

## LDA training on random subsets

```
library(ggplot2)
```

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

```
library(tidyr)
```

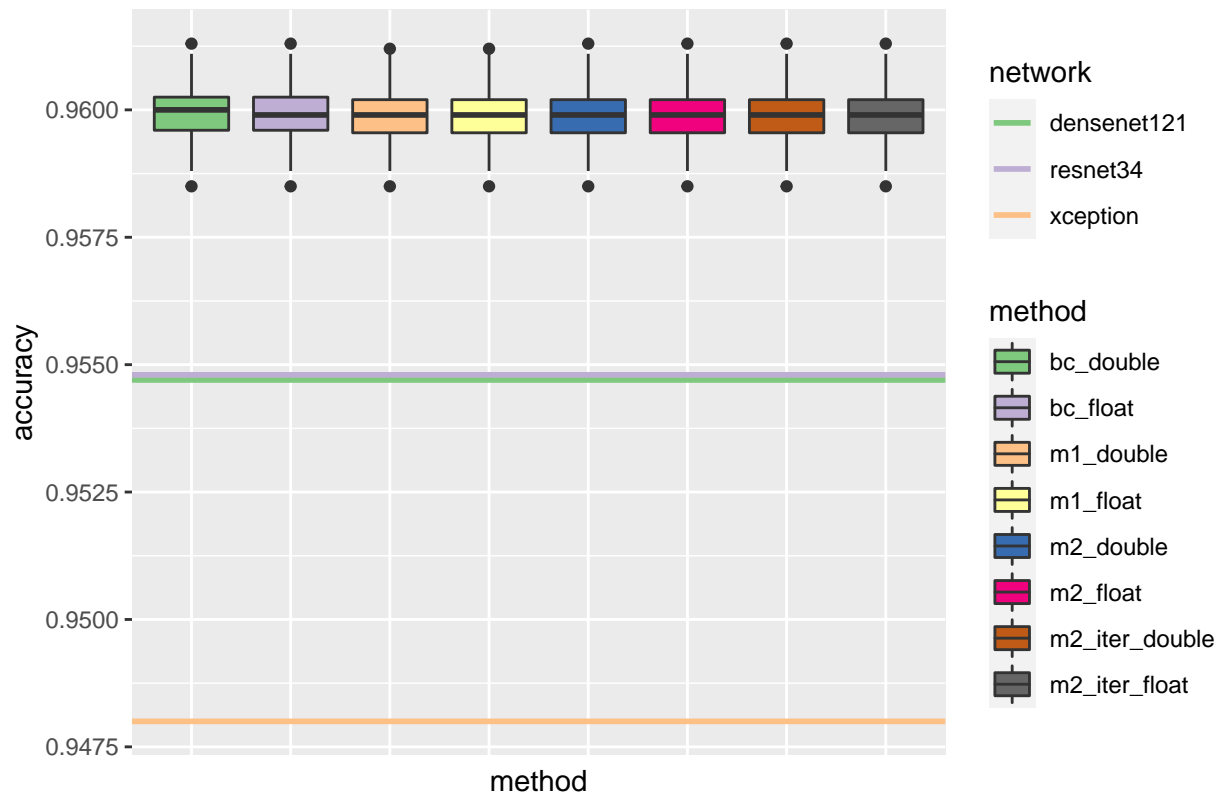
```
## Warning: package 'tidyr' was built under R version 4.0.5
```

Experiment code is in the file `training_subsets_experiment.py`. Experiment on CIFAR10 dataset. This experiment trains `WeightedLDAEnsemble` on various disjoint subsets of data on which neural networks were trained. Size of these subsets is 500, therefore each class has 50 samples and each LDA is trained on 100 samples. Three different coupling methods are used: method one and two from (Wu, Lin, and Weng 2004) and Bayes covariant method (Šuch and Barreda 2016). Goal of this experiment is to determine, whether usage of different training sets for LDA has important effects on the performance of the ensemble.

```
acc_ens_subsets <- read.csv("../data/data_train_val_c10/1/exp_subsets_train_outputs/accuracies.csv")
acc_nets <- read.csv("../data/data_train_val_c10/1/exp_subsets_train_outputs/net_accuracies.csv")
```

```
box_acc <- ggplot() + geom_boxplot(data=acc_ens_subsets, mapping=aes(x=method, y=accuracy, fill=method),
  geom_hline(data=acc_nets, mapping=aes(yintercept=accuracy, color=network), size=1) +
  scale_fill_brewer(type="qual") +
  scale_color_brewer(type="qual") +
  ggtitle("Accuracy of coupling methods compared to accuracy of networks") +
  theme(axis.ticks.x=element_blank(), axis.text.x = element_blank()))
box_acc
```

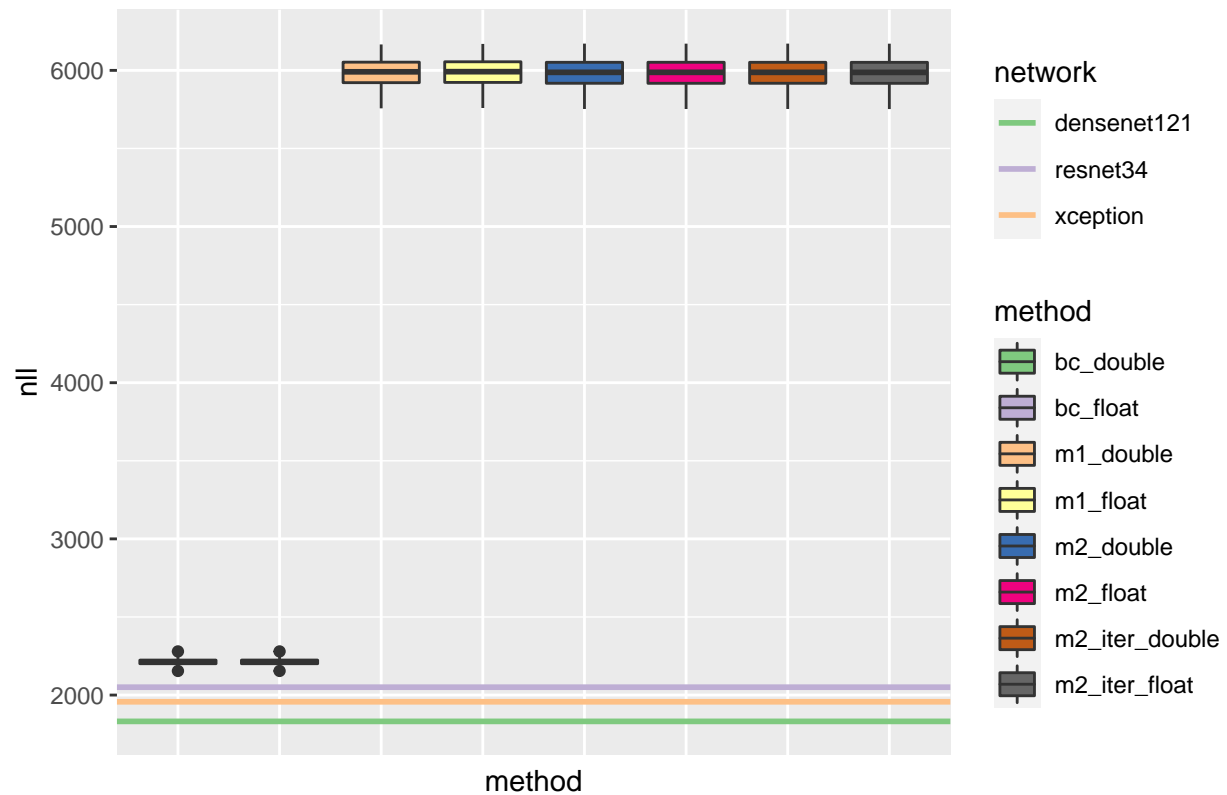
Accuracy of coupling methods compared to accuracy of networks



All coupling methods perform similarly. Also we can observe, that the median improvement of the ensemble against the best network is much larger than the 1.5 multiple of IQR.

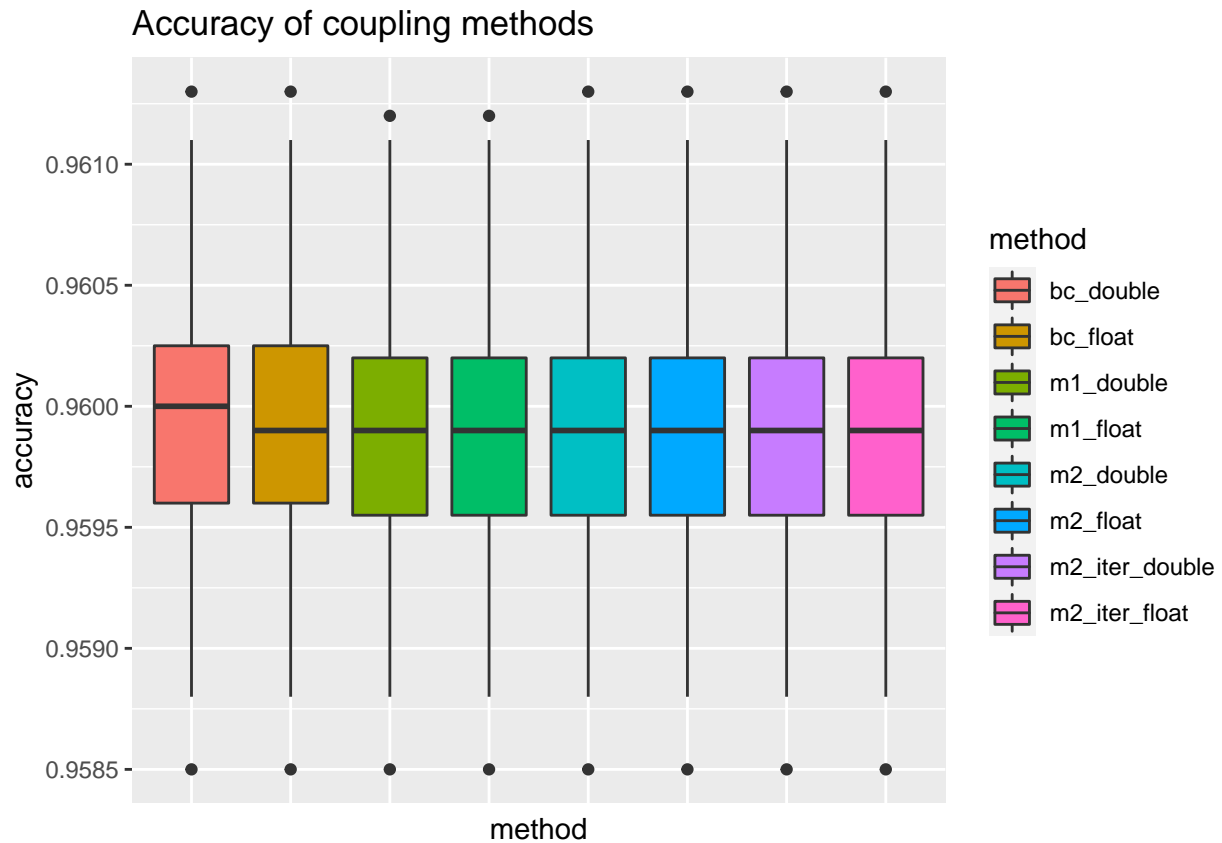
```
box_nll <- ggplot() + geom_boxplot(data=acc_ens_subsets, mapping=aes(x=method, y=nll, fill=method)) +
  geom_hline(data=acc_nets, mapping=aes(yintercept=nll, color=network), size=1) +
  scale_fill_brewer(type="qual") +
  scale_color_brewer(type="qual") +
  ggtitle("NLL of coupling methods compared to accuracy of networks") +
  theme(axis.ticks.x=element_blank(), axis.text.x = element_blank())
box_nll
```

NLL of coupling methods compared to accuracy of networks



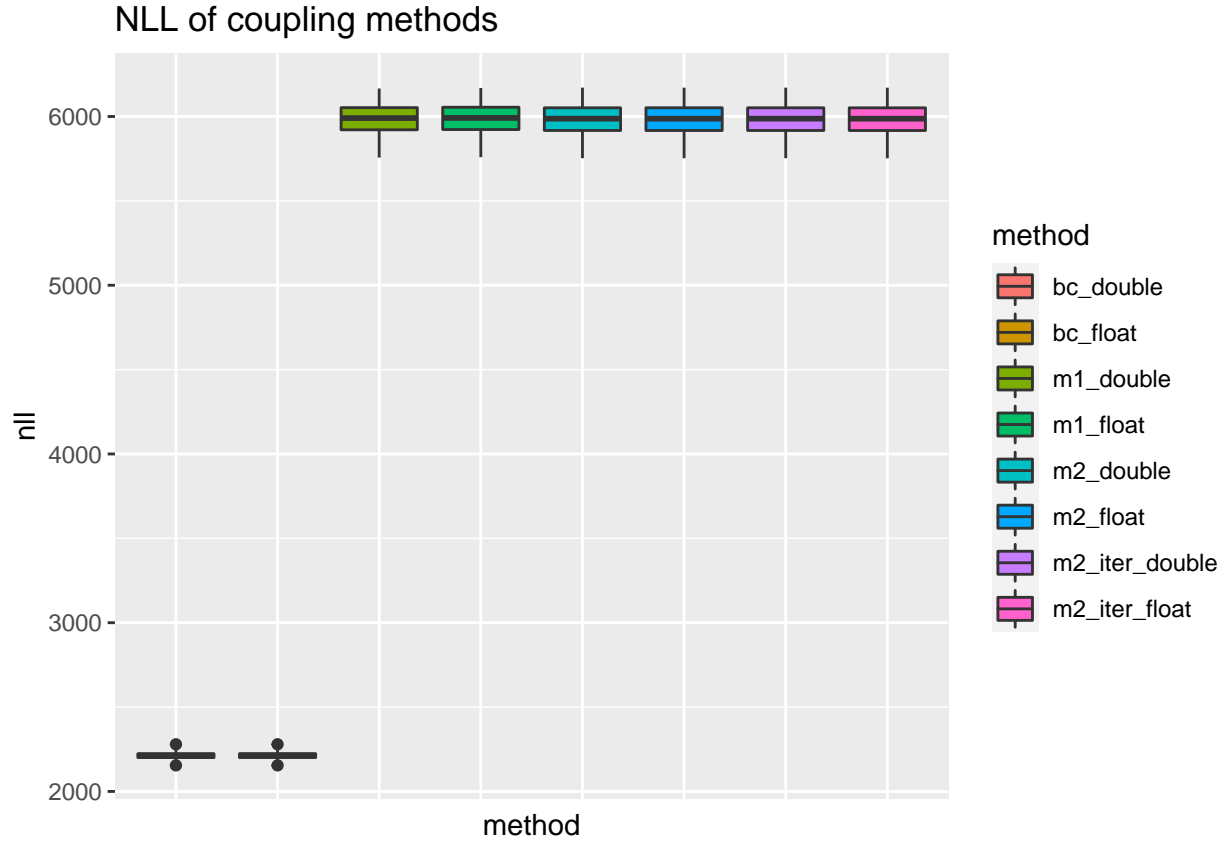
Ensembles have worse nll as the combined networks. This is due to ensemble producing close to zero or zero probability for the correct label in several test samples. Needs further attention.

```
box2_acc <- ggplot() + geom_boxplot(data=acc_ens_subsets, mapping=aes(x=method, y=accuracy, fill=method),
  ggtitle("Accuracy of coupling methods") +
  theme(axis.ticks.x=element_blank(), axis.text.x = element_blank())
box2_acc
```



All coupling methods perform similarly, regardless of the single or double precision. A little difference is visible for bc method, where median accuracy for double precision is higher than for float precision.

```
box2_nll <- ggplot() + geom_boxplot(data=acc_ens_subsets, mapping=aes(x=method, y=nll, fill=method)) +
  ggtitle("NLL of coupling methods") +
  theme(axis.ticks.x=element_blank(), axis.text.x = element_blank())
box2_nll
```



Bayes covariant coupling method seems to be doing much better than the remaining coupling methods. This may be due to fewer close to zero probabilities for the correct label produced by this method.

Šuch, Ondrej, and Santiago Barreda. 2016. “Bayes Covariant Multi-Class Classification.” *Pattern Recognition Letters* 84: 99–106.

Wu, Ting-Fan, Chih-Jen Lin, and Ruby C Weng. 2004. “Probability Estimates for Multi-Class Classification by Pairwise Coupling.” *Journal of Machine Learning Research* 5 (Aug): 975–1005.