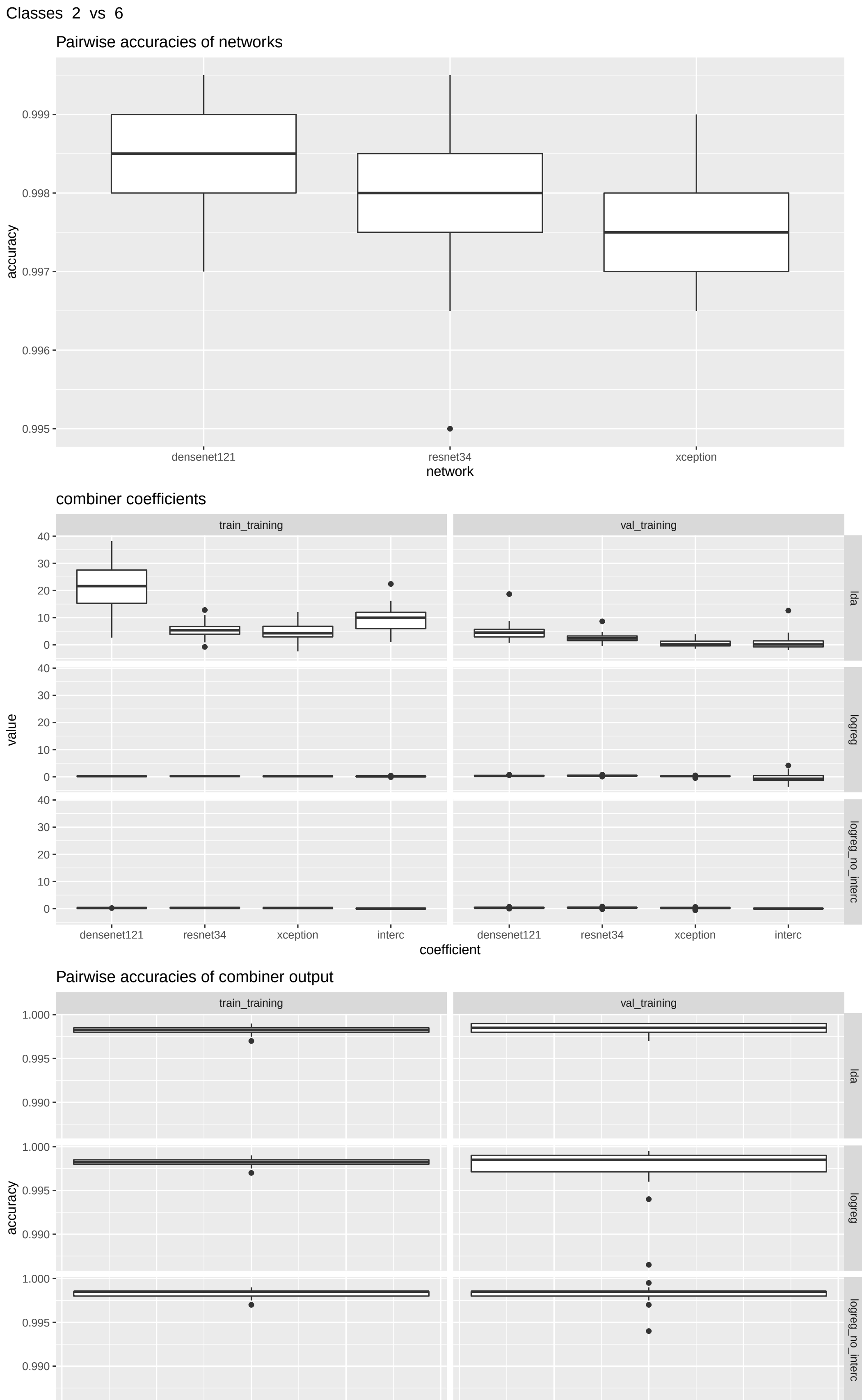

Classes 1 vs 9
Pairwise accuracies of networks
accuracy
0.9875
0.9850
0.9825
0.9800
0.9775
densenet121
combiner coefficients
value
30
20
10
0
densenet121
resnet34
0.985
0.980
0.975

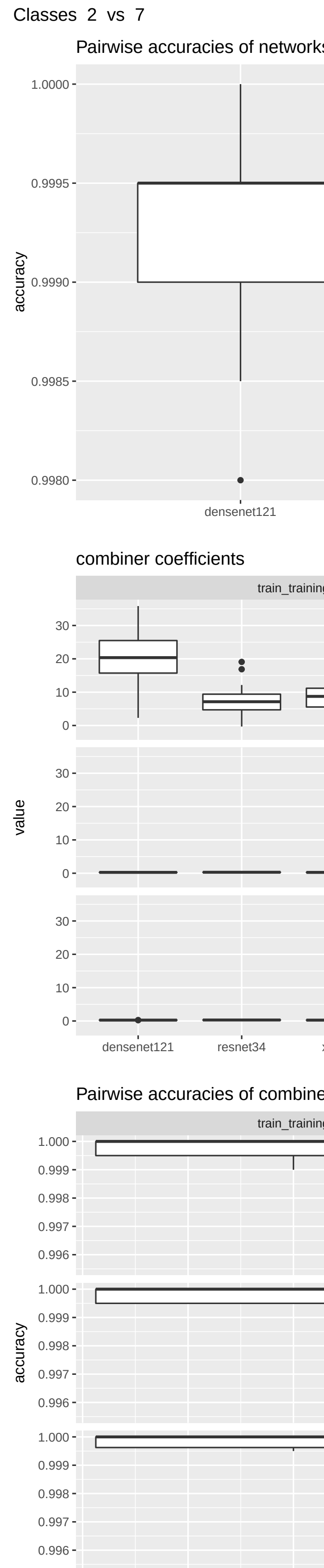

Classes 1 vs 10
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output


Classes 2 vs 3
Pairwise accuracies of networks
accuracy
densenet121
combiner coefficients
value
resnet34

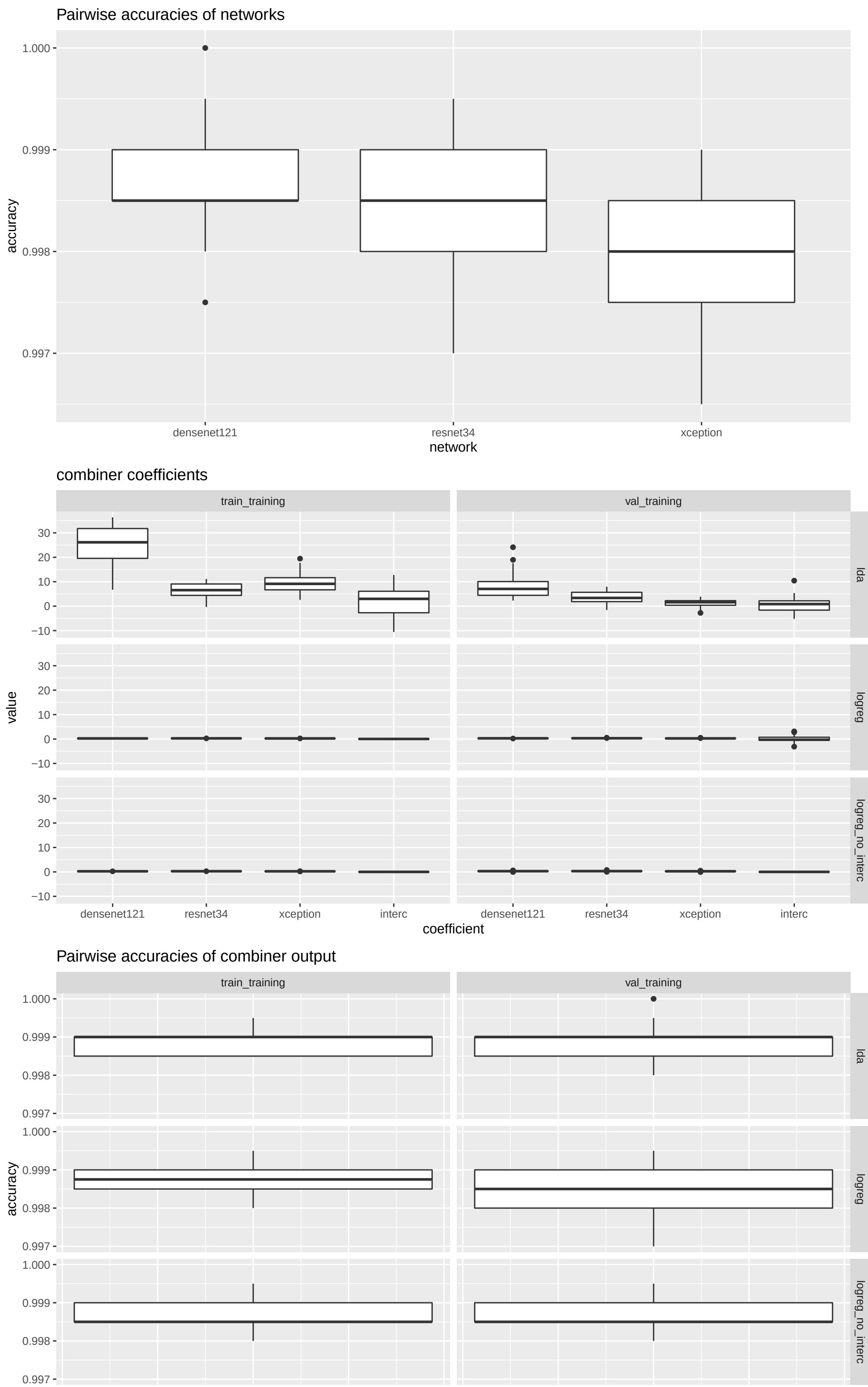

Classes 2 vs 4
Pairwise accuracies of networks
accuracy
0.999
0.998
0.997
0.996
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
30
20
10
0
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy
1.000
0.995
0.990


Classes 2 vs 5
accuracy
1.0000
0.9995
0.9990
0.9985
densenet121
combiner coefficients
value
40
30
20
10
0
−10
densenet121
resnet34
accuracy
1.00000
0.99975
0.99950
0.99925
0.99900

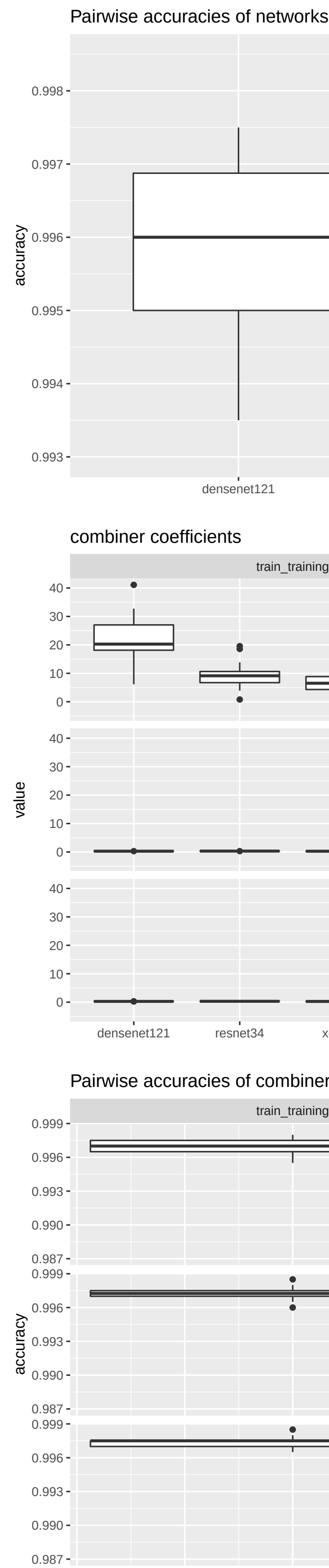

Classes 2 vs 6
Pairwise accuracies of networks
accuracy
0.999
0.998
0.997
0.996
0.995
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
40
30
20
10
0
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
1.000
0.995
0.990

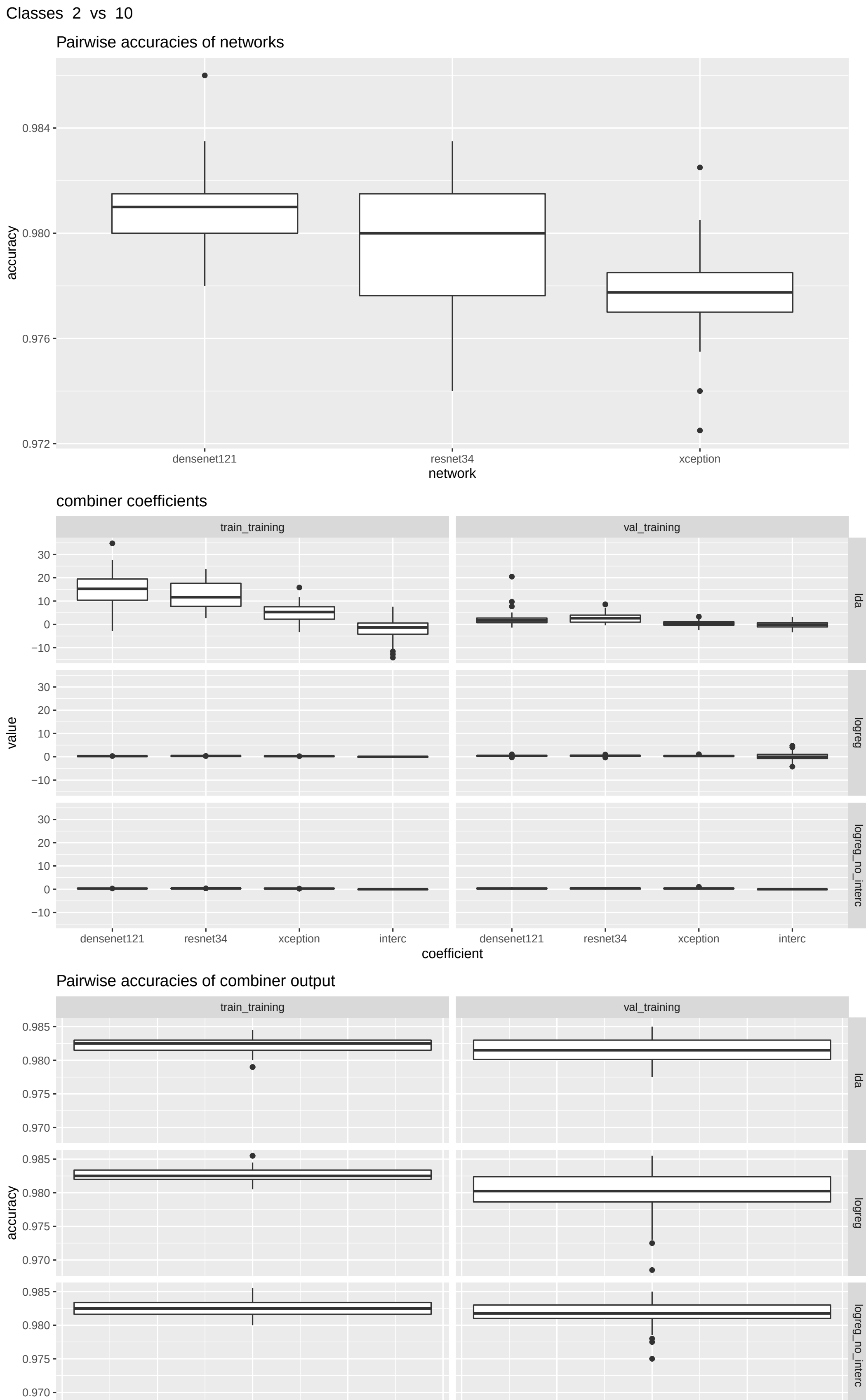

Classes 2 vs 7
Pairwise accuracies of networks
accuracy
1.0000
0.9995
0.9990
0.9985
0.9980
densenet121
combiner coefficients
train_training
value
30
20
10
0
densenet121
resnet34
Pairwise accuracies of combine
1.000
0.999
0.998
0.997
0.996

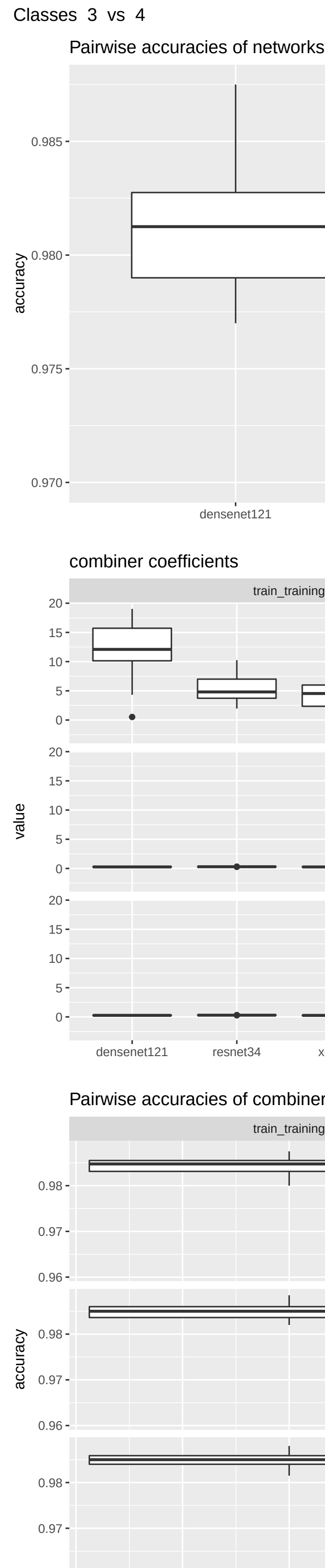
Classes 2 vs 8



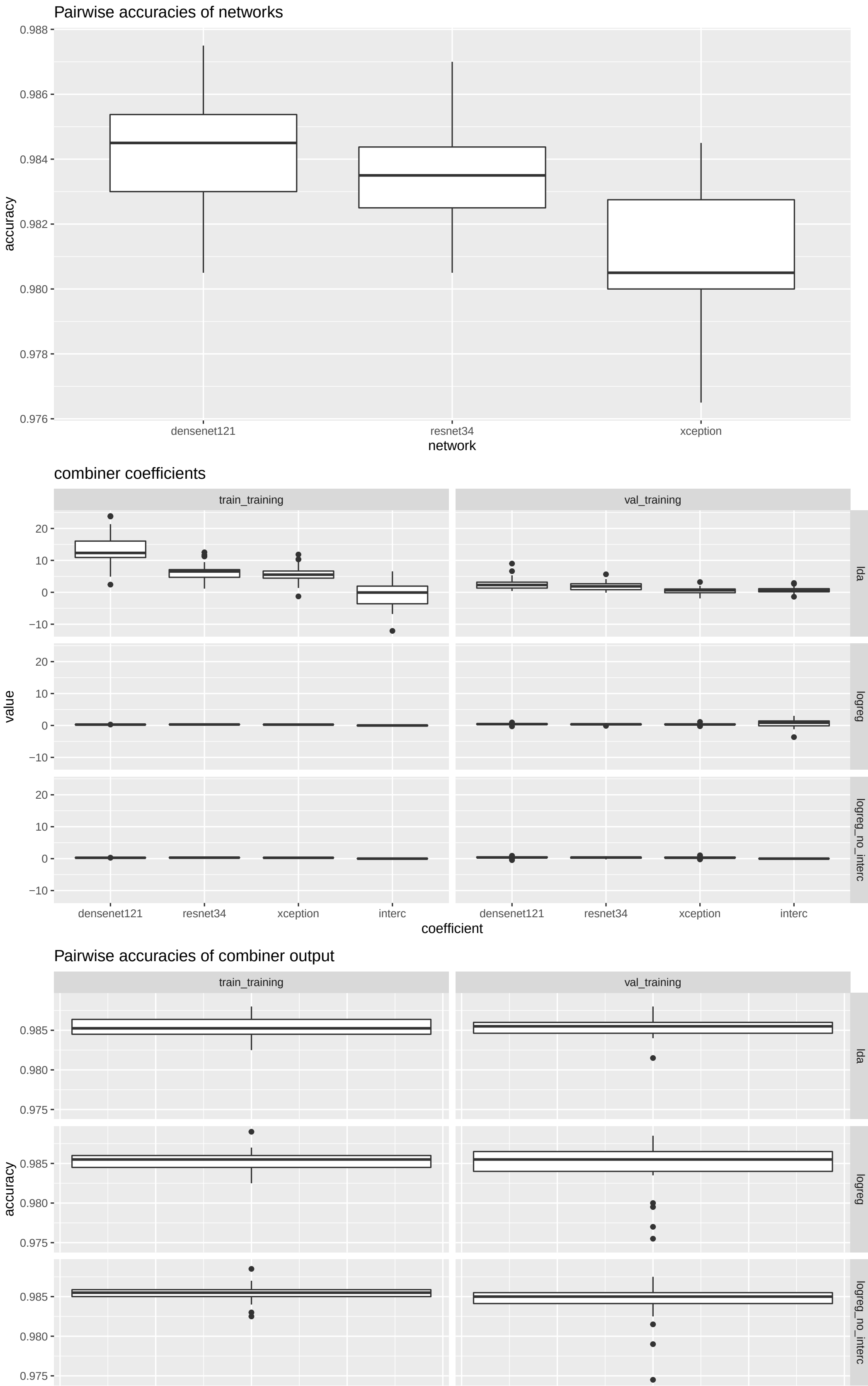
Classes 2 vs 9



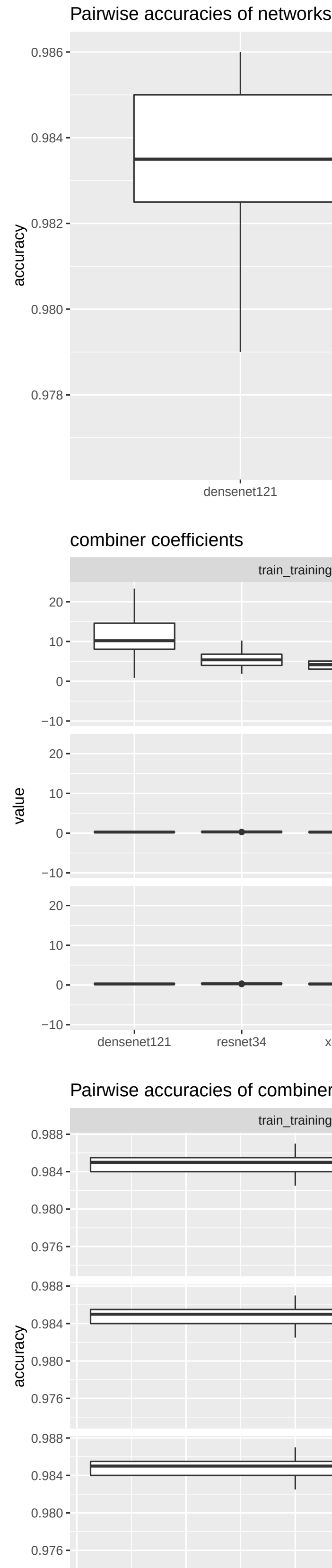

Classes 2 vs 10
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output

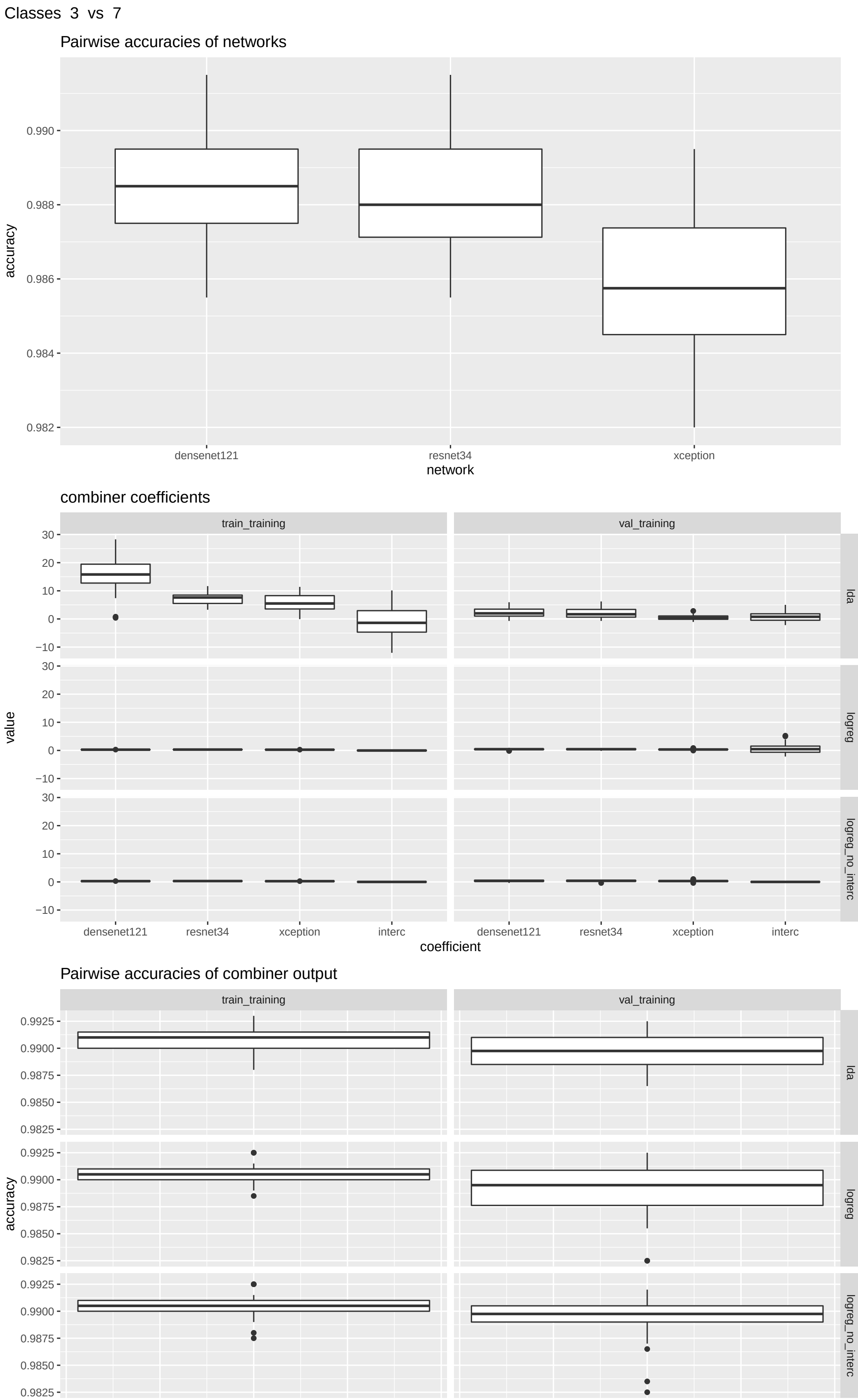

Classes 3 vs 4
Pairwise accuracies of networks
accuracy
densenet121
combiner coefficients
train_training
value
resnet34
Pairwise accuracies of combiner

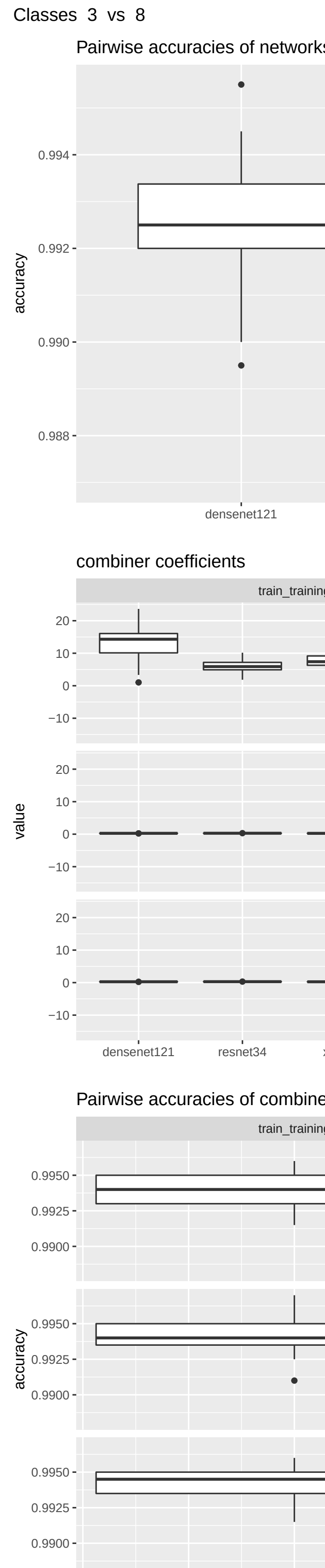
Classes 3 vs 5

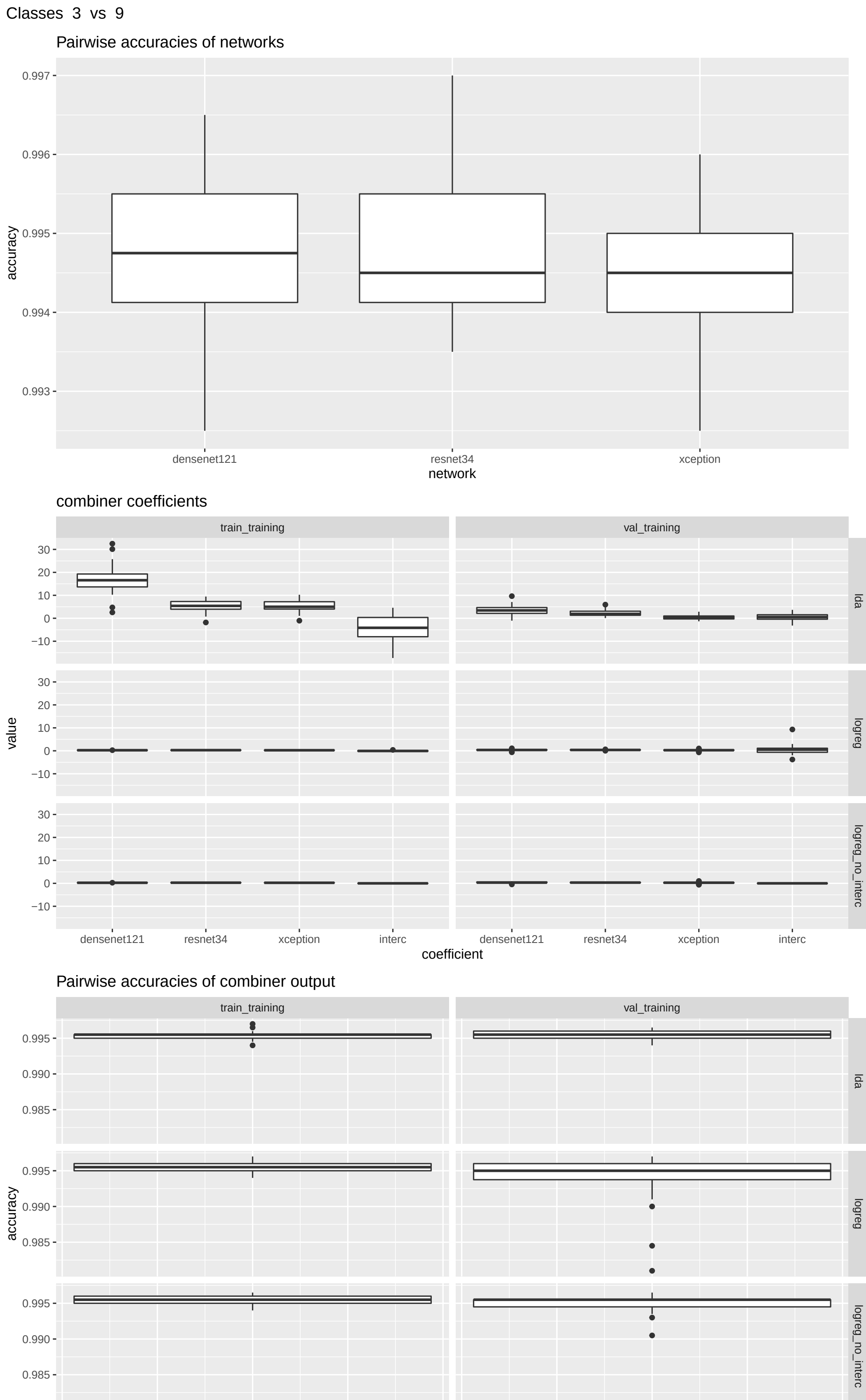


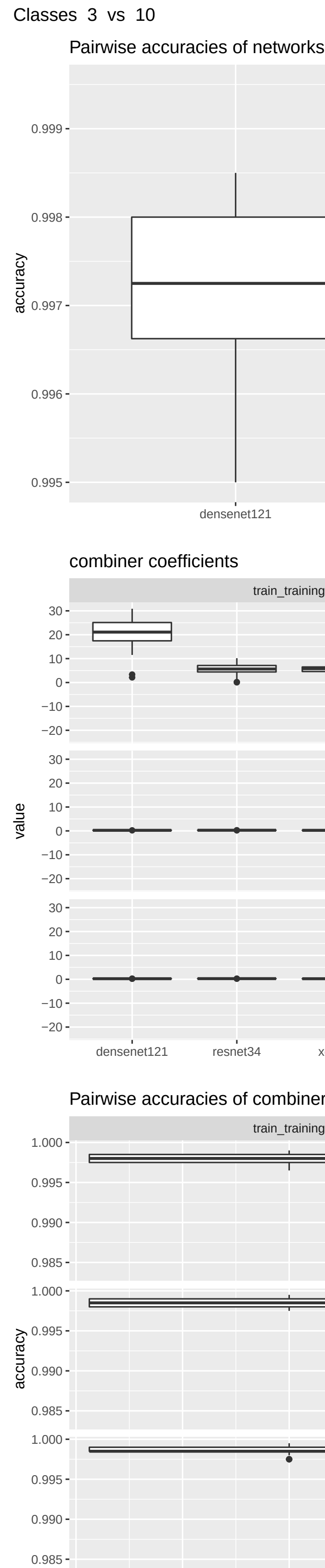
Classes 3 vs 6

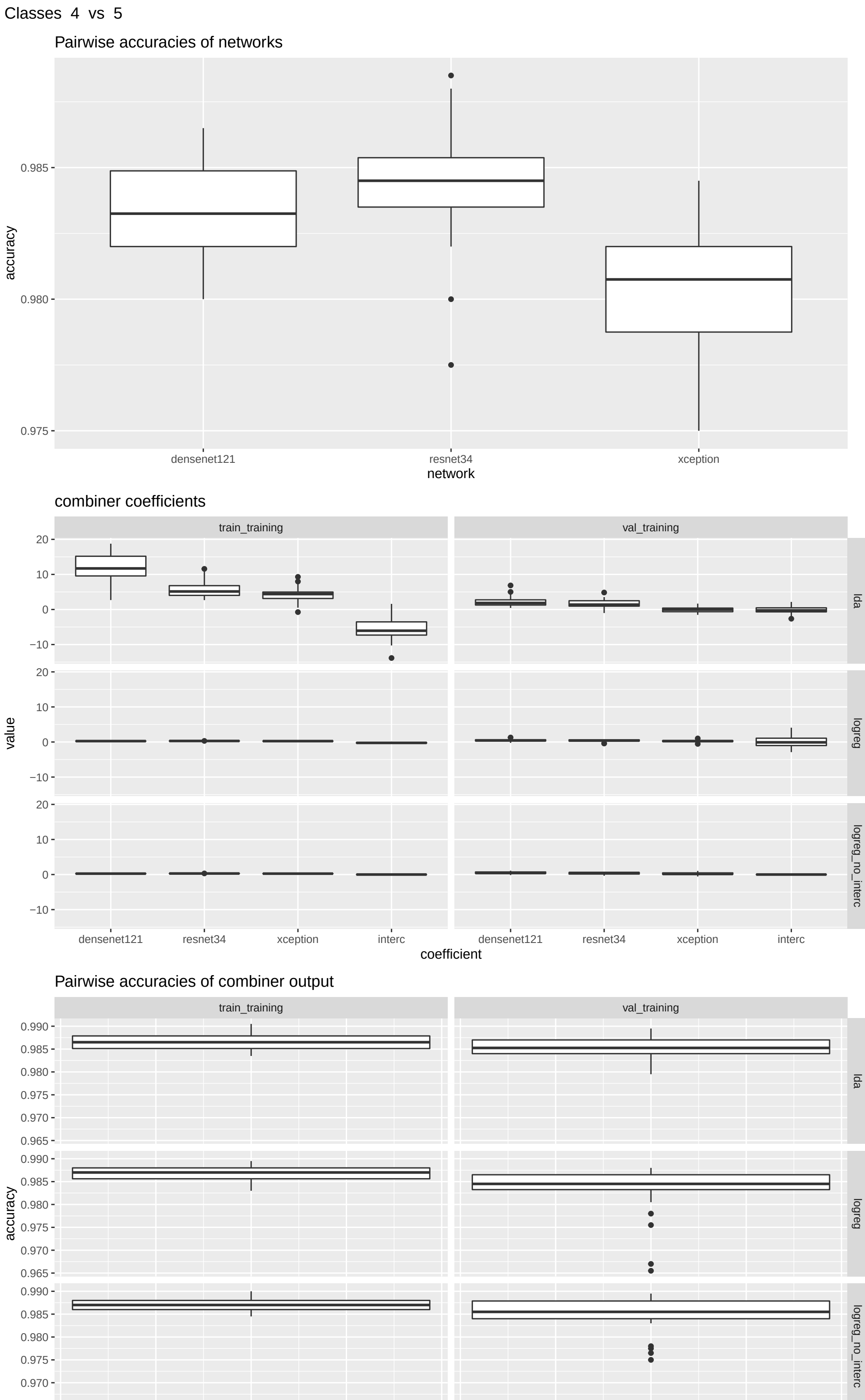


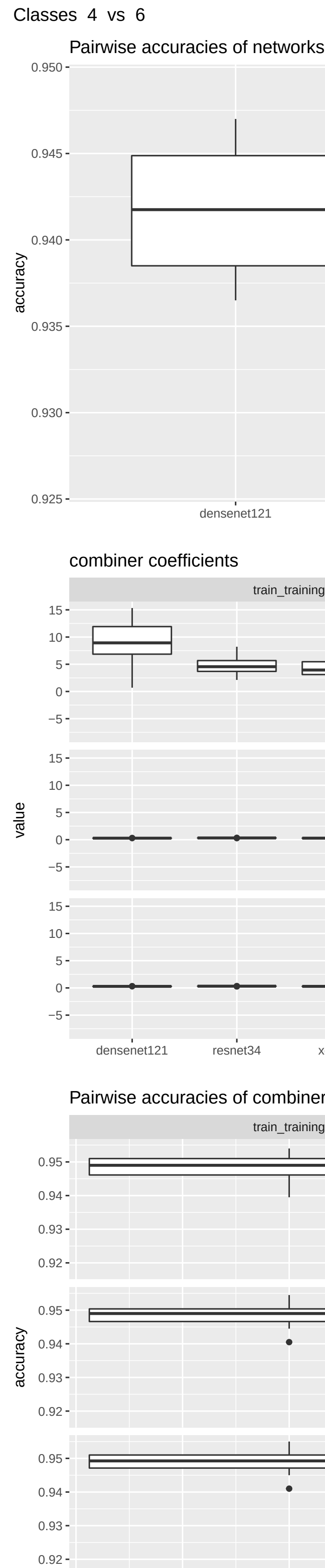

Classes 3 vs 7
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output

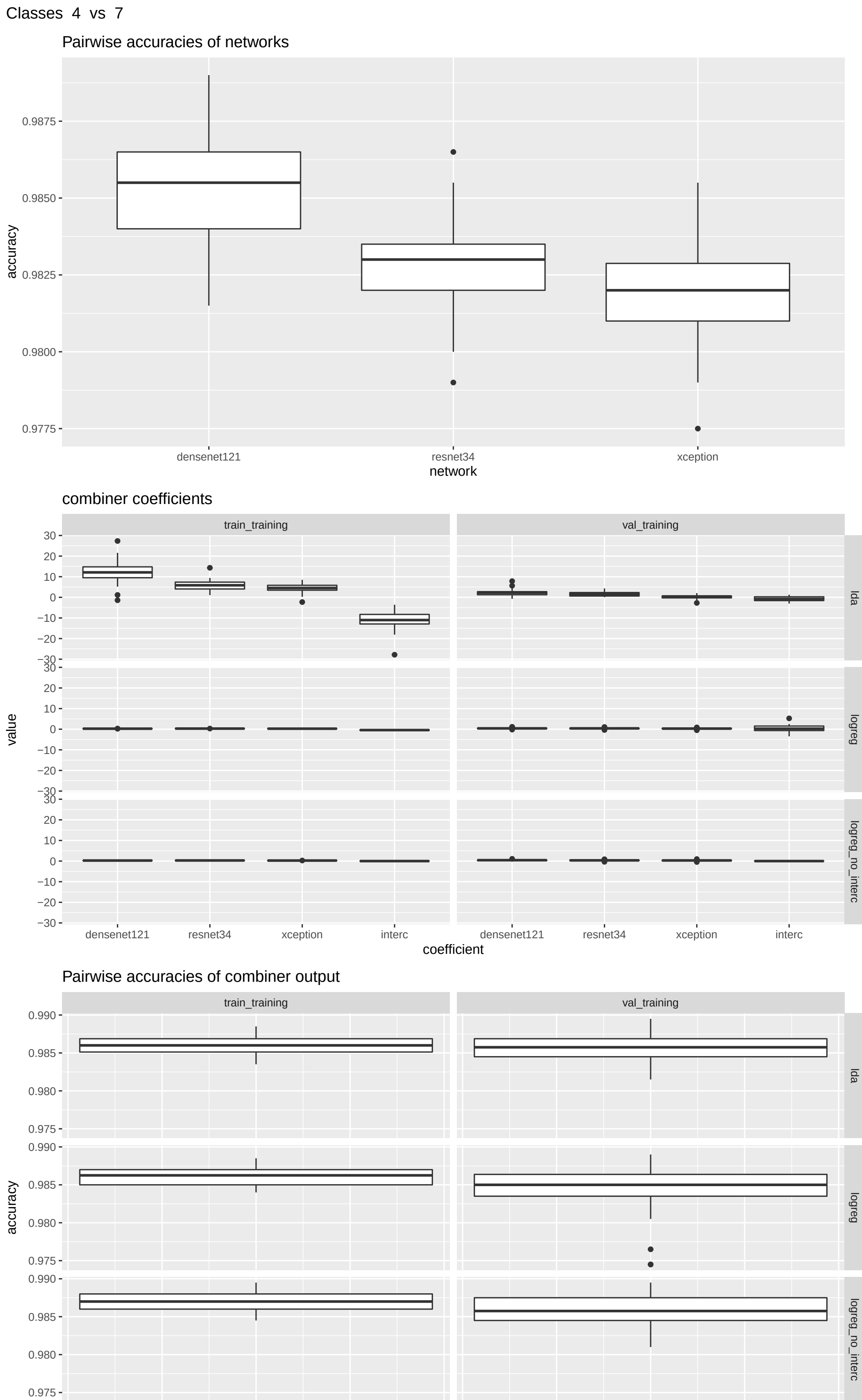

Classes 3 vs 8
combiner coefficients
accuracy
value
densenet121
resnet34


Classes 3 vs 9
Pairwise accuracies of networks
accuracy
0.997
0.996
0.995
0.994
0.993
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
30
20
10
0
−10
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
accuracy
0.995
0.990
0.985

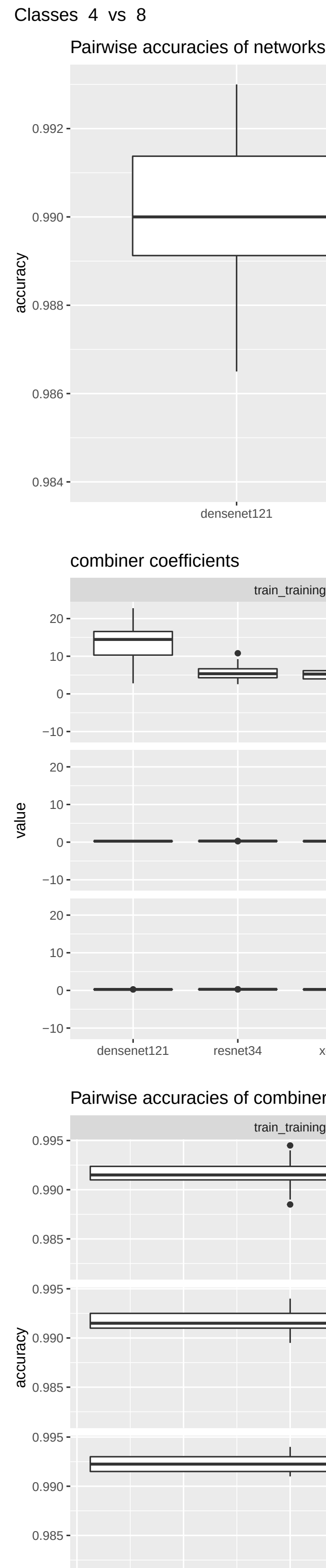

Classes 3 vs 10
Pairwise accuracies of networks
accuracy
0.999
0.998
0.997
0.996
0.995
densenet121
combiner coefficients
train_training
value
30
20
10
0
−10
−20
densenet121
resnet34
Pairwise accuracies of combiner
1.000
0.995
0.990
0.985


Classes 4 vs 5
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output


Classes 4 vs 6
Pairwise accuracies of networks
accuracy
densenet121
combiner coefficients
train_training
value
resnet34

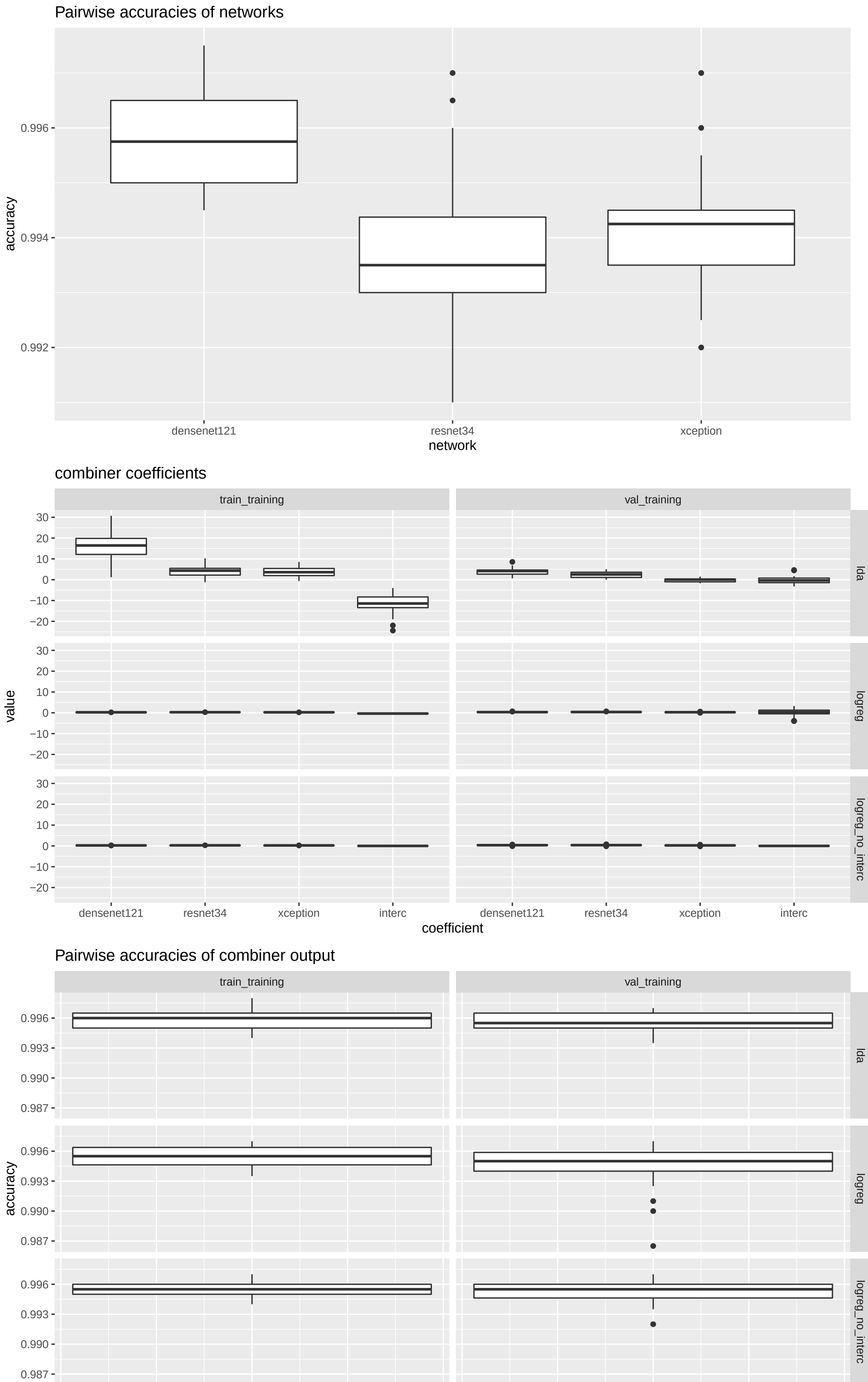


Classes 4 vs 7
Pairwise accuracies of networks
accuracy
0.9875
0.9850
0.9825
0.9800
0.9775
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
30
20
10
0
−10
−20
−30
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.990
0.985
0.980
0.975

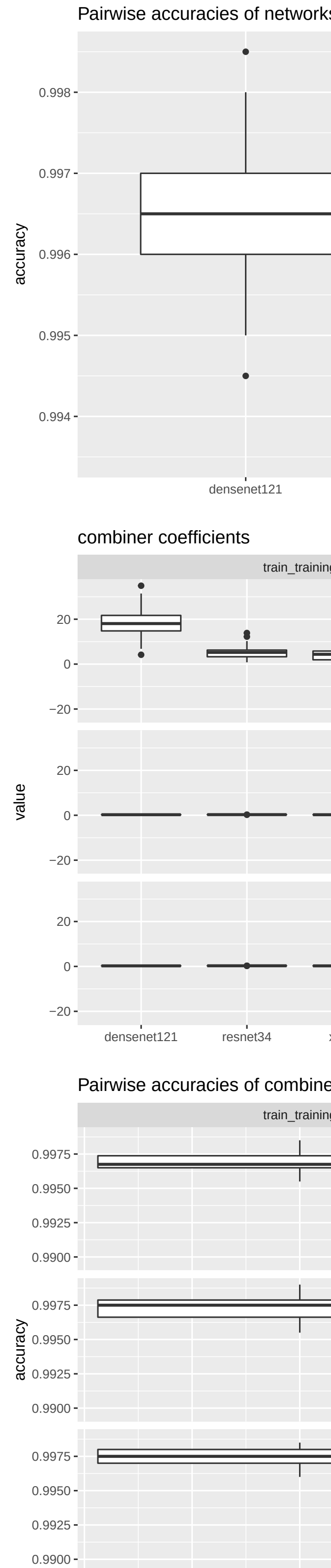


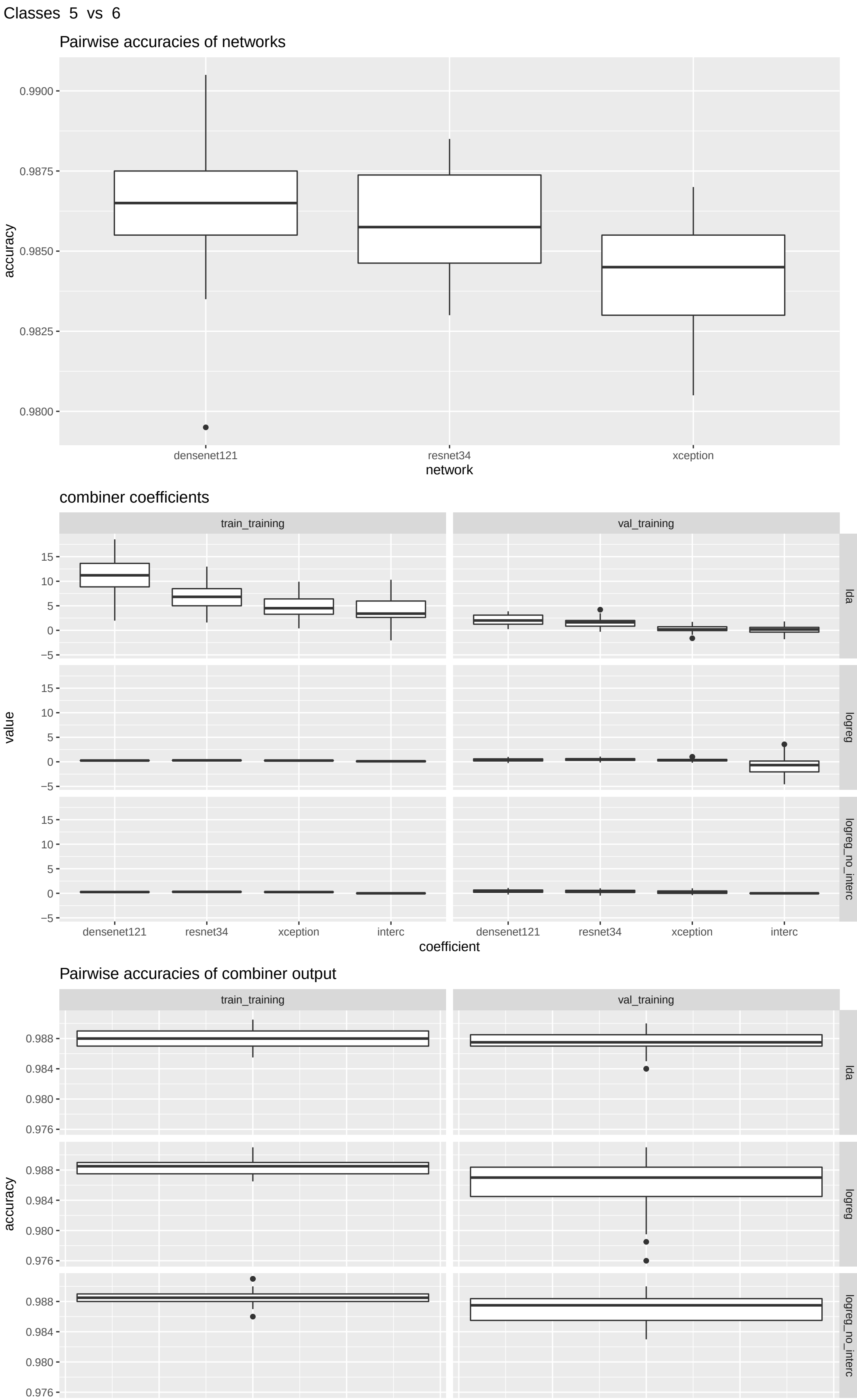
Classes 4 vs 8
Pairwise accuracies of networks
accuracy
0.992
0.990
0.988
0.986
0.984
densenet121
combiner coefficients
train_training
value
20
10
0
−10
densenet121
resnet34
Pairwise accuracies of combiner
0.995
0.990
0.985

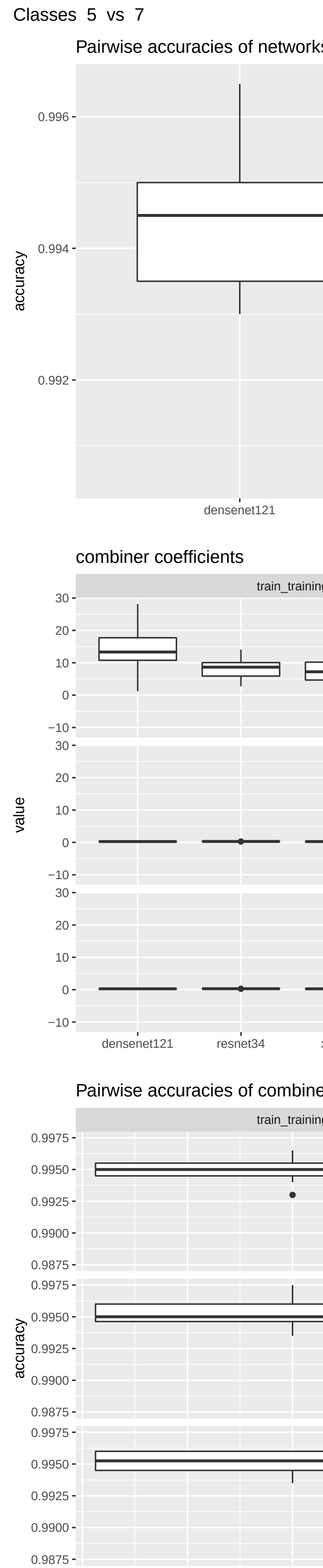
Classes 4 vs 9

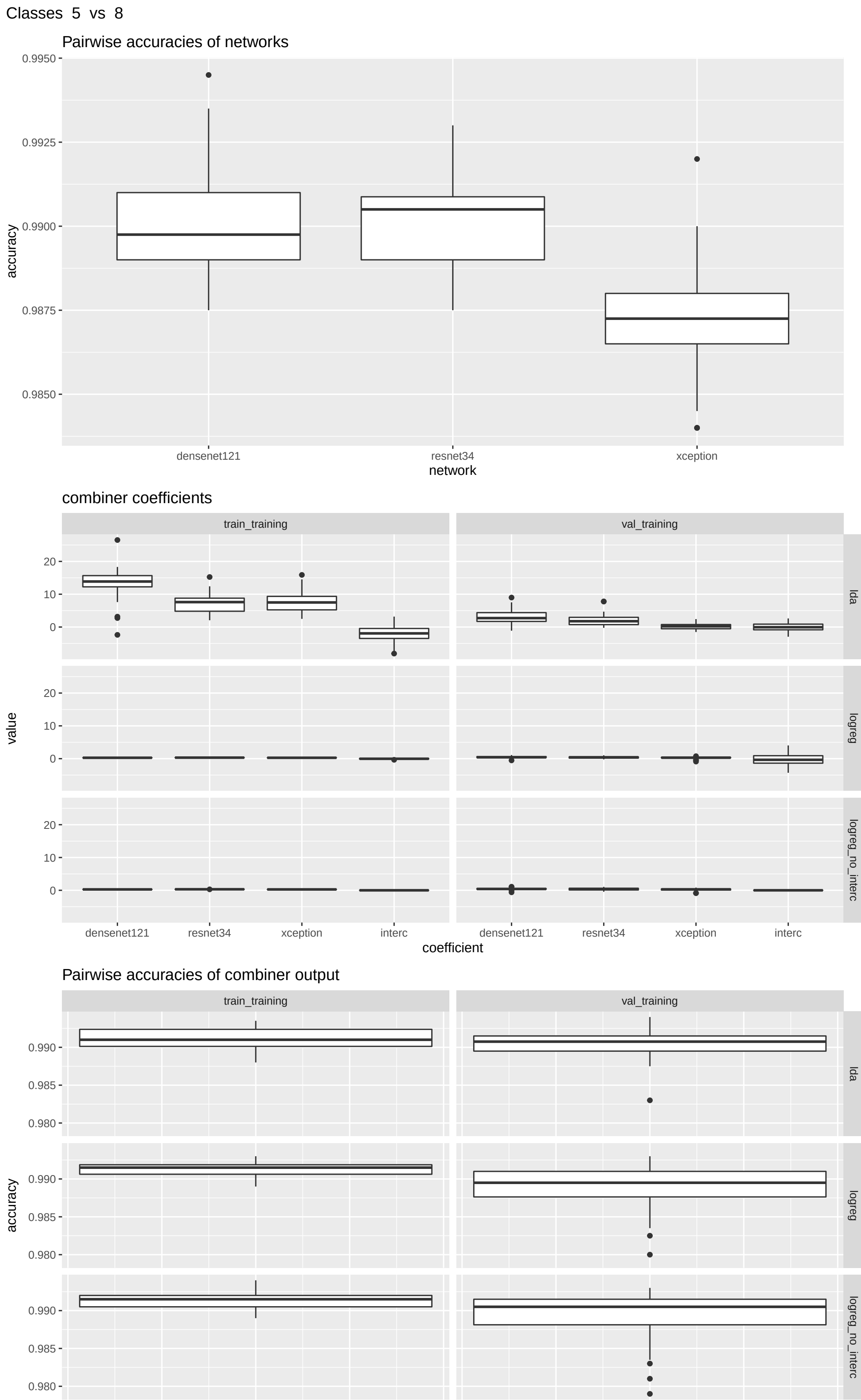


Classes 4 vs 10

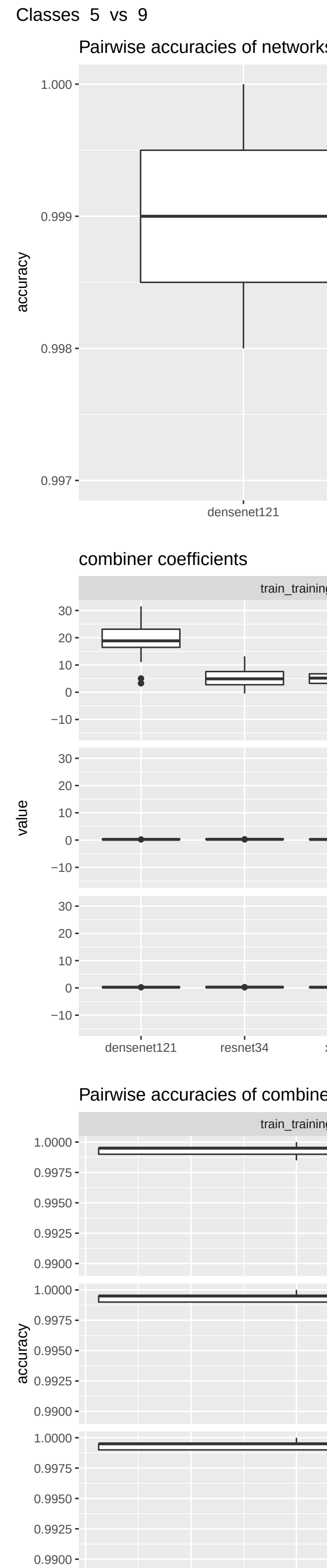



Classes 5 vs 6
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output


Classes 5 vs 7
Pairwise accuracies of networks
accuracy
densenet121
combiner coefficients
train_training
value
resnet34
Pairwise accuracies of combiner output

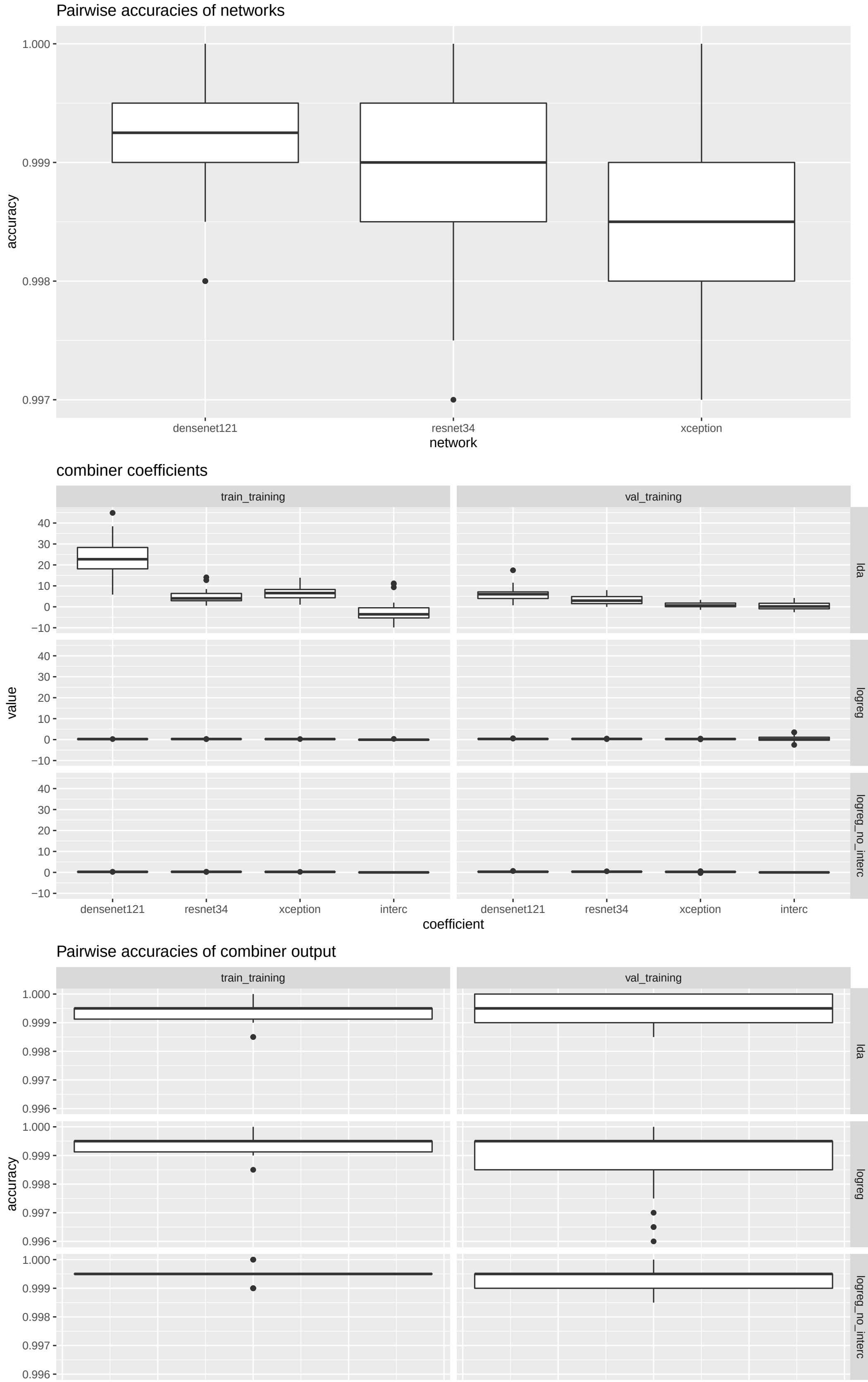


Classes 5 vs 8
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output

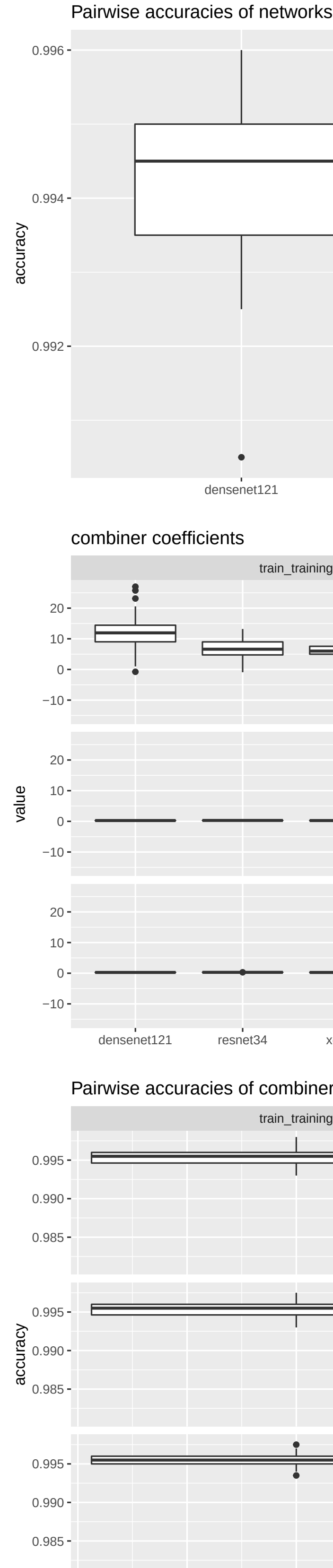


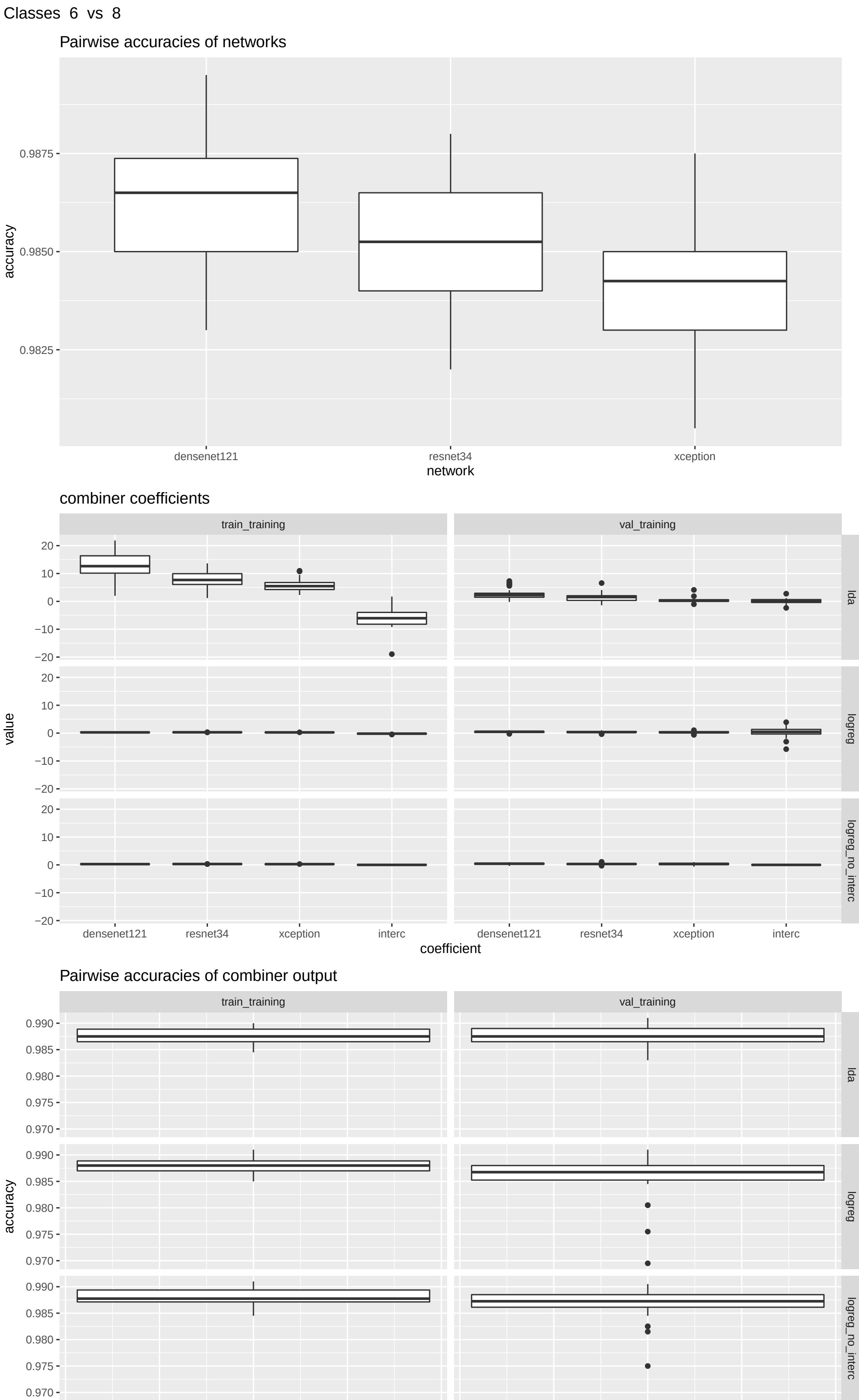
Classes 5 vs 9
accuracy
densenet121
combiner coefficients
value
resnet34

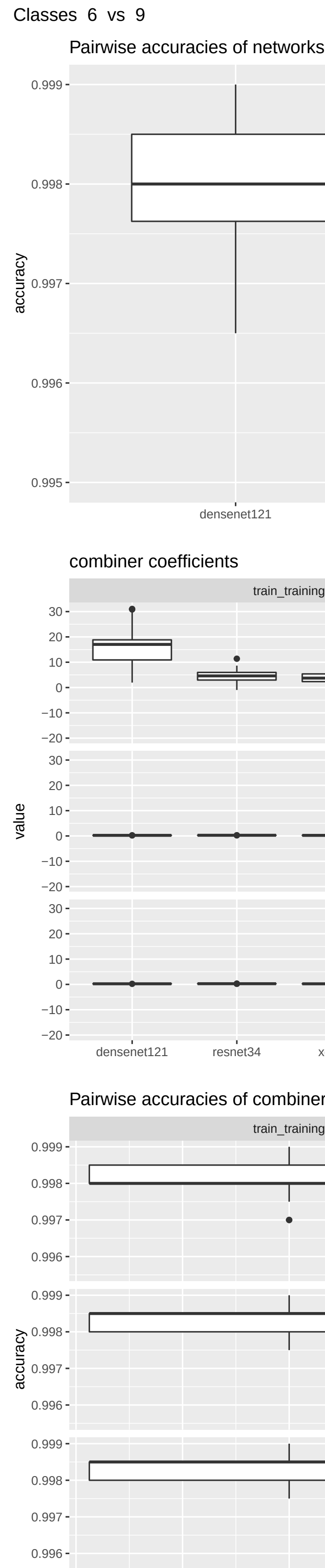
Classes 5 vs 10



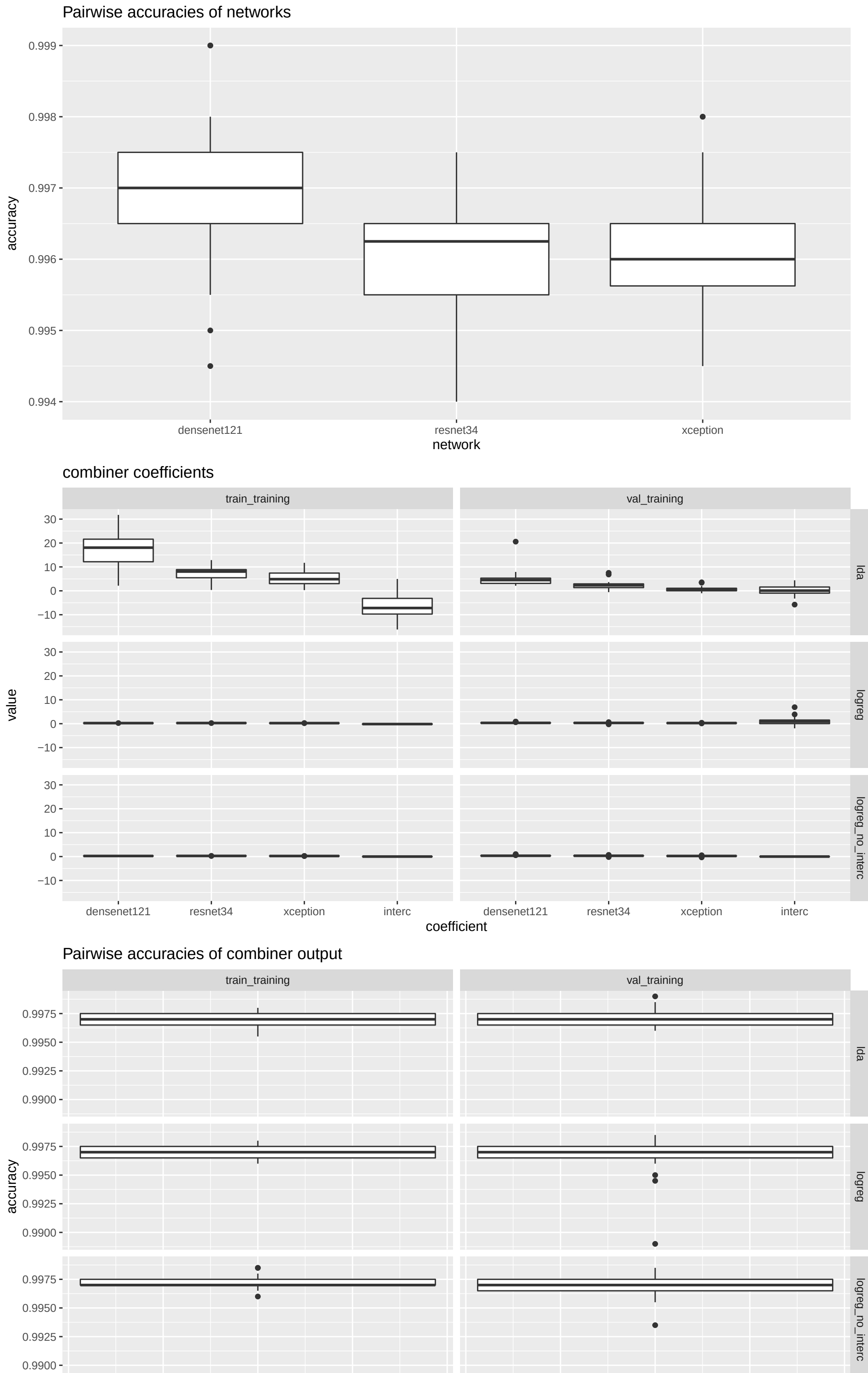
Classes 6 vs 7



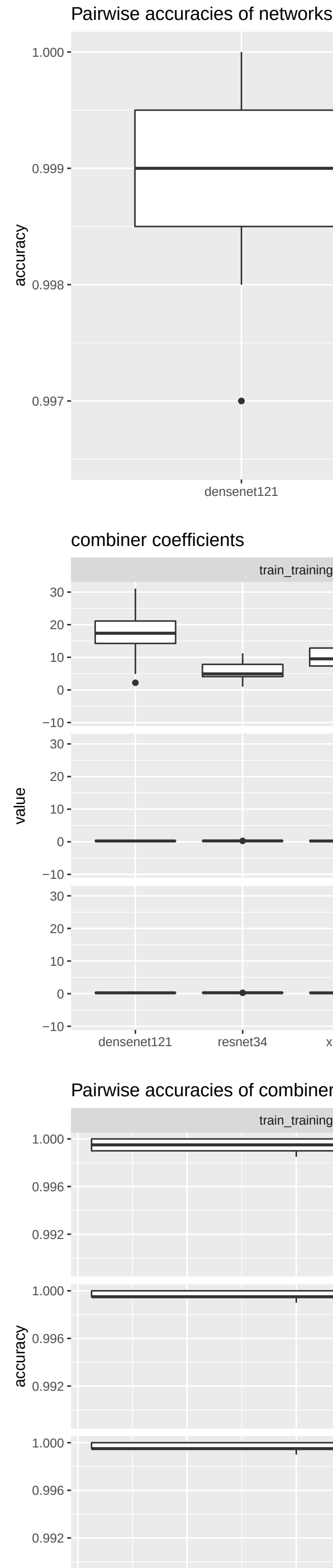

Classes 6 vs 8
Pairwise accuracies of networks
accuracy
0.9875
0.9850
0.9825
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
20
10
0
−10
−20
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
accuracy
0.990
0.985
0.980
0.975
0.970
lda
logreg
logreg_no_interc

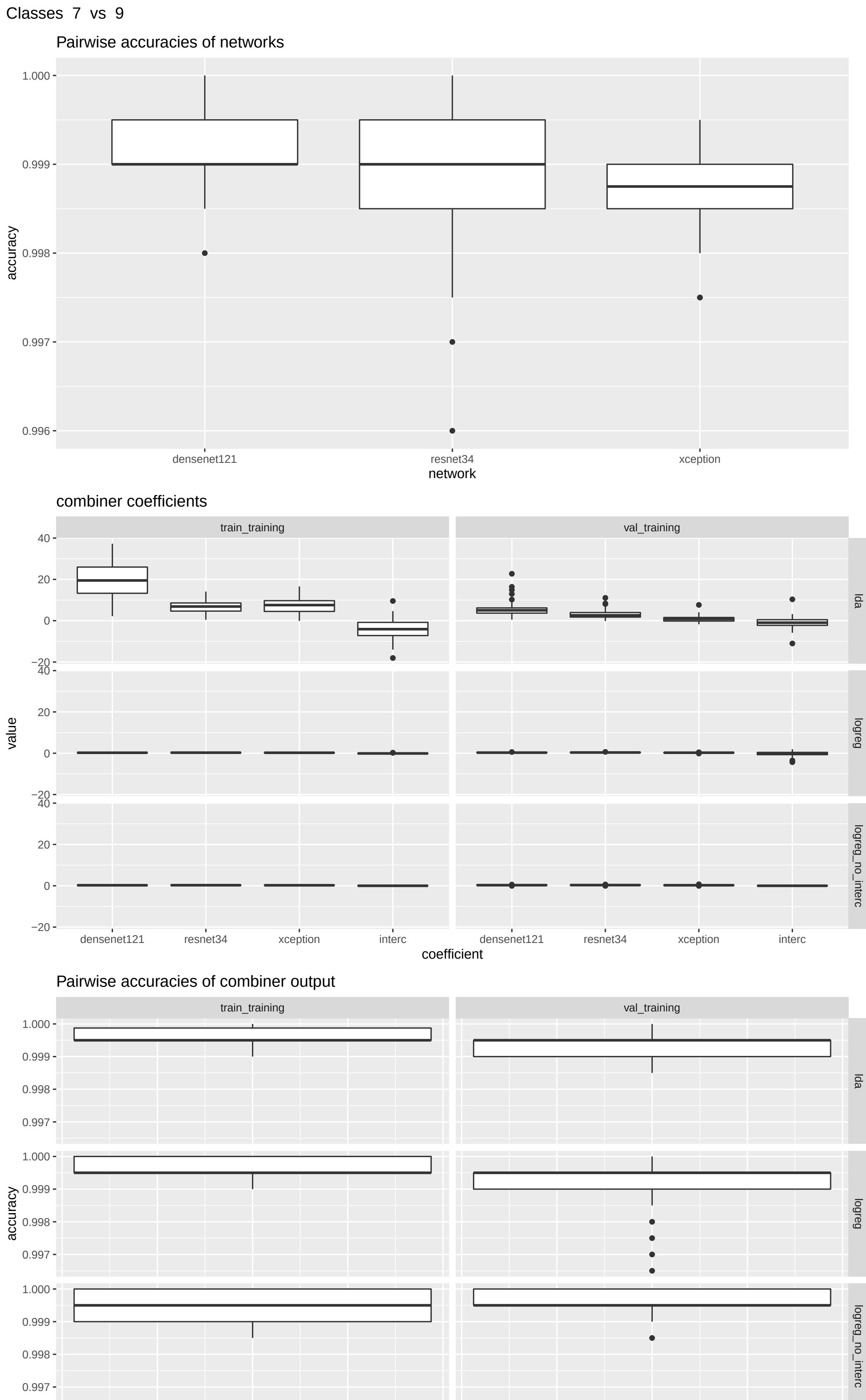

Classes 6 vs 9
Pairwise accuracies of networks
accuracy
0.999
0.998
0.997
0.996
0.995
densenet121
combiner coefficients
train_training
value
30
20
10
0
−10
−20
densenet121
resnet34
Pairwise accuracies of combiner
train_training
accuracy
0.999
0.998
0.997
0.996

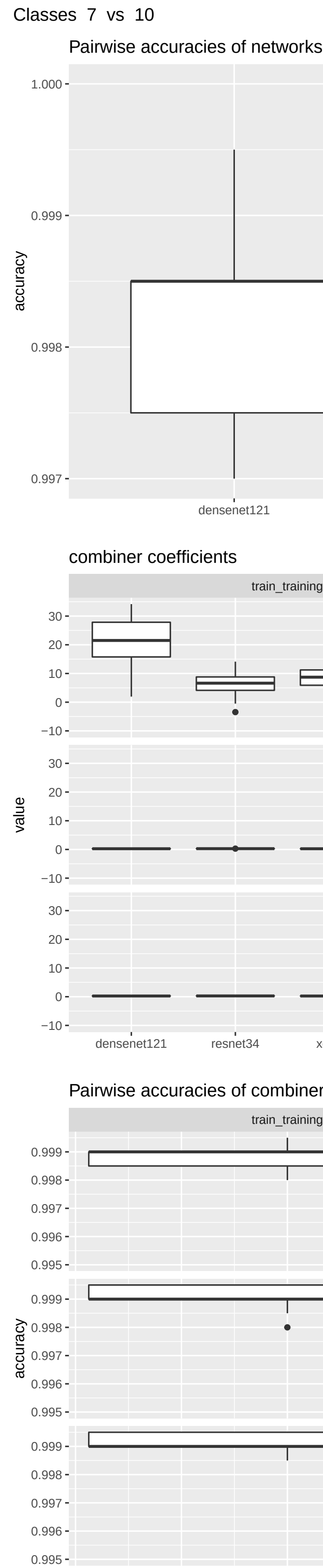
Classes 6 vs 10



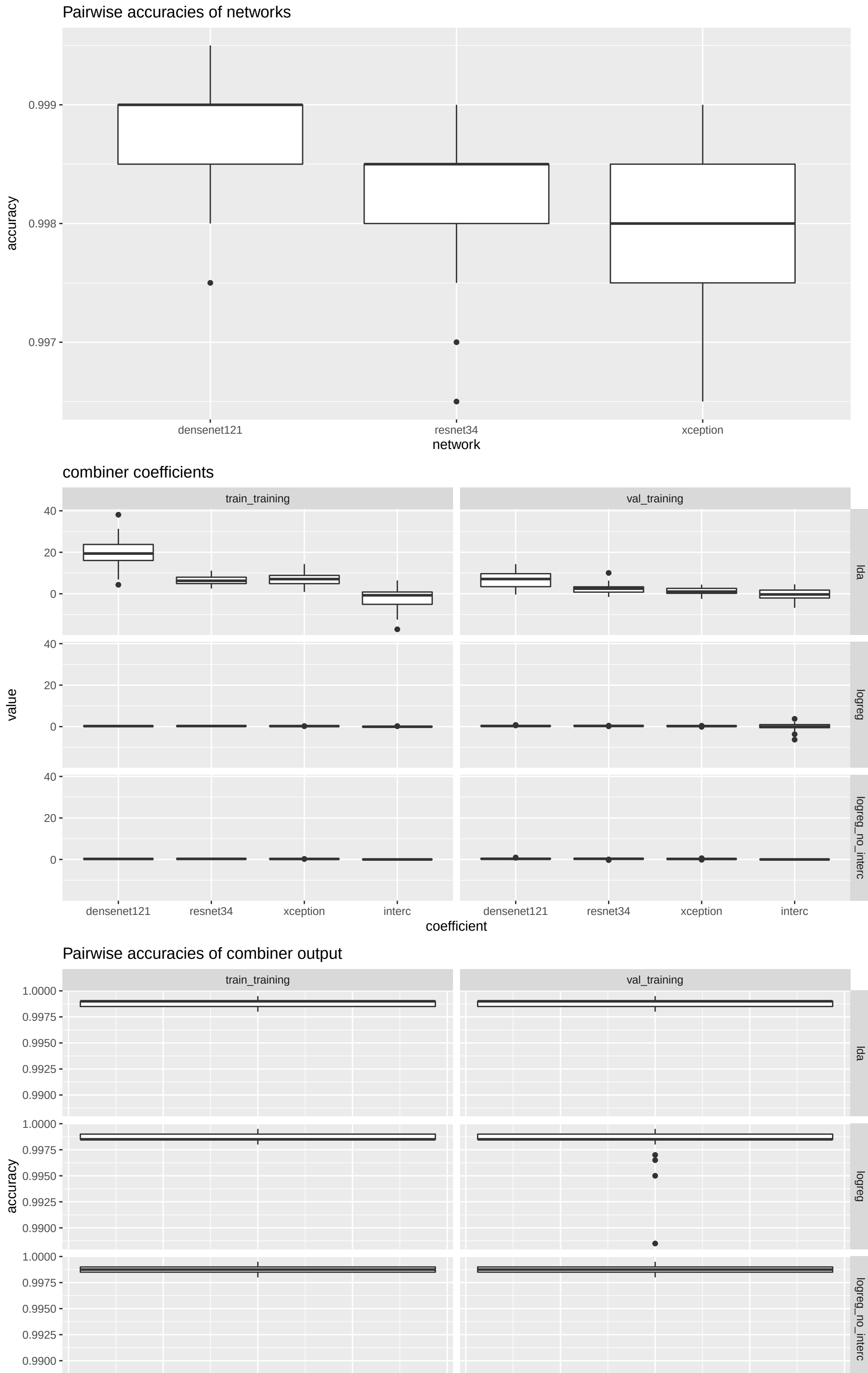
Classes 7 vs 8



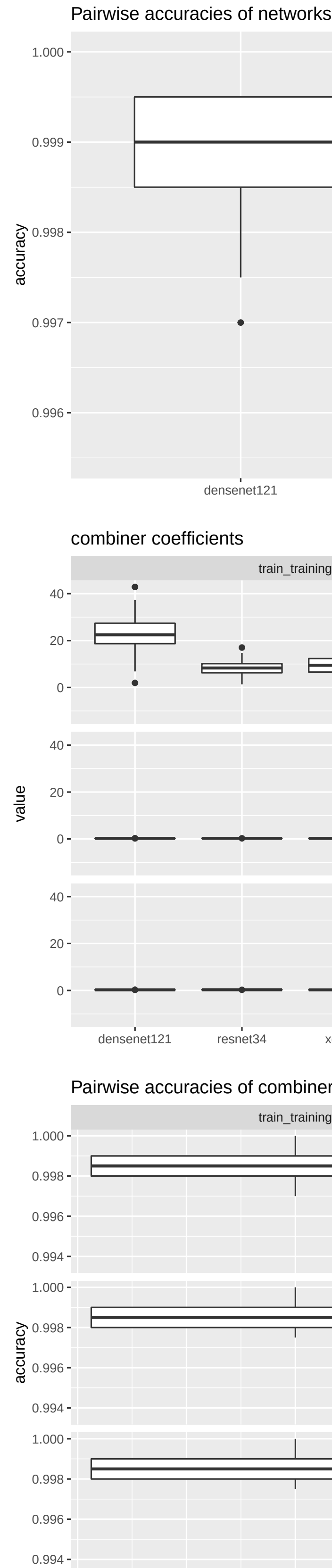

Classes 7 vs 9
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output


Classes 7 vs 10
Pairwise accuracies of networks
accuracy
densenet121
combiner coefficients
train_training
value
resnet34
Pairwise accuracies of combiner
train_training
accuracy

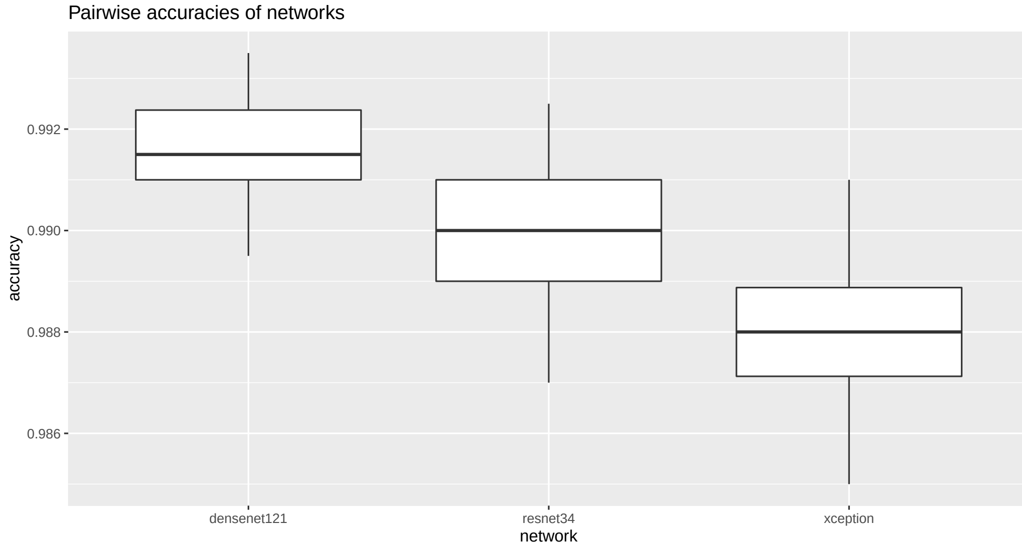
Classes 8 vs 9

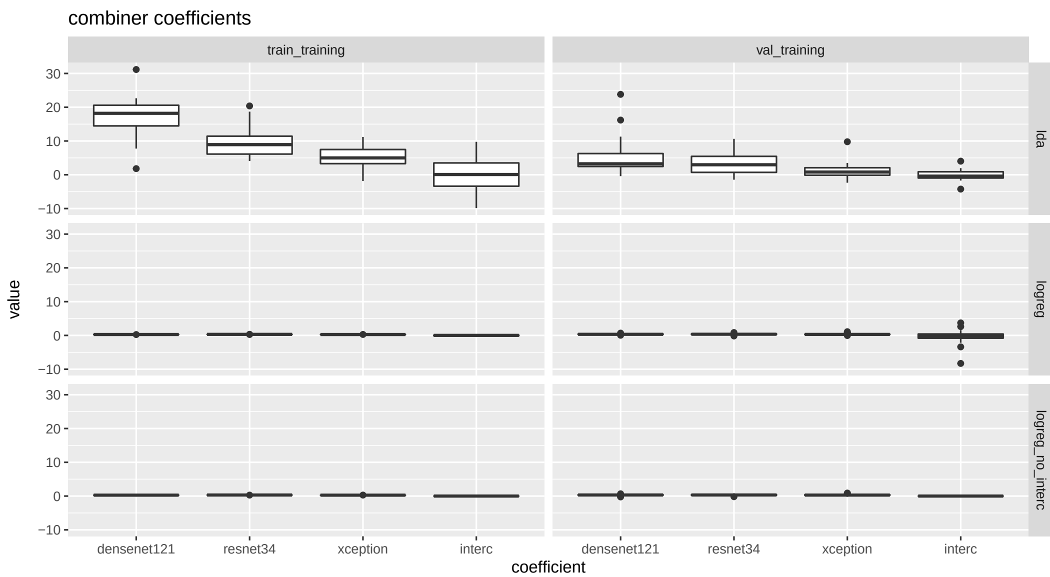


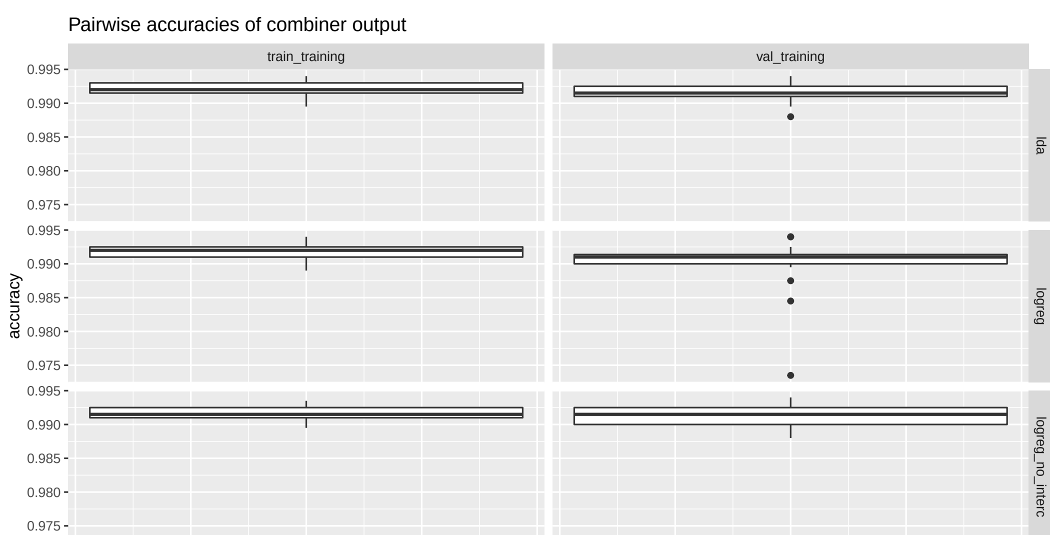
Classes 8 vs 10



Classes 9 vs 10





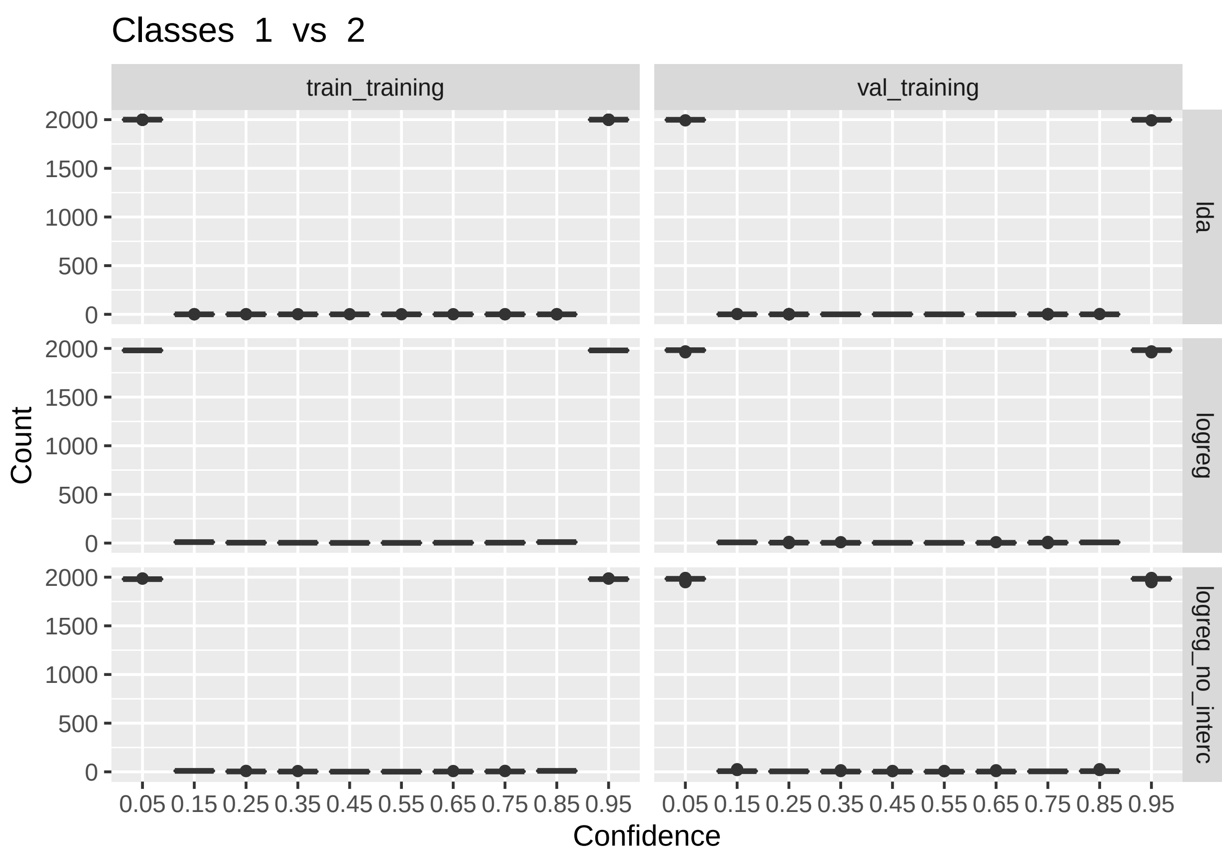


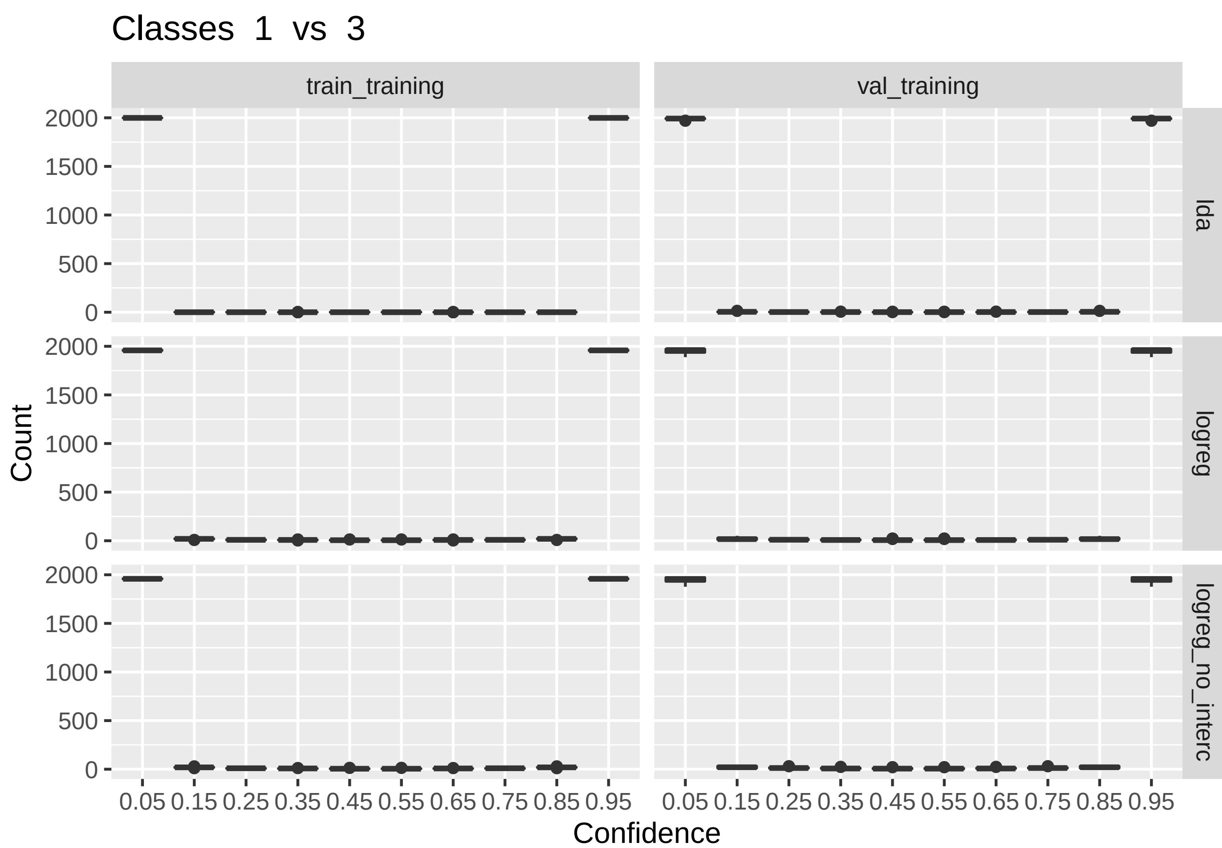
Val training

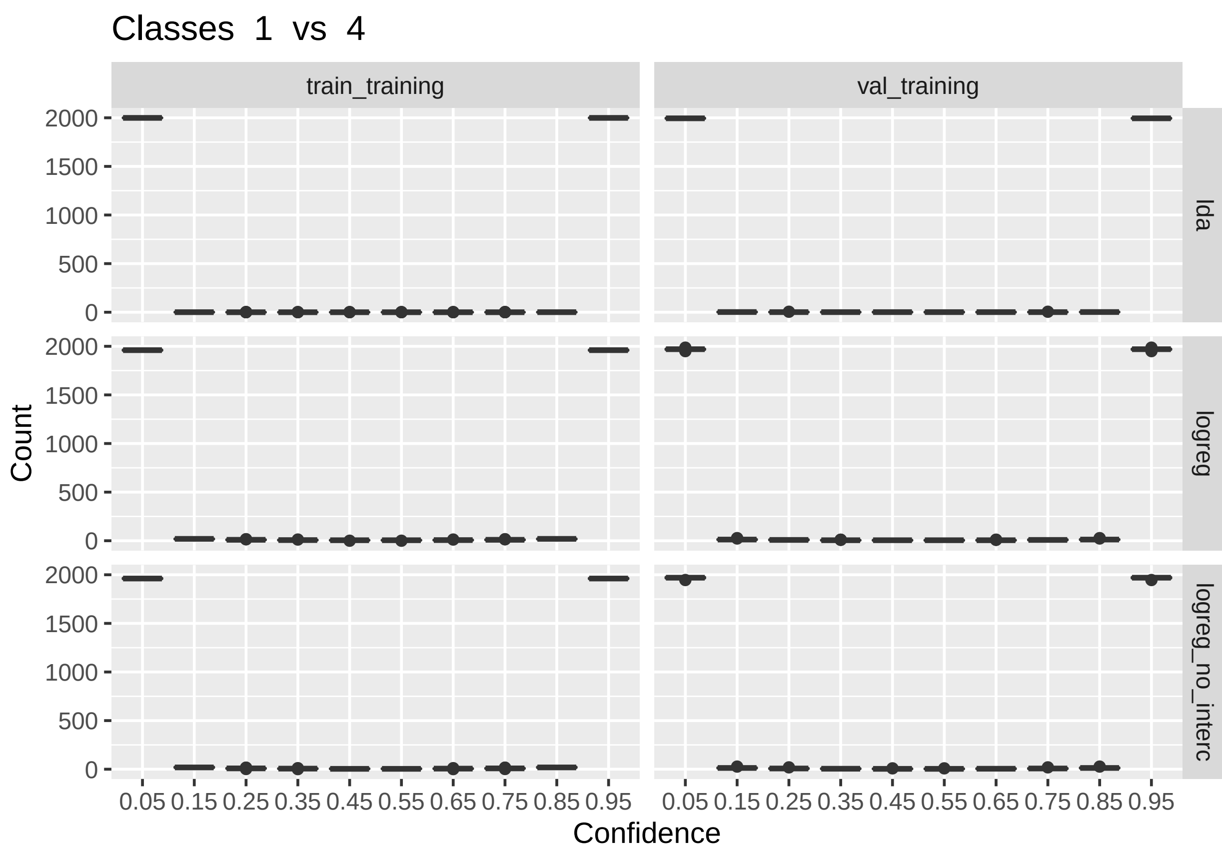
accuracies have higher variance.

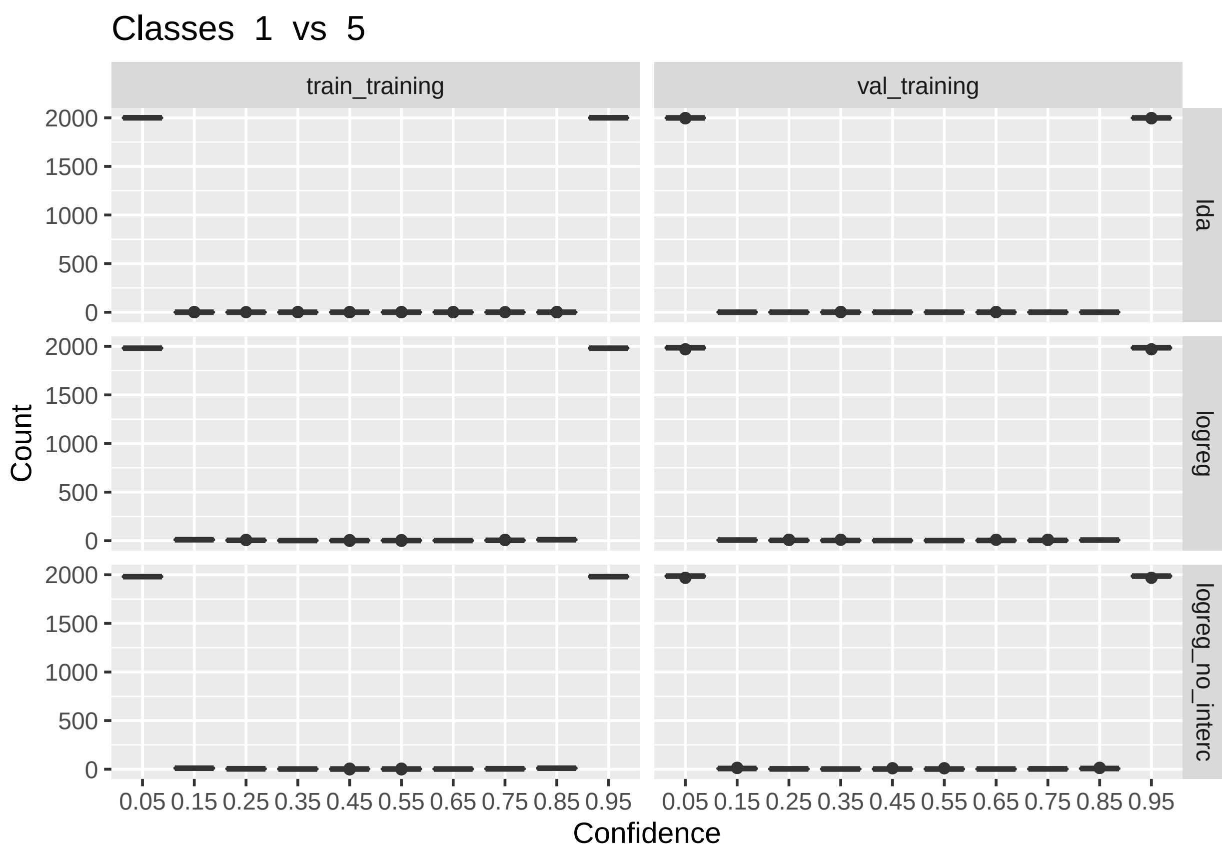
An attempt to visualize calibration of LDAs. Almost all elements are present near the zero and one probabilities, therefore the visualization is not very informative.

```
for (cl1 in 1:(classes - 1))
{
  for (cl2 in (cl1 + 1):classes)
  {
    cal_count_plt <- ens_pw_cal %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot() +
      geom_boxplot(mapping=aes(x=bin_c, y=bin_count)) + facet_grid(cols=vars(train_set), rows=vars(comb
      ggtitle(paste("Classes ", cl1, " vs ", cl2)) + xlab("Confidence") + ylab("Count")

    print(cal_count_plt)
  }
}
```

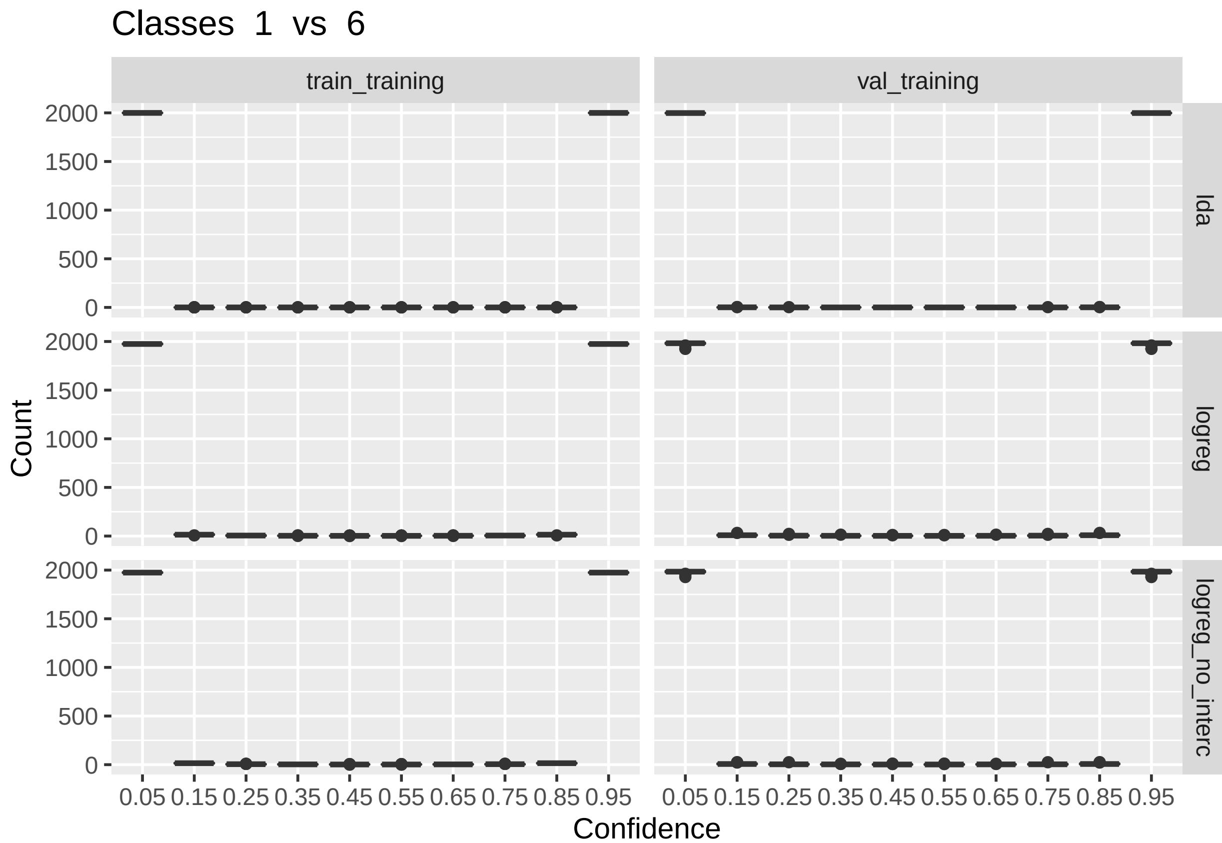


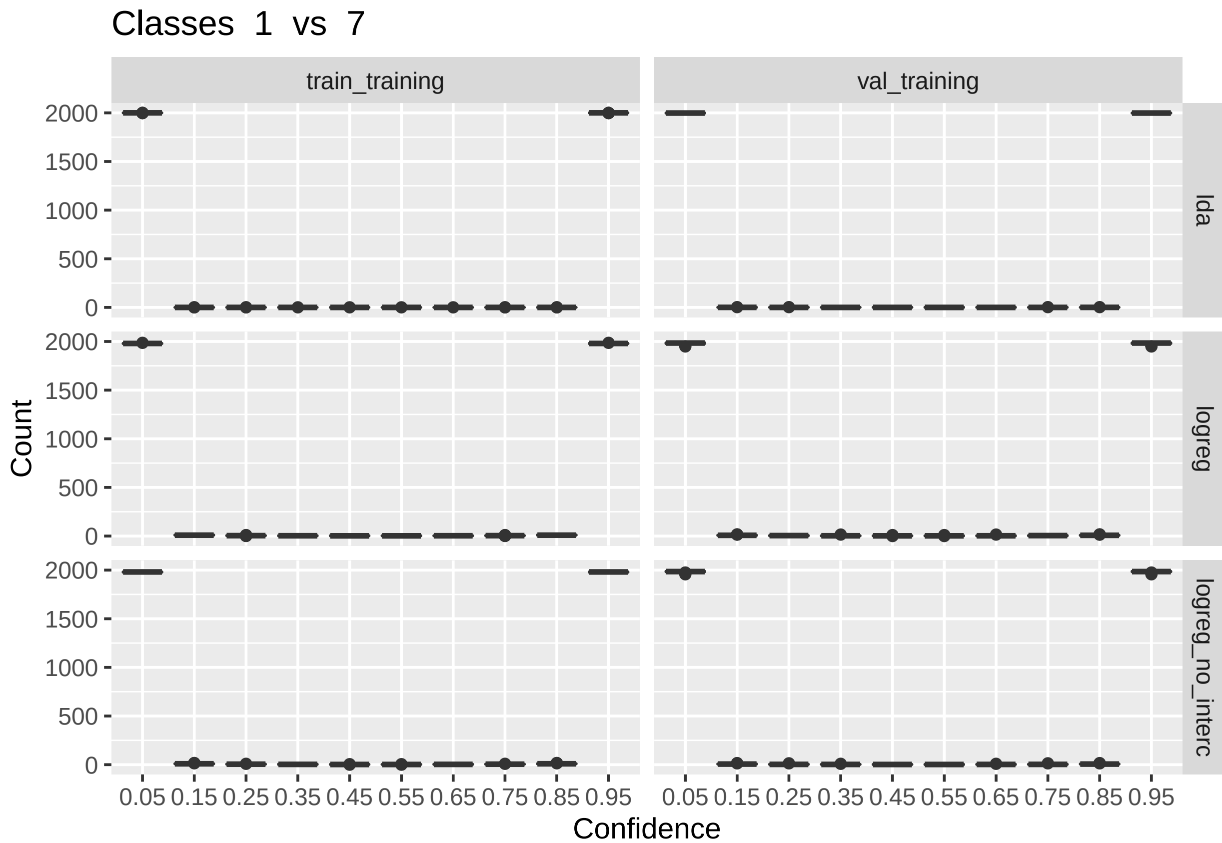

Classes 1 vs 3
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

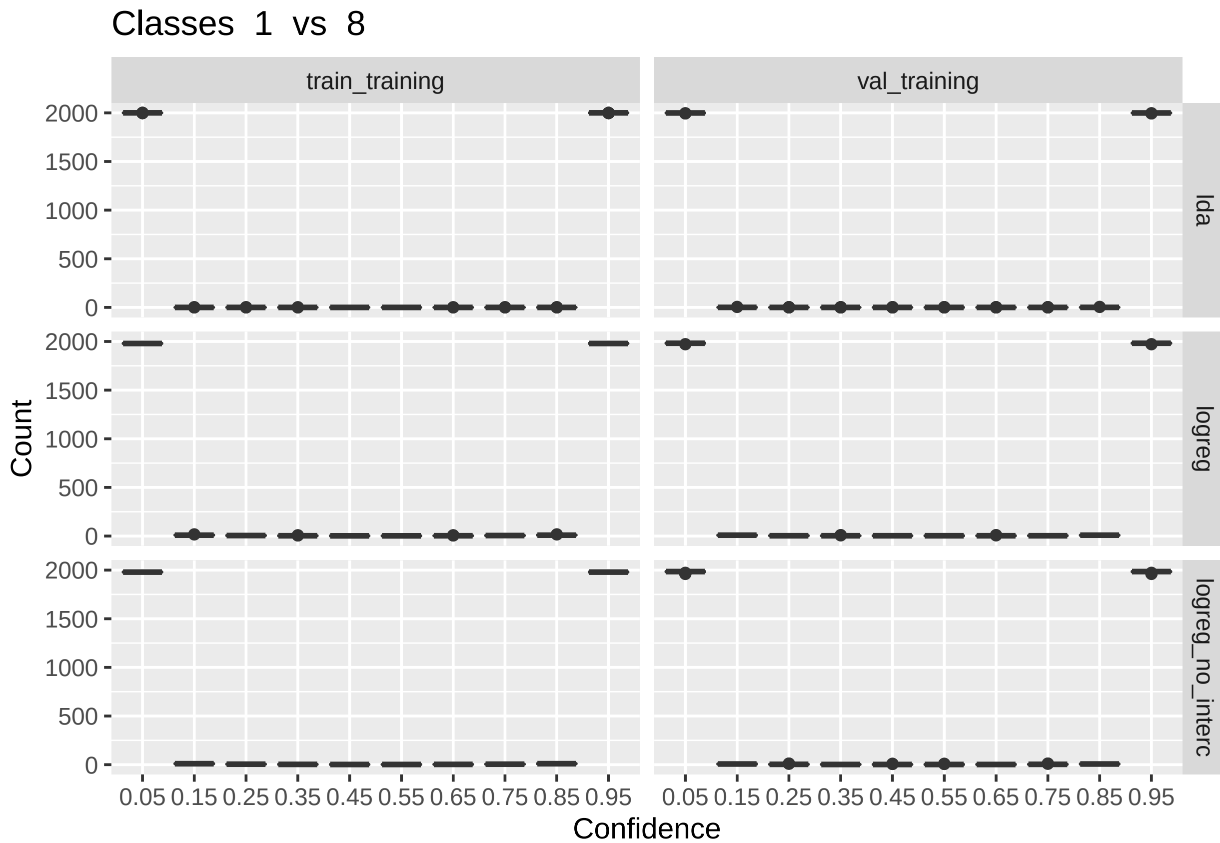

Classes 1 vs 4
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

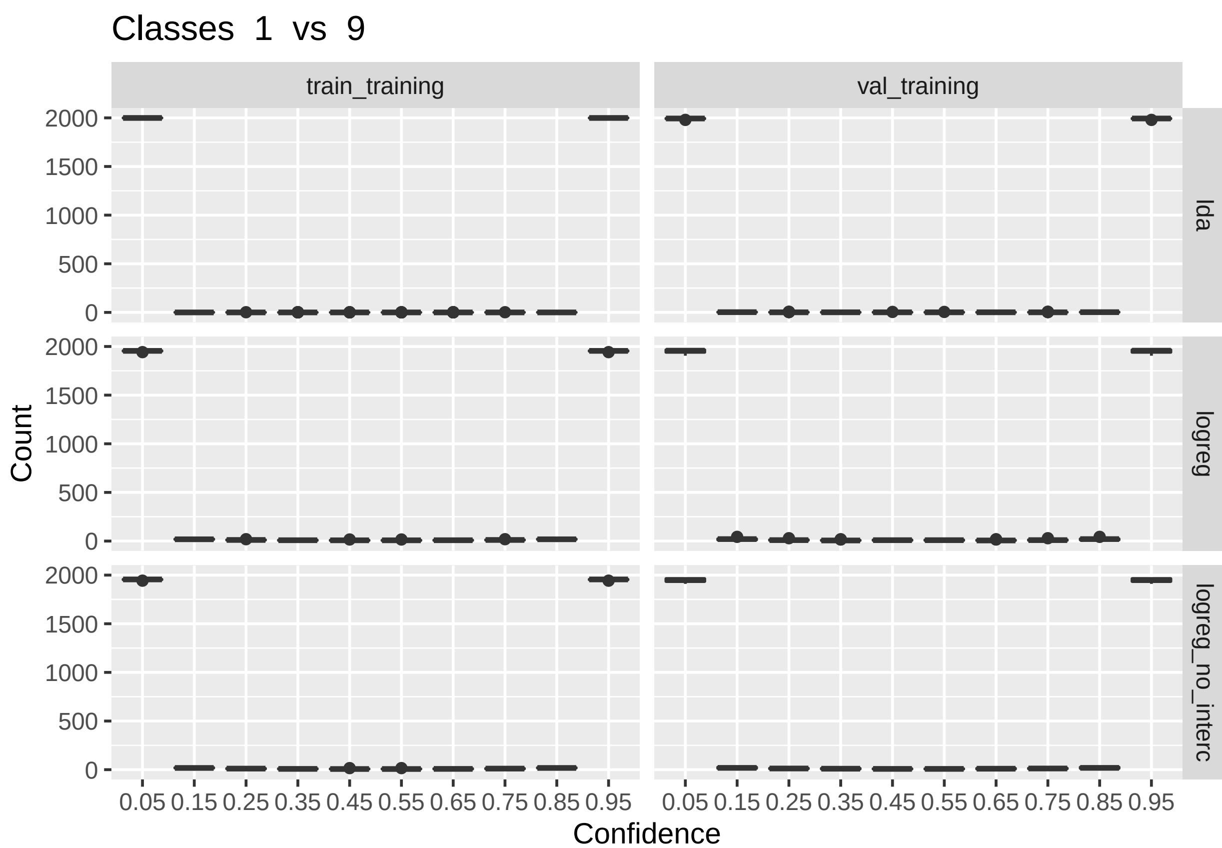


Classes 1 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
Confidence
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95

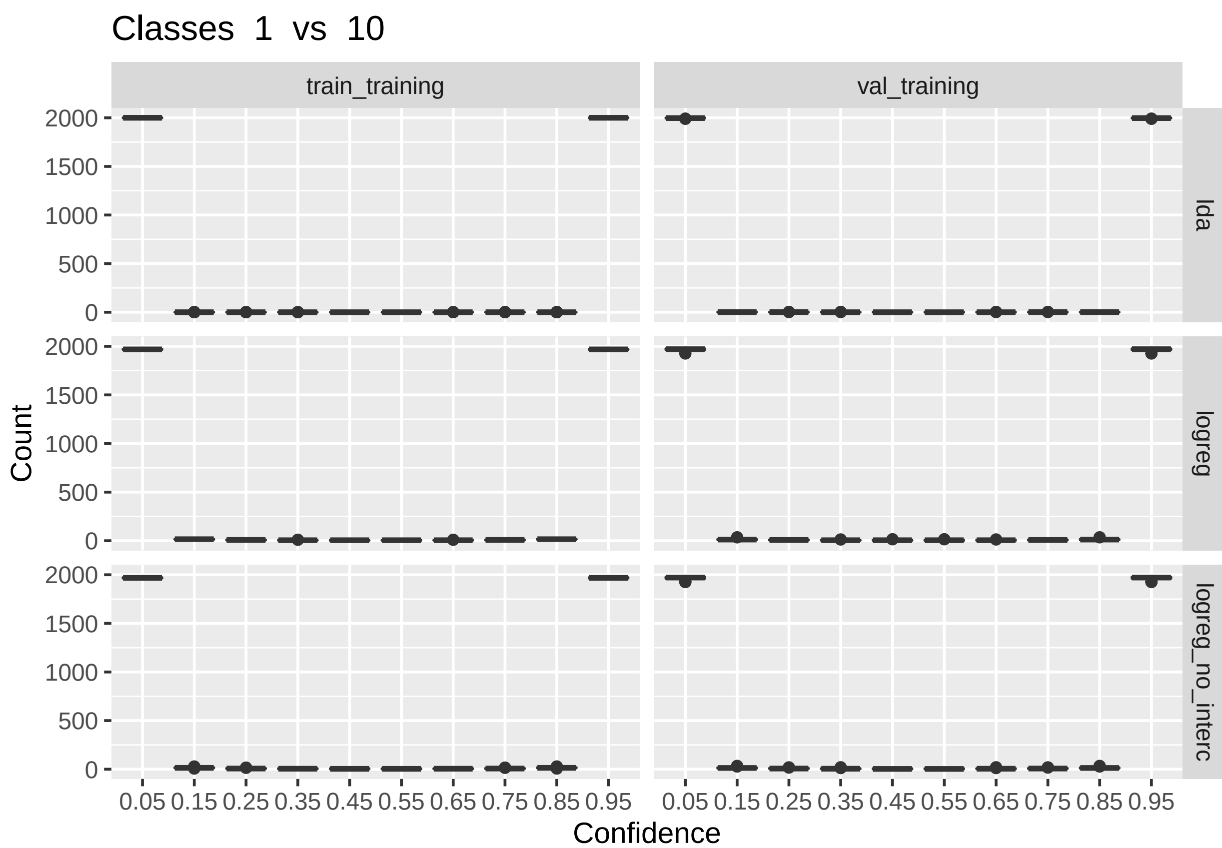

Classes 1 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence


Classes 1 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

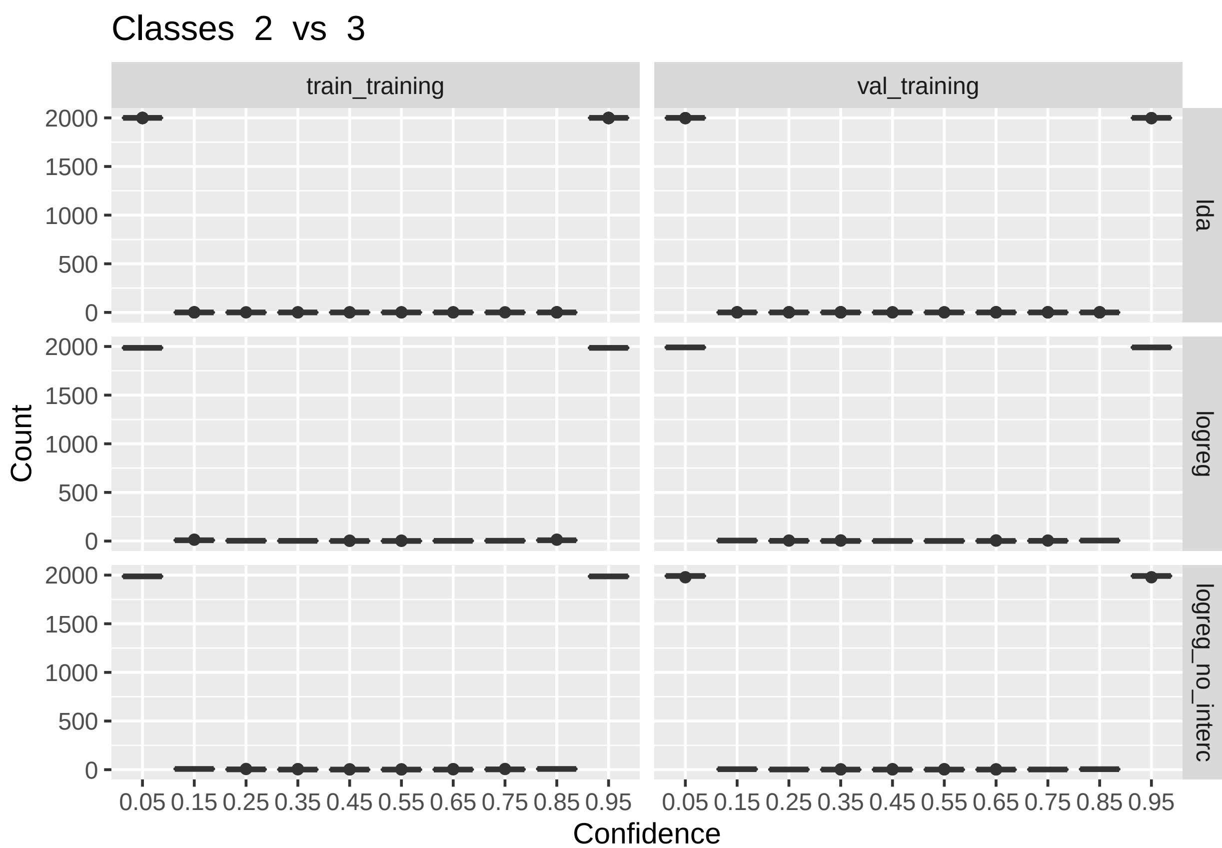

Classes 1 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence



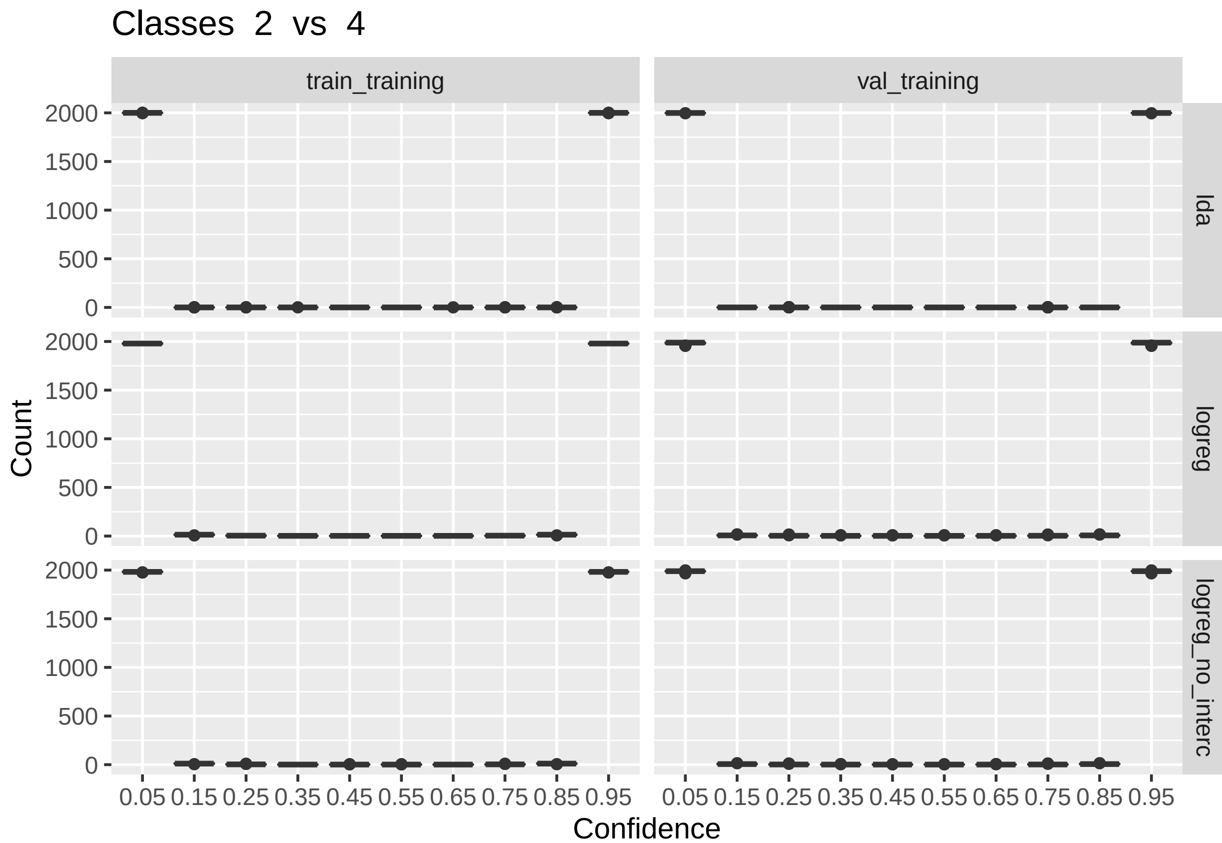
Classes 1 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

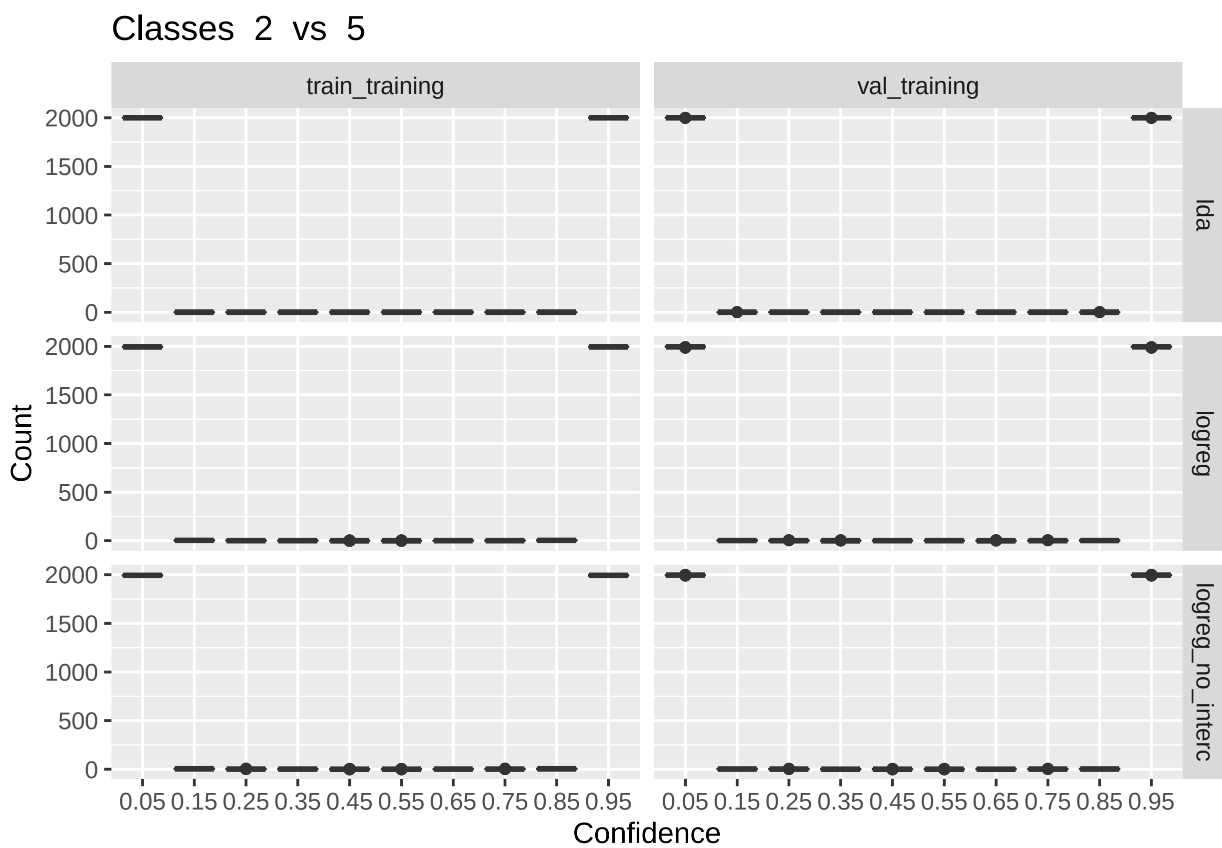


Classes 1 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

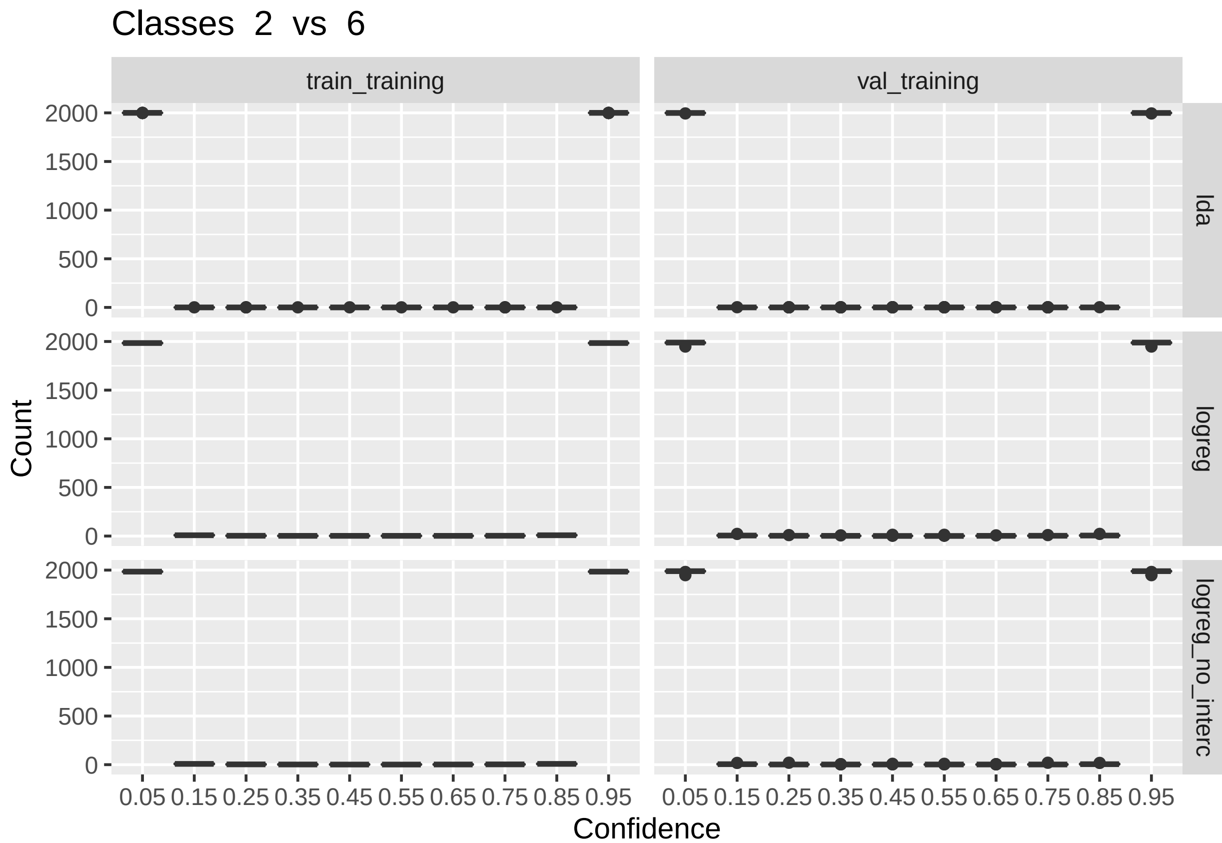


Classes 2 vs 3
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

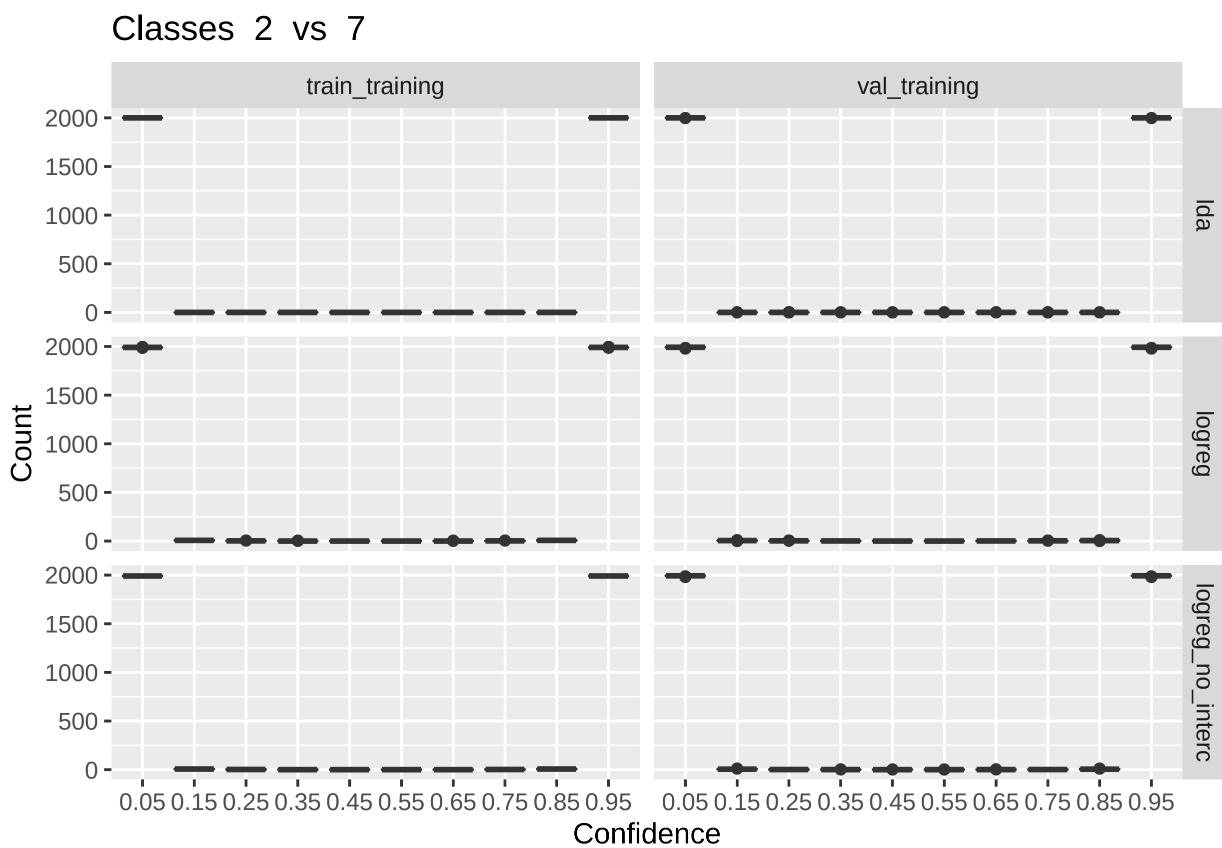

Classes 2 vs 4
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

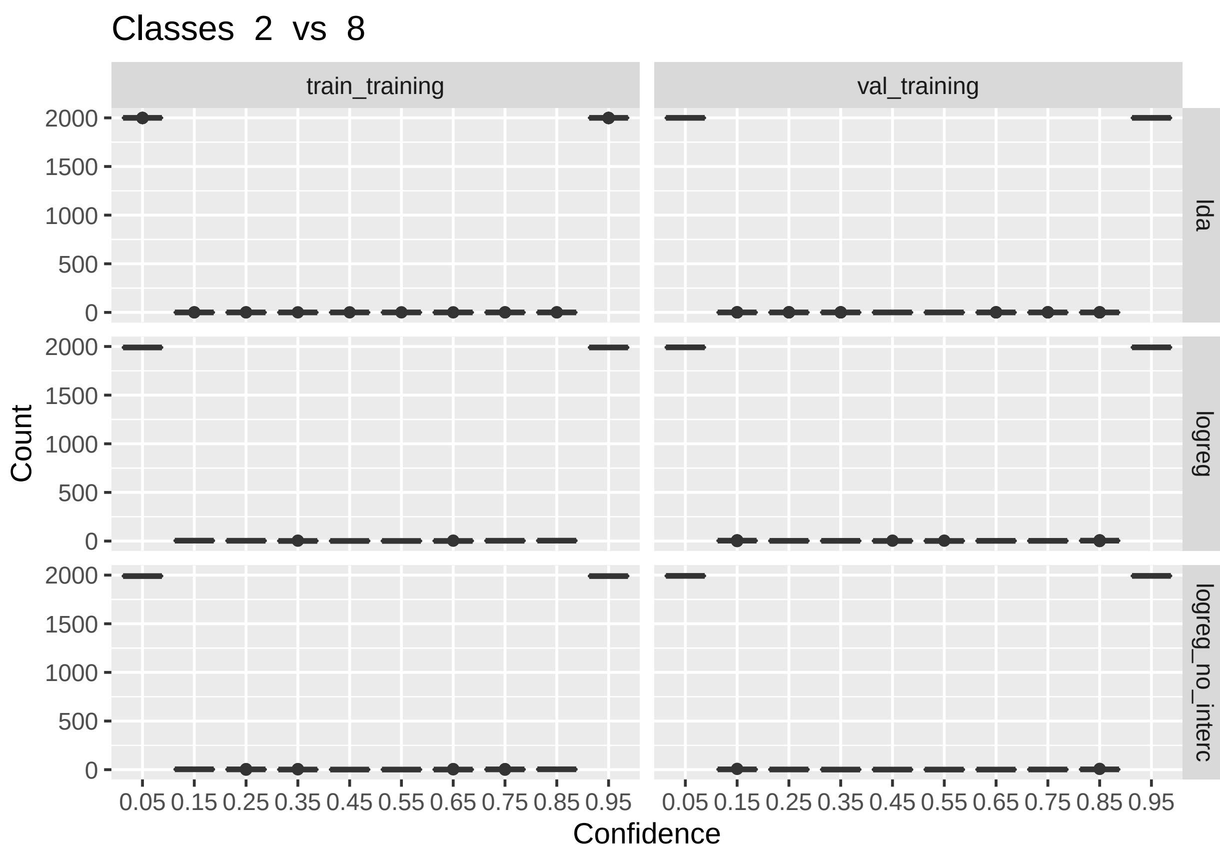


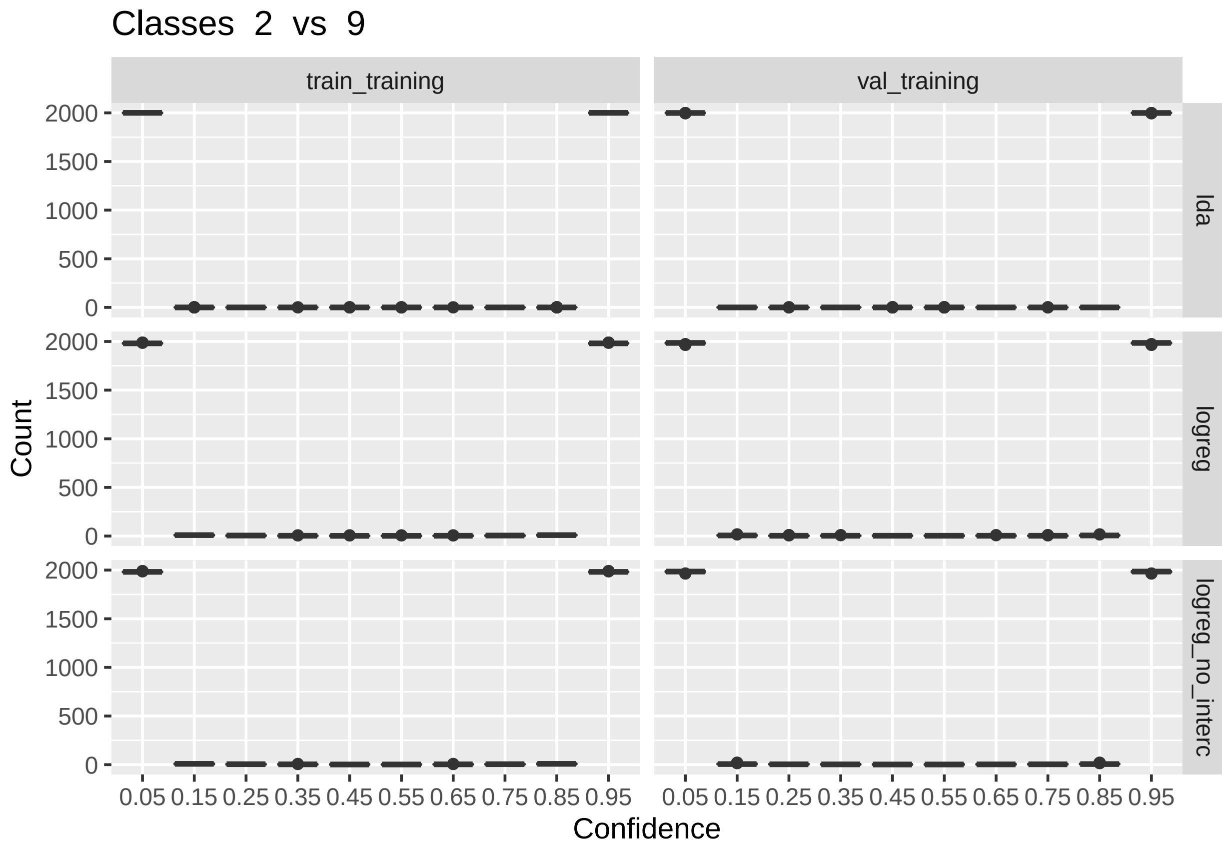
Classes 2 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence



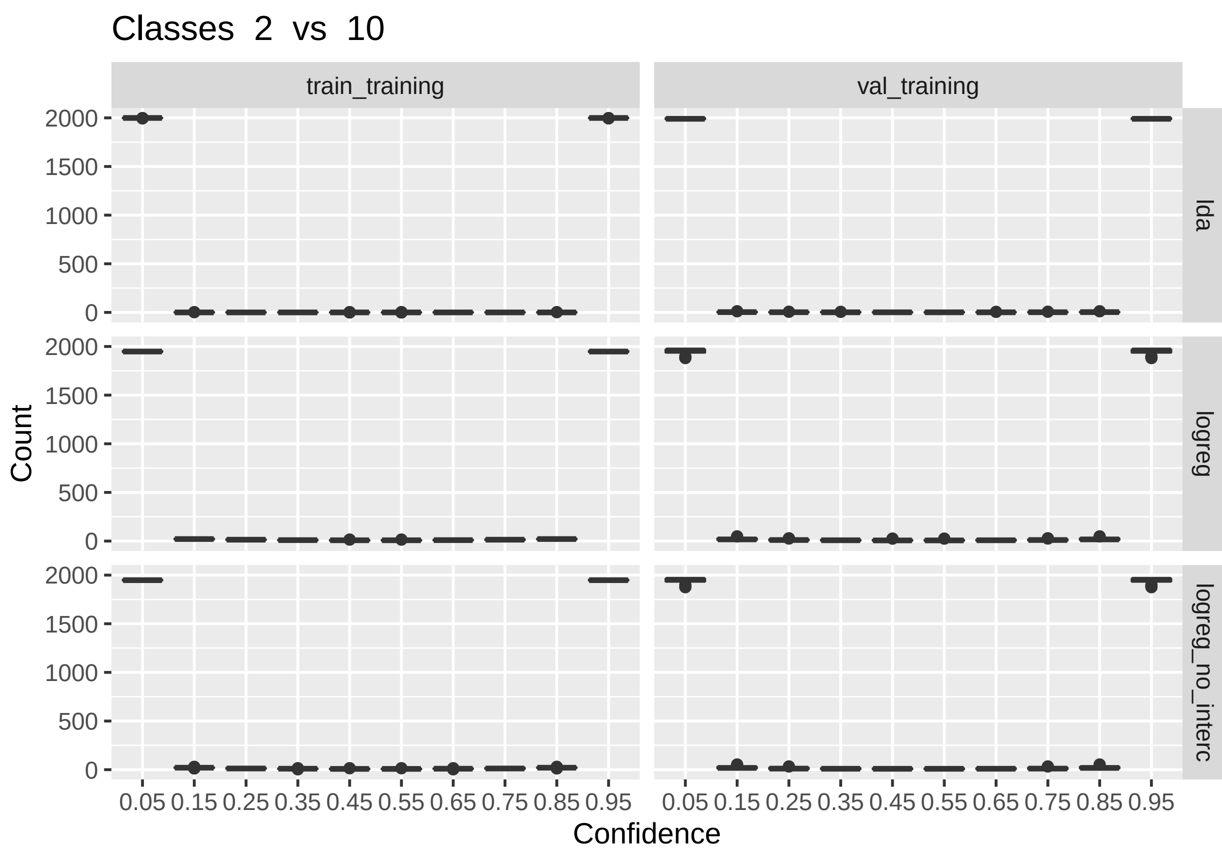
Classes 2 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence


Classes 2 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
Count
Confidence
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95

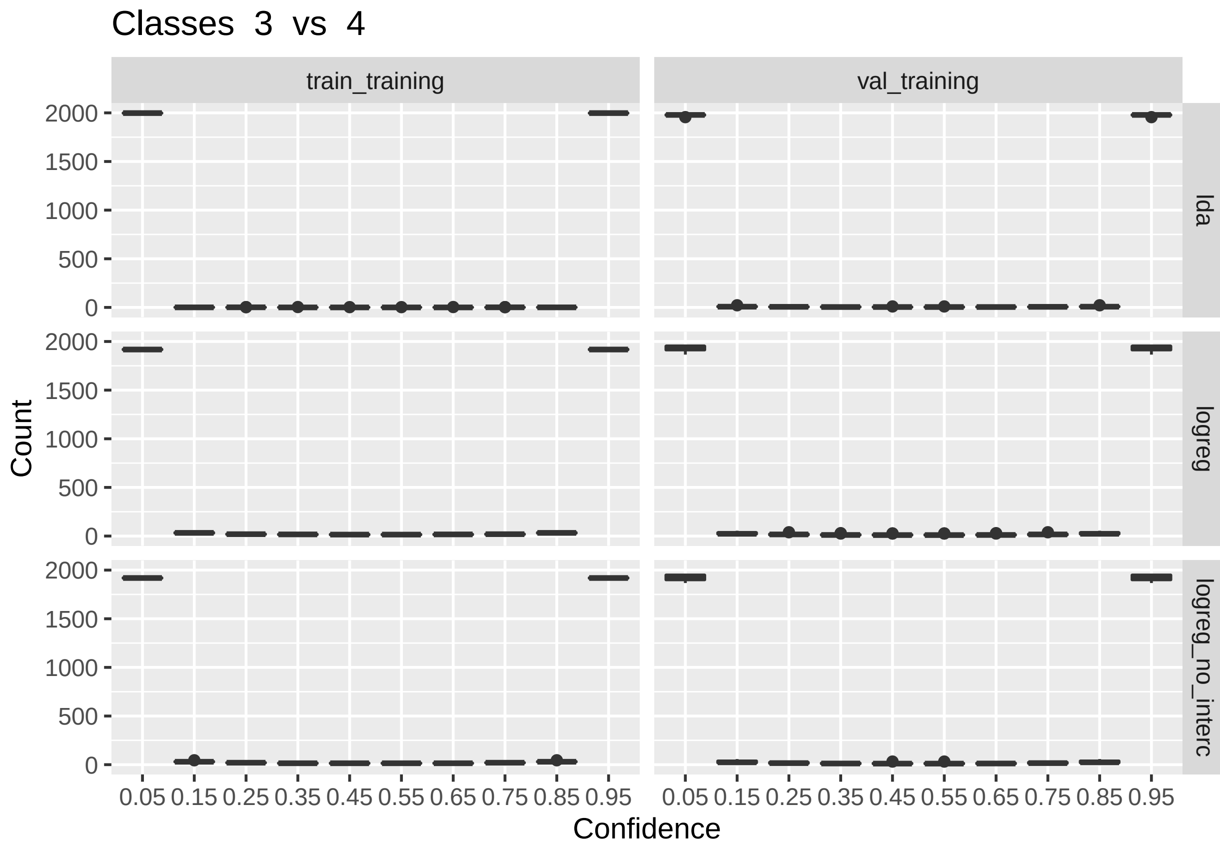

Classes 2 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence



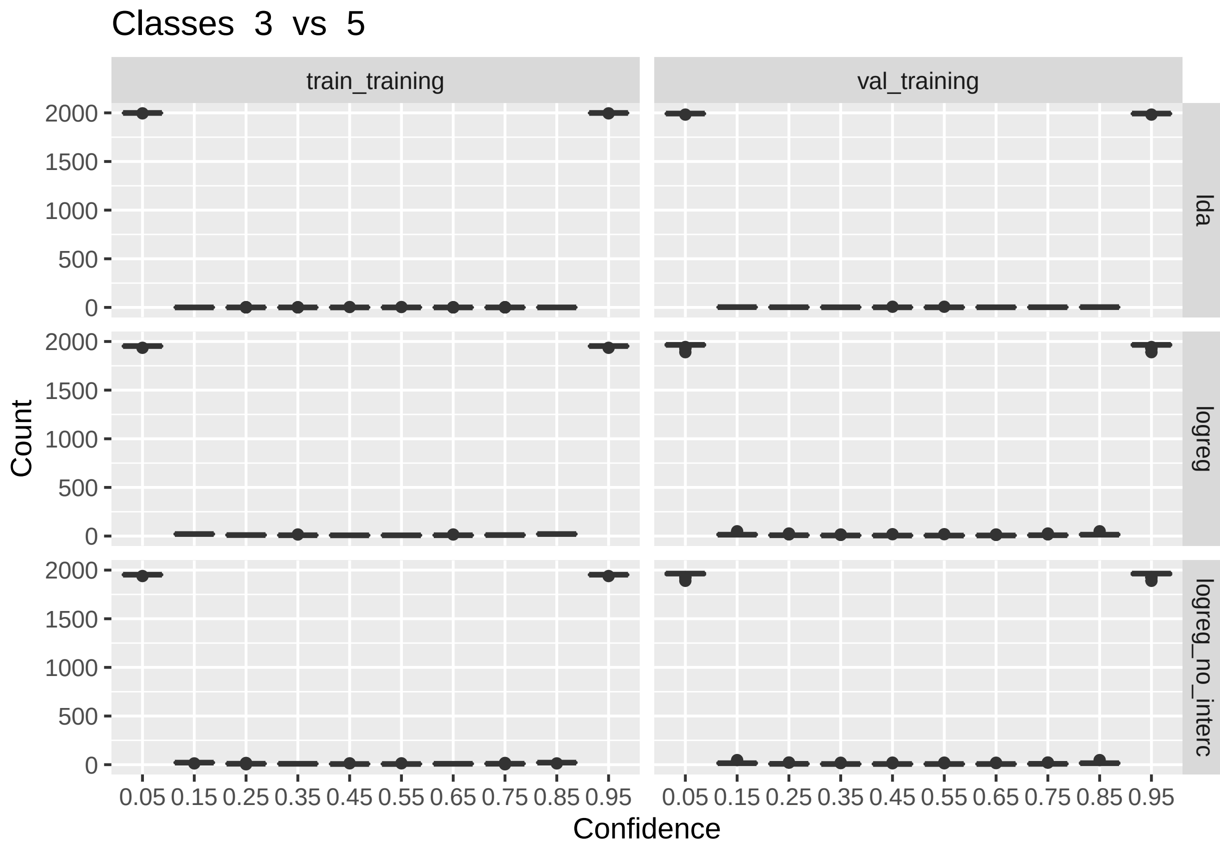
Classes 2 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

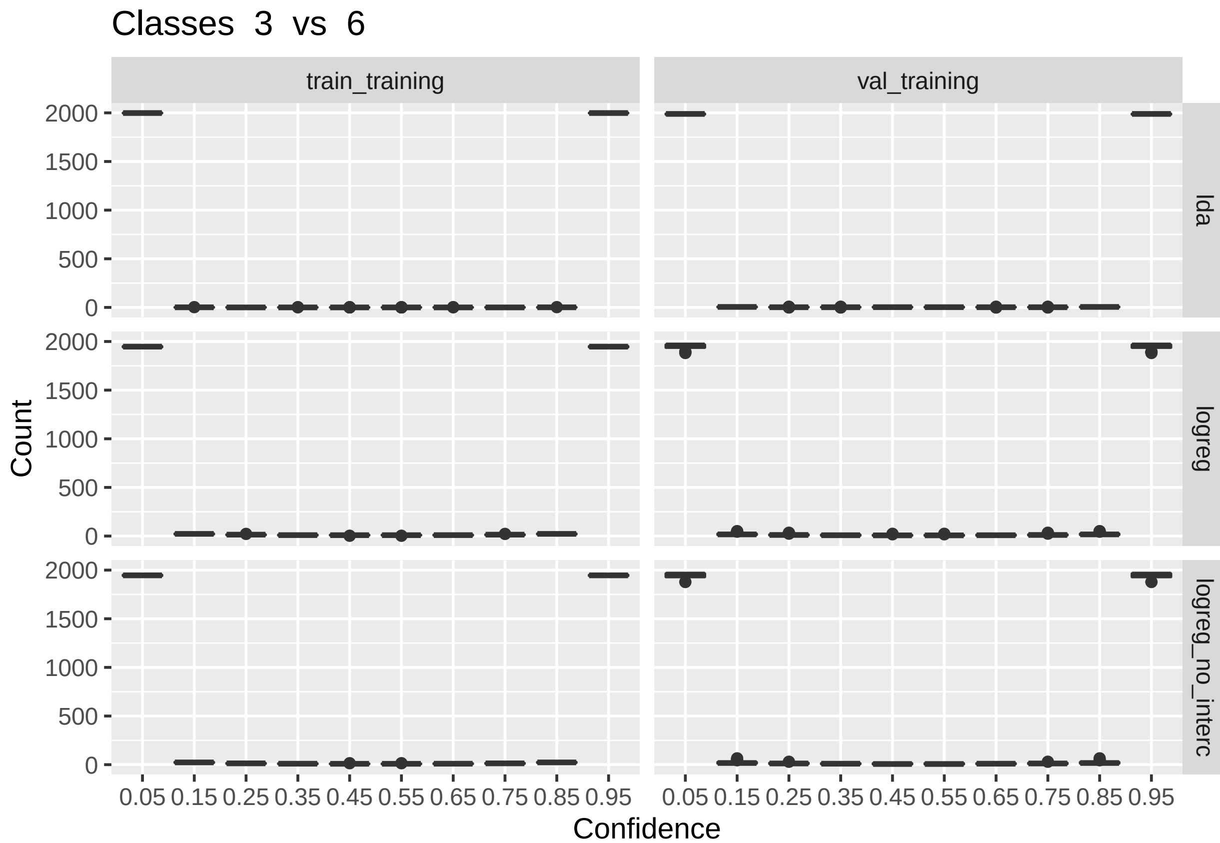


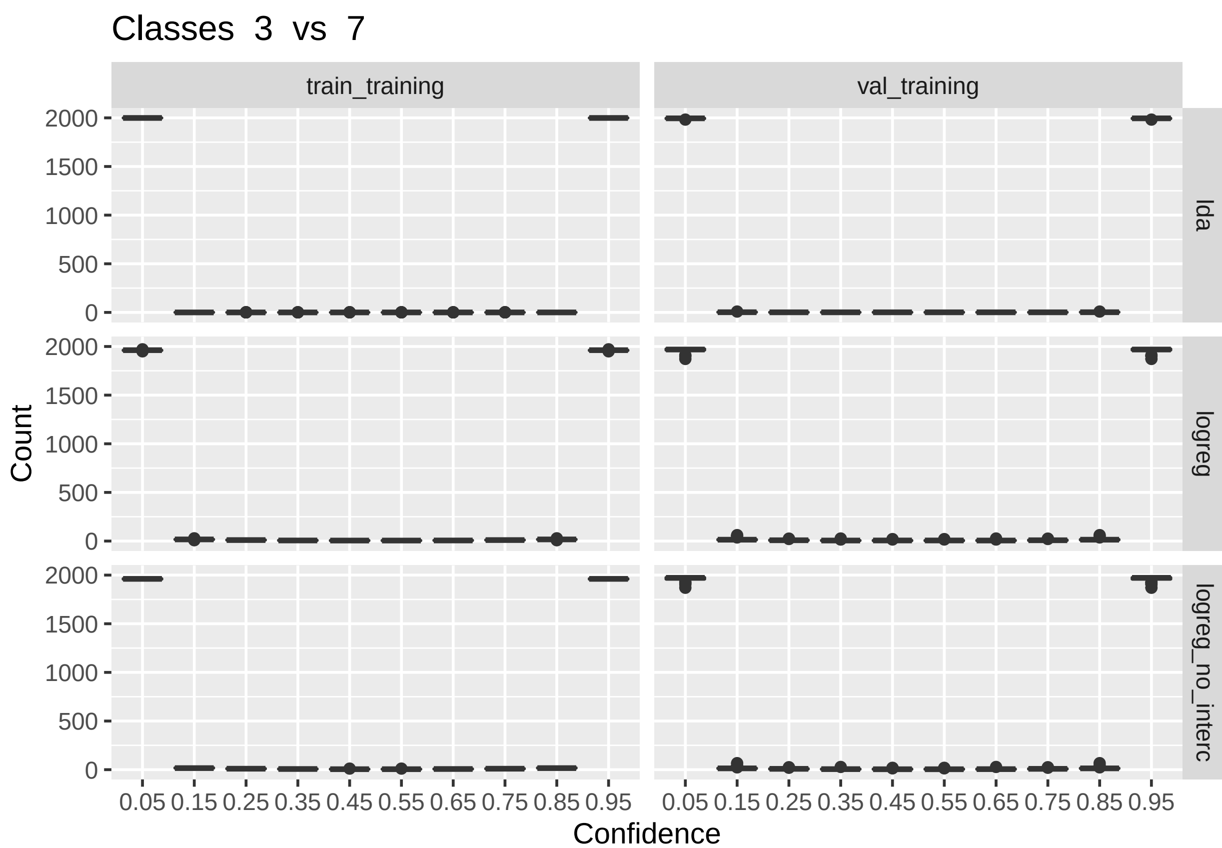
Classes 2 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
Confidence
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95



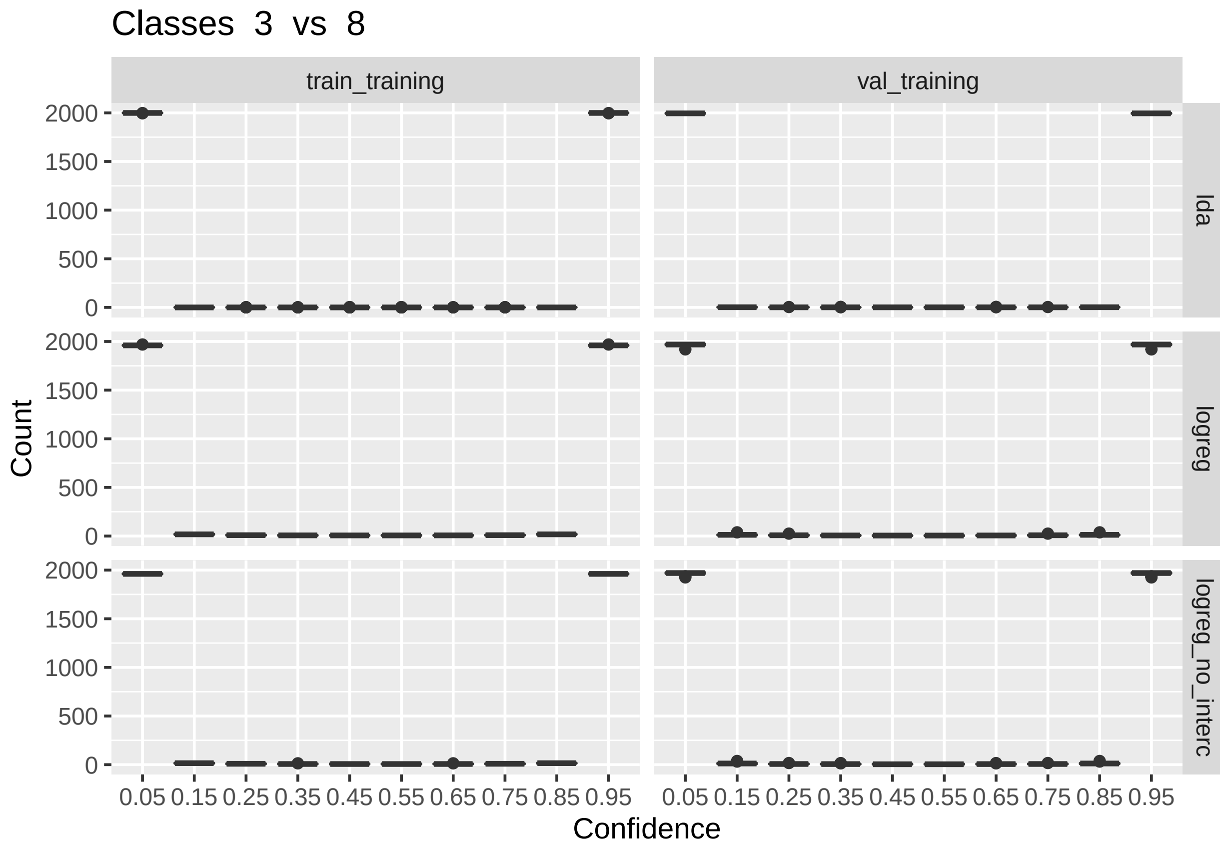
Classes 3 vs 4
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

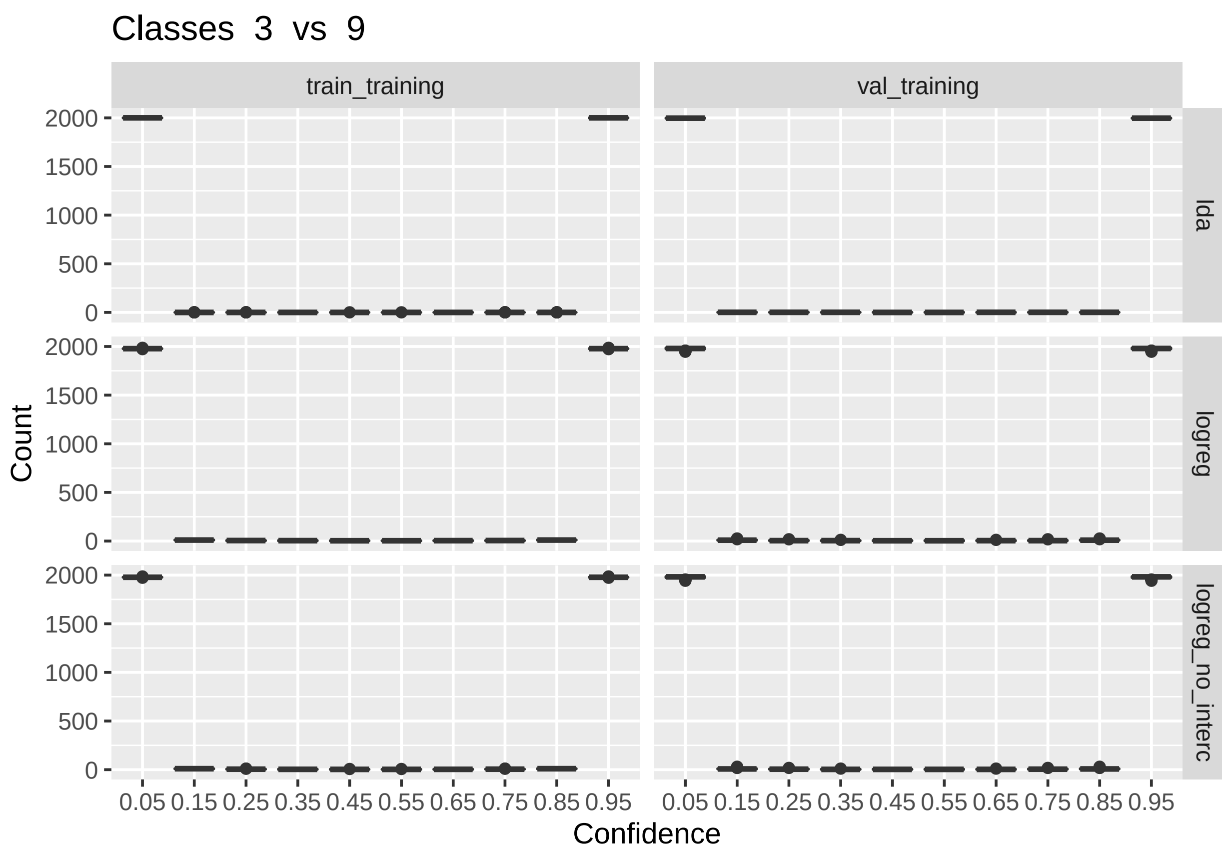

Classes 3 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence


Classes 3 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

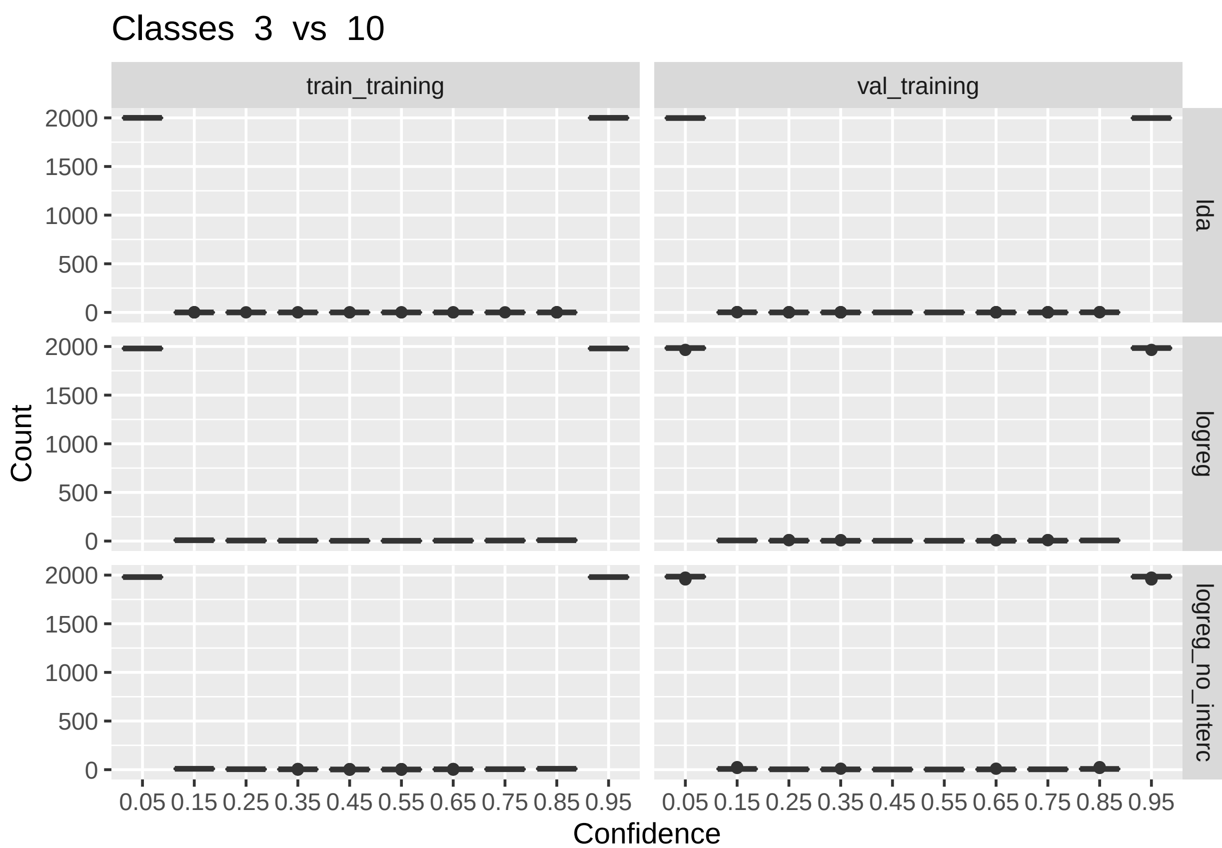


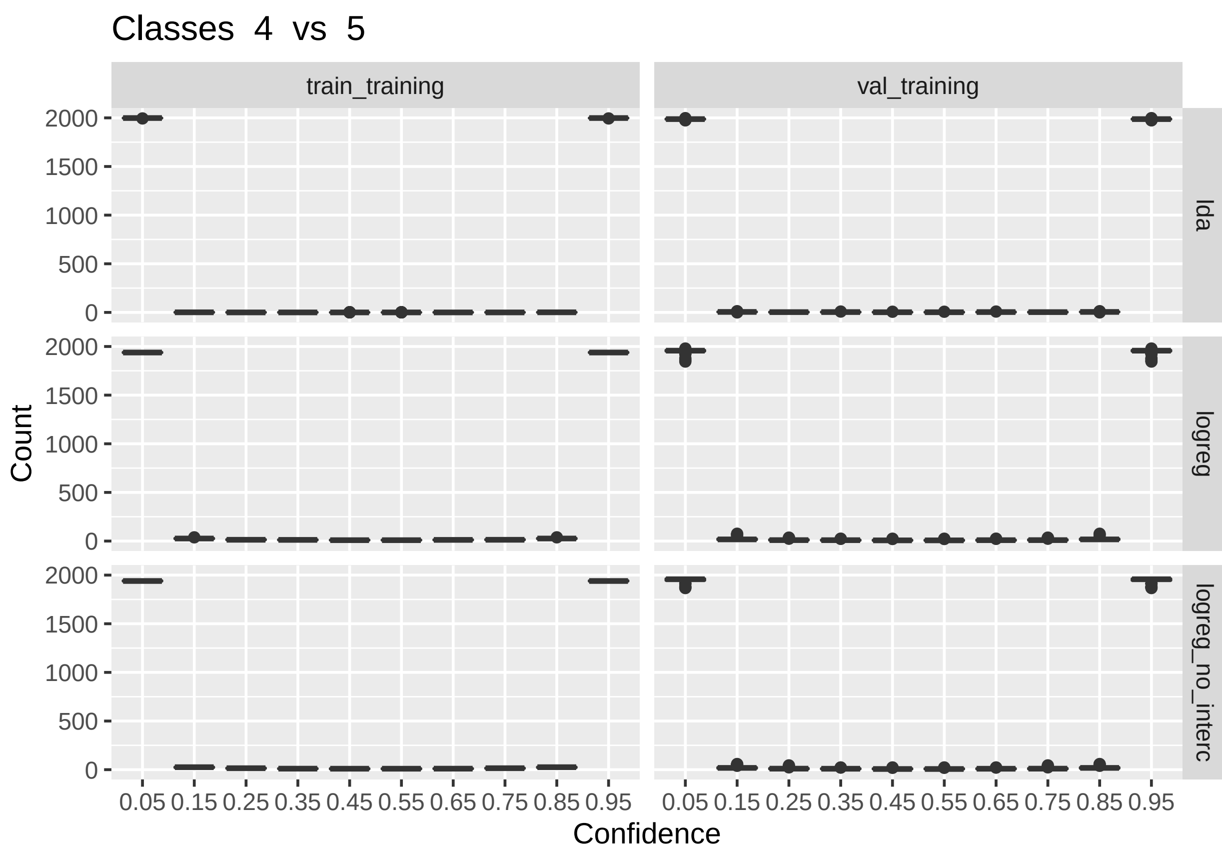
Classes 3 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

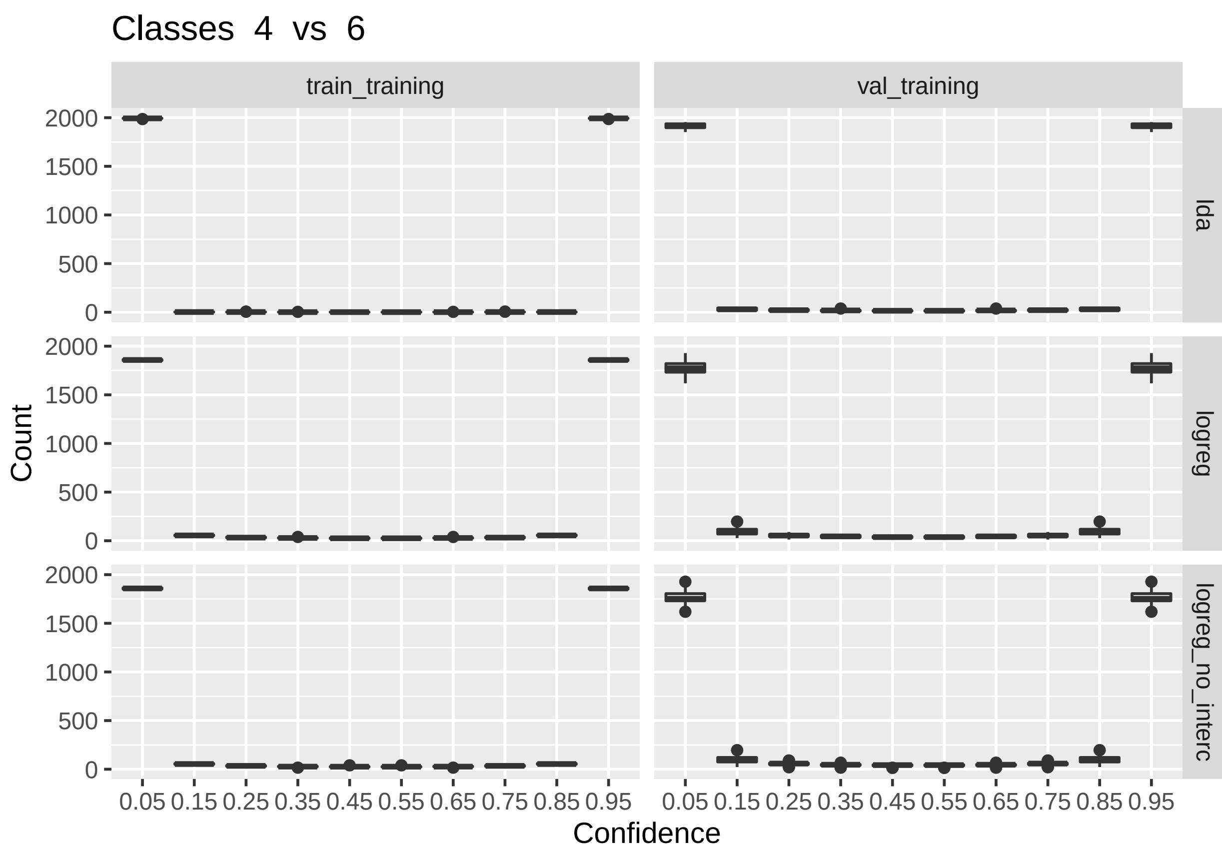

Classes 3 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
Confidence
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95

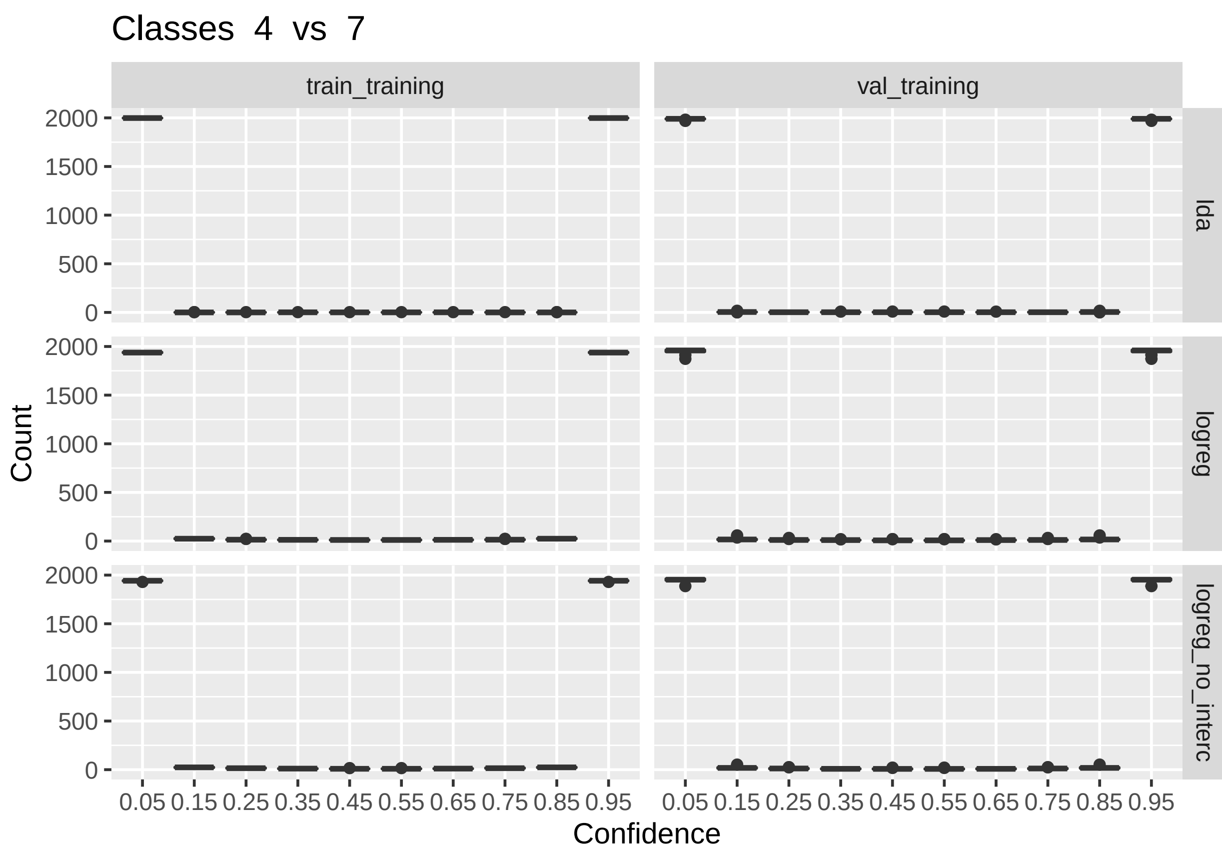


Classes 3 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

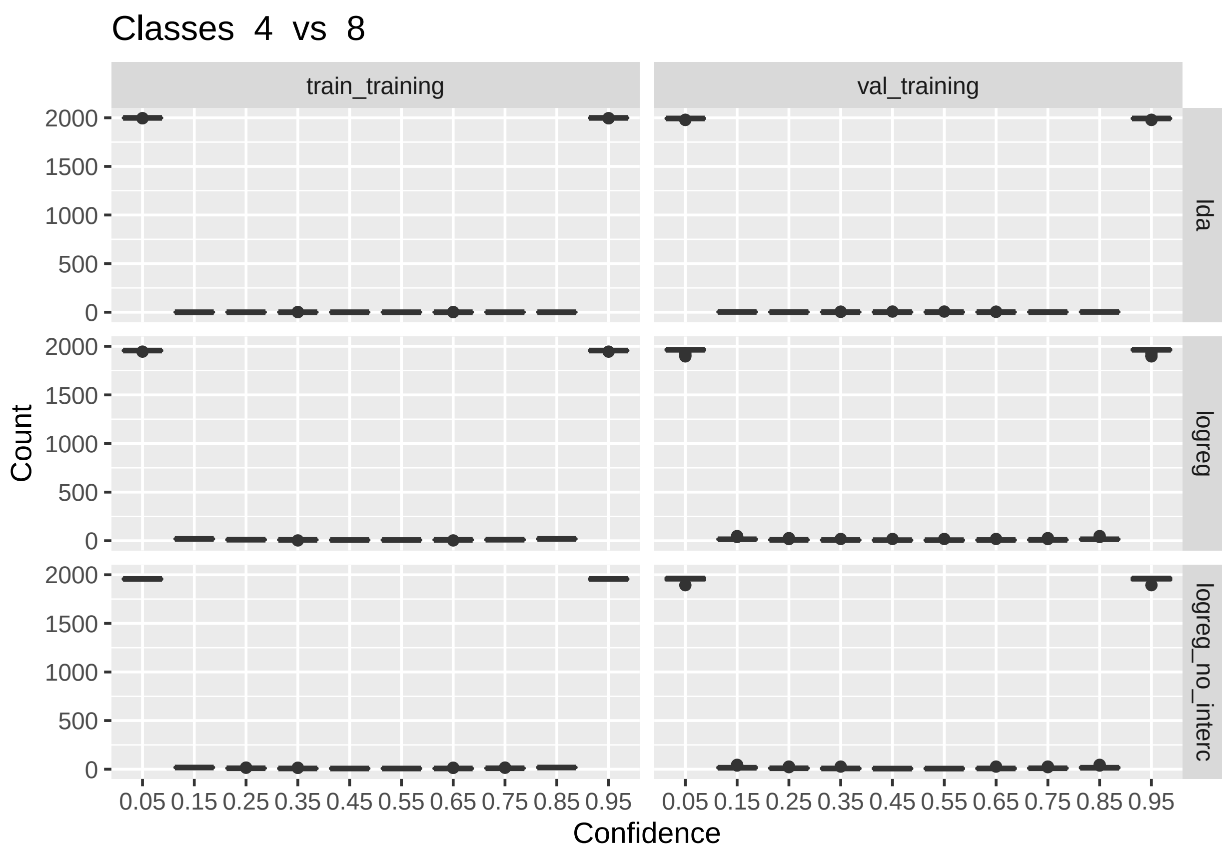

Classes 3 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
Confidence
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95

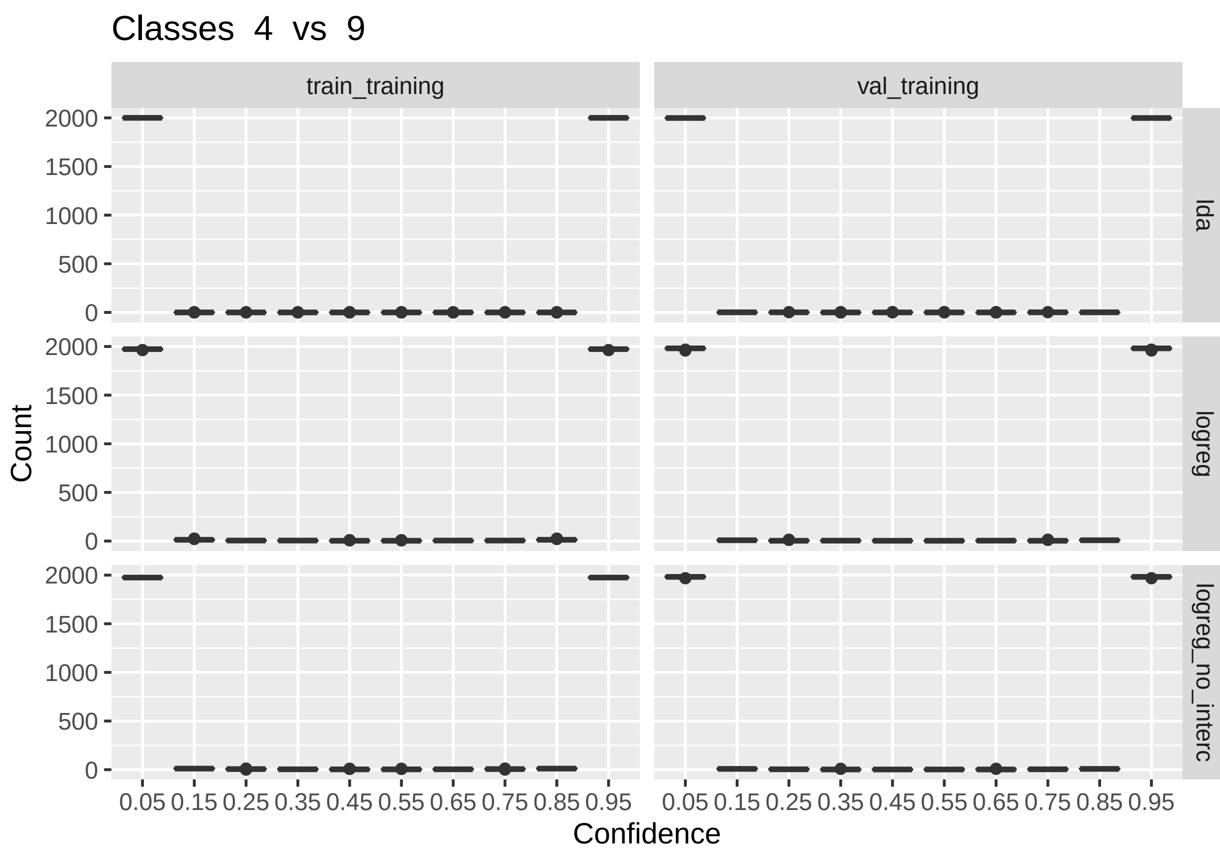

Classes 4 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence


Classes 4 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

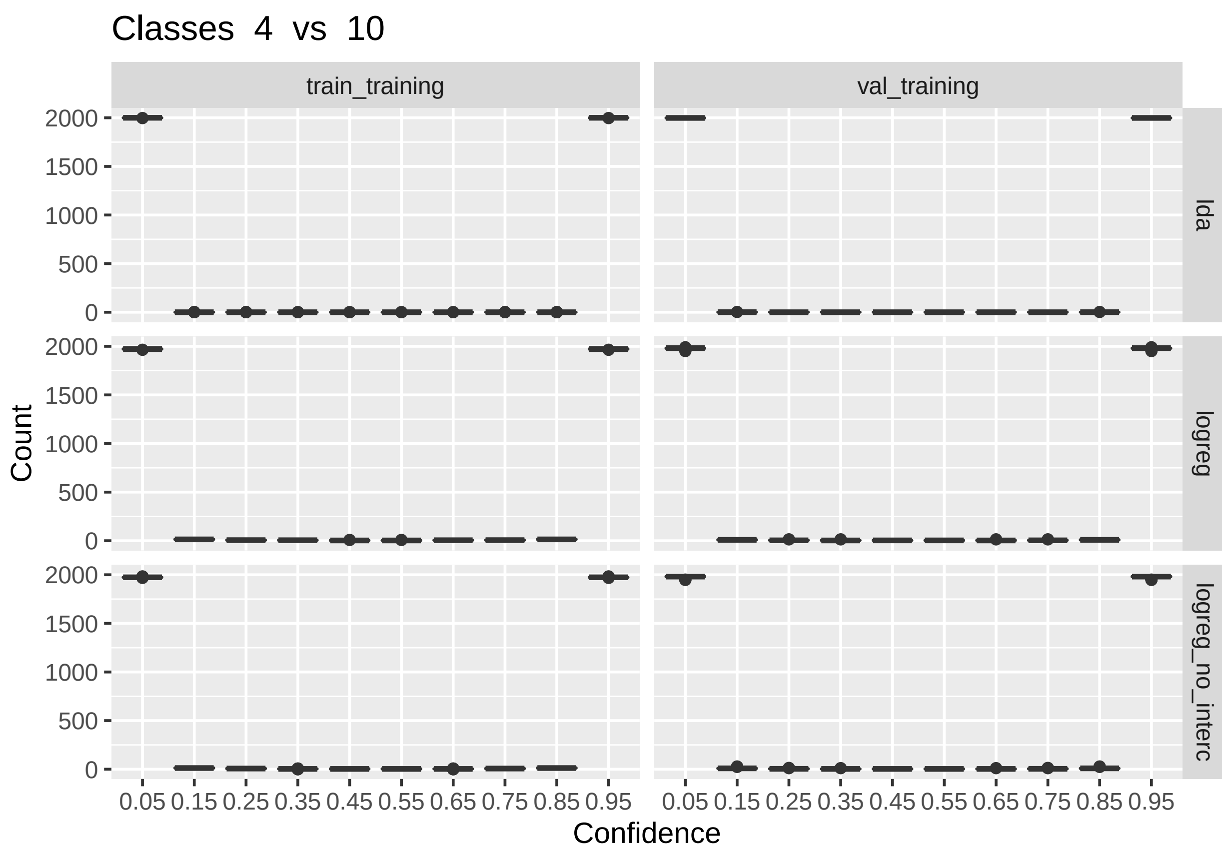


Classes 4 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

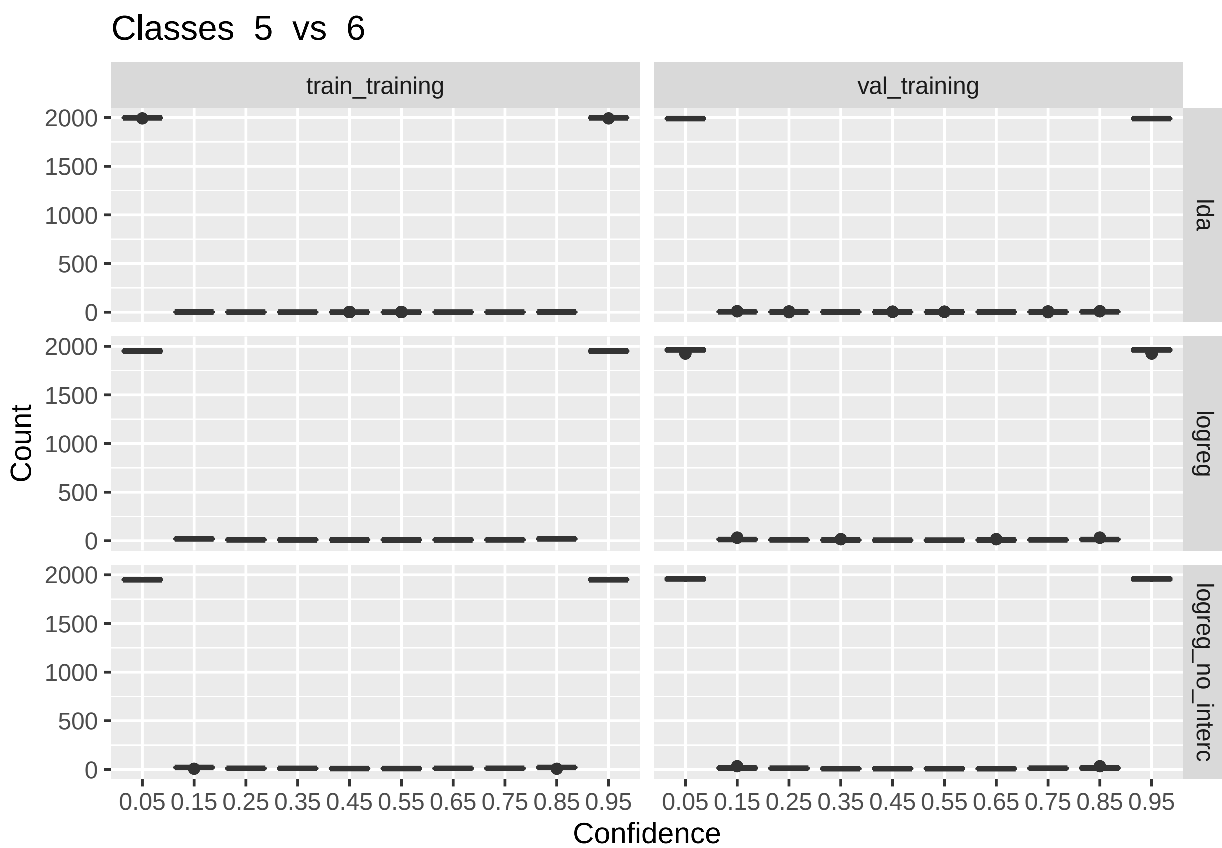

Classes 4 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

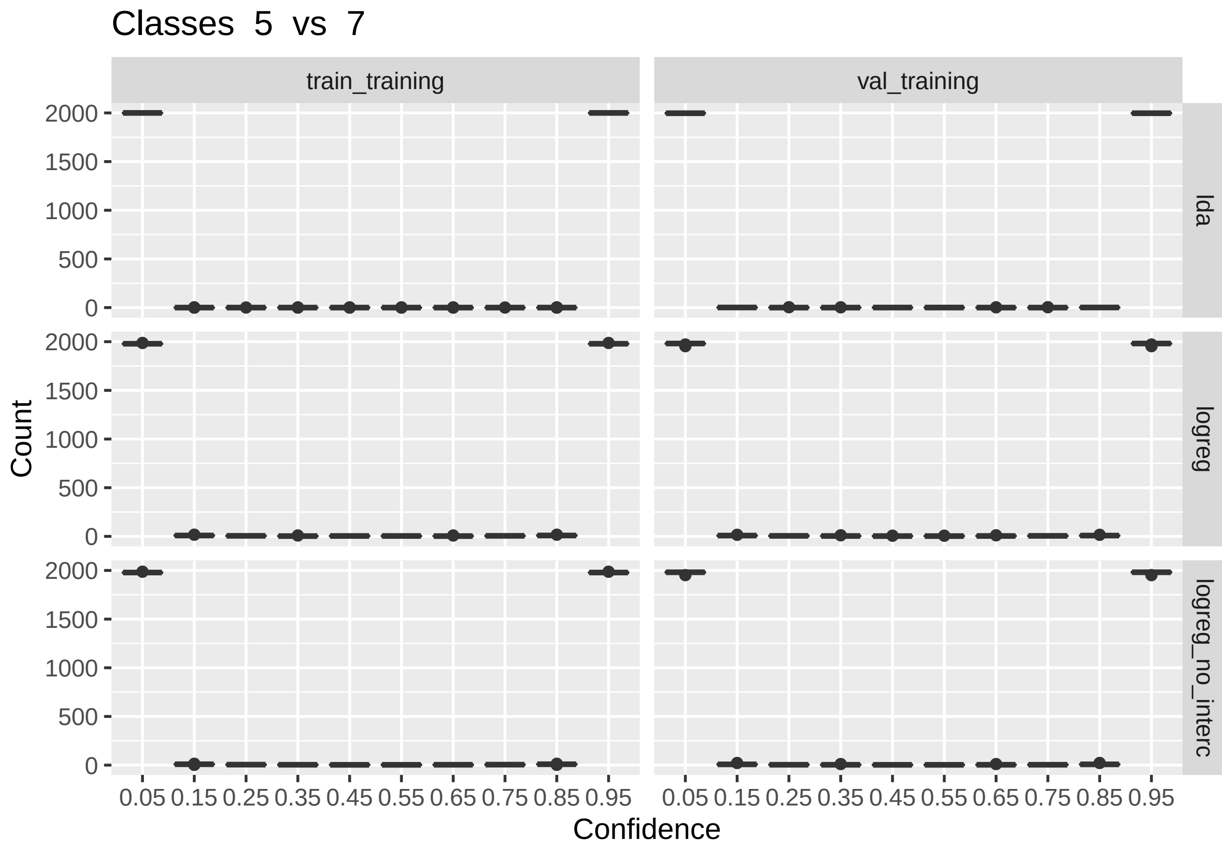


Classes 4 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

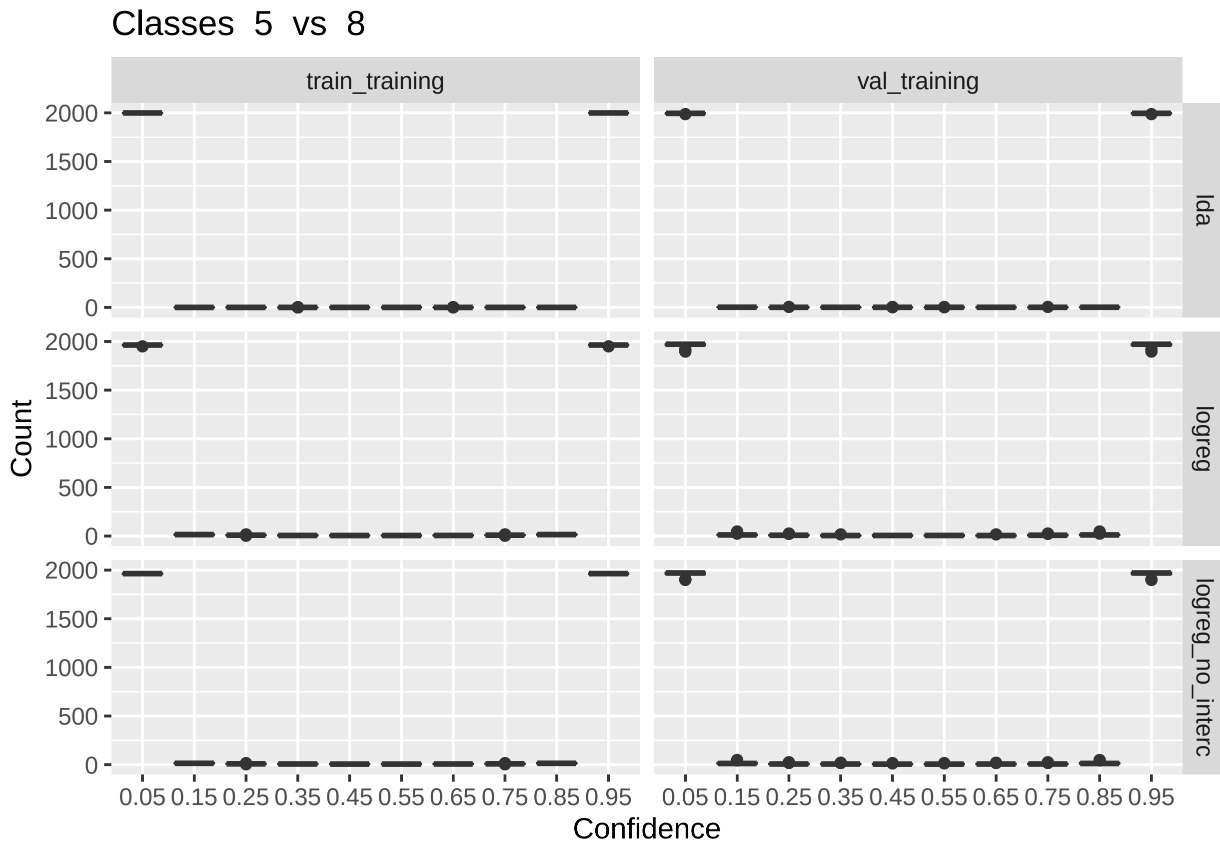


Classes 4 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

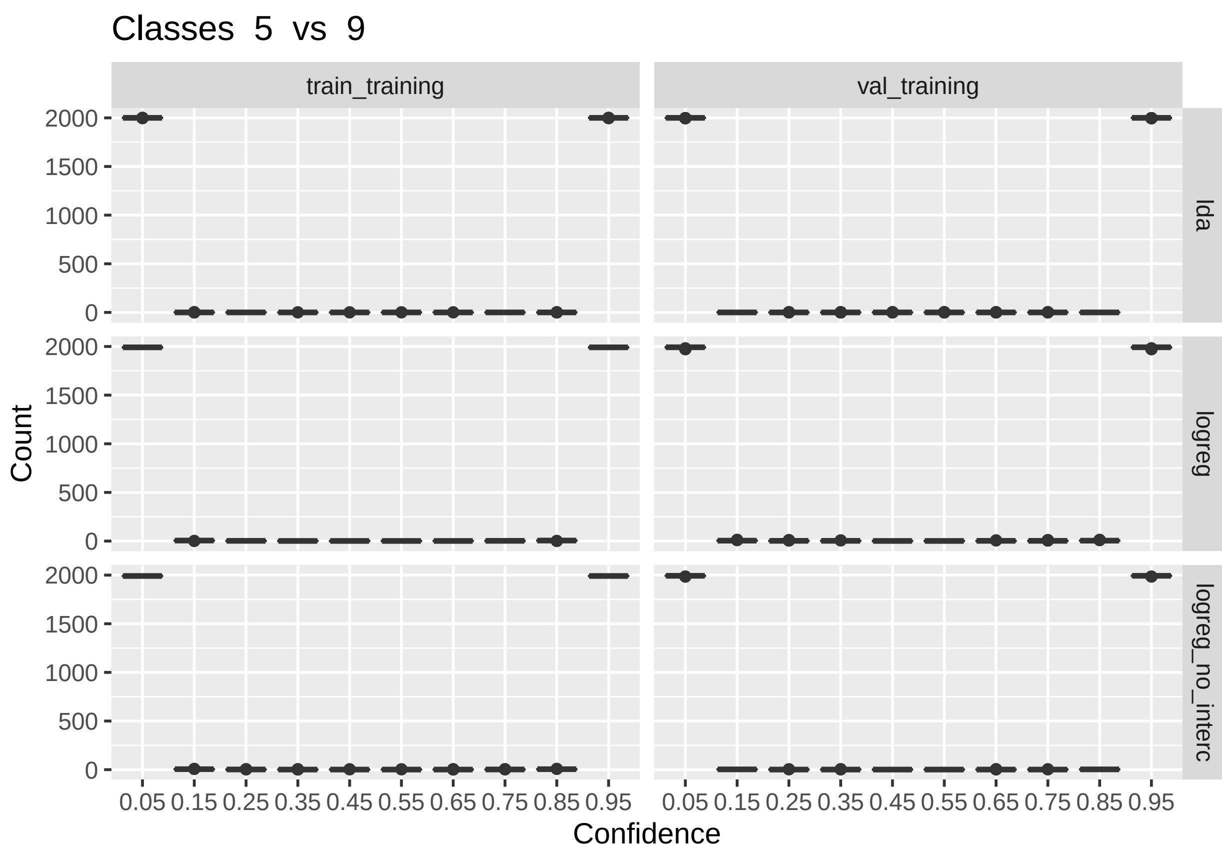

Classes 5 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence



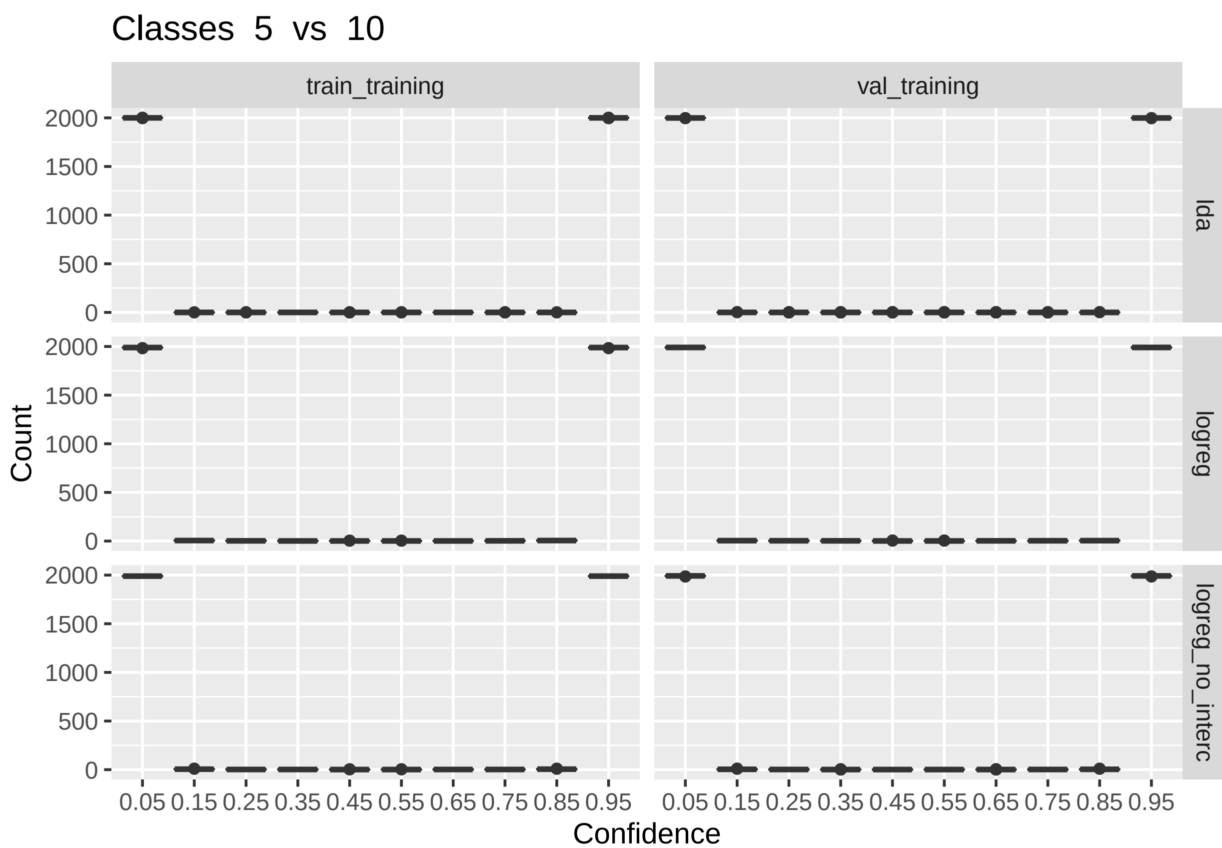
Classes 5 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

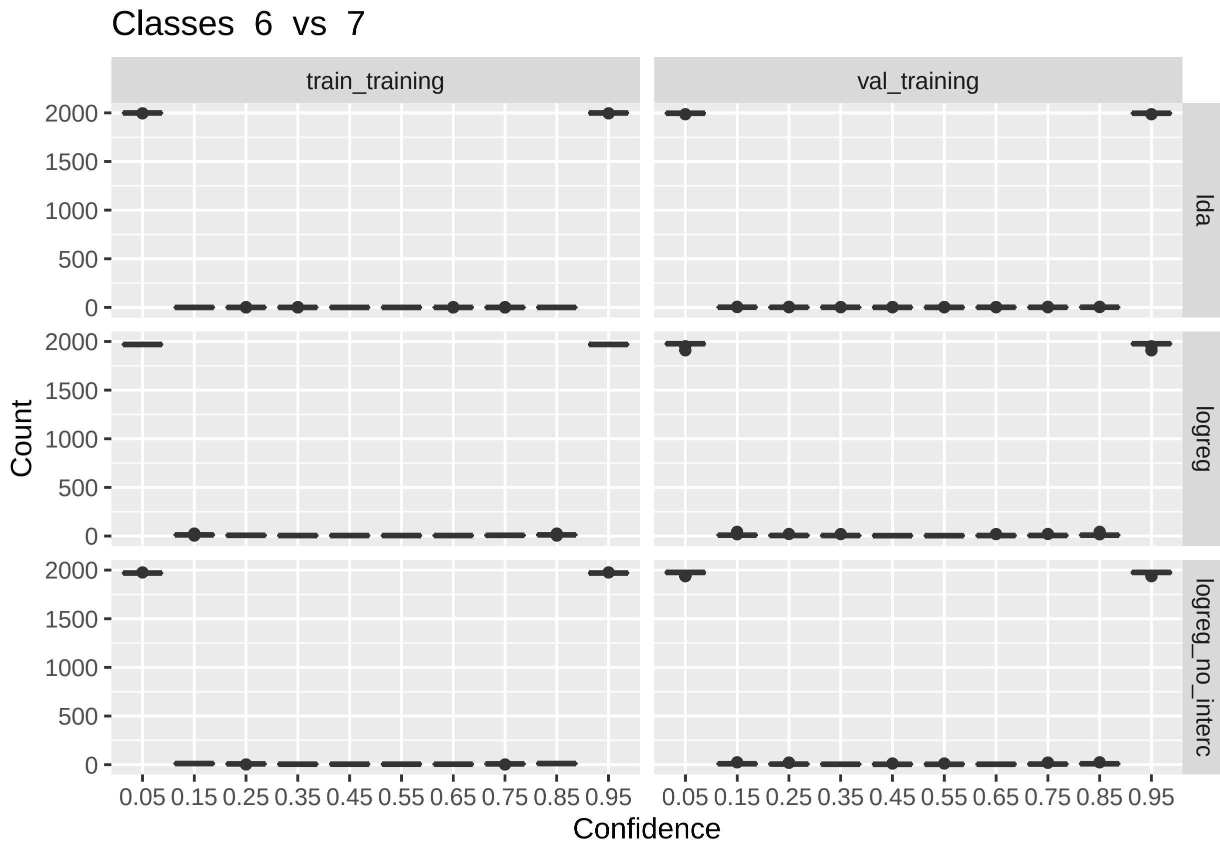


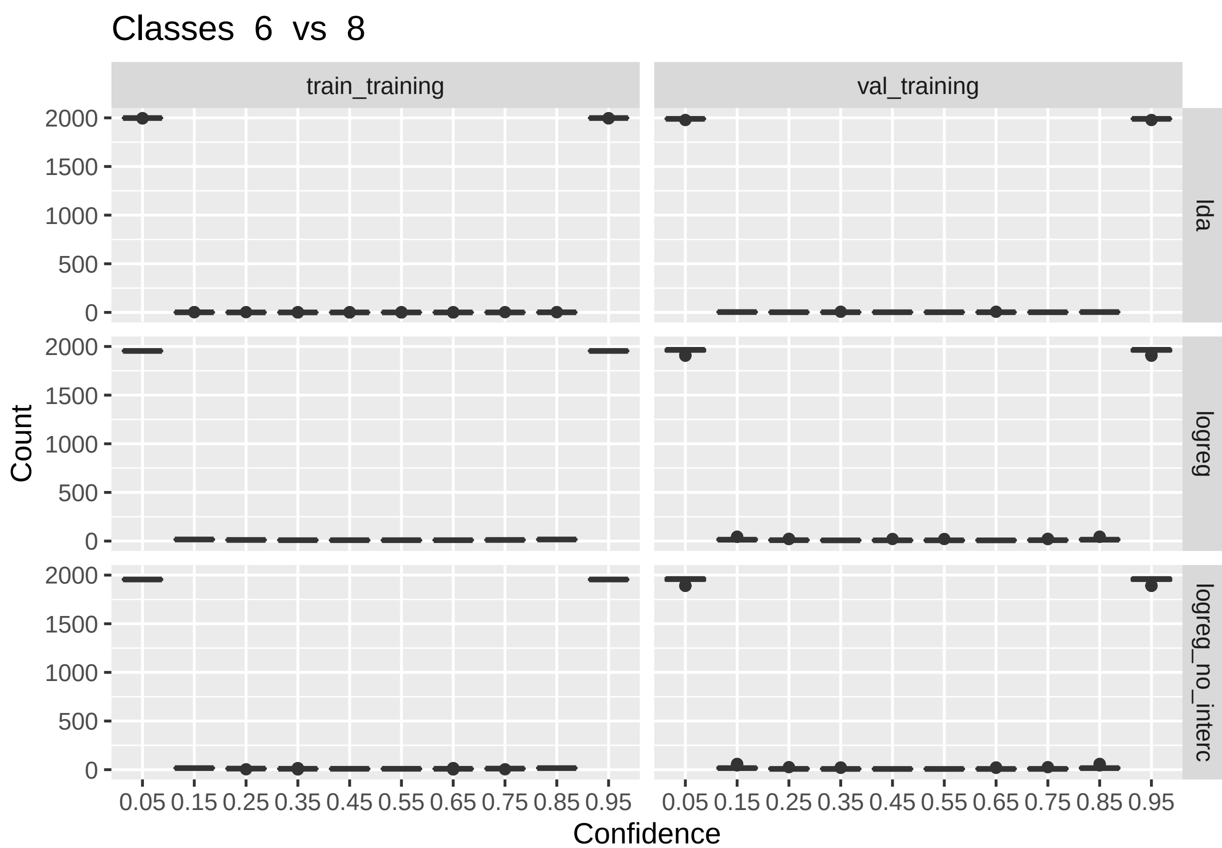
Classes 5 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

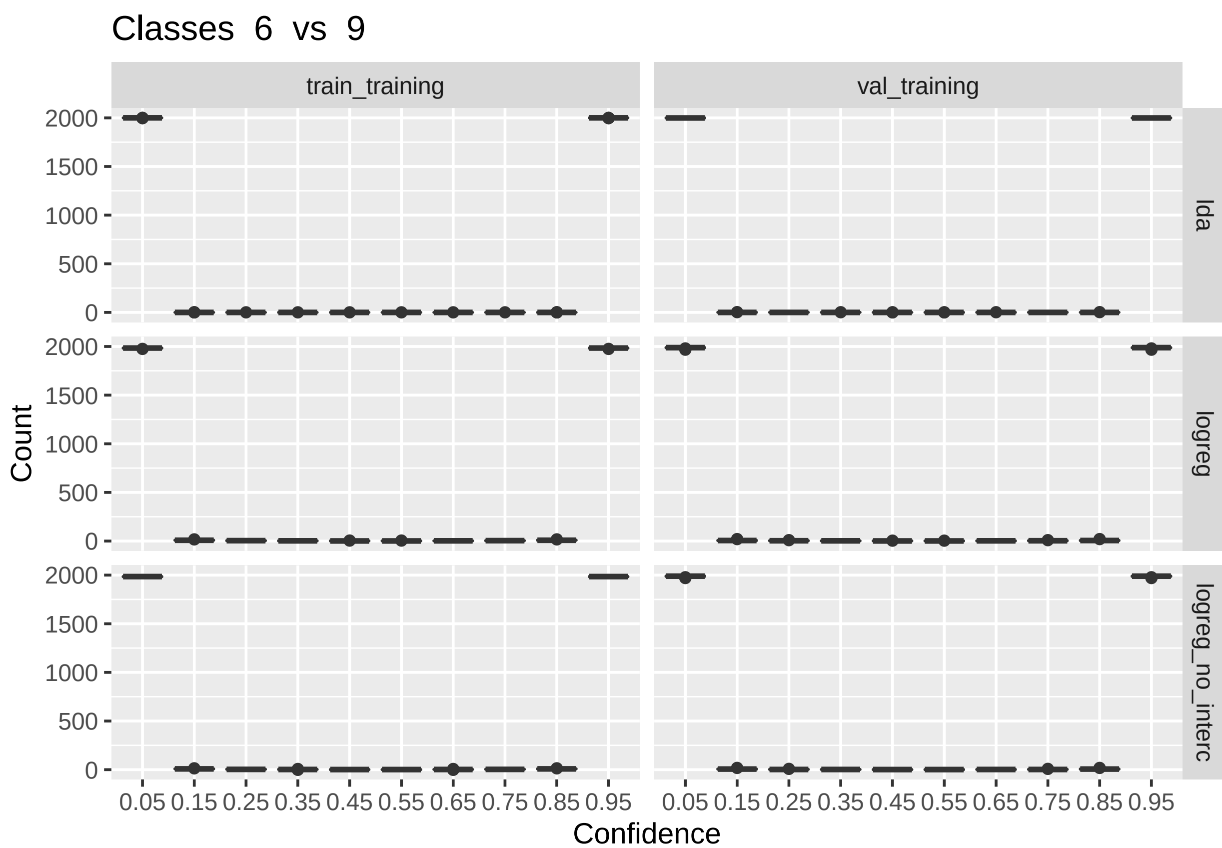


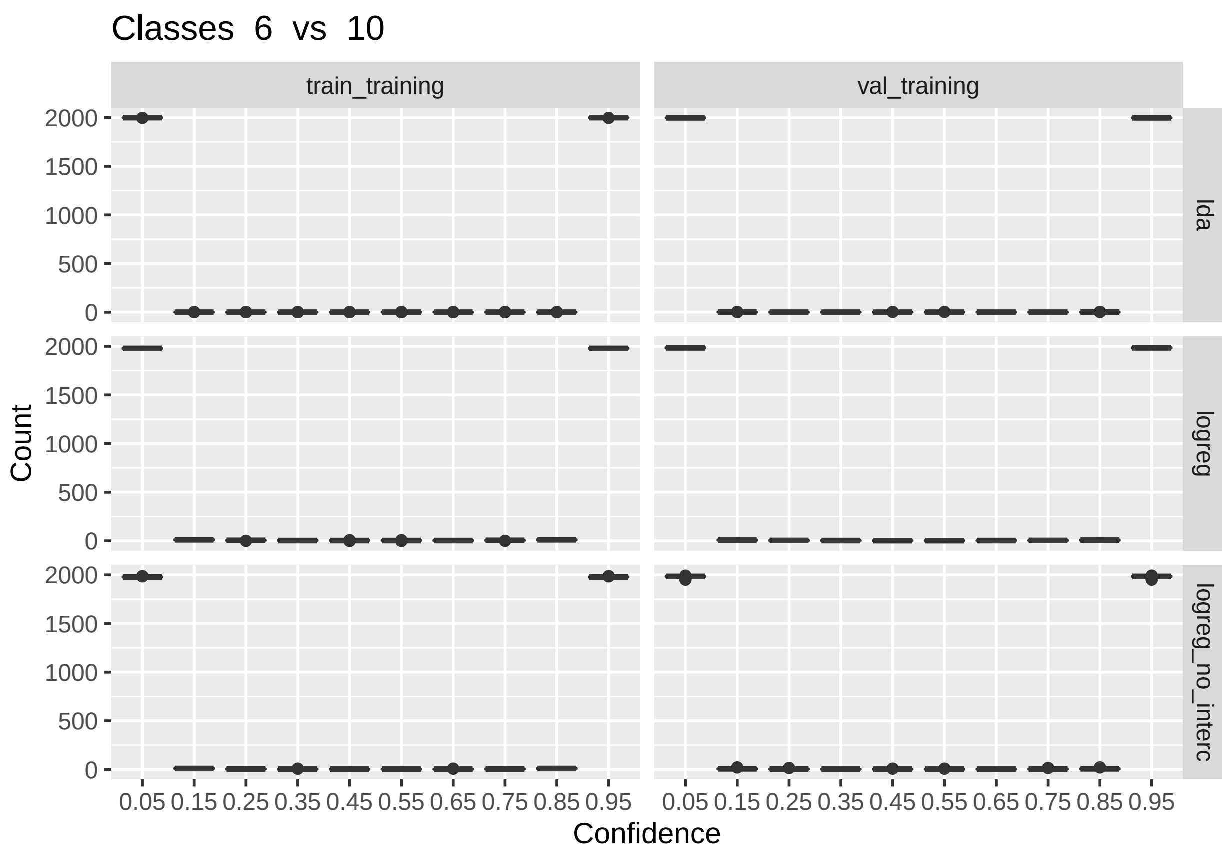
Classes 5 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
Confidence
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95

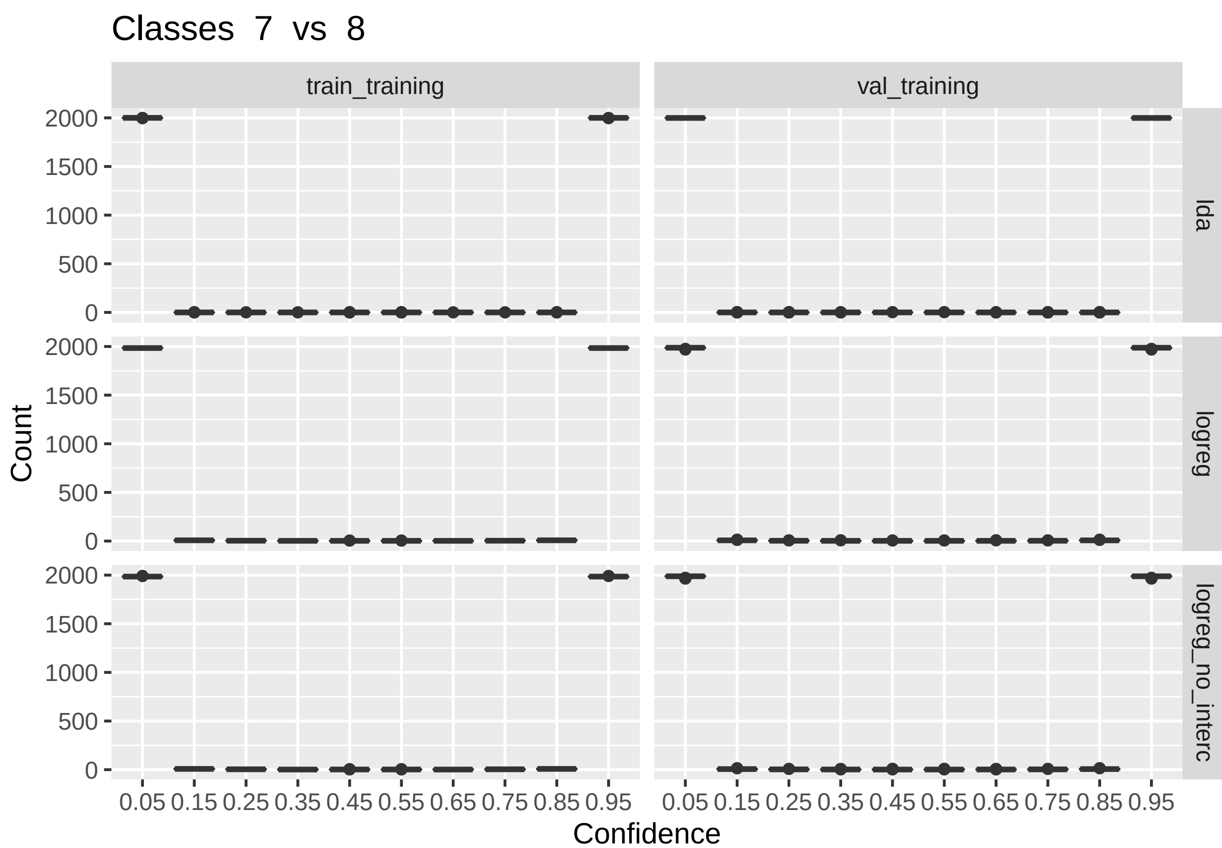

Classes 5 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

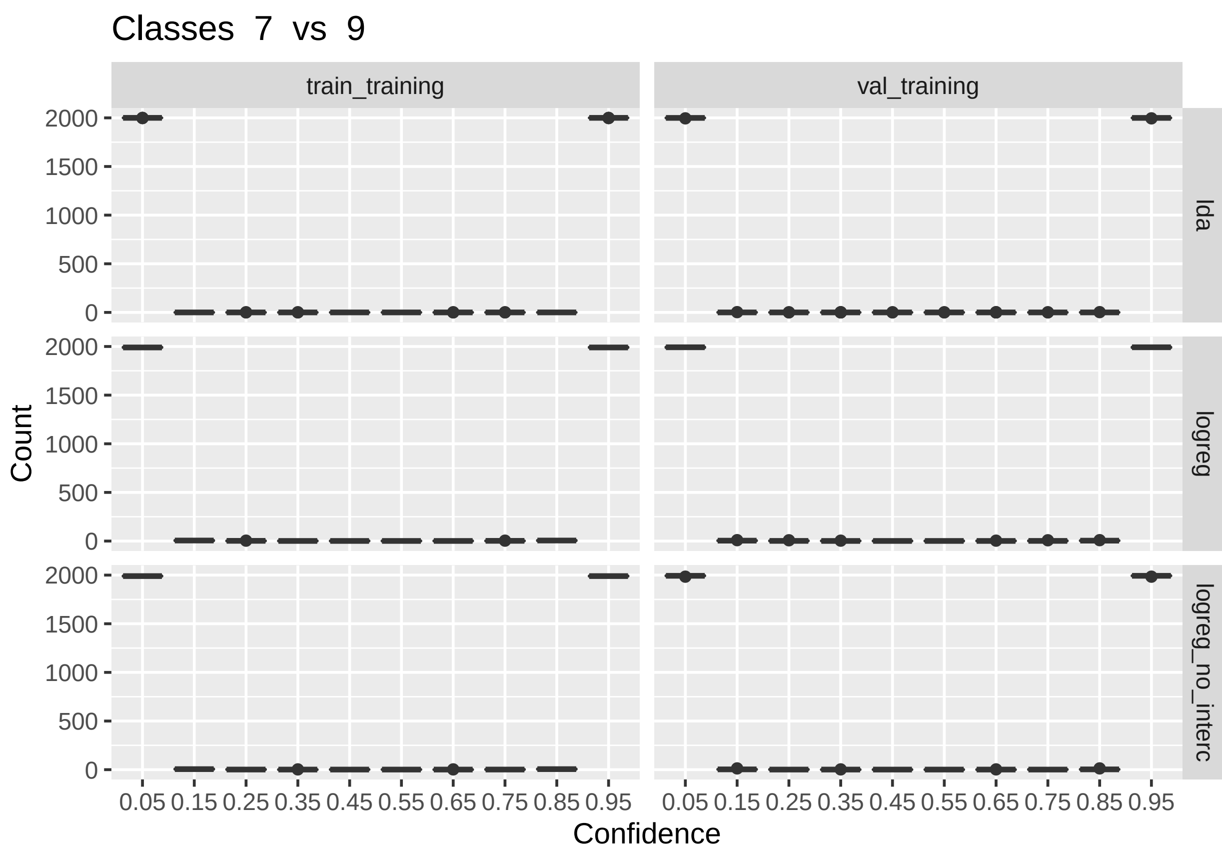

Classes 6 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence


Classes 6 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

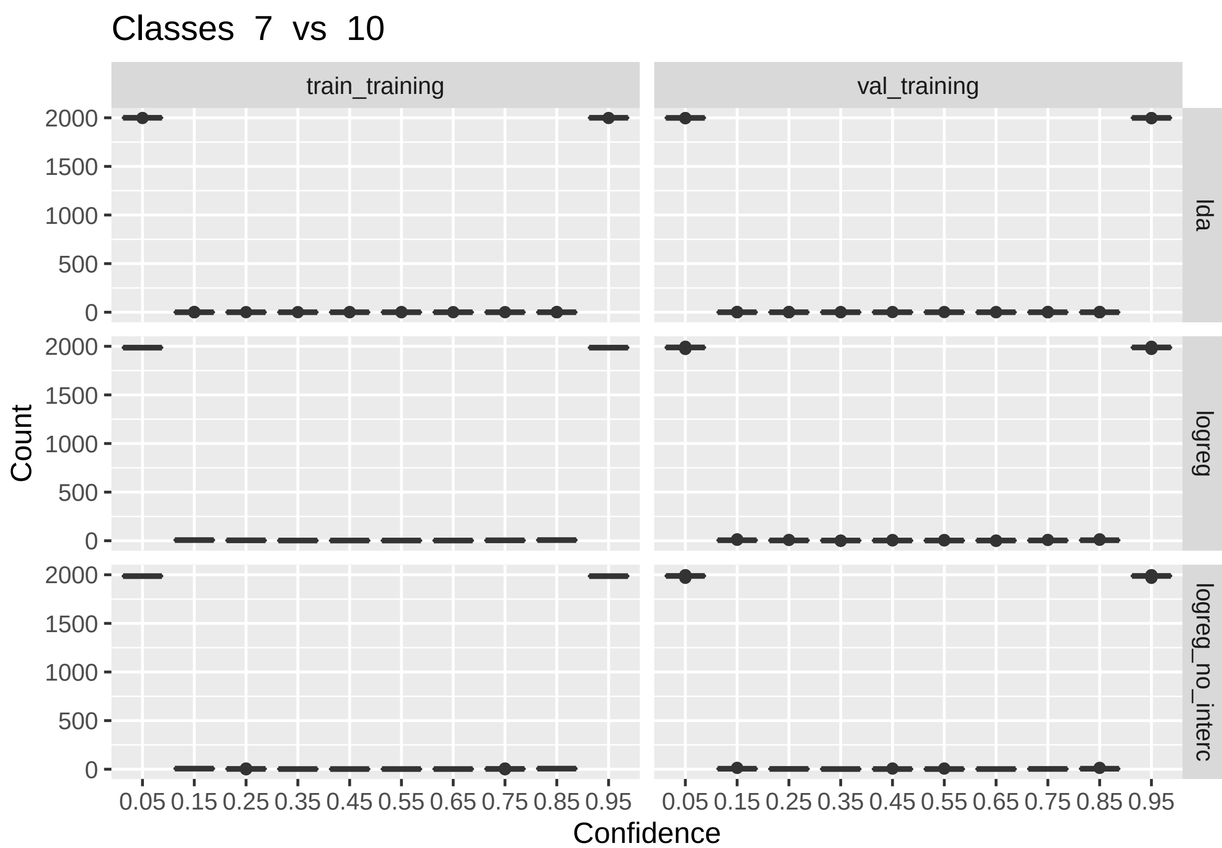

Classes 6 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence


Classes 6 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
Confidence
0
500
1000
1500
2000
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95


Classes 7 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

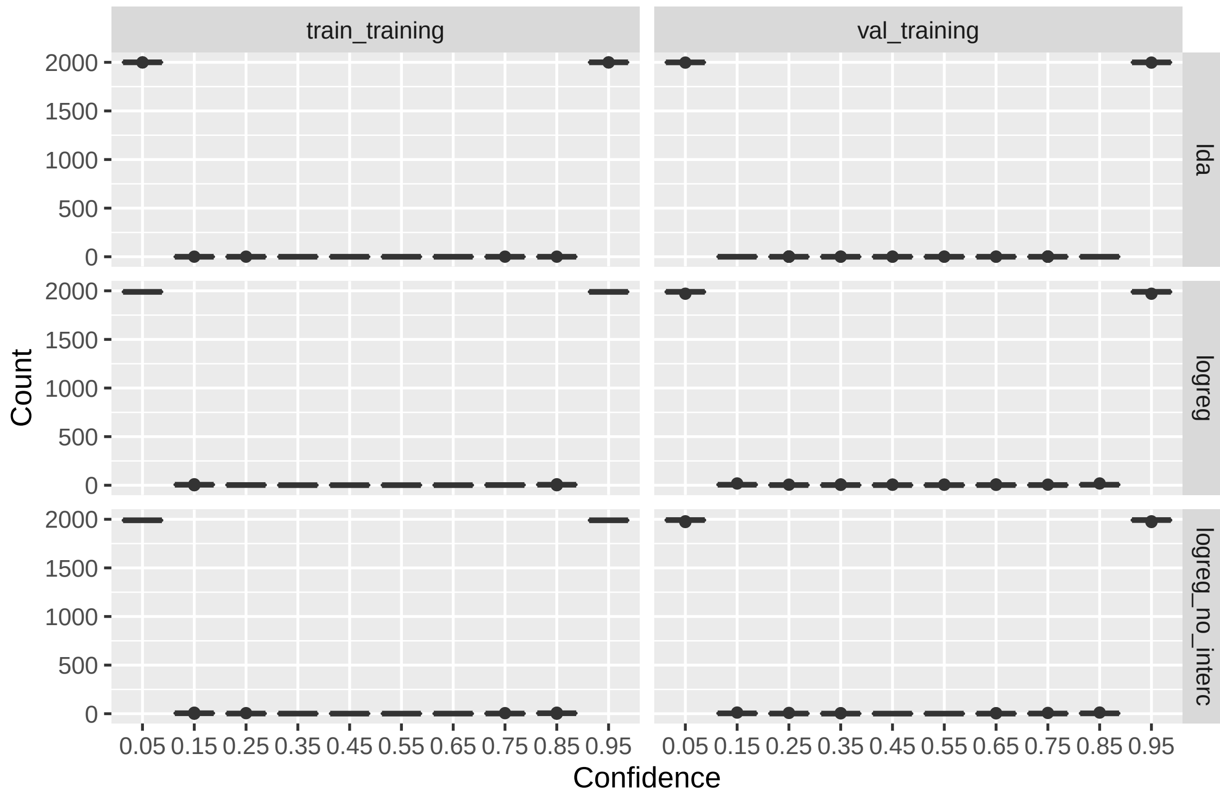


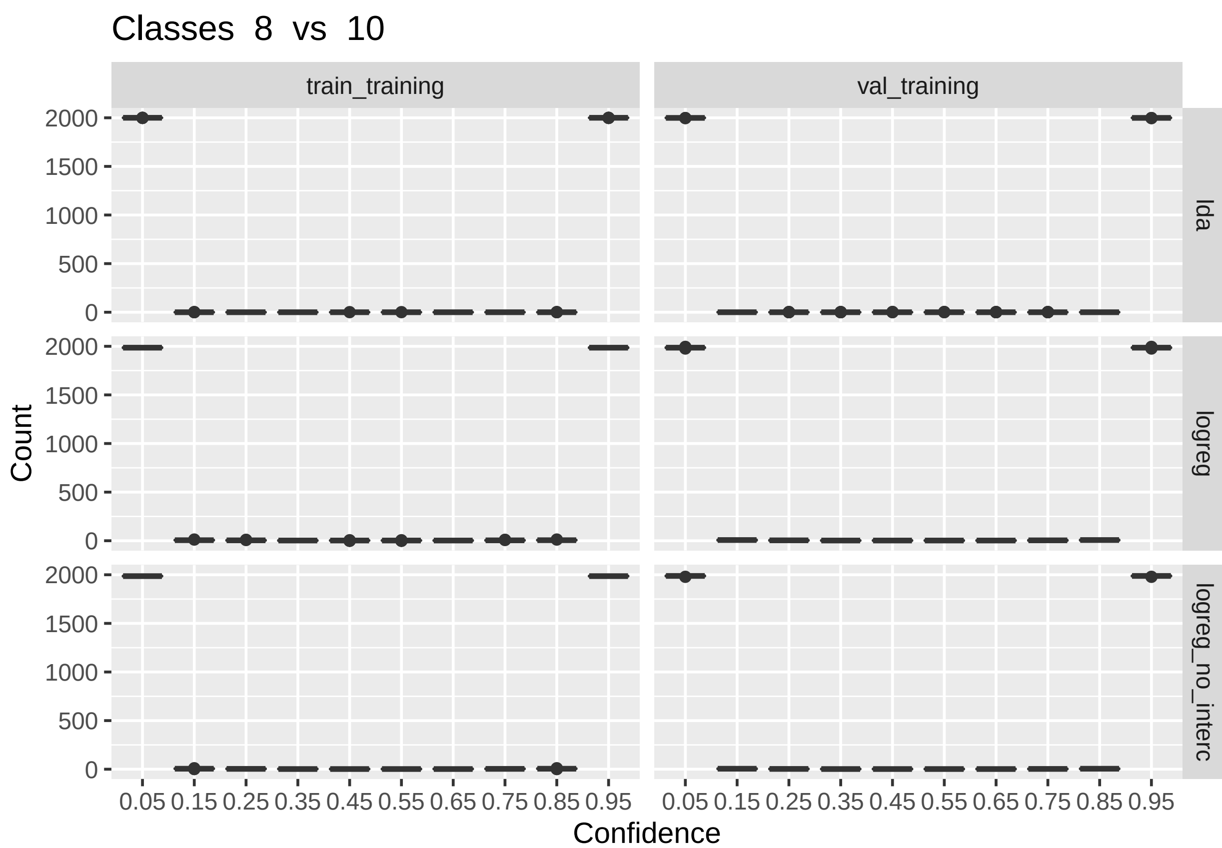
Classes 7 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

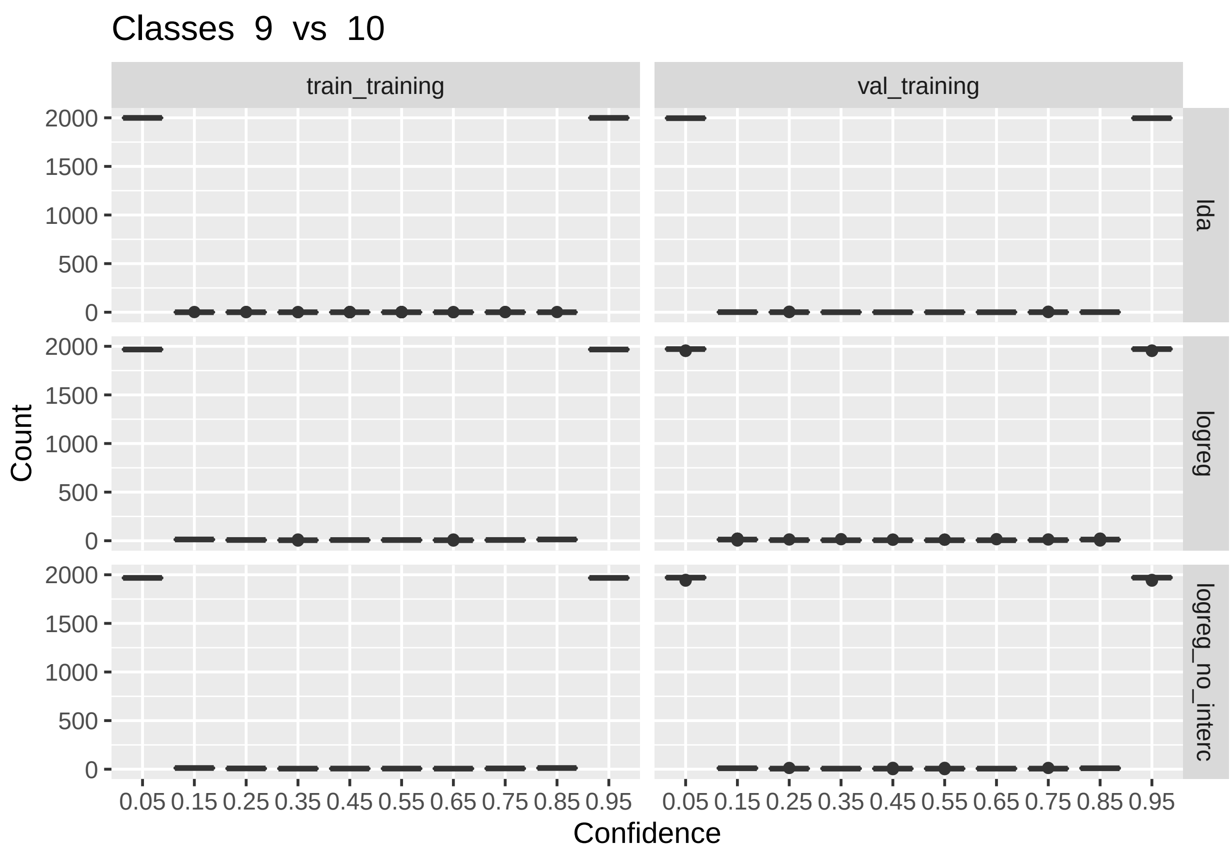


Classes 7 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

Classes 8 vs 9

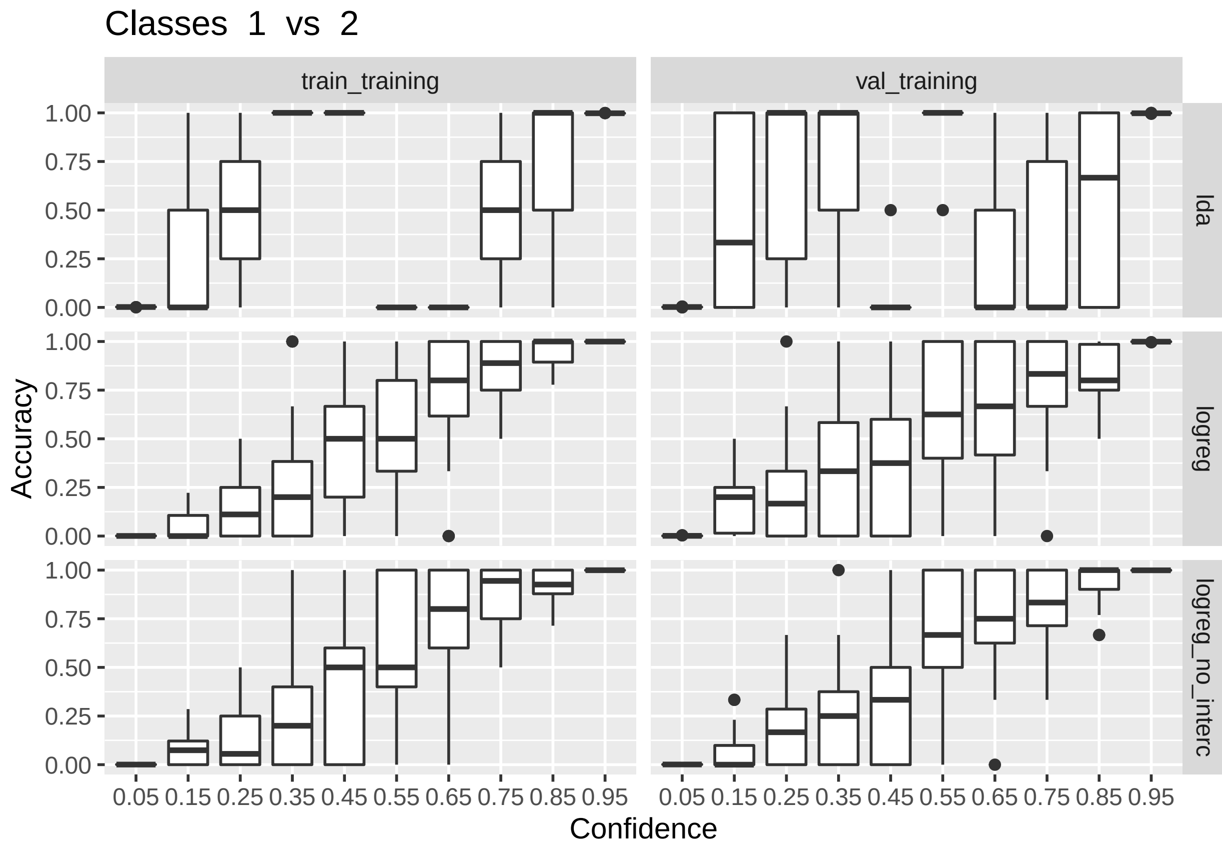



Classes 8 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
Count
2000
1500
1000
500
0
0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95
Confidence

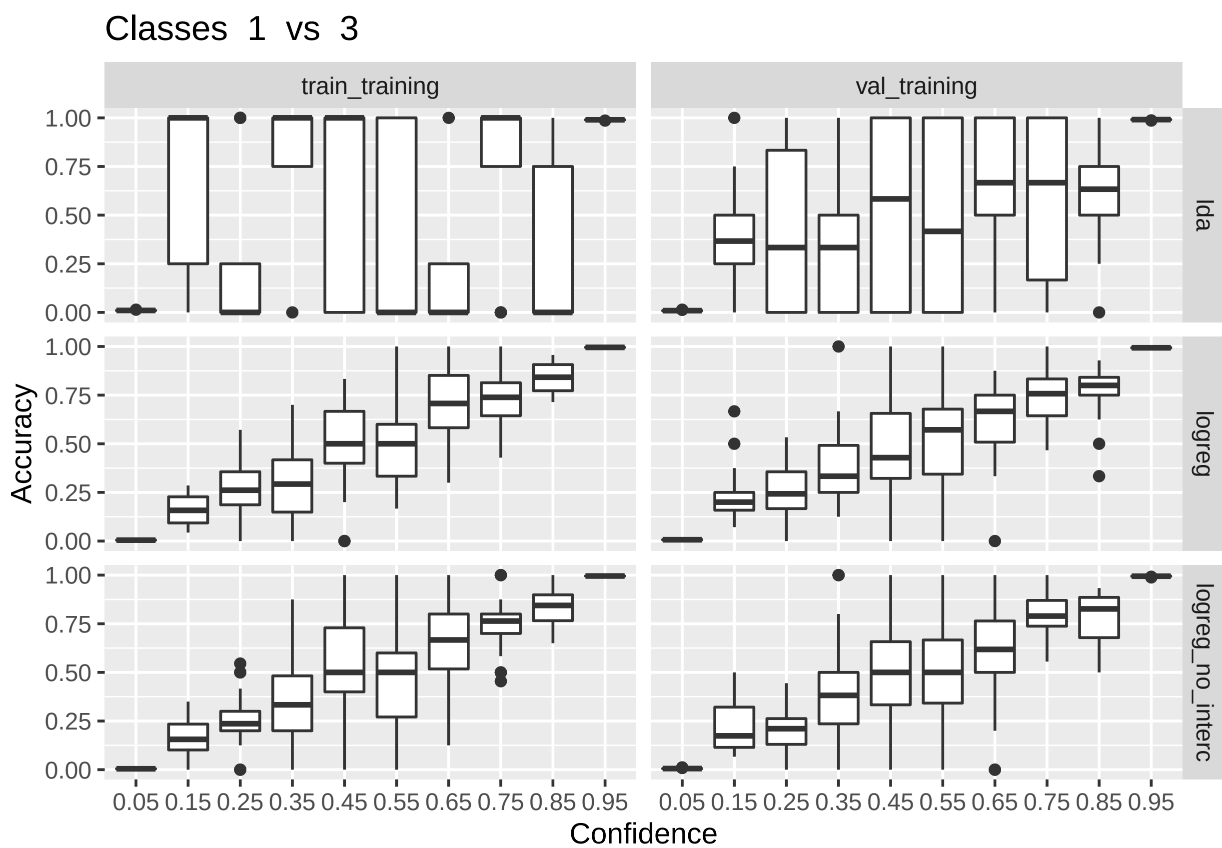



```
for (cl1 in 1:(classes - 1))
{
  for (cl2 in (cl1 + 1):classes)
  {
    cal_plt <- ens_pw_cal %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot() +
      geom_boxplot(mapping=aes(x=bin_c, y=bin_accuracy)) + facet_grid(cols=vars(train_set), rows=vars(c
      ggtitle(paste("Classes ", cl1, " vs ", cl2)) + xlab("Confidence") + ylab("Accuracy")

    print(cal_plt)
  }
}
```

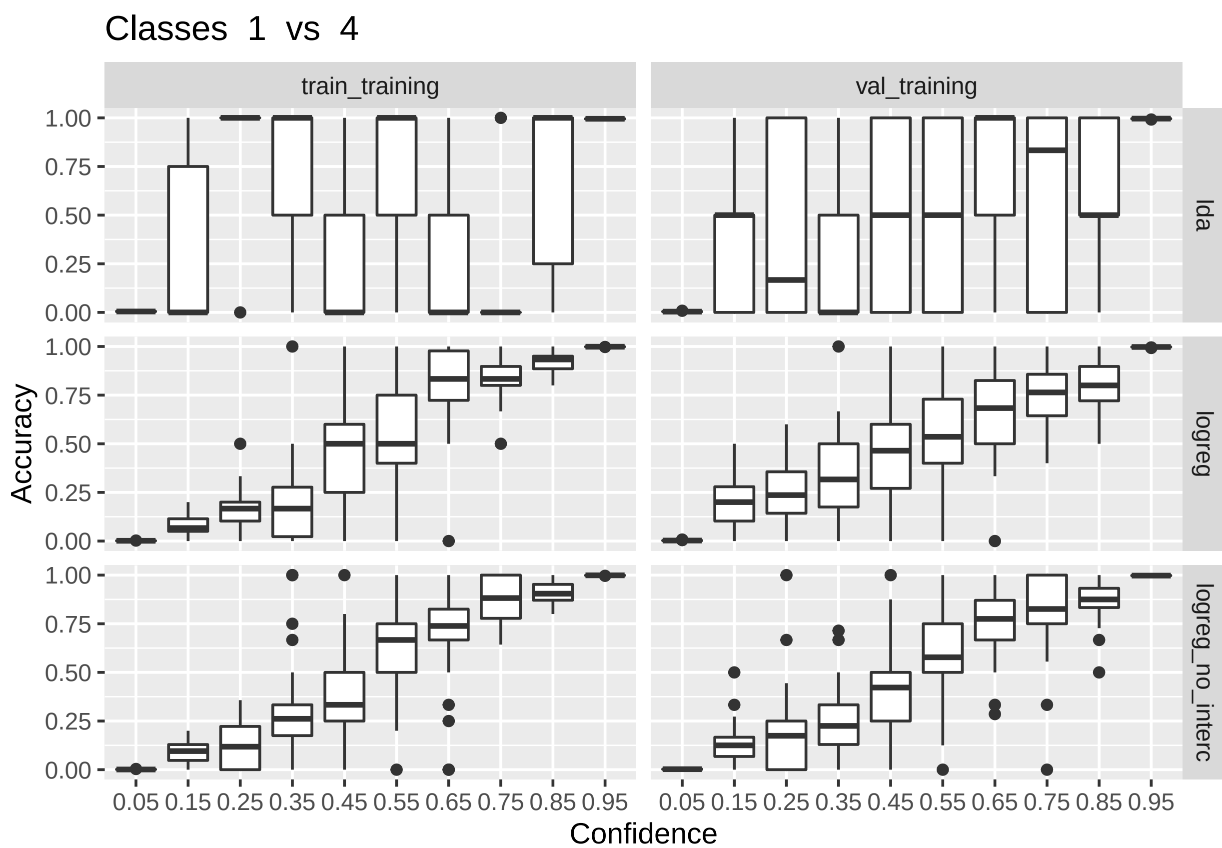
```
## Warning: Removed 438 rows containing non-finite values (stat_boxplot).
```



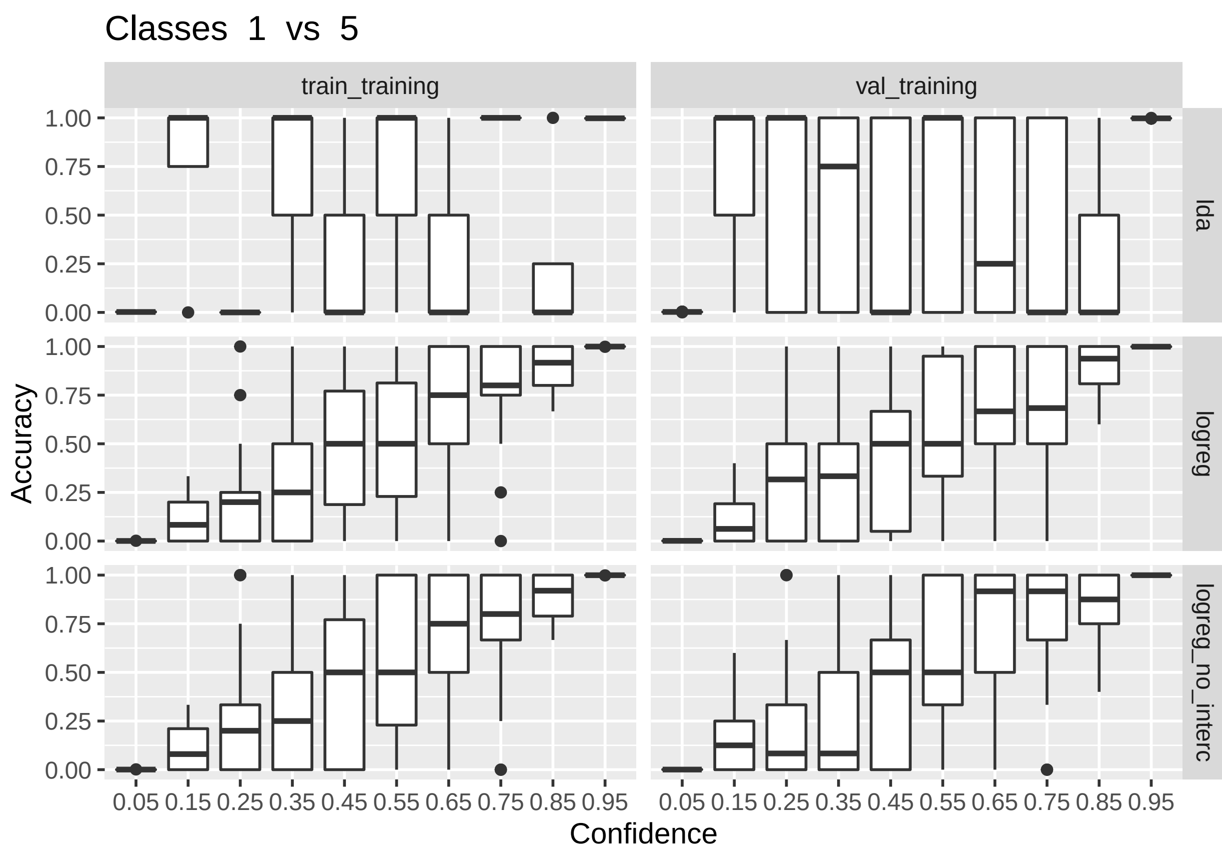

```
## Warning: Removed 216 rows containing non-finite values (stat_boxplot).
```



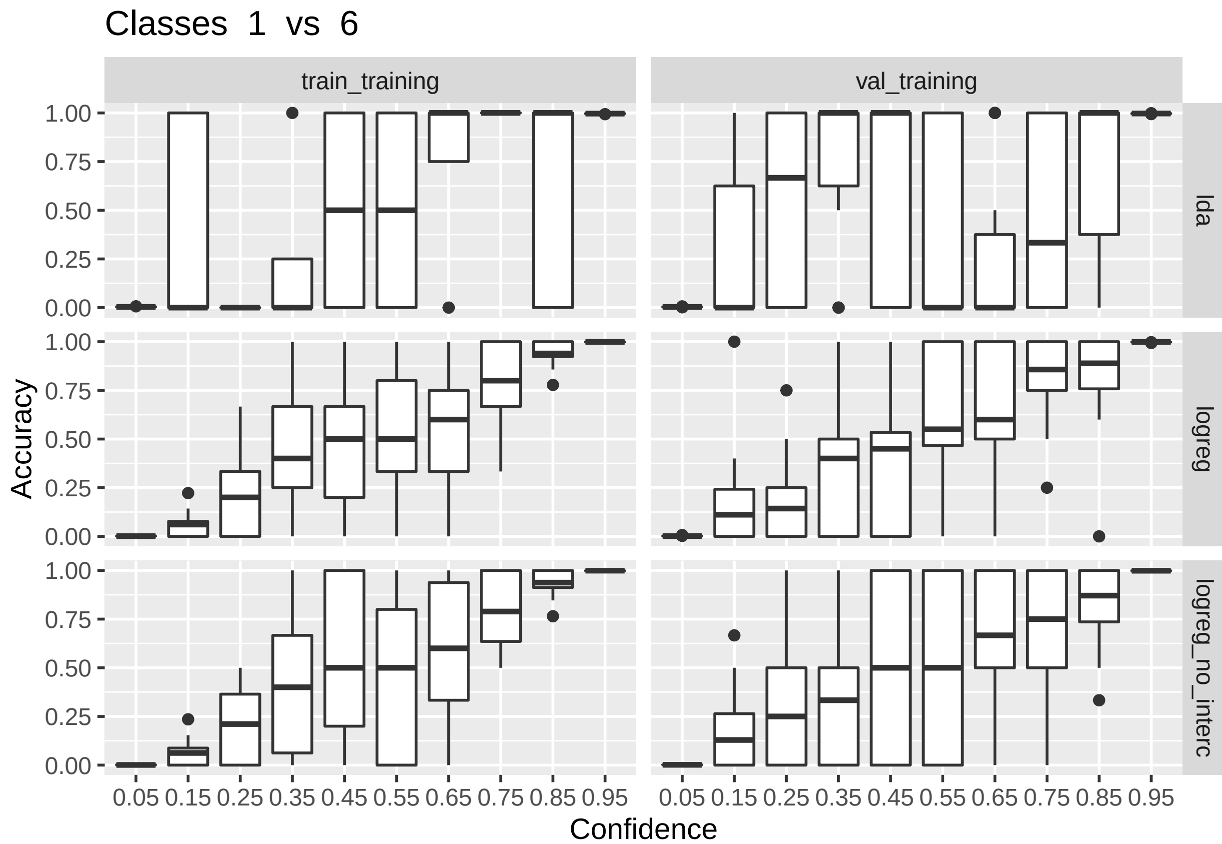

```
## Warning: Removed 262 rows containing non-finite values (stat_boxplot).
```



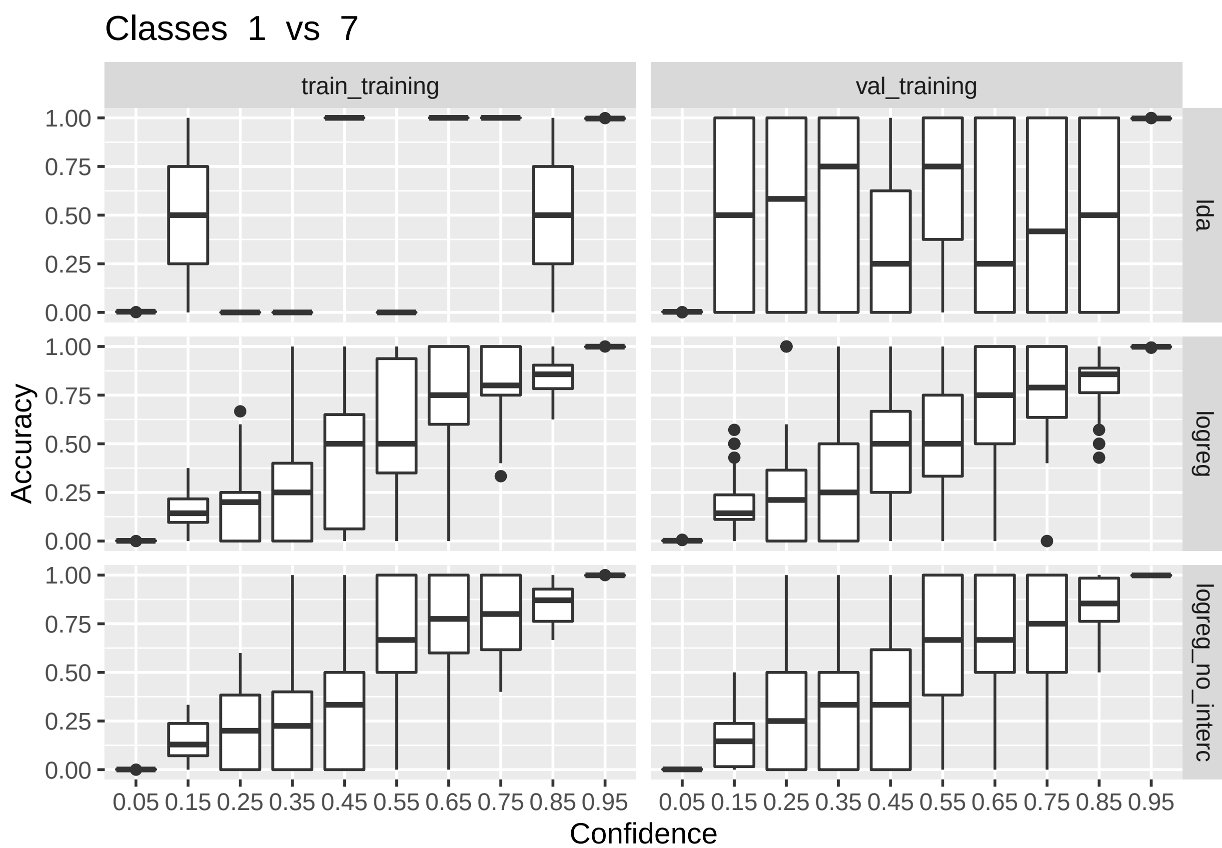

```
## Warning: Removed 428 rows containing non-finite values (stat_boxplot).
```



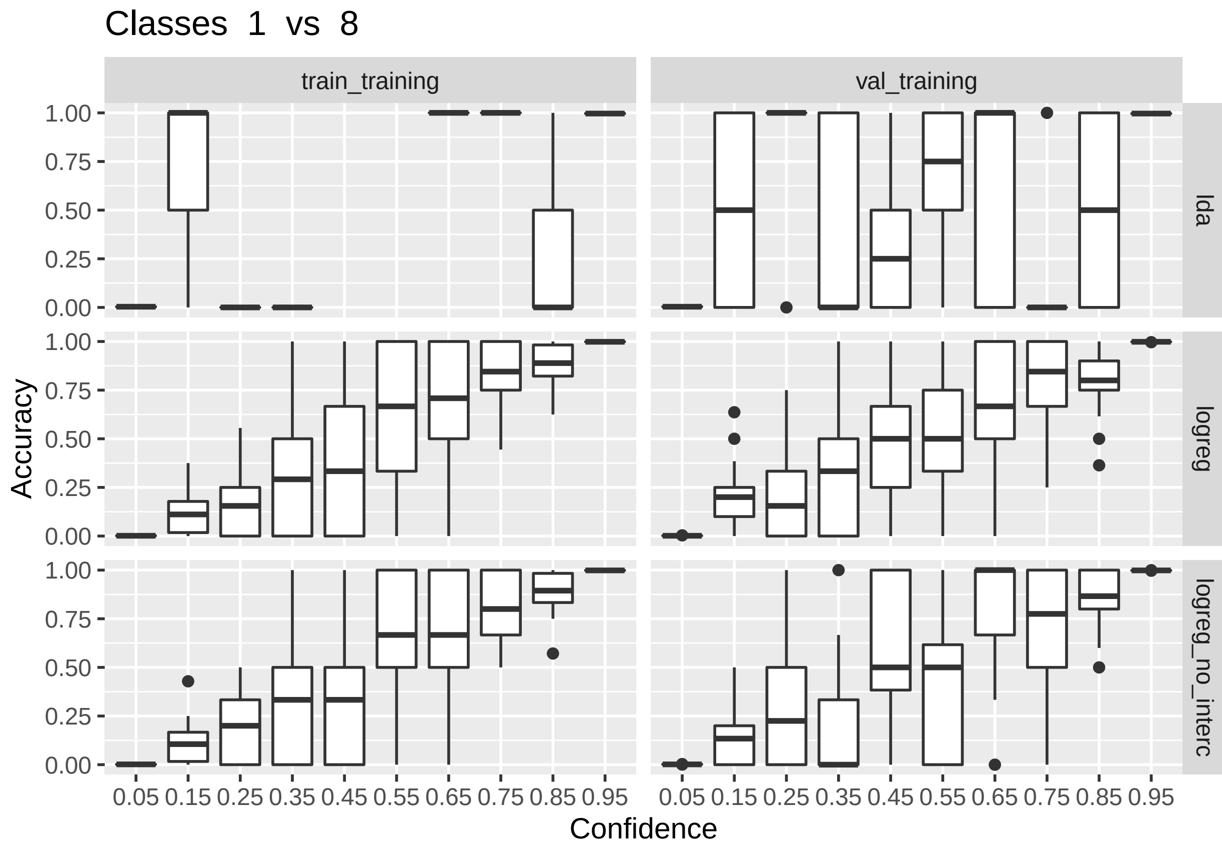

```
## Warning: Removed 380 rows containing non-finite values (stat_boxplot).
```



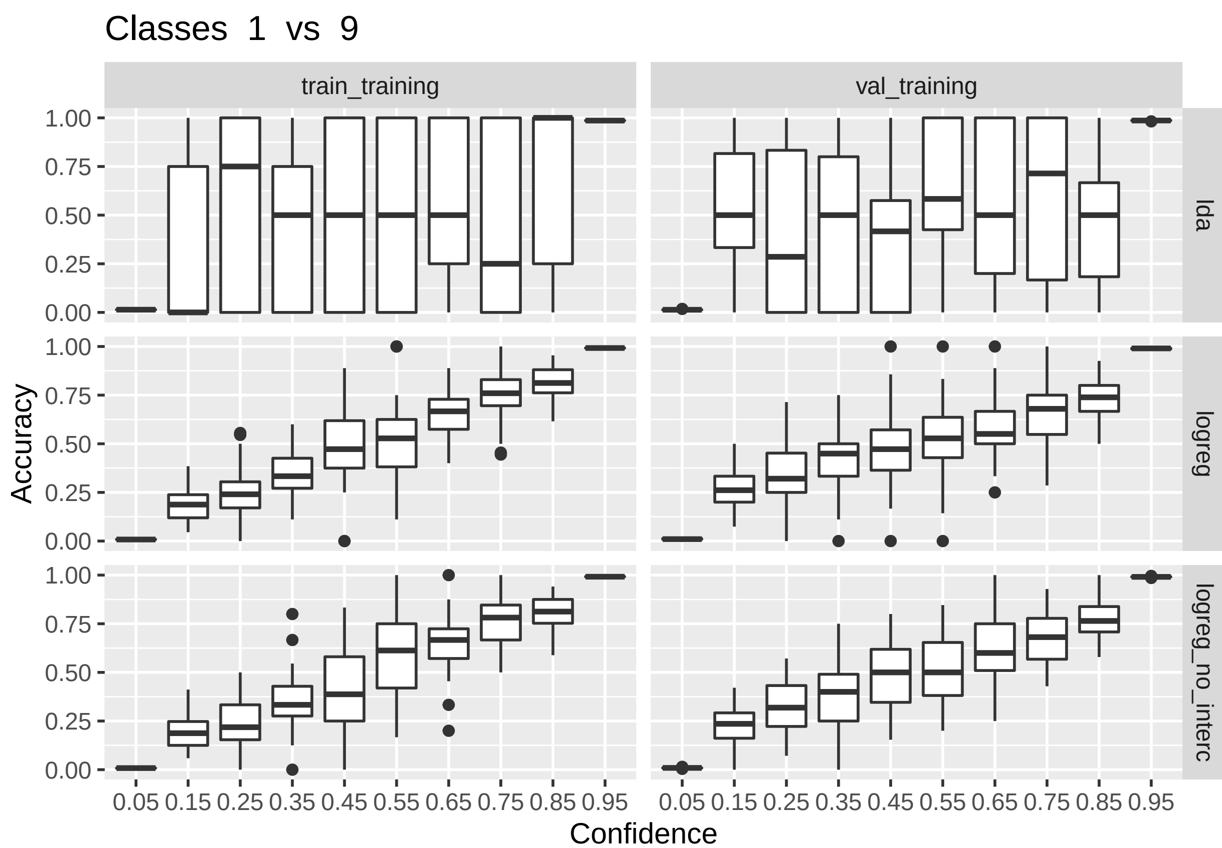

```
## Warning: Removed 388 rows containing non-finite values (stat_boxplot).
```



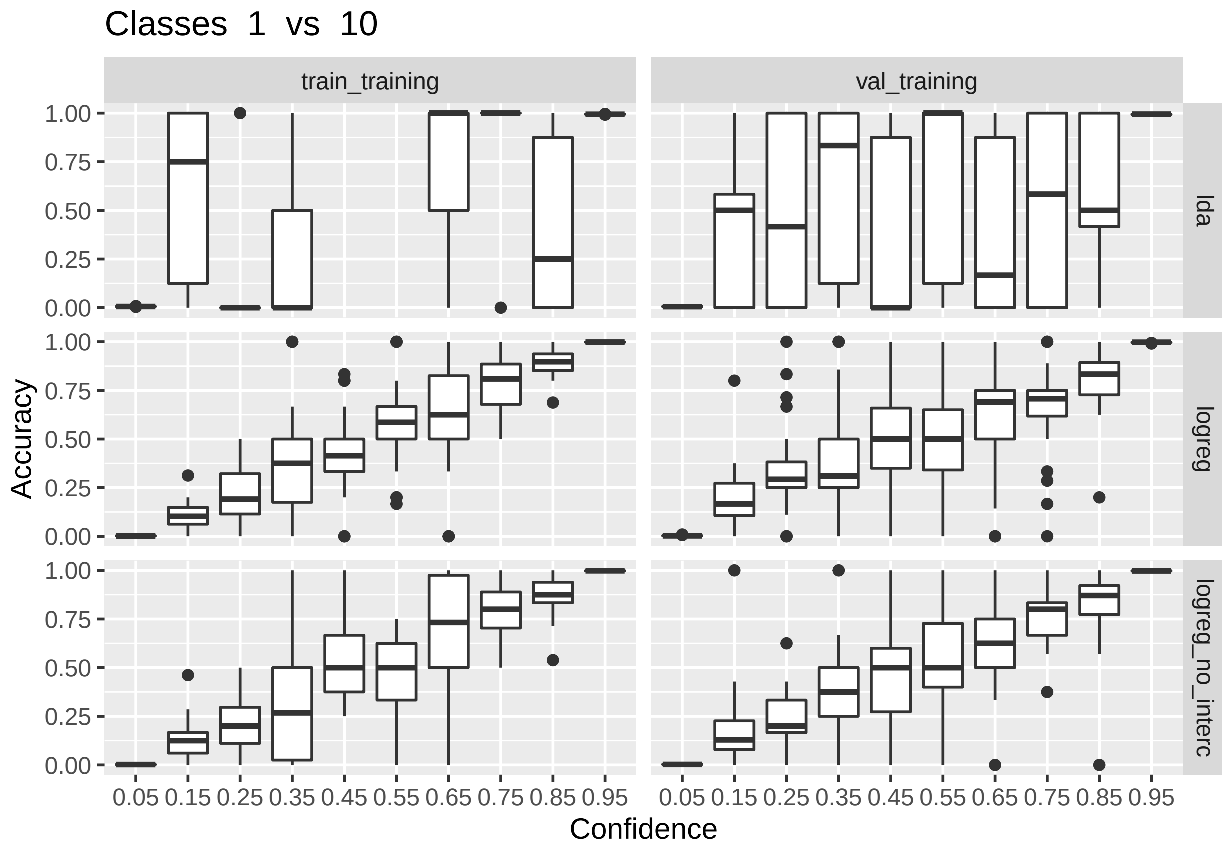

```
## Warning: Removed 420 rows containing non-finite values (stat_boxplot).
```



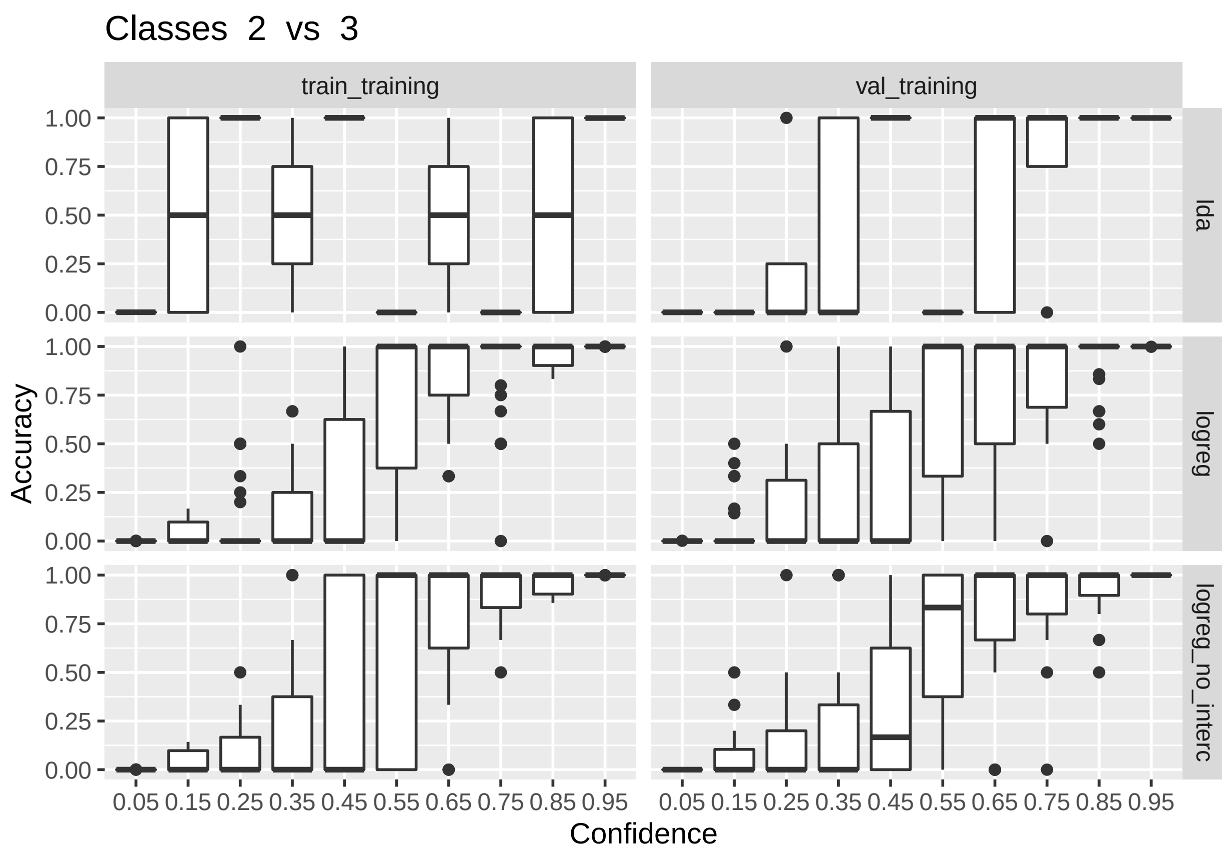

```
## Warning: Removed 230 rows containing non-finite values (stat_boxplot).
```



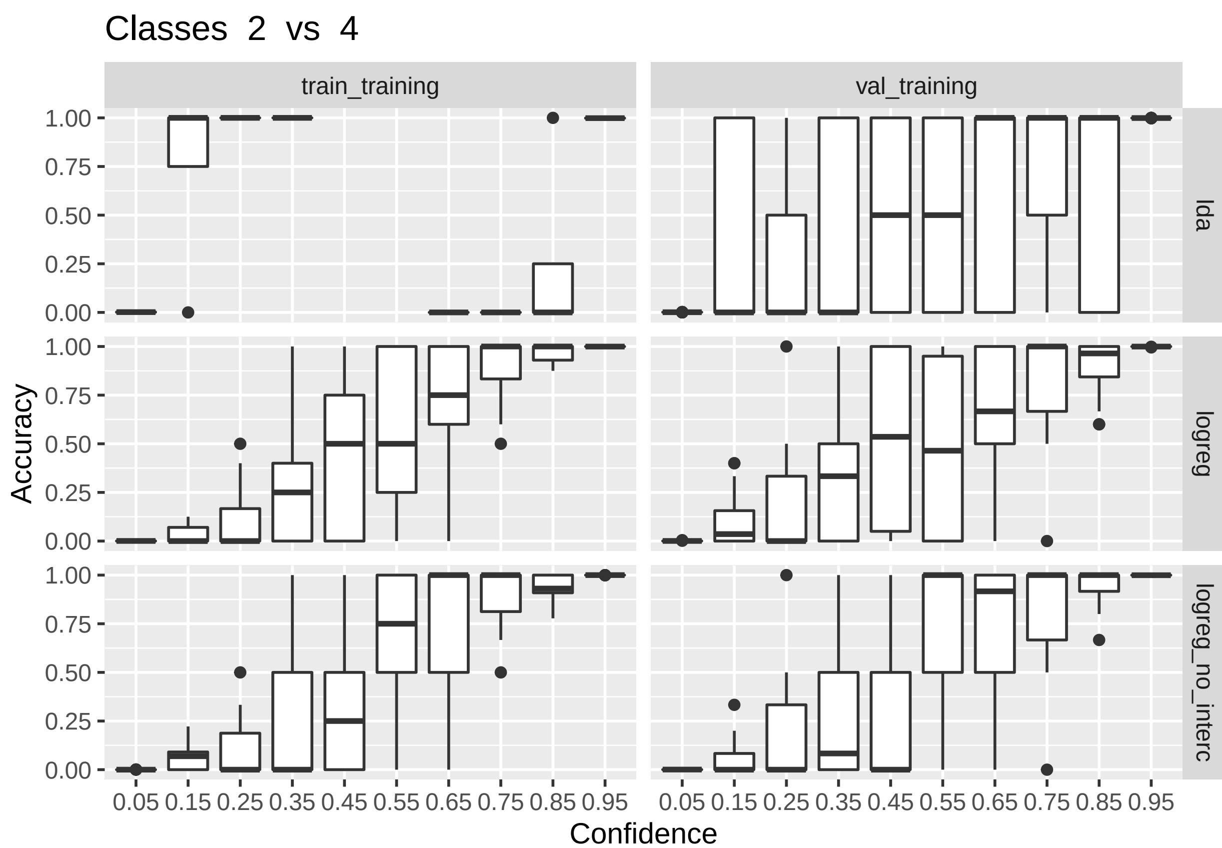

```
## Warning: Removed 344 rows containing non-finite values (stat_boxplot).
```



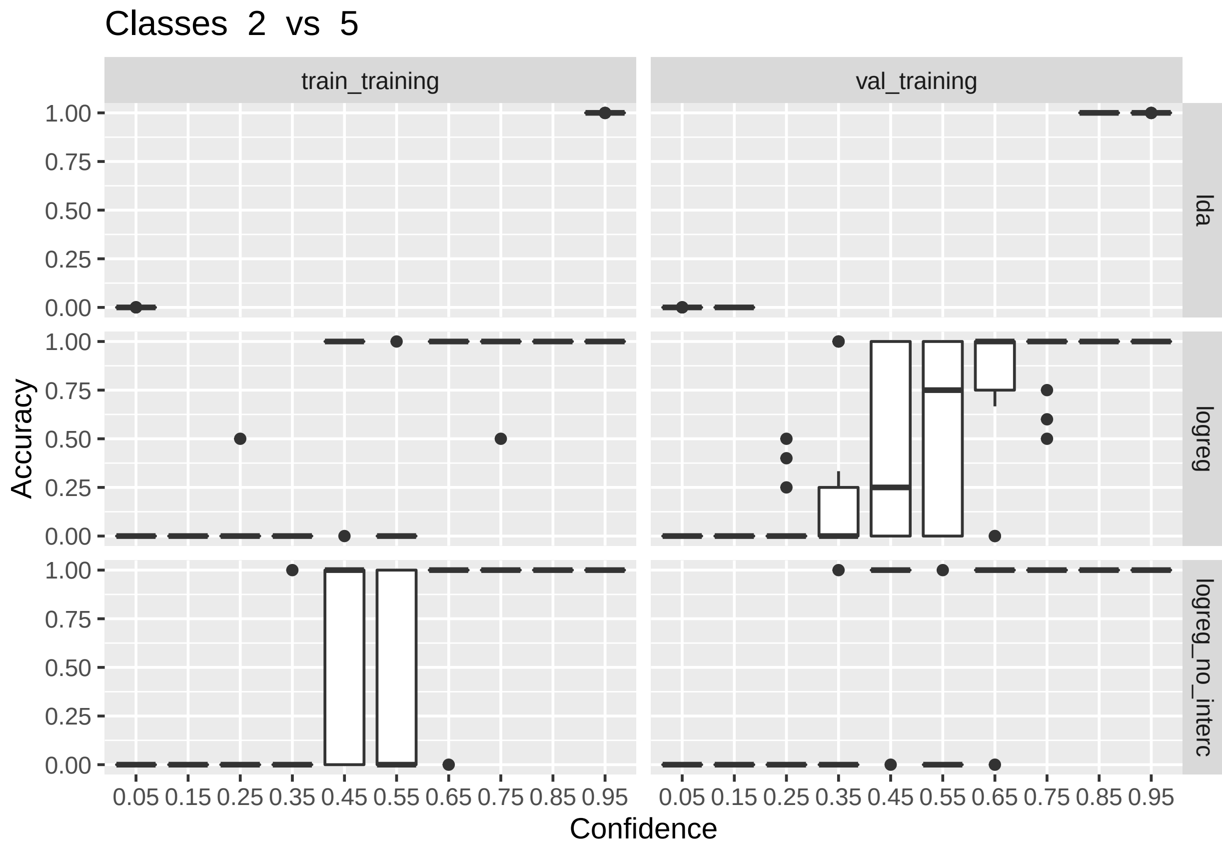

```
## Warning: Removed 598 rows containing non-finite values (stat_boxplot).
```



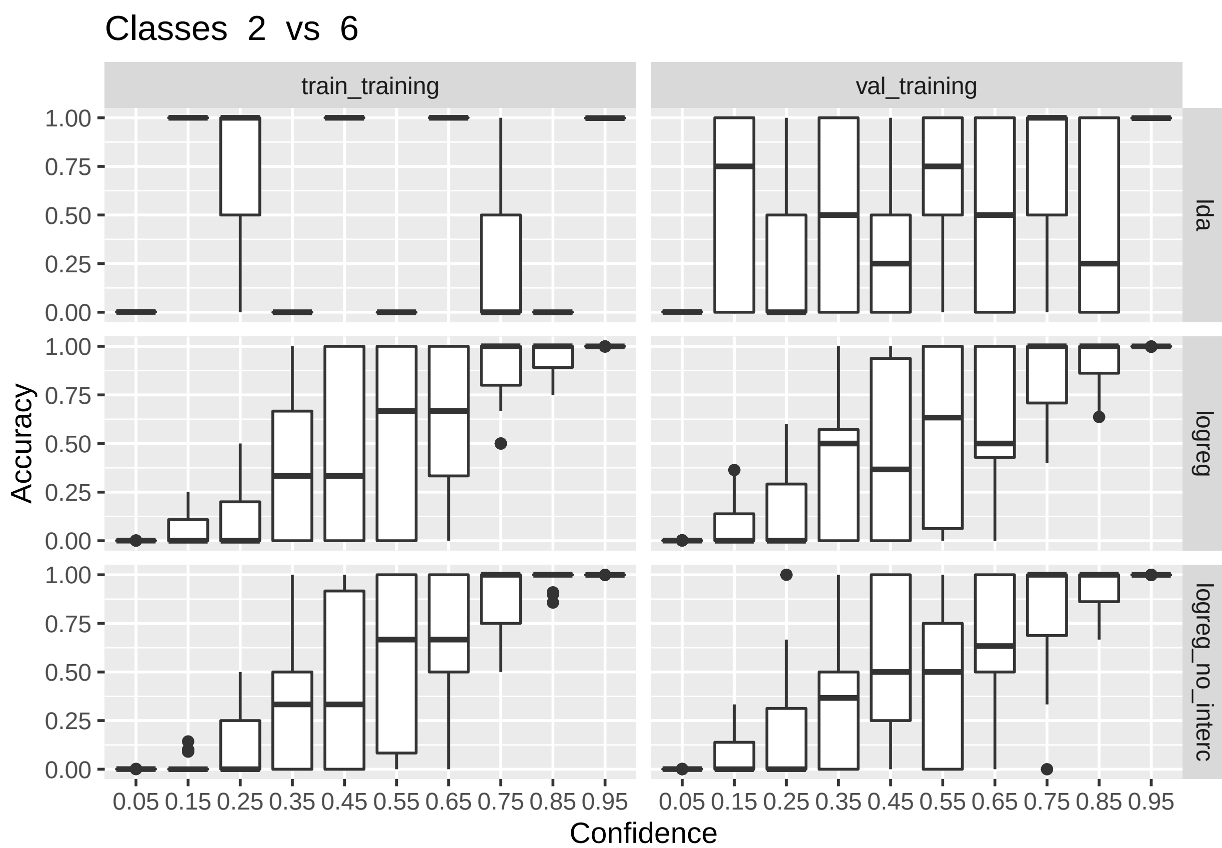

```
## Warning: Removed 486 rows containing non-finite values (stat_boxplot).
```



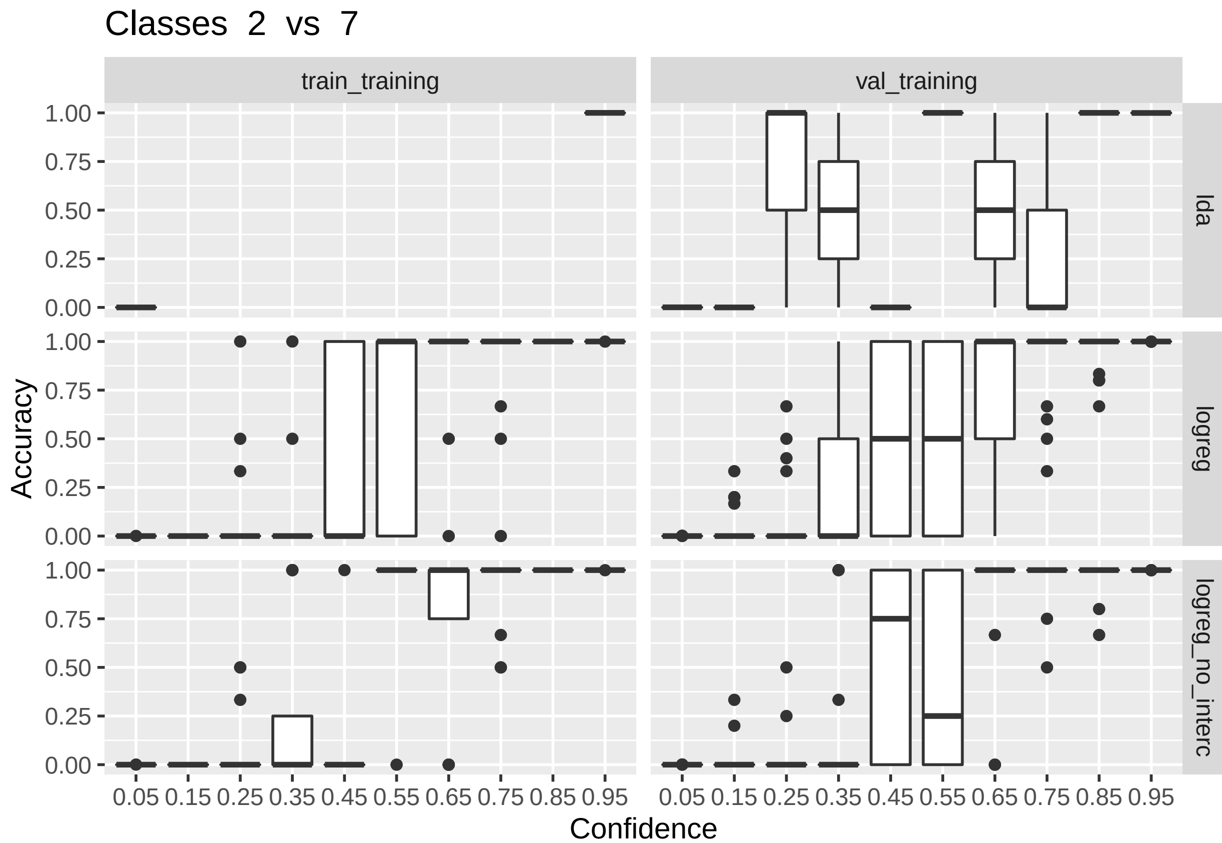

```
## Warning: Removed 828 rows containing non-finite values (stat_boxplot).
```



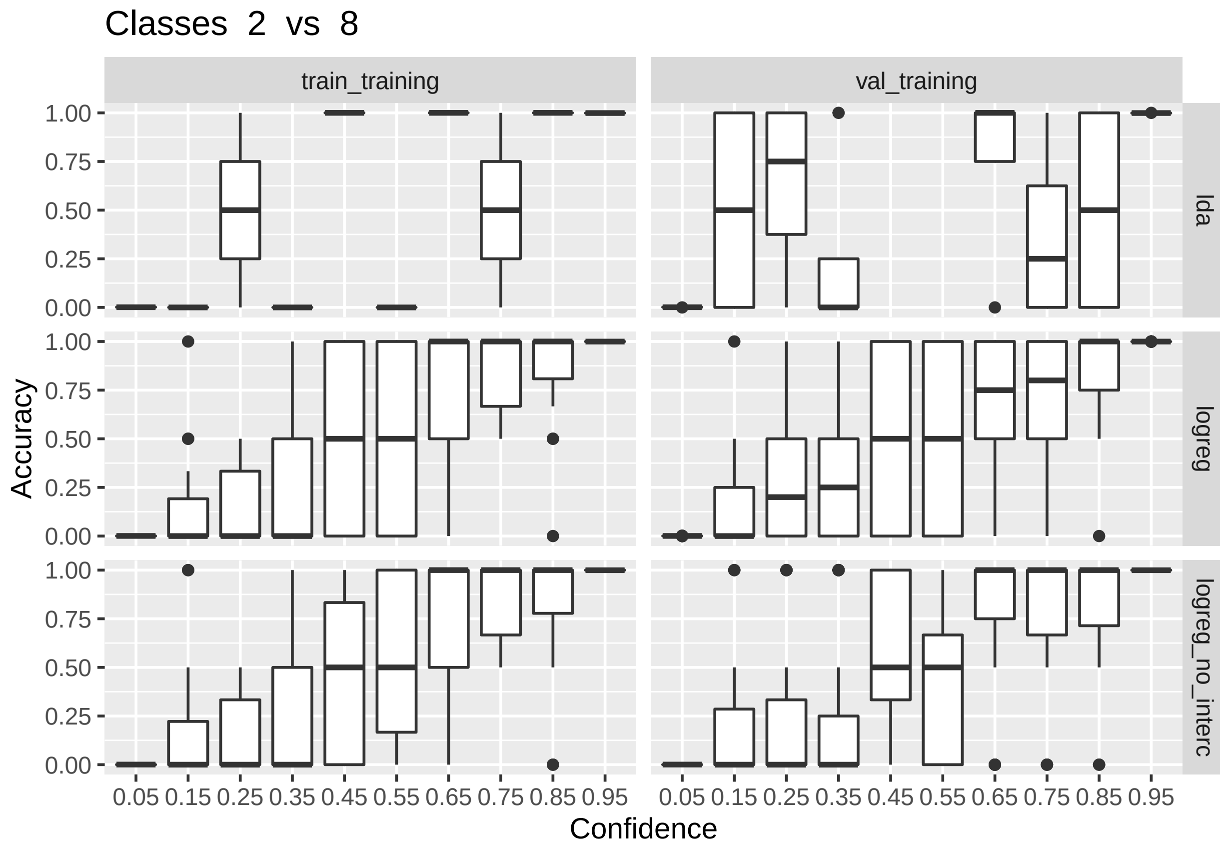

```
## Warning: Removed 472 rows containing non-finite values (stat_boxplot).
```



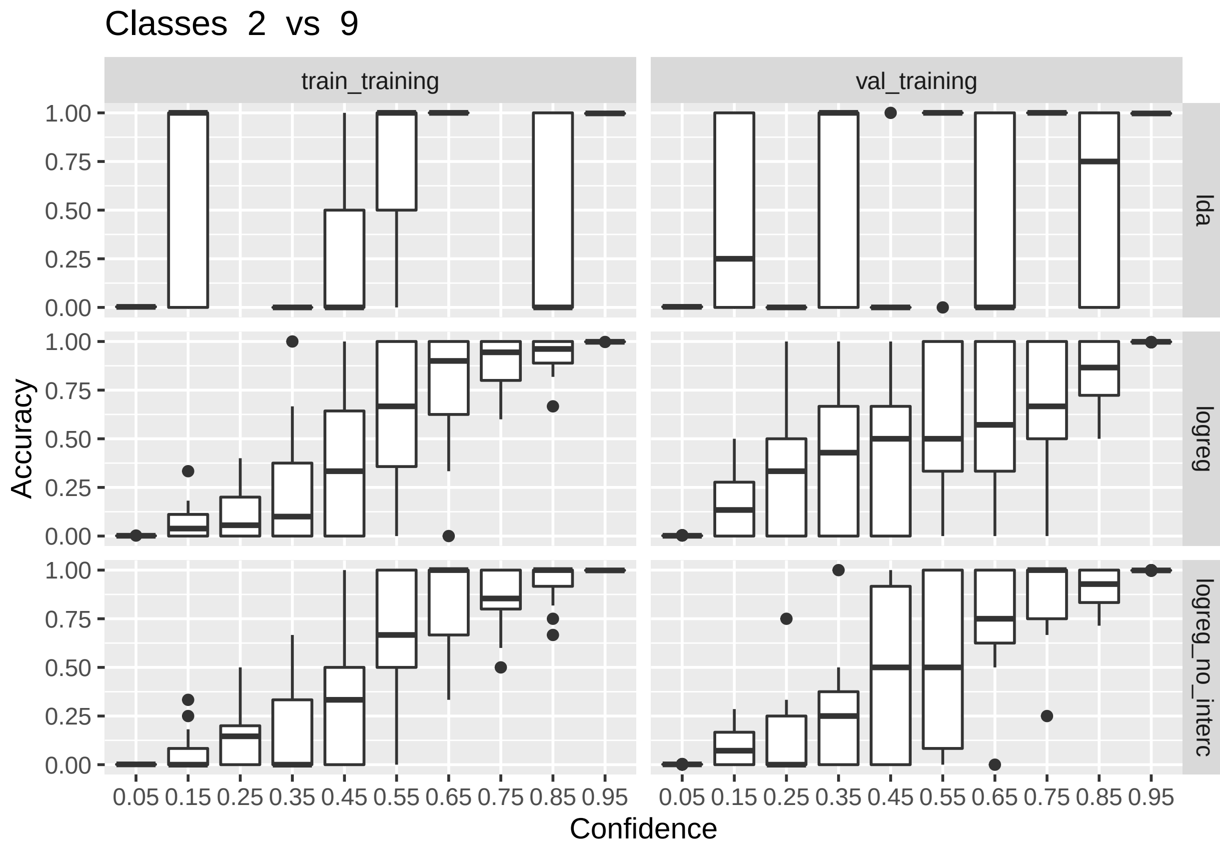

```
## Warning: Removed 772 rows containing non-finite values (stat_boxplot).
```



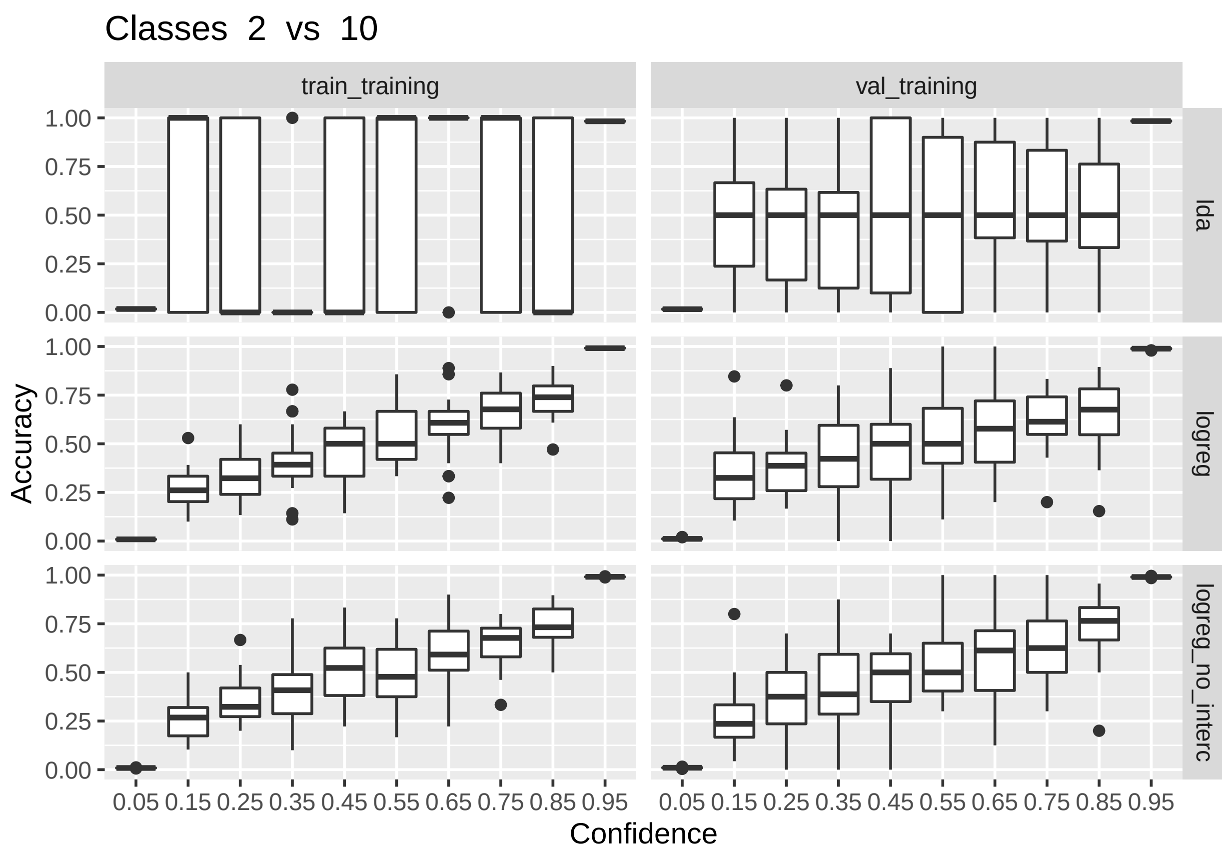

```
## Warning: Removed 542 rows containing non-finite values (stat_boxplot).
```



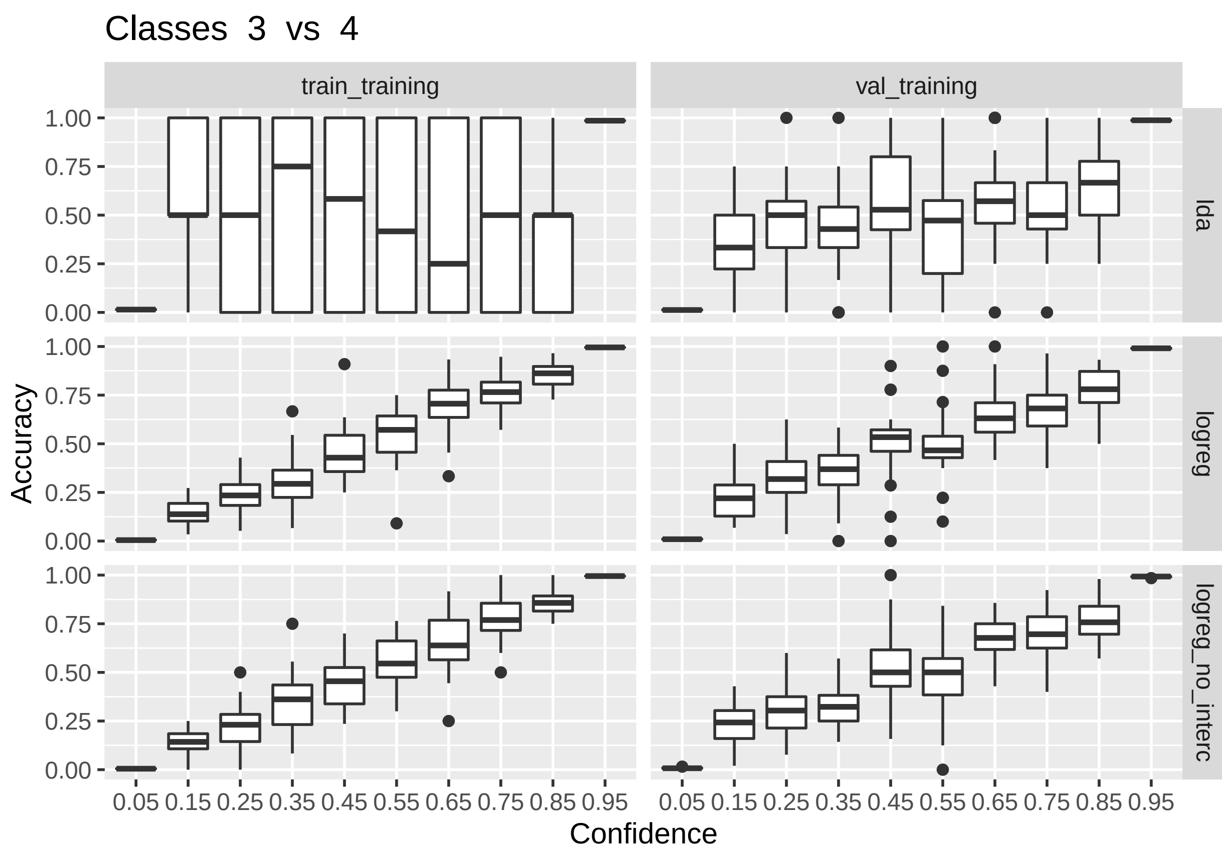

```
## Warning: Removed 448 rows containing non-finite values (stat_boxplot).
```



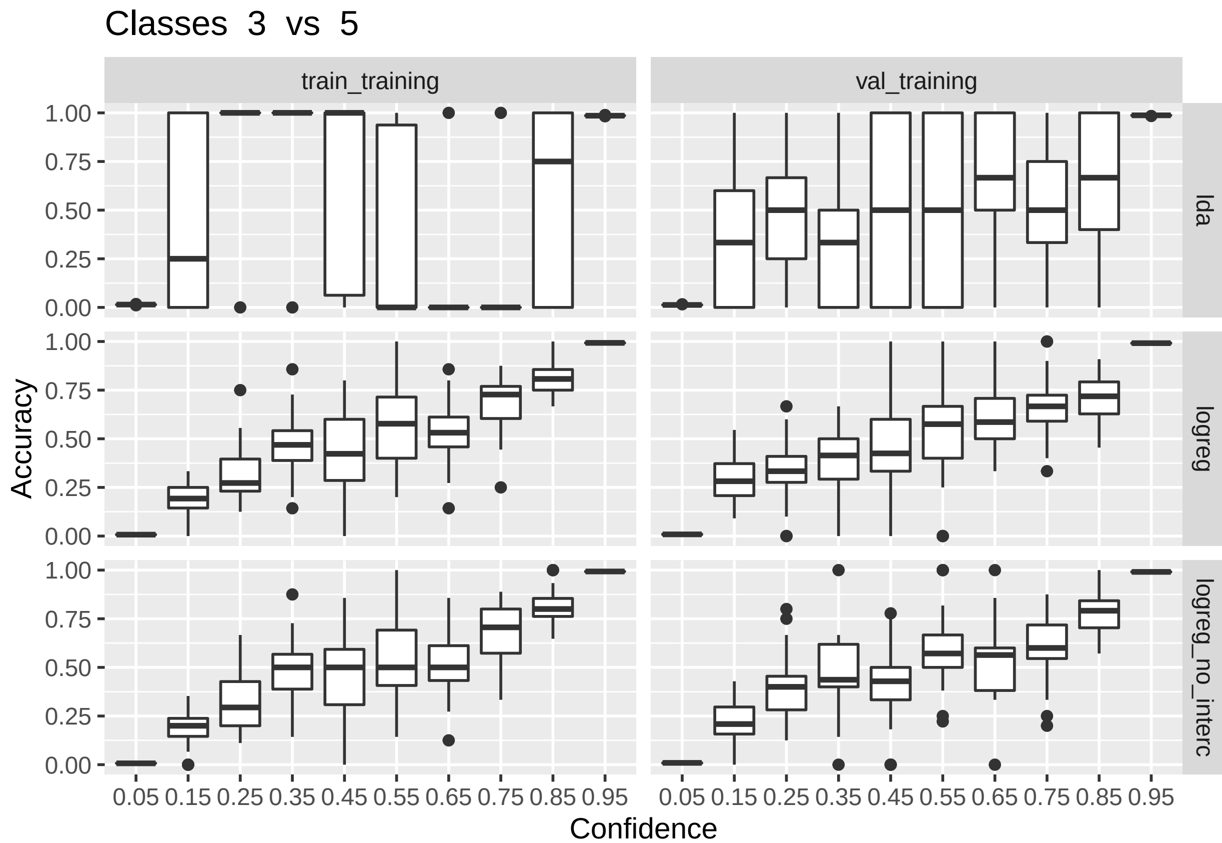

```
## Warning: Removed 220 rows containing non-finite values (stat_boxplot).
```



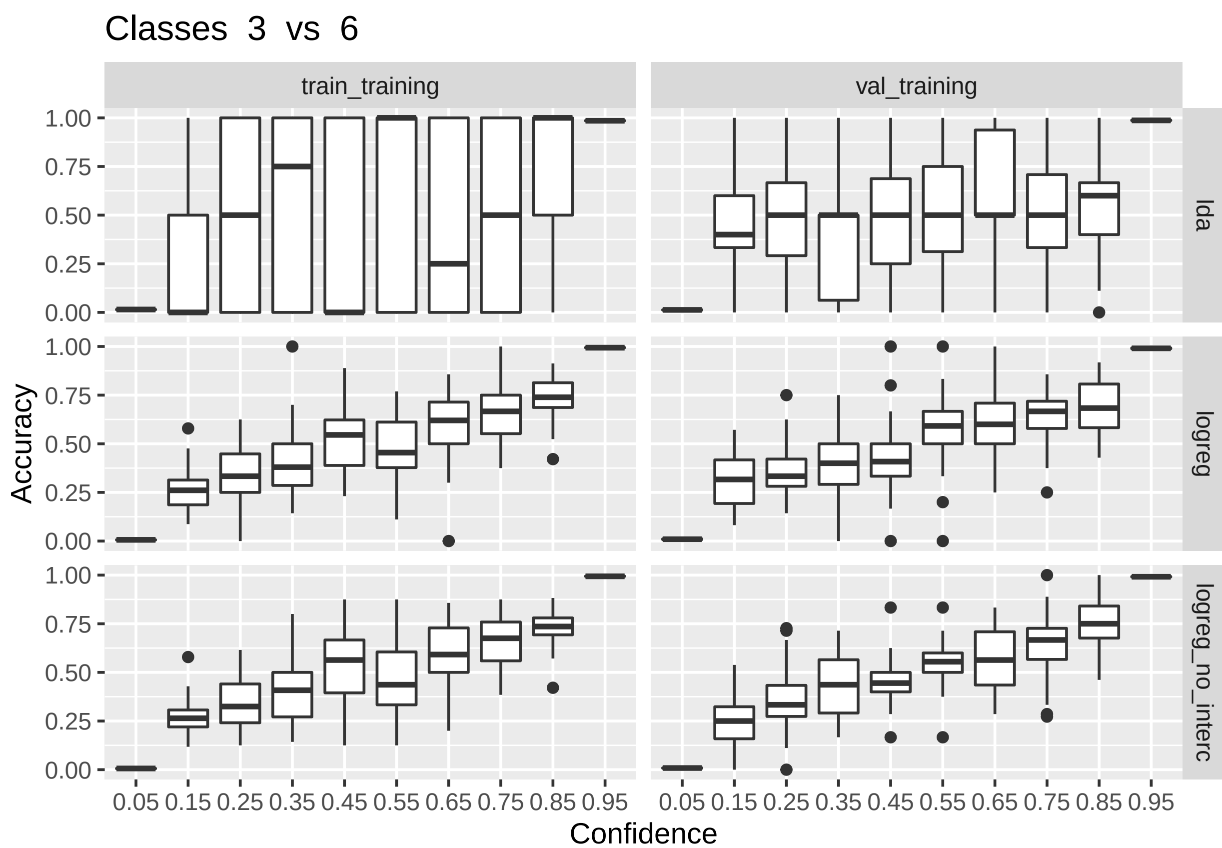

```
## Warning: Removed 124 rows containing non-finite values (stat_boxplot).
```



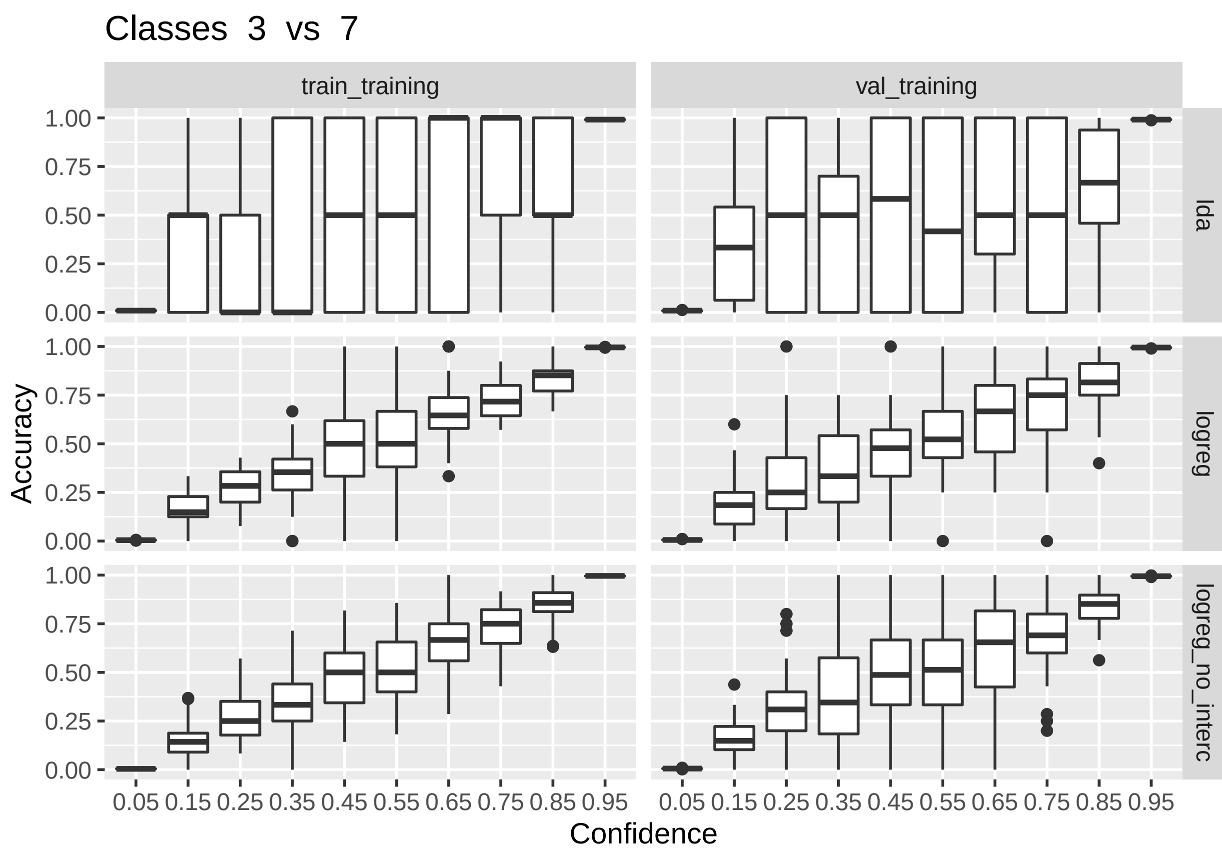

```
## Warning: Removed 226 rows containing non-finite values (stat_boxplot).
```



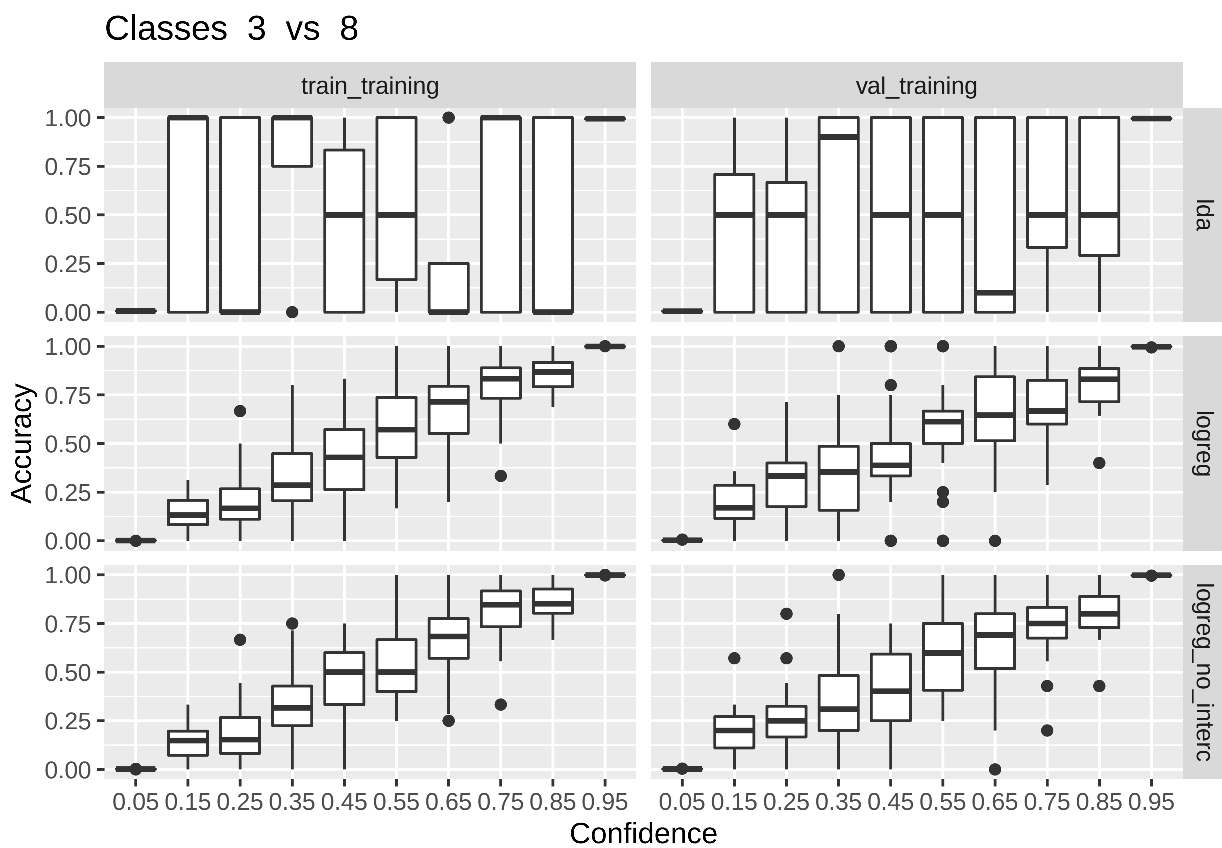

```
## Warning: Removed 170 rows containing non-finite values (stat_boxplot).
```



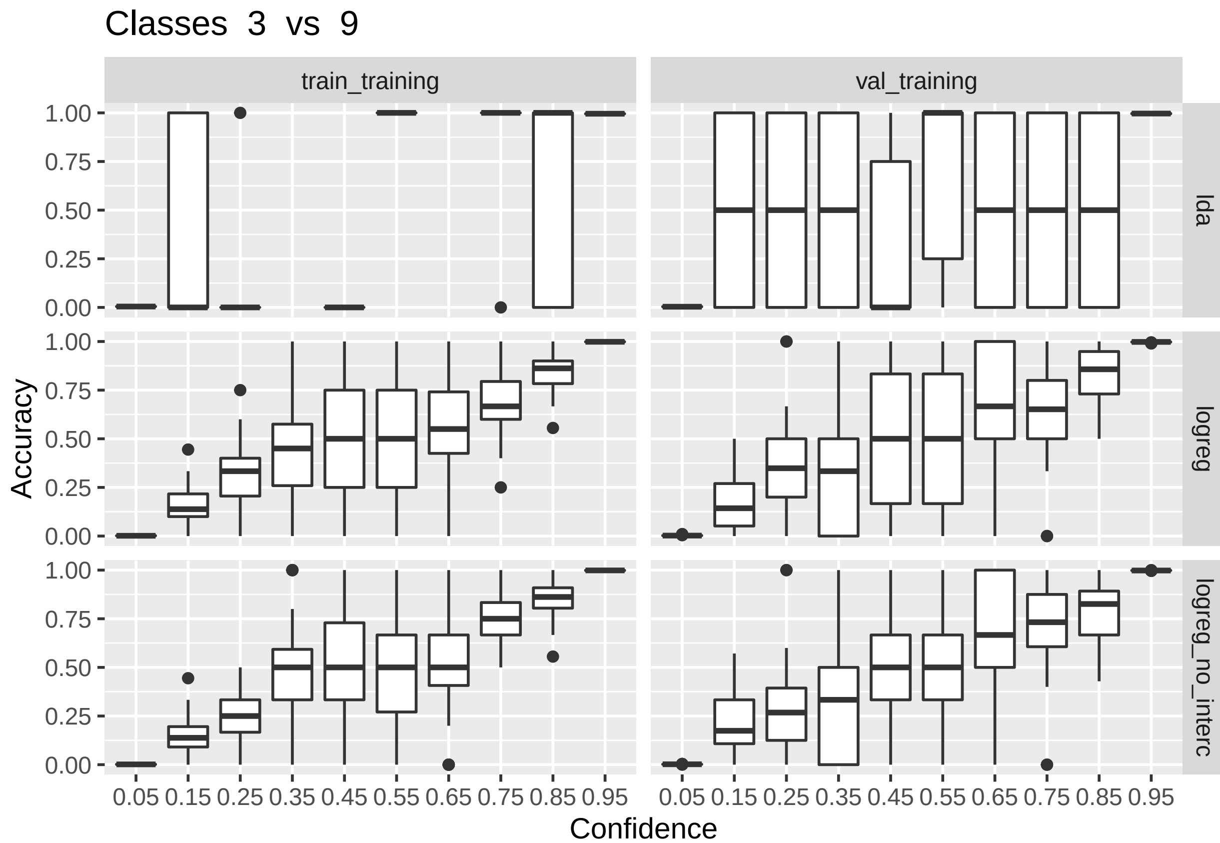

```
## Warning: Removed 260 rows containing non-finite values (stat_boxplot).
```



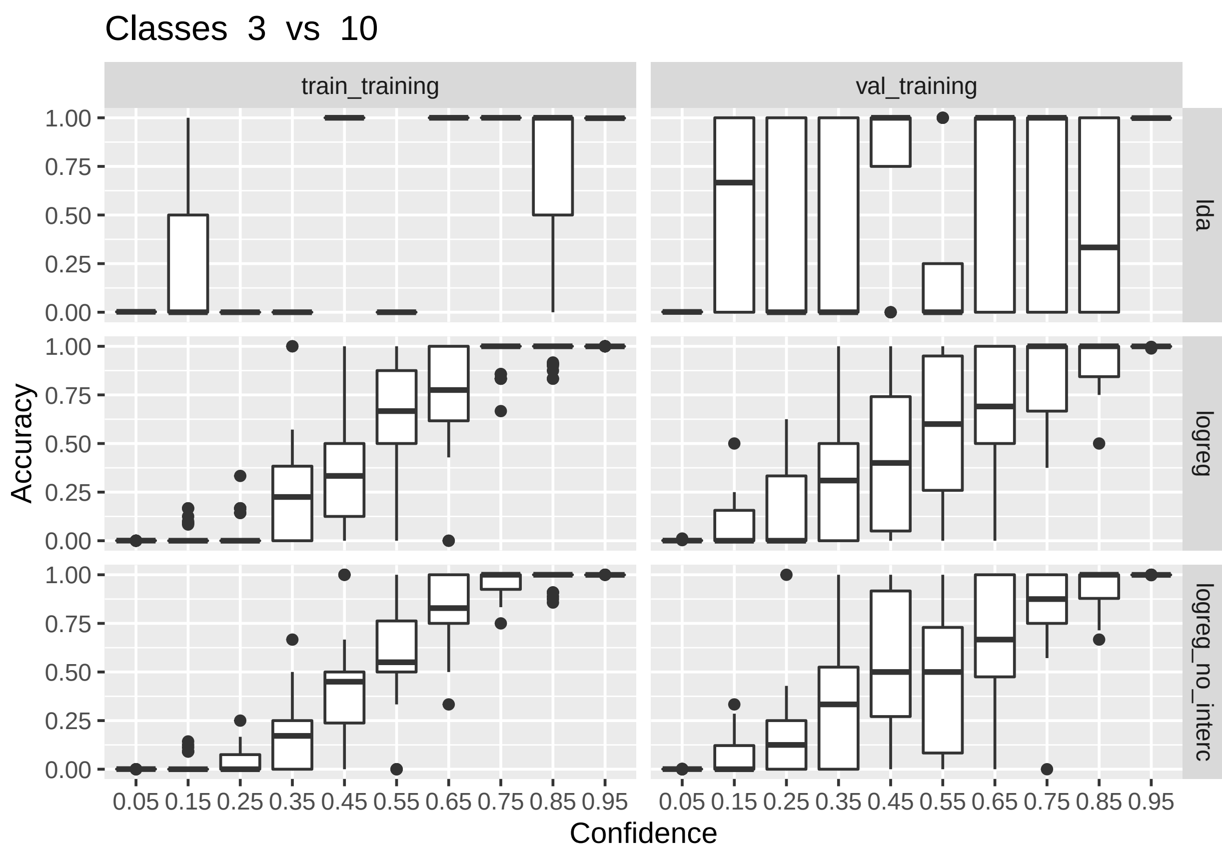

```
## Warning: Removed 258 rows containing non-finite values (stat_boxplot).
```



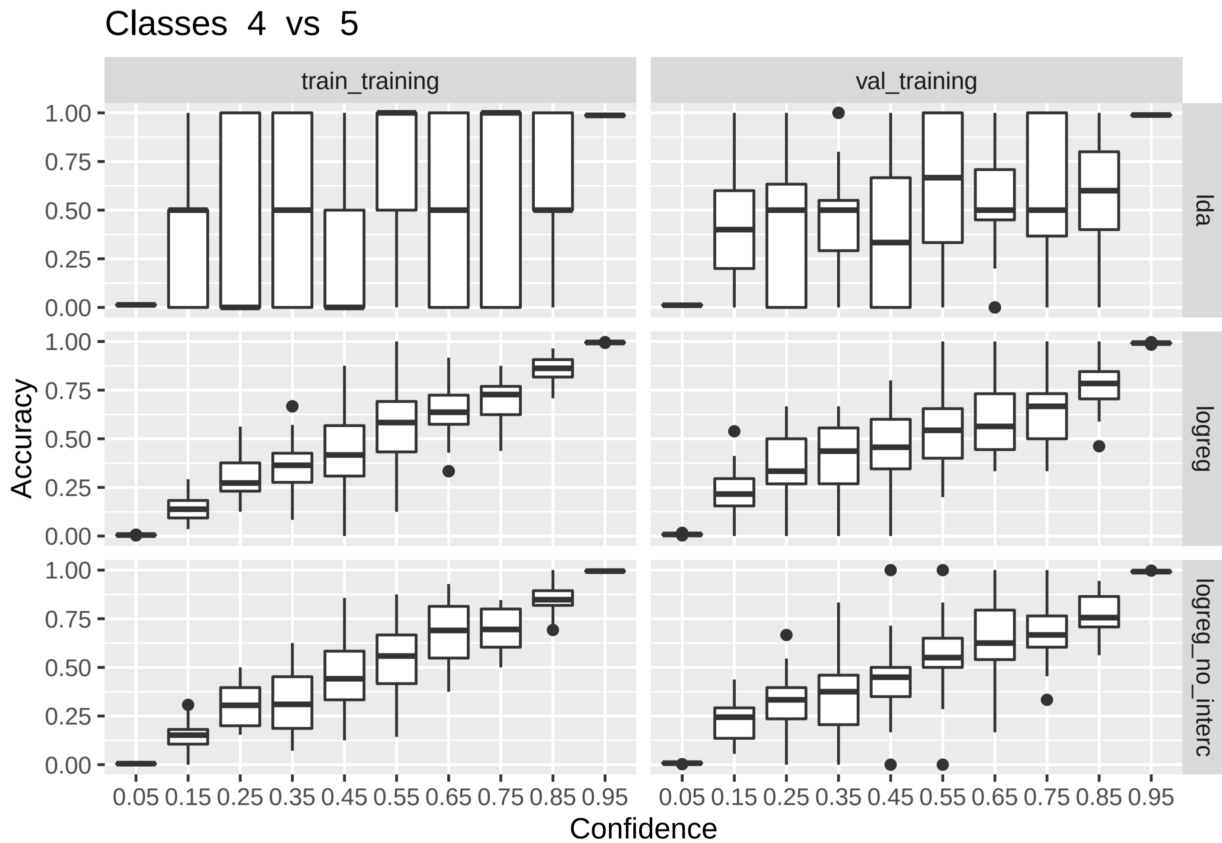

```
## Warning: Removed 332 rows containing non-finite values (stat_boxplot).
```



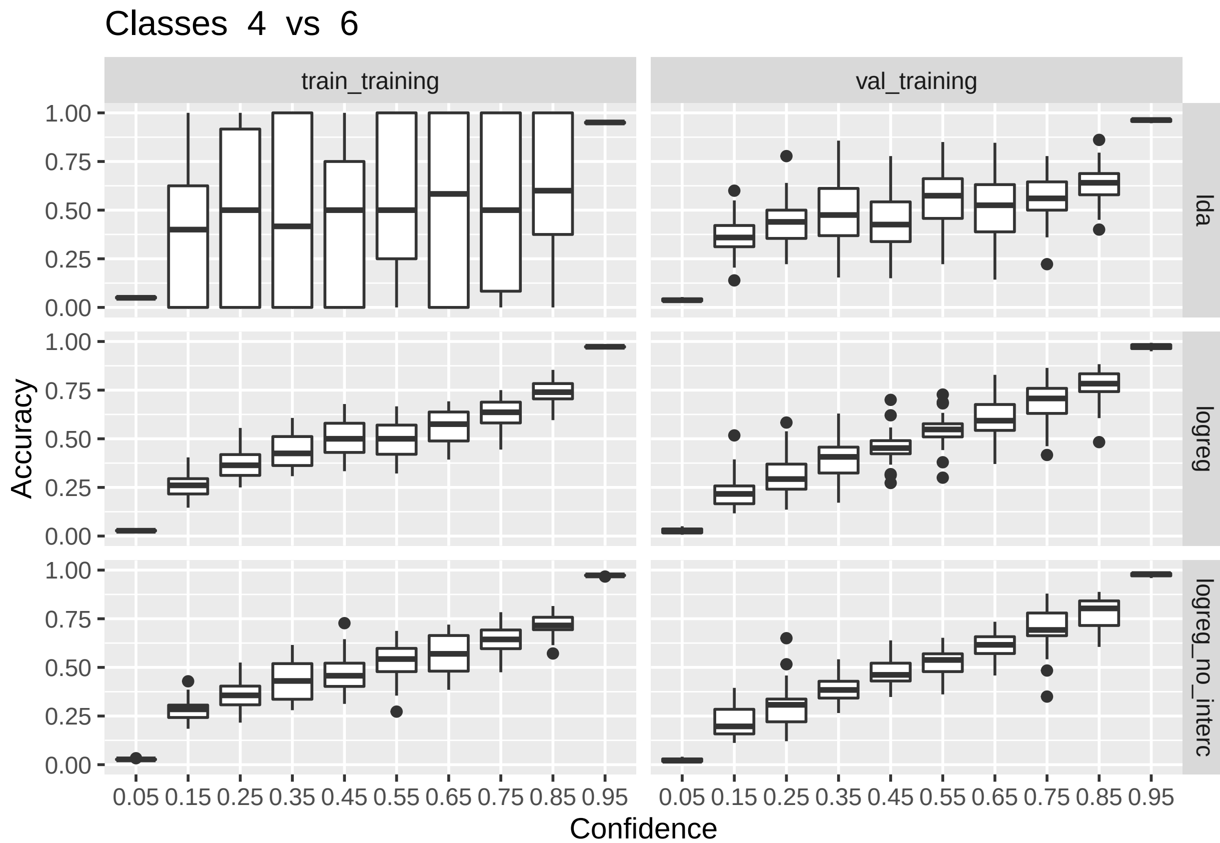

```
## Warning: Removed 404 rows containing non-finite values (stat_boxplot).
```



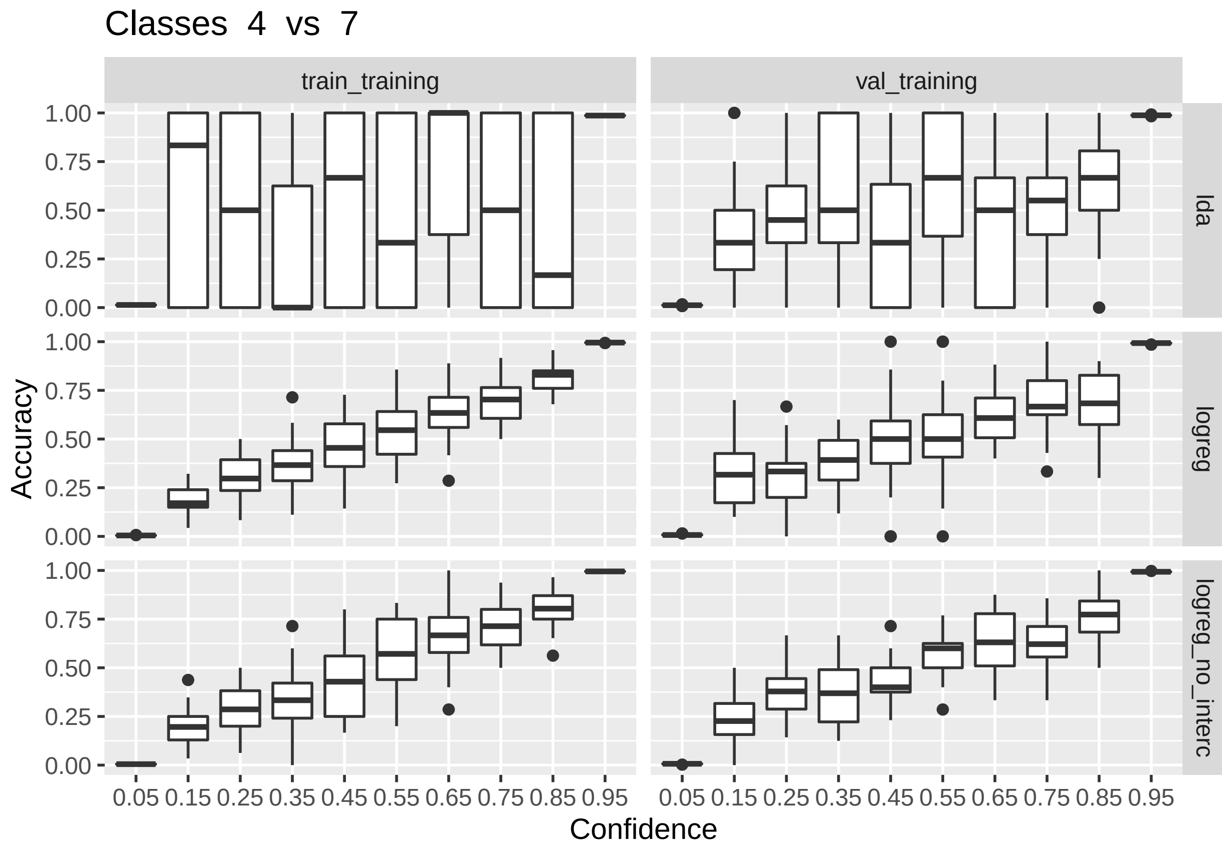

```
## Warning: Removed 172 rows containing non-finite values (stat_boxplot).
```



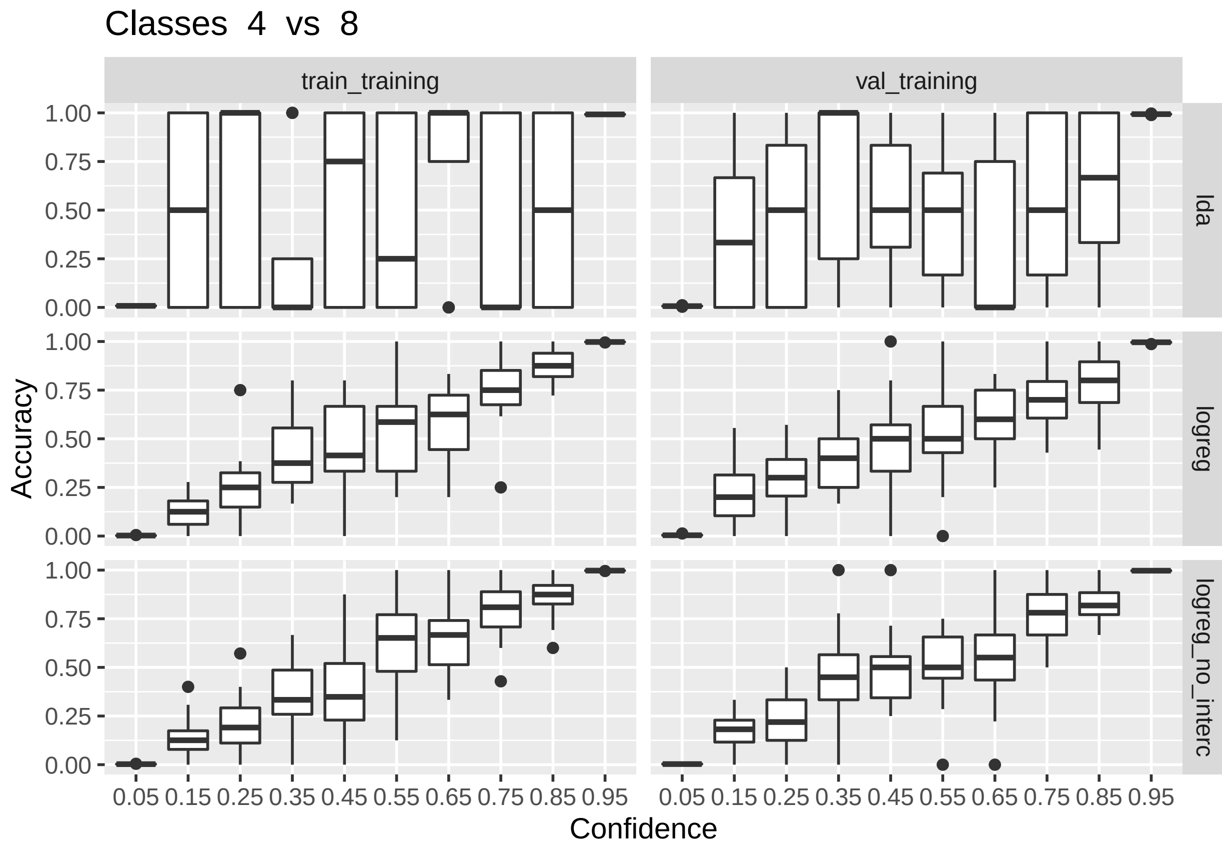

```
## Warning: Removed 42 rows containing non-finite values (stat_boxplot).
```



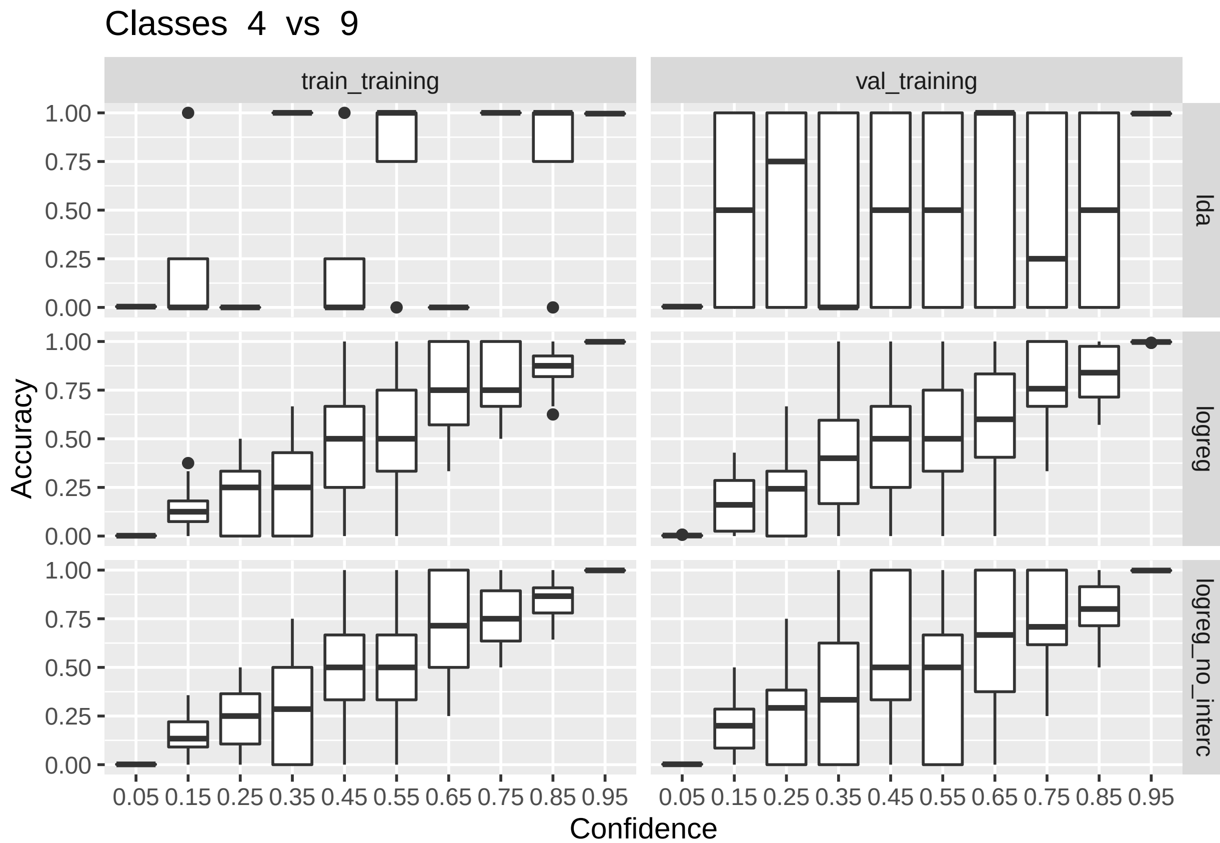

```
## Warning: Removed 152 rows containing non-finite values (stat_boxplot).
```



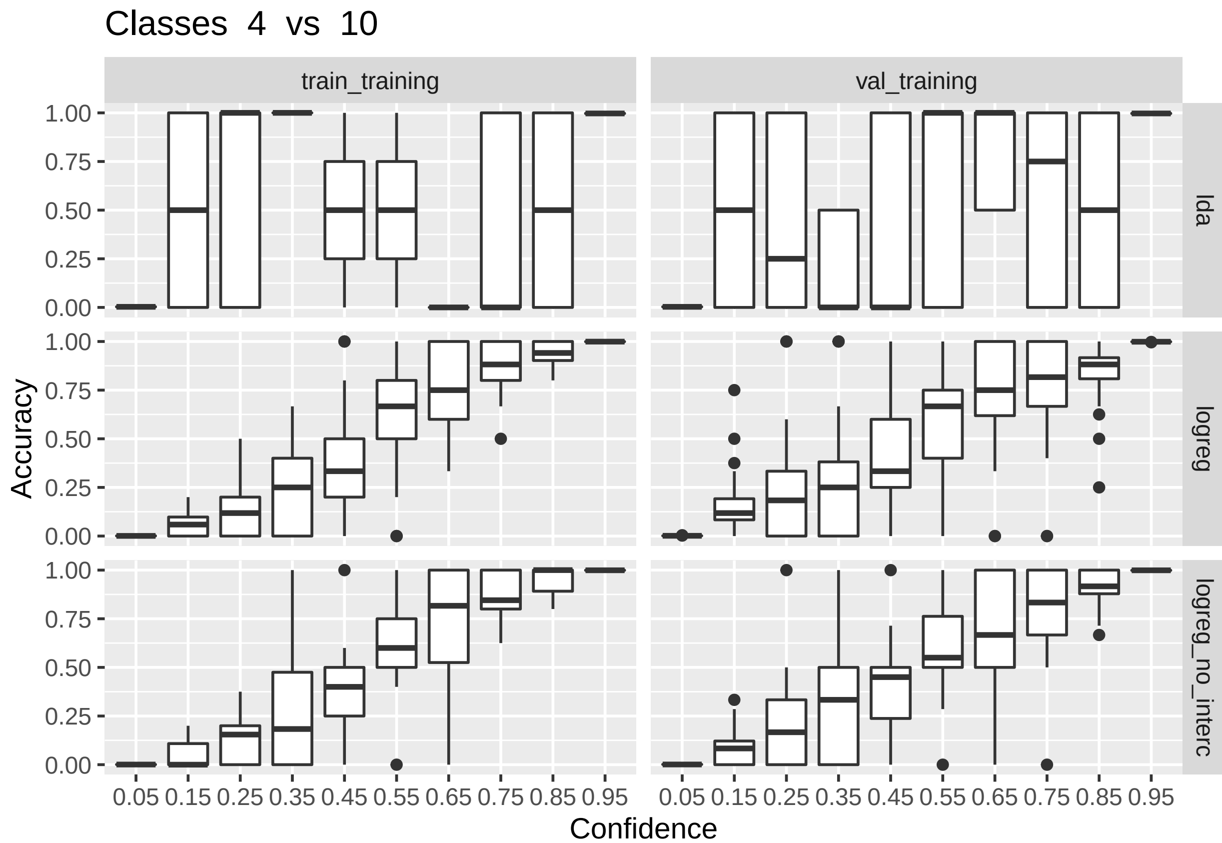

```
## Warning: Removed 220 rows containing non-finite values (stat_boxplot).
```



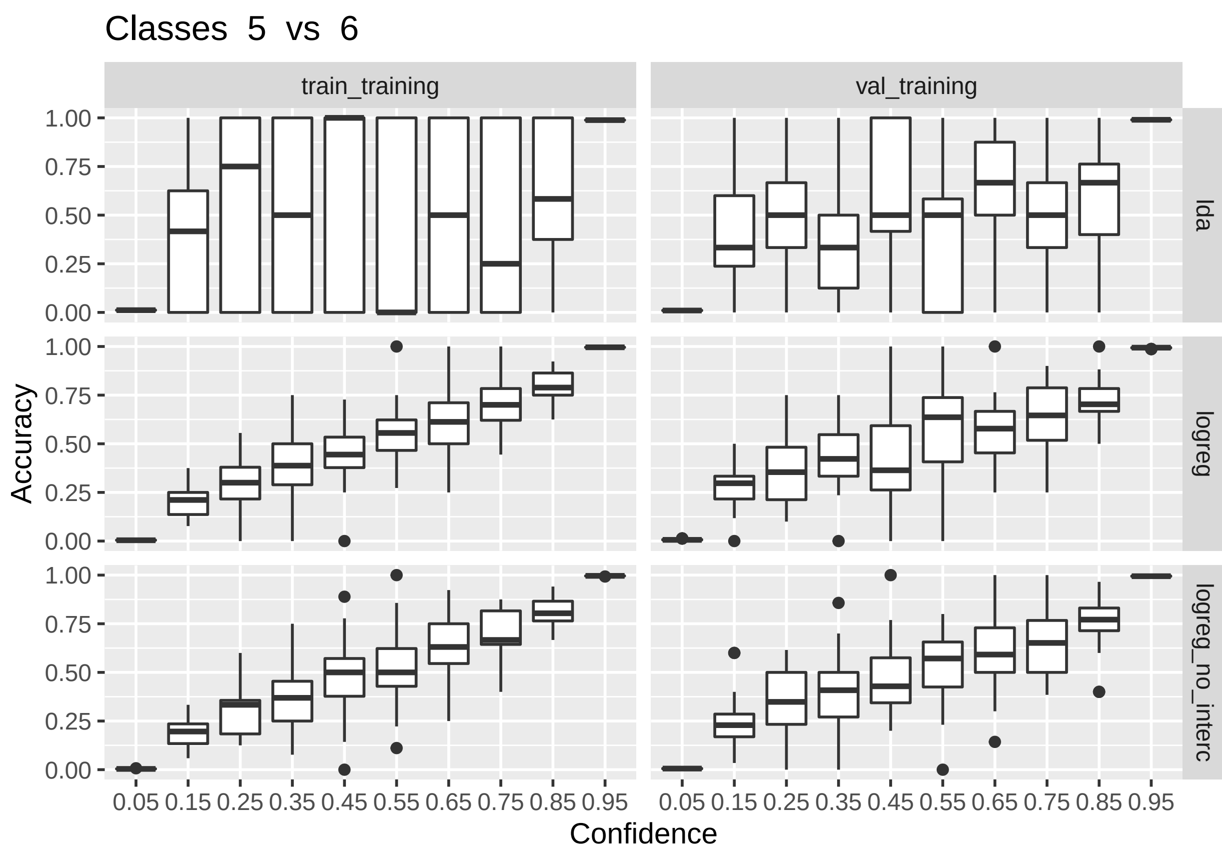

```
## Warning: Removed 386 rows containing non-finite values (stat_boxplot).
```



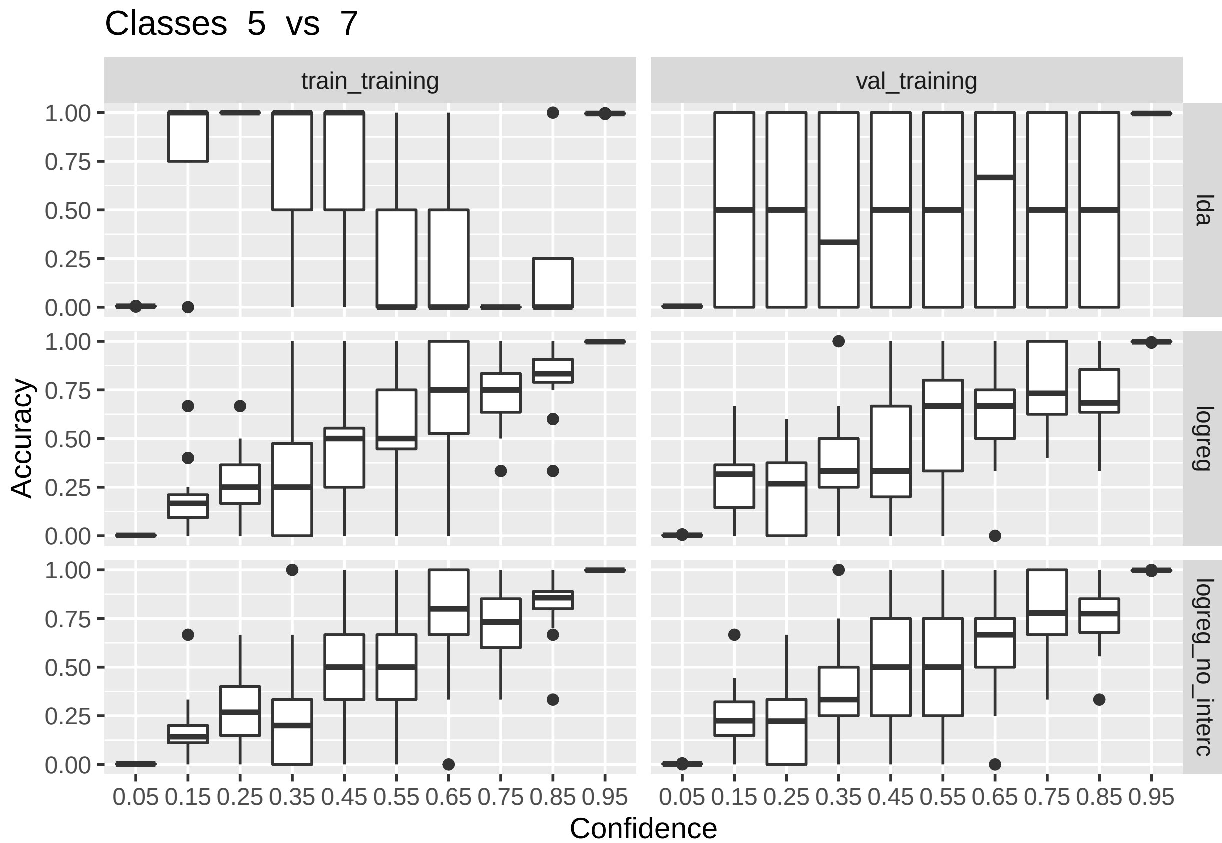

```
## Warning: Removed 380 rows containing non-finite values (stat_boxplot).
```



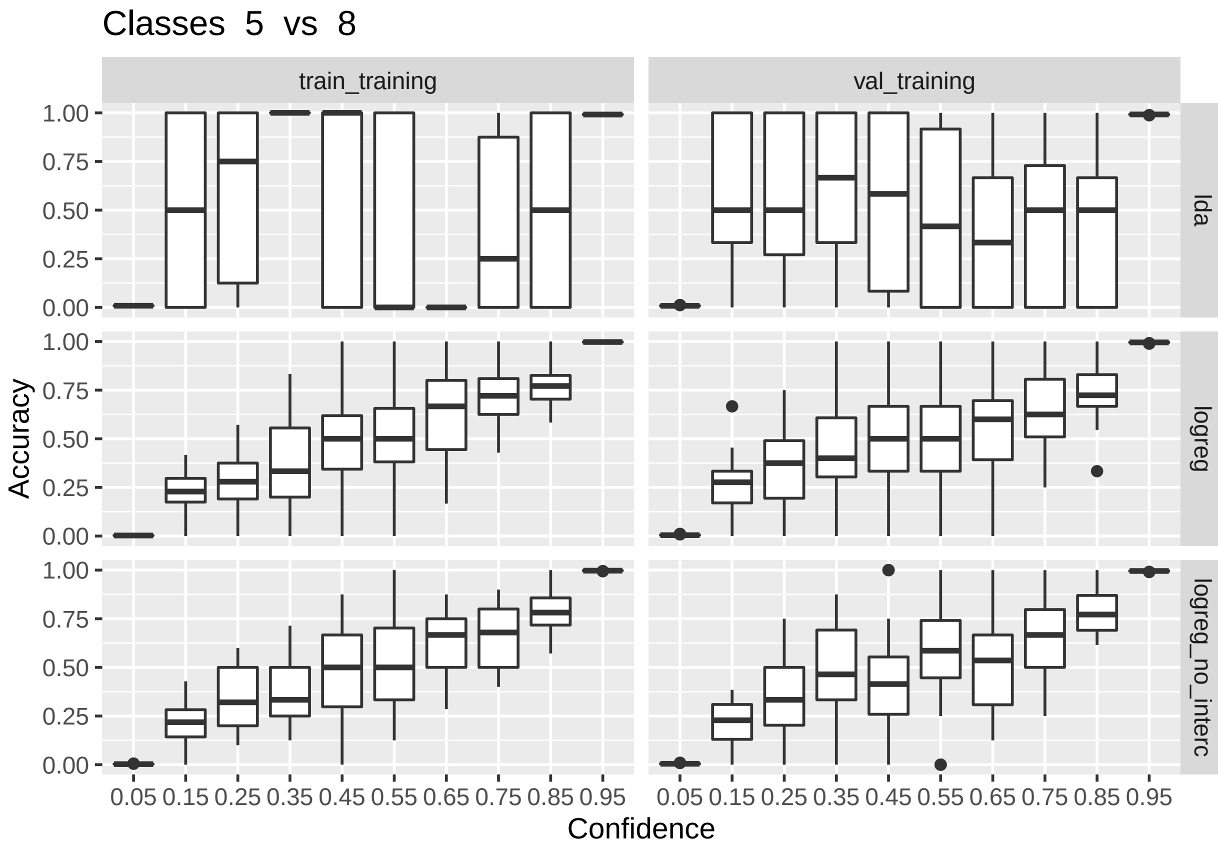

```
## Warning: Removed 174 rows containing non-finite values (stat_boxplot).
```



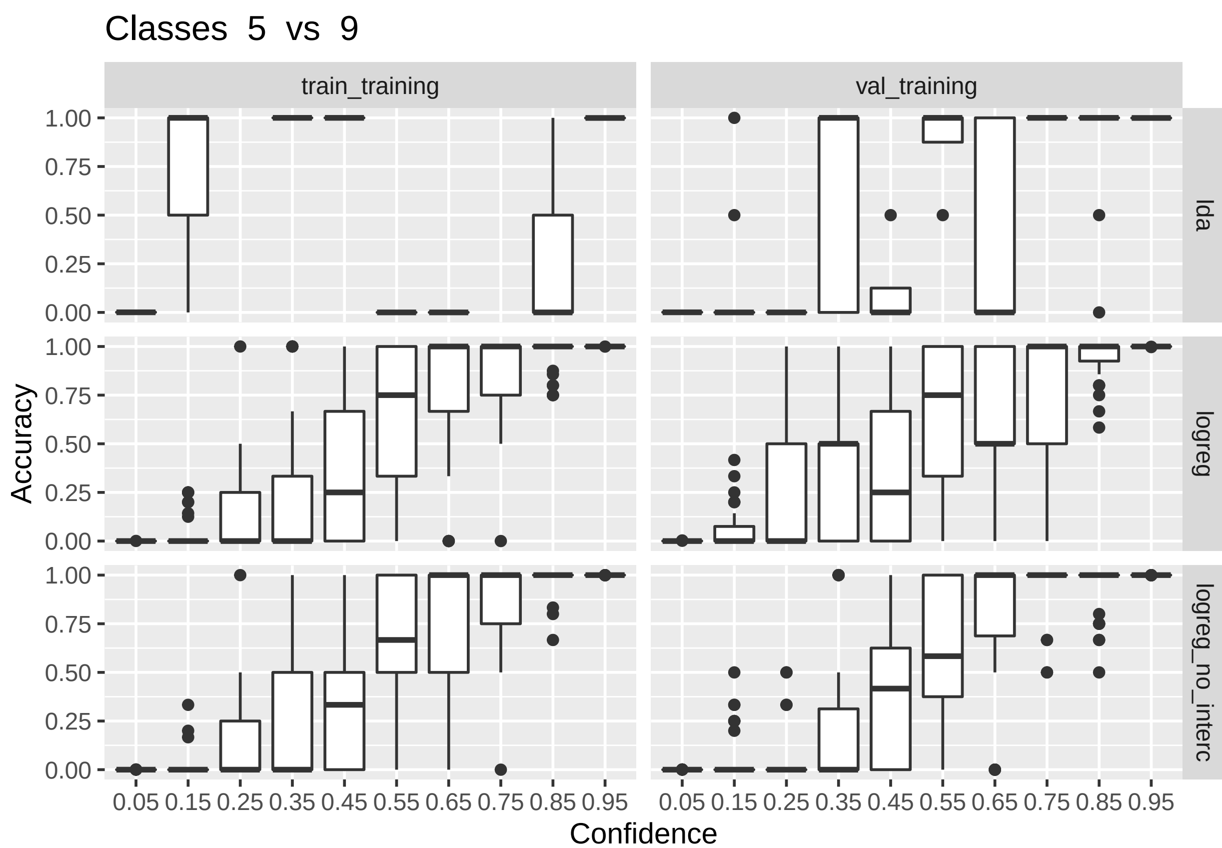

```
## Warning: Removed 360 rows containing non-finite values (stat_boxplot).
```



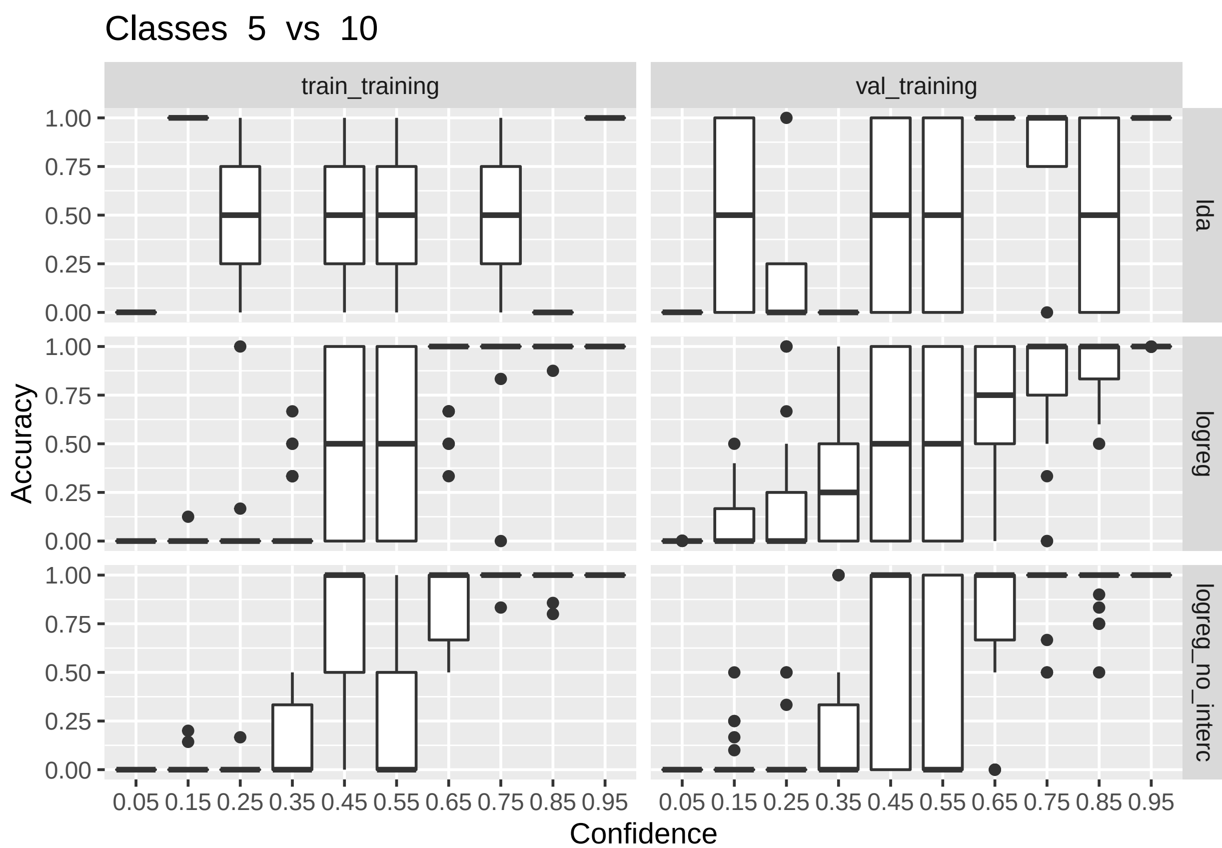

```
## Warning: Removed 236 rows containing non-finite values (stat_boxplot).
```



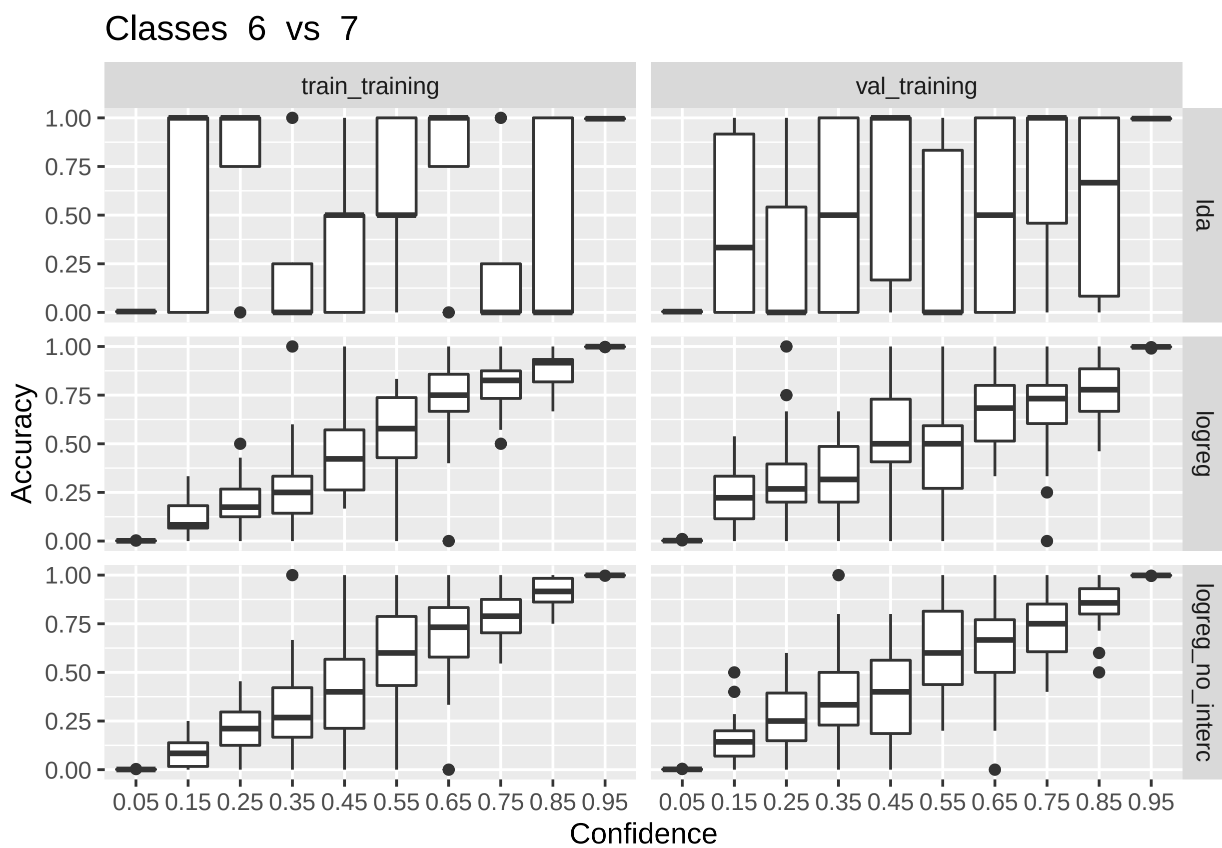

```
## Warning: Removed 544 rows containing non-finite values (stat_boxplot).
```



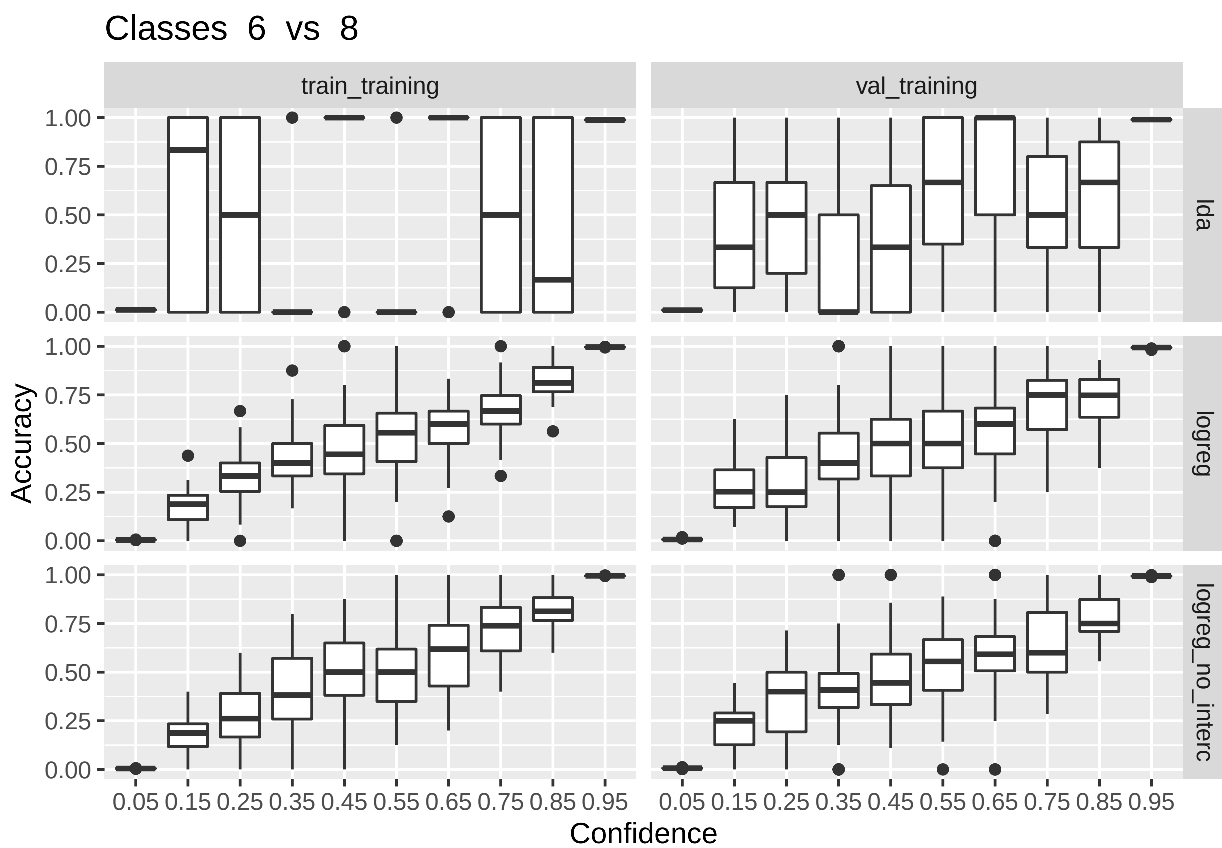

```
## Warning: Removed 516 rows containing non-finite values (stat_boxplot).
```



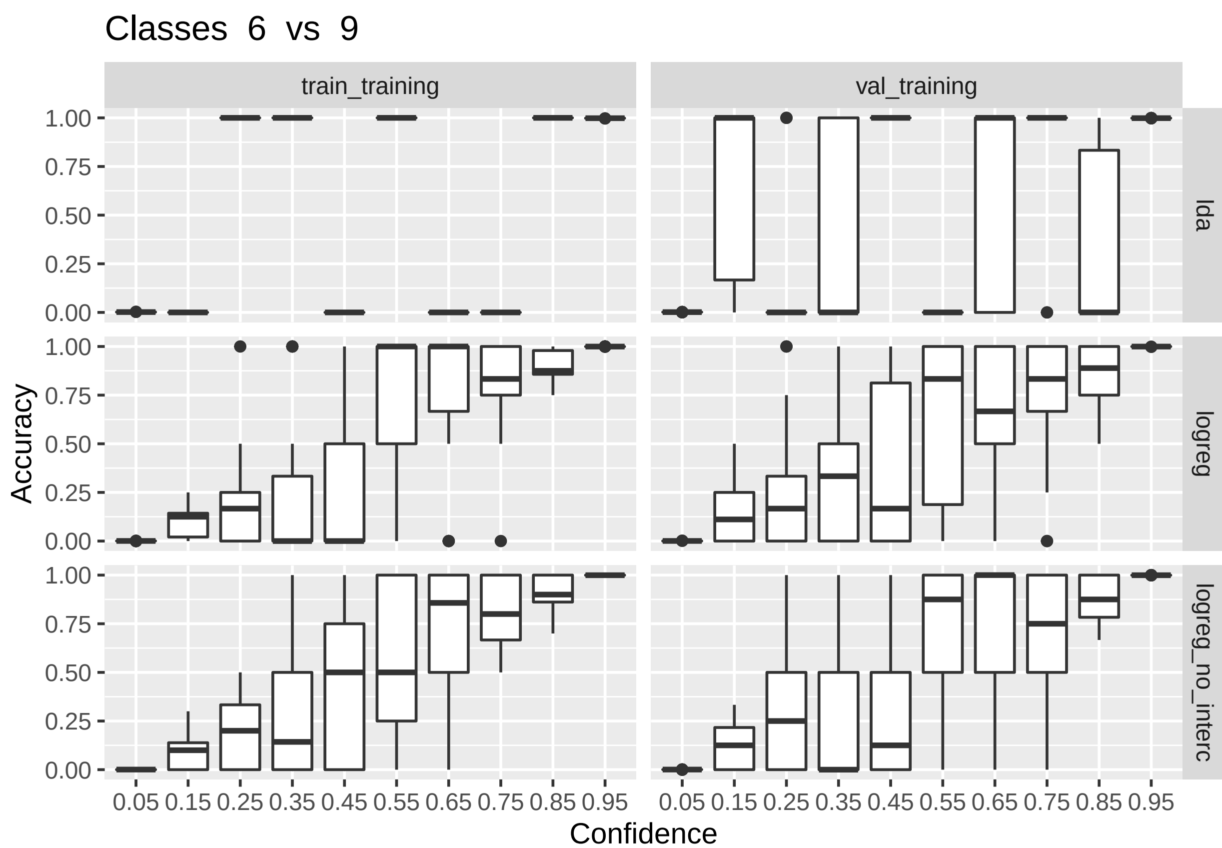

```
## Warning: Removed 284 rows containing non-finite values (stat_boxplot).
```



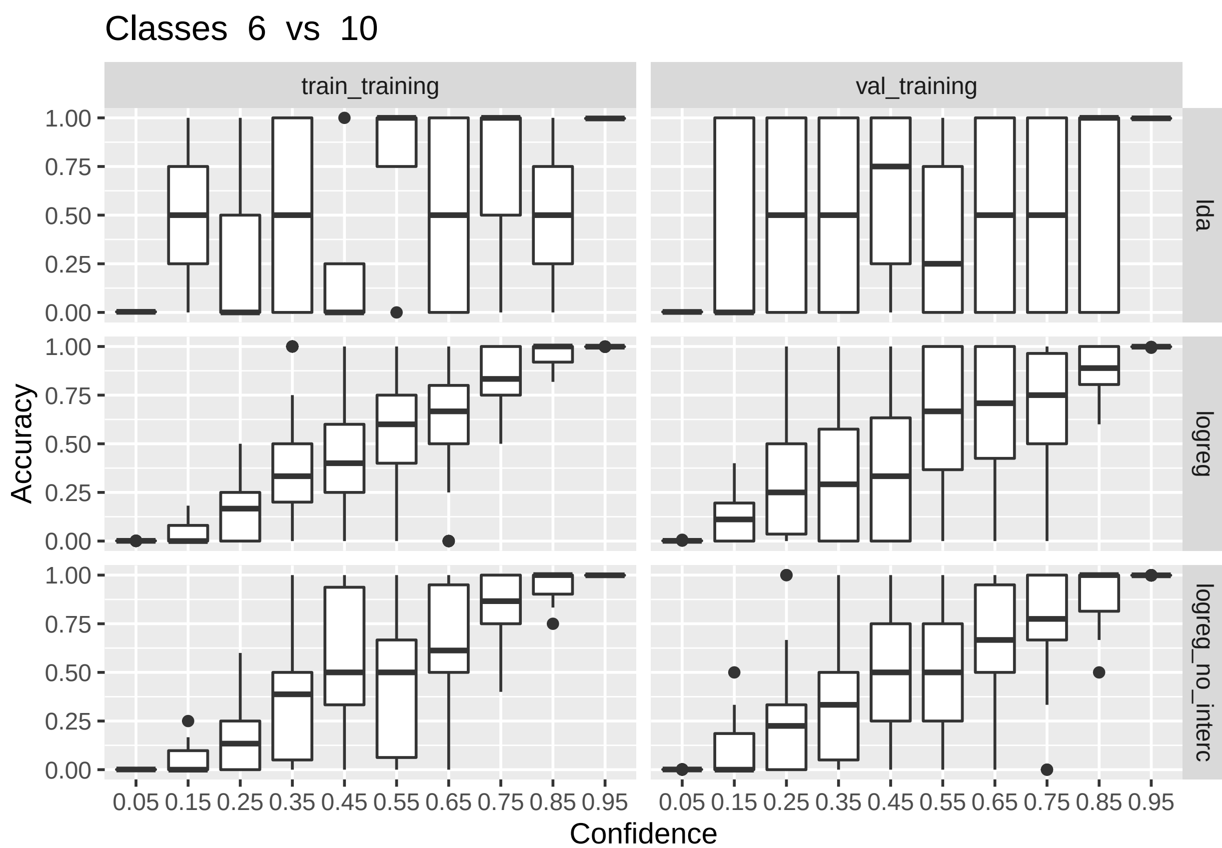

```
## Warning: Removed 190 rows containing non-finite values (stat_boxplot).
```



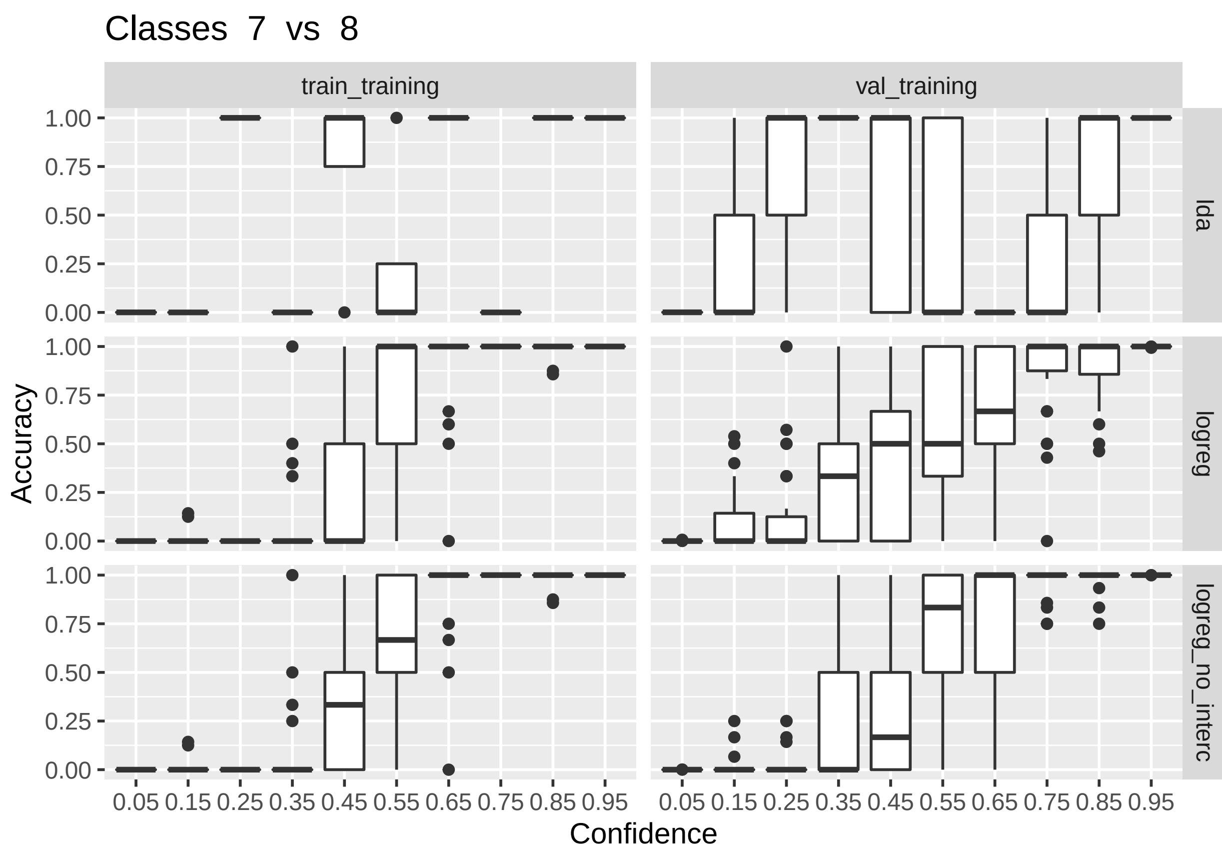

```
## Warning: Removed 498 rows containing non-finite values (stat_boxplot).
```



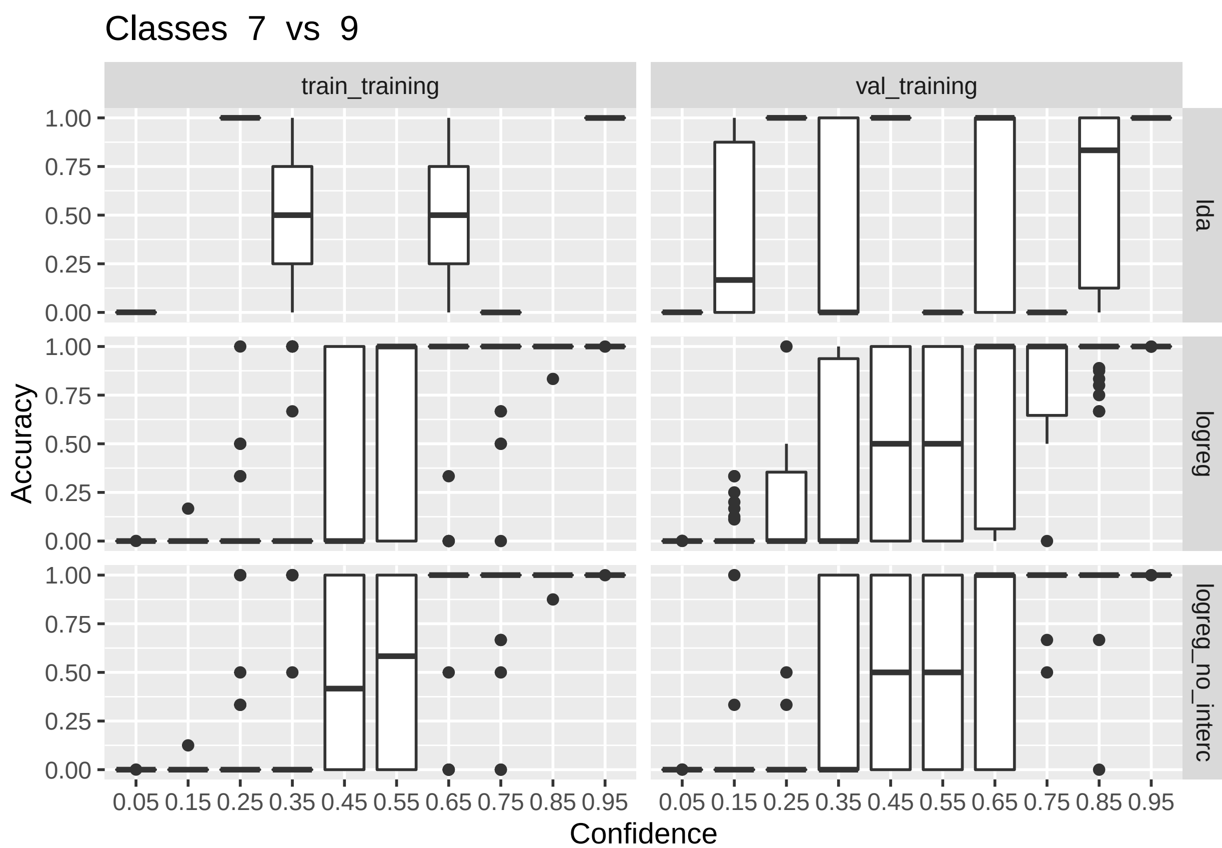

```
## Warning: Removed 362 rows containing non-finite values (stat_boxplot).
```




```
## Warning: Removed 502 rows containing non-finite values (stat_boxplot).
```

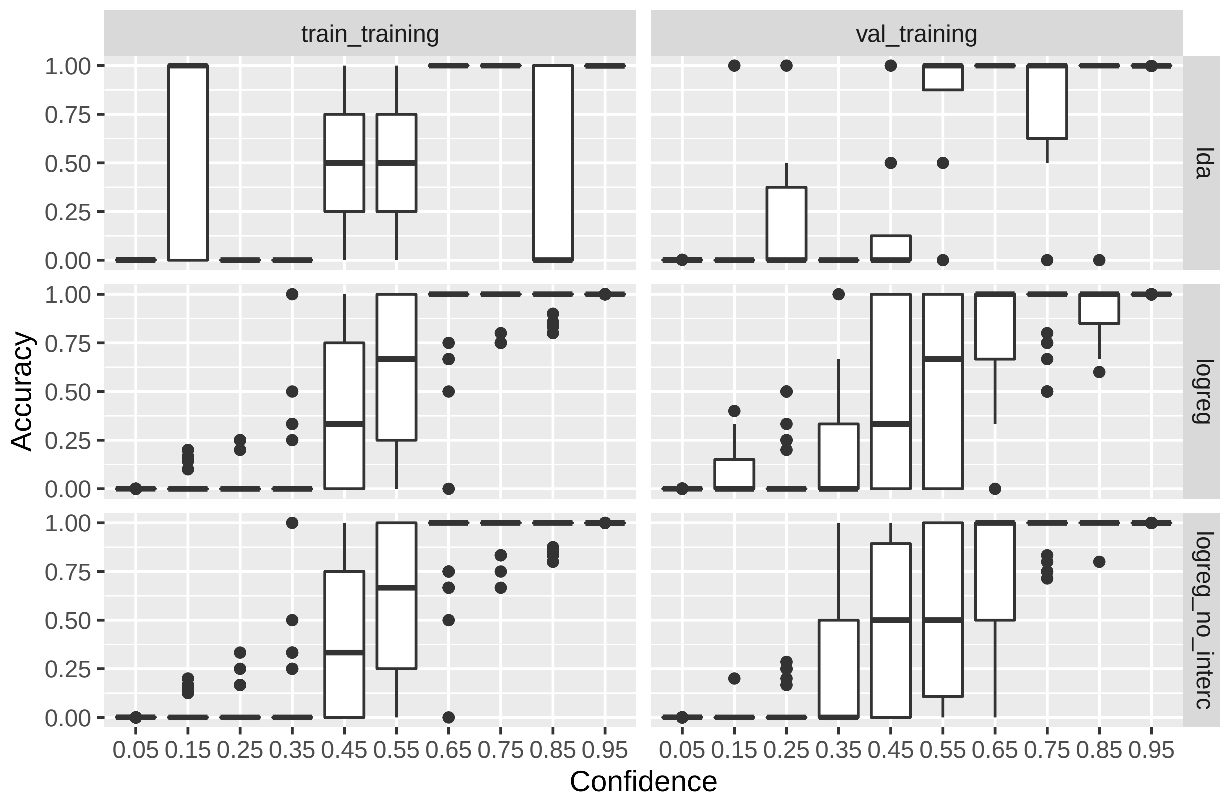



```
## Warning: Removed 610 rows containing non-finite values (stat_boxplot).
```

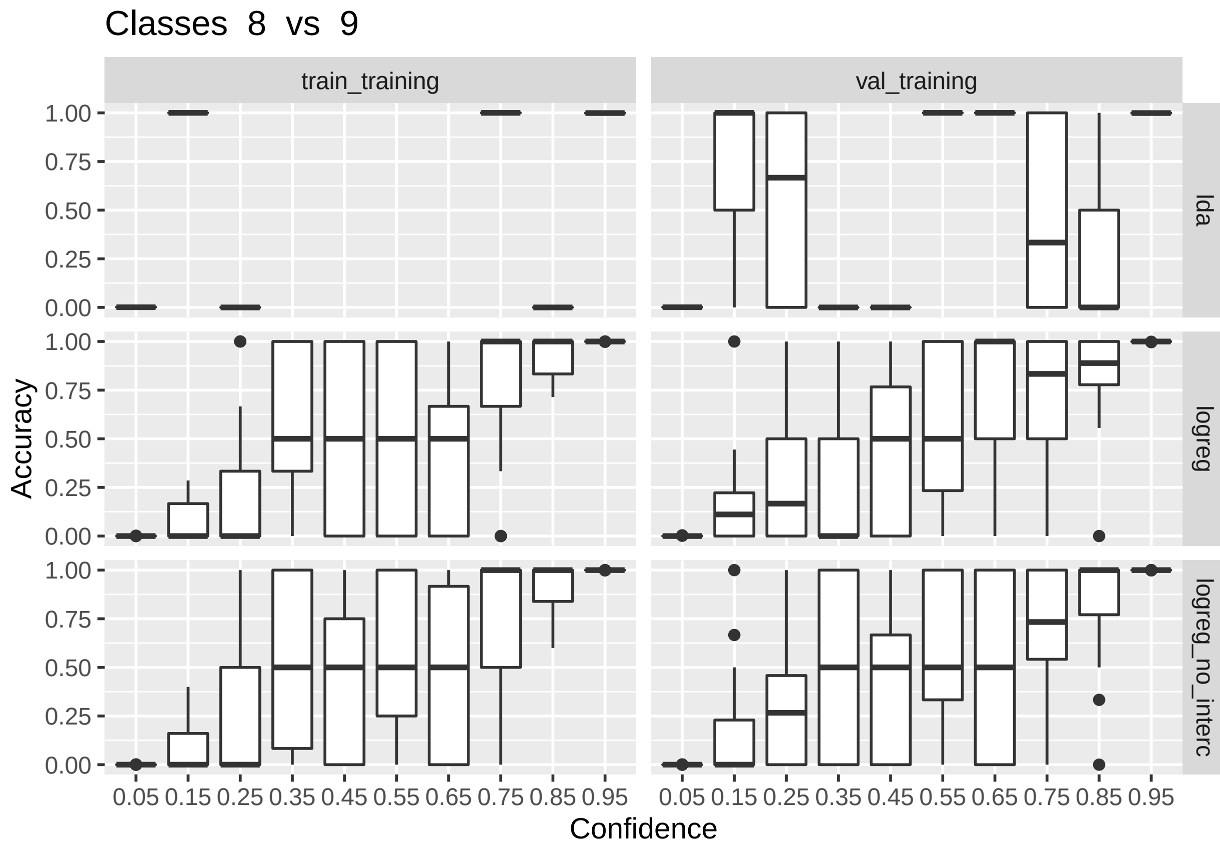



```
## Warning: Removed 458 rows containing non-finite values (stat_boxplot).
```

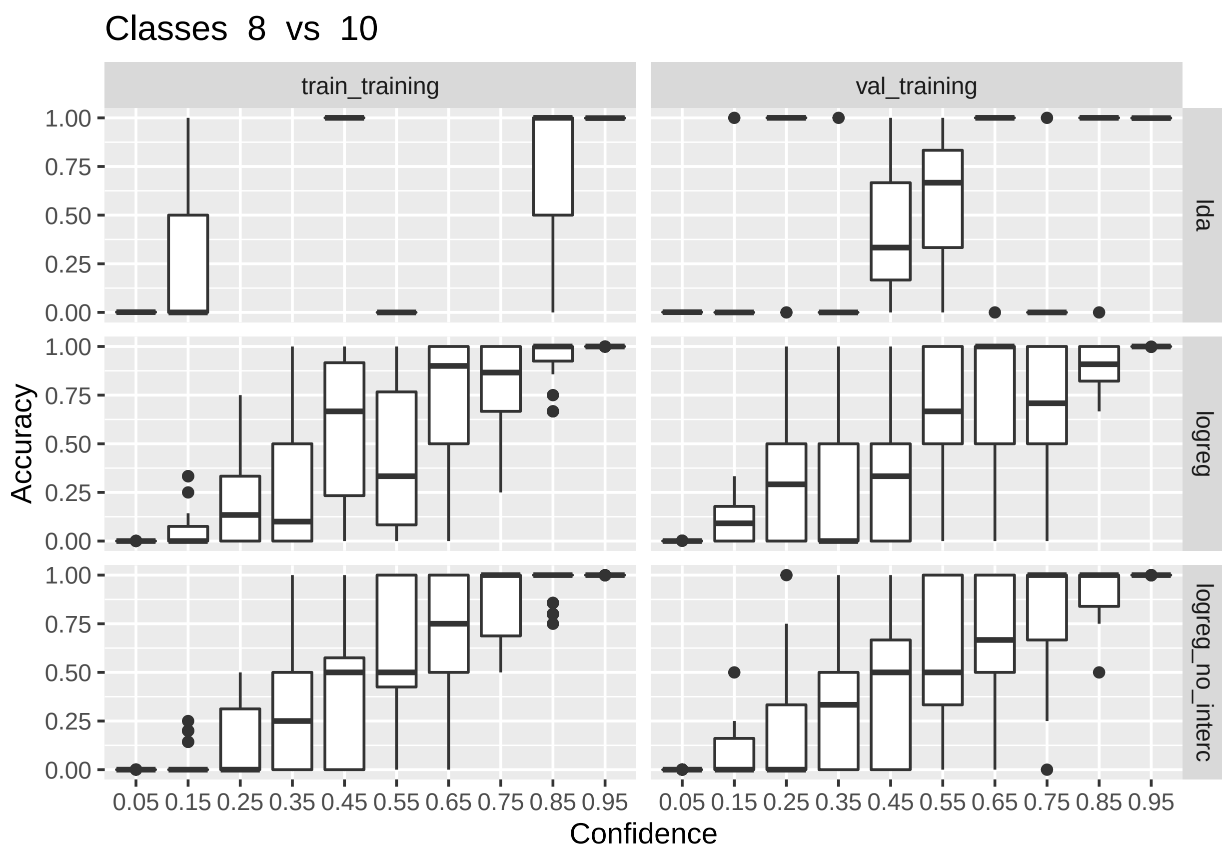
Classes 7 vs 10



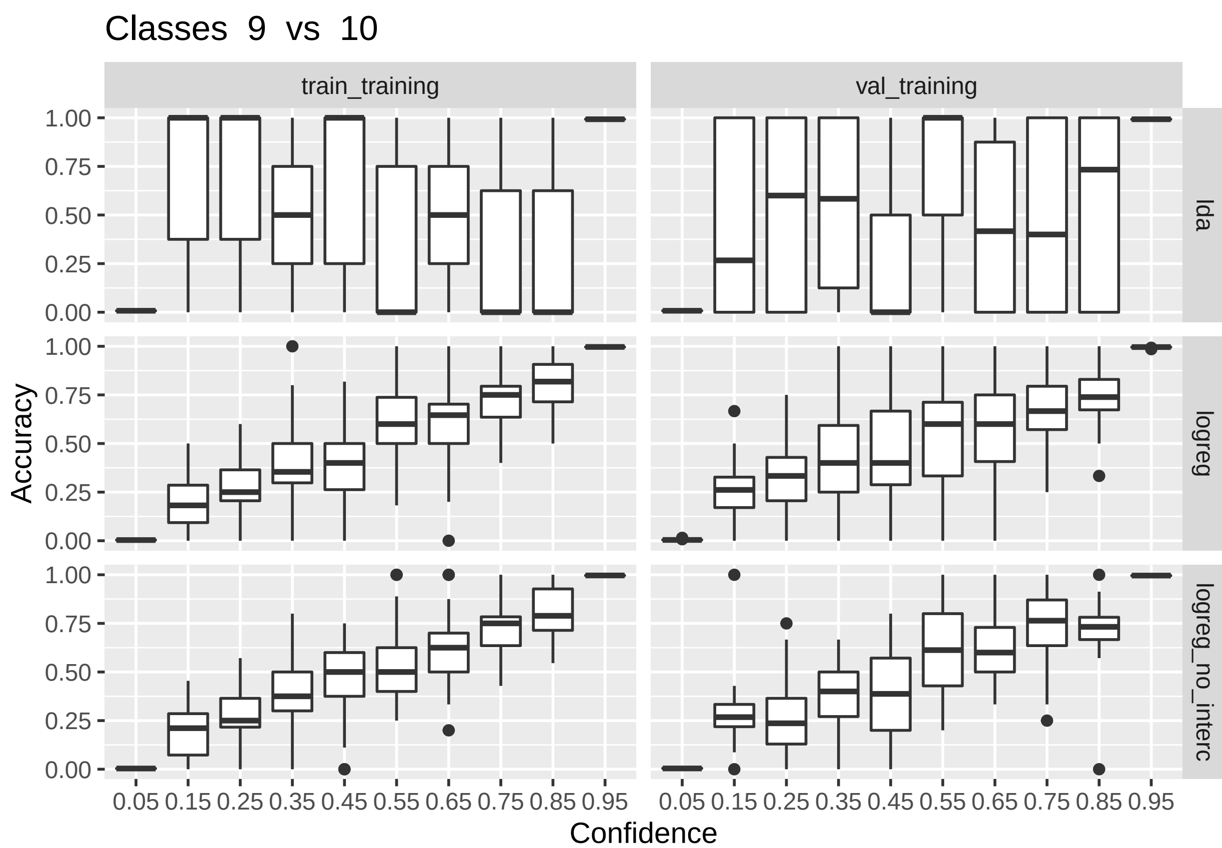

```
## Warning: Removed 562 rows containing non-finite values (stat_boxplot).
```




```
## Warning: Removed 482 rows containing non-finite values (stat_boxplot).
```

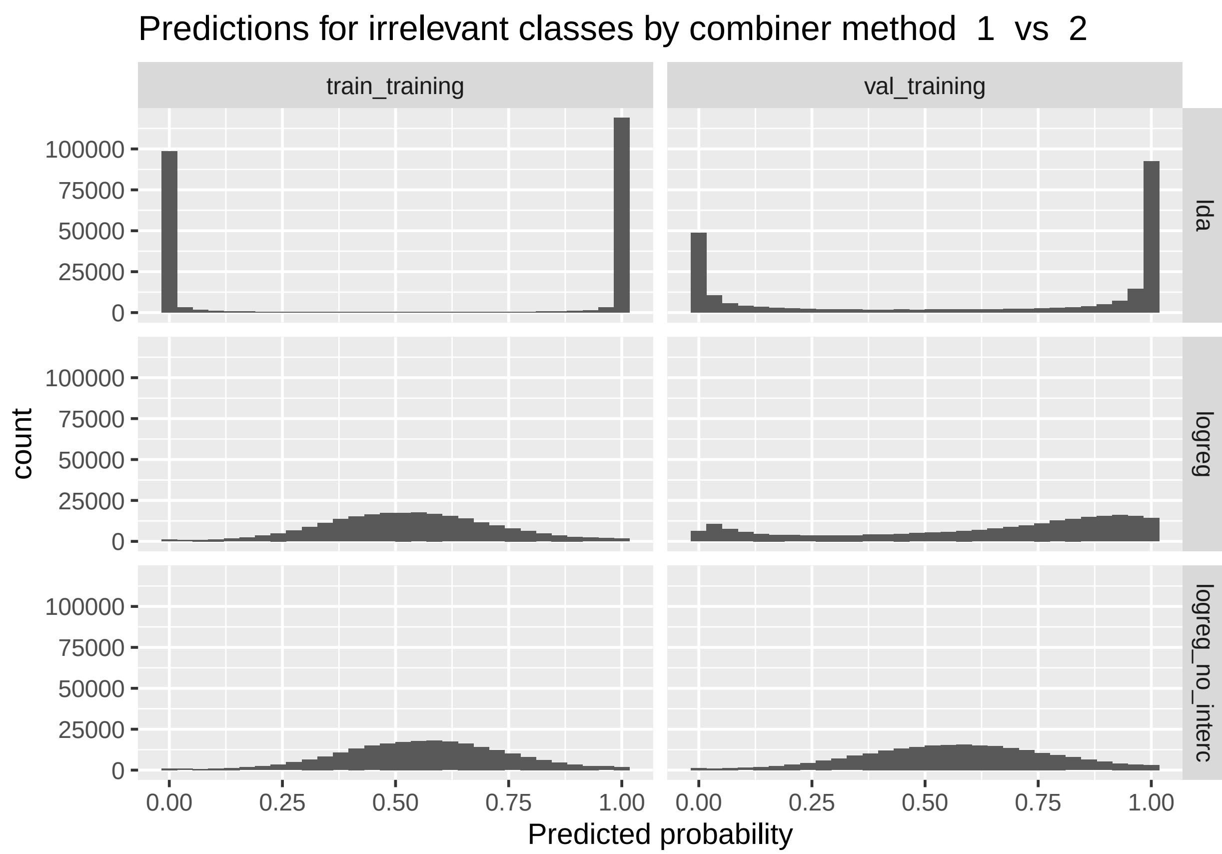


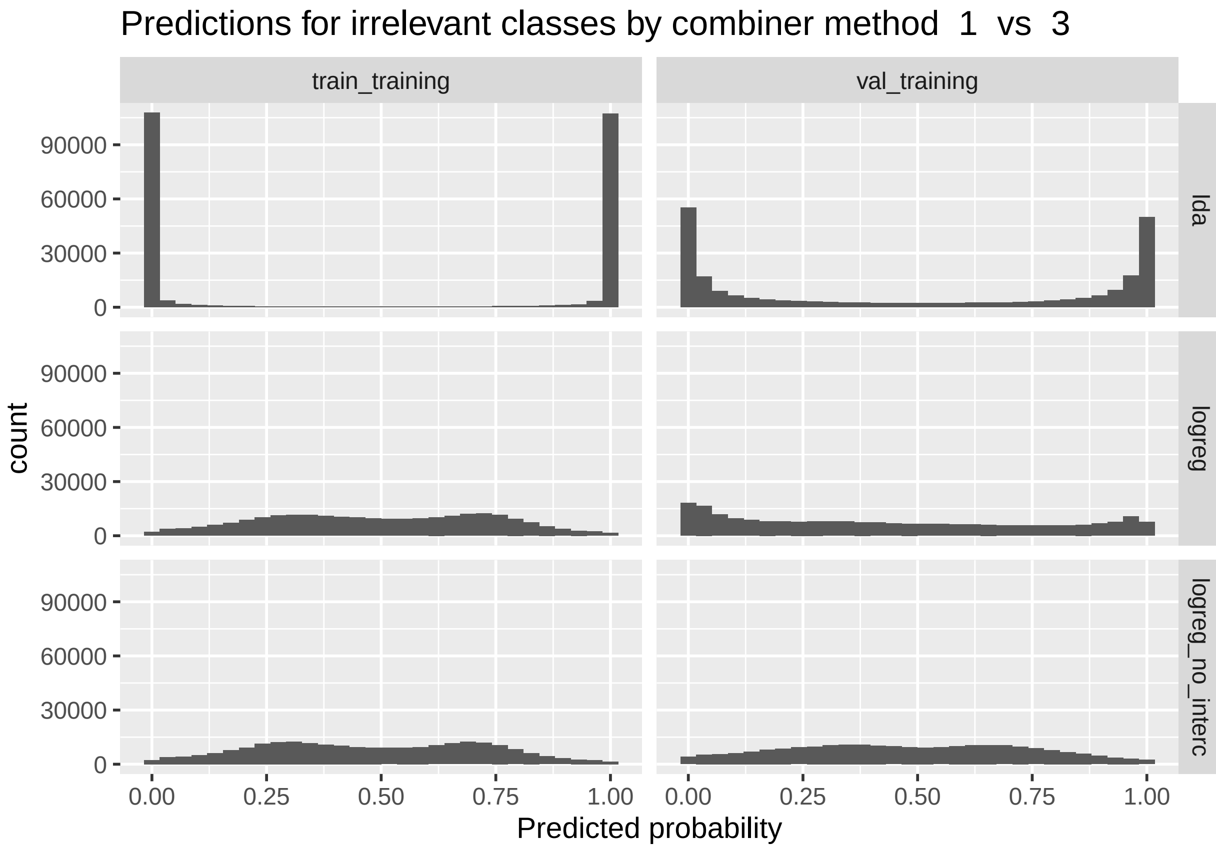

```
## Warning: Removed 308 rows containing non-finite values (stat_boxplot).
```

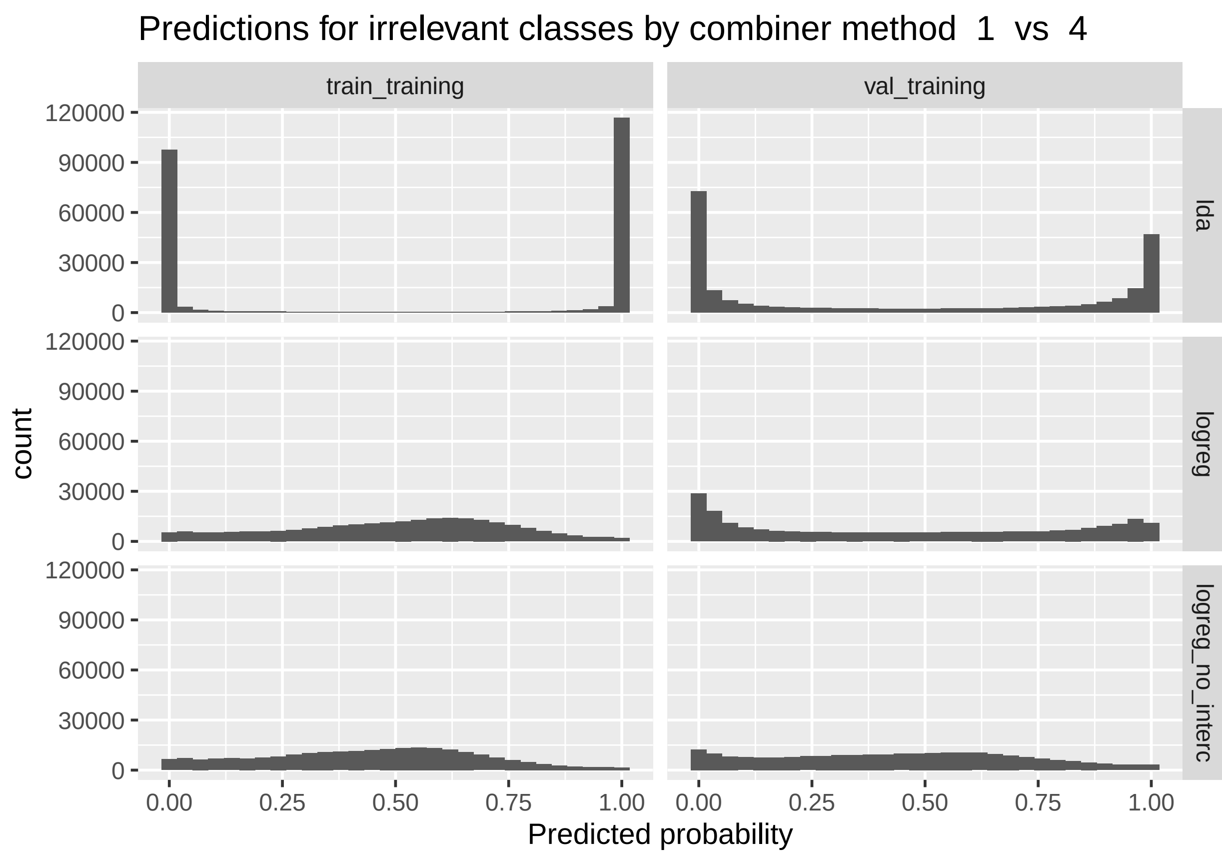


Even with low bin counts, better calibration can be observed for logreg than for lda.

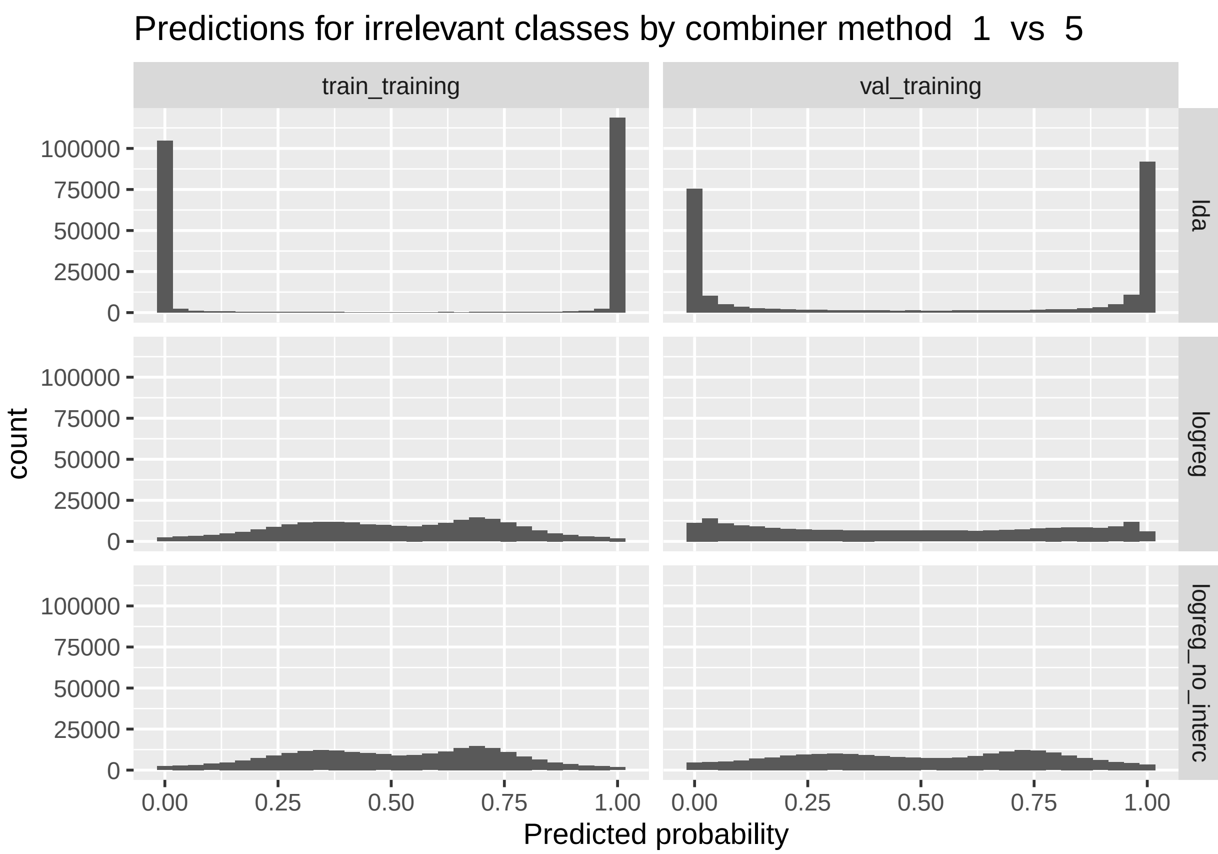
```
for (cl1 in 1:(classes - 1))
{
  for (cl2 in (cl1 + 1):classes)
  {
    irrel_plt <- ens_pw_irrel %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot() +
      geom_histogram(mapping=aes(x=pred1), bins=30) + facet_grid(cols=vars(train_set), rows=vars(combin
      ggtitle(paste("Predictions for irrelevant classes by combiner method ", cl1, " vs ", cl2)) +
      xlab("Predicted probability")

    print(irrel_plt)
  }
}
```


Predictions for irrelevant classes by combiner method 1 vs 2
train_training
val_training
lda
logreg
logreg_no_interc
count
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

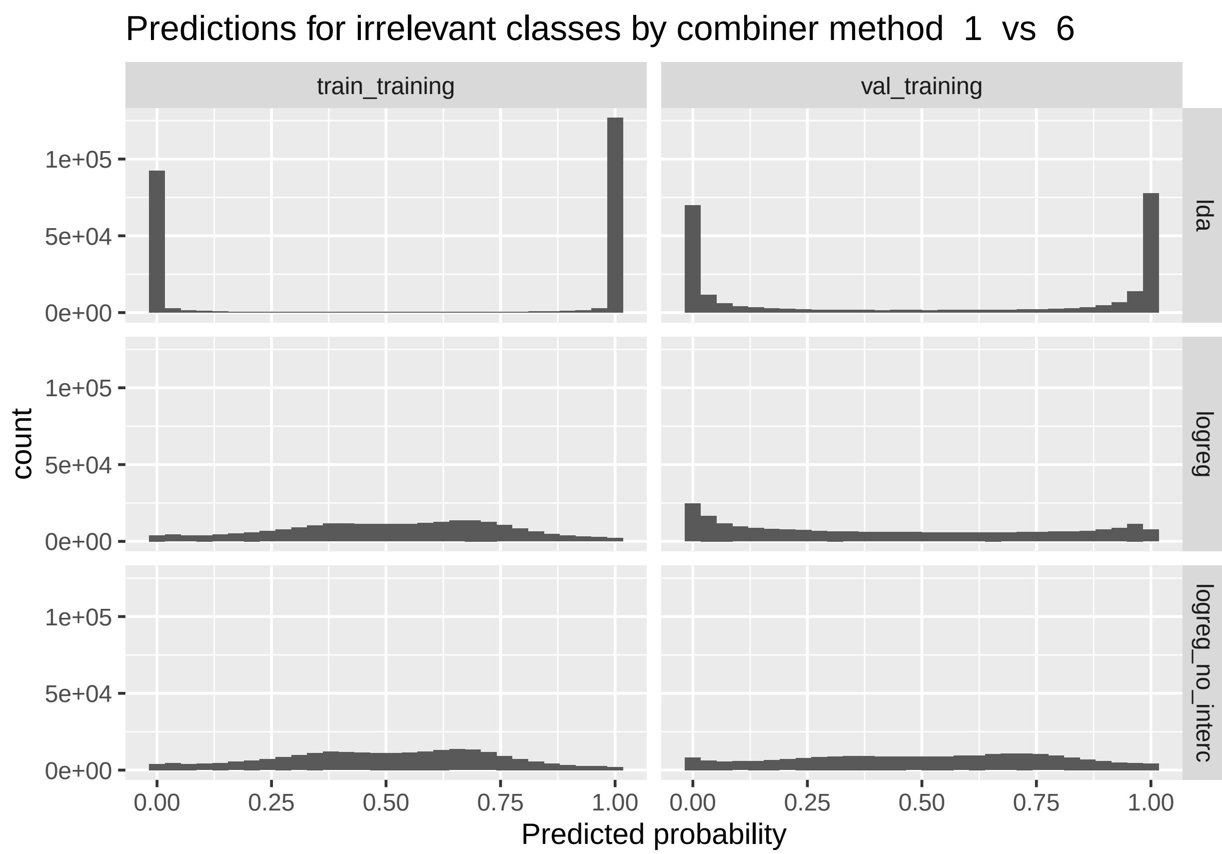

Predictions for irrelevant classes by combiner method 1 vs 3
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

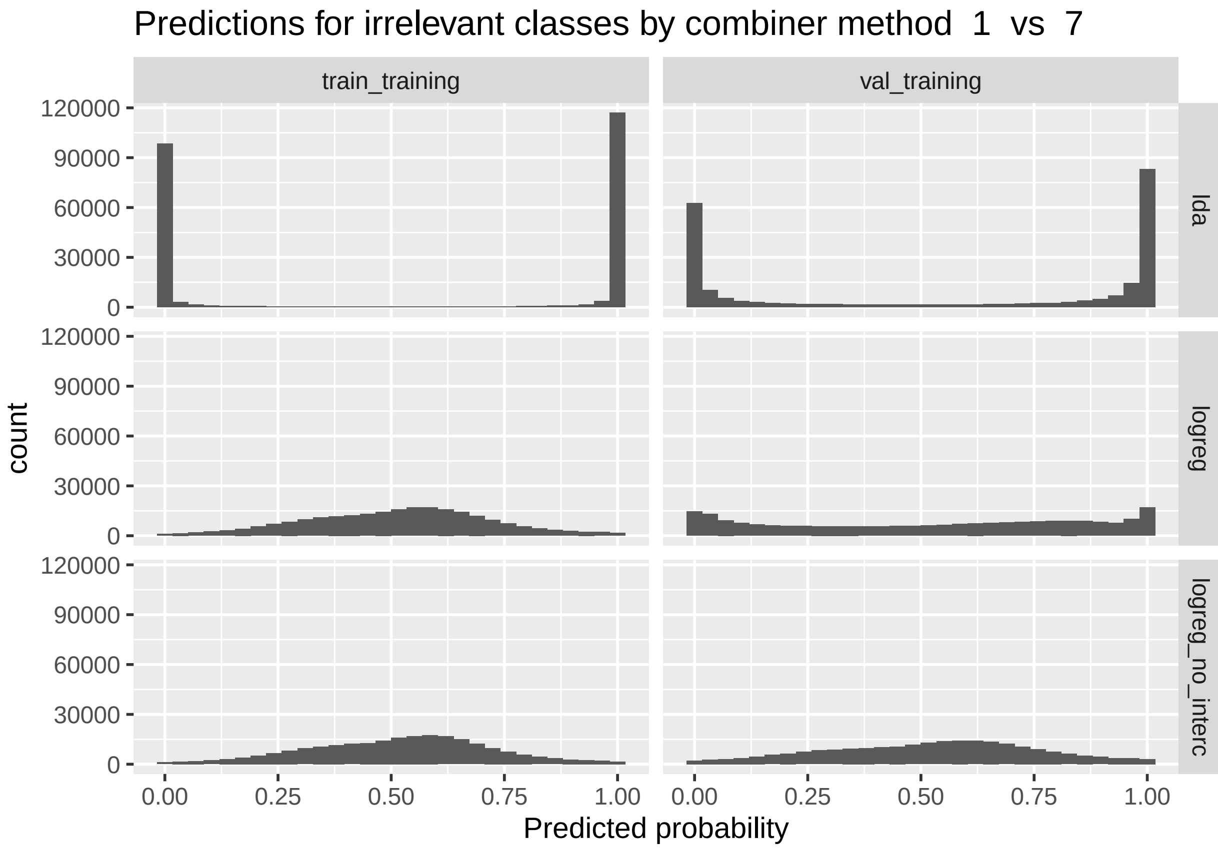


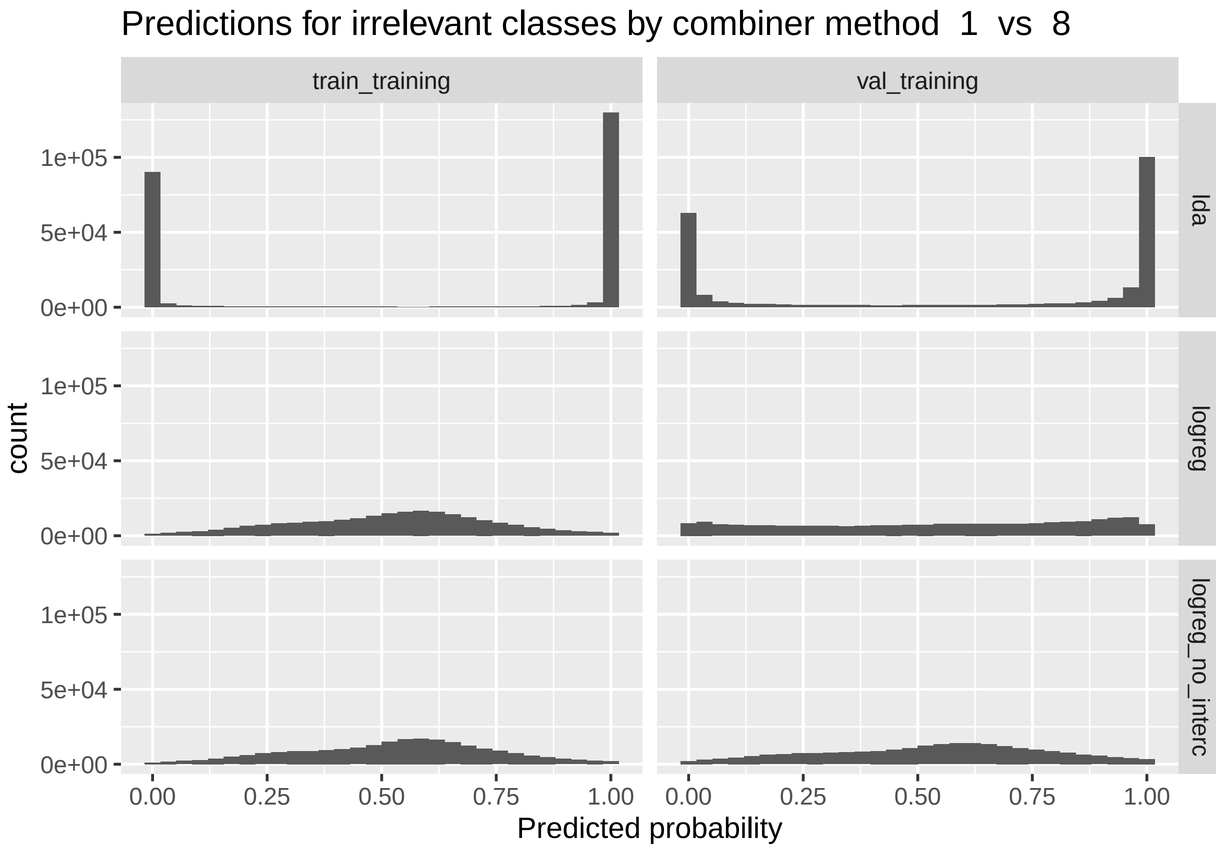
Predictions for irrelevant classes by combiner method 1 vs 4
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

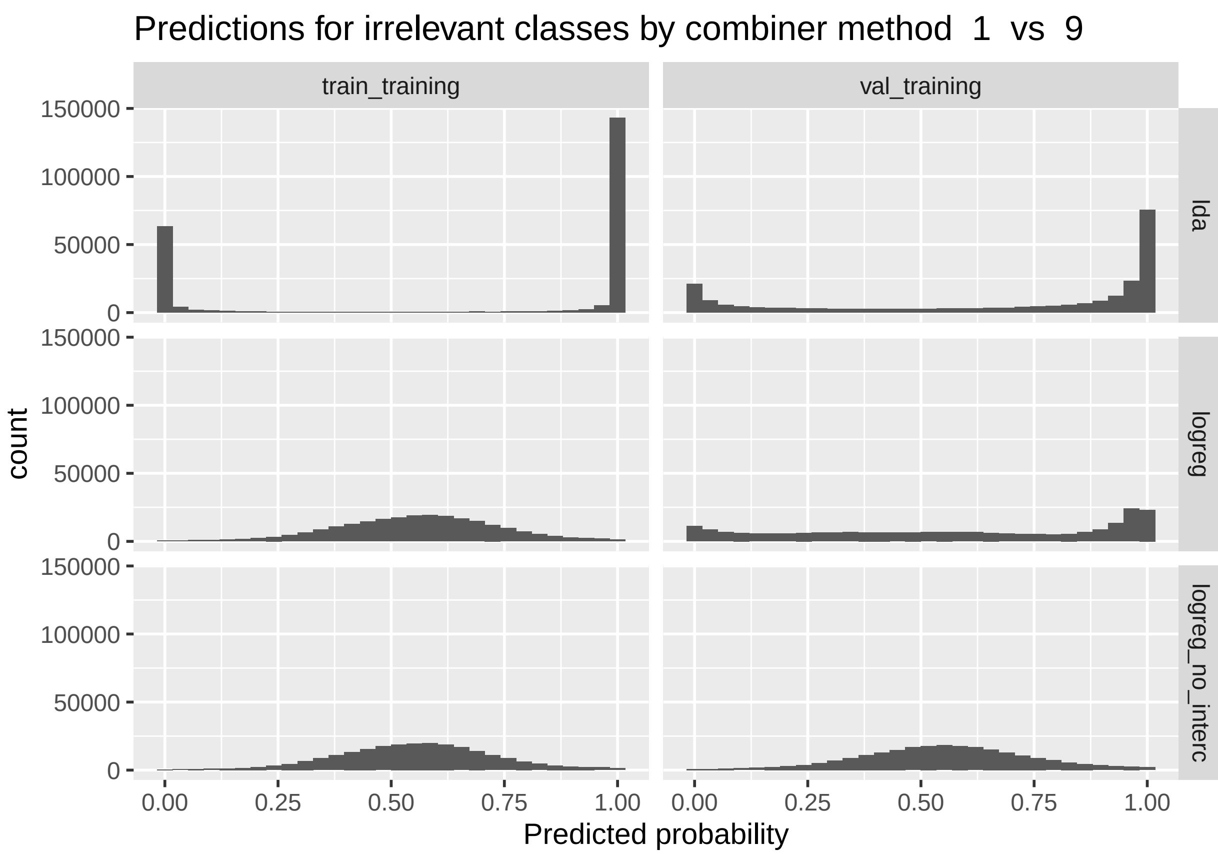


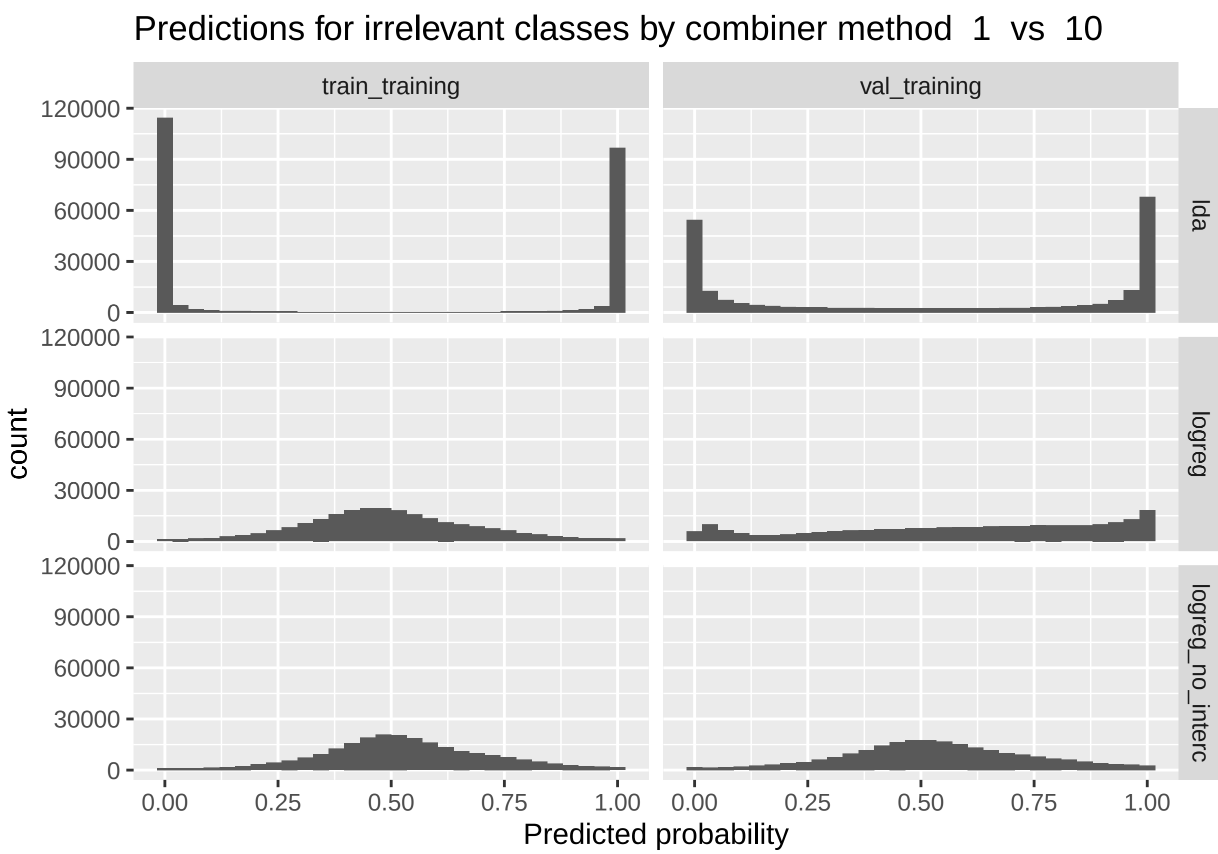
Predictions for irrelevant classes by combiner method 1 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
count
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 1 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

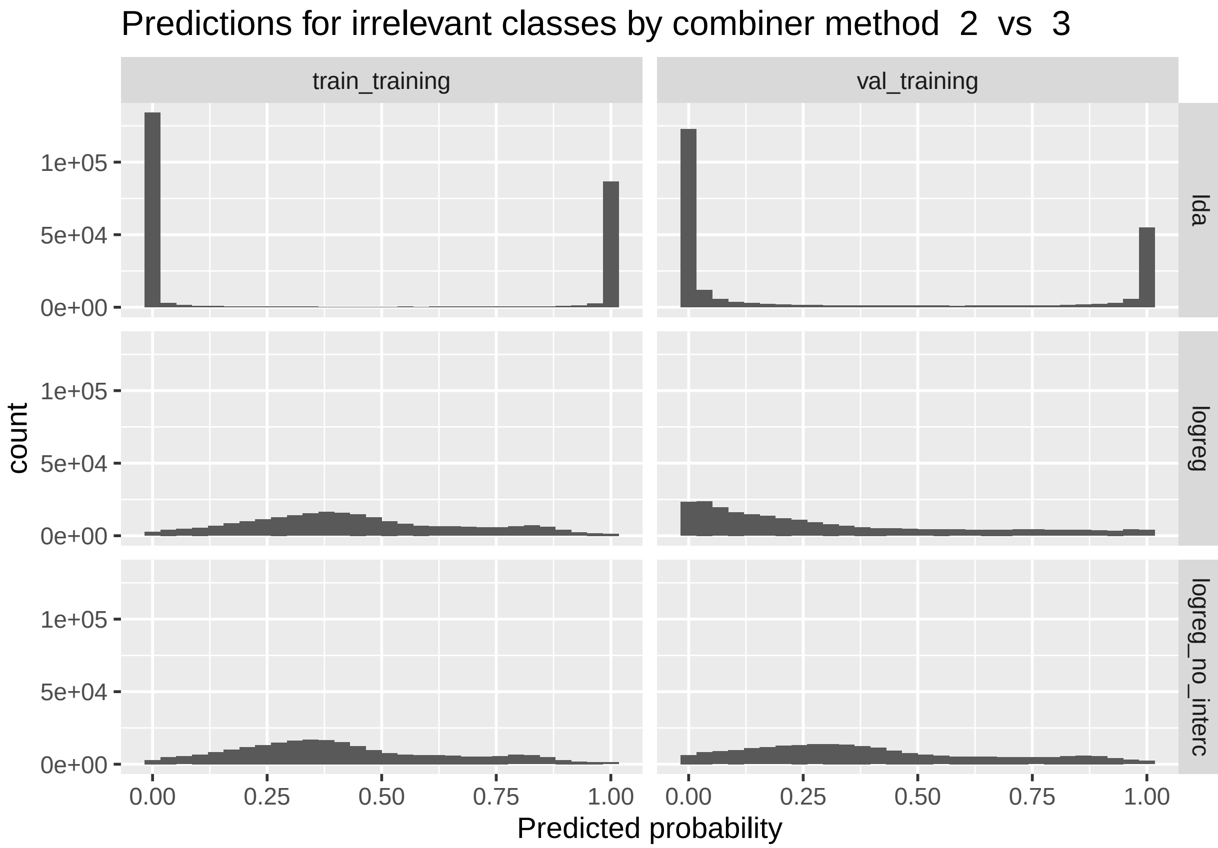

Predictions for irrelevant classes by combiner method 1 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

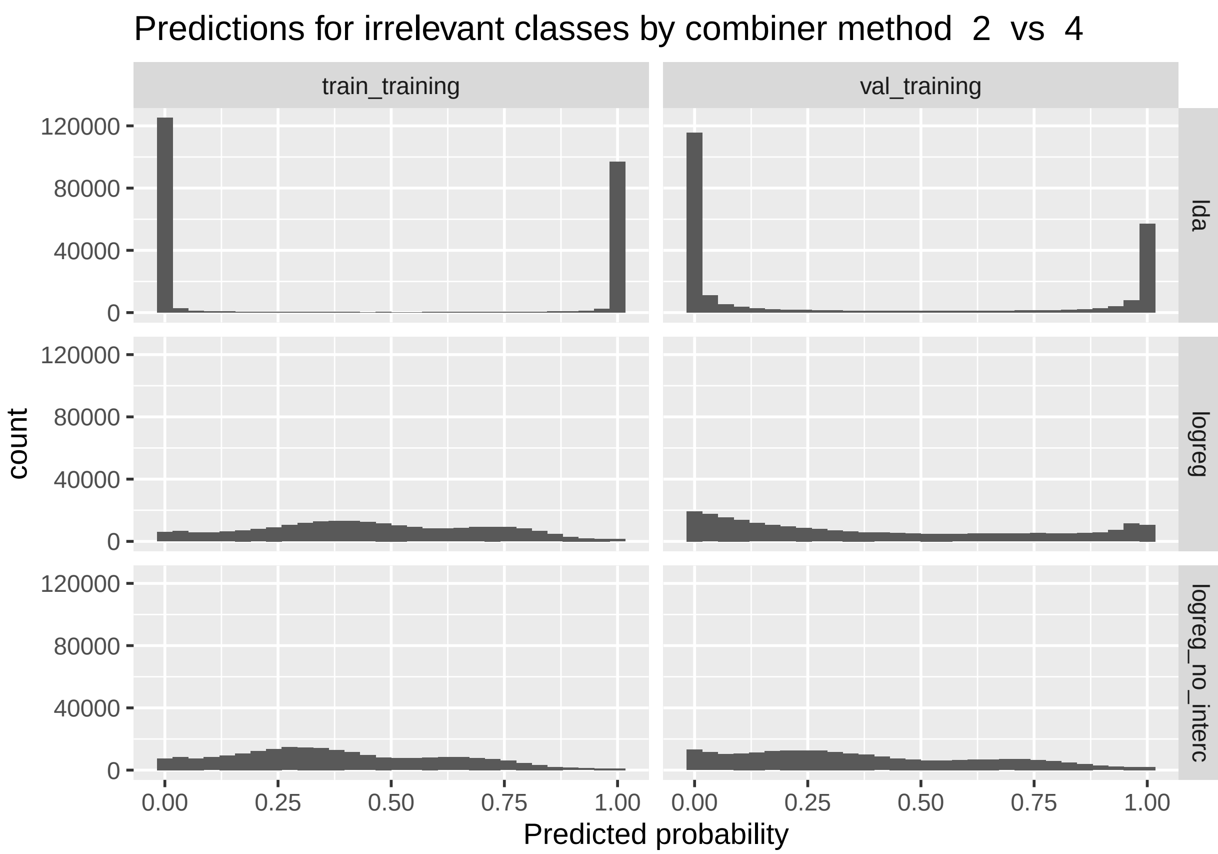

Predictions for irrelevant classes by combiner method 1 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

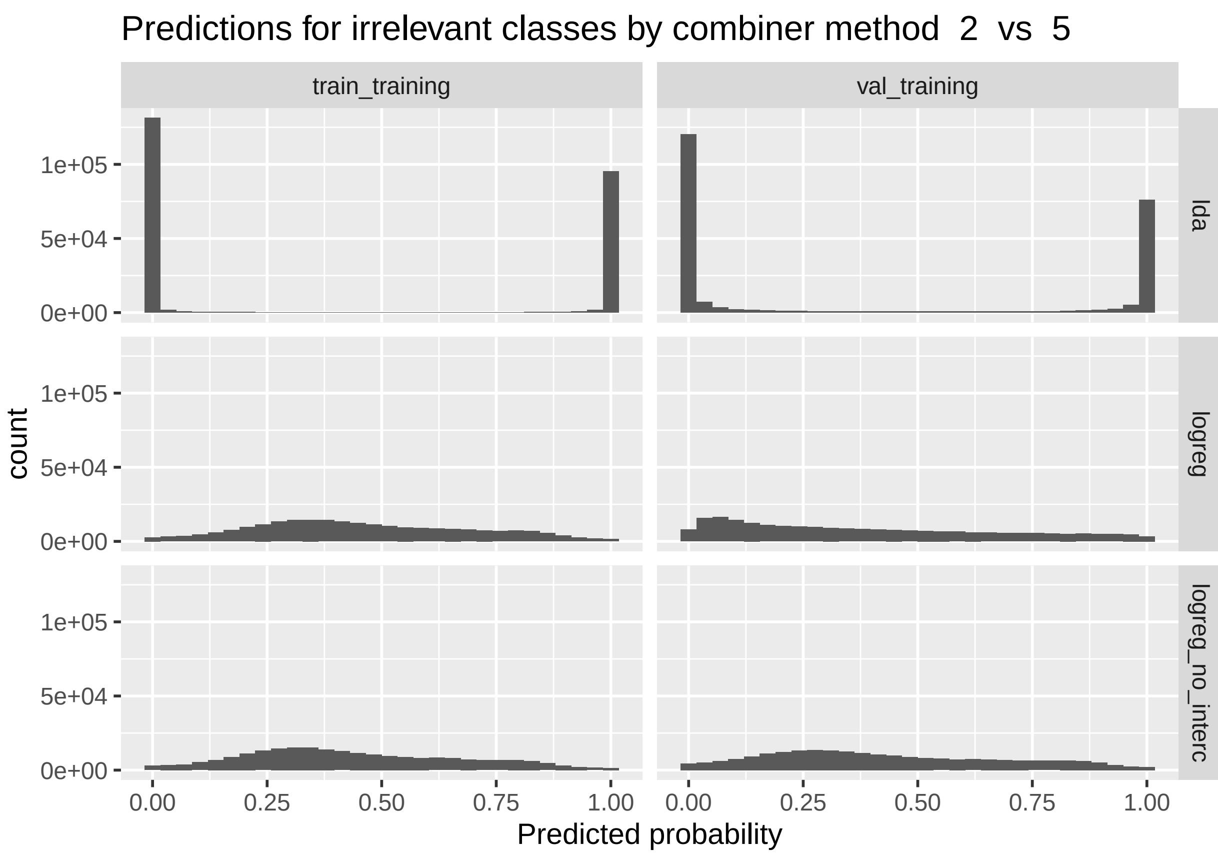

Predictions for irrelevant classes by combiner method 1 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
150000
100000
50000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



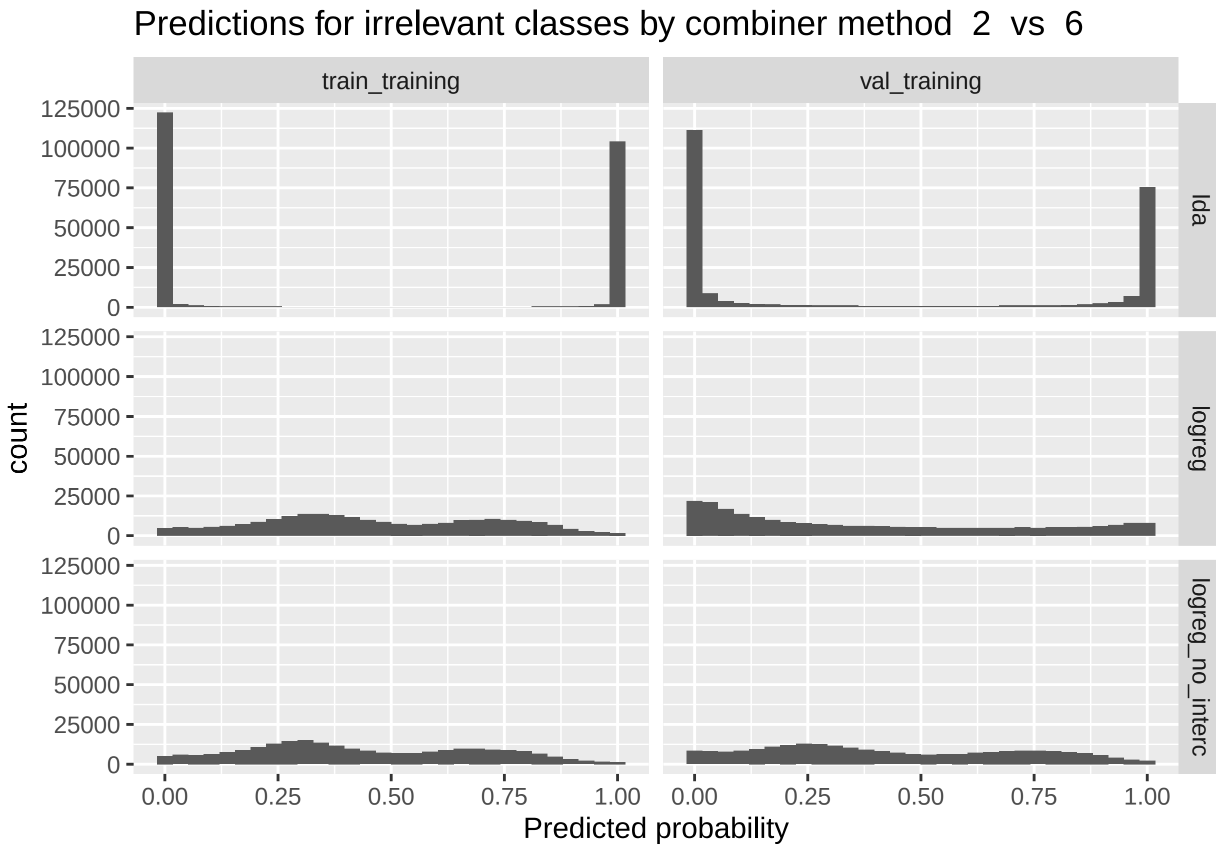
Predictions for irrelevant classes by combiner method 1 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
0
30000
60000
90000
120000
0.00
0.25
0.50
0.75
1.00
Predicted probability

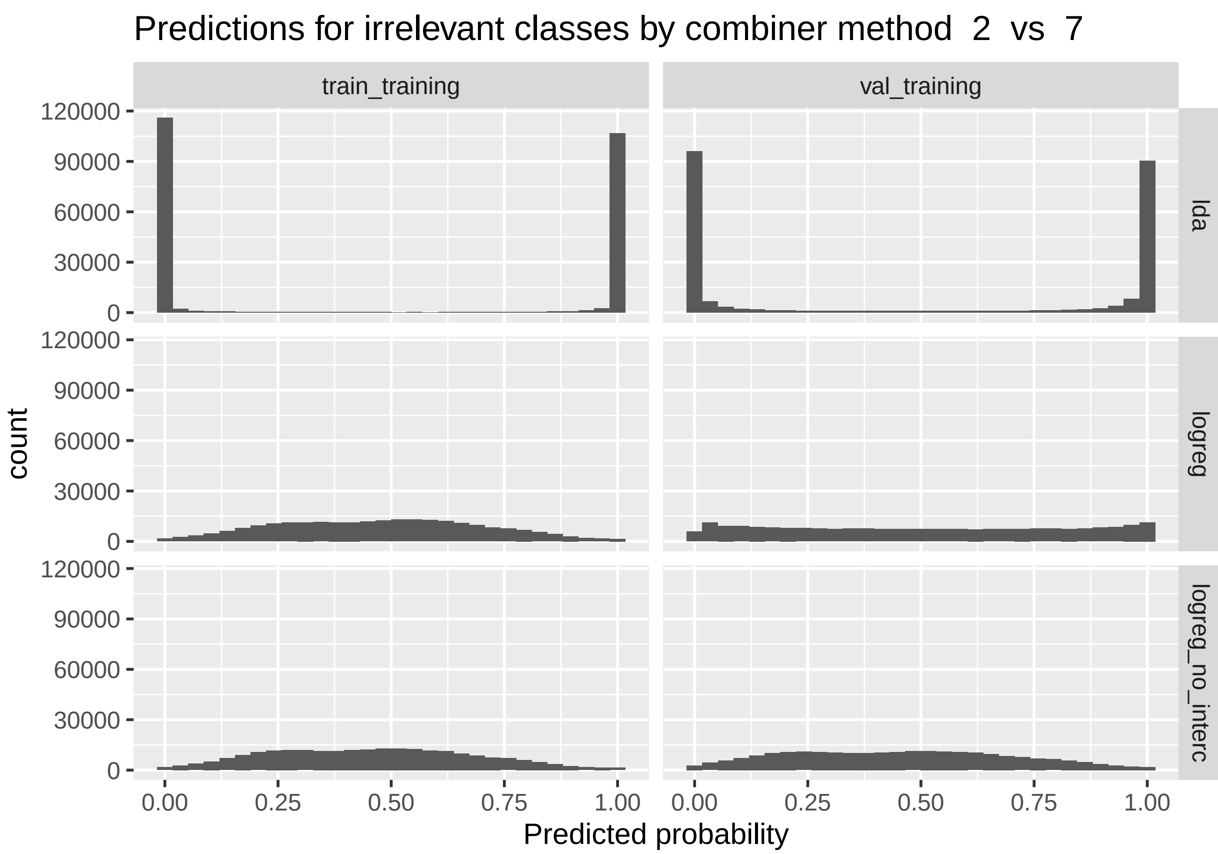

Predictions for irrelevant classes by combiner method 2 vs 3
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 2 vs 4
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
80000
40000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

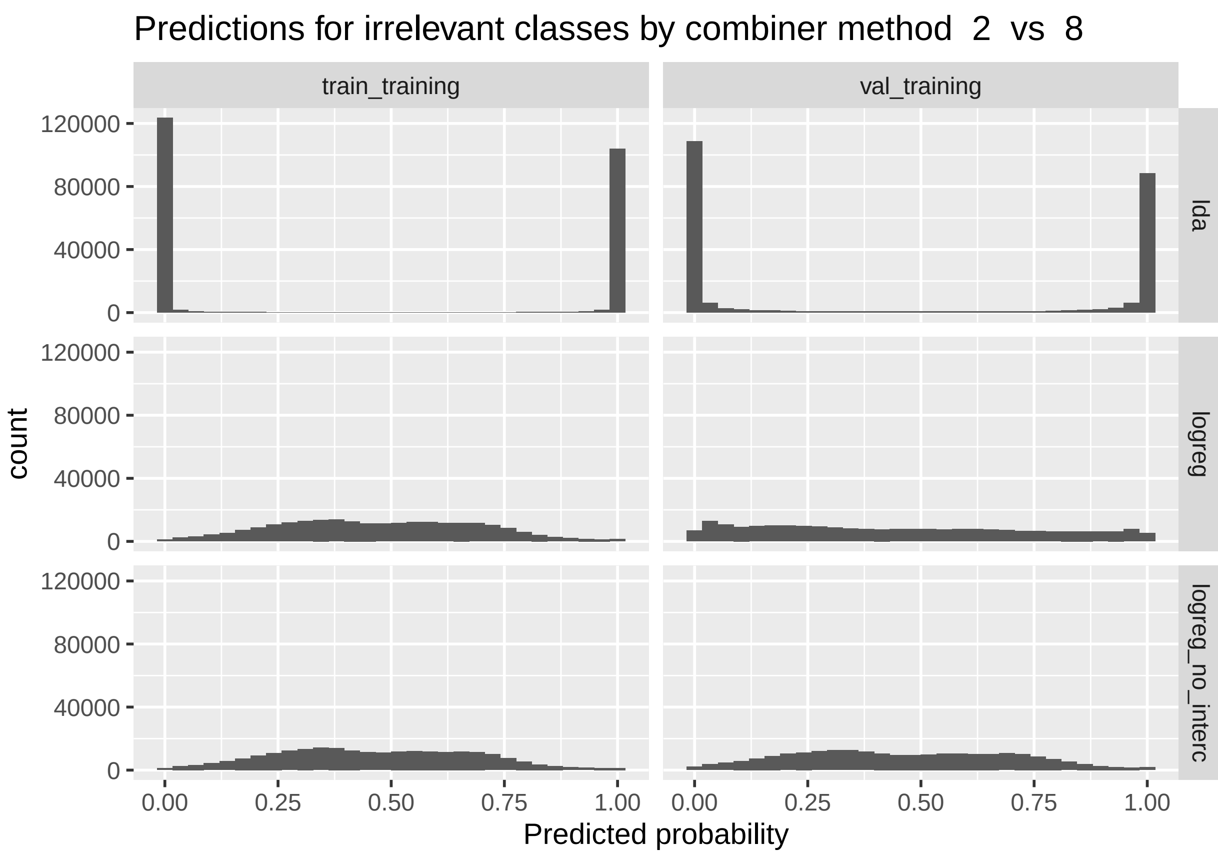


Predictions for irrelevant classes by combiner method 2 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

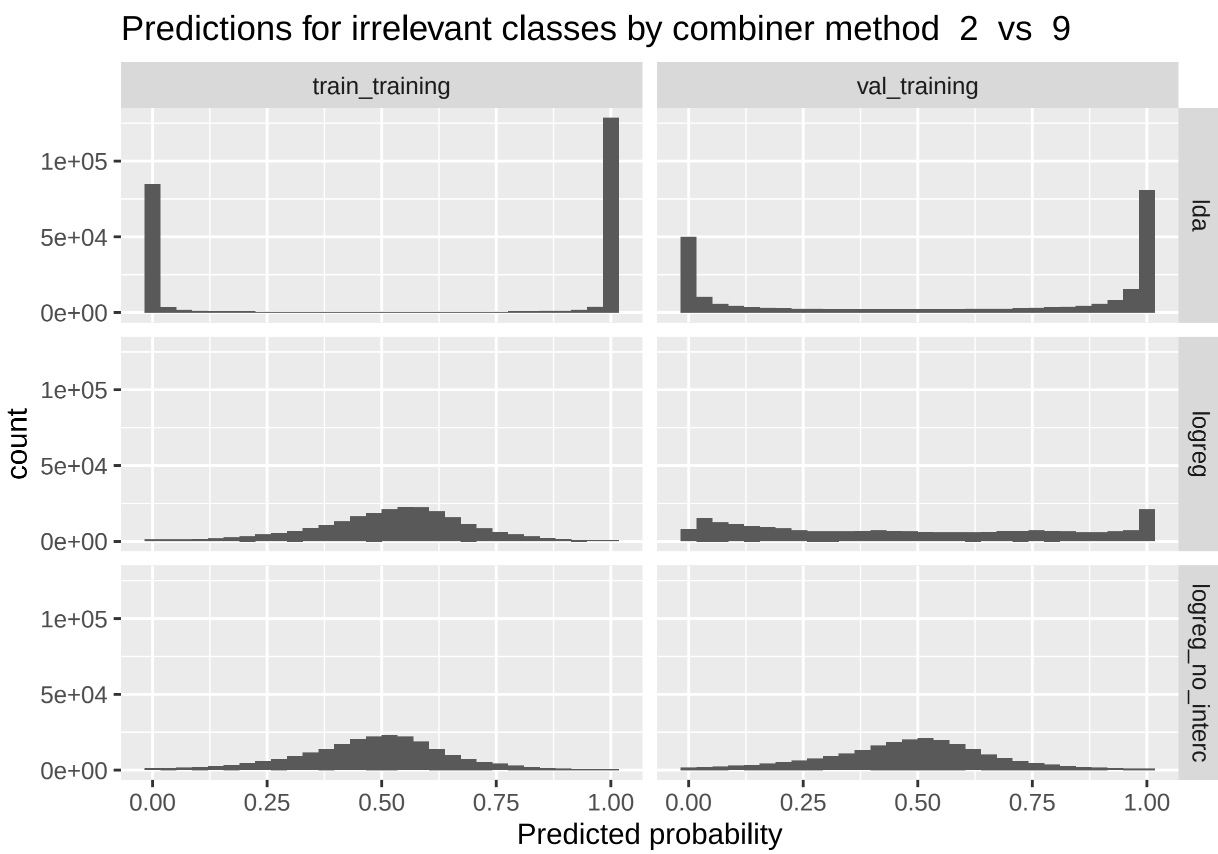

Predictions for irrelevant classes by combiner method 2 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
count
125000
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

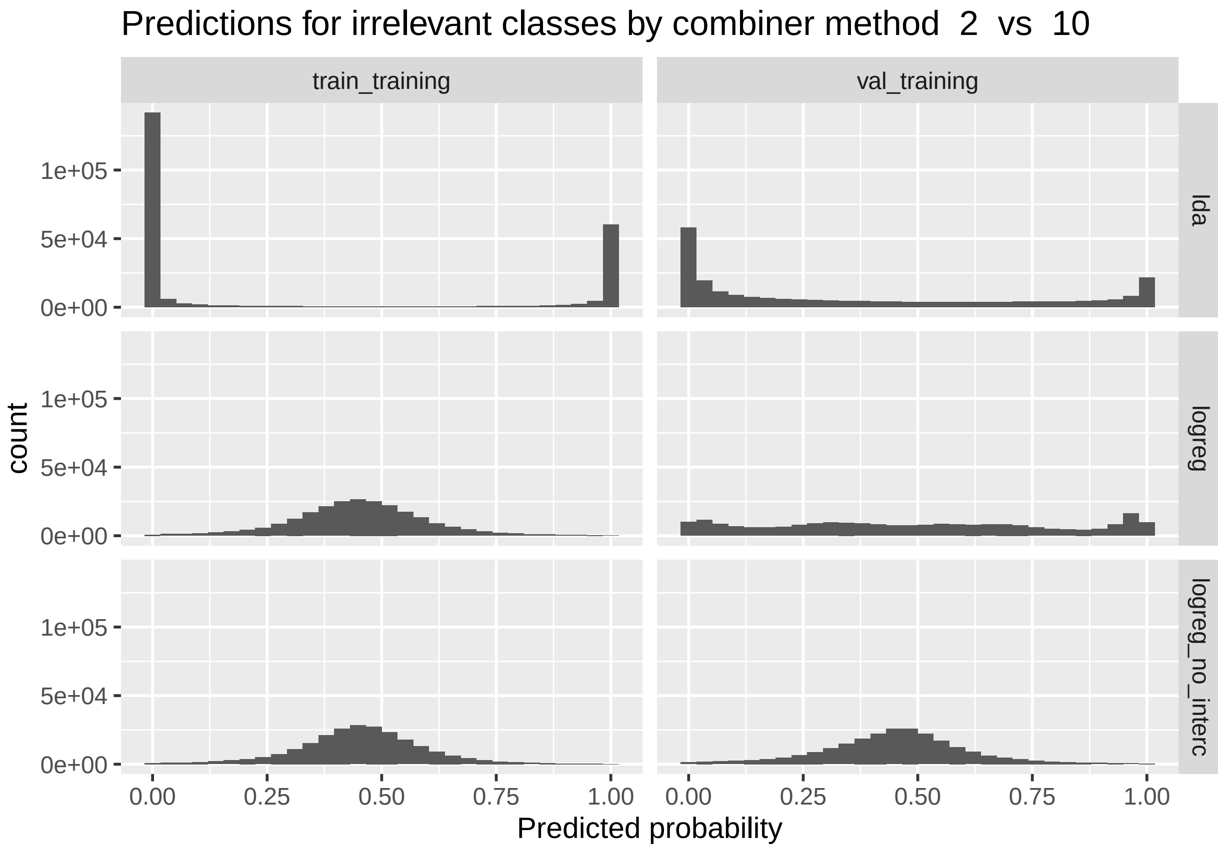


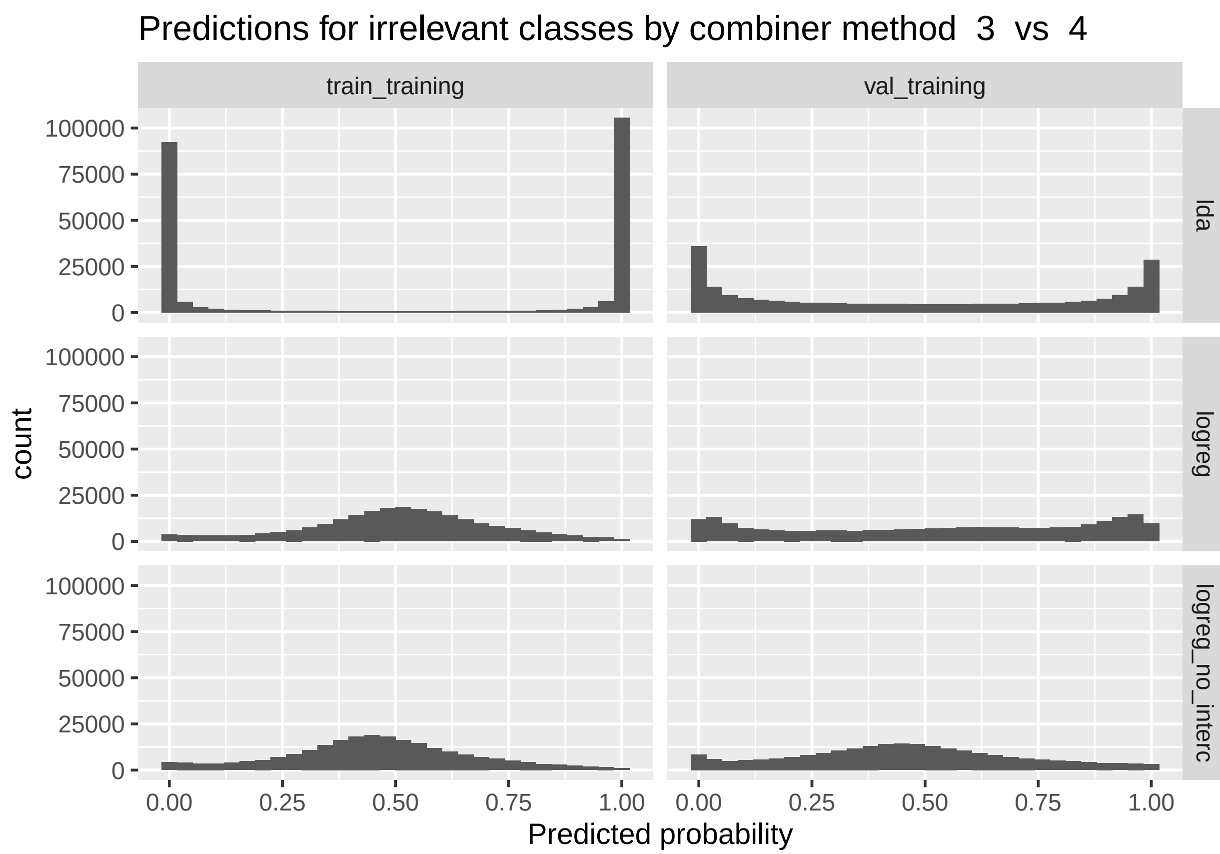
Predictions for irrelevant classes by combiner method 2 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



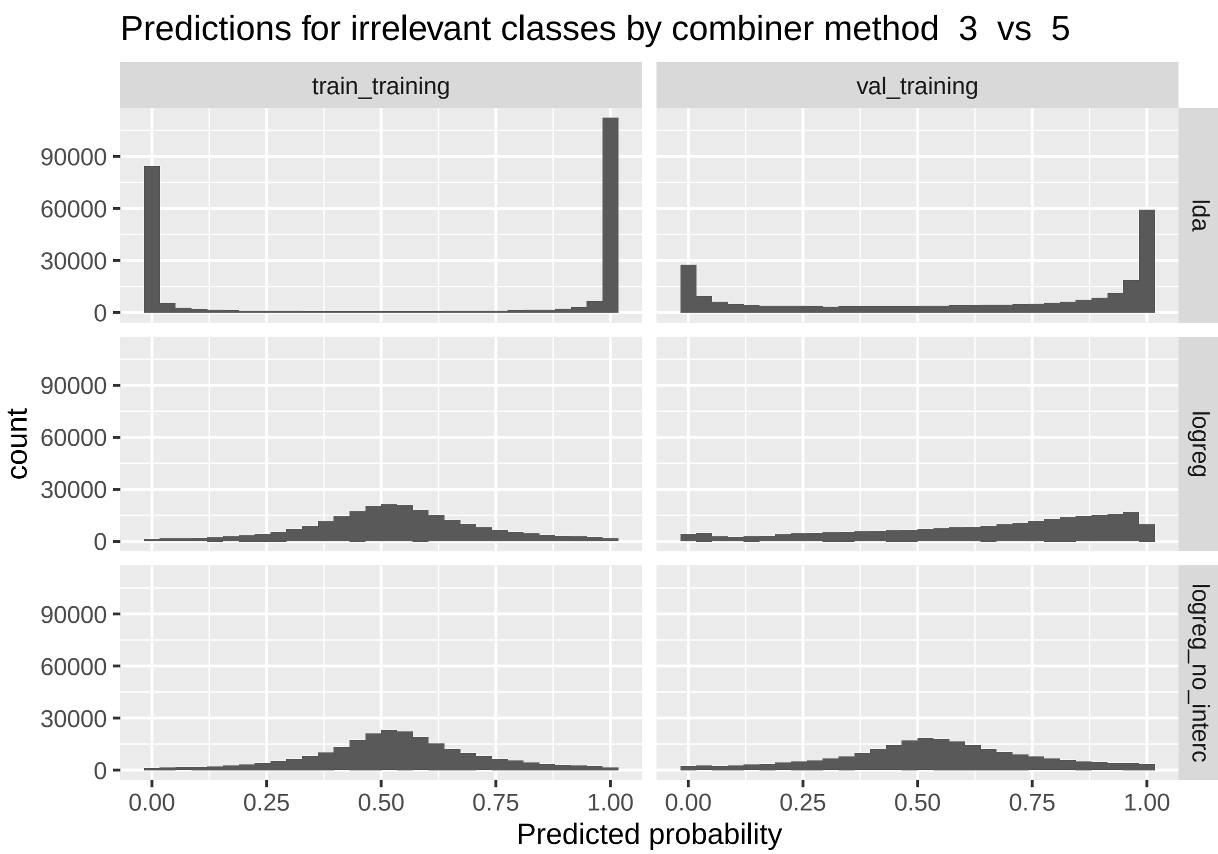
Predictions for irrelevant classes by combiner method 2 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
80000
40000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

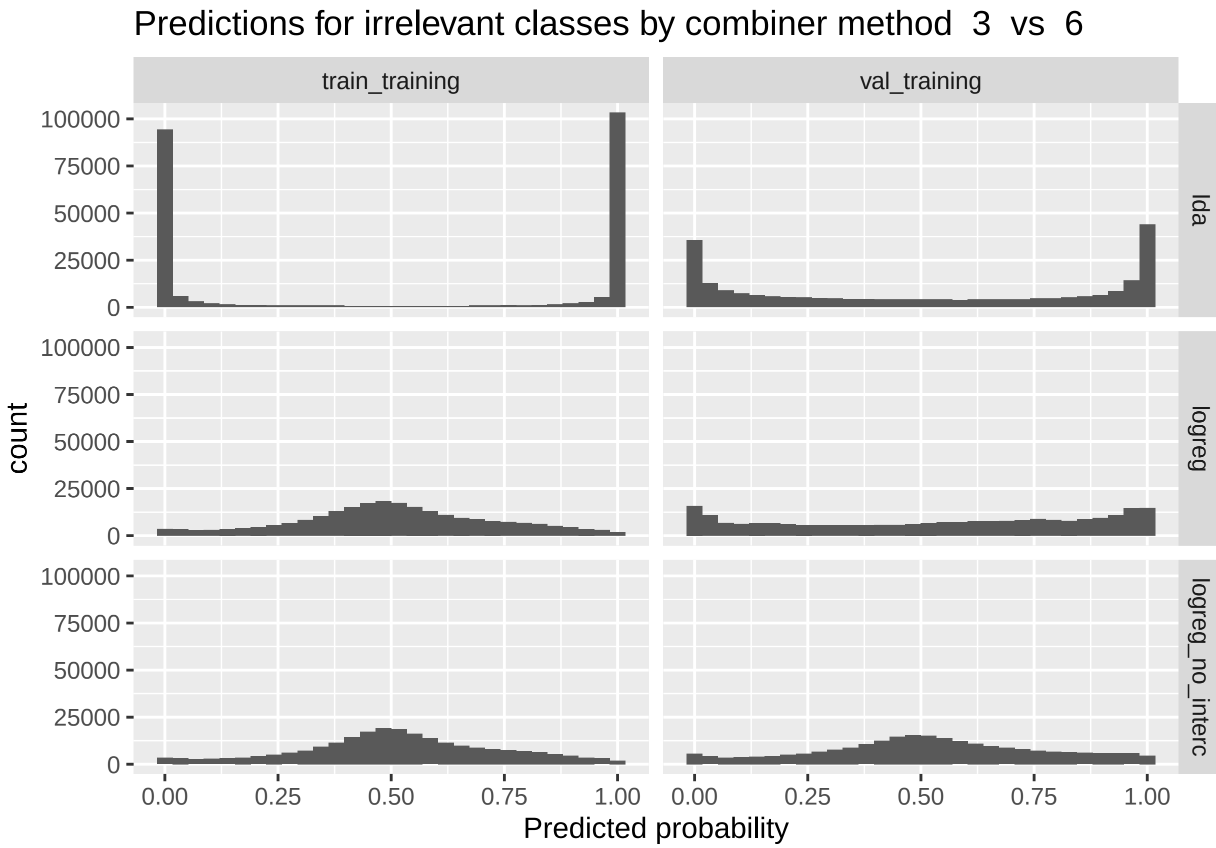

Predictions for irrelevant classes by combiner method 2 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

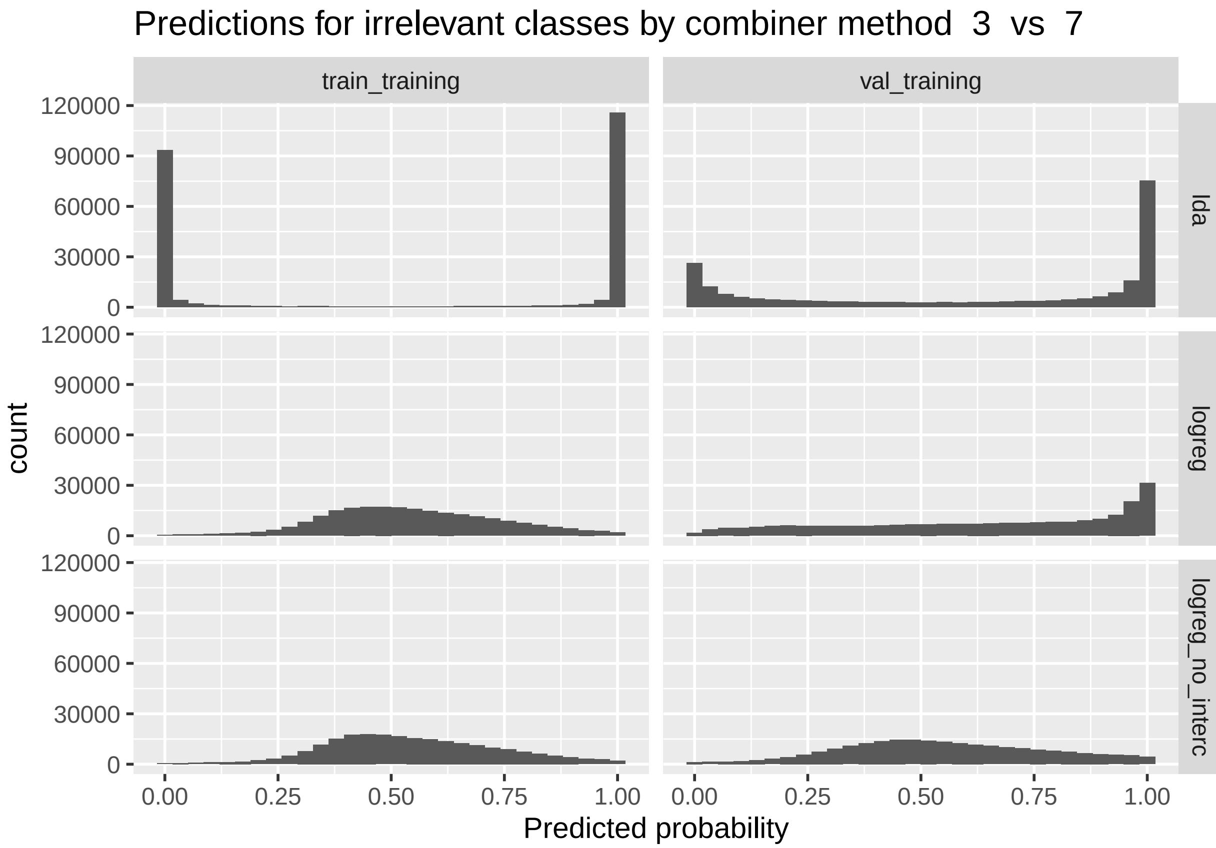

Predictions for irrelevant classes by combiner method 2 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability



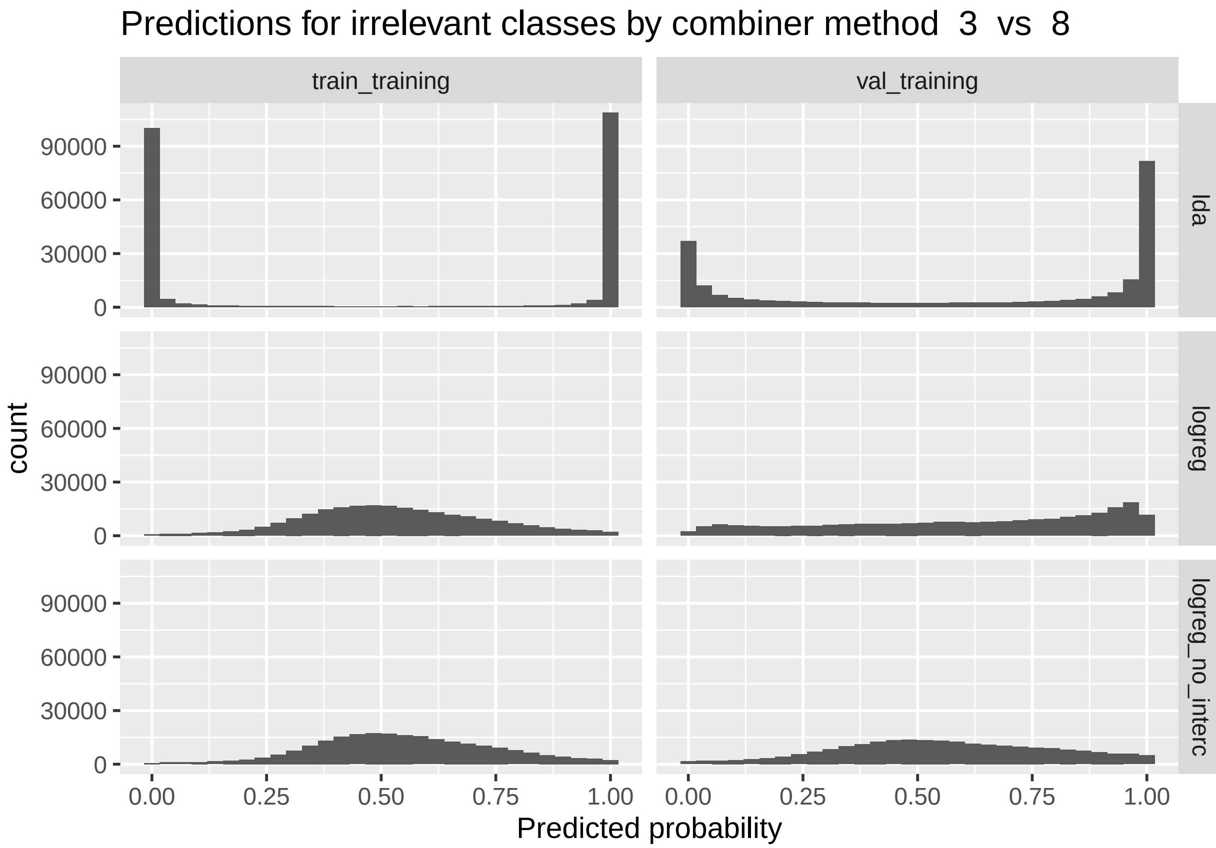
Predictions for irrelevant classes by combiner method 3 vs 4
train_training
val_training
lda
logreg
logreg_no_interc
count
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

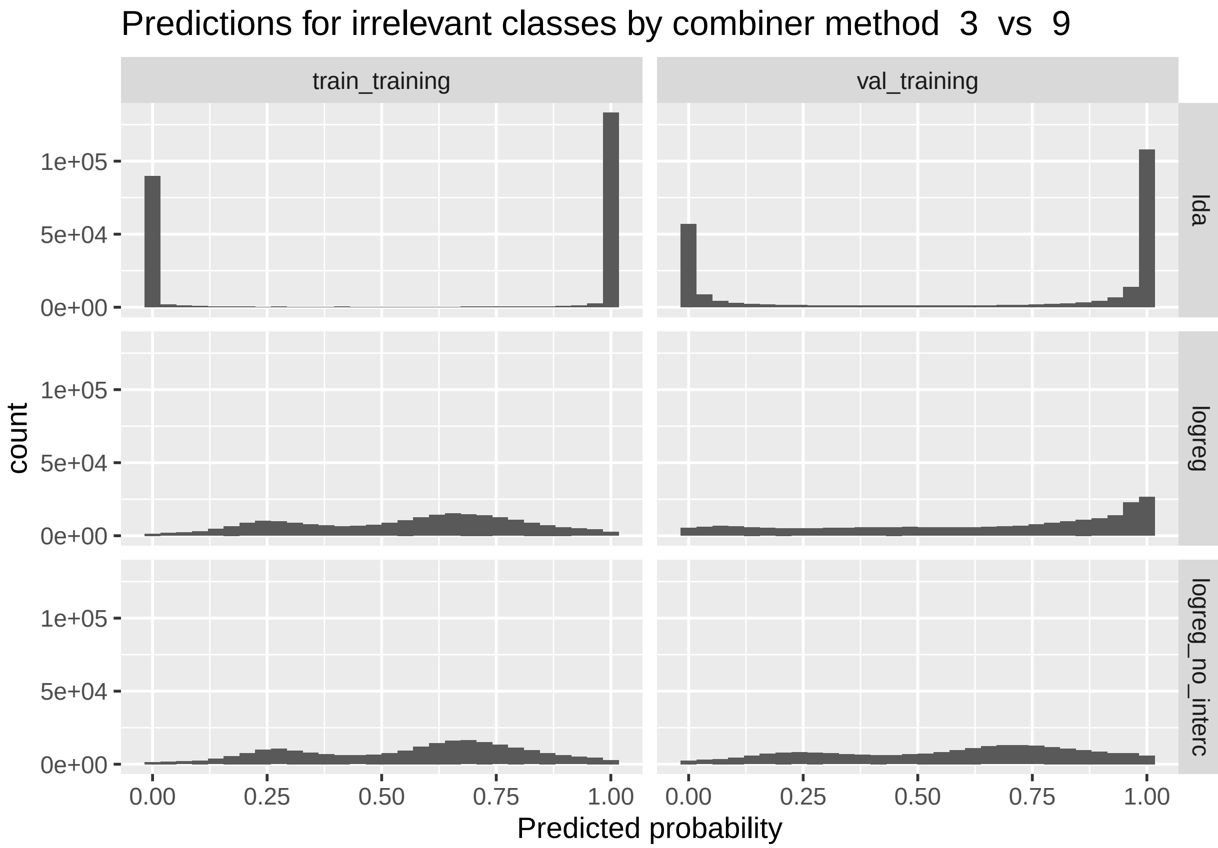

Predictions for irrelevant classes by combiner method 3 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 3 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
count
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

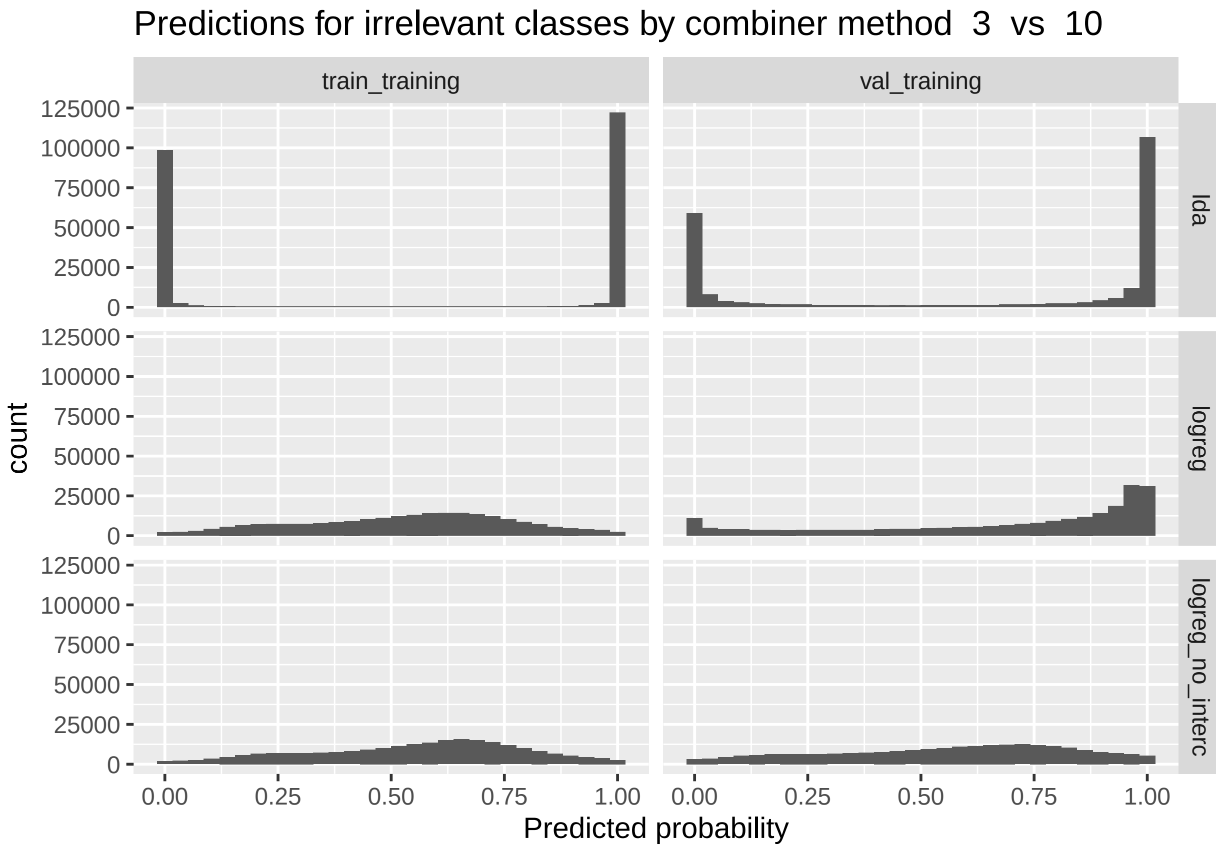


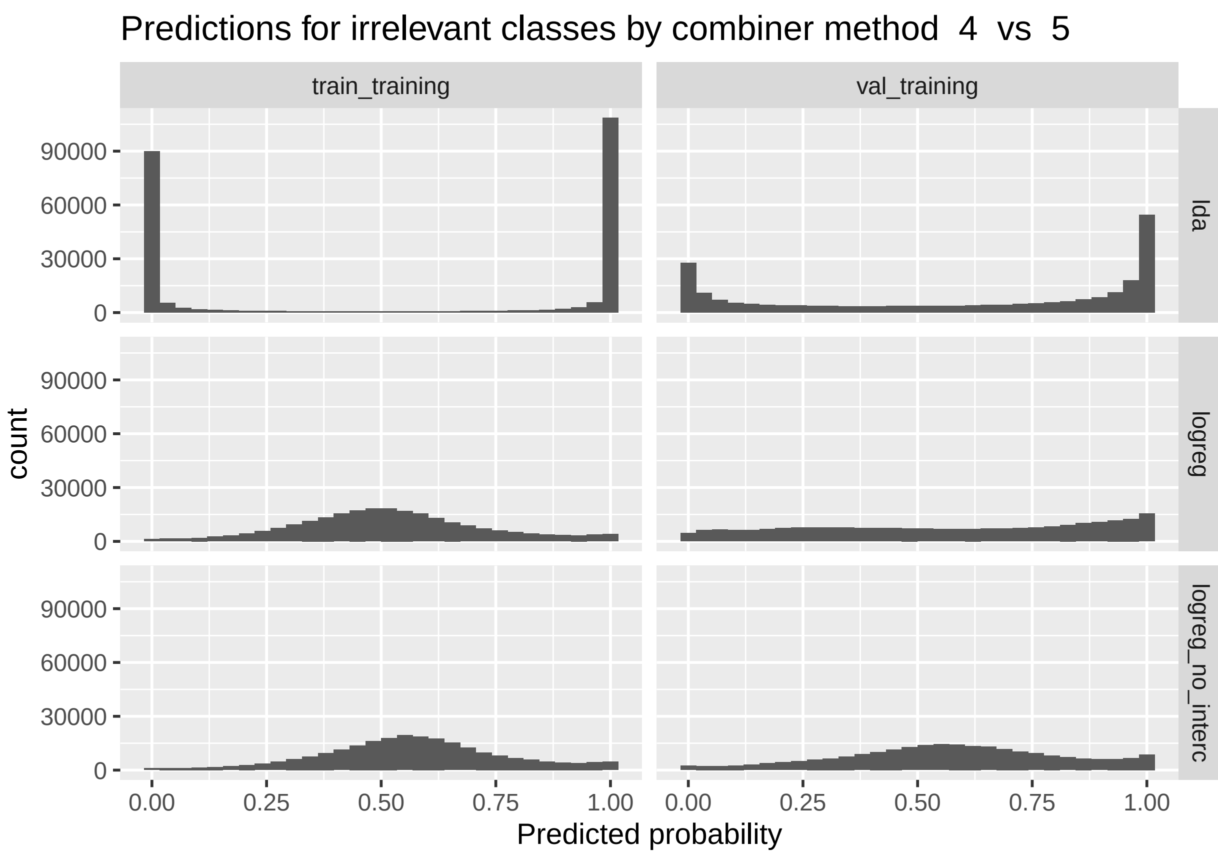
Predictions for irrelevant classes by combiner method 3 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 3 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

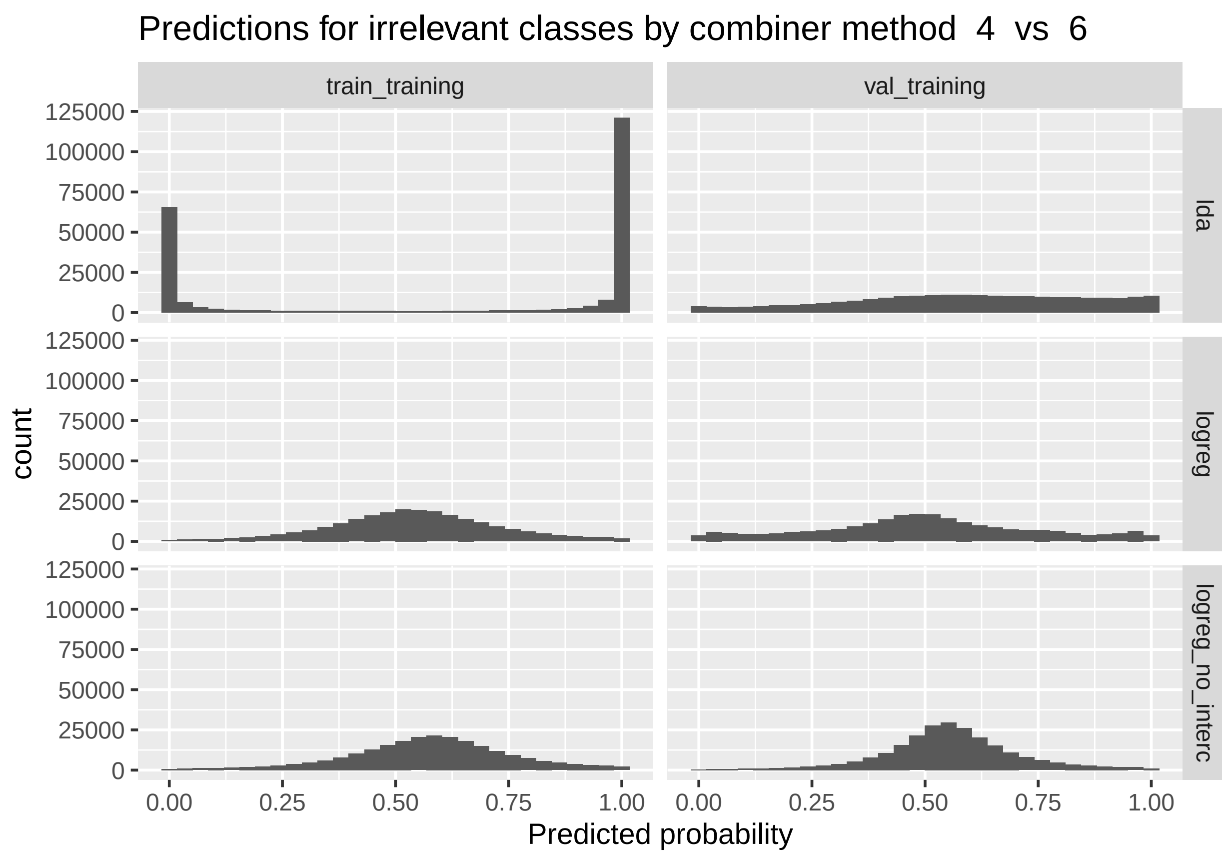


Predictions for irrelevant classes by combiner method 3 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

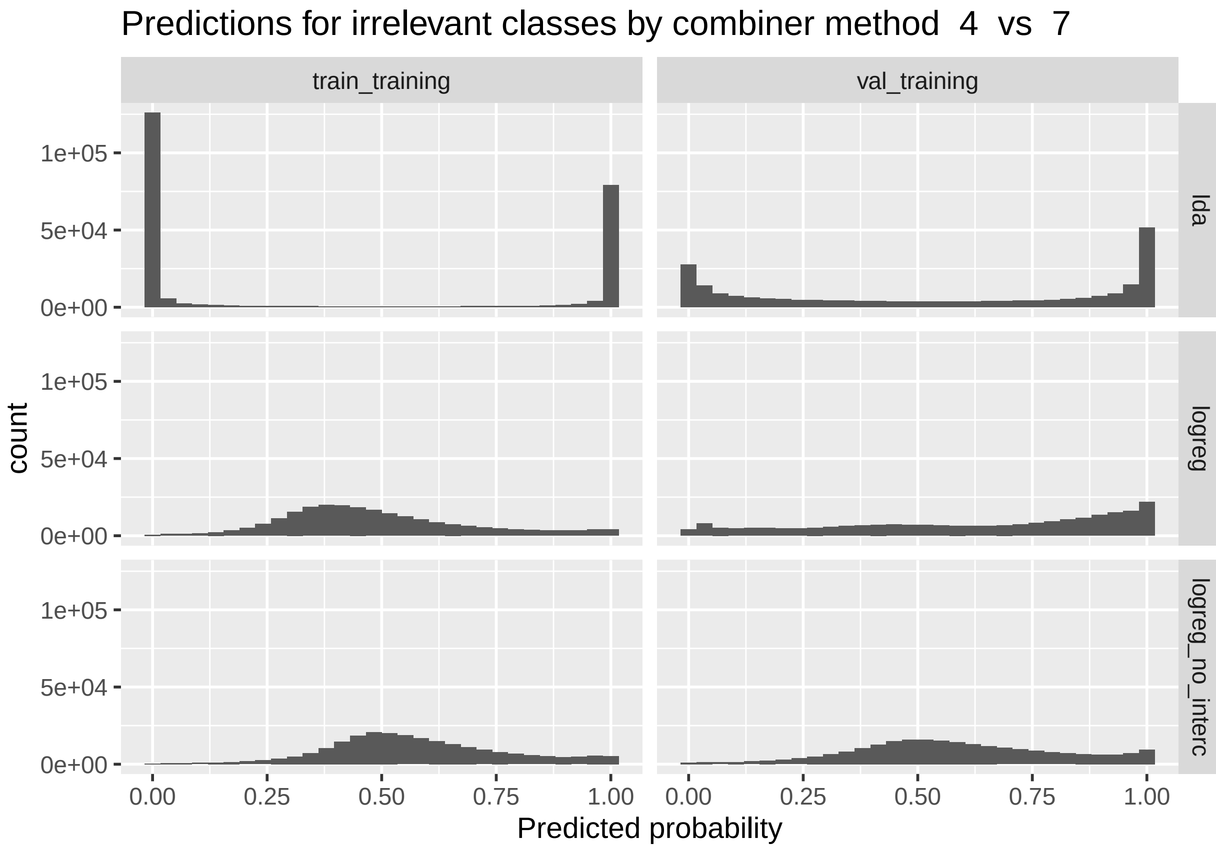

Predictions for irrelevant classes by combiner method 3 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
125000
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



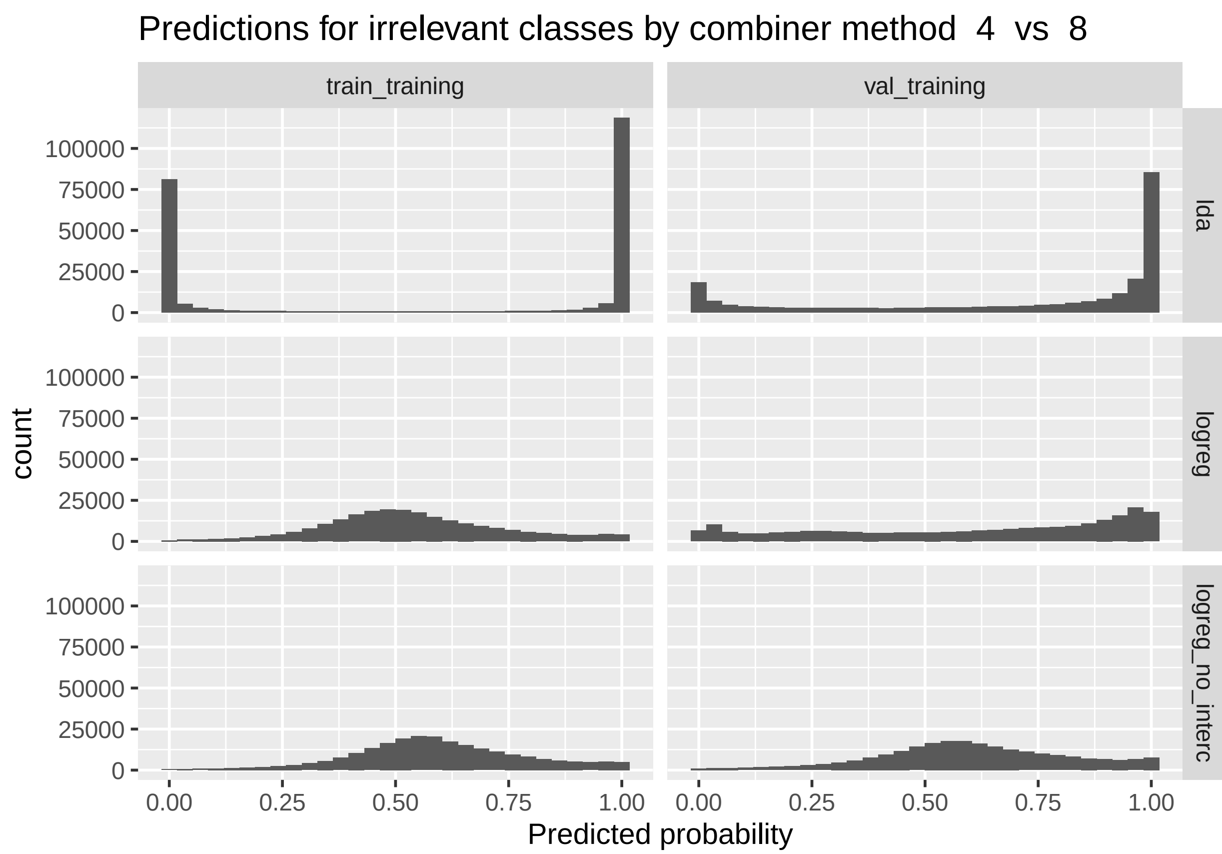
Predictions for irrelevant classes by combiner method 4 vs 5
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



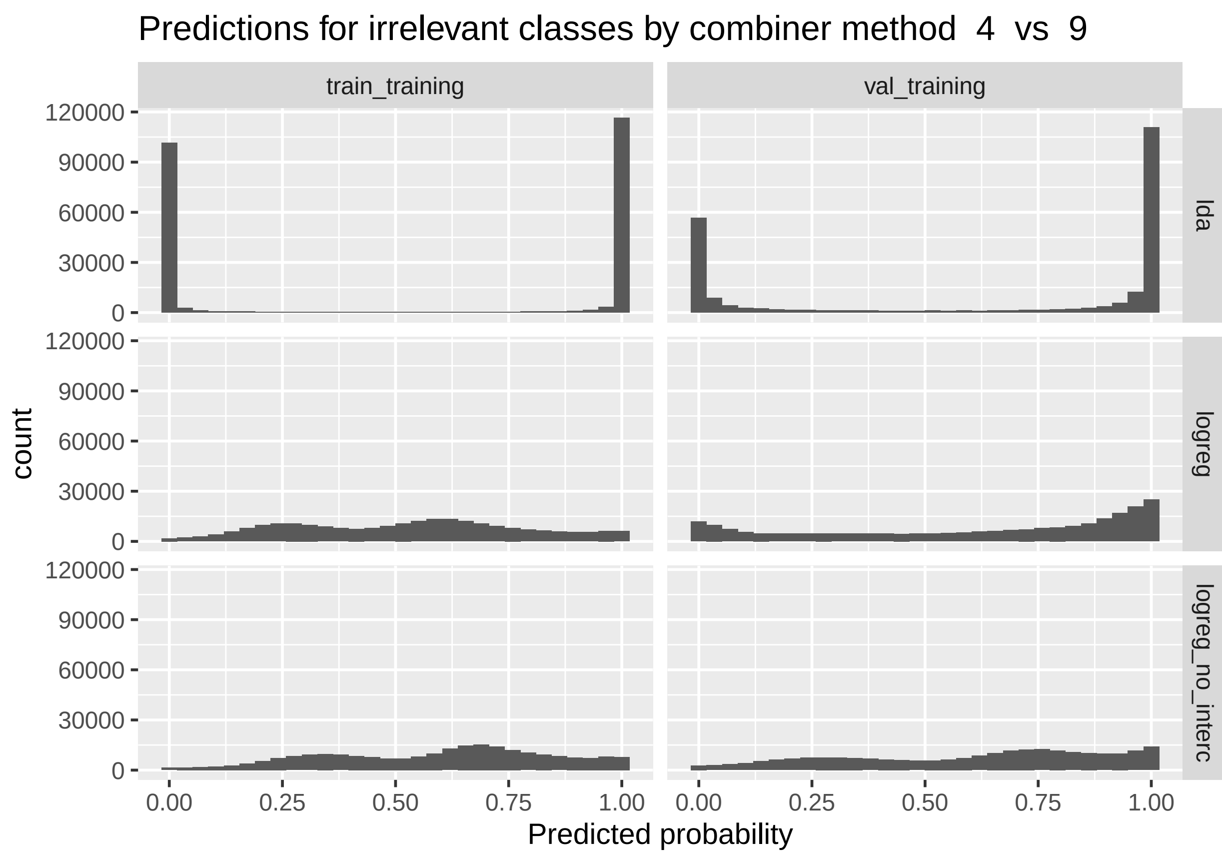
Predictions for irrelevant classes by combiner method 4 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
count
125000
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

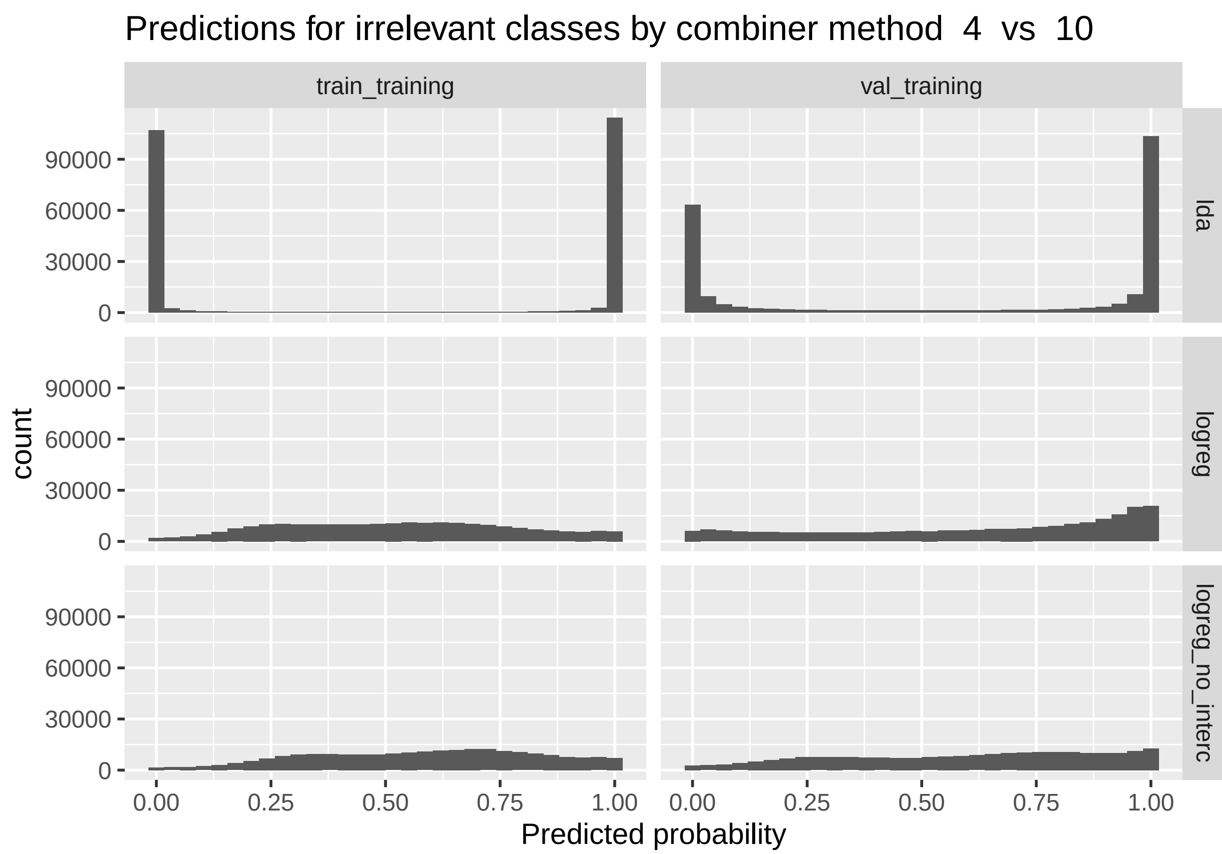


Predictions for irrelevant classes by combiner method 4 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

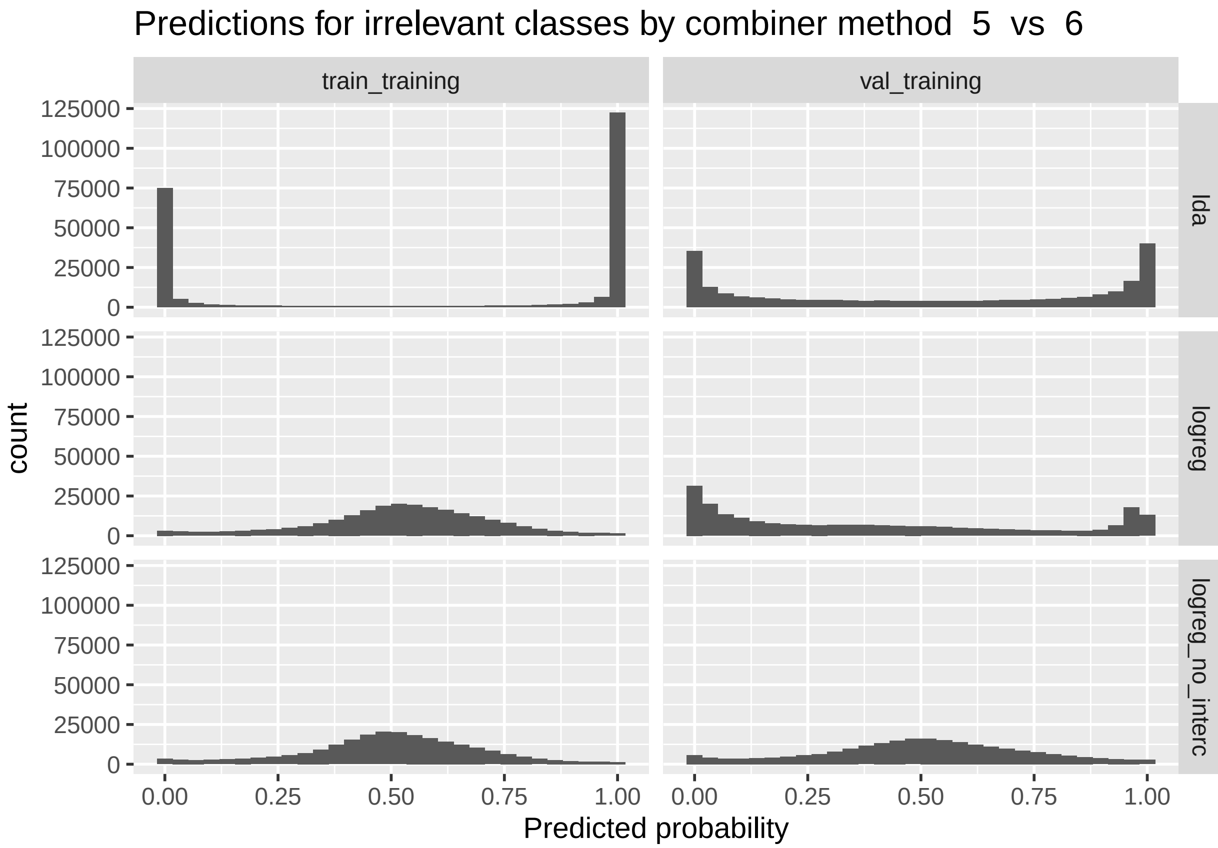


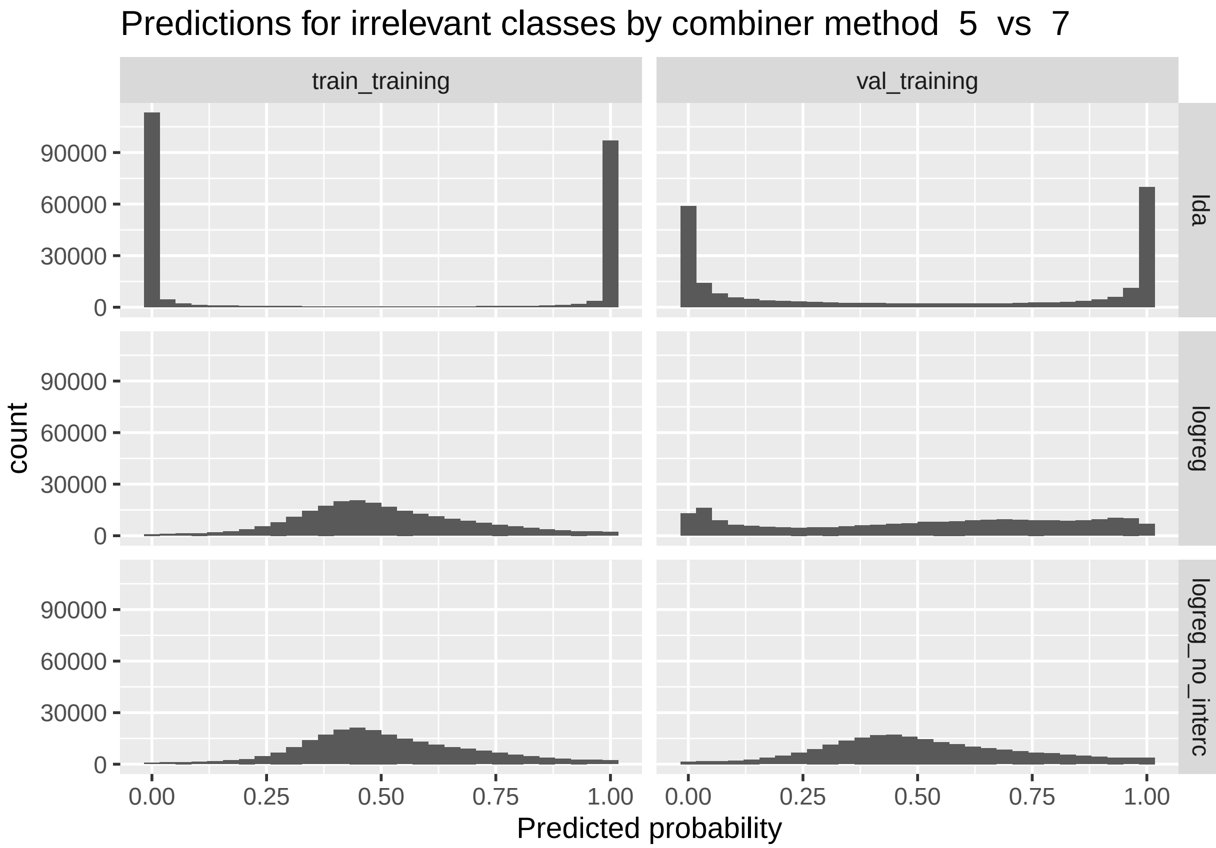
Predictions for irrelevant classes by combiner method 4 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 4 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

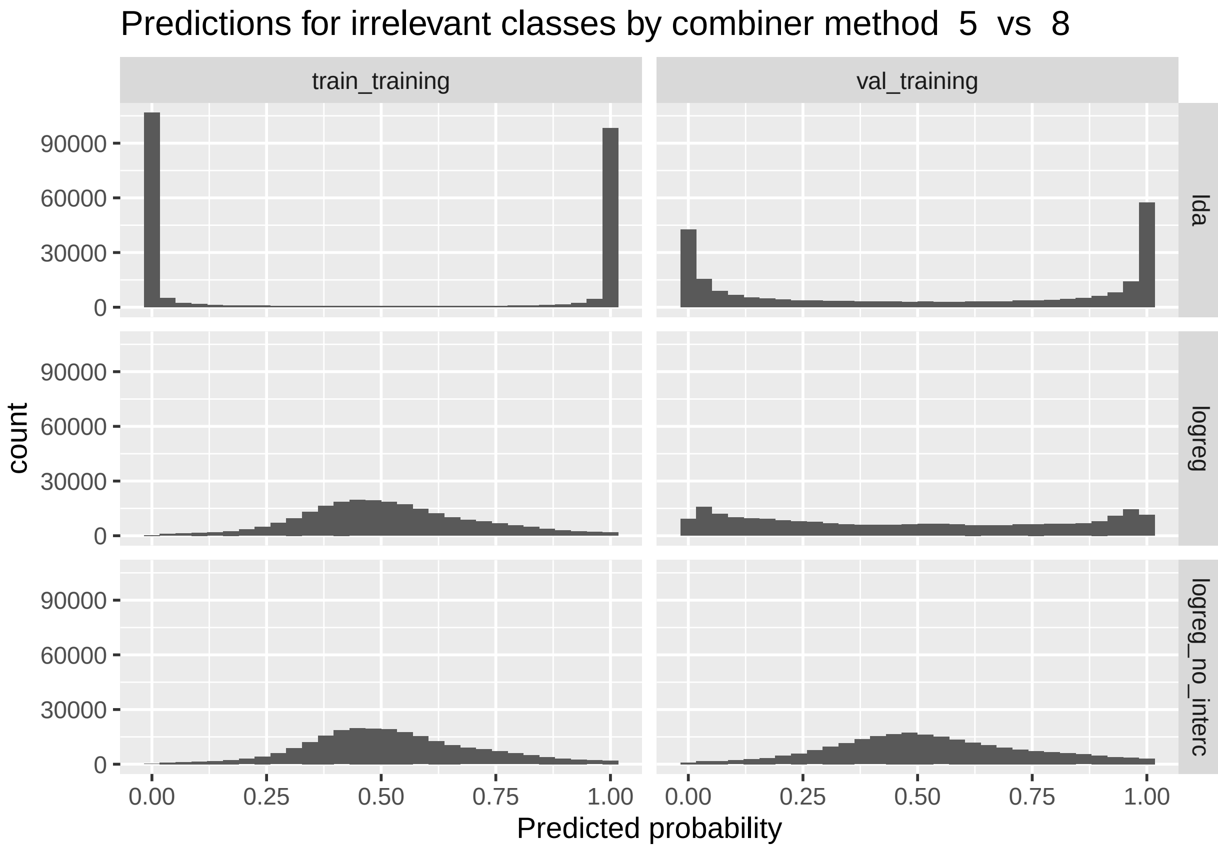


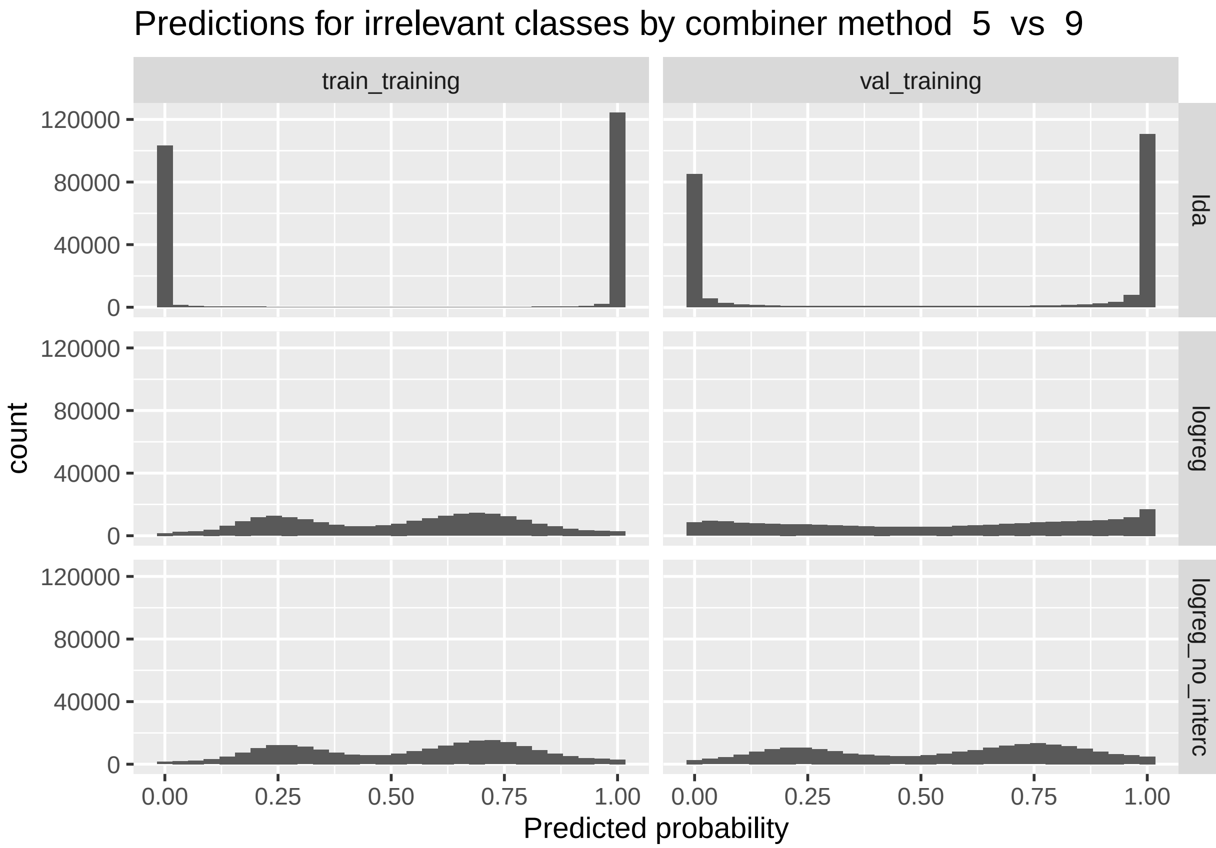
Predictions for irrelevant classes by combiner method 4 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 5 vs 6
train_training
val_training
lda
logreg
logreg_no_interc
count
125000
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

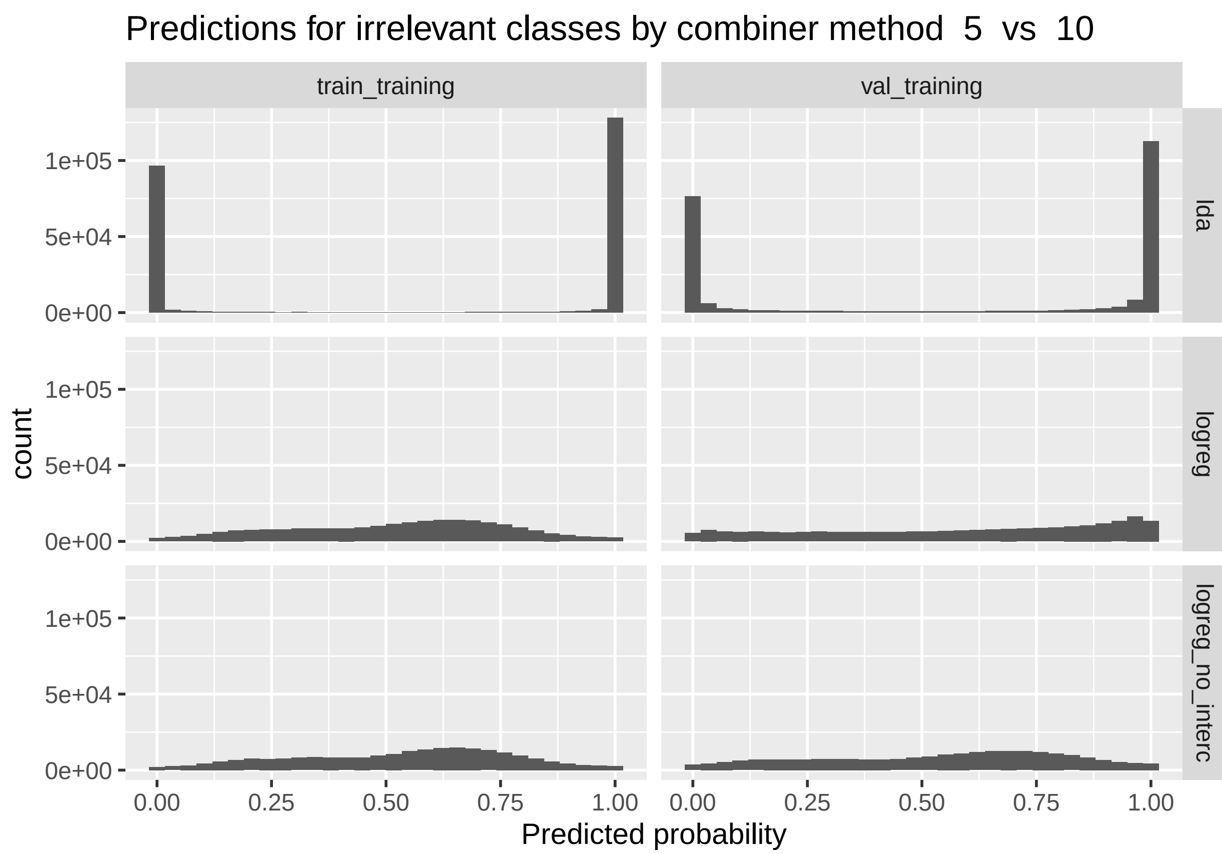


Predictions for irrelevant classes by combiner method 5 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

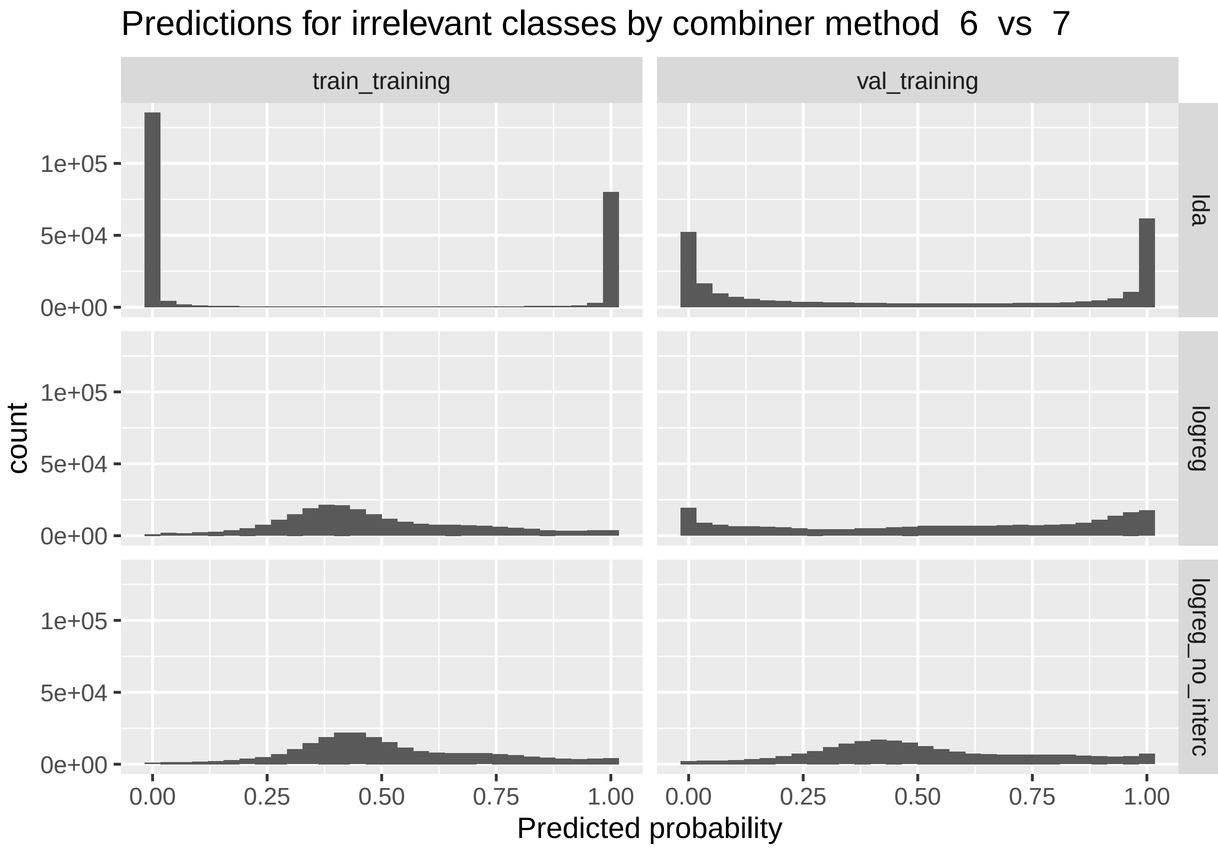

Predictions for irrelevant classes by combiner method 5 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

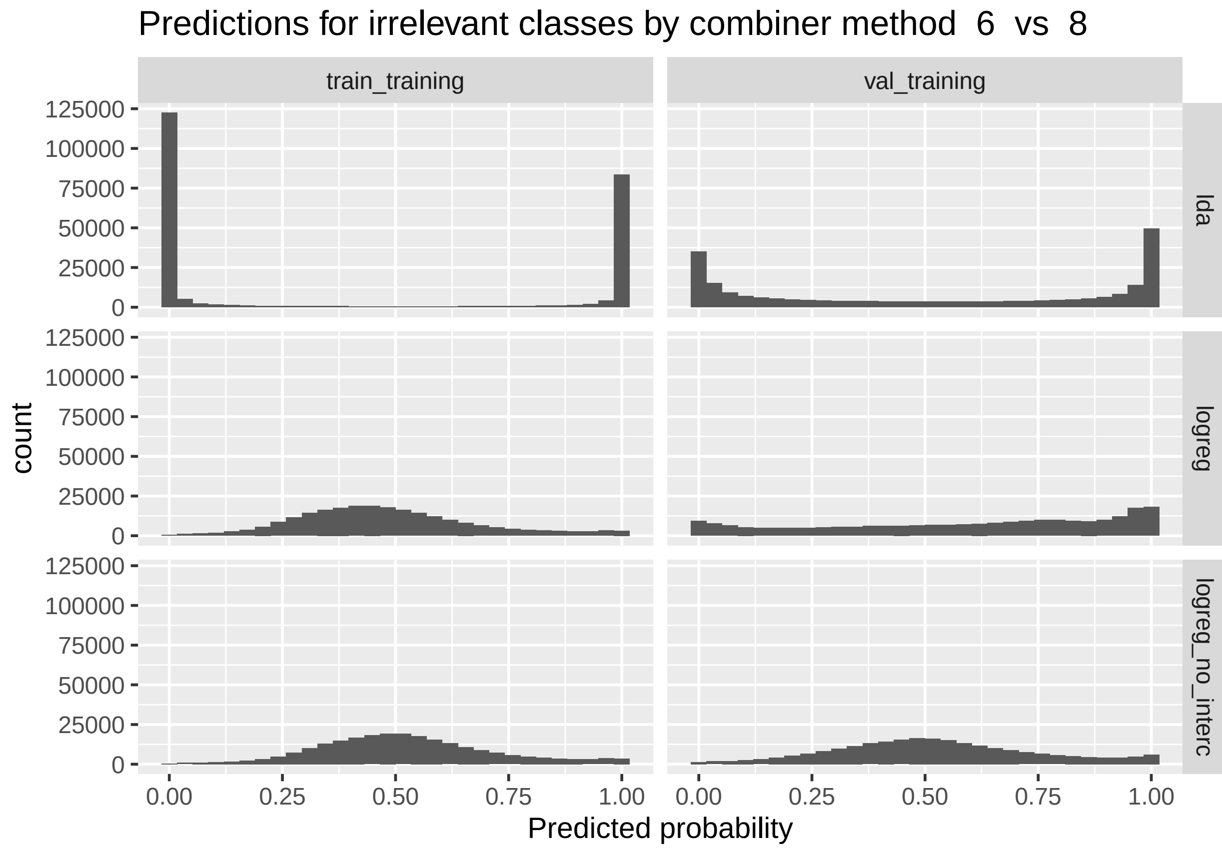


Predictions for irrelevant classes by combiner method 5 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
80000
40000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

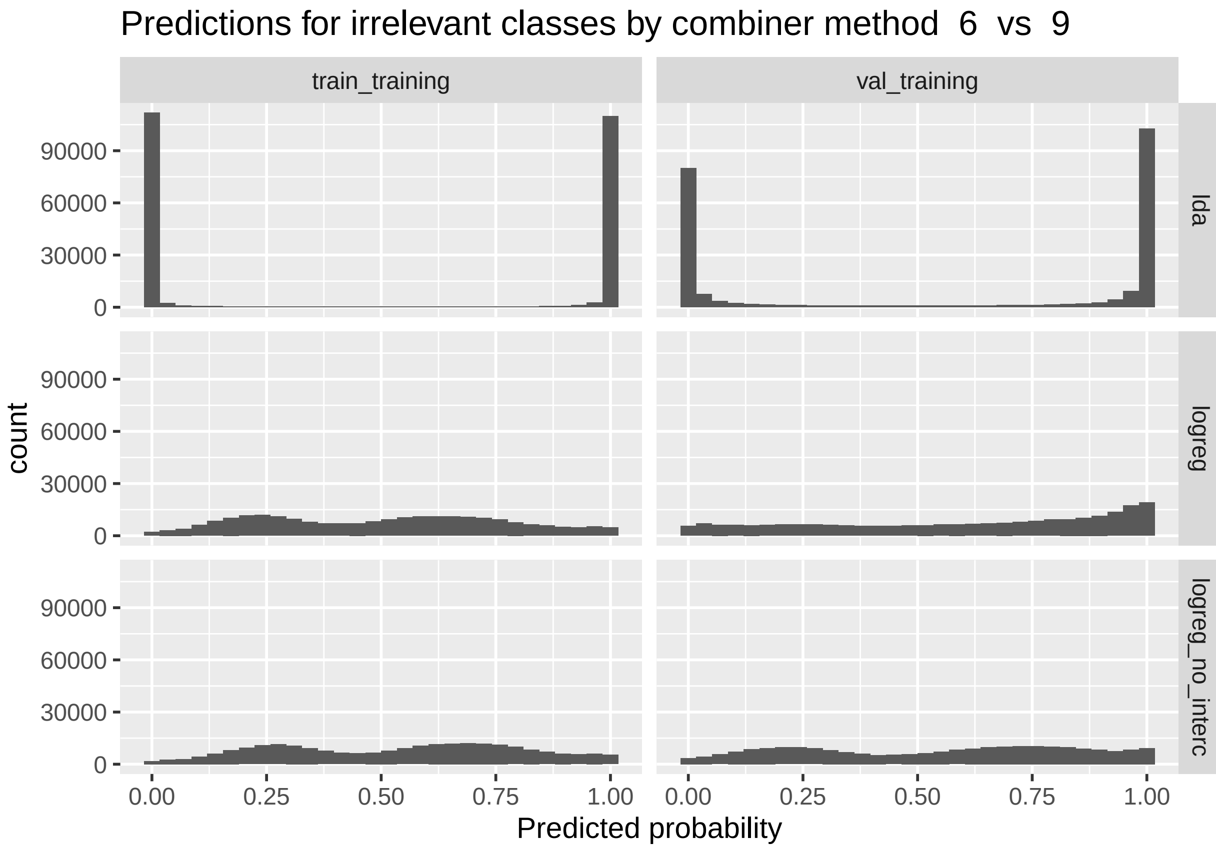


Predictions for irrelevant classes by combiner method 5 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
1e+05
5e+04
0e+00
count
0.00
0.25
0.50
0.75
1.00
Predicted probability

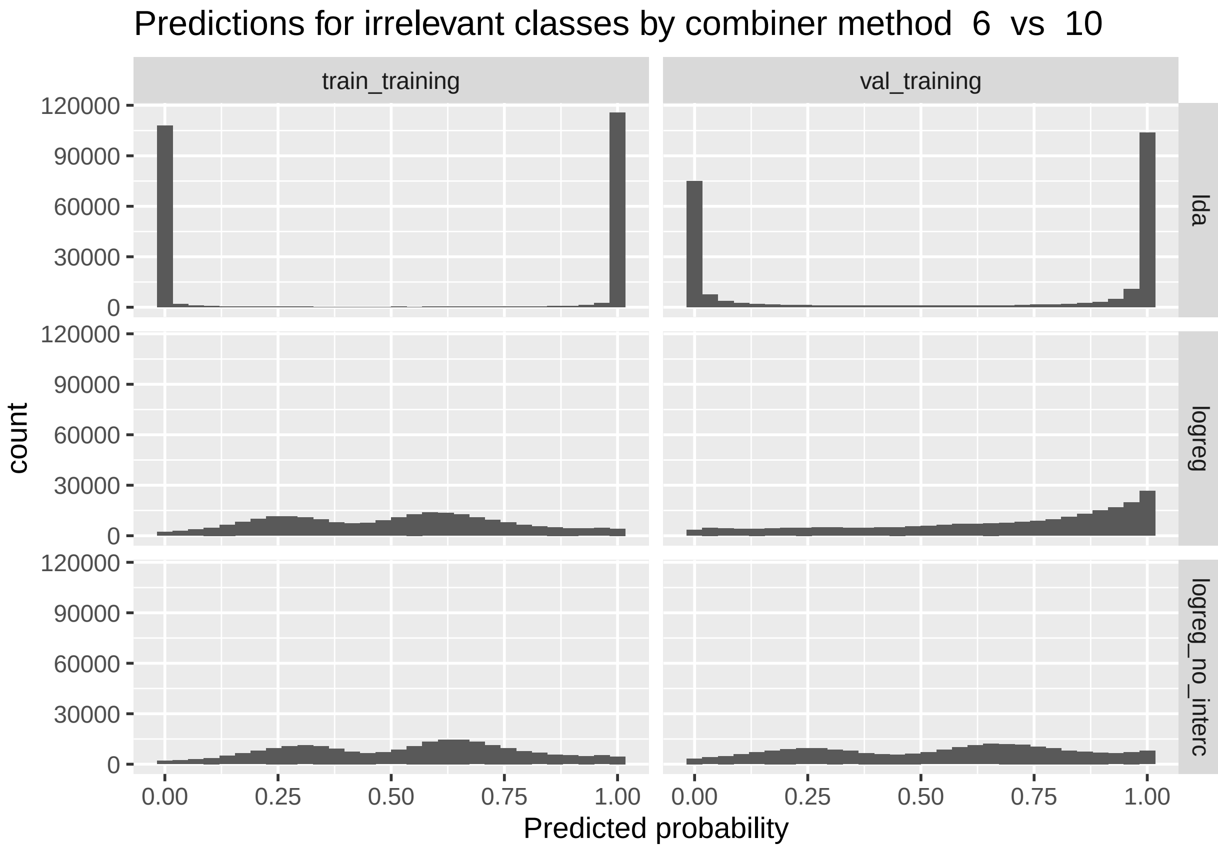

Predictions for irrelevant classes by combiner method 6 vs 7
train_training
val_training
lda
logreg
logreg_no_interc
count
1e+05
5e+04
0e+00
0.00
0.25
0.50
0.75
1.00
Predicted probability

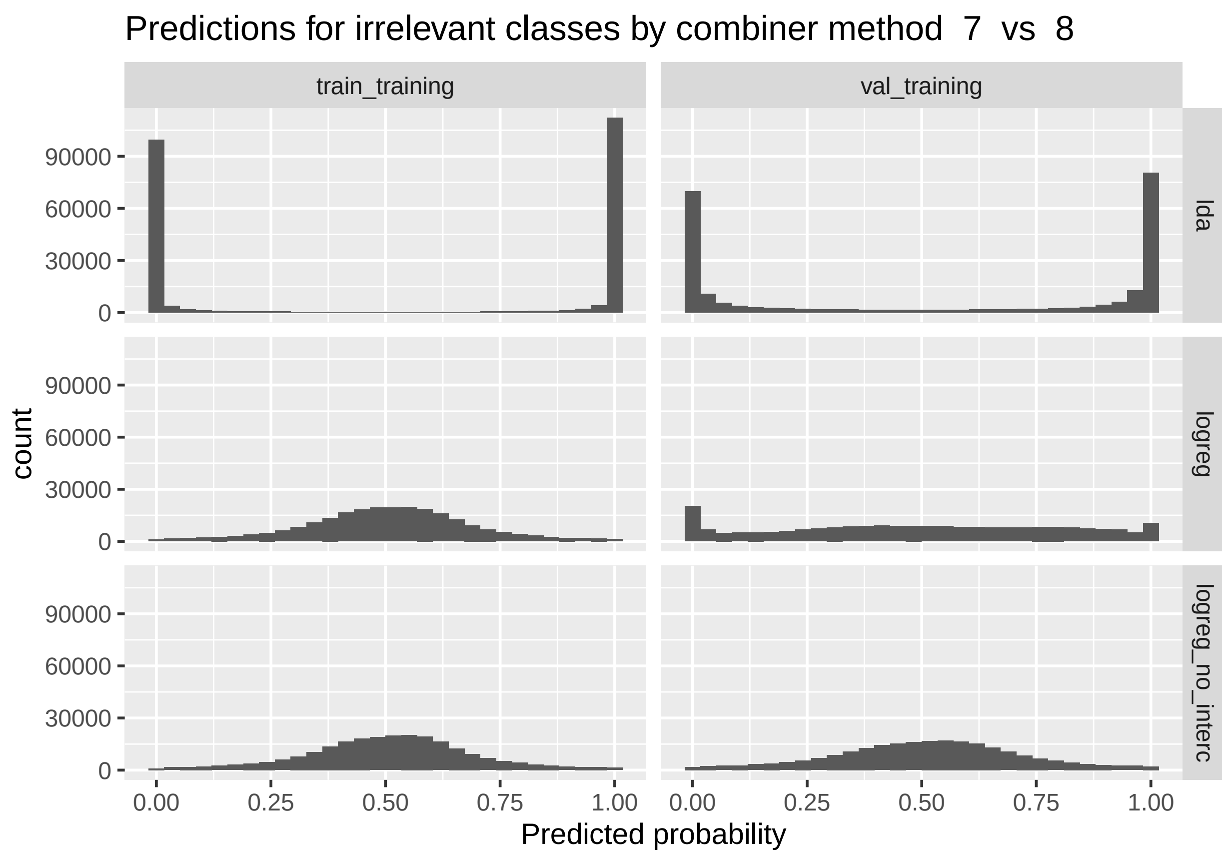


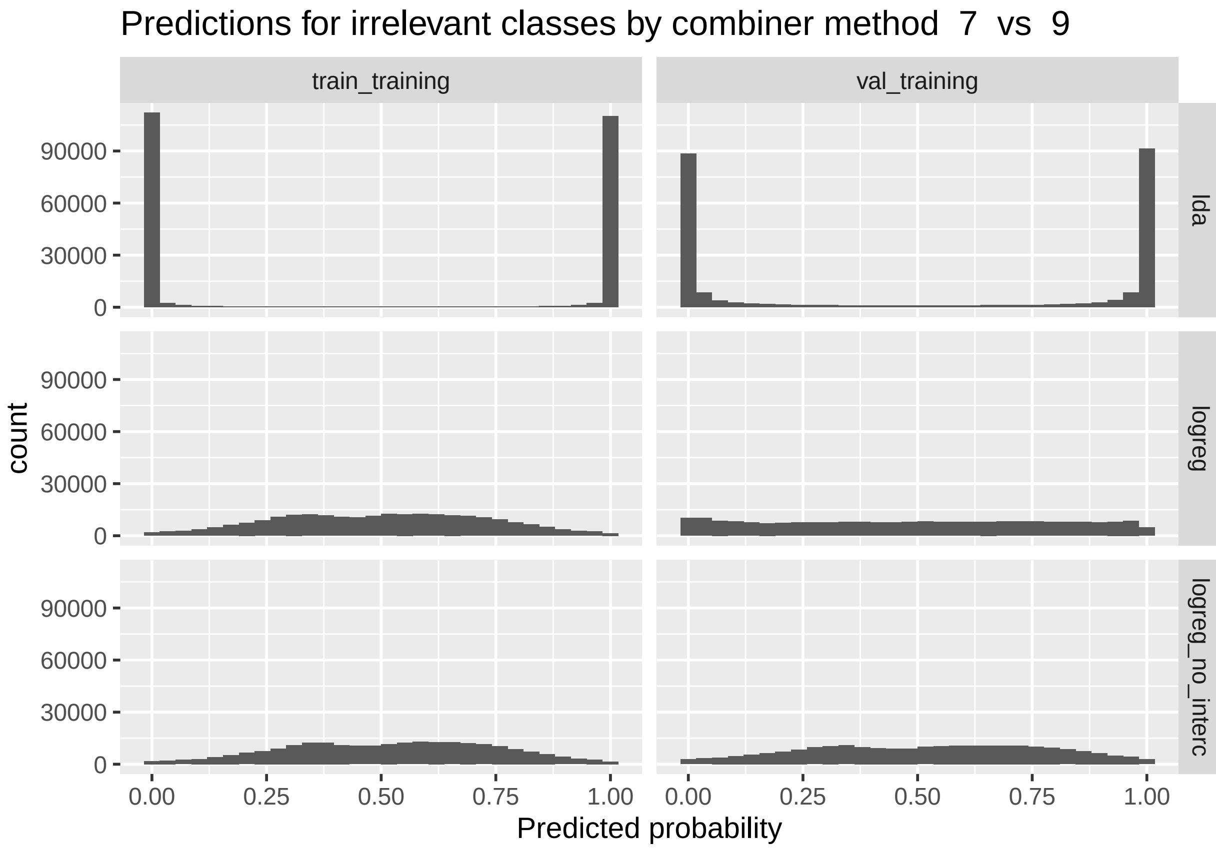
Predictions for irrelevant classes by combiner method 6 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
125000
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

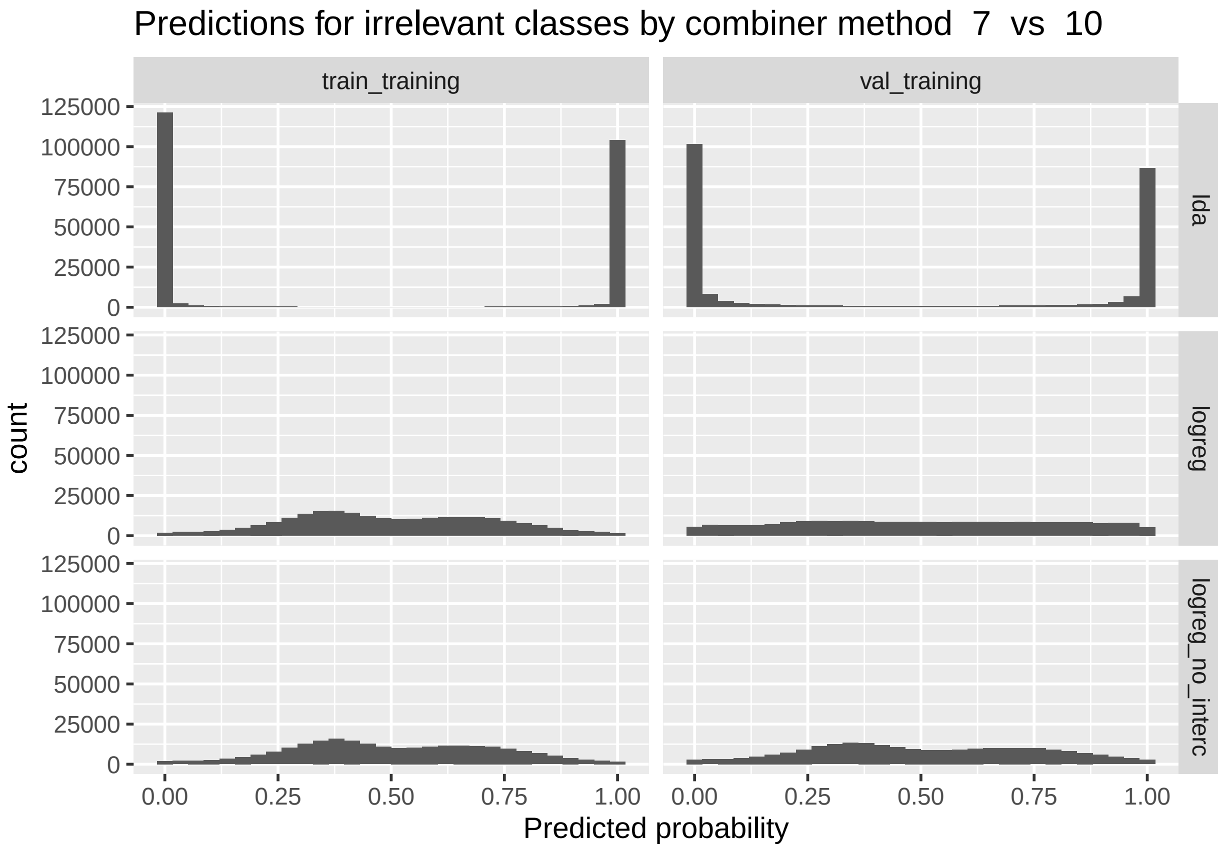


Predictions for irrelevant classes by combiner method 6 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

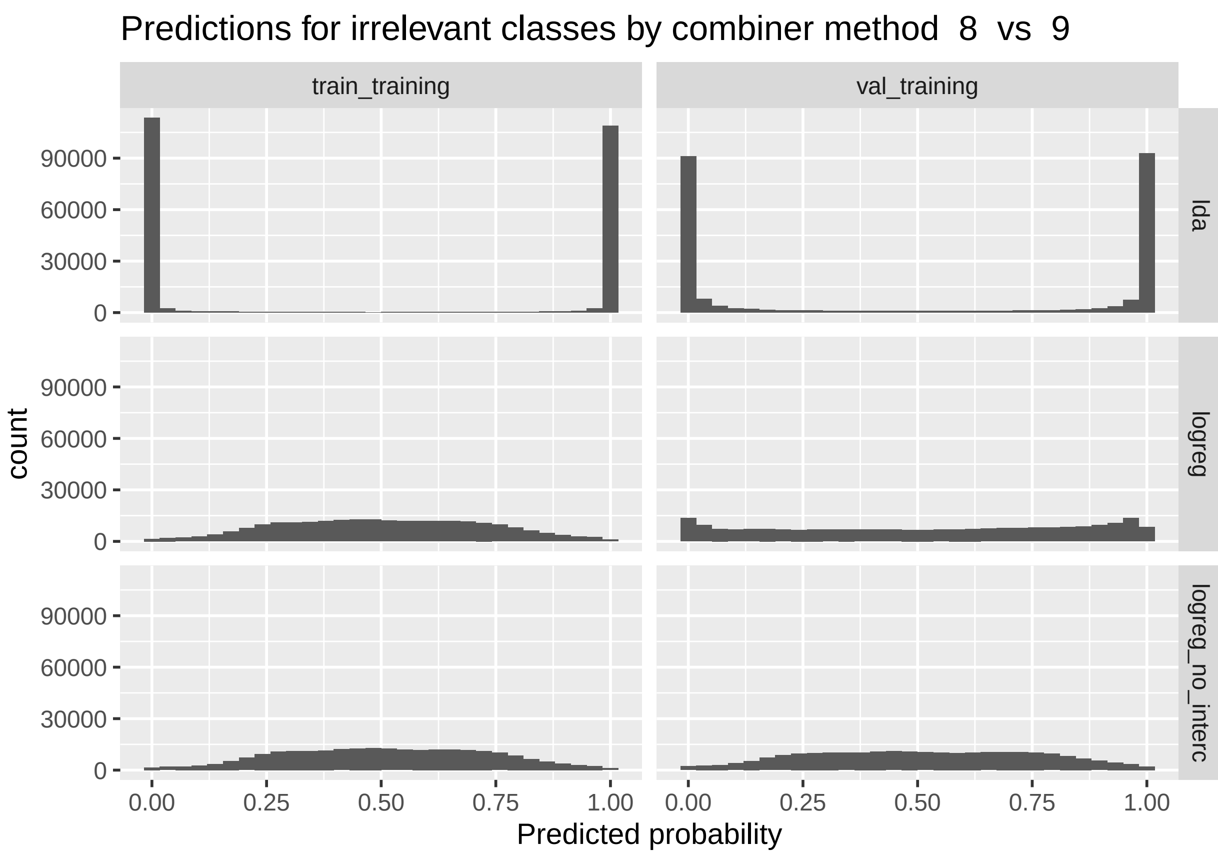

Predictions for irrelevant classes by combiner method 6 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
120000
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability


Predictions for irrelevant classes by combiner method 7 vs 8
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability

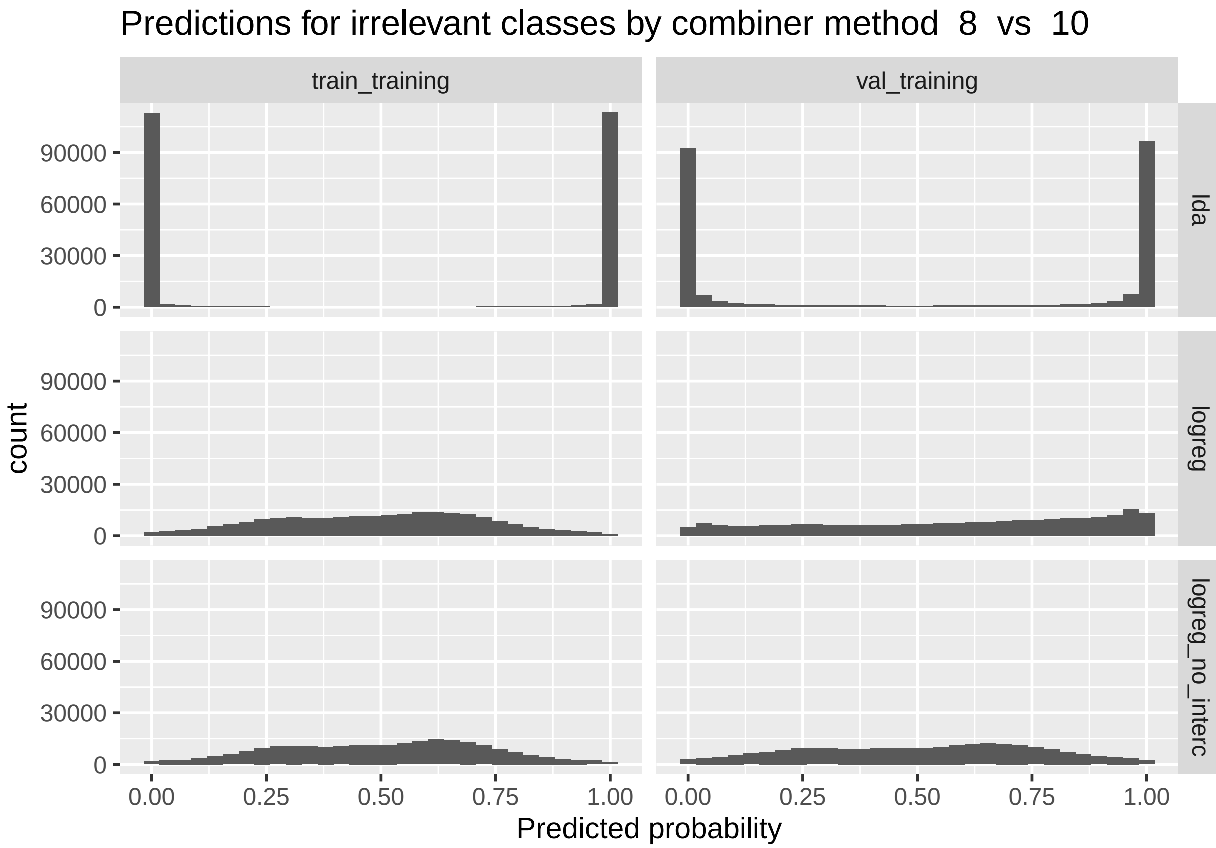

Predictions for irrelevant classes by combiner method 7 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



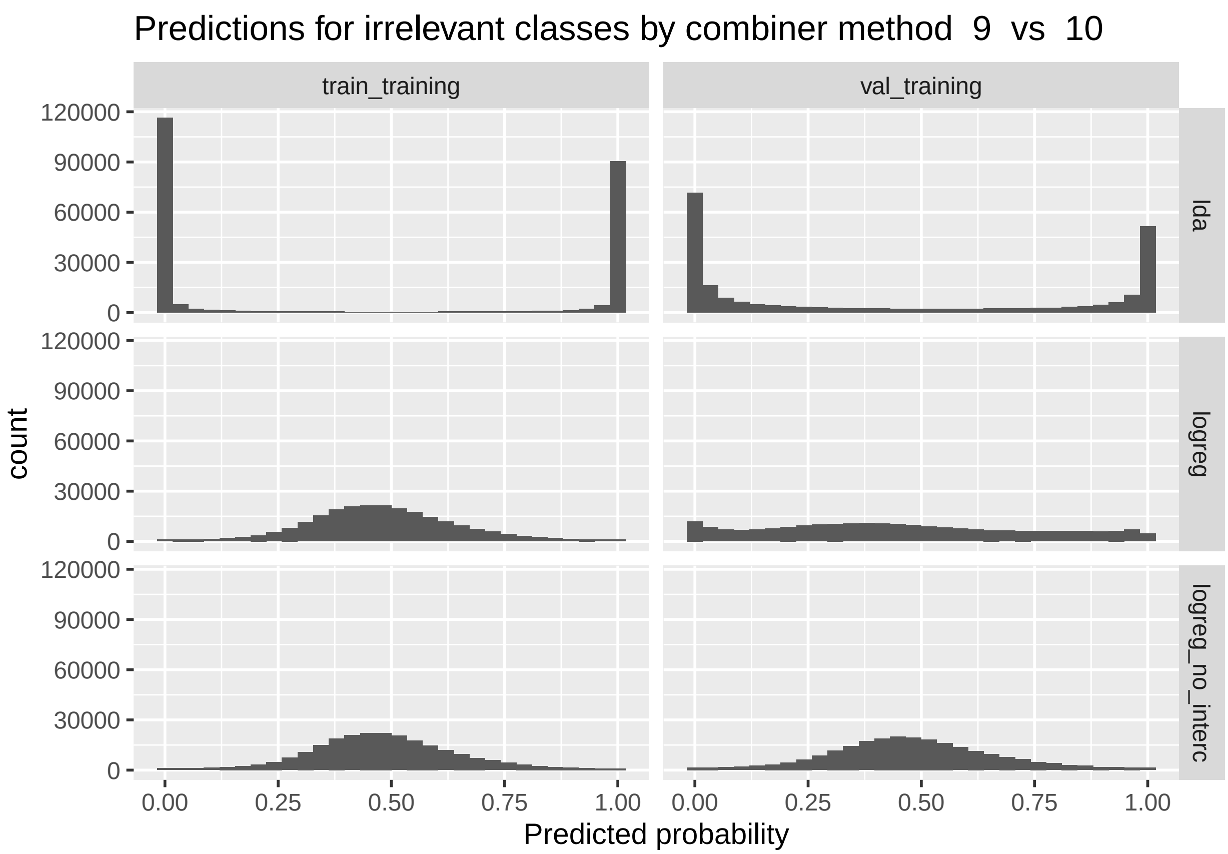
Predictions for irrelevant classes by combiner method 7 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
125000
100000
75000
50000
25000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



Predictions for irrelevant classes by combiner method 8 vs 9
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



Predictions for irrelevant classes by combiner method 8 vs 10
train_training
val_training
lda
logreg
logreg_no_interc
count
90000
60000
30000
0
0.00
0.25
0.50
0.75
1.00
Predicted probability



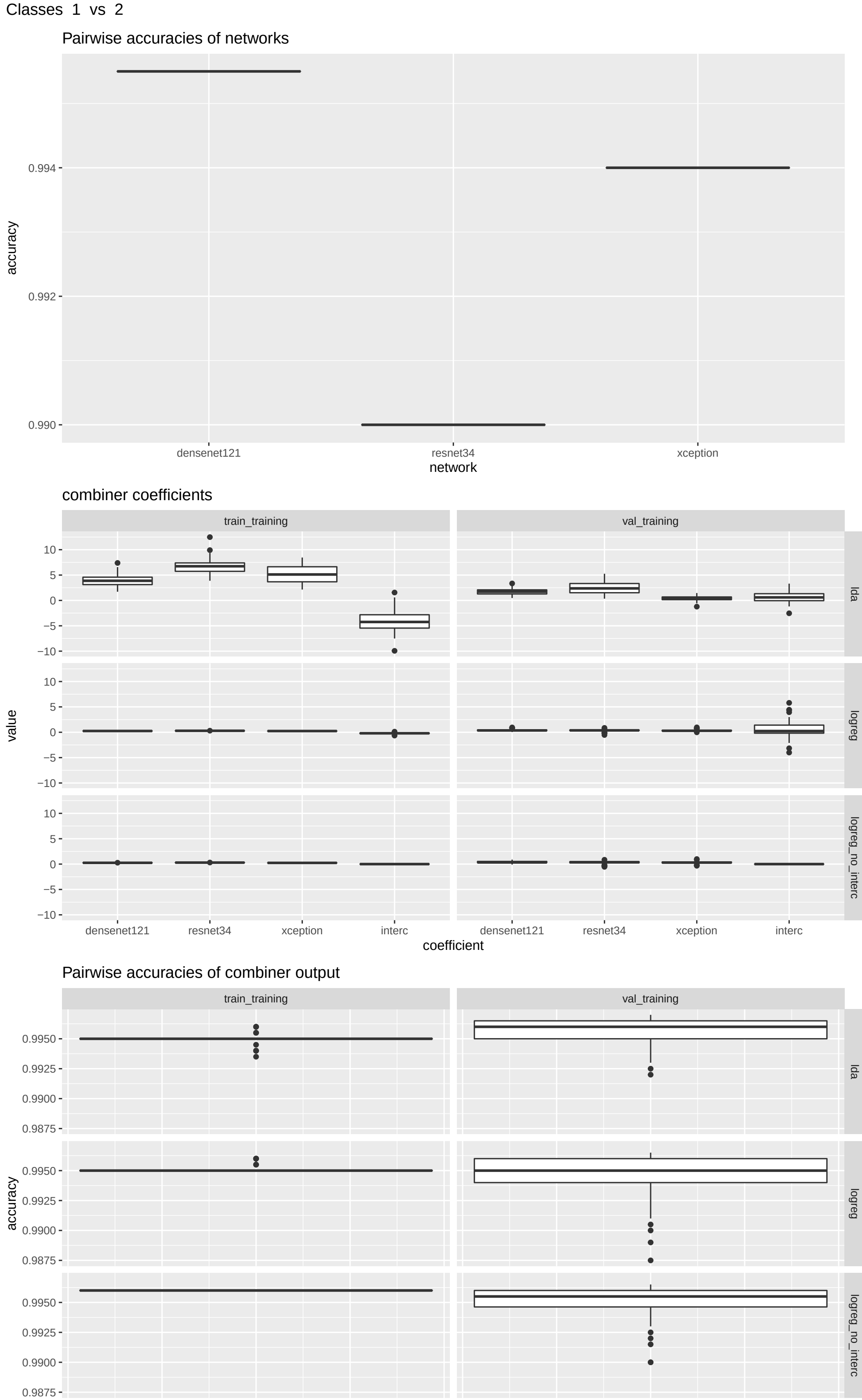
These histograms show, that even for unknown classes, LDA most often produces probabilities close to one and zero. One exception is val_train LDA for class pair 4 and 6, where probabilities are divided evenly. Also the confidence plot for this LDA looks sensibly. Histograms for method logreg, on the other hand produce more evenly distributed probabilities.

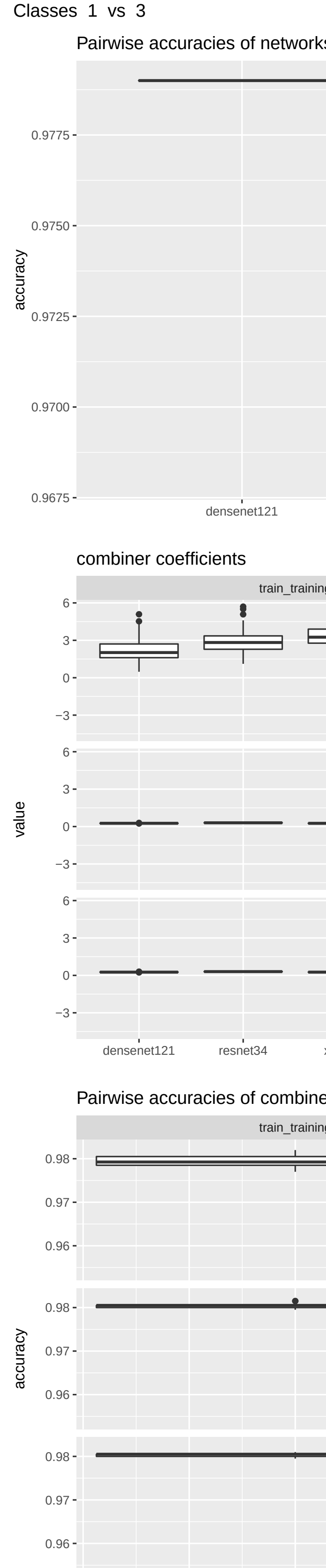
CIFAR 10 half

```
base_dir <- "../data/data_train_val_half_c10"
repls <- 0:0
folds <- 0:49
classes <- 10
combiner_coefs <- load_combiner_coefs(base_dir, repls, folds)
net_pw_results <- read.csv(file.path(base_dir, "net_pw_accuracies.csv"))
ens_pw_results <- read.csv(file.path(base_dir, "ensemble_pw_accuracies.csv"))
net_pw_results[, c("class1", "class2")] <- lapply(net_pw_results[, c("class1", "class2")], as.factor)
ens_pw_results[, c("class1", "class2", "combining_method")] <- lapply(ens_pw_results[, c("class1", "cla
```

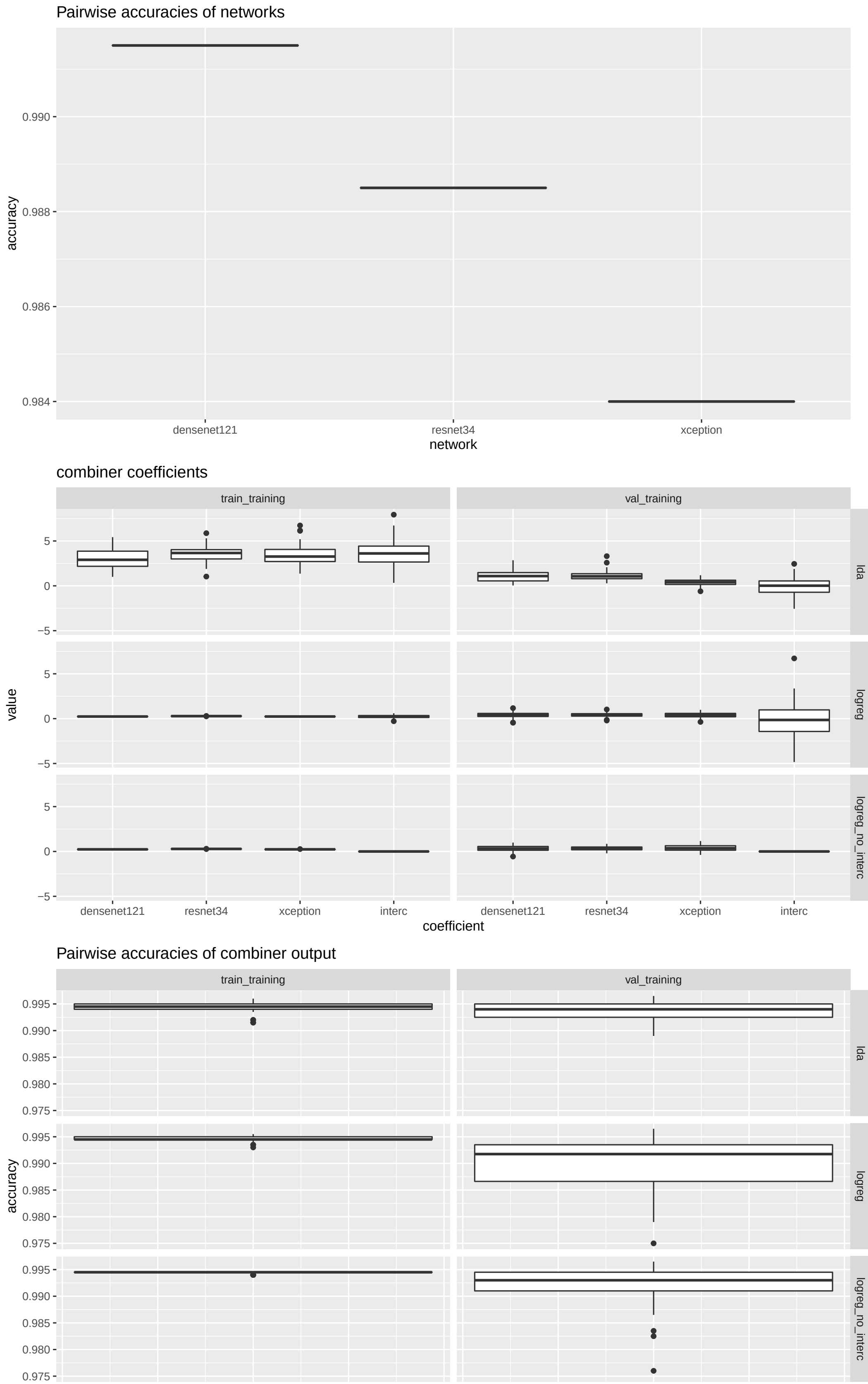
```
for (cl1 in 1:(classes - 1))
{
  for (cl2 in (cl1 + 1):classes)
  {
    combiner_plt <- combiner_coefs %>% filter(class1 == cl1 & class2 == cl2) %>% ggplot() +
      geom_boxplot(aes(x=coefficient, y=value)) +
      facet_grid(cols=vars(train_type), rows=vars(combining_method)) + ggtitle("combiner coefficients")
```

```
    acc_plt_net <- net_pw_results %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot(mappi
      geom_boxplot() + ggtitle("Pairwise accuracies of networks")

    acc_plt_ens <- ens_pw_results %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot(mappi
      geom_boxplot() + facet_grid(cols=vars(train_set), rows=vars(combining_method)) + ggtitle("Pairwise
      theme(axis.ticks.x=element_blank(), axis.text.x=element_blank())


    print((acc_plt_net/combiner_plt/acc_plt_ens) + plot_annotation(title=paste("Classes ", cl1, " vs ",
  }
}
```

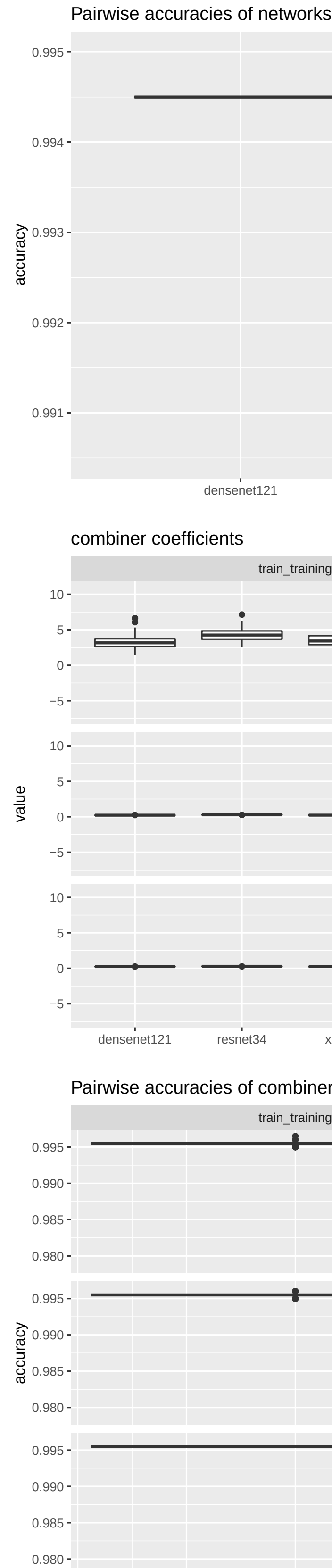

Classes 1 vs 2
Pairwise accuracies of networks
accuracy
0.994
0.992
0.990
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
10
5
0
−5
−10
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.9950
0.9925
0.9900
0.9875

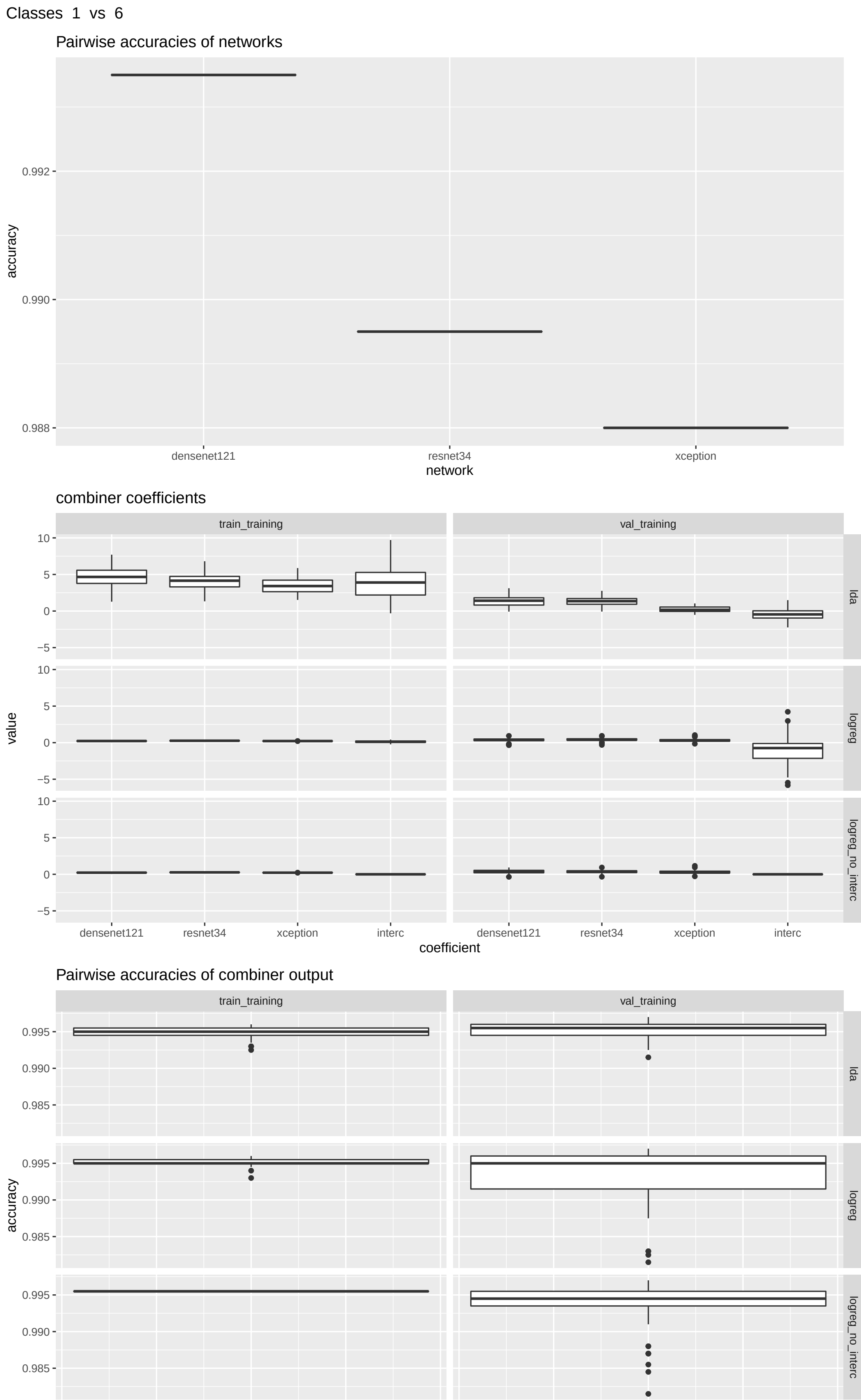

Classes 1 vs 3
Pairwise accuracies of networks
accuracy
0.9775
0.9750
0.9725
0.9700
0.9675
densenet121
combiner coefficients
value
6
3
0
−3
densenet121
resnet34
accuracy
0.98
0.97
0.96

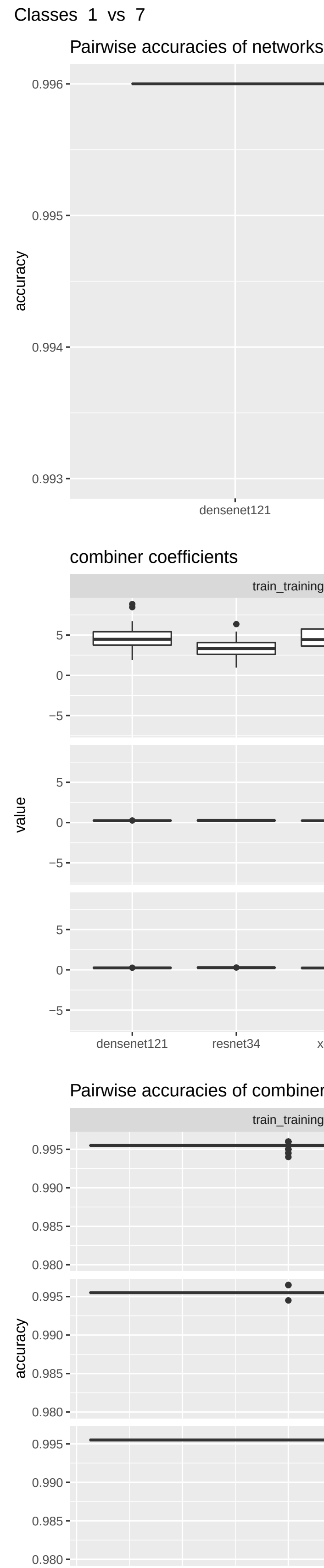
Classes 1 vs 4



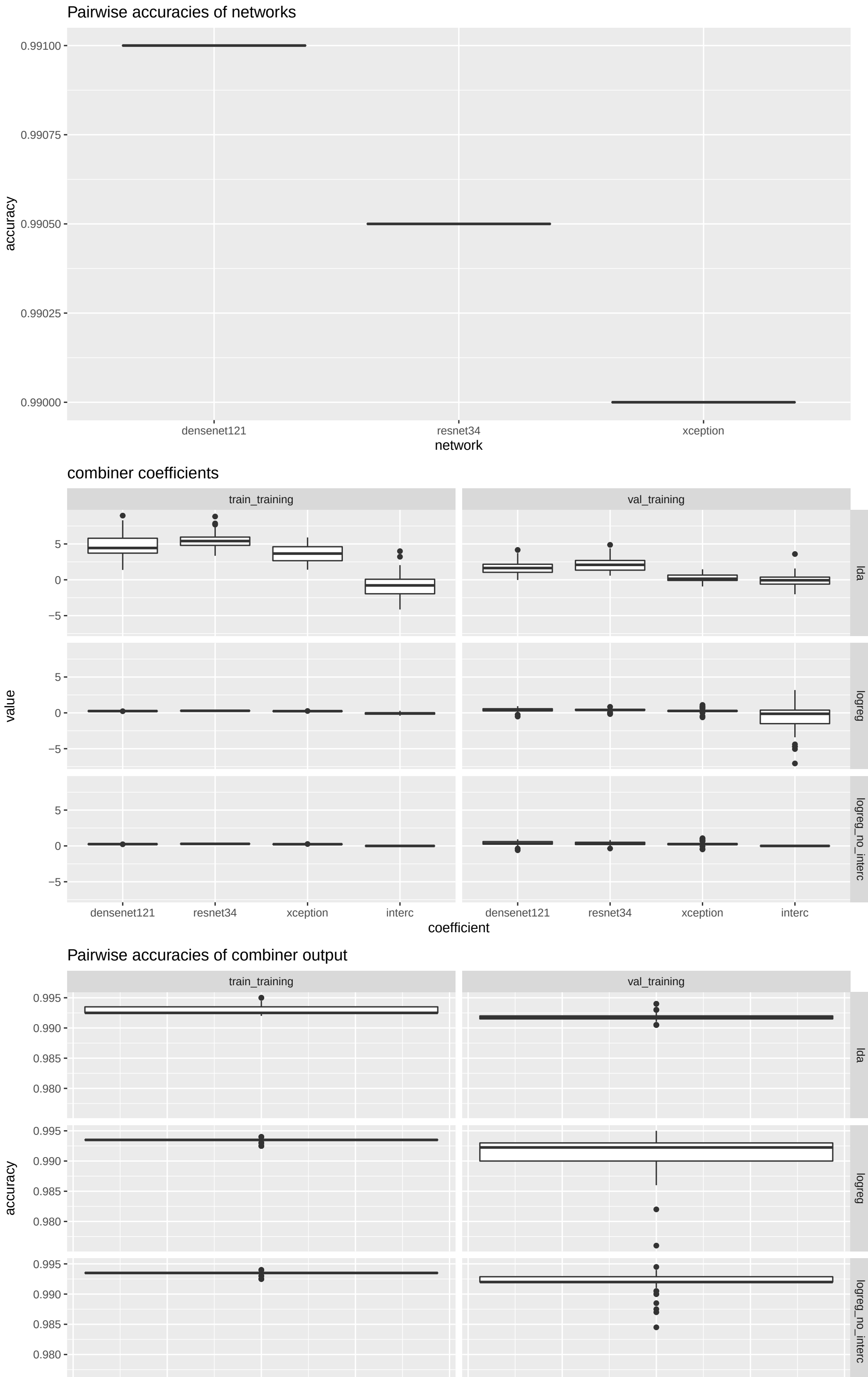
Classes 1 vs 5



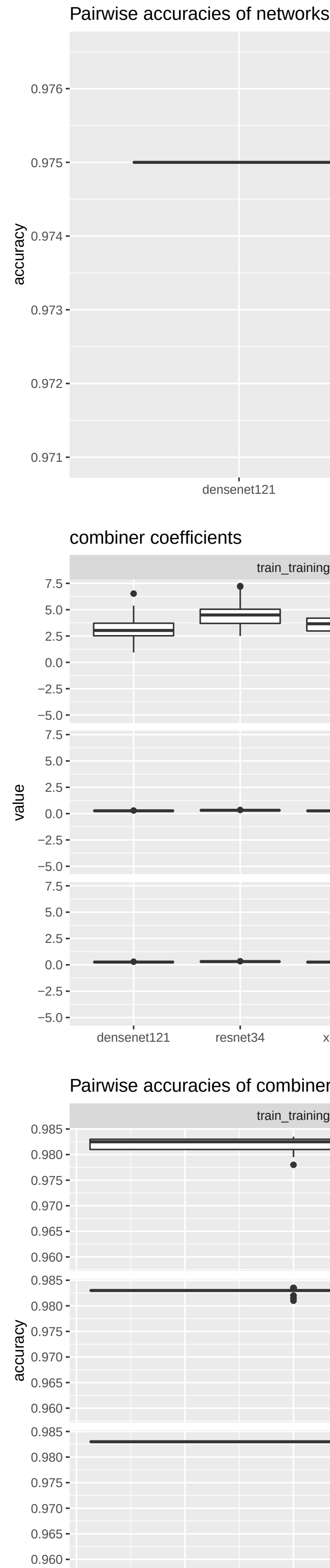

Classes 1 vs 6
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output

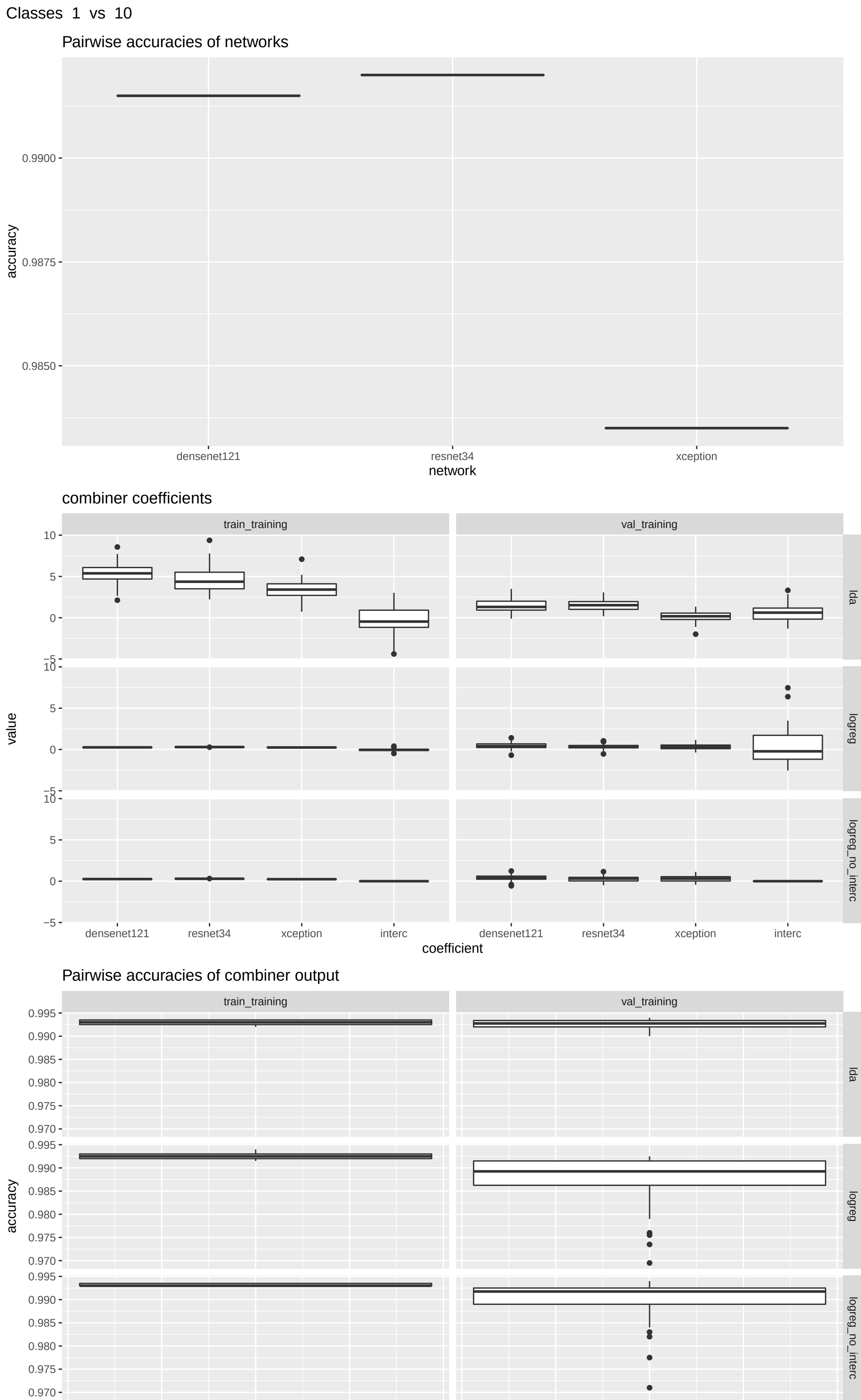

Classes 1 vs 7
Pairwise accuracies of networks
accuracy
densenet121
combiner coefficients
train_training
value
resnet34

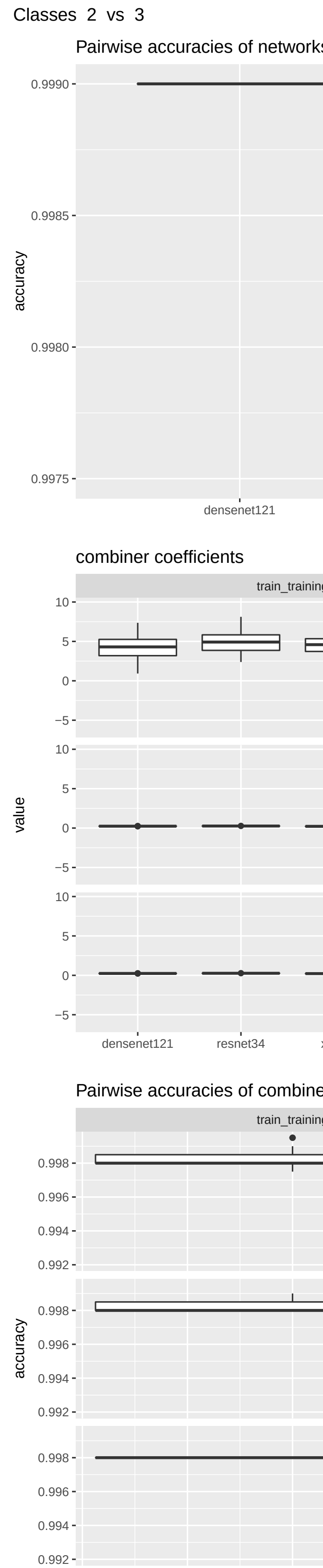
Classes 1 vs 8



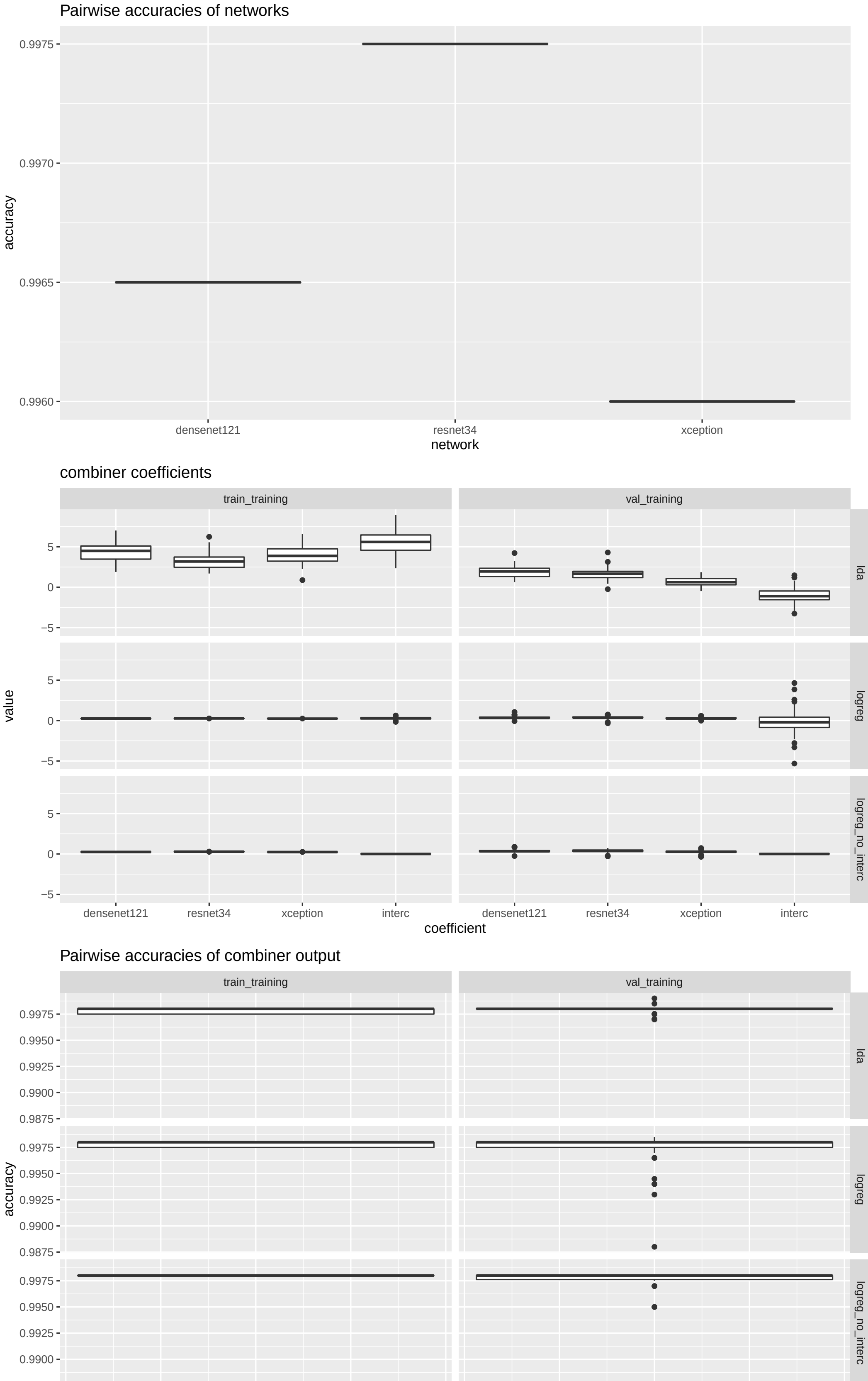
Classes 1 vs 9



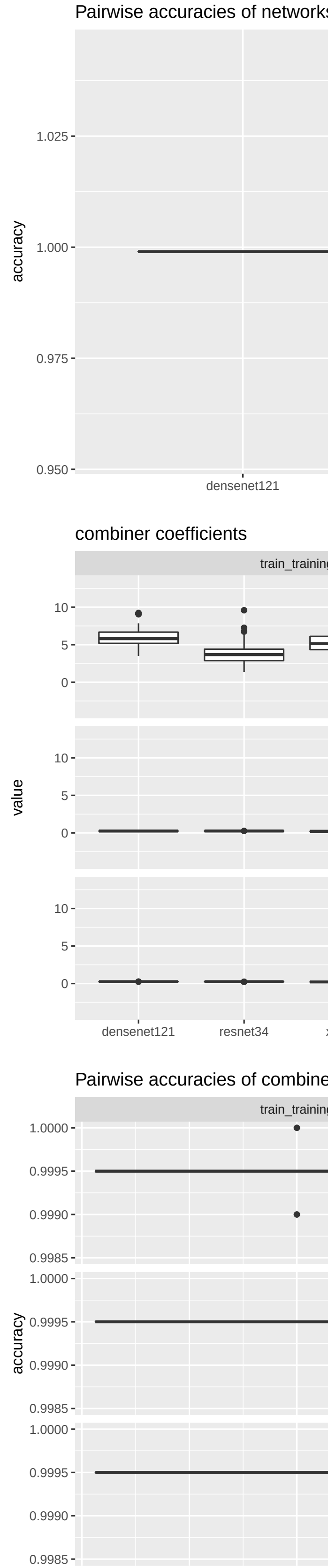

Classes 1 vs 10
Pairwise accuracies of networks
accuracy
0.9900
0.9875
0.9850
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
10
5
0
−5
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy
0.995
0.990
0.985
0.980
0.975
0.970


Classes 2 vs 3
Pairwise accuracies of networks
accuracy
0.9990
0.9985
0.9980
0.9975
densenet121
combiner coefficients
value
10
5
0
−5
densenet121
resnet34
accuracy
0.998
0.996
0.994
0.992

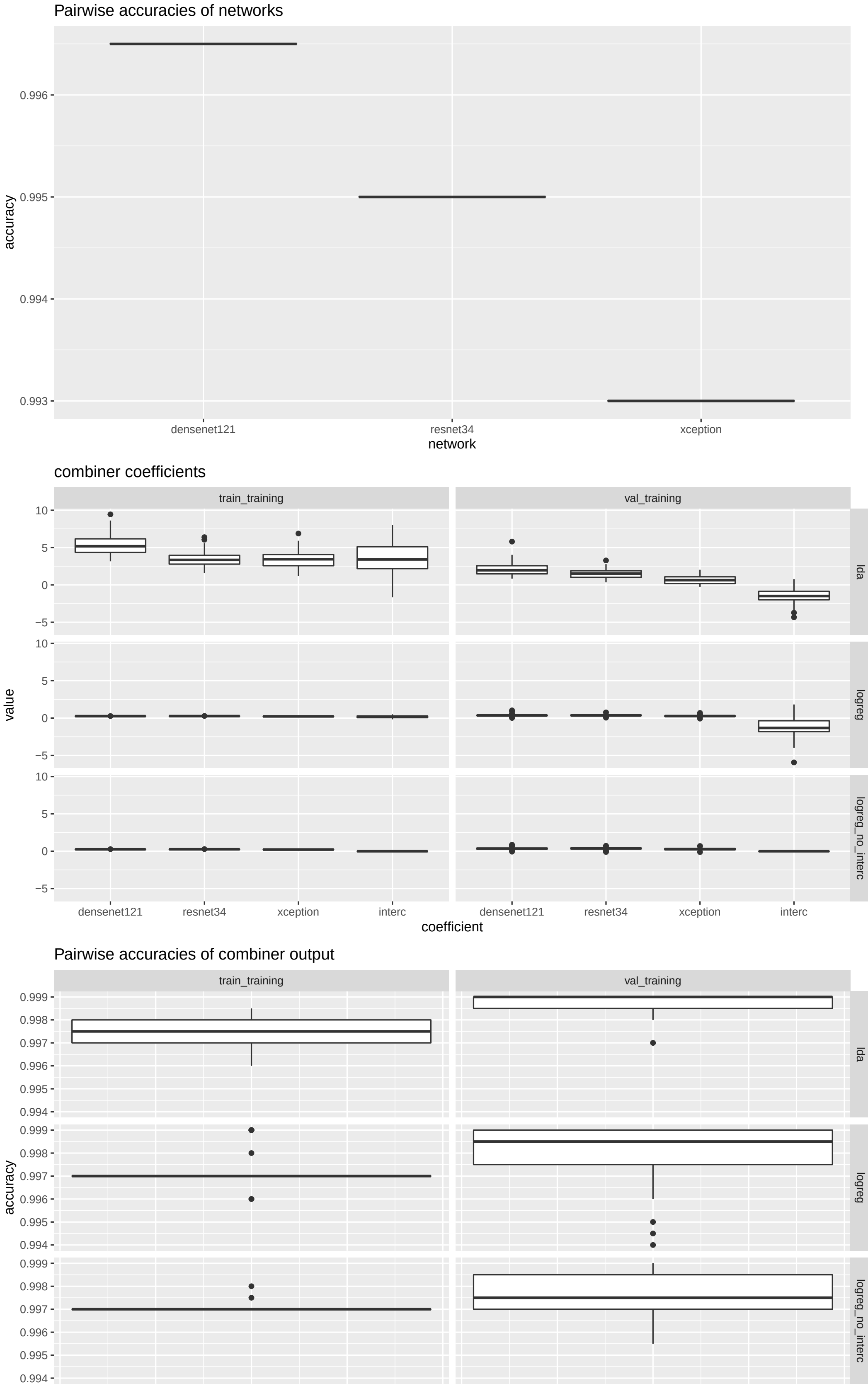
Classes 2 vs 4



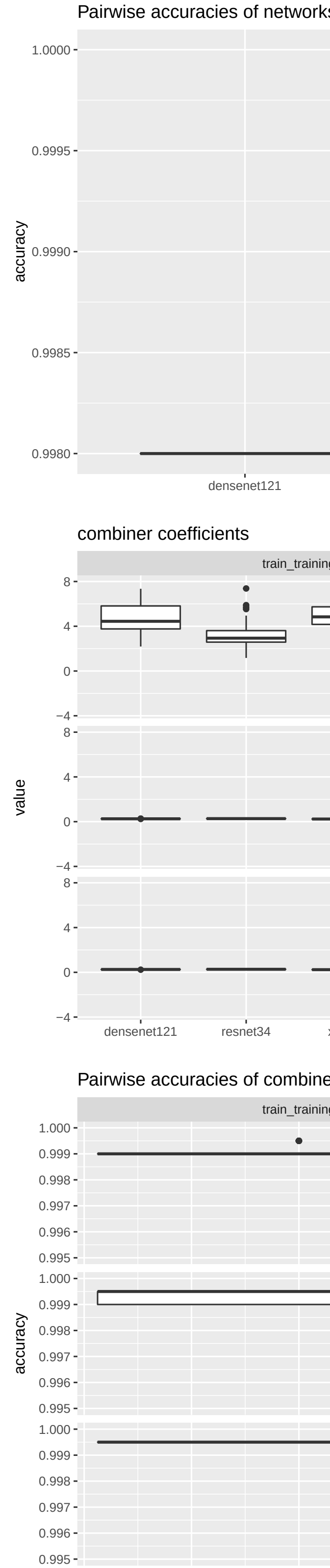
Classes 2 vs 5



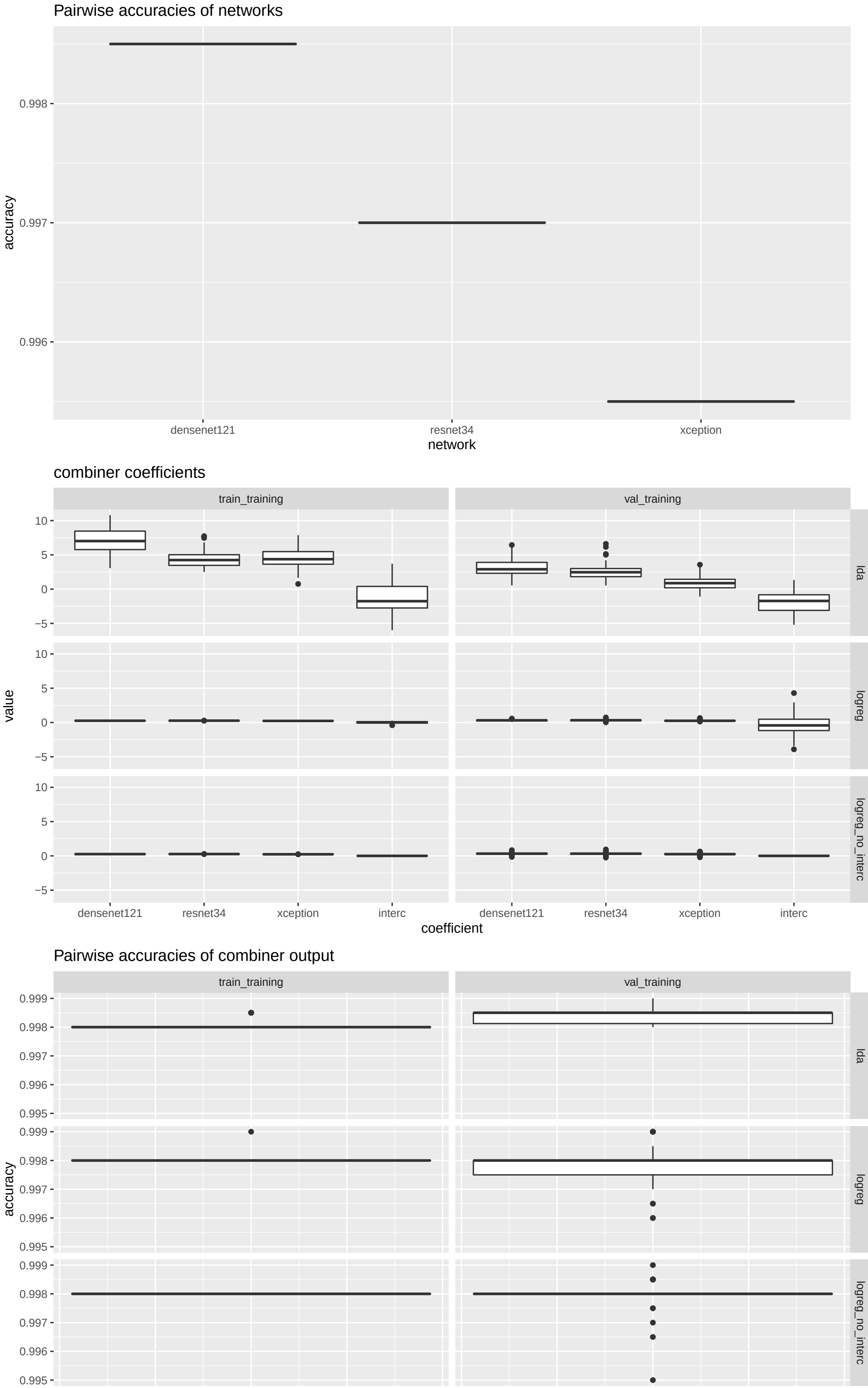
Classes 2 vs 6



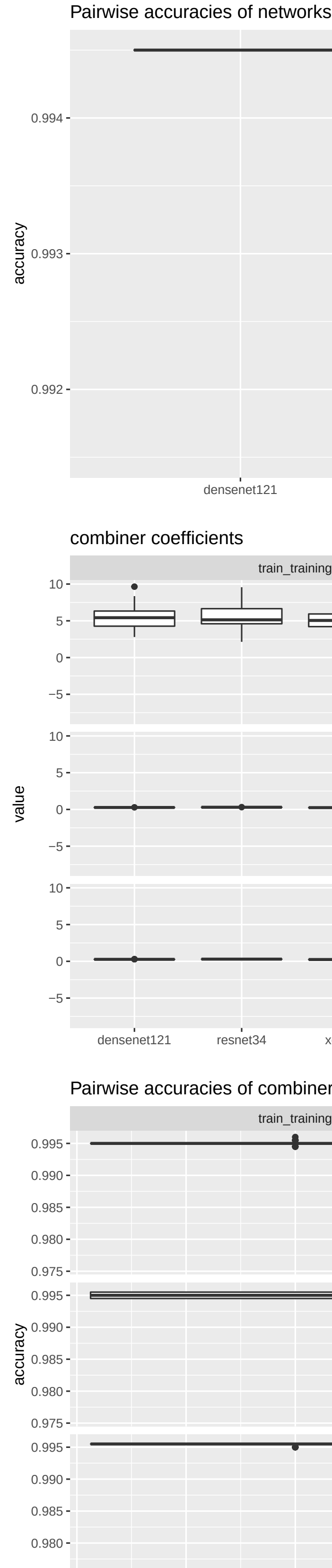
Classes 2 vs 7



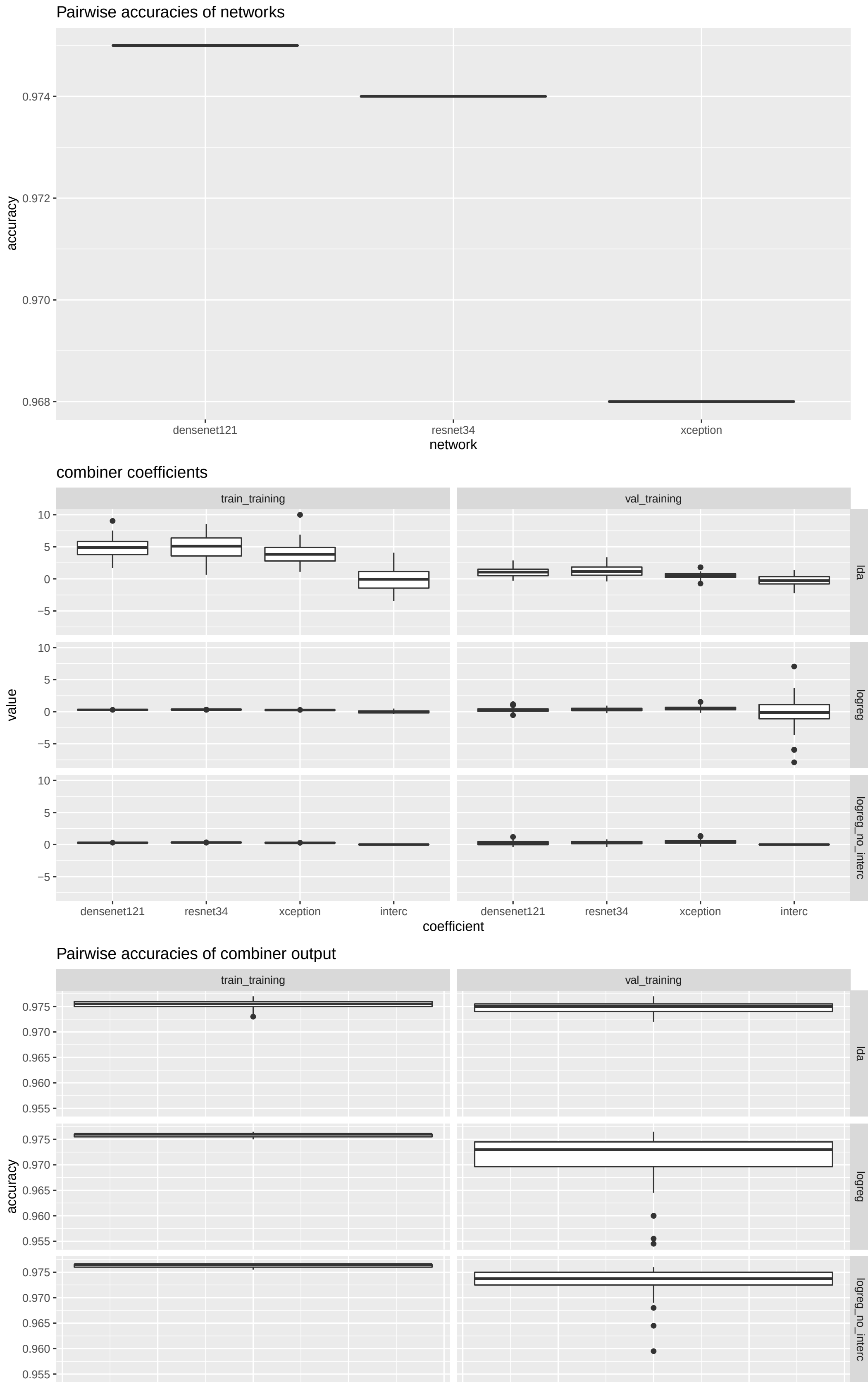
Classes 2 vs 8



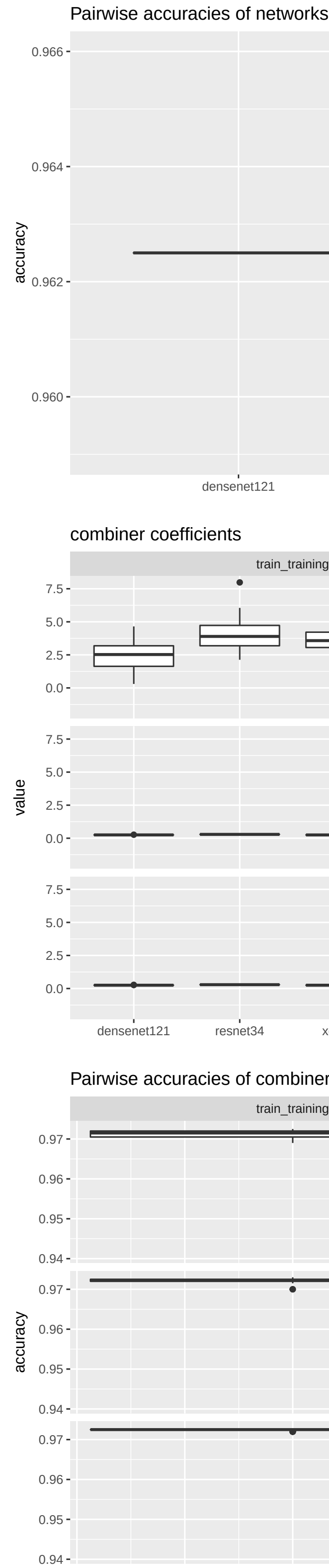
Classes 2 vs 9



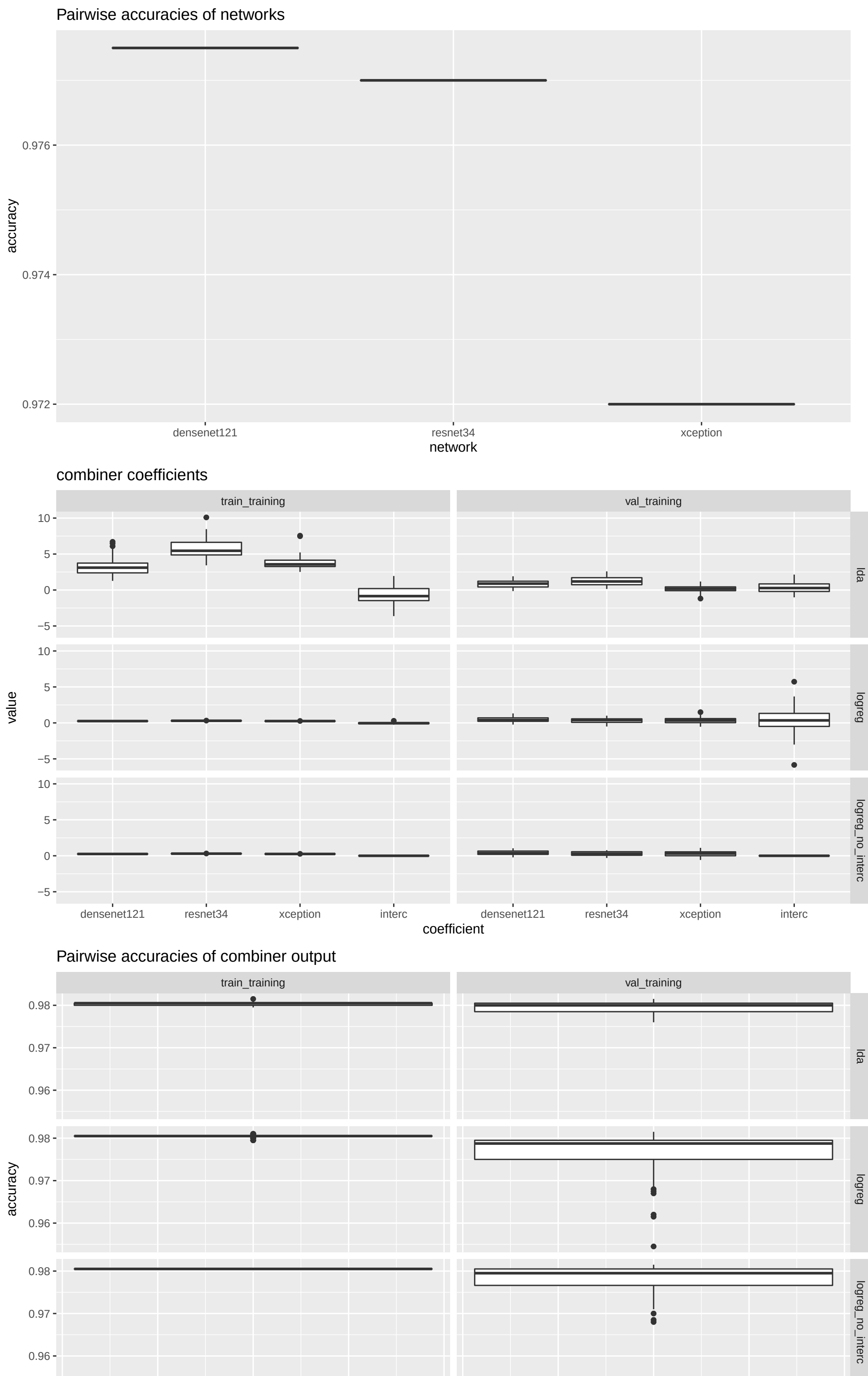
Classes 2 vs 10



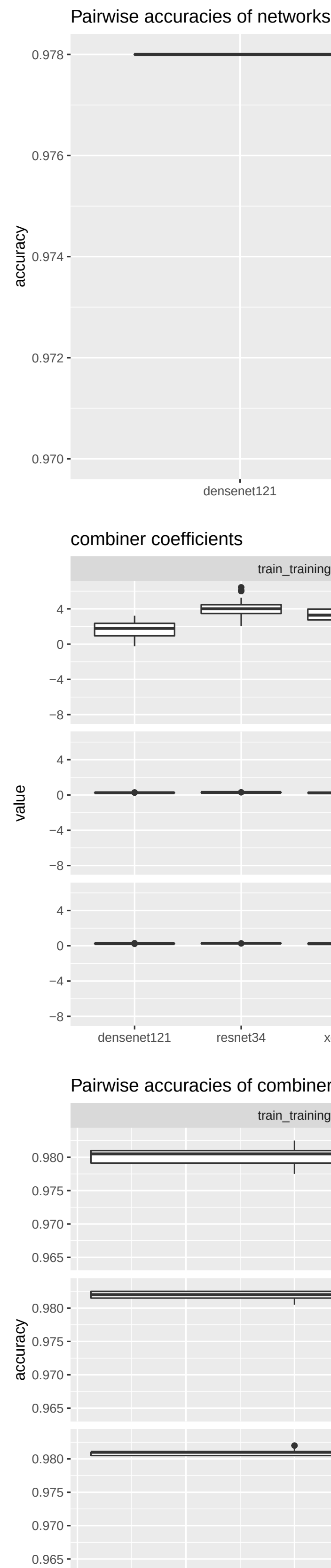
Classes 3 vs 4

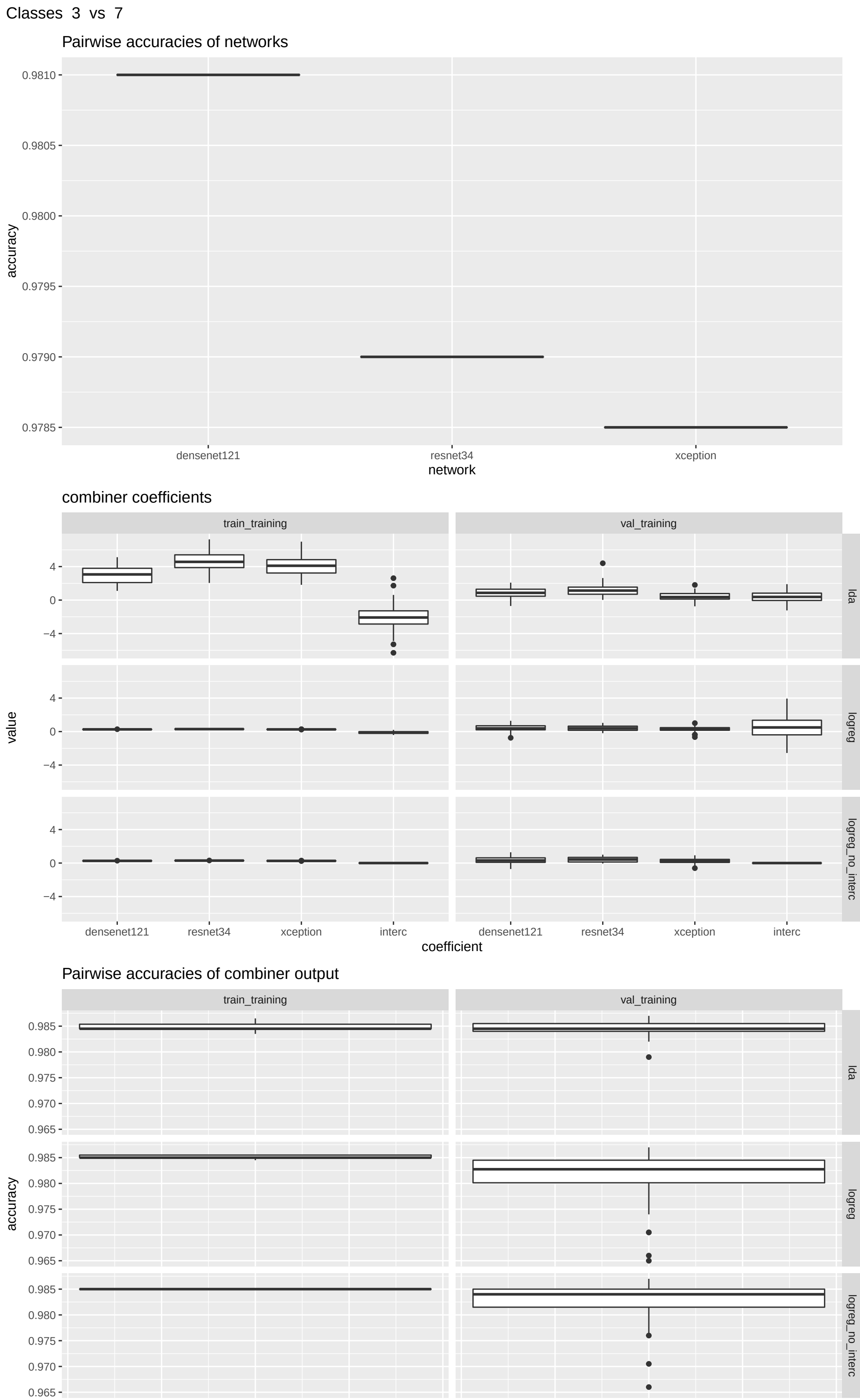


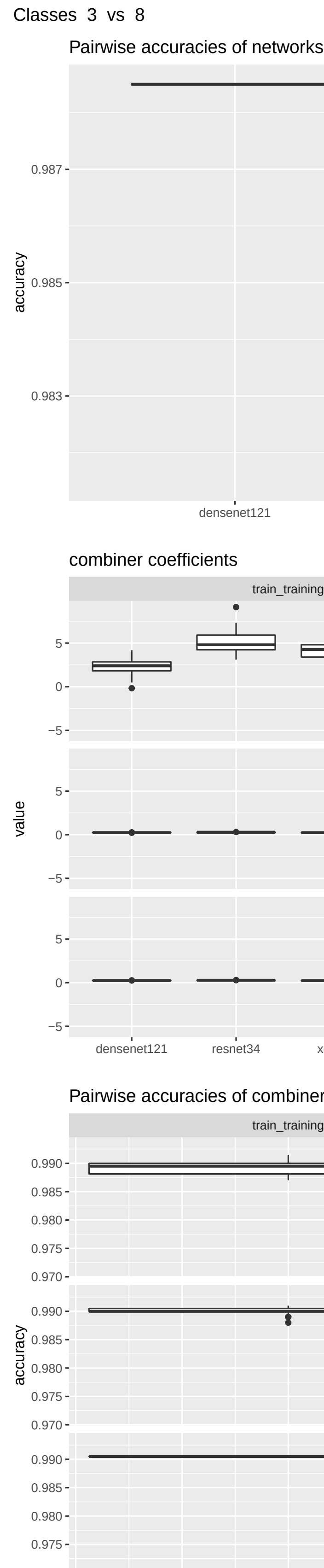
Classes 3 vs 5

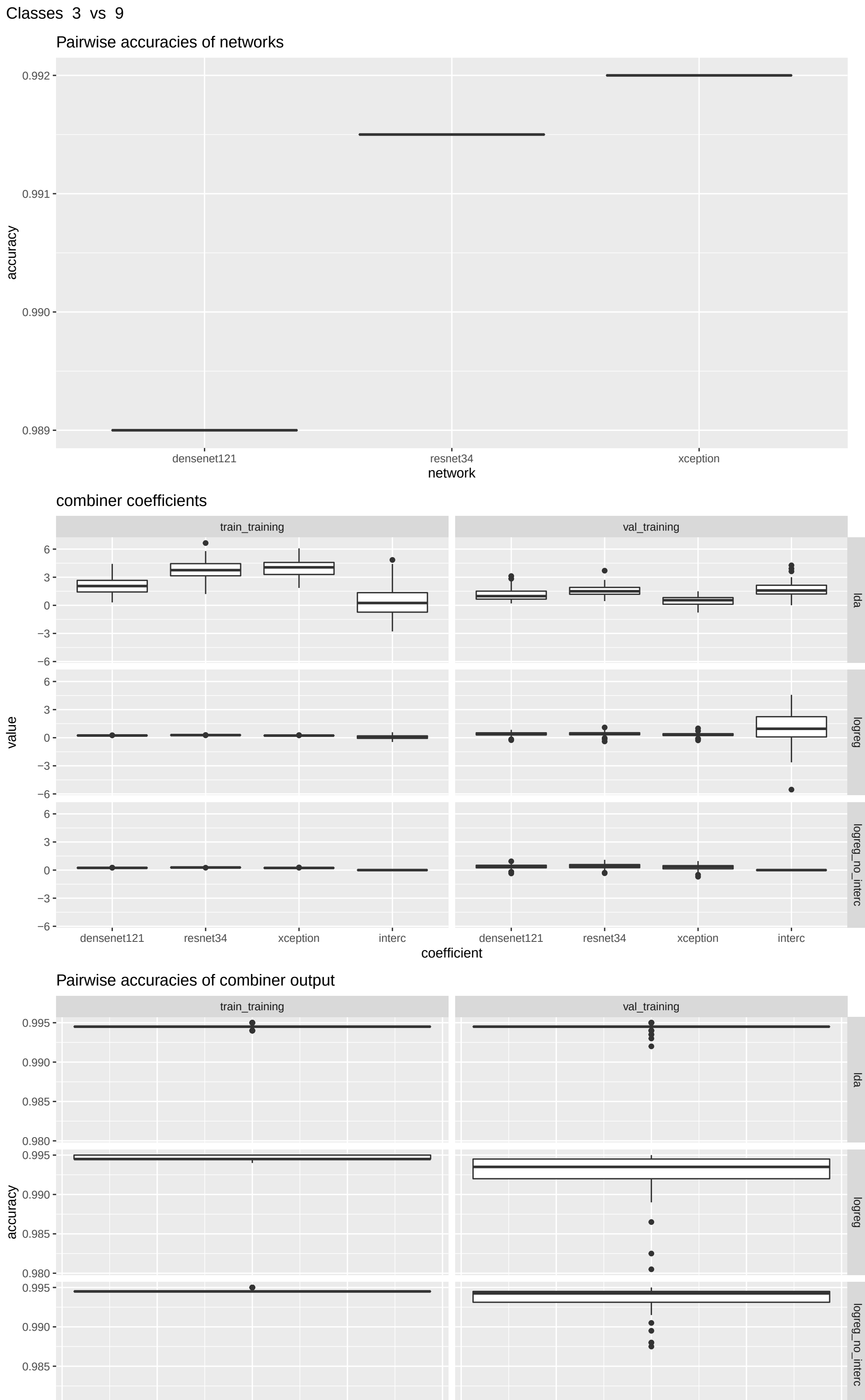


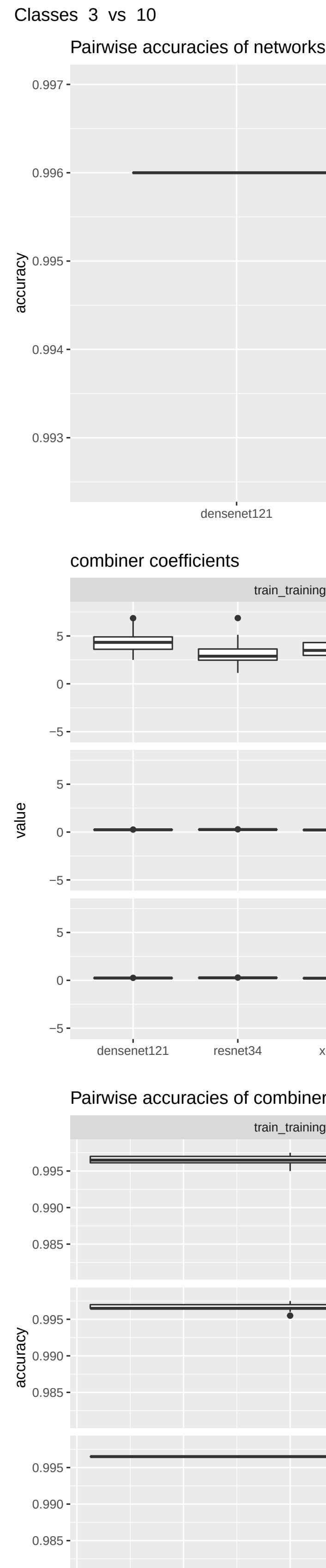
Classes 3 vs 6

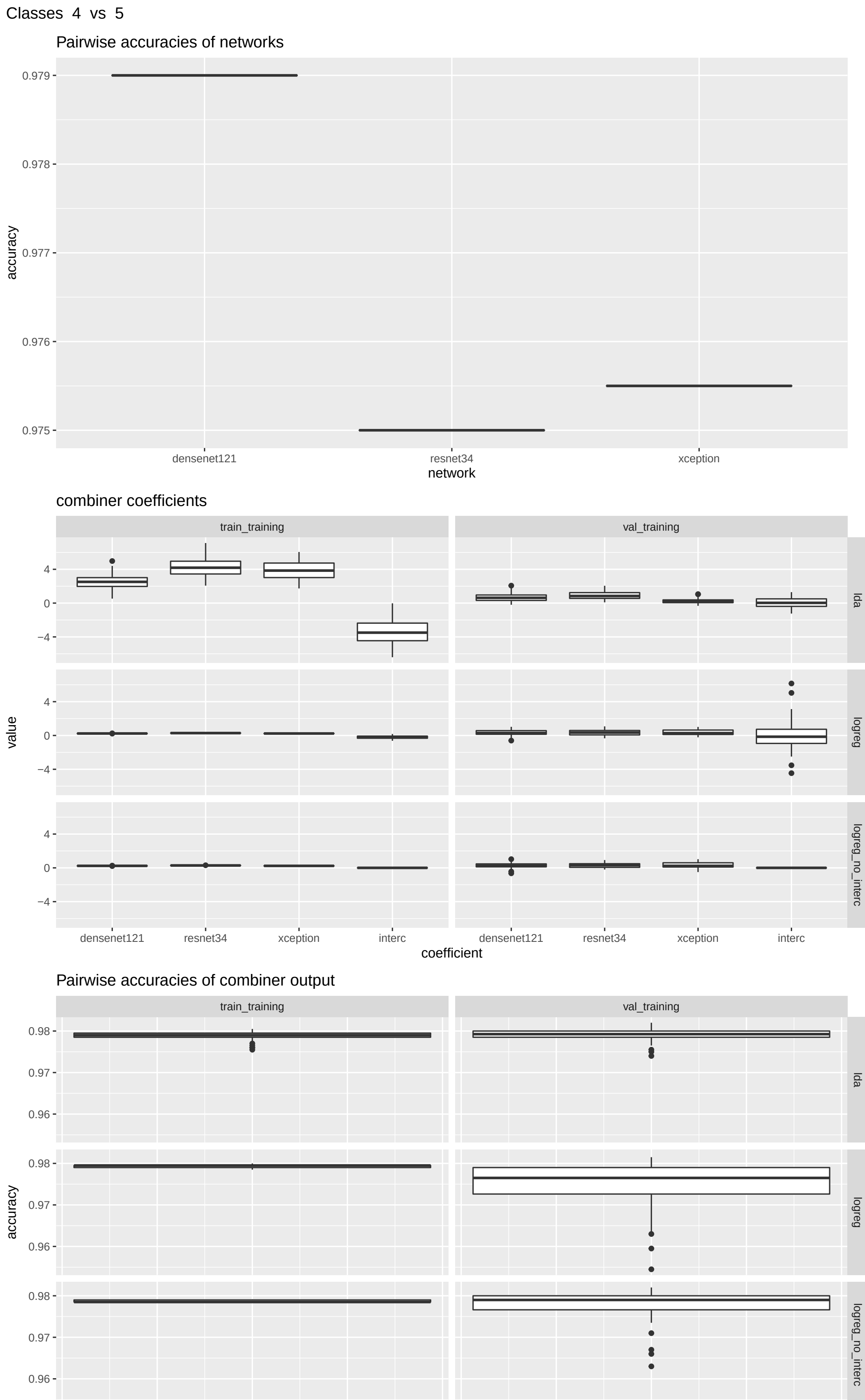


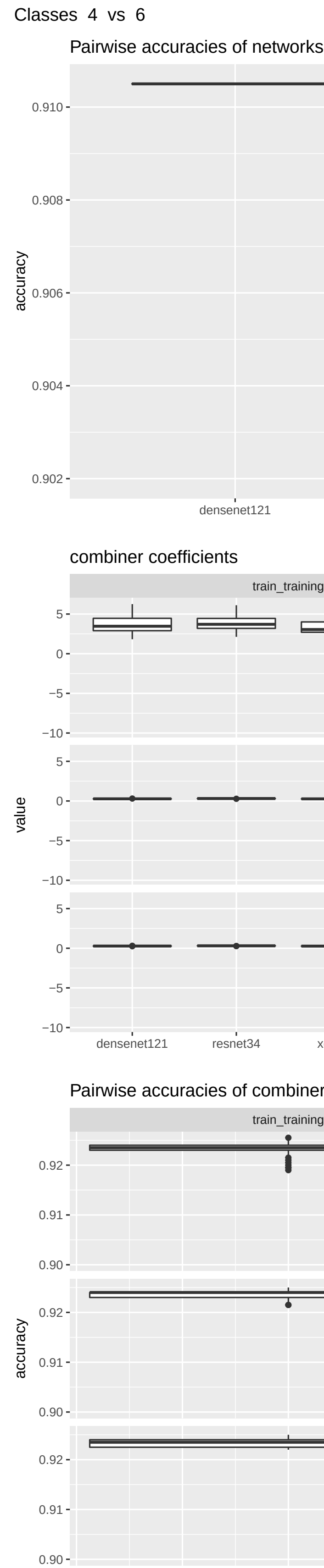

Classes 3 vs 7
Pairwise accuracies of networks
accuracy
0.9810
0.9805
0.9800
0.9795
0.9790
0.9785
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
4
0
−4
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.985
0.980
0.975
0.970
0.965

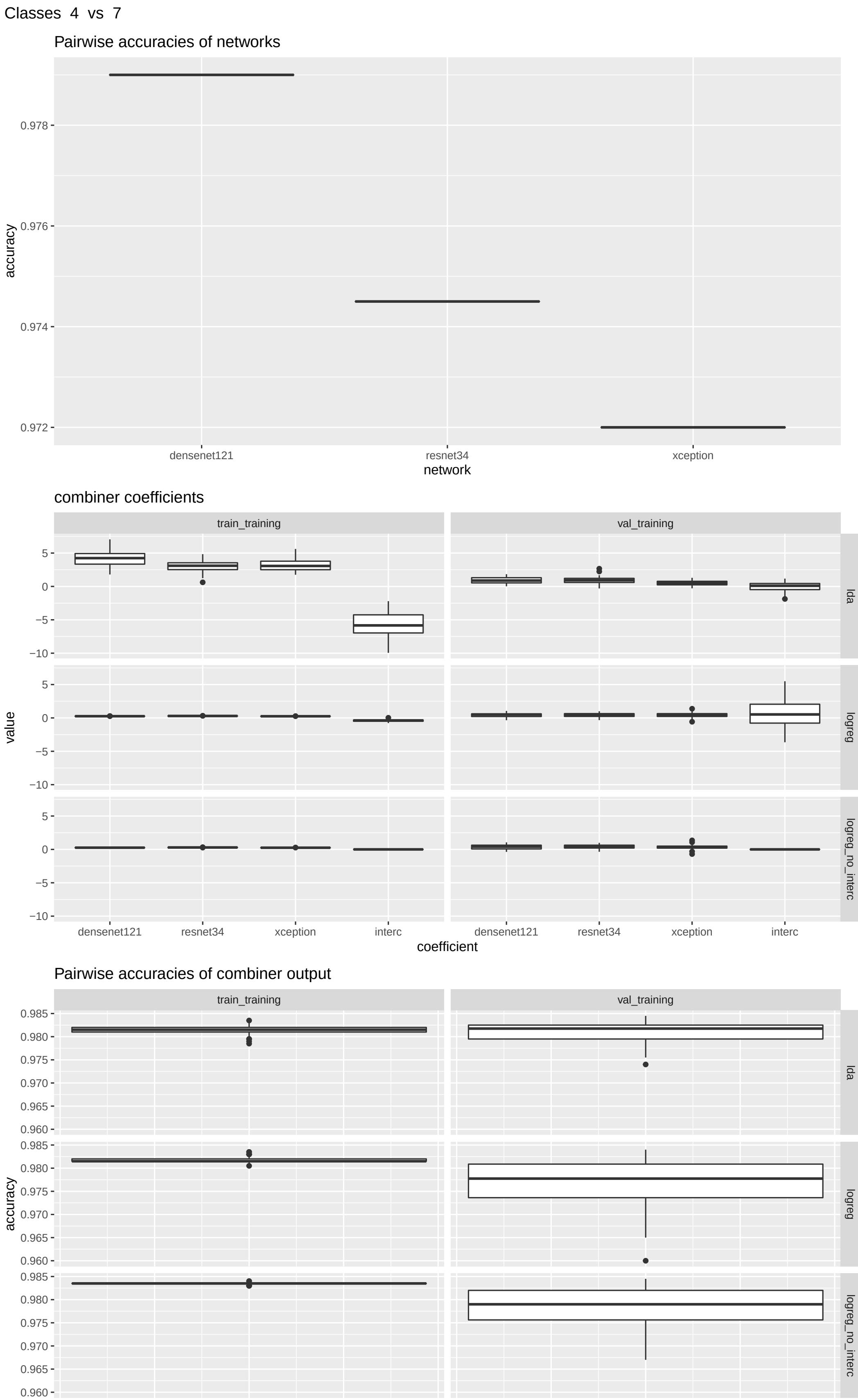

Classes 3 vs 8
Pairwise accuracies of networks
accuracy
0.987
0.985
0.983
densenet121
combiner coefficients
train_training
value
5
0
−5
densenet121
resnet34
Pairwise accuracies of combiner
0.990
0.985
0.980
0.975
0.970

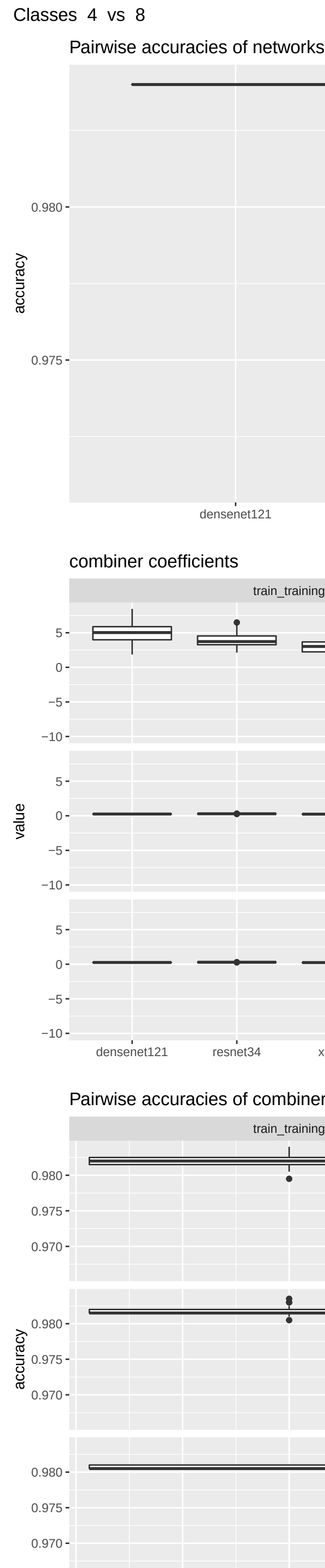

Classes 3 vs 9
Pairwise accuracies of networks
accuracy
0.992
0.991
0.990
0.989
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
6
3
0
−3
−6
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy
0.995
0.990
0.985
0.980


Classes 3 vs 10
Pairwise accuracies of networks
accuracy
0.997
0.996
0.995
0.994
0.993
densenet121
combiner coefficients
train_training
value
5
0
−5
densenet121
resnet34
Pairwise accuracies of combiner
train_training
accuracy
0.995
0.990
0.985

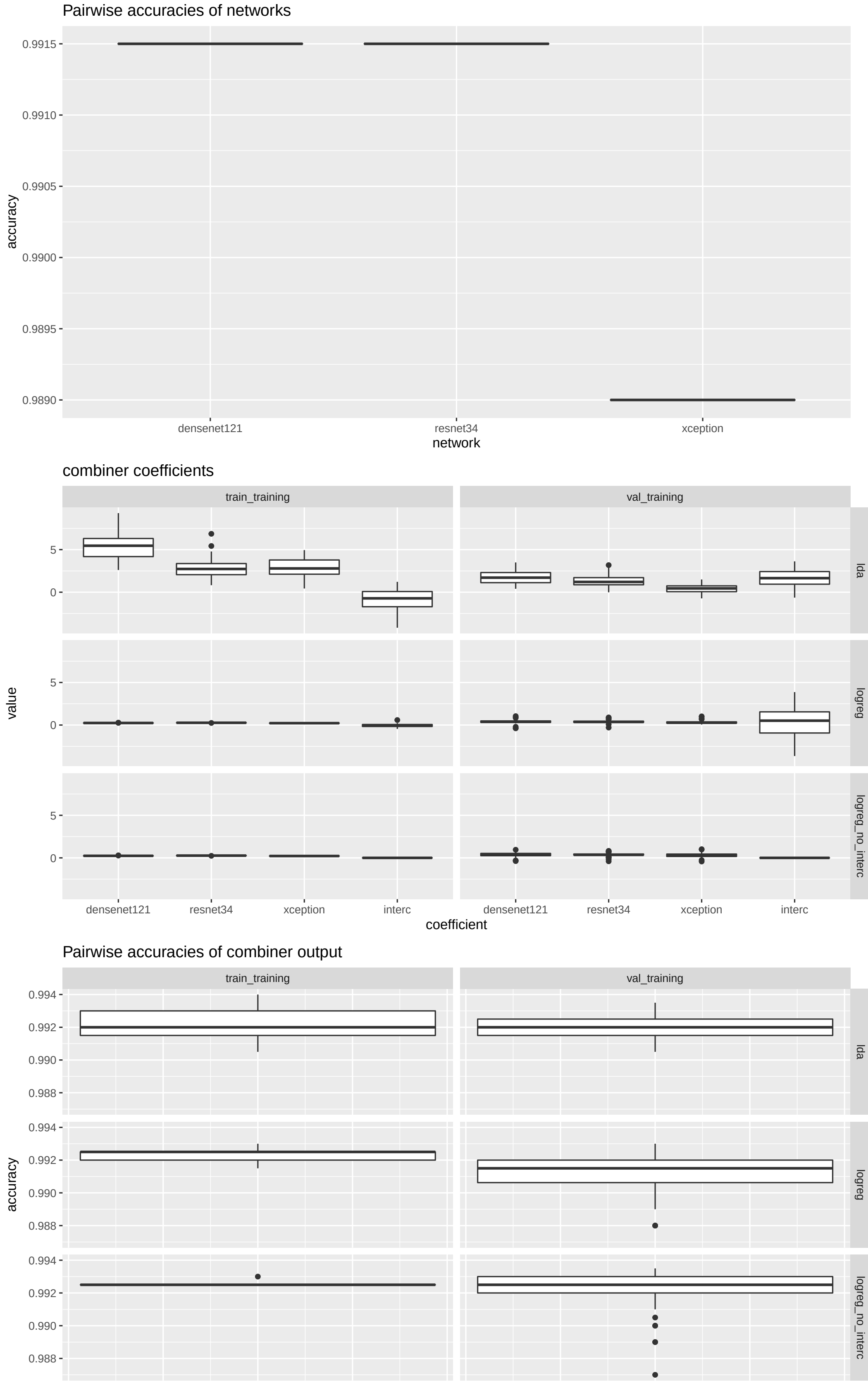

Classes 4 vs 5
Pairwise accuracies of networks
accuracy
0.979
0.978
0.977
0.976
0.975
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
4
0
−4
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.98
0.97
0.96


Classes 4 vs 6
Pairwise accuracies of networks
accuracy
0.910
0.908
0.906
0.904
0.902
densenet121
combiner coefficients
train_training
value
5
0
−5
−10
densenet121
resnet34
Pairwise accuracies of combiner
0.92
0.91
0.90

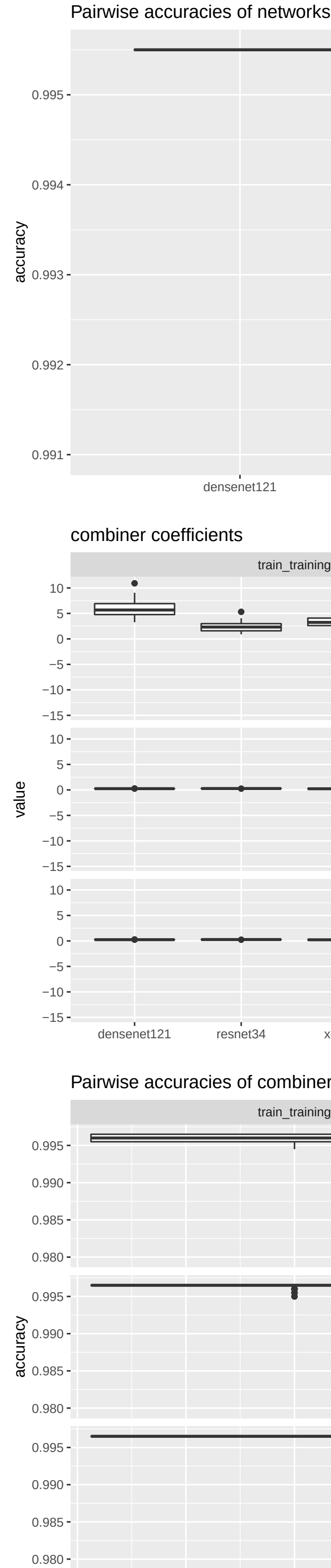

Classes 4 vs 7
Pairwise accuracies of networks
accuracy
0.978
0.976
0.974
0.972
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
5
0
−5
−10
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy
0.985
0.980
0.975
0.970
0.965
0.960

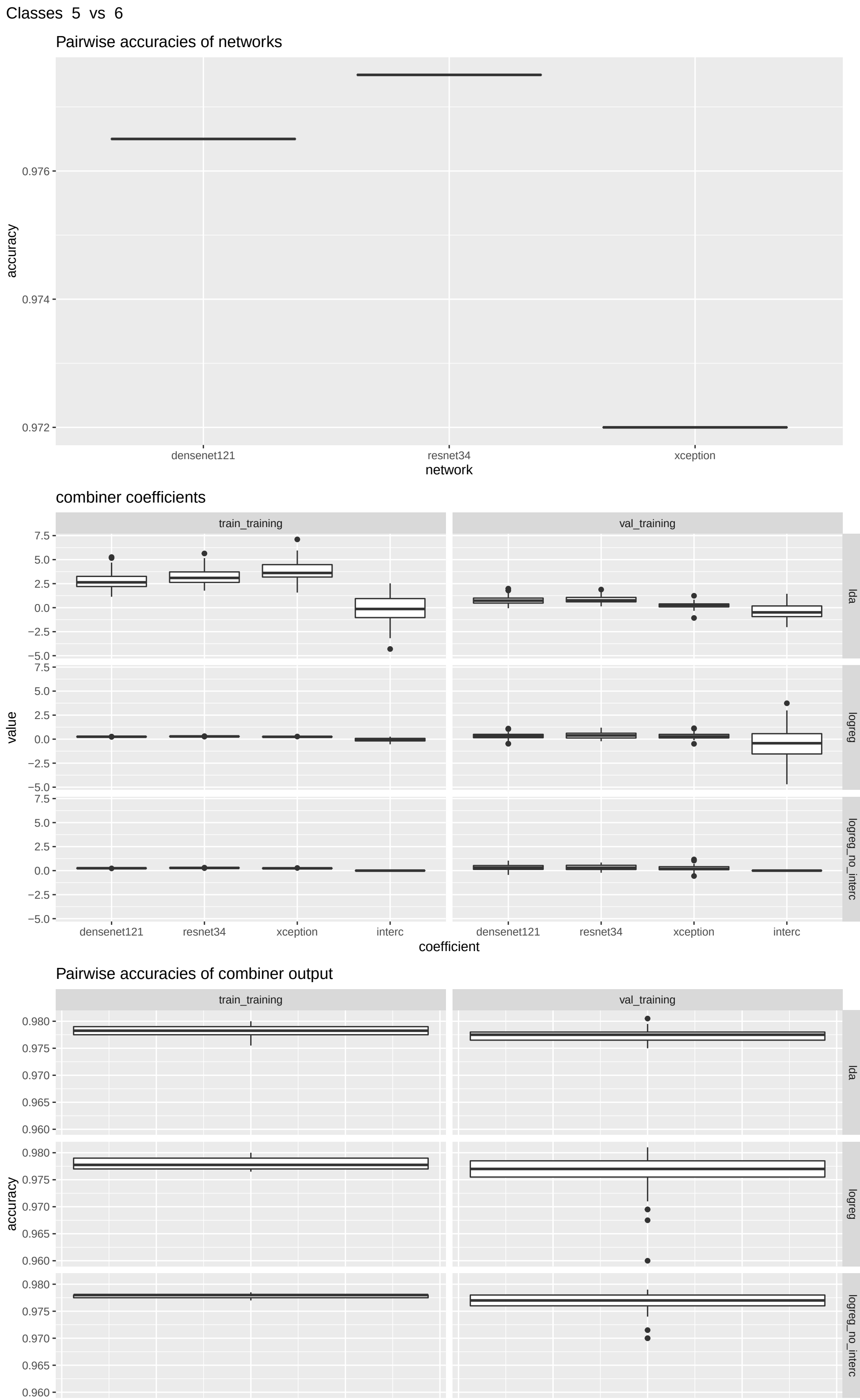

Classes 4 vs 8
Pairwise accuracies of networks
accuracy
0.980
0.975
densenet121
combiner coefficients
train_training
value
5
0
−5
−10
densenet121
resnet34
Pairwise accuracies of combiner
train_training
accuracy
0.980
0.975
0.970

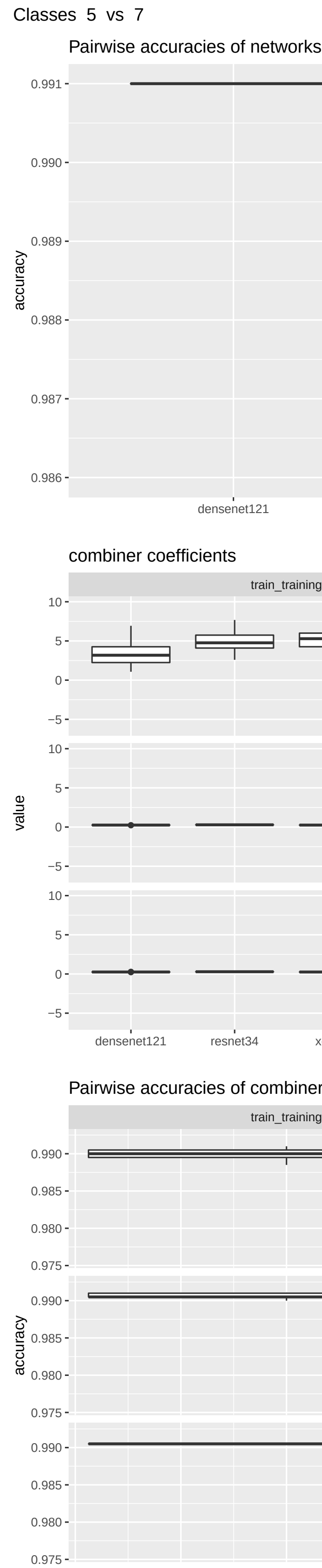
Classes 4 vs 9

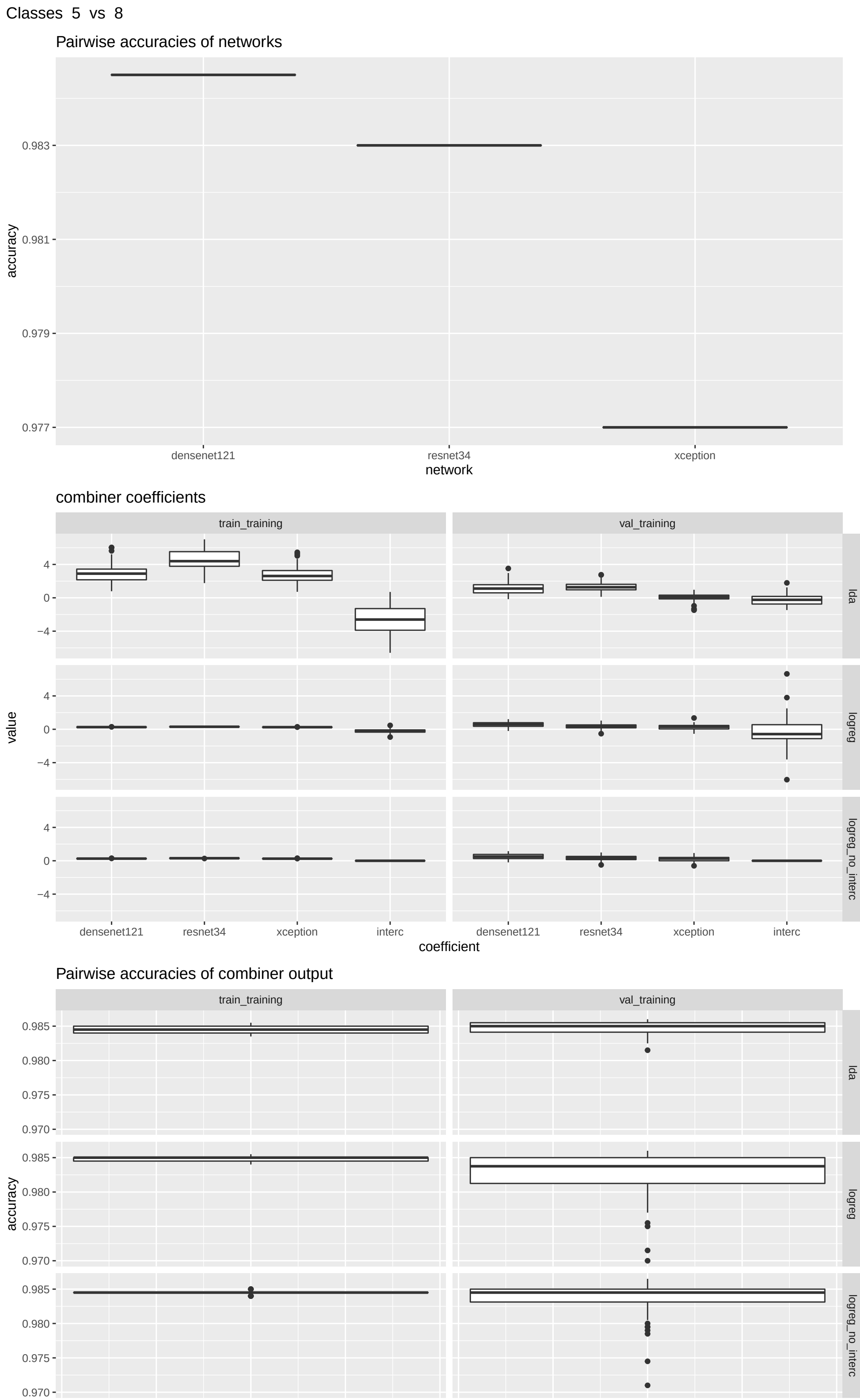


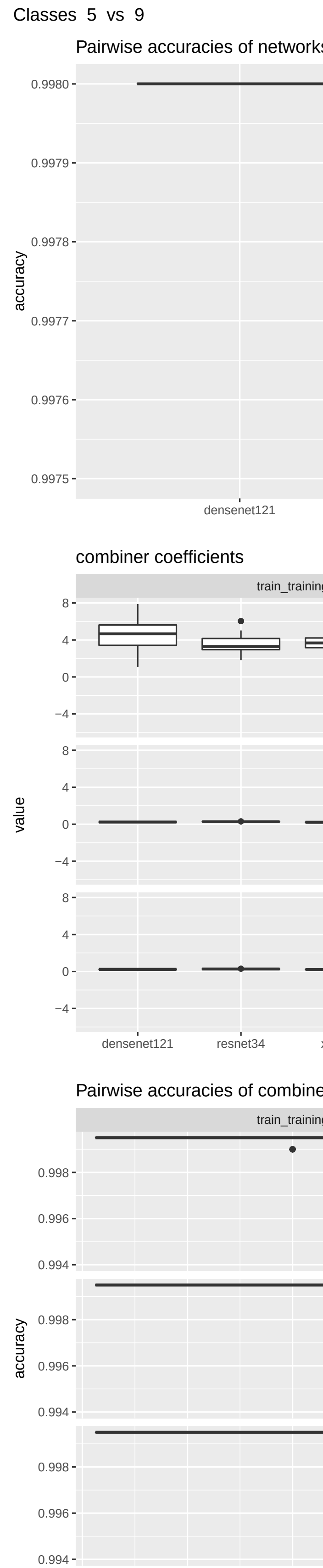
Classes 4 vs 10

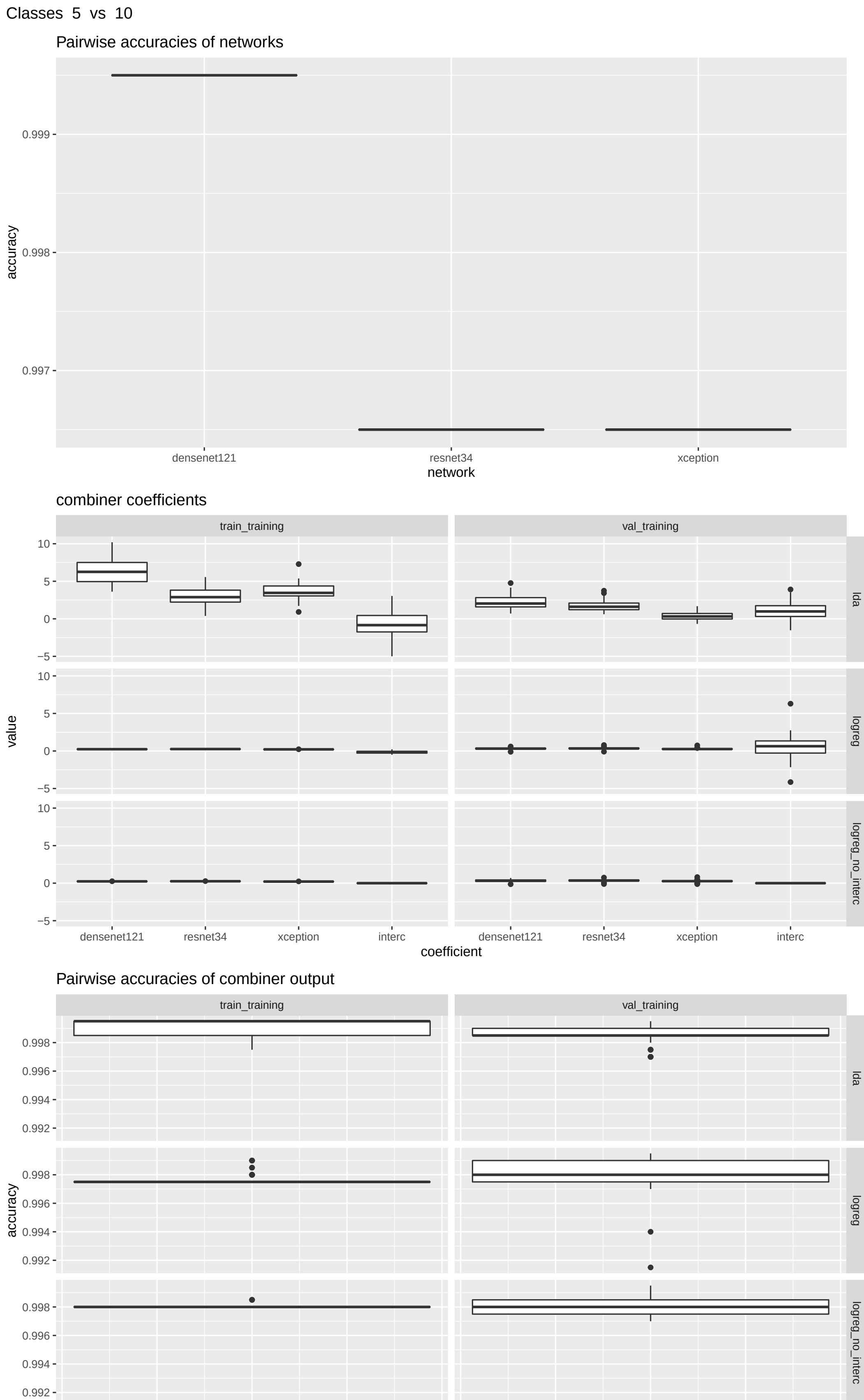


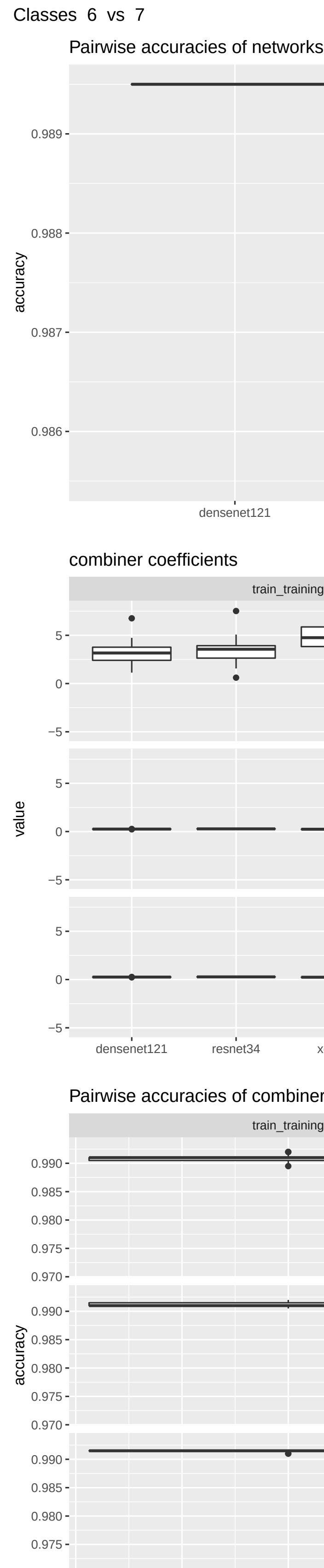

Classes 5 vs 6
Pairwise accuracies of networks
accuracy
0.976
0.974
0.972
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
7.5
5.0
2.5
0.0
−2.5
−5.0
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.980
0.975
0.970
0.965
0.960


Classes 5 vs 7
Pairwise accuracies of networks
accuracy
0.991
0.990
0.989
0.988
0.987
0.986
densenet121
combiner coefficients
train_training
value
10
5
0
−5
densenet121
resnet34
Pairwise accuracies of combiner
0.990
0.985
0.980
0.975

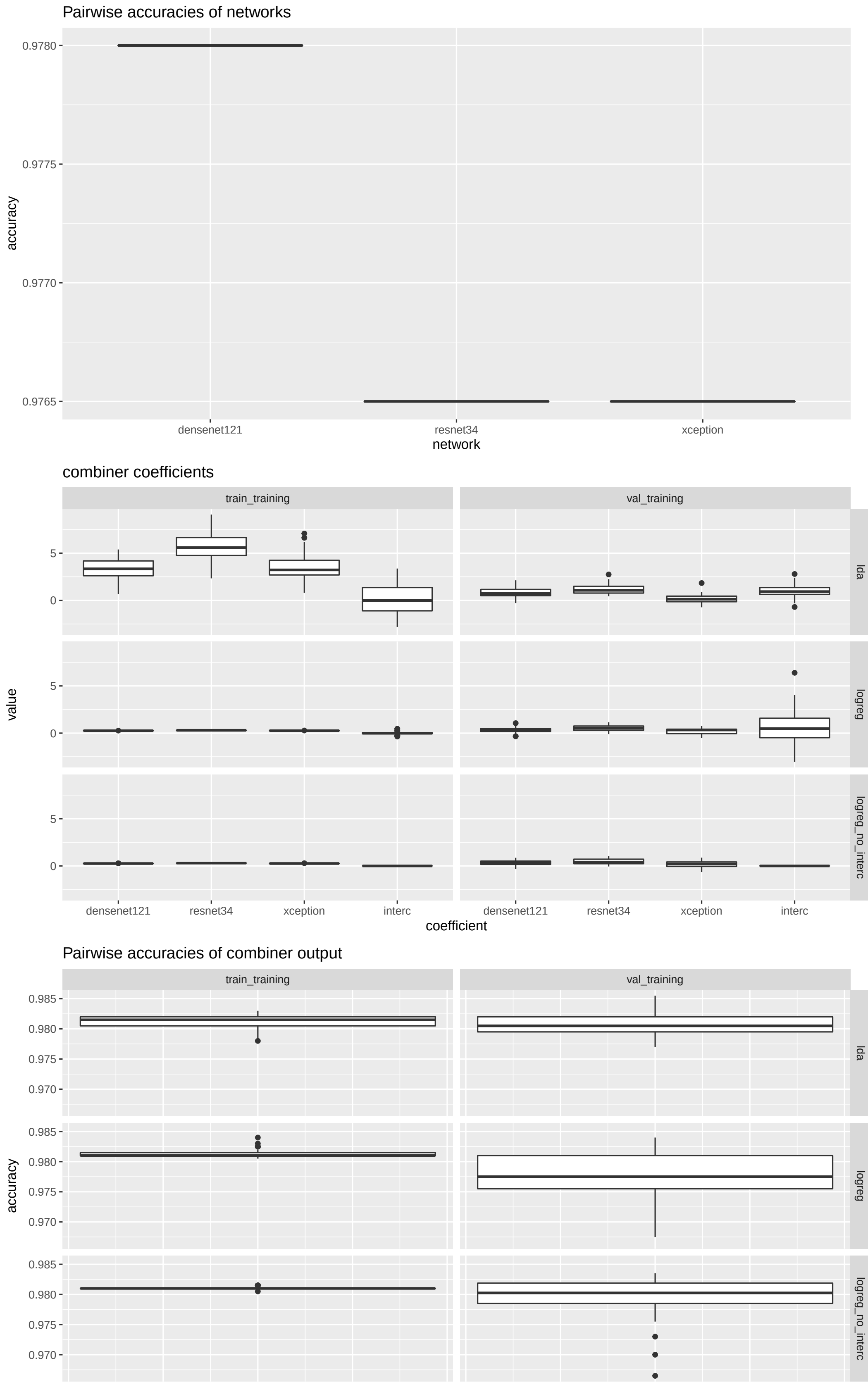

Classes 5 vs 8
Pairwise accuracies of networks
accuracy
0.983
0.981
0.979
0.977
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
4
0
−4
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
0.985
0.980
0.975
0.970
accuracy


Classes 5 vs 9
Pairwise accuracies of networks
accuracy
0.9980
0.9979
0.9978
0.9977
0.9976
0.9975
densenet121
combiner coefficients
value
8
4
0
−4
densenet121
resnet34
Pairwise accuracies of combiner output
accuracy
0.998
0.996
0.994

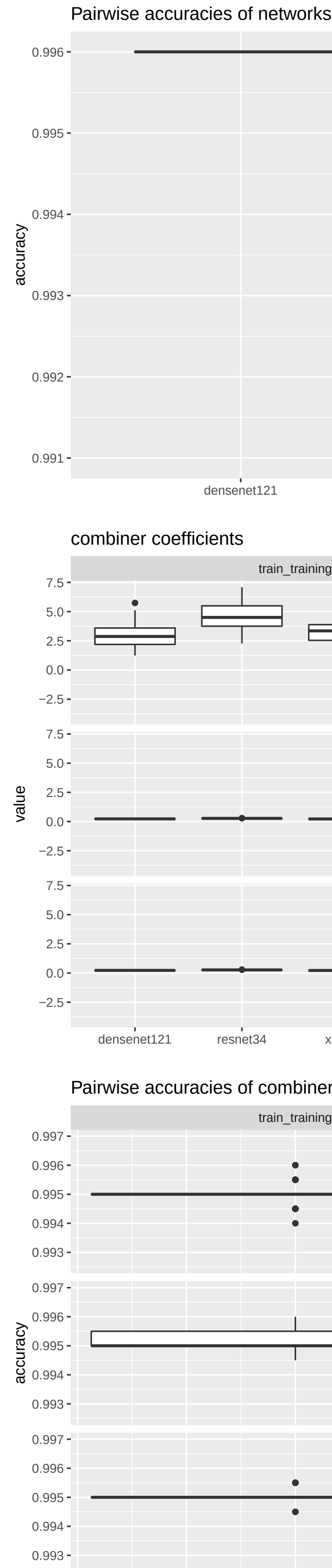

Classes 5 vs 10
Pairwise accuracies of networks
accuracy
0.999
0.998
0.997
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
10
5
0
−5
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
accuracy
0.998
0.996
0.994
0.992

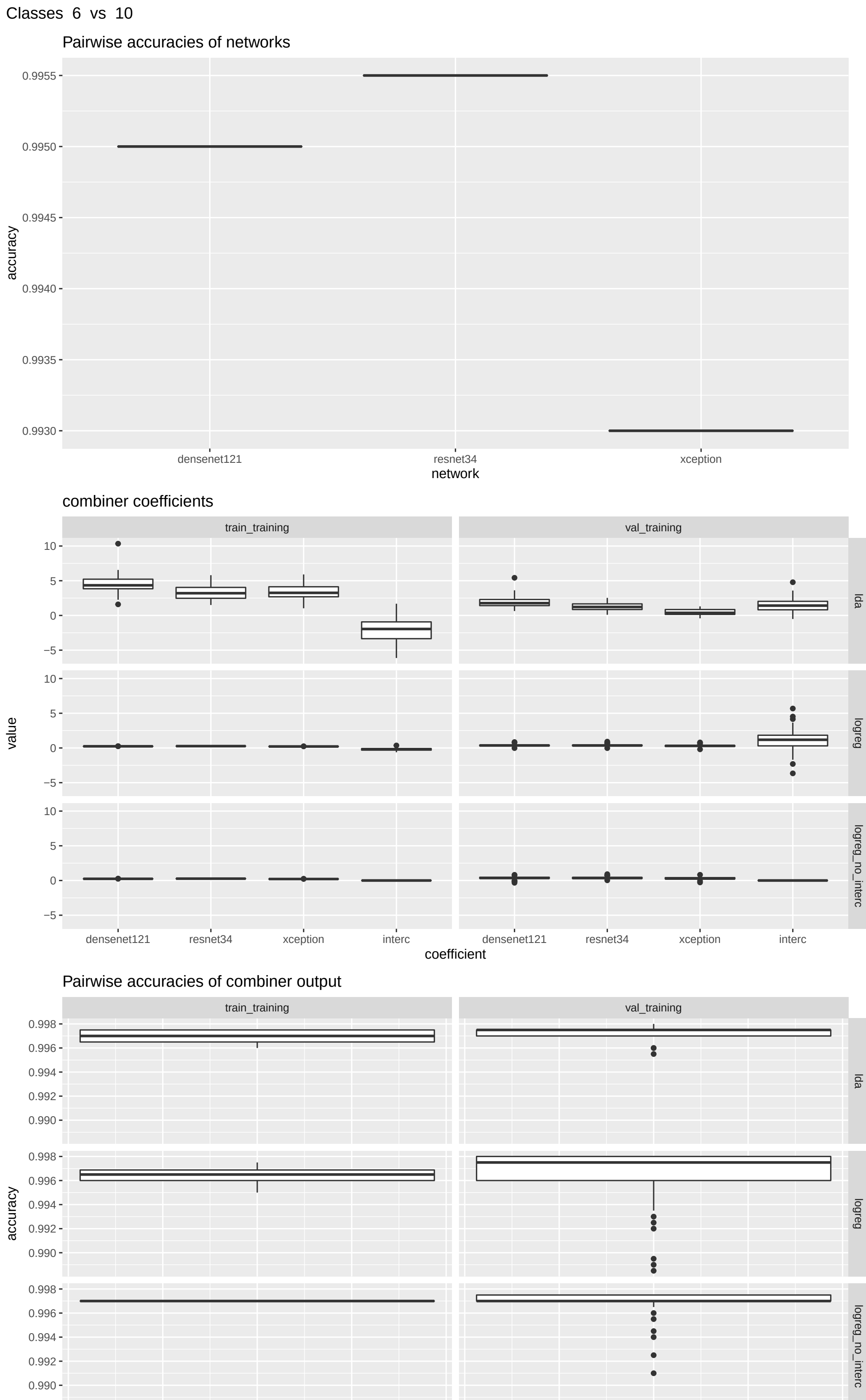

Classes 6 vs 7
Pairwise accuracies of networks
accuracy
0.989
0.988
0.987
0.986
densenet121
combiner coefficients
train_training
value
5
0
−5
densenet121
resnet34
Pairwise accuracies of combiner
train_training
accuracy
0.990
0.985
0.980
0.975
0.970

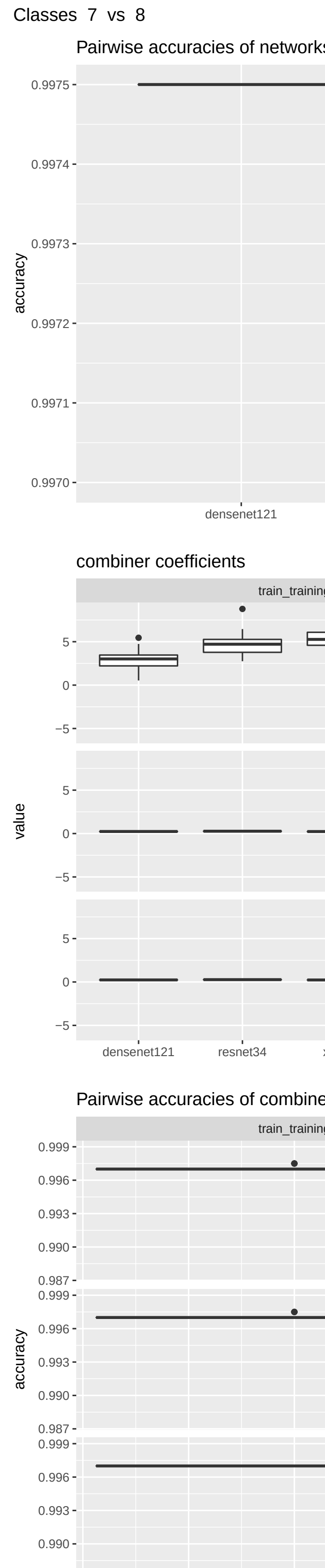
Classes 6 vs 8



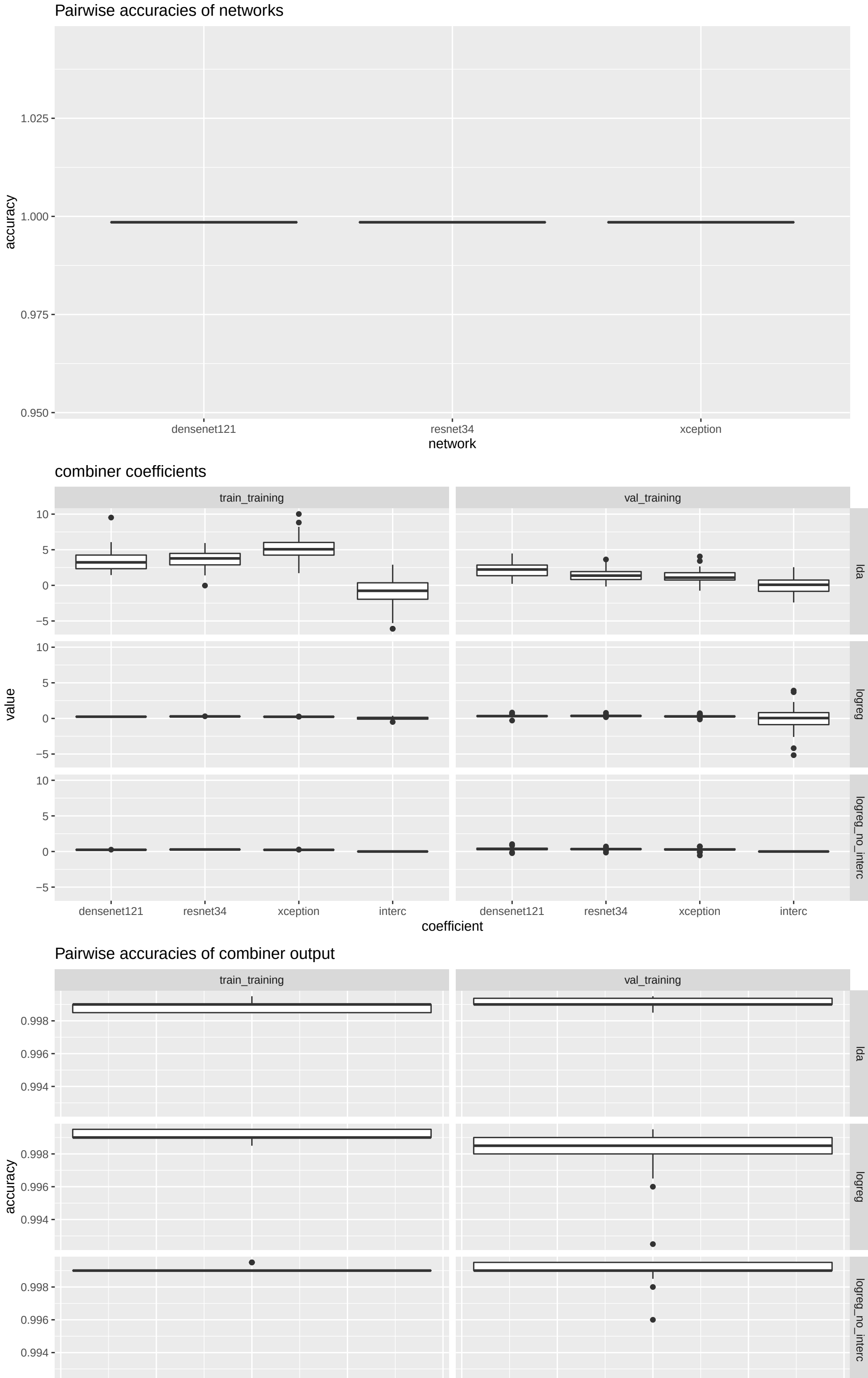
Classes 6 vs 9



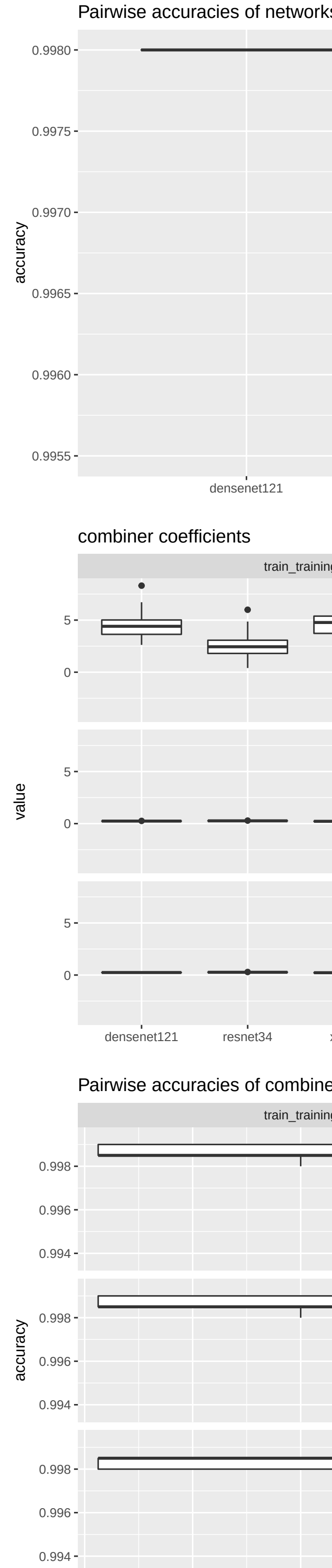

Classes 6 vs 10
Pairwise accuracies of networks
accuracy
0.9955
0.9950
0.9945
0.9940
0.9935
0.9930
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
10
5
0
−5
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
accuracy
0.998
0.996
0.994
0.992
0.990

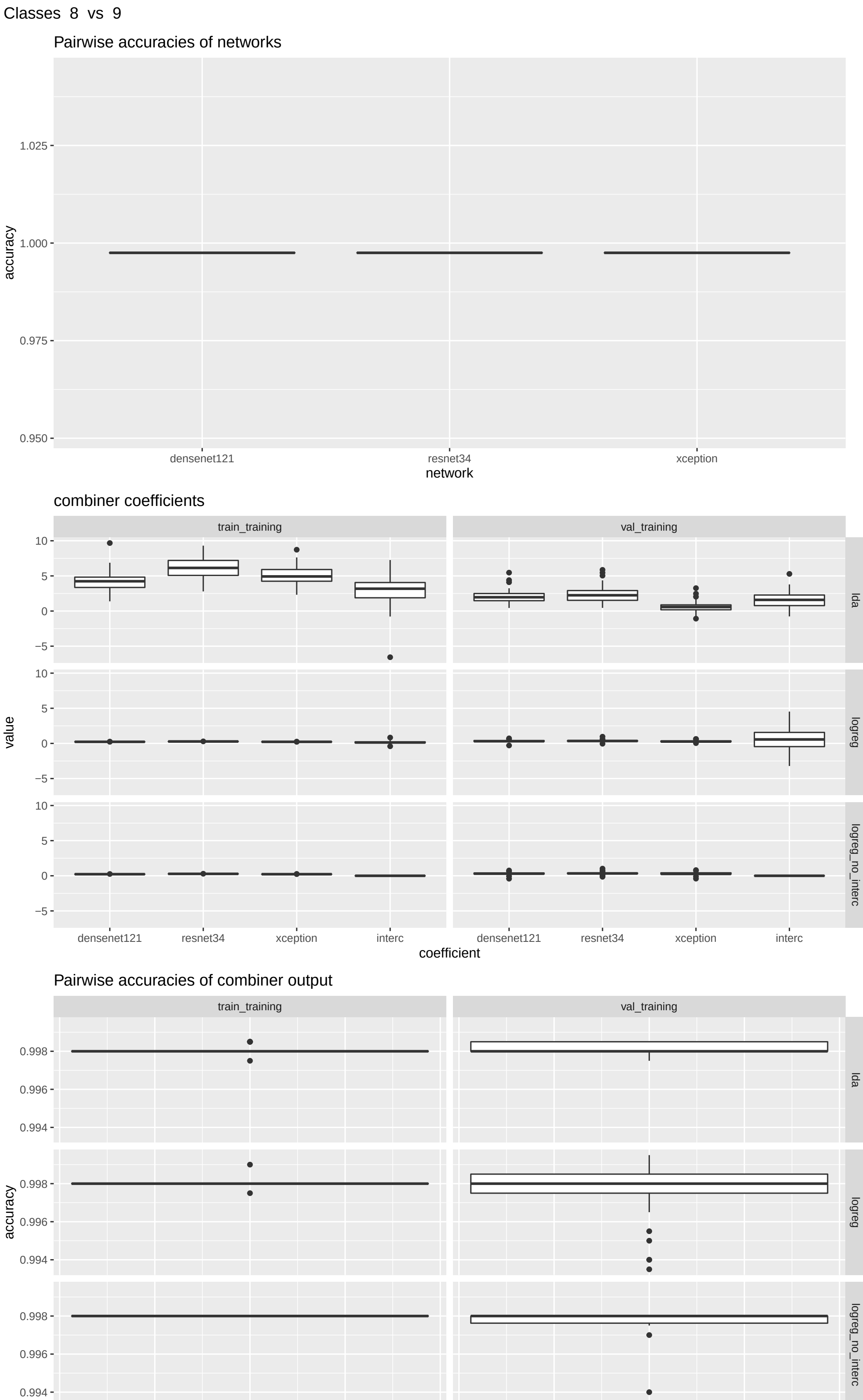

Classes 7 vs 8
Pairwise accuracies of networks
accuracy
0.9975
0.9974
0.9973
0.9972
0.9971
0.9970
densenet121
combiner coefficients
value
5
0
−5
densenet121
resnet34
accuracy
0.999
0.996
0.993
0.990
0.987

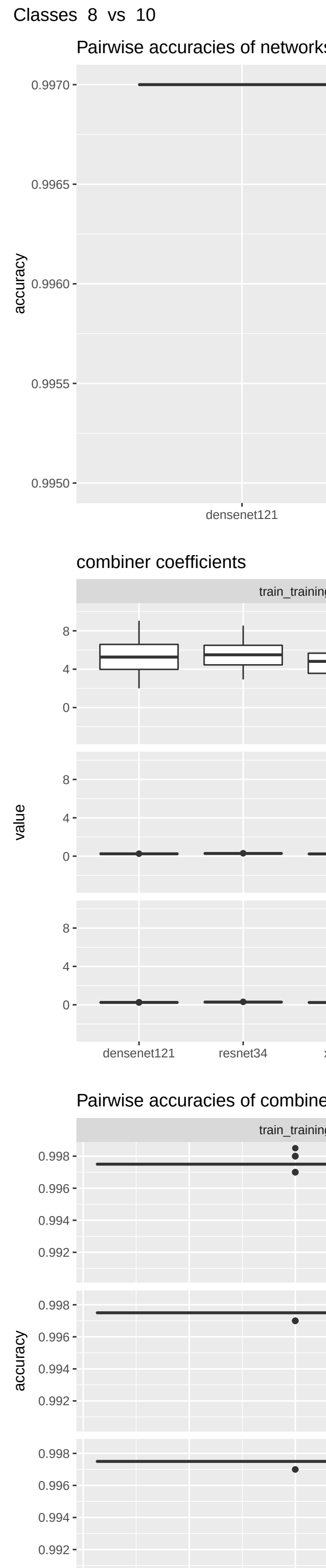
Classes 7 vs 9



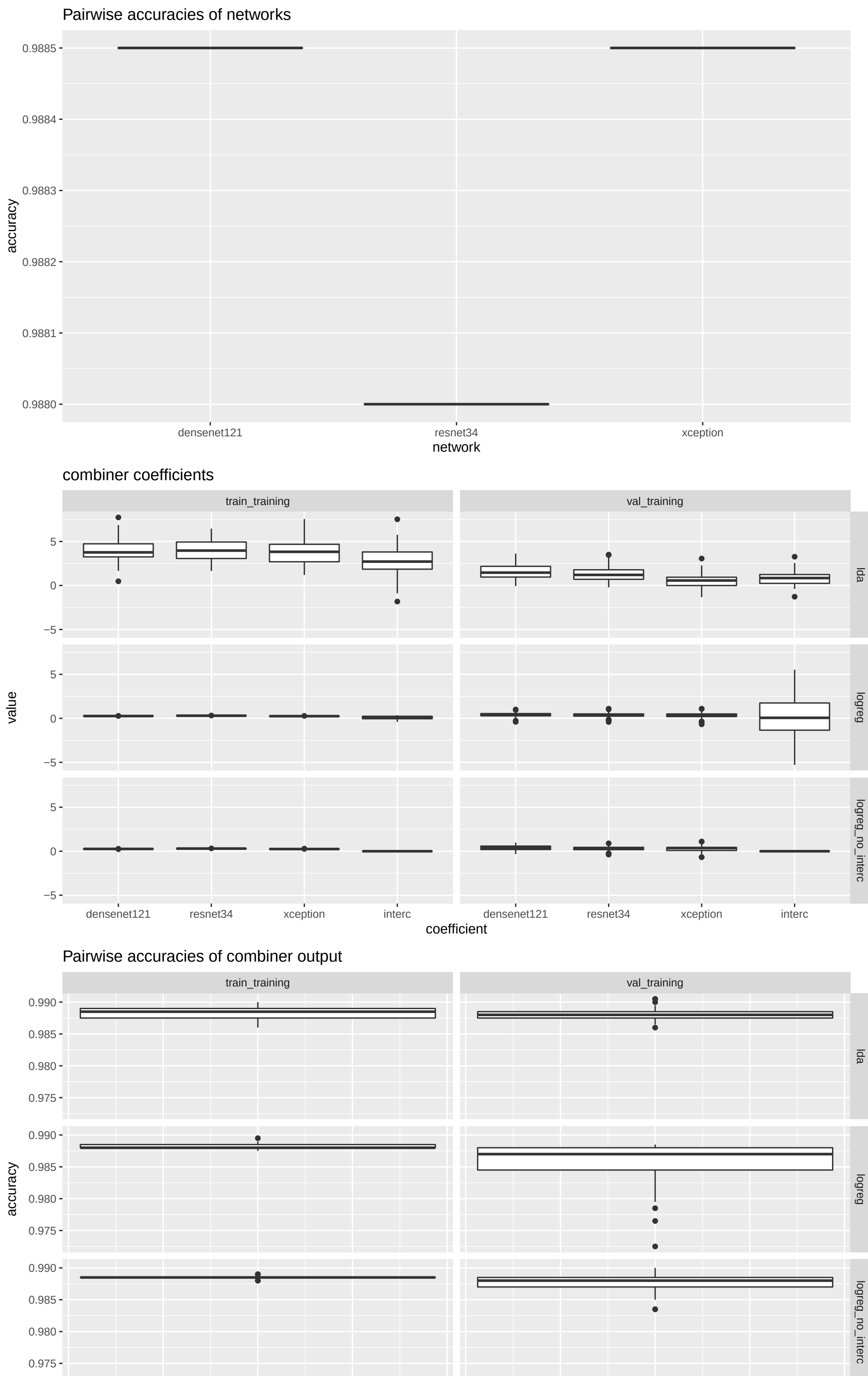
Classes 7 vs 10




Classes 8 vs 9
Pairwise accuracies of networks
accuracy
1.025
1.000
0.975
0.950
densenet121
resnet34
xception
network
combiner coefficients
train_training
val_training
value
10
5
0
−5
lda
logreg
logreg_no_interc
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
accuracy
0.998
0.996
0.994
lda
logreg
logreg_no_interc


Classes 8 vs 10
0.9970
0.9965
0.9960
0.9955
0.9950
accuracy
densenet121
combiner coefficients
train_training
value
8
4
0
densenet121
resnet34
accuracy
0.998
0.996
0.994
0.992

Classes 9 vs 10



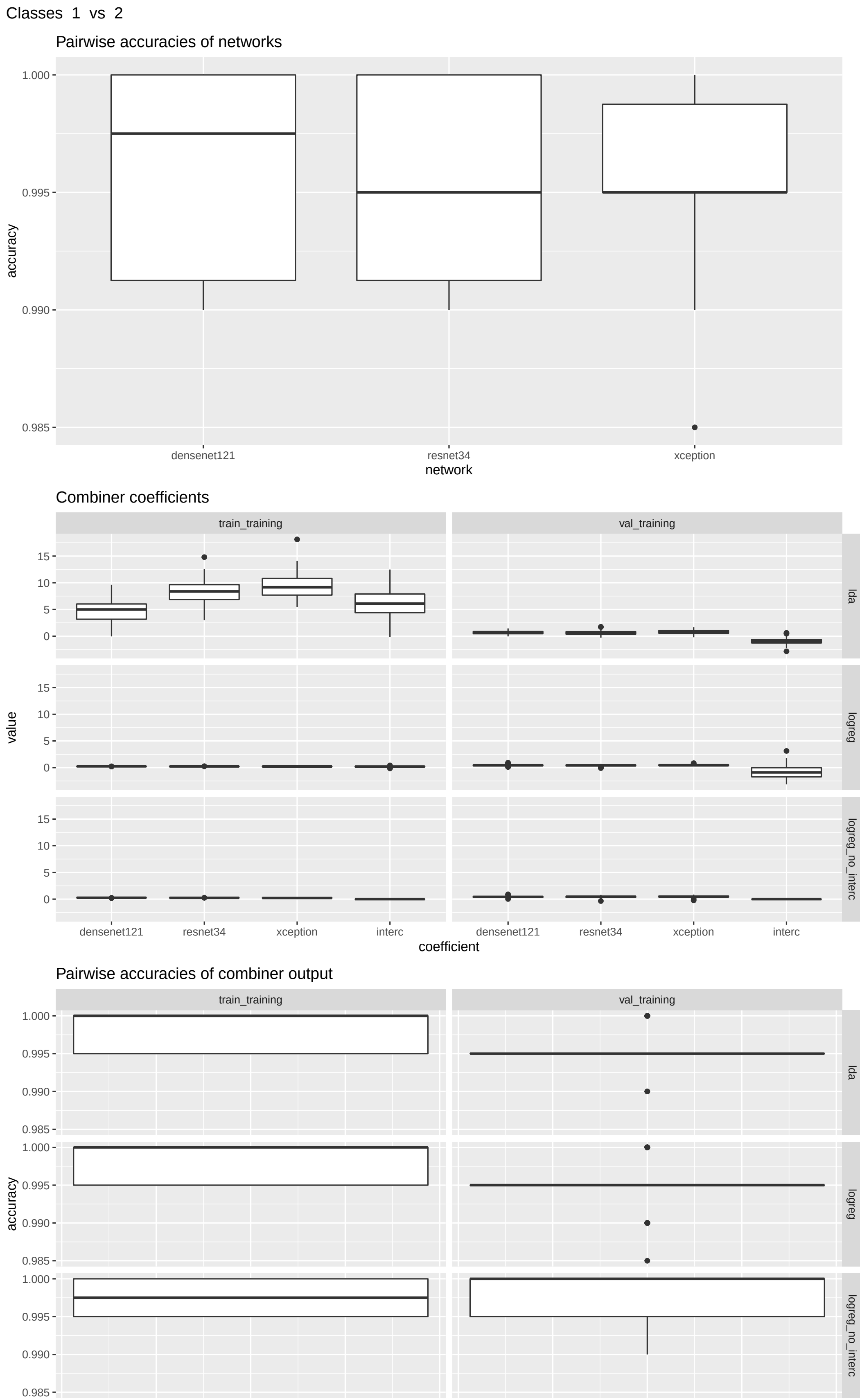
Val training

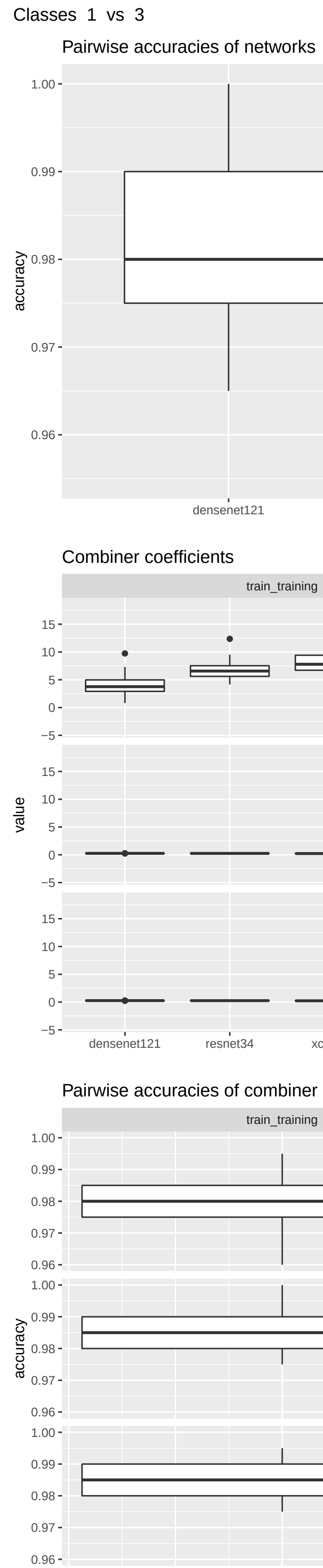
accuracies have higher variance.

CIFAR 100 half

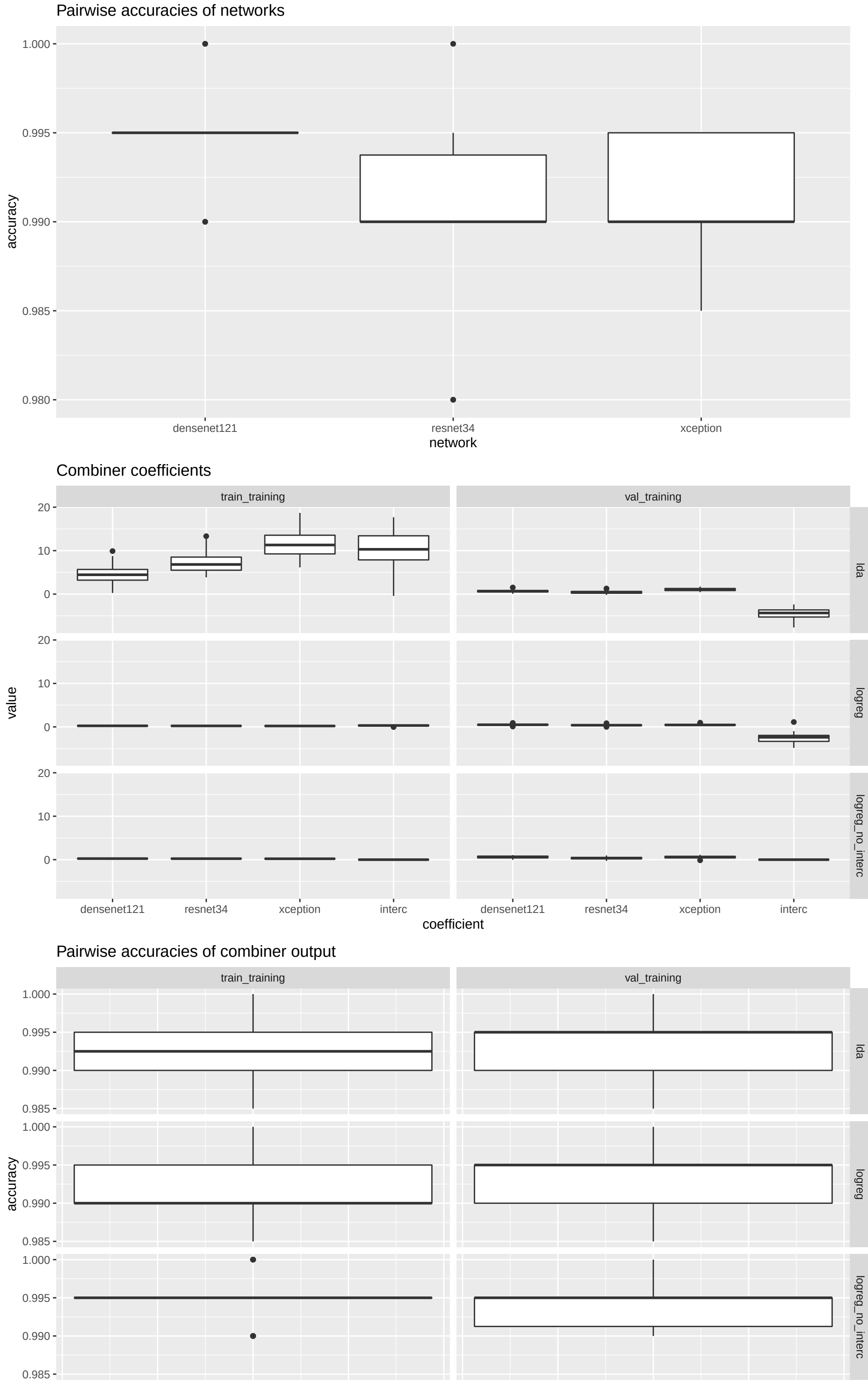
```
base_dir <- "../data/data_train_val_half_c100"
repls <- 0:9
folds <- 0:4
disp_classes <- 10
combiner_coefs <- load_combiner_coefs(base_dir, repls, folds)
net_pw_results <- read.csv(file.path(base_dir, "net_pw_accuracies.csv"))
ens_pw_results <- read.csv(file.path(base_dir, "ensemble_pw_accuracies.csv"))
net_pw_results[, c("class1", "class2")] <- lapply(net_pw_results[, c("class1", "class2")], as.factor)
ens_pw_results[, c("class1", "class2", "combining_method")] <- lapply(ens_pw_results[, c("class1", "cla
```

```
for (cl1 in 1:(disp_classes - 1))
{
  for (cl2 in (cl1 + 1):disp_classes)
  {
    combiner_plt <- combiner_coefs %>% filter(class1 == cl1 & class2 == cl2) %>% ggplot() + geom_boxplot
      facet_grid(cols=vars(train_type), rows=vars(combining_method)) + ggtitle("Combiner coefficients")

    acc_plt_net <- net_pw_results %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot(mappi
      geom_boxplot() + ggtitle("Pairwise accuracies of networks")

    acc_plt_ens <- ens_pw_results %>% filter(class1 == (cl1 - 1) & class2 == (cl2 - 1)) %>% ggplot(mappi
      geom_boxplot() + facet_grid(cols=vars(train_set), rows=vars(combining_method)) +
      ggtitle("Pairwise accuracies of combiner output") +
      theme(axis.ticks.x=element_blank(), axis.text.x=element_blank())


    print((acc_plt_net/combiner_plt/acc_plt_ens) + plot_annotation(title=paste("Classes ", cl1, " vs ",
  }
}
```

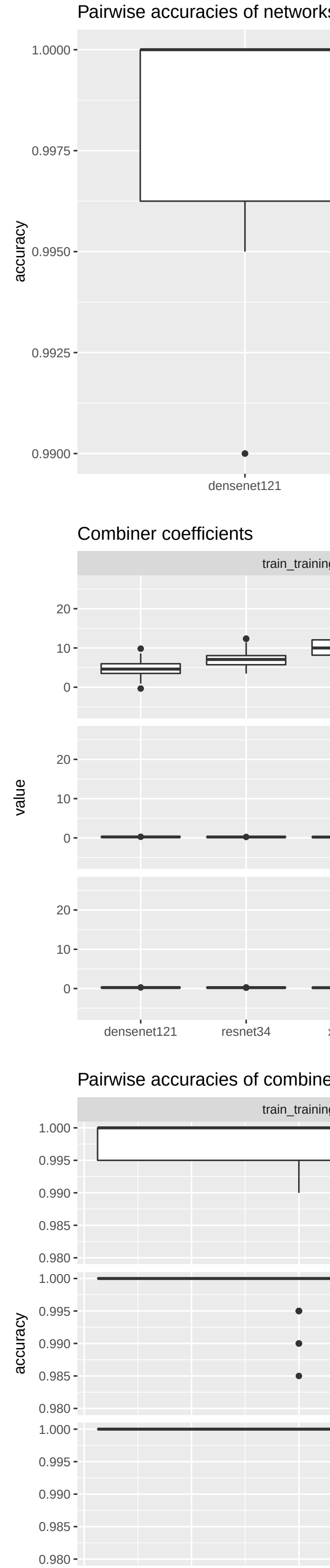

Classes 1 vs 2
Pairwise accuracies of networks
accuracy
1.000
0.995
0.990
0.985
densenet121
resnet34
xception
network
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
15
10
5
0
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy
1.000
0.995
0.990
0.985

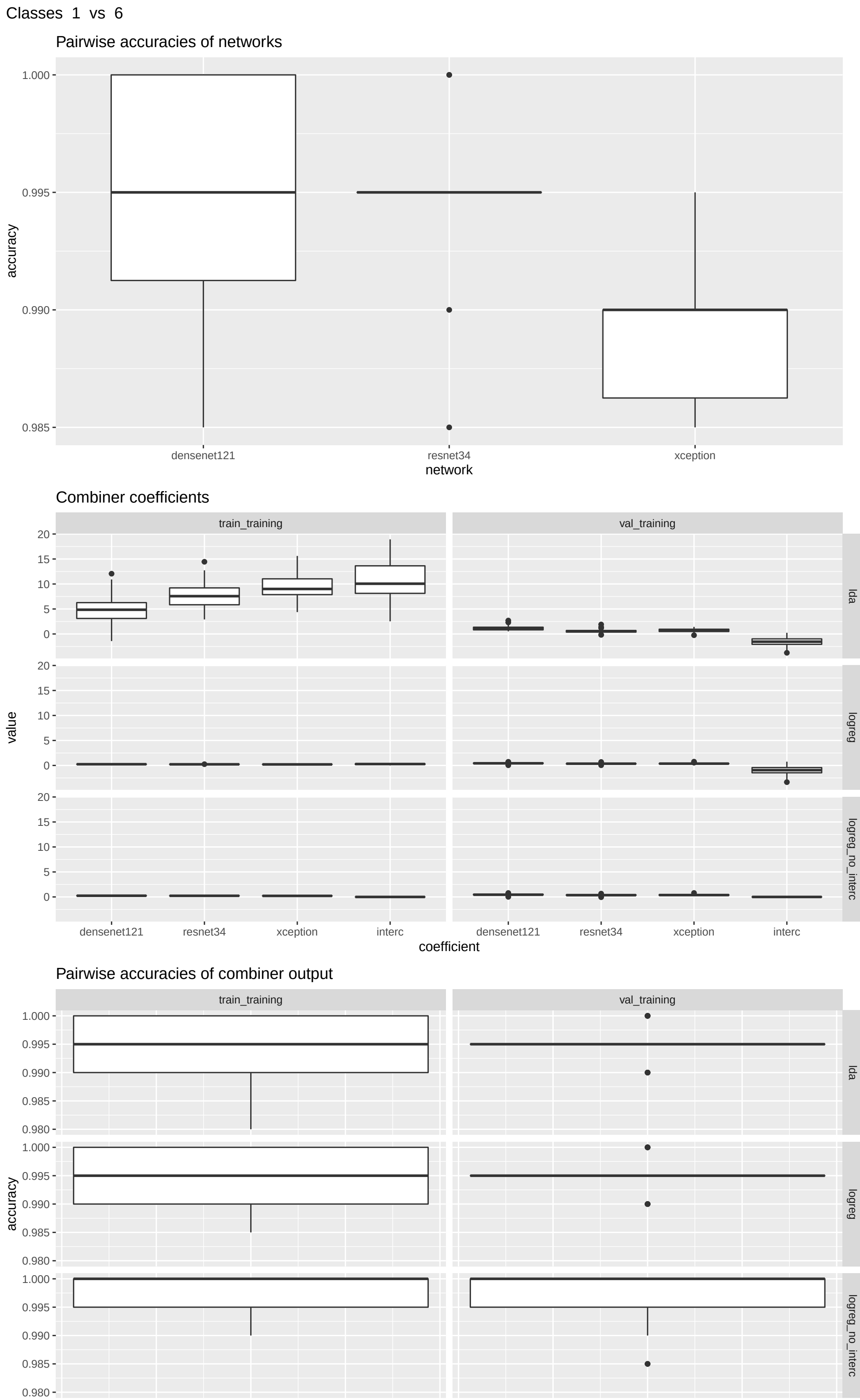

Classes 1 vs 3
Pairwise accuracies of networks
accuracy
1.00
0.99
0.98
0.97
0.96
densenet121
Combiner coefficients
train_training
value
15
10
5
0
−5
densenet121
resnet34
Pairwise accuracies of combiner
train_training
accuracy
1.00
0.99
0.98
0.97
0.96

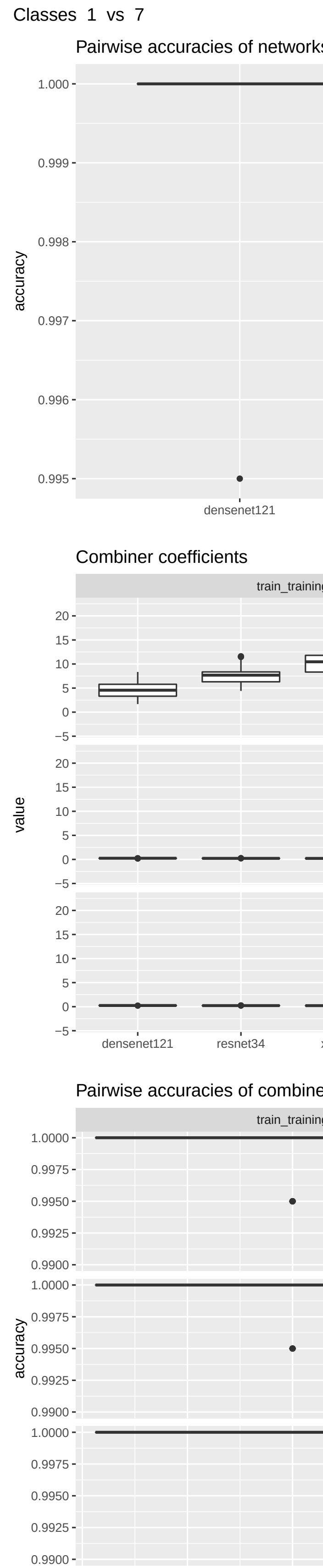
Classes 1 vs 4

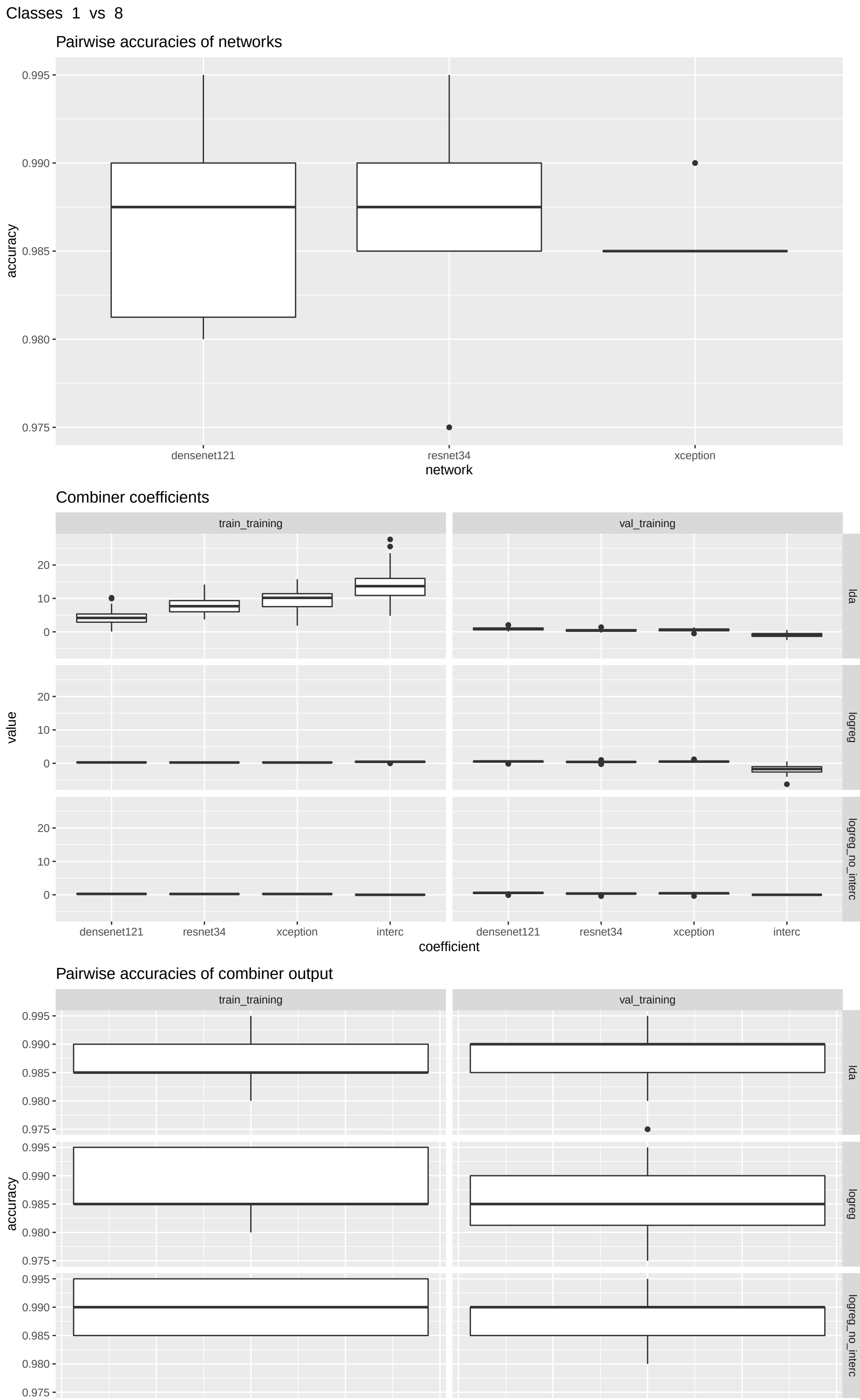


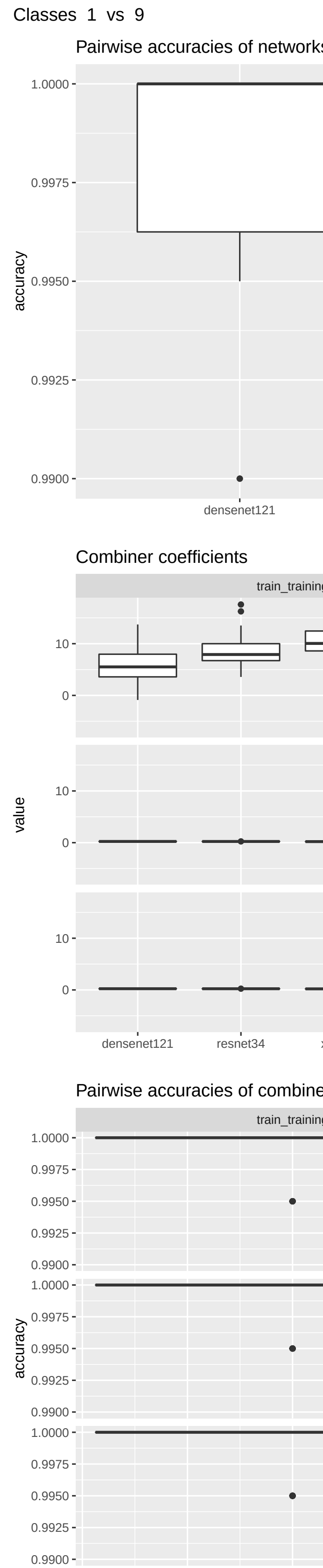
Classes 1 vs 5



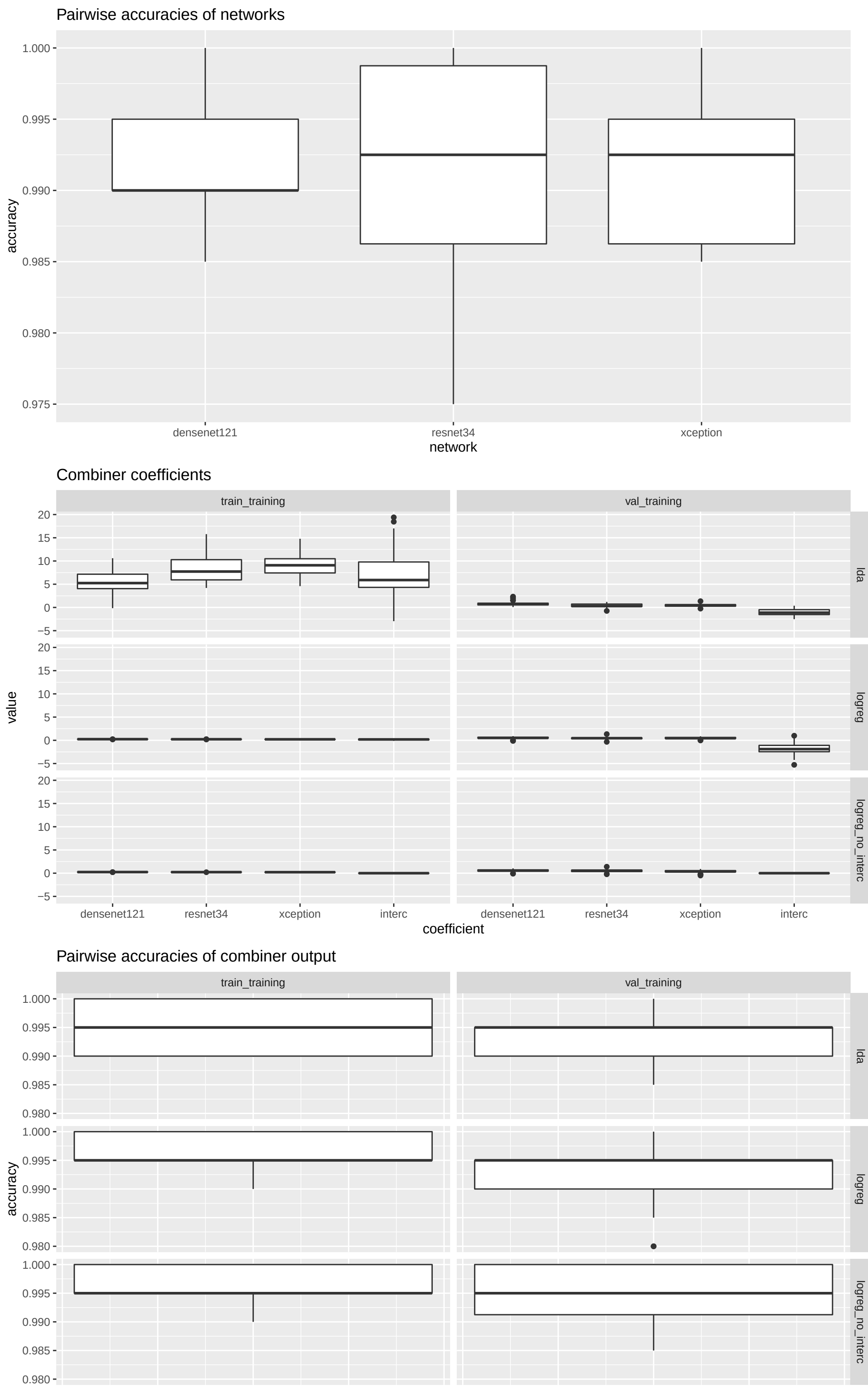

Classes 1 vs 6
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output


Classes 1 vs 7
accuracy
densenet121
Combiner coefficients
value
resnet34

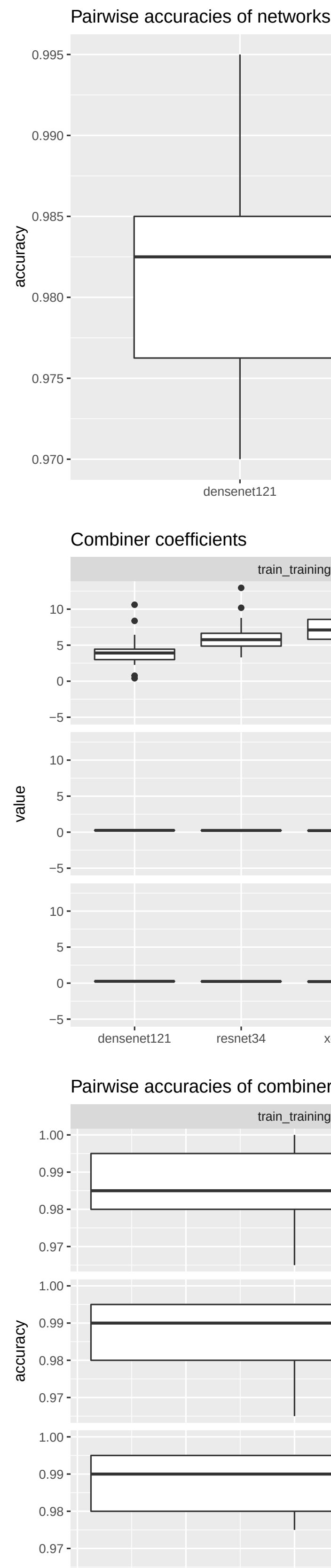

Classes 1 vs 8
Pairwise accuracies of networks
accuracy
0.995
0.990
0.985
0.980
0.975
densenet121
resnet34
xception
network
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
20
10
0
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy
0.995
0.990
0.985
0.980
0.975


Classes 1 vs 9
Pairwise accuracies of networks
accuracy
1.0000
0.9975
0.9950
0.9925
0.9900
densenet121
Combiner coefficients
value
10
0
densenet121
resnet34
Pairwise accuracies of combine
accuracy
1.0000
0.9975
0.9950
0.9925
0.9900

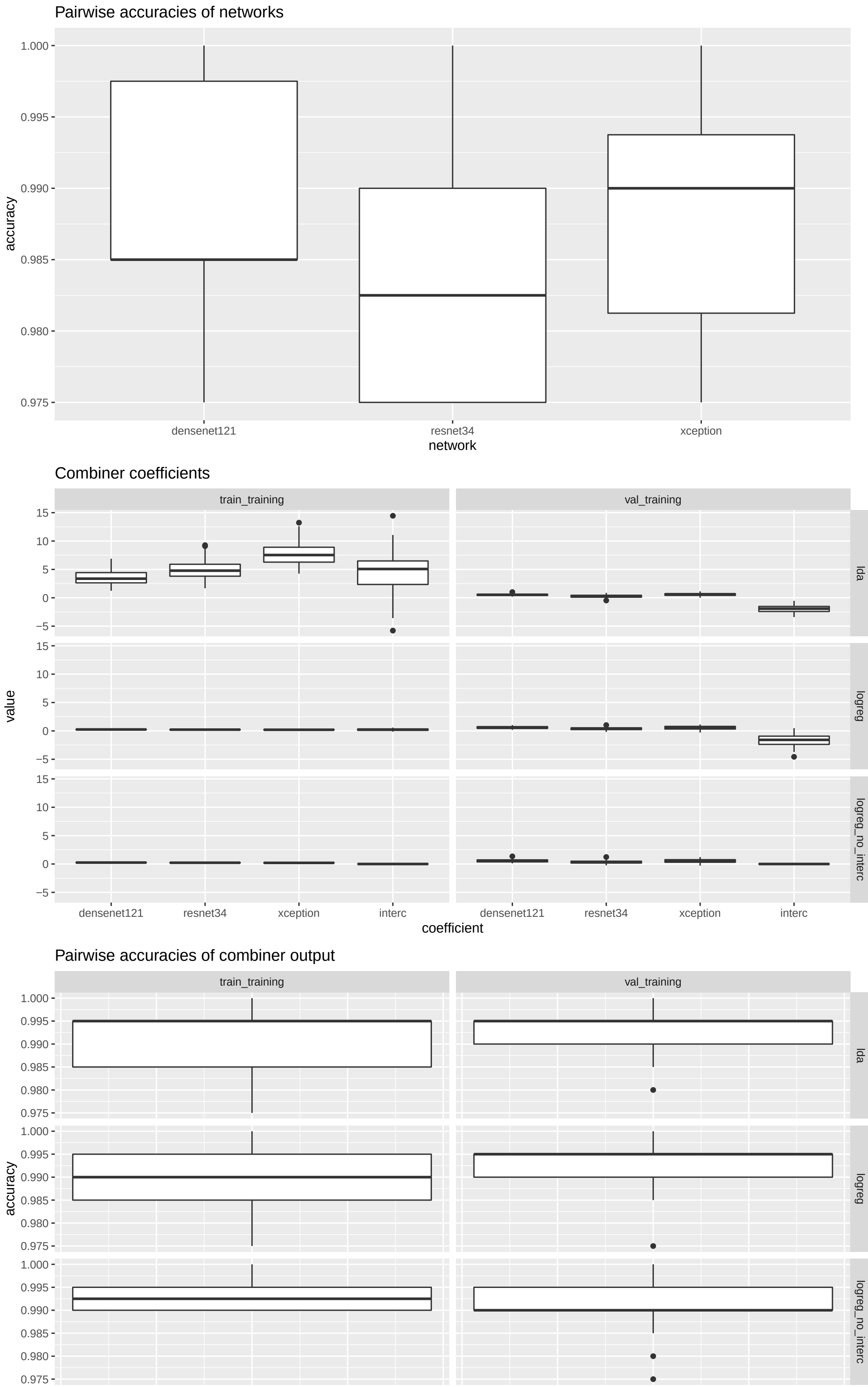
Classes 1 vs 10



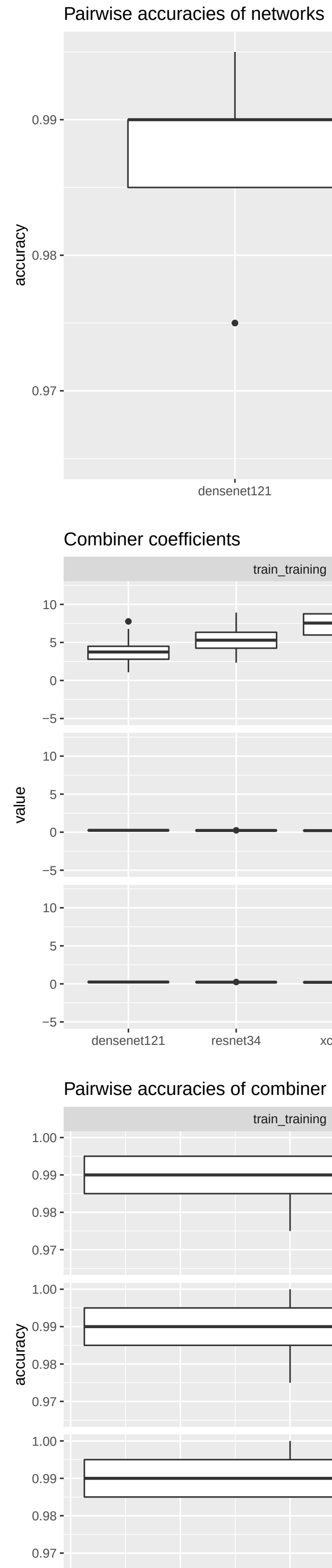
Classes 2 vs 3



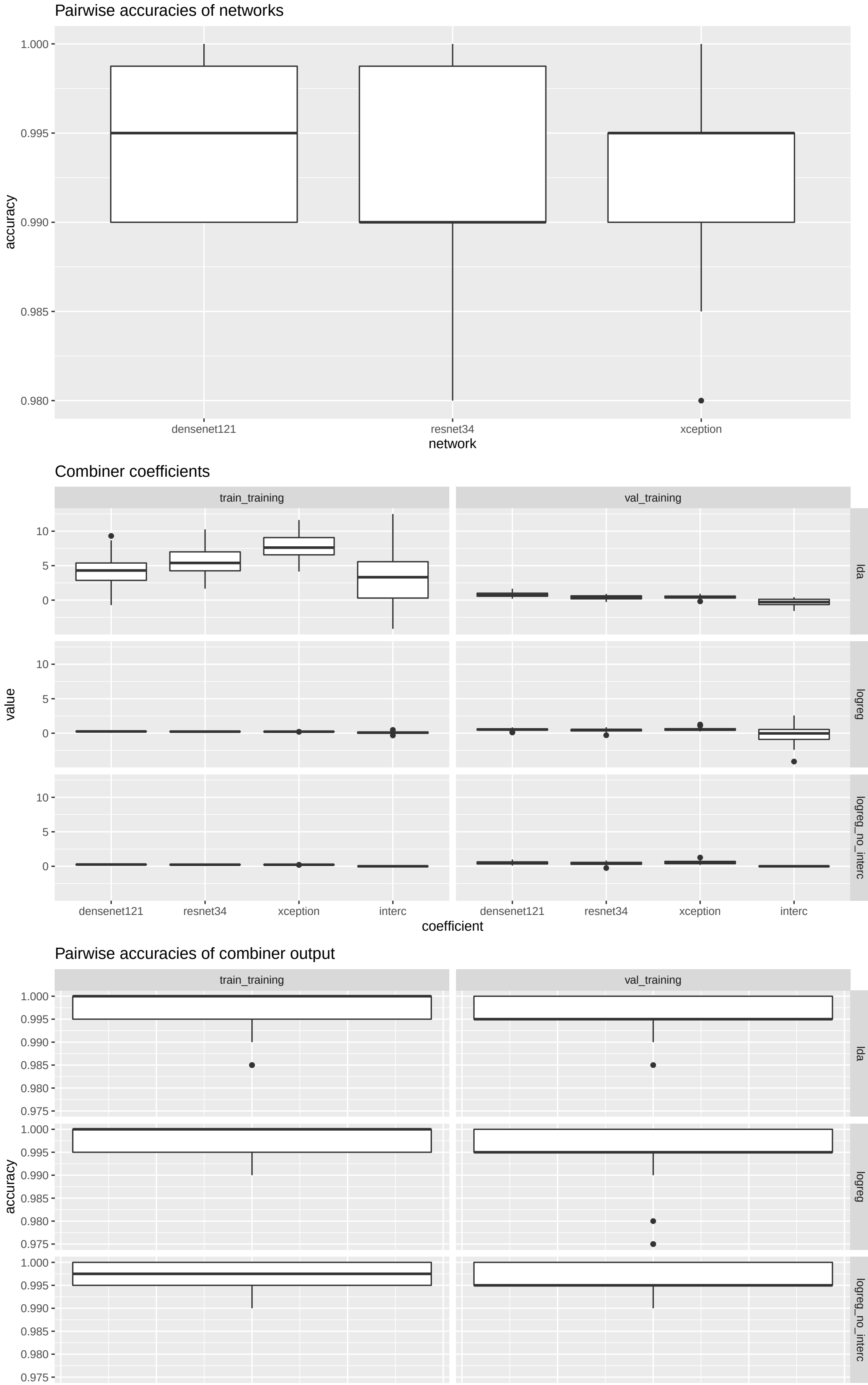
Classes 2 vs 4



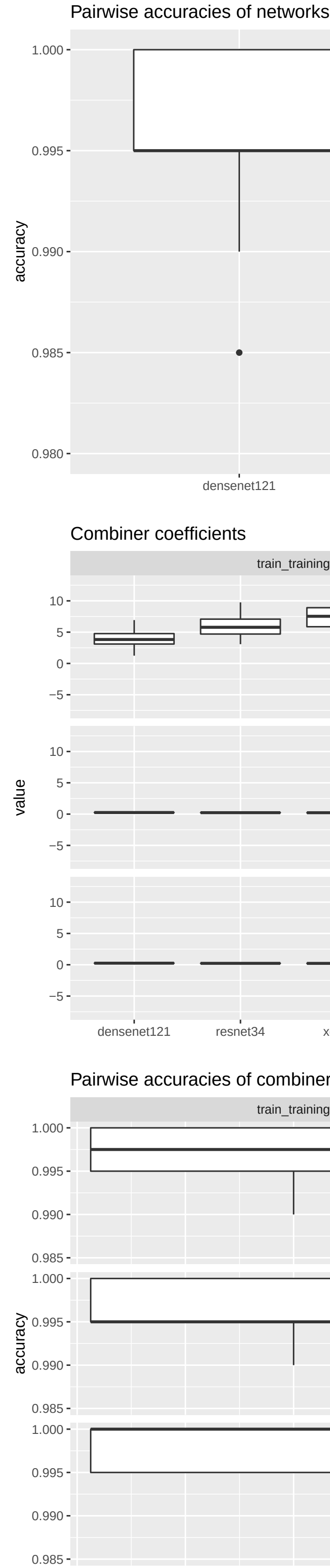
Classes 2 vs 5



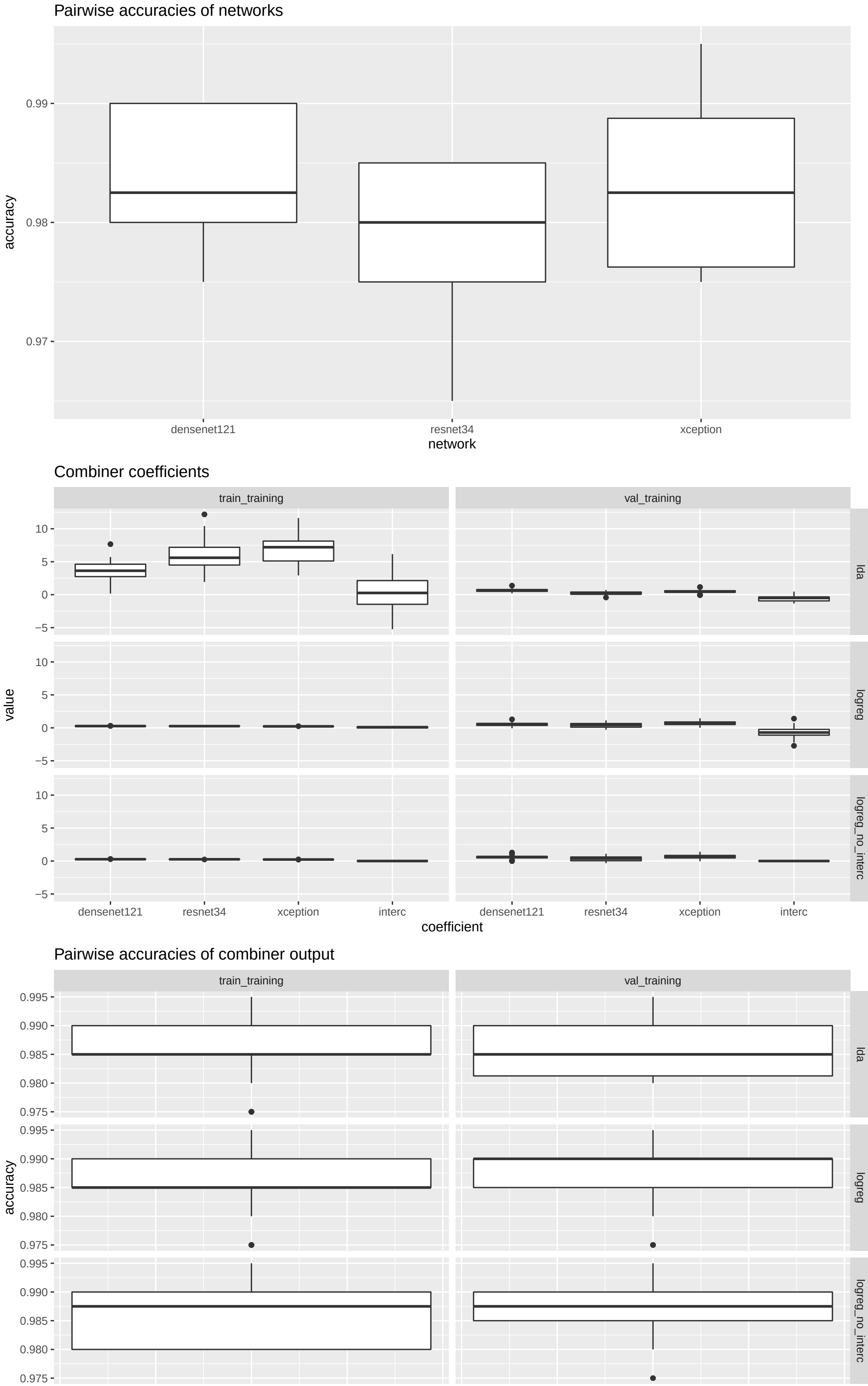
Classes 2 vs 6



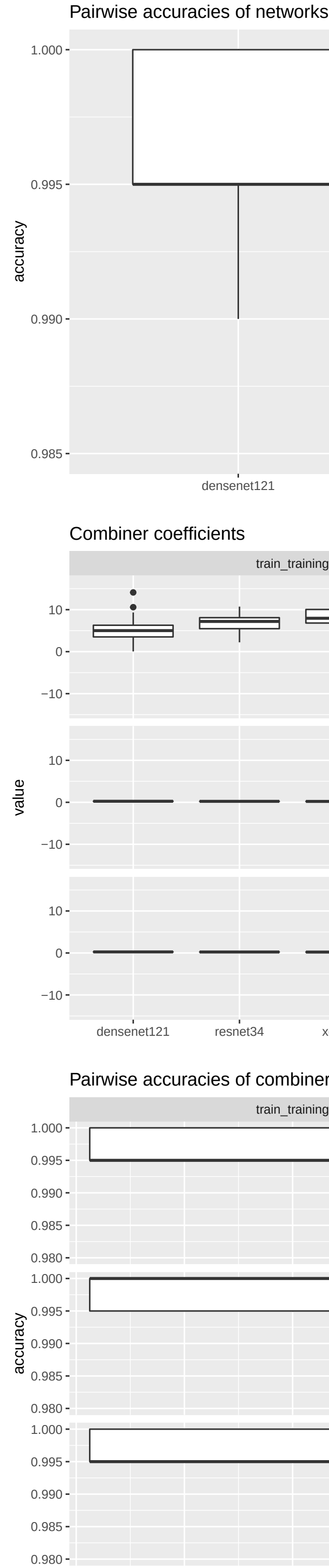
Classes 2 vs 7

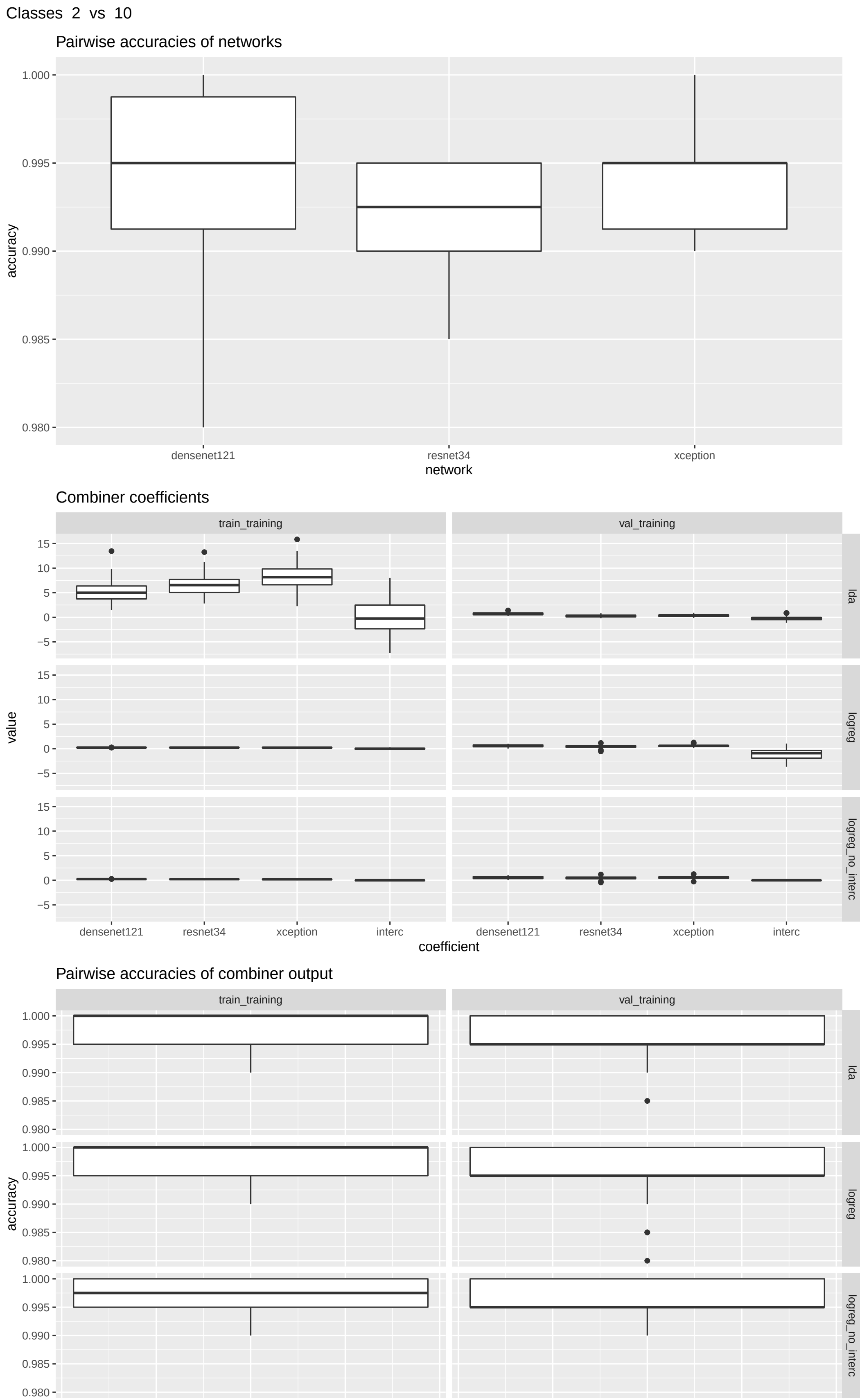


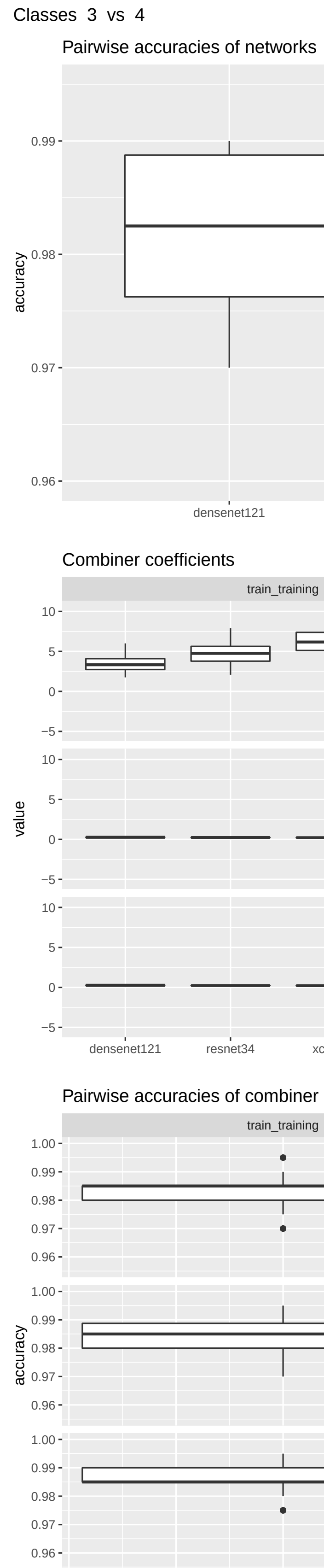
Classes 2 vs 8

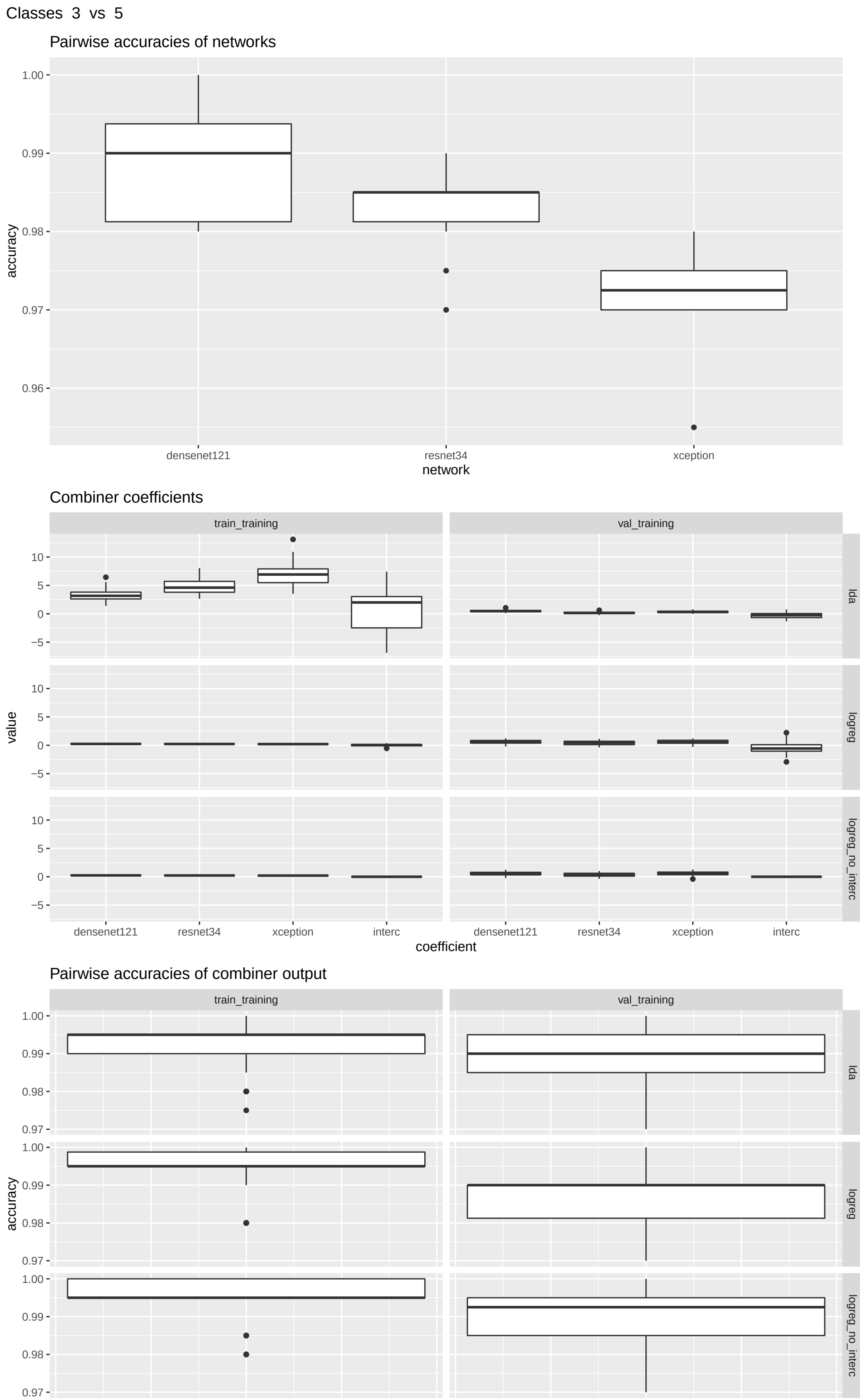


Classes 2 vs 9

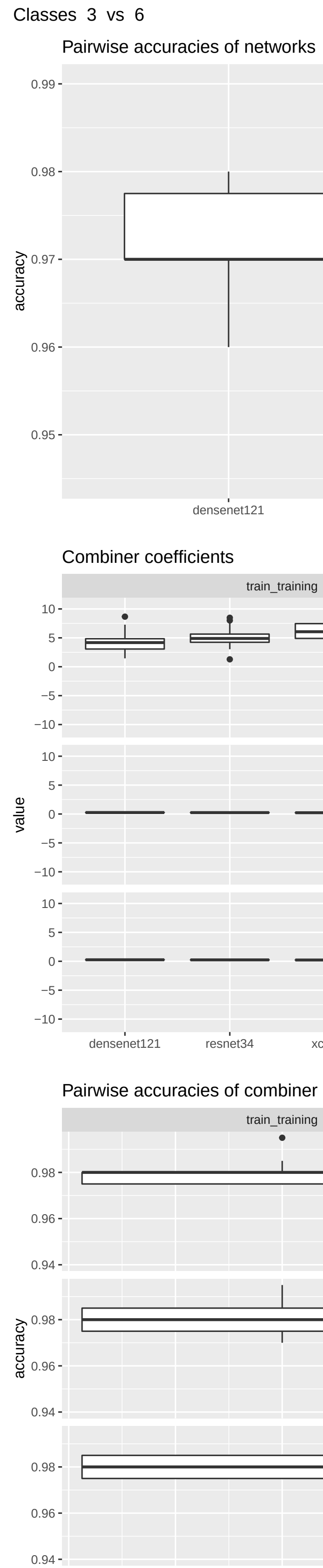



Classes 2 vs 10
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
interc
Pairwise accuracies of combiner output

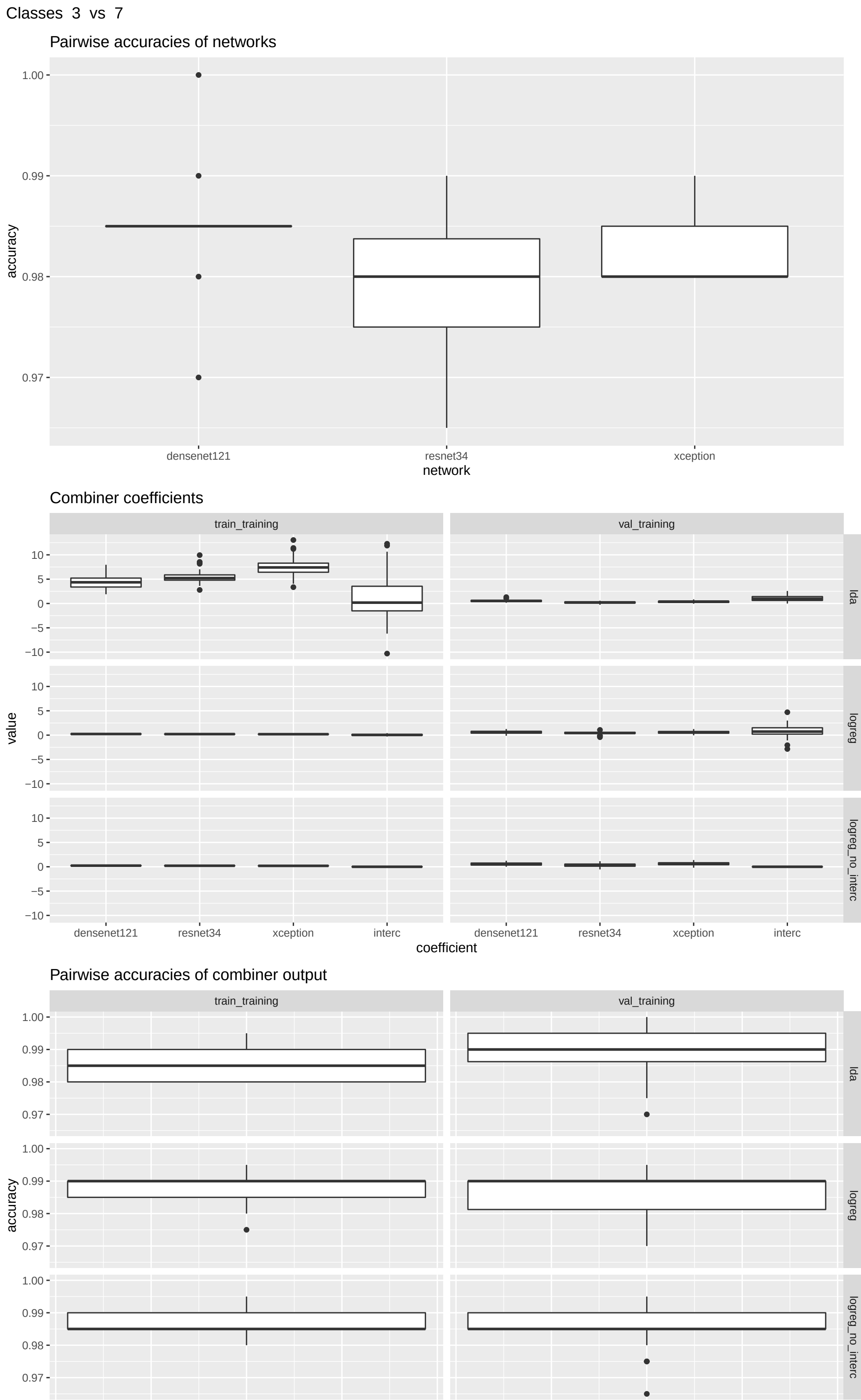

Classes 3 vs 4
Pairwise accuracies of networks
accuracy
densenet121
Combiner coefficients
train_training
value
resnet34
Pairwise accuracies of combiner

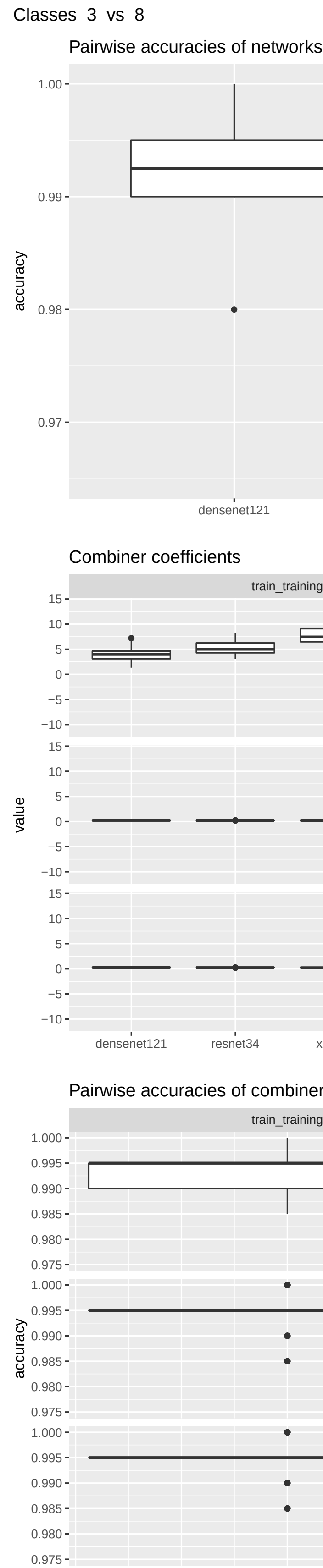


Classes 3 vs 5
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
Combiner coefficients
train_training
val_training
value
coefficient
lda
logreg
logreg_no_interc
interc
Pairwise accuracies of combiner output

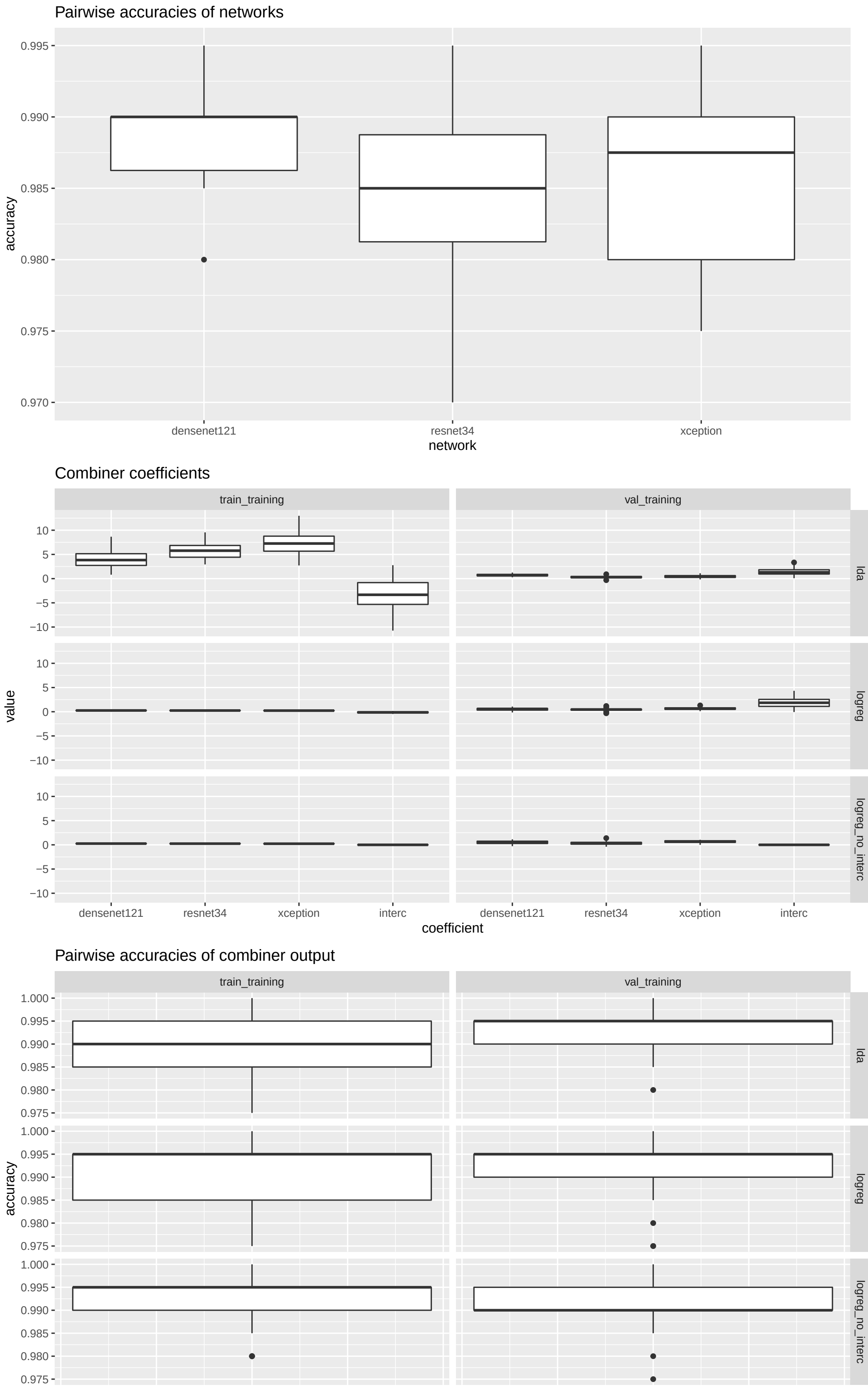


Classes 3 vs 6
Pairwise accuracies of networks
accuracy
densenet121
Combiner coefficients
train_training
value
resnet34
Pairwise accuracies of combiner

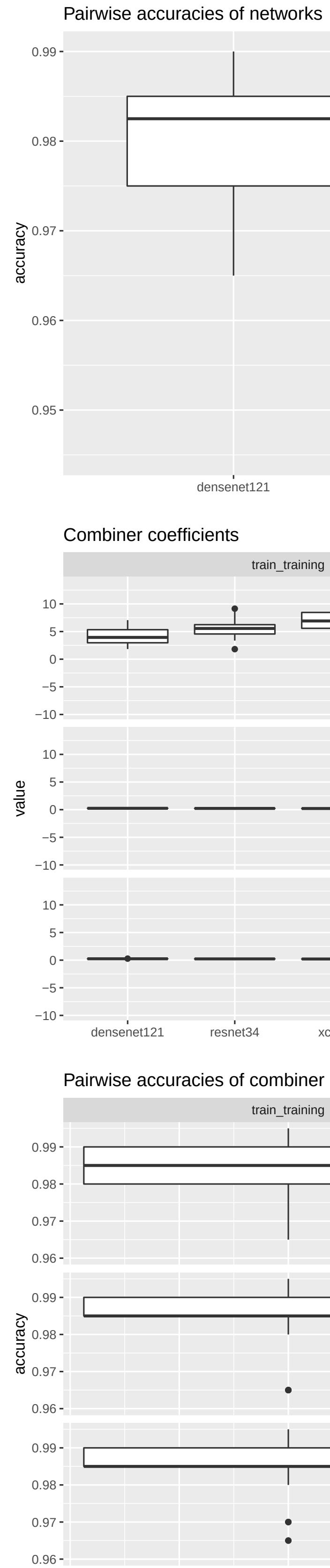

Classes 3 vs 7
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
densenet121
resnet34
xception
interc
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy


Classes 3 vs 8
Pairwise accuracies of networks
accuracy
densenet121
Combiner coefficients
train_training
value
densenet121
resnet34
Pairwise accuracies of combiner
train_training
accuracy

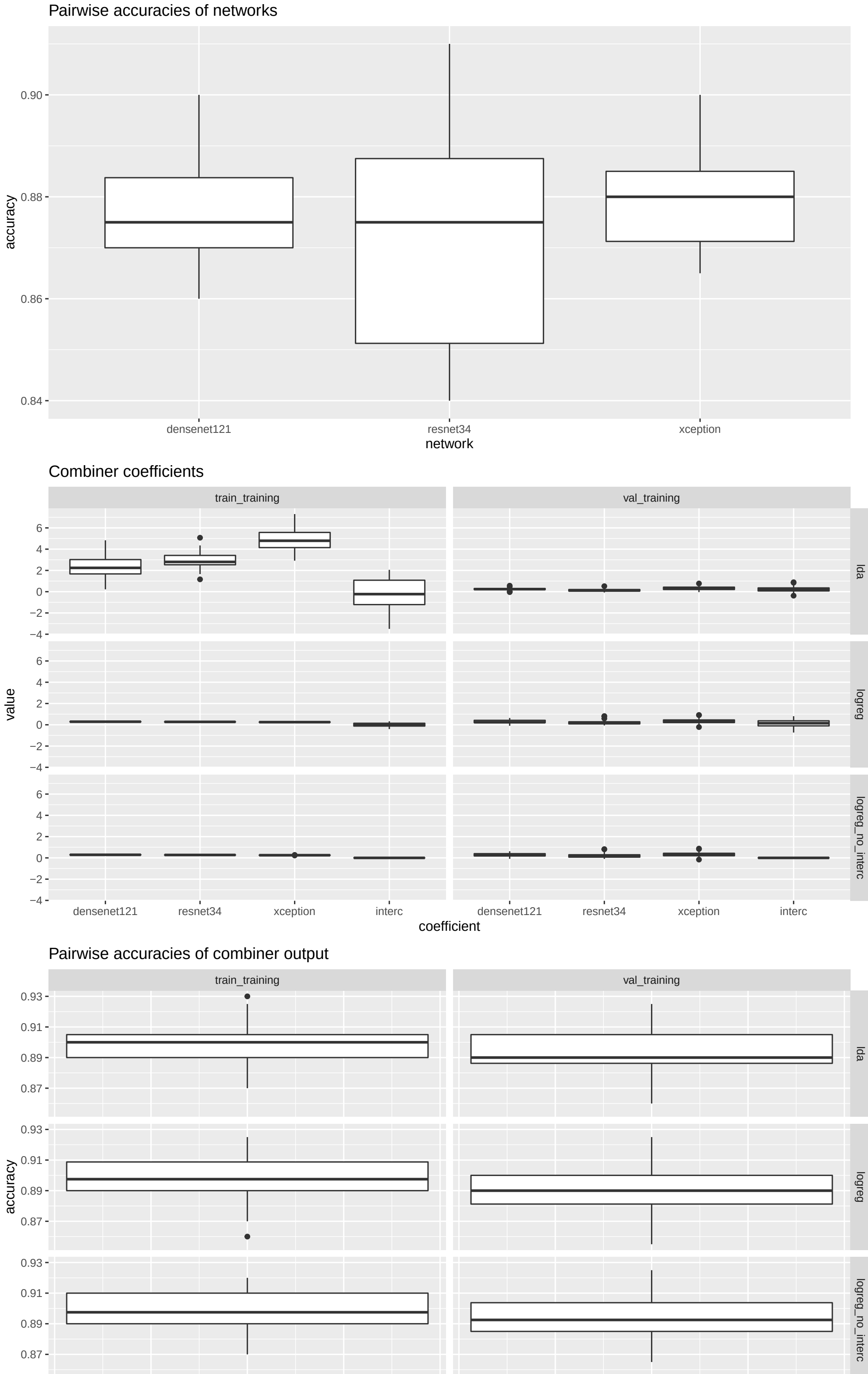
Classes 3 vs 9



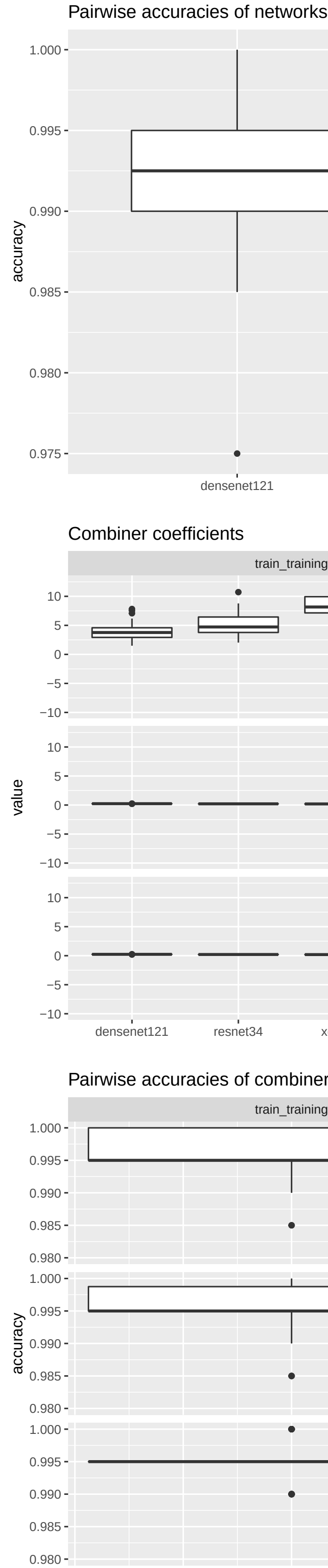
Classes 3 vs 10



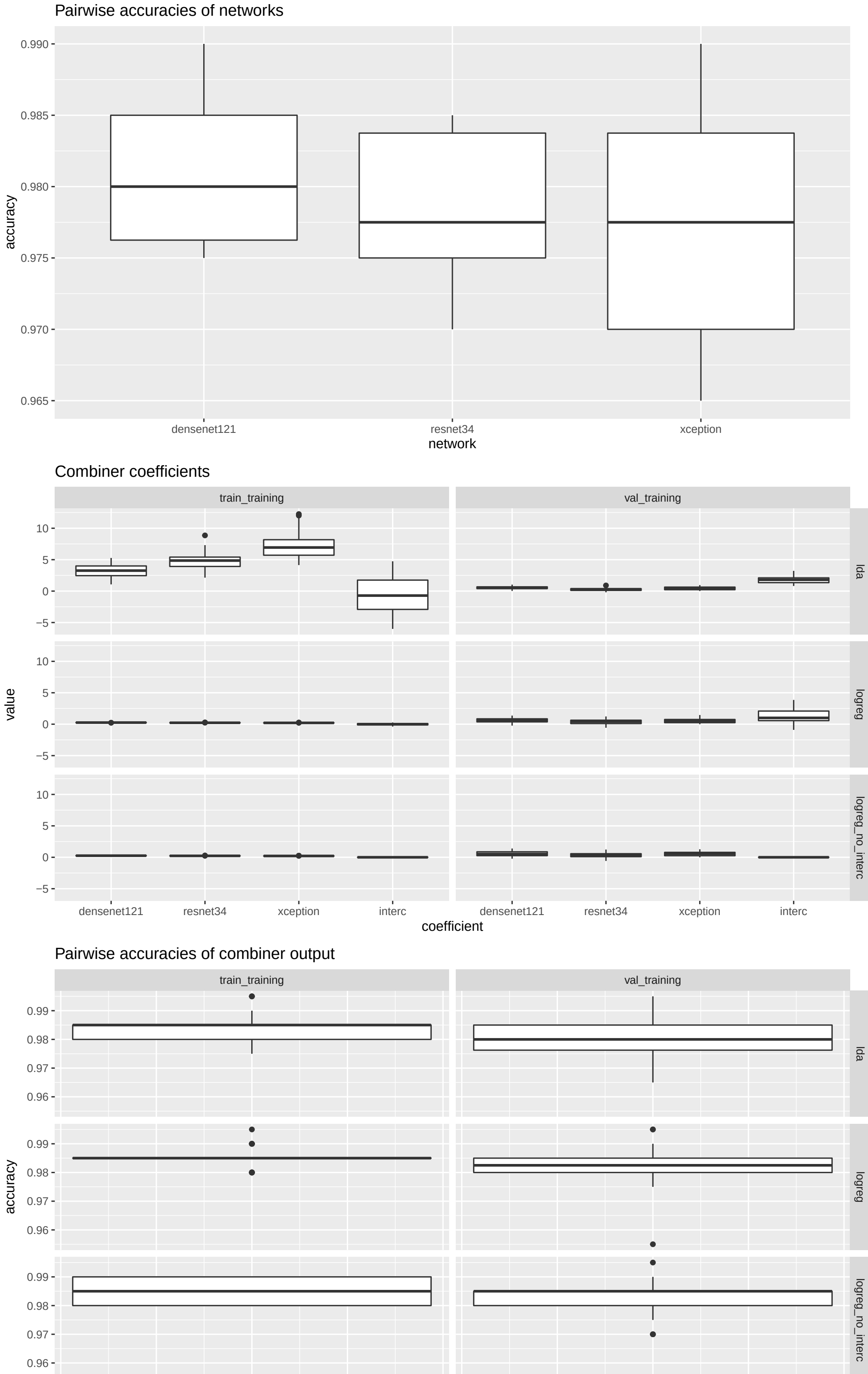
Classes 4 vs 5



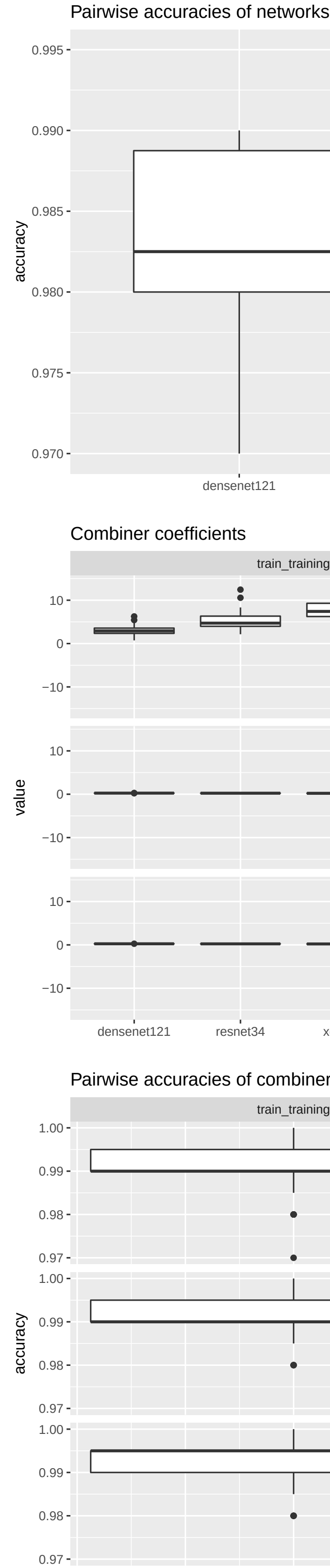
Classes 4 vs 6

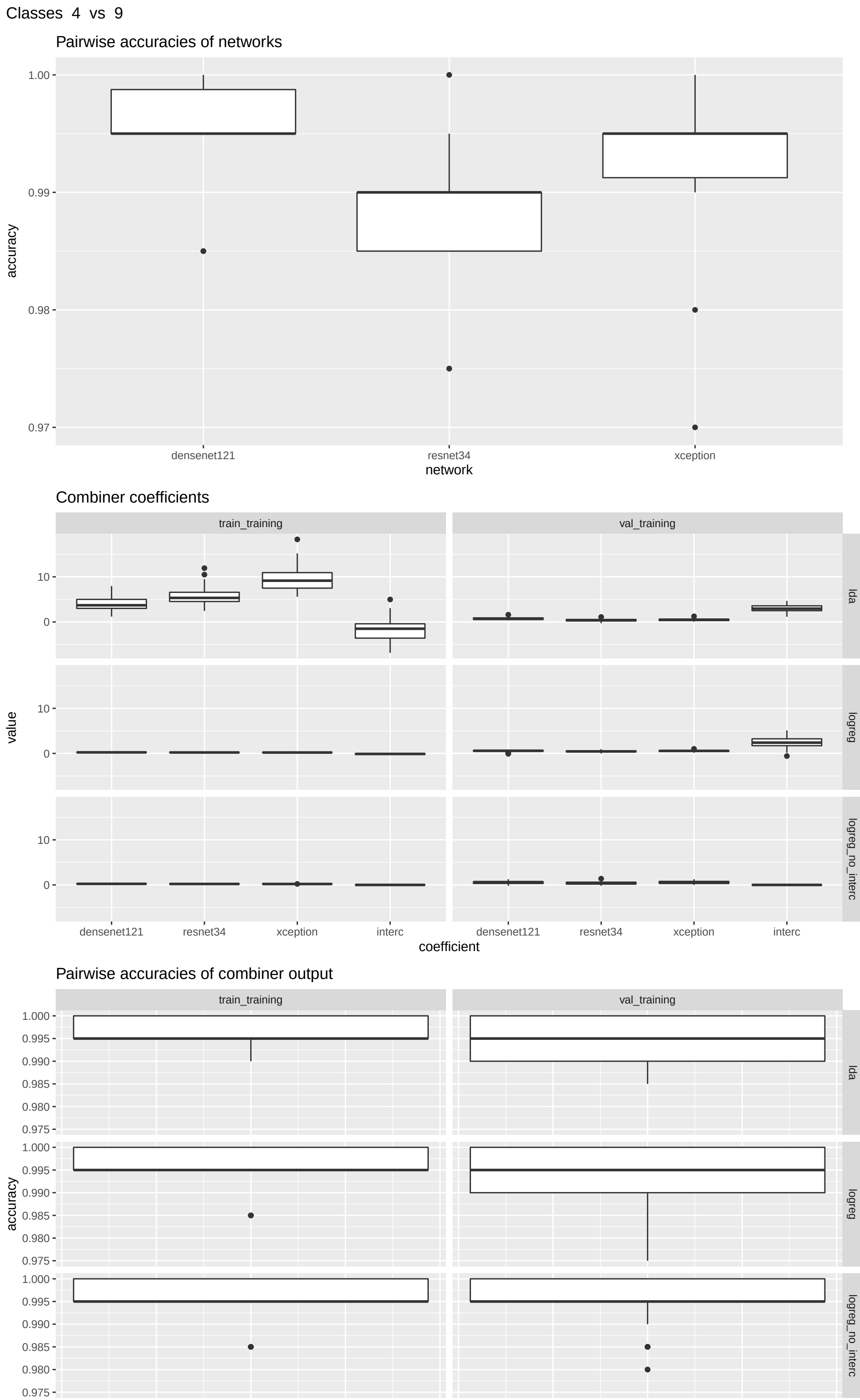


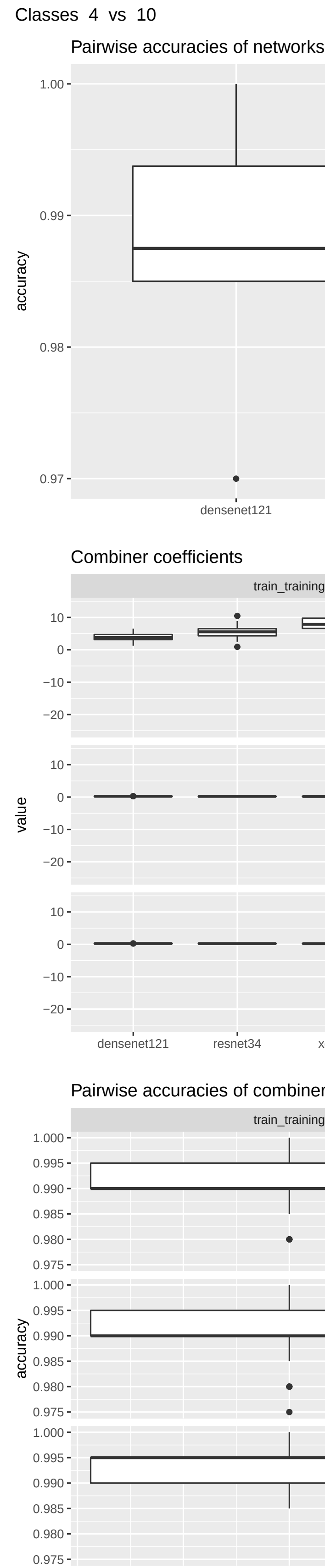
Classes 4 vs 7



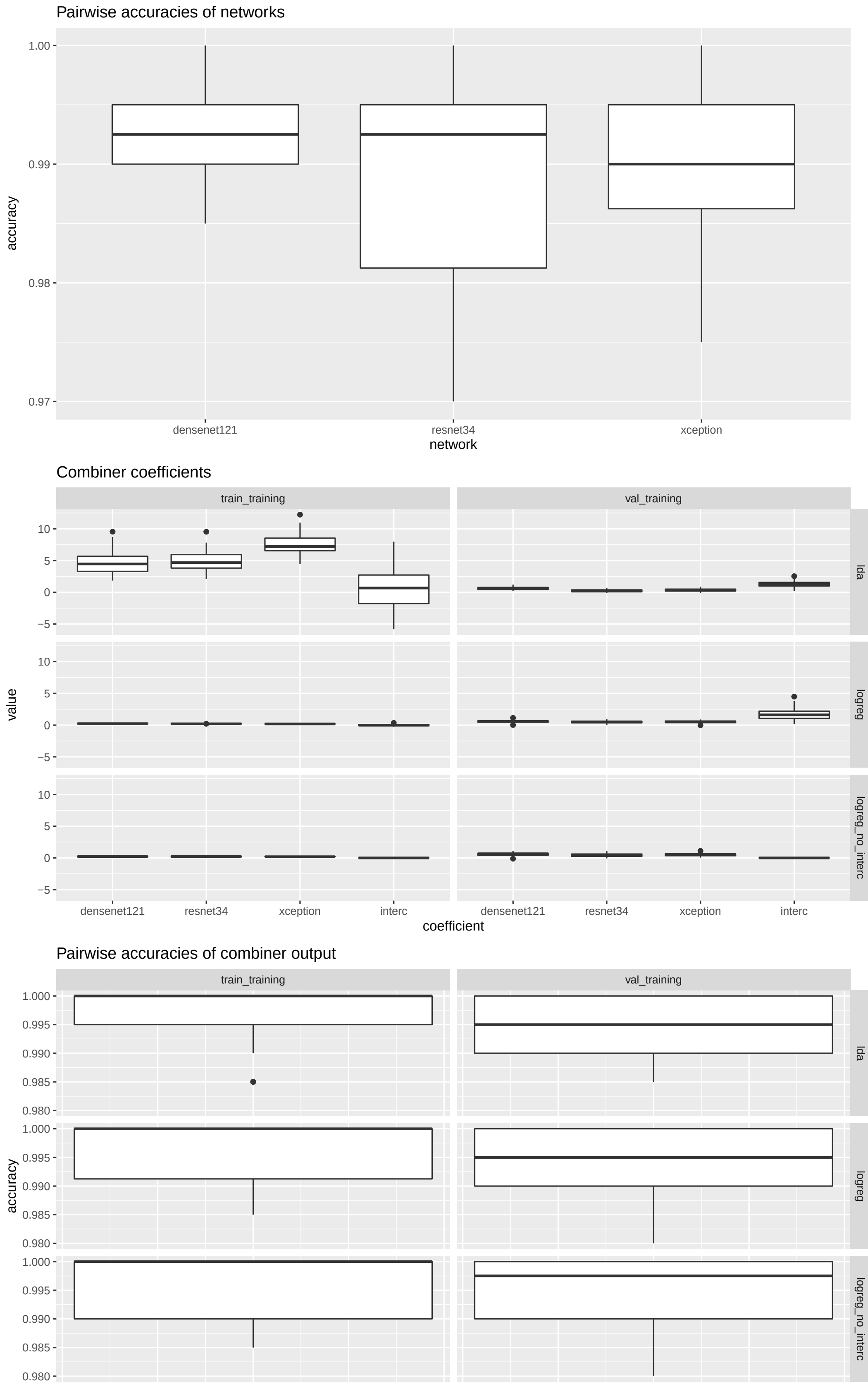
Classes 4 vs 8



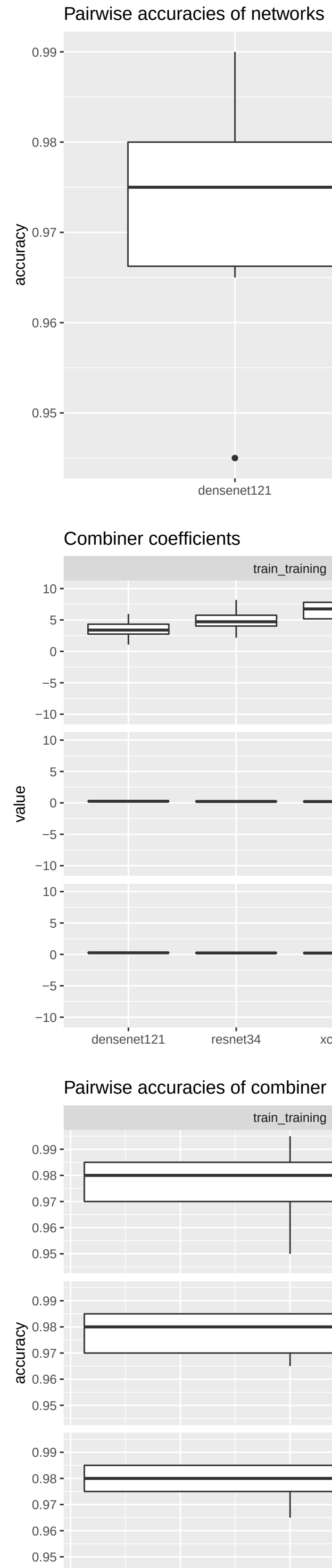

Classes 4 vs 9
Pairwise accuracies of networks
accuracy
1.00
0.99
0.98
0.97
densenet121
resnet34
xception
network
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
10
0
densenet121
resnet34
xception
interc
coefficient
Pairwise accuracies of combiner output
accuracy
1.000
0.995
0.990
0.985
0.980
0.975


Classes 4 vs 10
Pairwise accuracies of networks
accuracy
1.00
0.99
0.98
0.97
densenet121
Combiner coefficients
train_training
value
10
0
−10
−20
densenet121
resnet34
Pairwise accuracies of combiner
accuracy
1.000
0.995
0.990
0.985
0.980
0.975

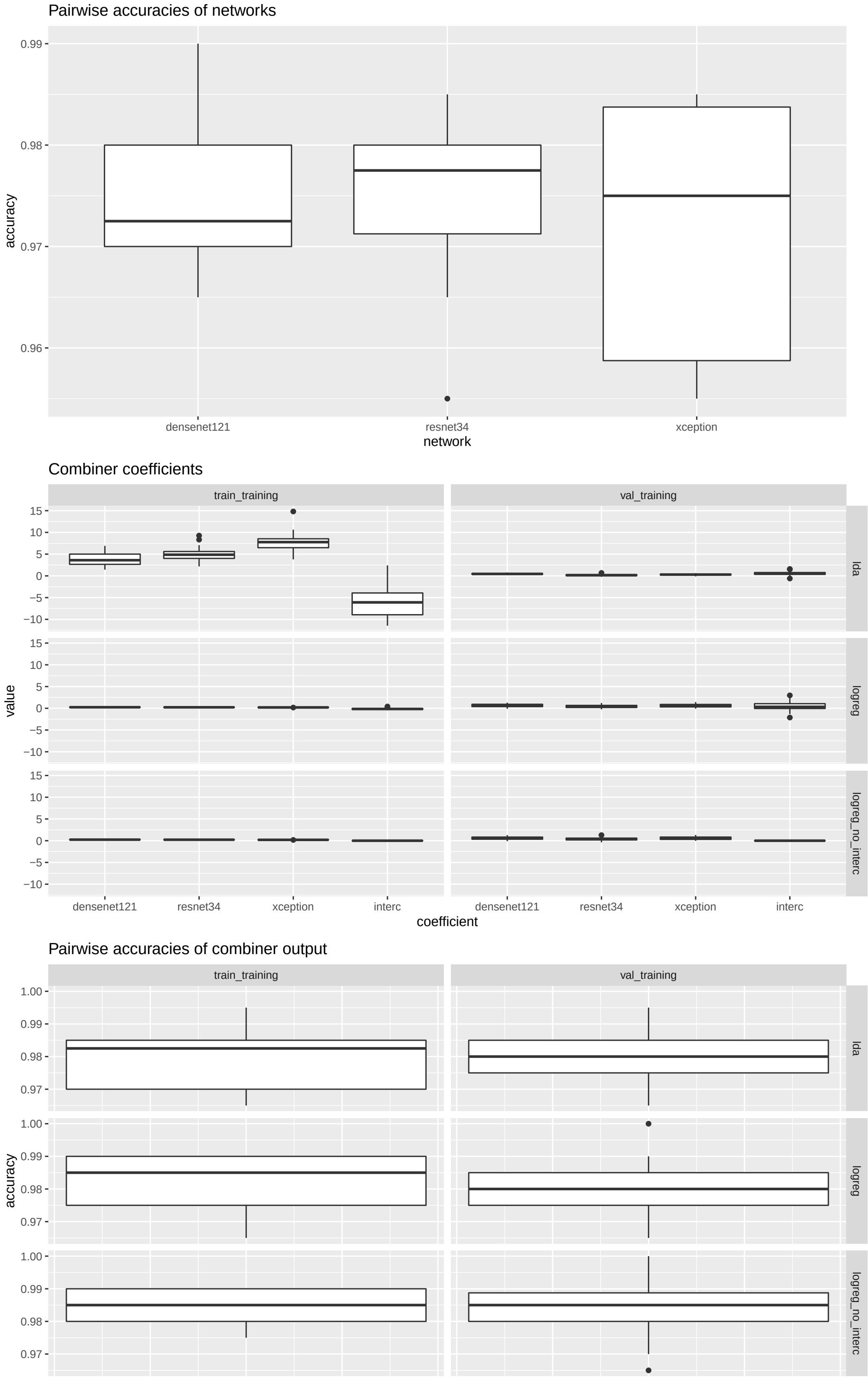
Classes 5 vs 6



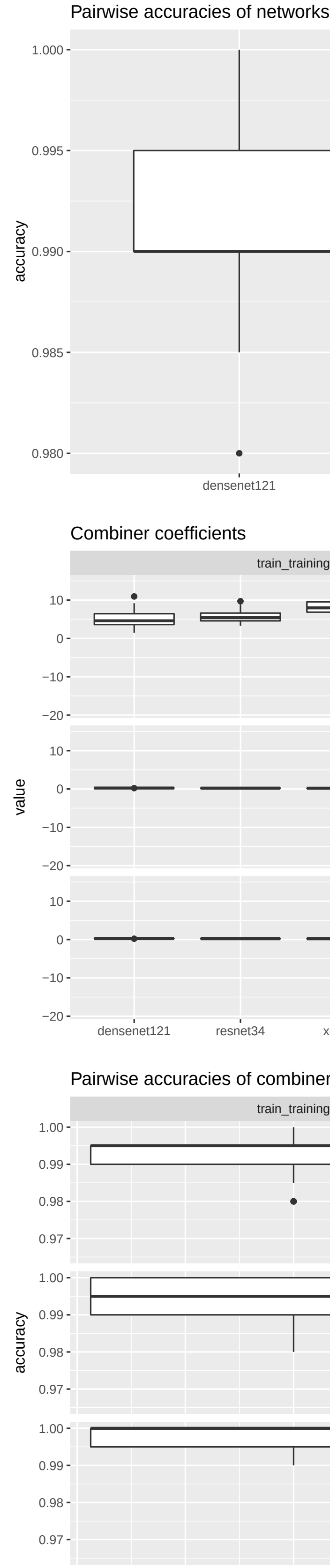
Classes 5 vs 7

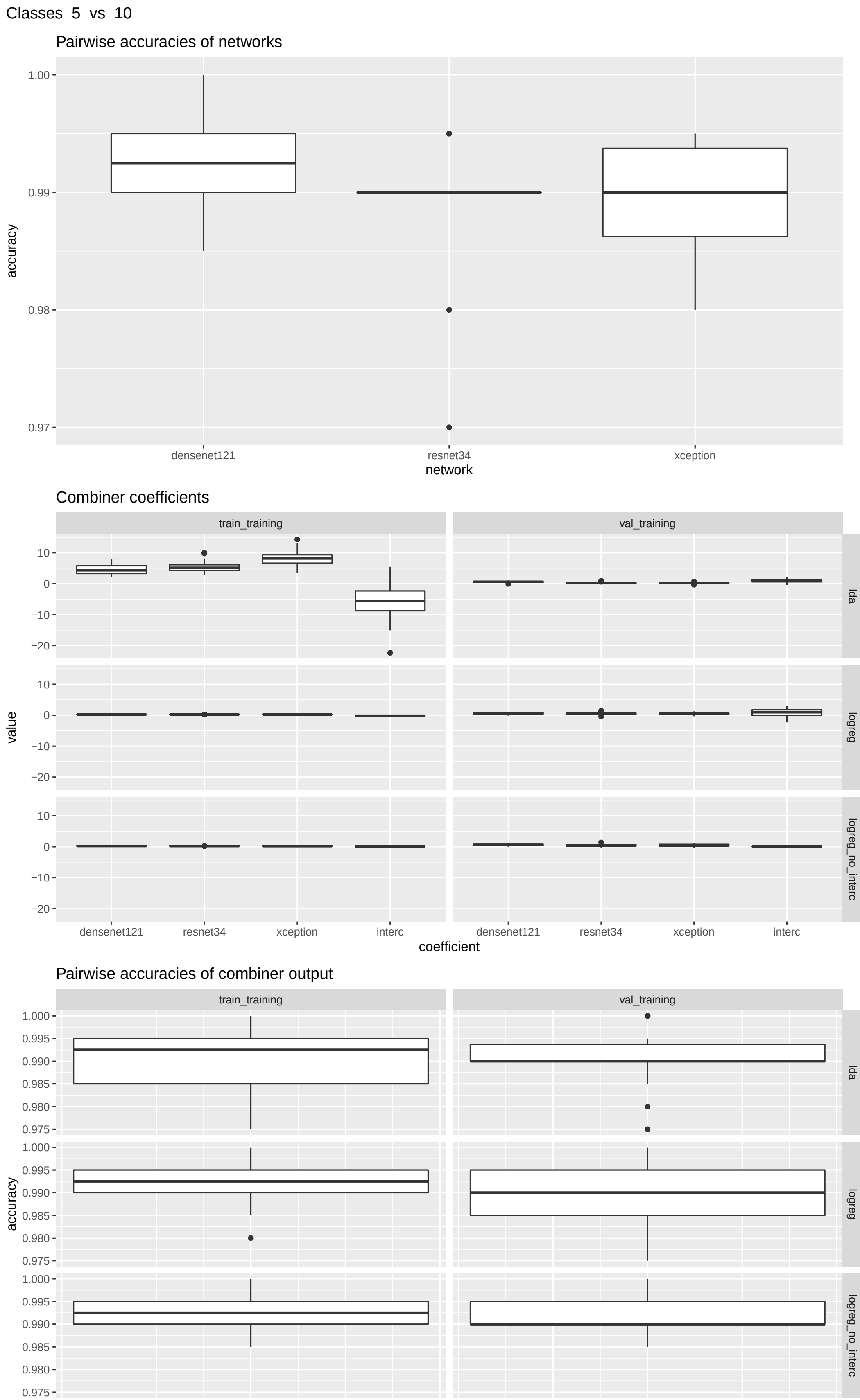


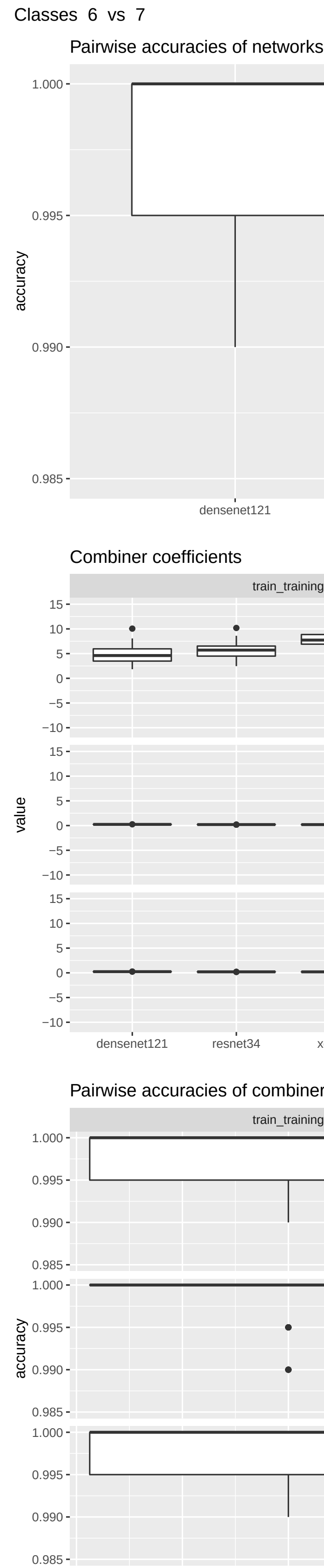
Classes 5 vs 8

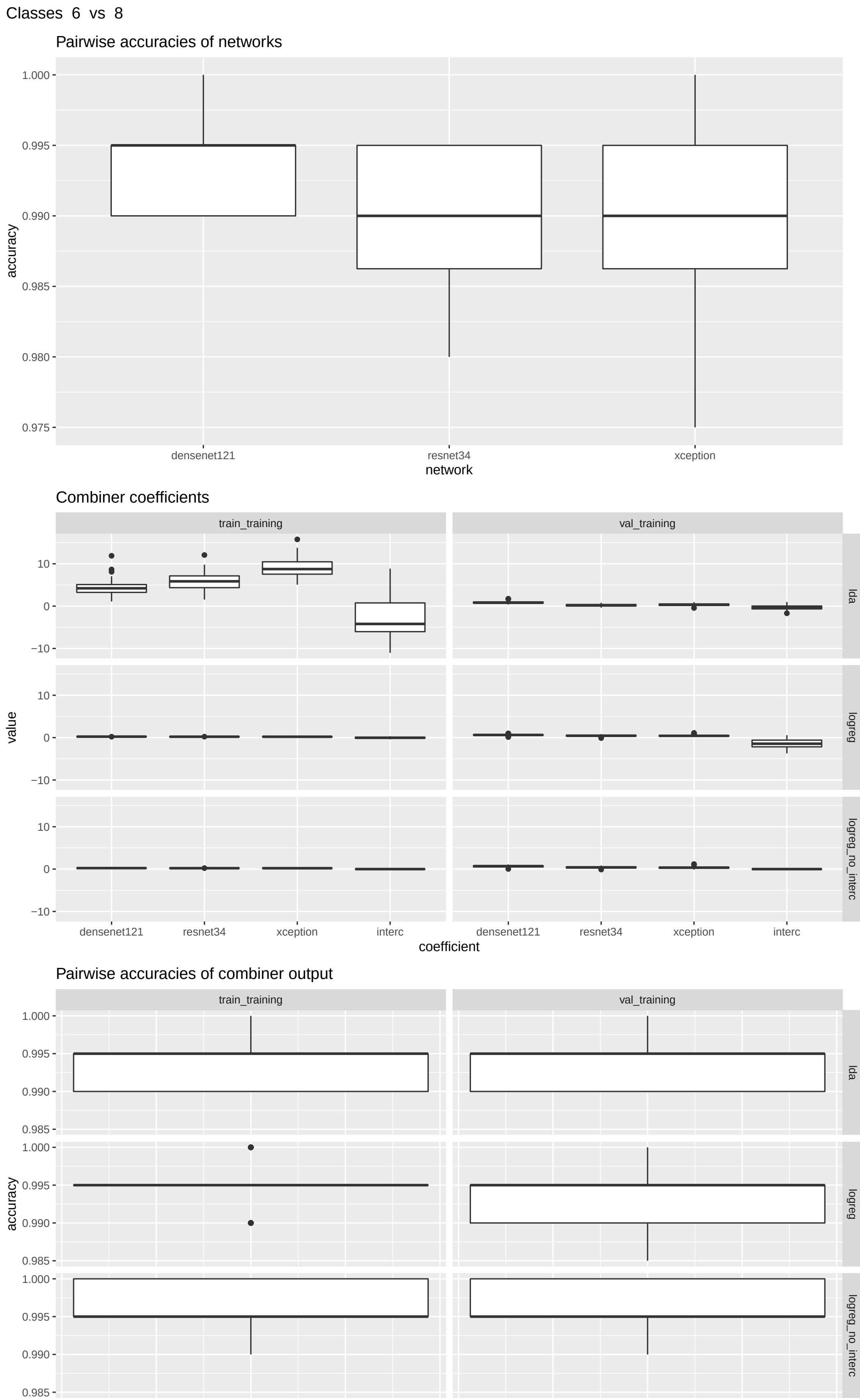


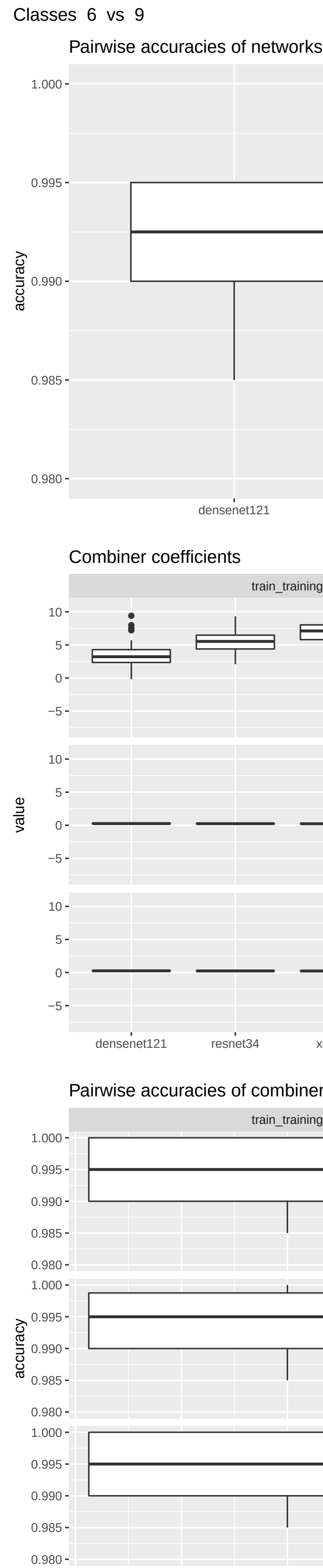
Classes 5 vs 9



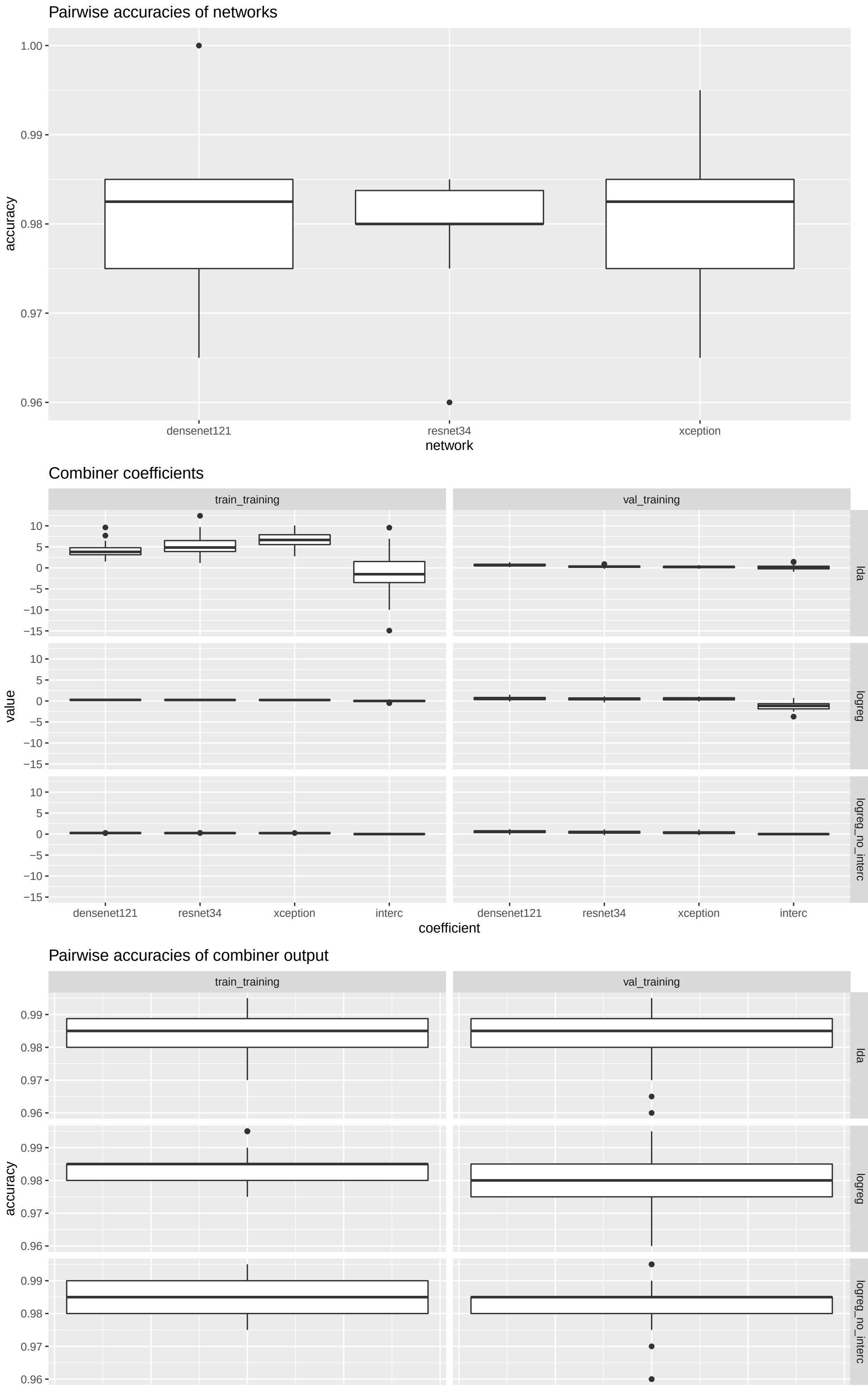

Classes 5 vs 10
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
Combiner coefficients
train_training
val_training
value
coefficient
interc
lda
logreg
logreg_no_interc
Pairwise accuracies of combiner output


Classes 6 vs 7
Pairwise accuracies of networks
accuracy
densenet121
Combiner coefficients
train_training
value
resnet34
Pairwise accuracies of combiner

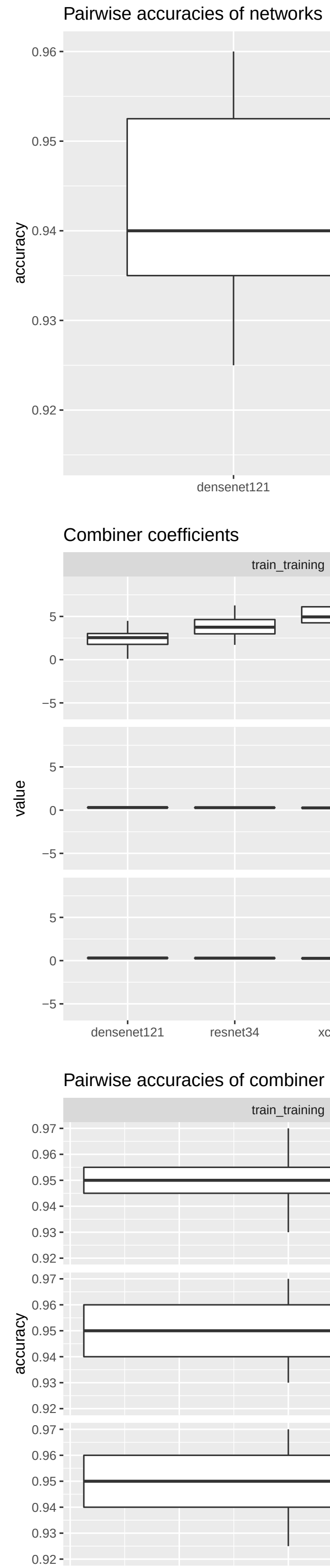

Classes 6 vs 8
Pairwise accuracies of networks
accuracy
network
densenet121
resnet34
xception
Combiner coefficients
train_training
val_training
lda
logreg
logreg_no_interc
value
coefficient
densenet121
resnet34
xception
interc
Pairwise accuracies of combiner output
train_training
val_training
lda
logreg
logreg_no_interc
accuracy


Classes 6 vs 9
Pairwise accuracies of networks
accuracy
densenet121
Combiner coefficients
train_training
value
densenet121
resnet34
accuracy

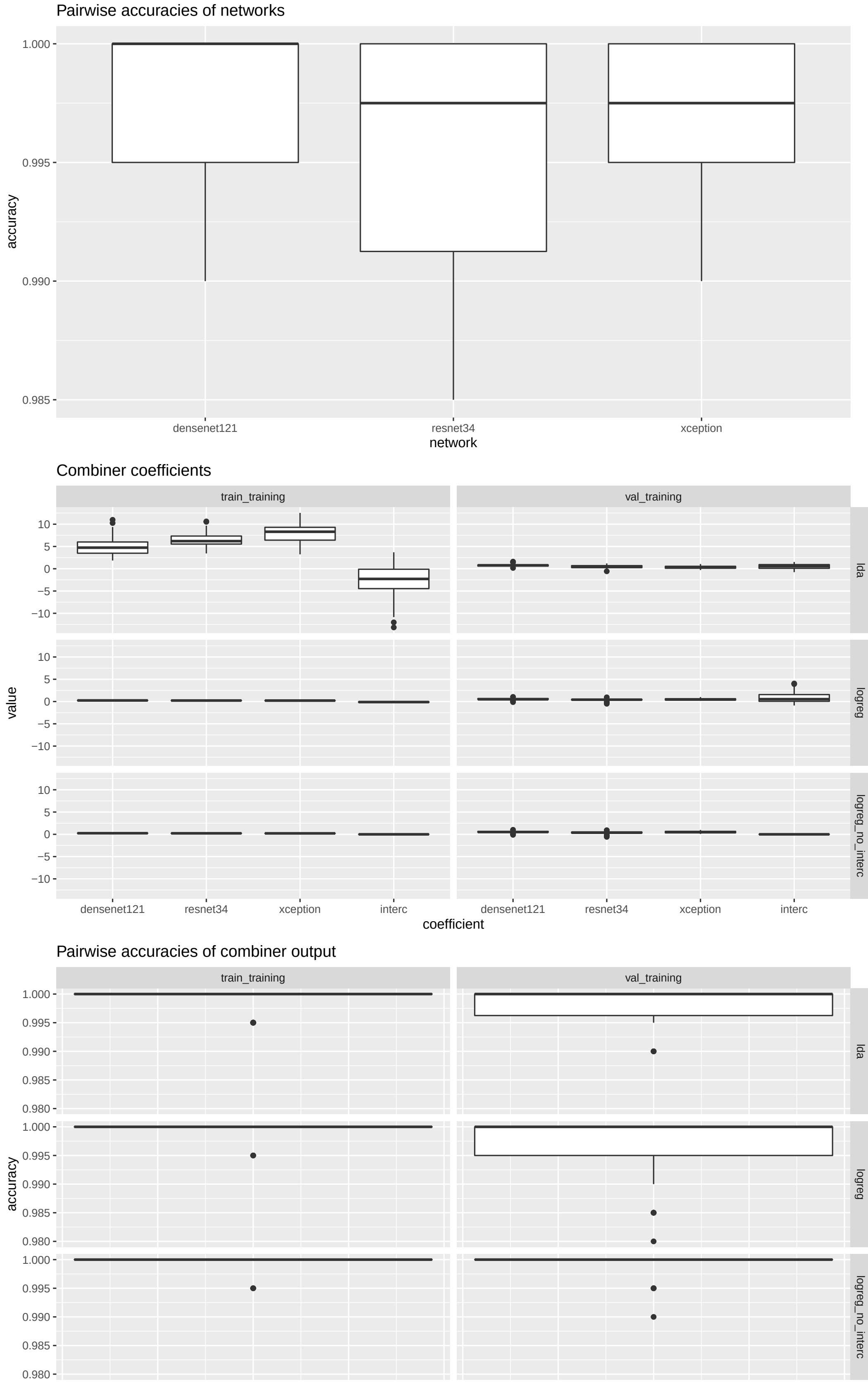
Classes 6 vs 10



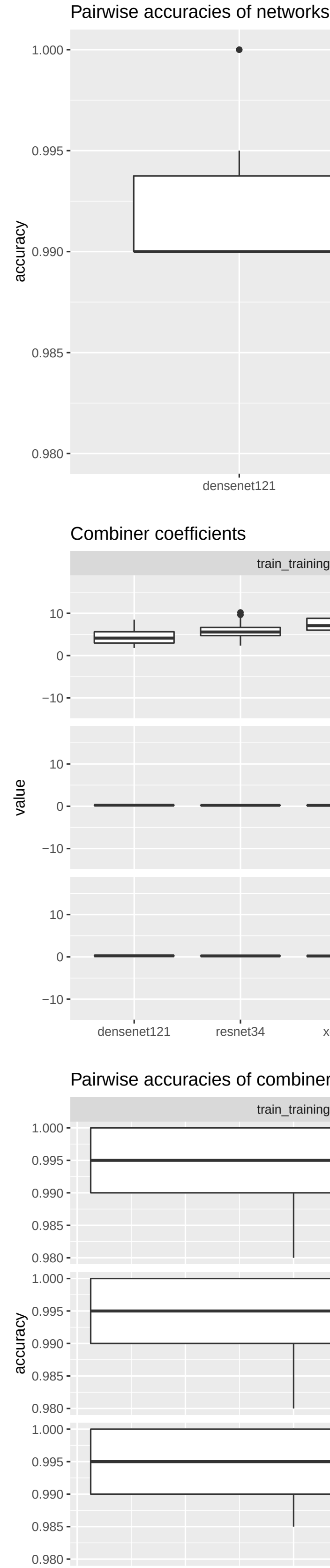
Classes 7 vs 8



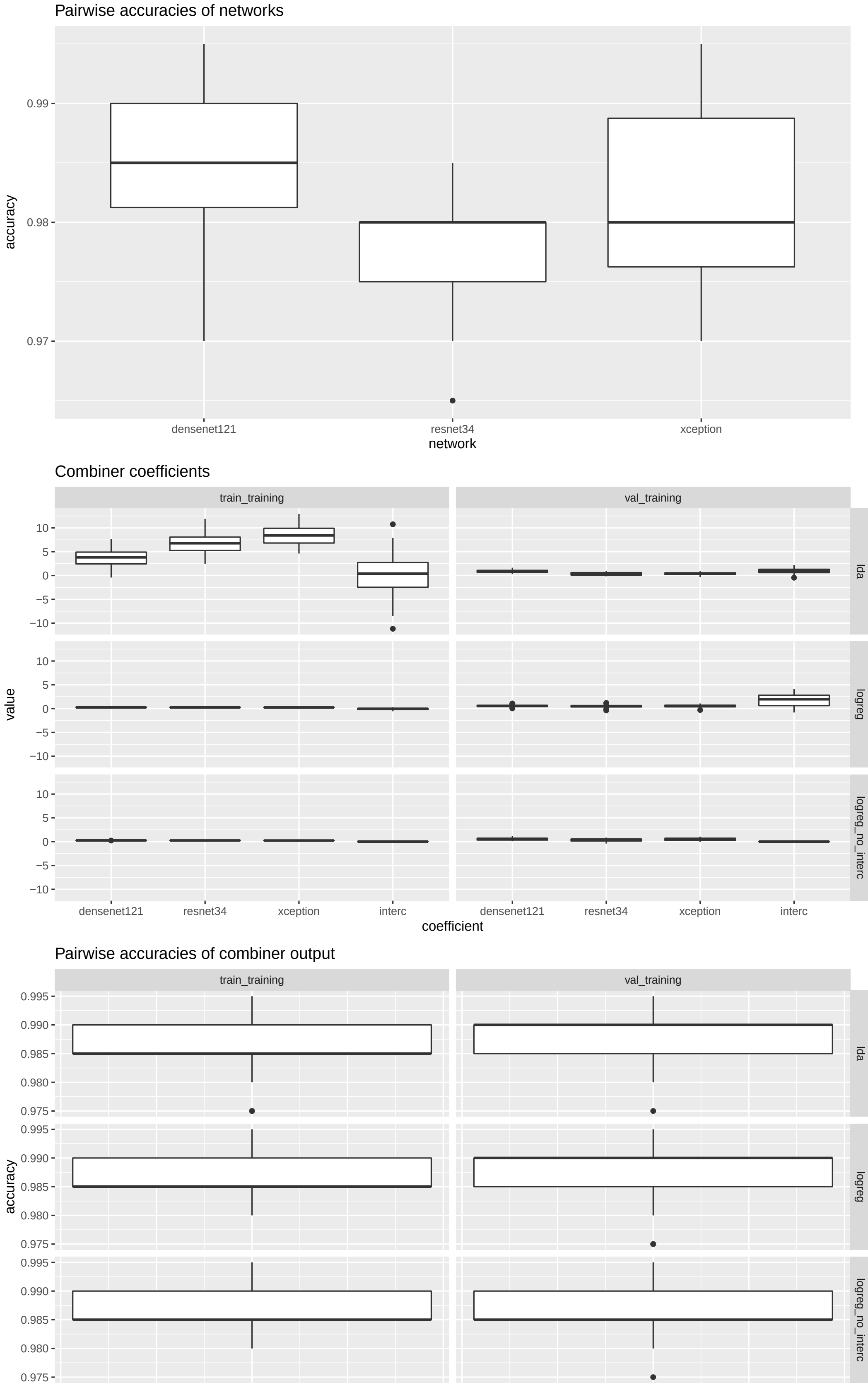
Classes 7 vs 9



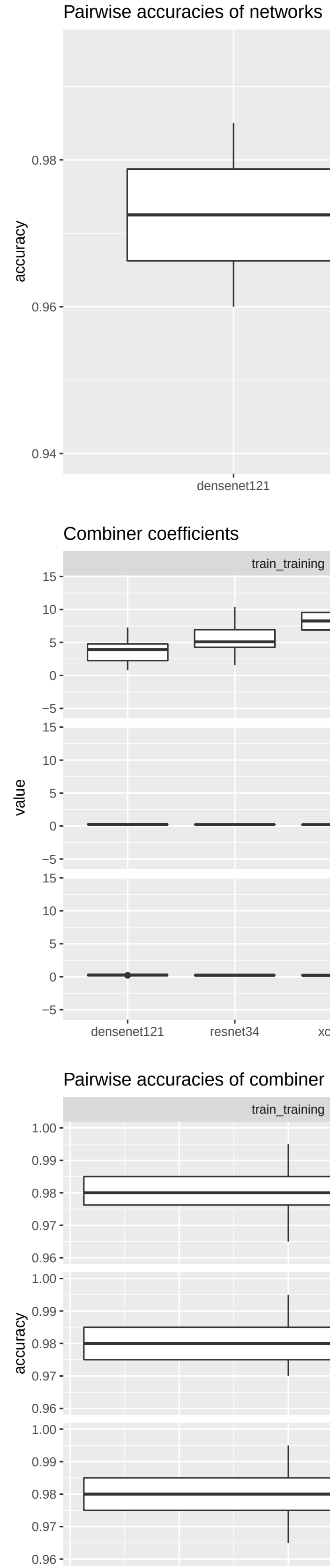
Classes 7 vs 10



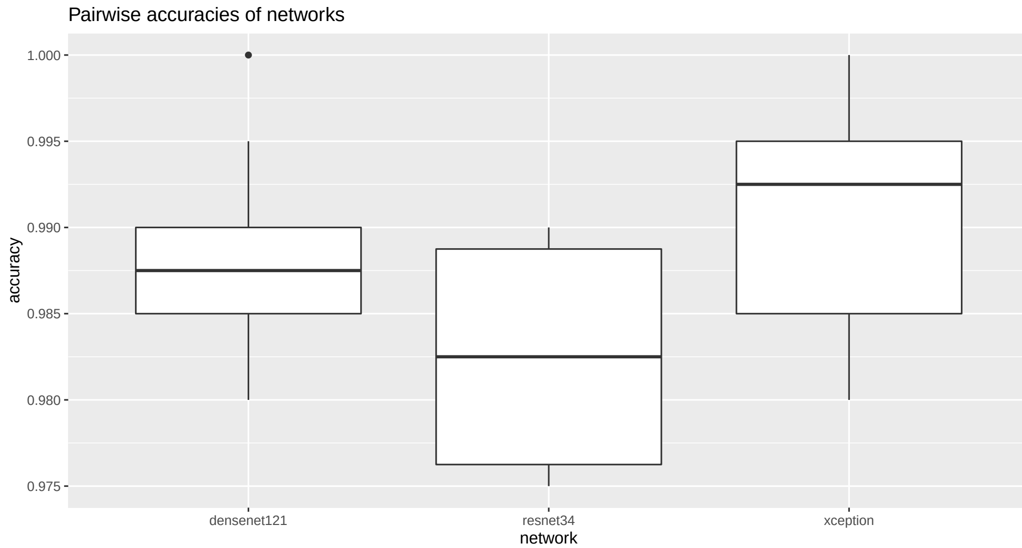
Classes 8 vs 9

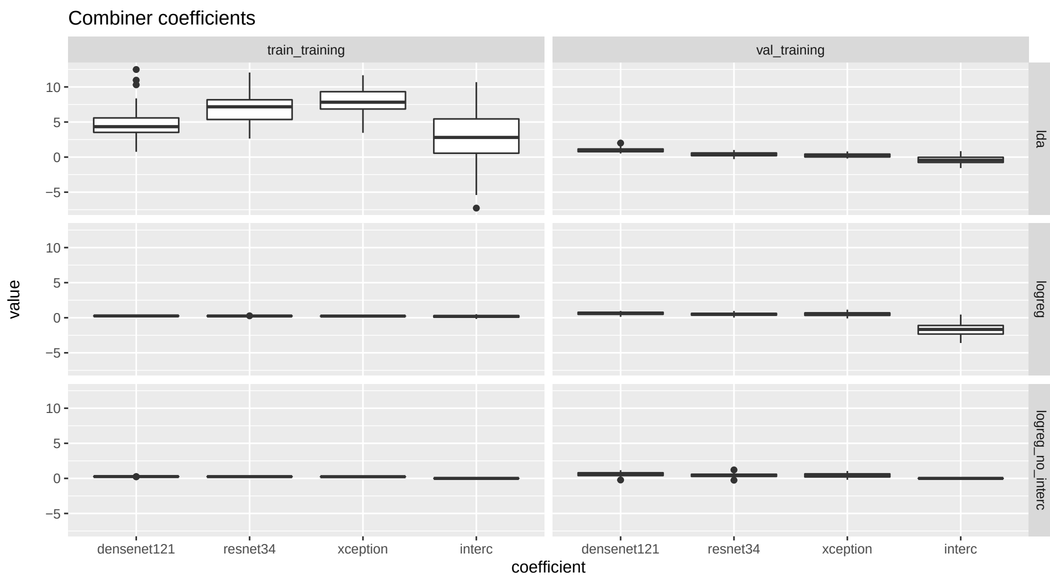


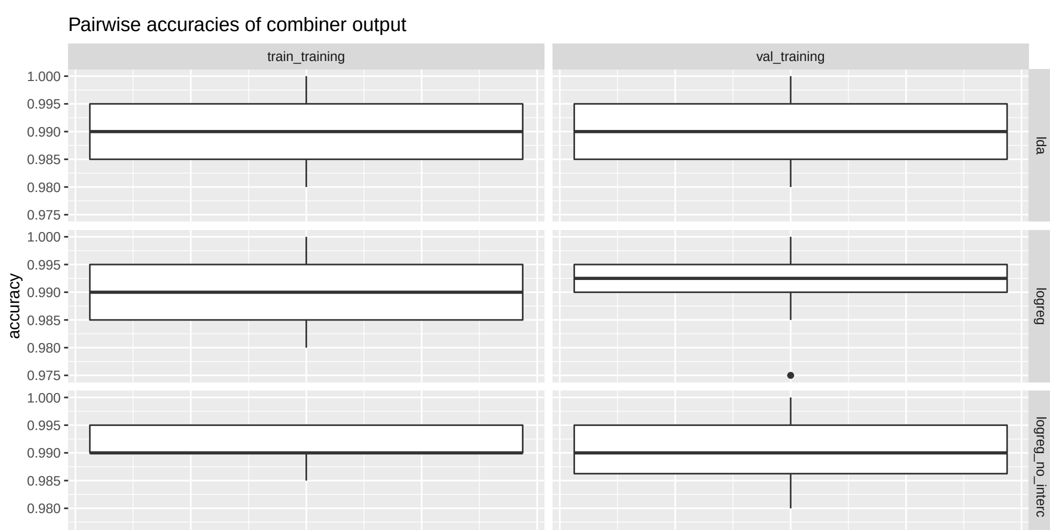
Classes 8 vs 10



Classes 9 vs 10







Here are too many class pair combinations to draw them all. TODO: think of some criteria to draw interesting ones.