

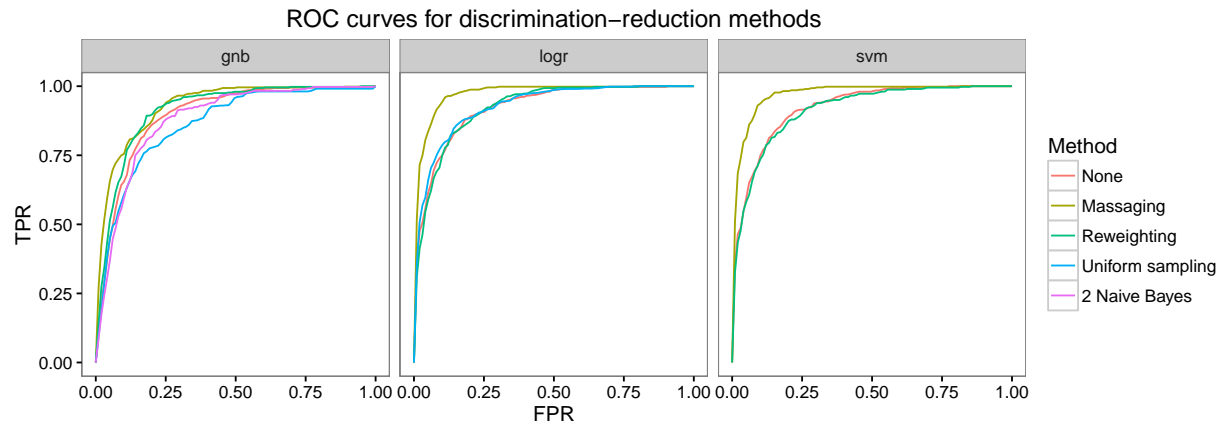
analysis

Selin Jessa

April 16, 2017

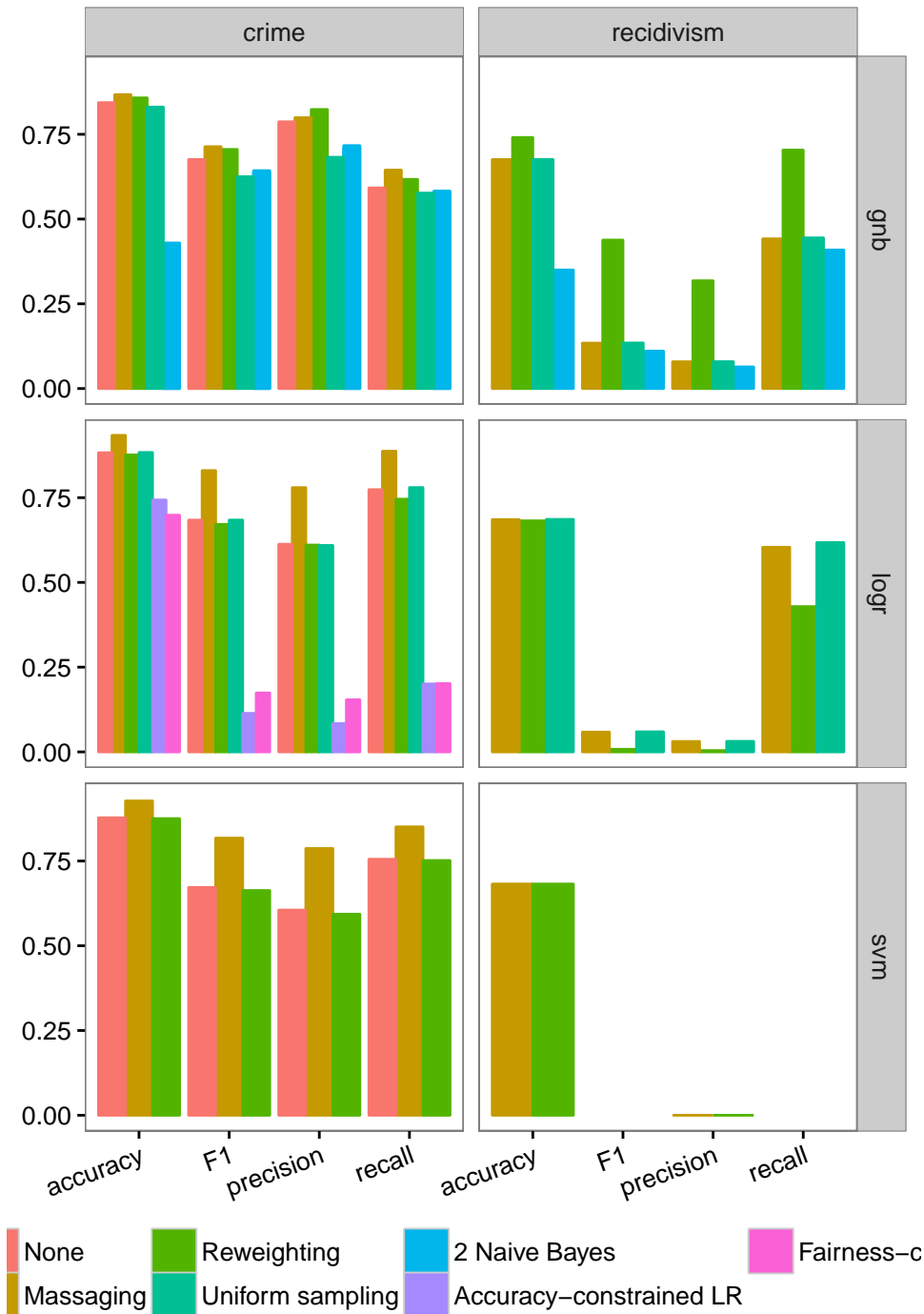
ROC curves

```
roc %>% ggplot(aes(x = fpr, y = tpr, colour = method)) +  
  geom_line() +  
  facet_wrap(~ model, drop = TRUE) +  
  xlab("FPR") + ylab("TPR") +  
  ggtitle("ROC curves for discrimination-reduction methods") +  
  guides(colour = guide_legend(title = "Method")) +  
  theme_bw() +  
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),  
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```



Accuracy

```
acc %>% gather(statistic, value, accuracy, precision, recall, F1) %>%  
  ggplot(aes(x = statistic, y = value)) +  
  geom_bar(stat = "identity", aes(fill = method, colour = method), position = "dodge") +  
  xlab(NULL) + ylab(NULL) +  
  guides(fill = guide_legend(title = "Method", nrow = 2), colour = guide_legend(title = "Method")) +  
  facet_grid(model ~ dataset, drop = TRUE) +  
  theme_bw() +  
  theme(axis.text.x = element_text(angle = 20, hjust = 1),  
        legend.position = "bottom", panel.grid.major = element_blank(),  
        panel.grid.minor = element_blank(),  
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```



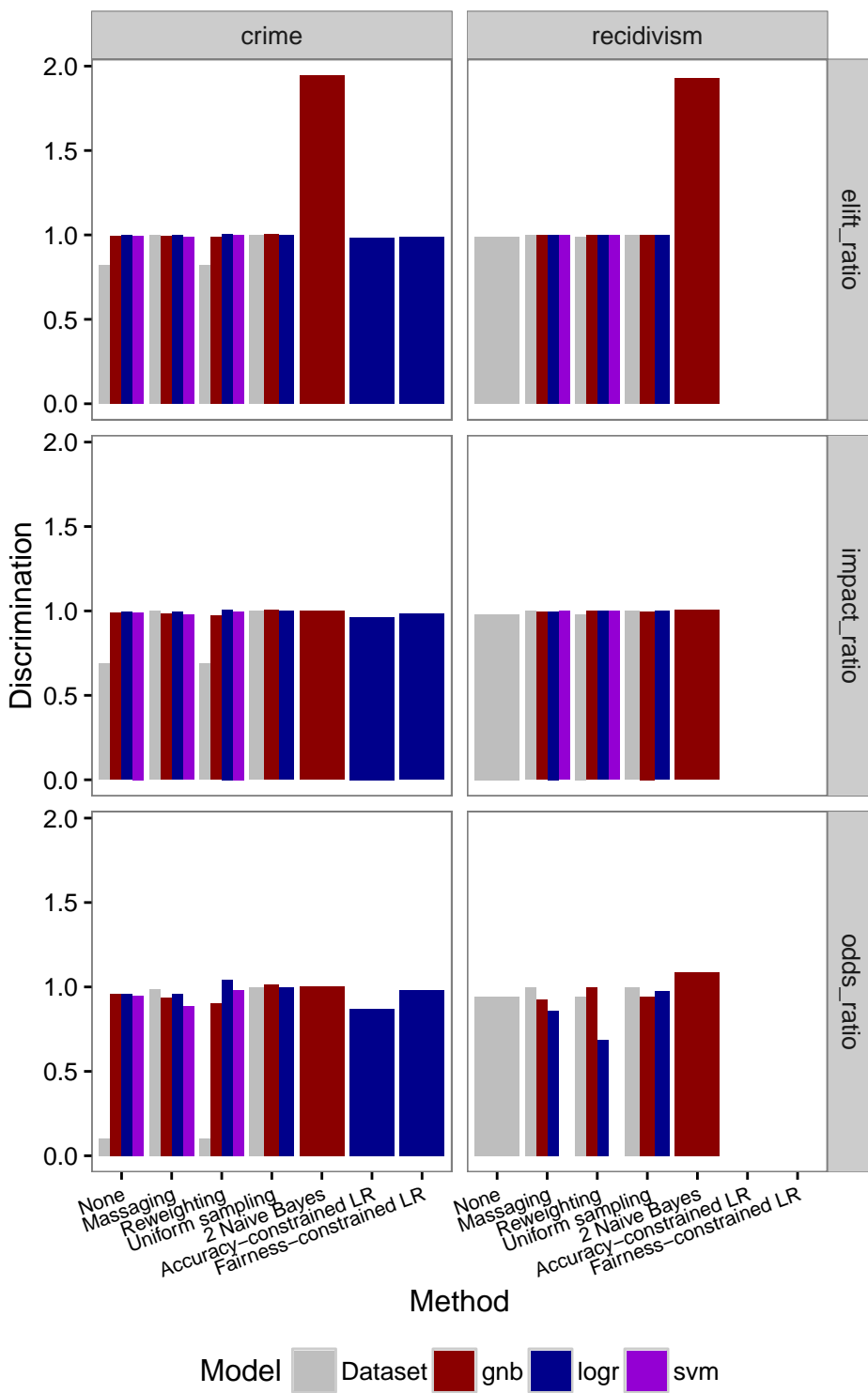
Discrimination

```
comp_tidy %>% gather(measure, value, impact_ratio, odds_ratio, elift_ratio) %>%
  ggplot(aes(x = method, y = value)) +
  geom_bar(stat = "identity", aes(fill = model), position = "dodge") +
```

```

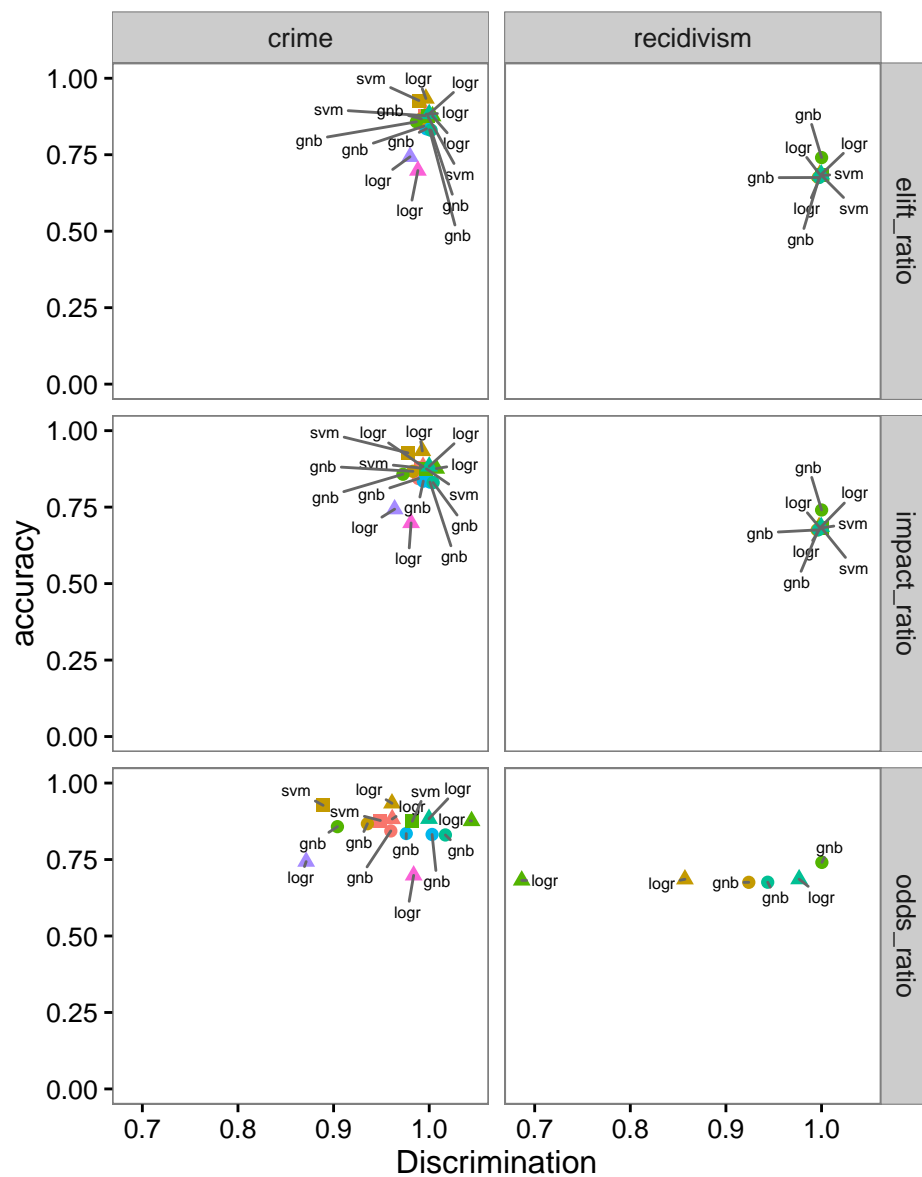
guides(fill = guide_legend(title = "Model")) +
scale_fill_manual(values = c("gray", "darkred", "darkblue", "darkviolet")) +
xlab("Method") + ylab("Discrimination") +
facet_grid(measure ~ dataset, drop = TRUE) +
theme_bw() +
theme(axis.text.x = element_text(angle = 20, hjust = 1, size = 8),
      legend.position = "bottom", panel.grid.major = element_blank(),
      panel.grid.minor = element_blank(),
      panel.background = element_blank(), axis.line = element_line(colour = "black"))

```



Accuracy vs. discrimination

```
all_tidy %>%
  ggplot(aes(x = value, y = accuracy)) +
  geom_point(aes(colour = method, shape = model), size = 2) +
  guides(colour = guide_legend(title = "Method", ncol = 2)) +
  guides(shape = guide_legend(title = "Model")) +
  ylim(0, 1) +
  xlab("Discrimination") +
  geom_text_repel(aes(x = value, y = accuracy, label = model), size = 2) +
  facet_grid(measure ~ dataset, drop = TRUE) +
  theme_bw() +
  theme(legend.position = "bottom") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```



Model ● gnb ▲ logr ■ svm

Method

- None
- Massaging
- Reweighting
- Uniform sampling
- 2 Naive Bayes
- Accuracy-constrained LR
- Fairness-constrained LR

Impact of threshold

```
thresh %>% gather(statistic, value, elift, impact, odds, acc) %>%  
  ggplot(aes(x = thresholds, y = value)) +  
  xlab("Decision threshold") + ylab(NULL) +  
  geom_line(aes(colour = method, linetype = model)) +  
  guides(colour = guide_legend(title = "Method", nrow = 3)) +  
  guides(linetype = guide_legend(title = "Model")) +  
  facet_grid(statistic ~ dataset, switch = "y", drop = TRUE) +  
  theme_bw() +  
  theme(legend.position = "bottom", panel.grid.major = element_blank(),  
        panel.grid.minor = element_blank(),  
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```

