```r
library(readxl)
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v tibble  3.1.7     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
## v purrr   0.3.4

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(reshape2)
```

```
##
## Attache Paket: 'reshape2'

## Das folgende Objekt ist maskiert 'package:tidyr':
##
##     smiths
```

```r
library(jtools)
```

```r
#thesis_data_preds = read_excel("preds/thesis_data_predictions_tuned_models.xlsx")
DS1_preds = read_excel("preds/DS1_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```r
DS2_preds = read_excel("preds/DS2_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```r
DS2_2_preds = read_excel("preds/DS2_2_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```r
DS3_preds = read_excel("preds/DS3_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```
DS4_preds = read_excel("preds/DS4_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```
DS5_preds = read_excel("preds/DS5_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```
DS6_preds = read_excel("preds/DS6_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```
DS7_preds = read_excel("preds/DS7_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

```
DS8_preds = read_excel("preds/DS8_predictions_tuned_models.xlsx")
```

```
## New names:
## * `` -> `...1`
```

Thesis data

“ “ ”

```
#tdp <- thesis_data_preds %>% select(y_test, ends_with("preds")) %>%
#  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
#         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
#  melt(id.vars = 'y_test', variable.name = "model")
#tdp
```

```
#t_nn <- tdp %>% filter(model %in% c("NN")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("NN") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
#t_nn
#ggsave("new_pred_plot/thesis_data_true_vs_preds_NN.png")
```

```
#t_lasso <- tdp %>% filter(model %in% c("Lasso")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("Lasso") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
#t_lasso
#ggsave("new_pred_plot/thesis_data_true_vs_preds_lasso.png")
```

```
#t_ridge <- tdp %>% filter(model %in% c("Ridge")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("Ridge") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
#t_ridge
#ggsave("new_pred_plot/thesis_data_true_vs_preds_ridge.png")
```

```
#t_lr <- tdp %>% filter(model %in% c("LR")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("LR") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
#t_lr
#ggsave("new_pred_plot/thesis_data_true_vs_preds_lr.png")
```

```
#t_rf <- tdp %>% filter(model %in% c("RF")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("RF") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
```

```
#t_rf
#ggsave("new_pred_plot/thesis_data_true_vs_preds_rf.png")
```

```
#t_rt <- tdp %>% filter(model %in% c("RT")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("RT") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
#t_rt
#ggsave("new_pred_plot/thesis_data_true_vs_preds_rt.png")
```

```
#t_ert <- tdp %>% filter(model %in% c("ERT")) %>%
#  ggplot(aes(value, y_test, alpha = .5)) +
#  geom_point() +
#  geom_abline(col = "red") +
#  ylim(6.0, 10.0) +
#  xlim(6.0, 10.0) +
#  ggtitle("ERT") +
#  ylab("True grade") +
#  xlab("Predicted grade") +
#  theme_apa() +
#  theme(legend.position = "none")
#t_ert
#ggsave("new_pred_plot/thesis_data_true_vs_preds_ert.png")
```

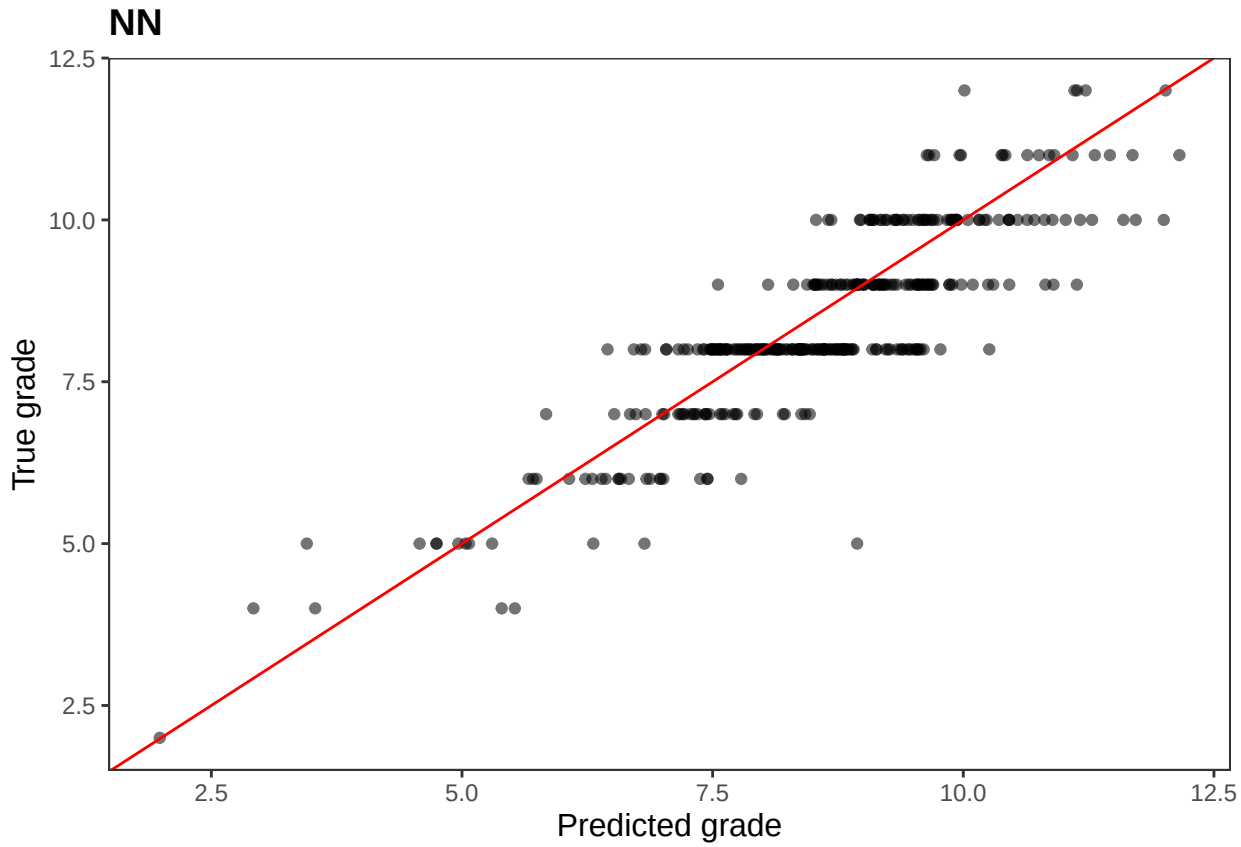
Hewlett datasets

DS1

```
DS1 <- DS1_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS1)
```

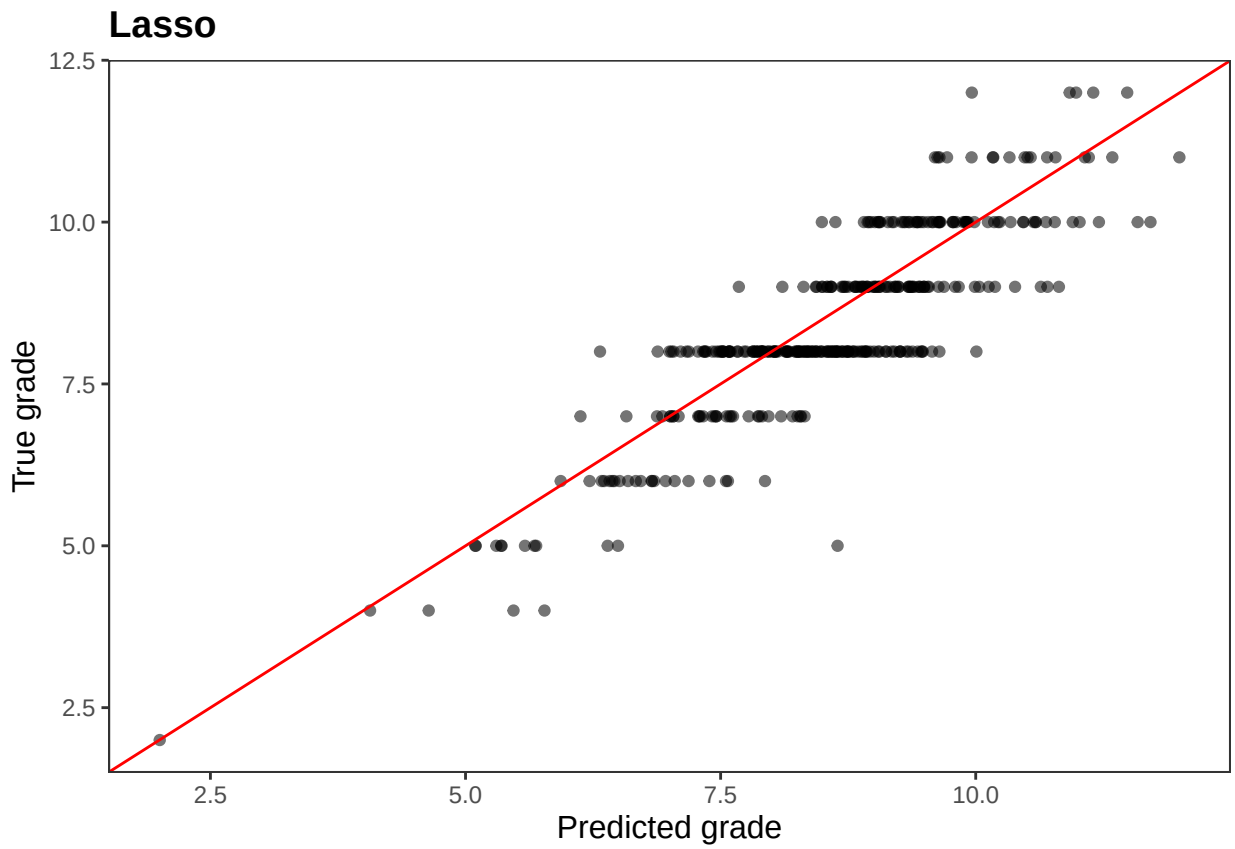
```
##   y_test model    value
## 1      9    RF 9.769424
## 2      9    RF 8.100251
## 3      8    RF 8.255639
## 4      8    RF 8.020050
## 5     10    RF 9.403509
## 6      8    RF 7.972431
```

```
DS1 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



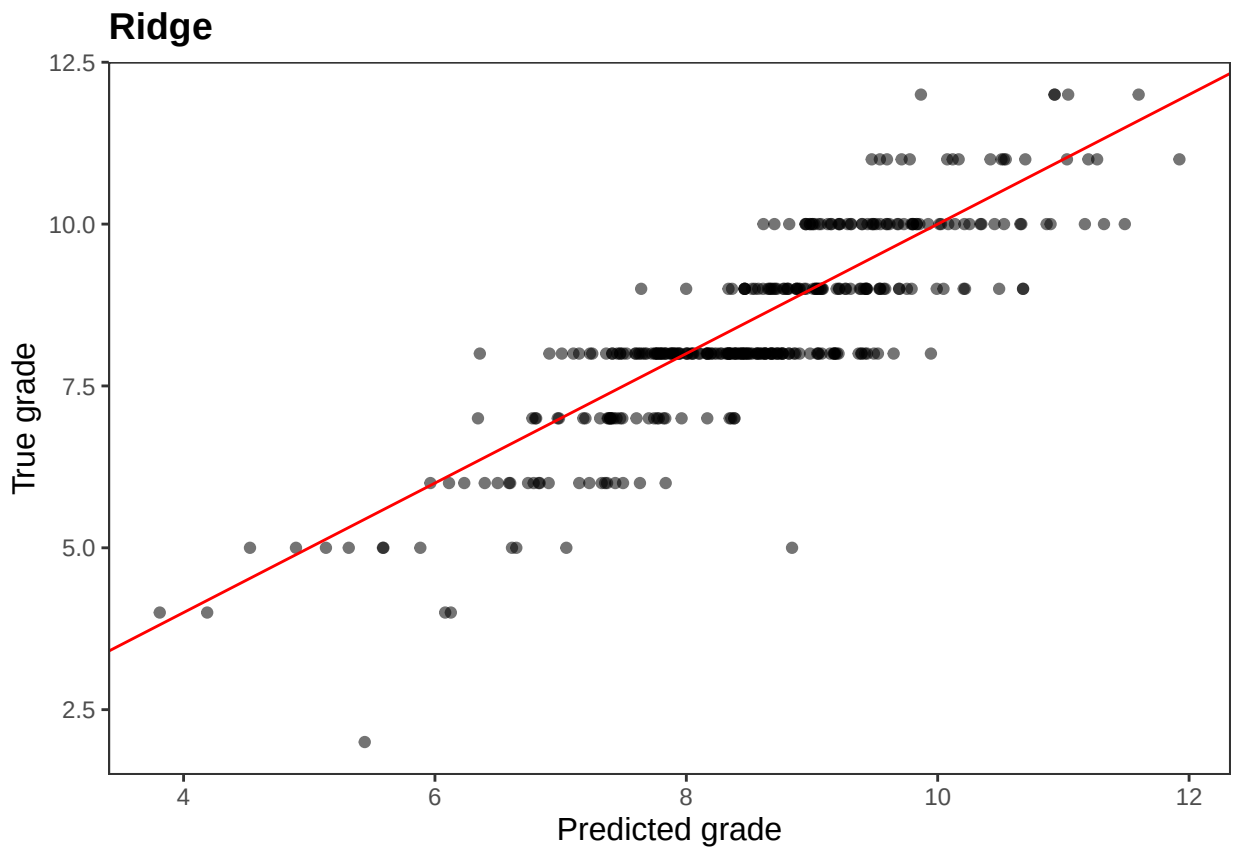

```
#ggsave("new_pred_plot/DS1_true_vs_preds_NN.png")
```

```
DS1 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



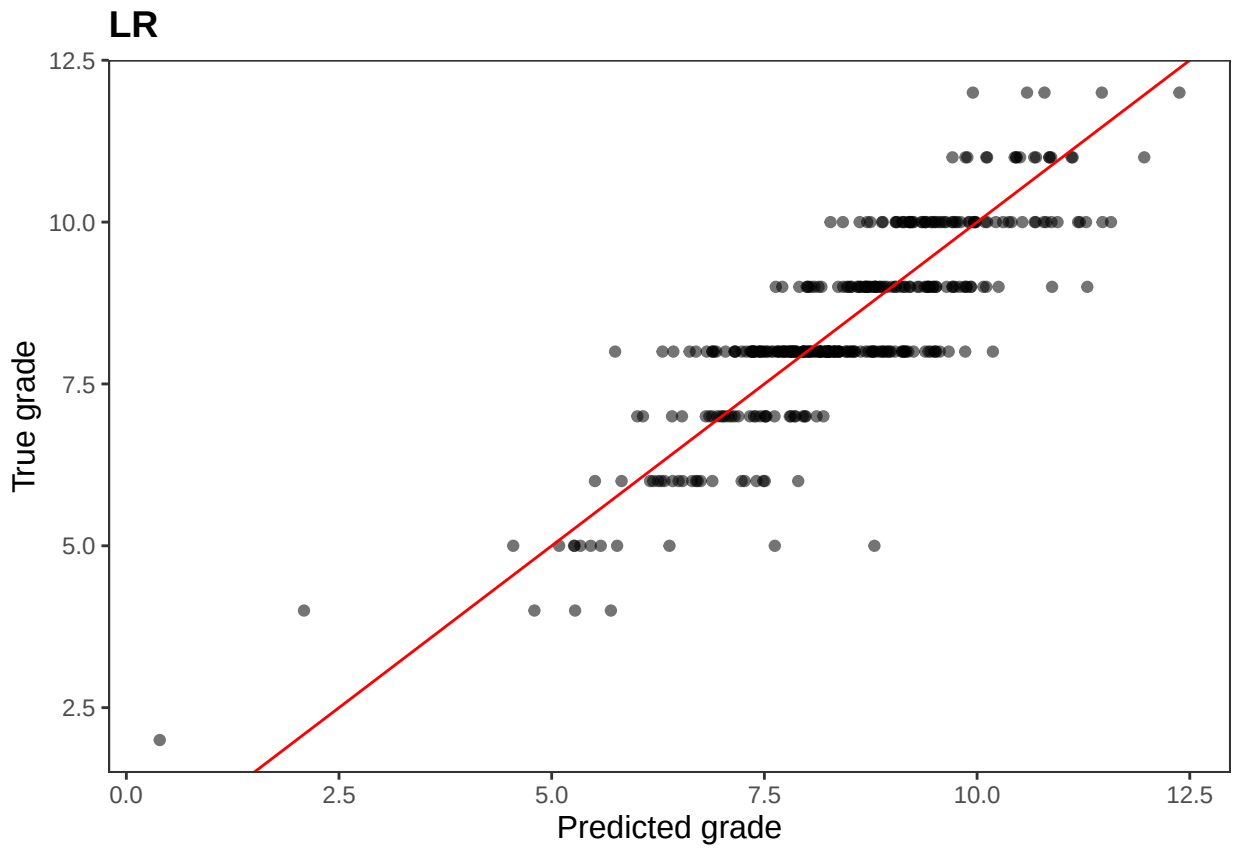

```
#ggsave("new_pred_plot/DS1_true_vs_preds_Lasso.png")
```

```
DS1 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



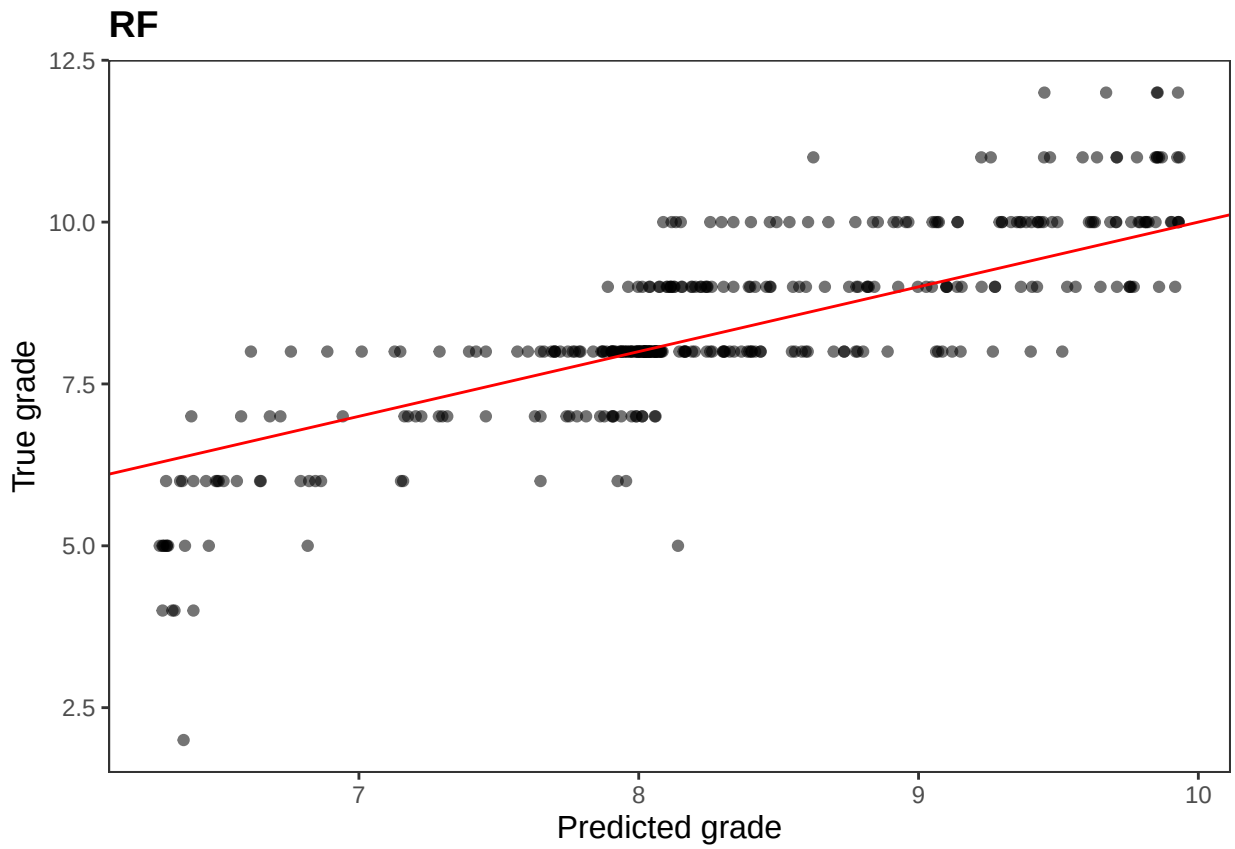

```
#ggsave("new_pred_plot/DS1_true_vs_preds_Ridge.png")
```

```
DS1 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



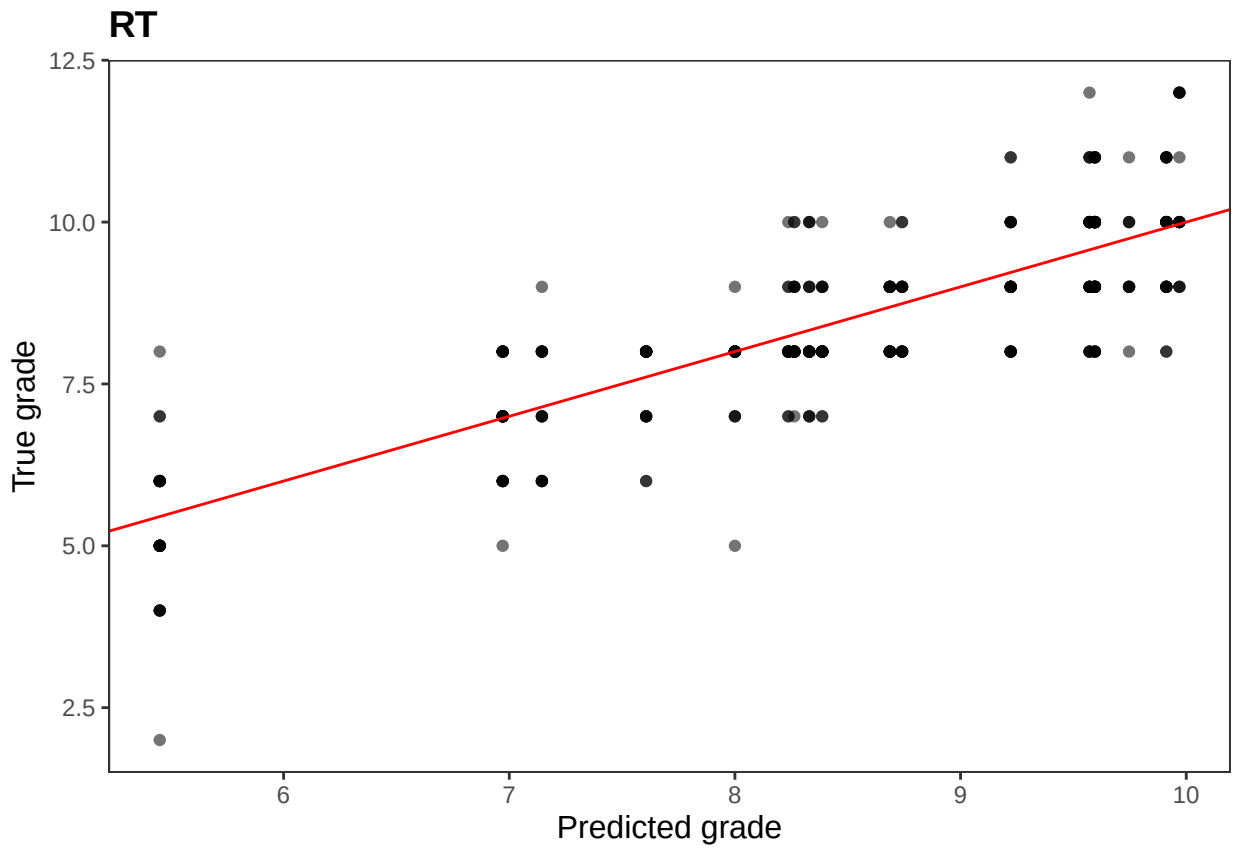

```
#ggsave("new_pred_plot/DS1_true_vs_preds_LR.png")
```

```
DS1 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



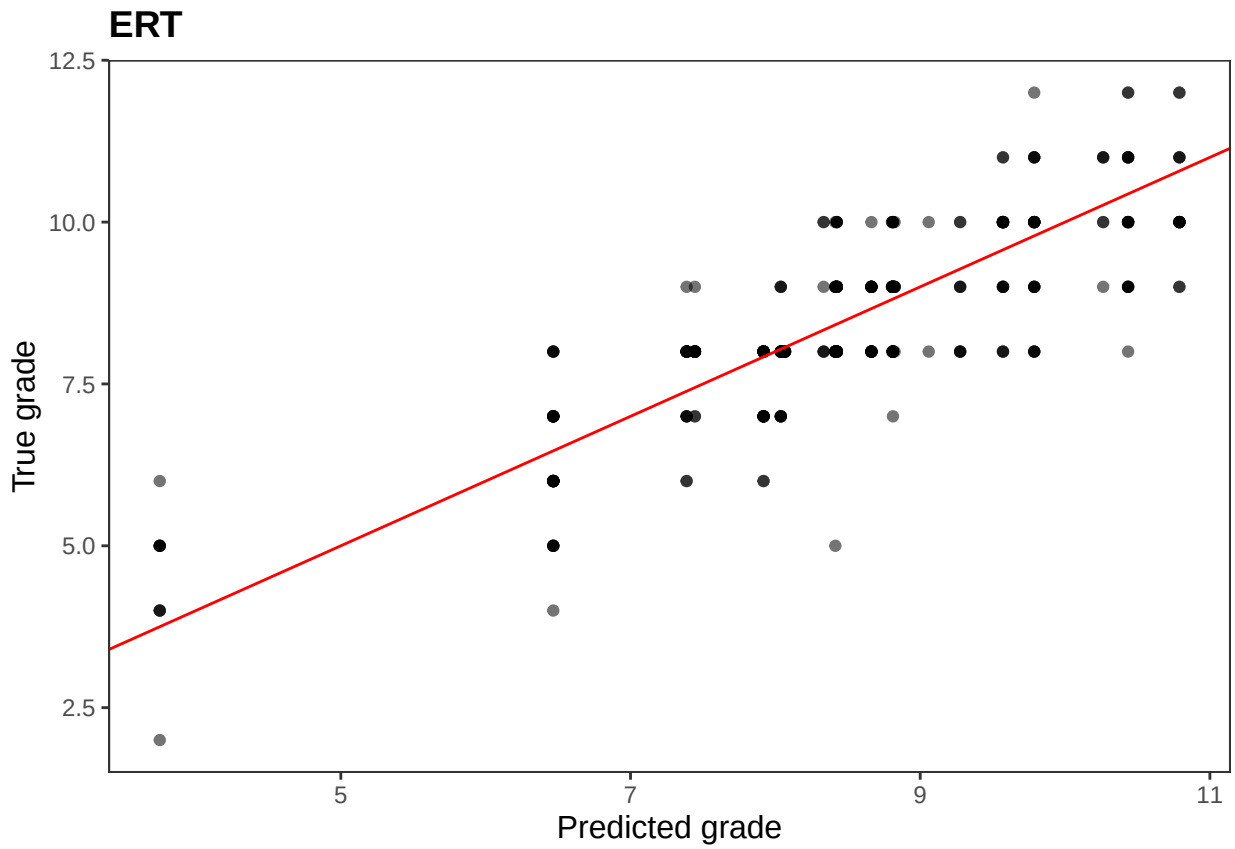

```
#ggsave("new_pred_plot/DS1_true_vs_preds_RF.png")
```

```
DS1 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS1_true_vs_preds_RT.png")
```

```
DS1 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS1_true_vs_preds_ERT.png")
```

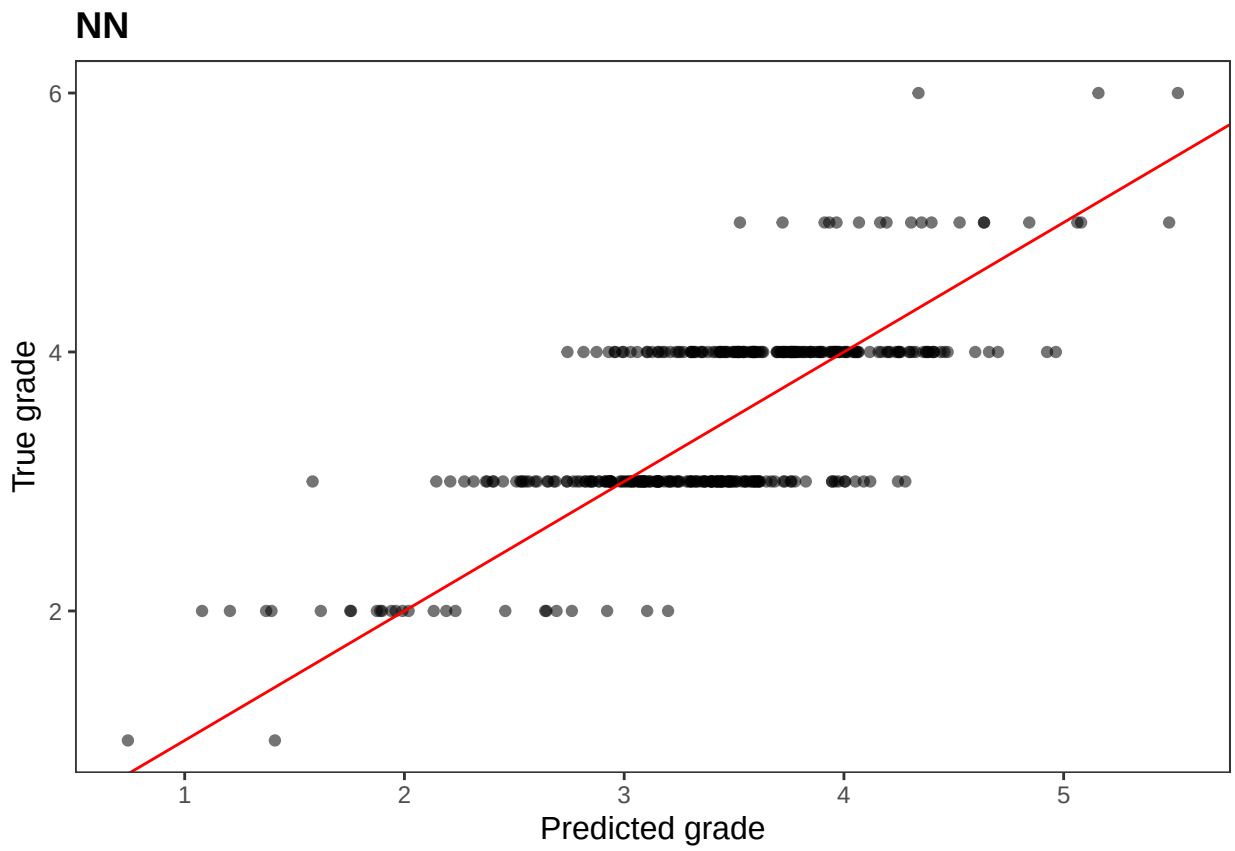
DS2A

```
DS2A <- DS2_2_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS2A)
```

```
##   y_test model    value
## 1      3    RF 3.204261
## 2      3    RF 3.667920
## 3      4    RF 3.795739
## 4      4    RF 3.309524
## 5      3    RF 3.487469
## 6      3    RF 3.672932
```

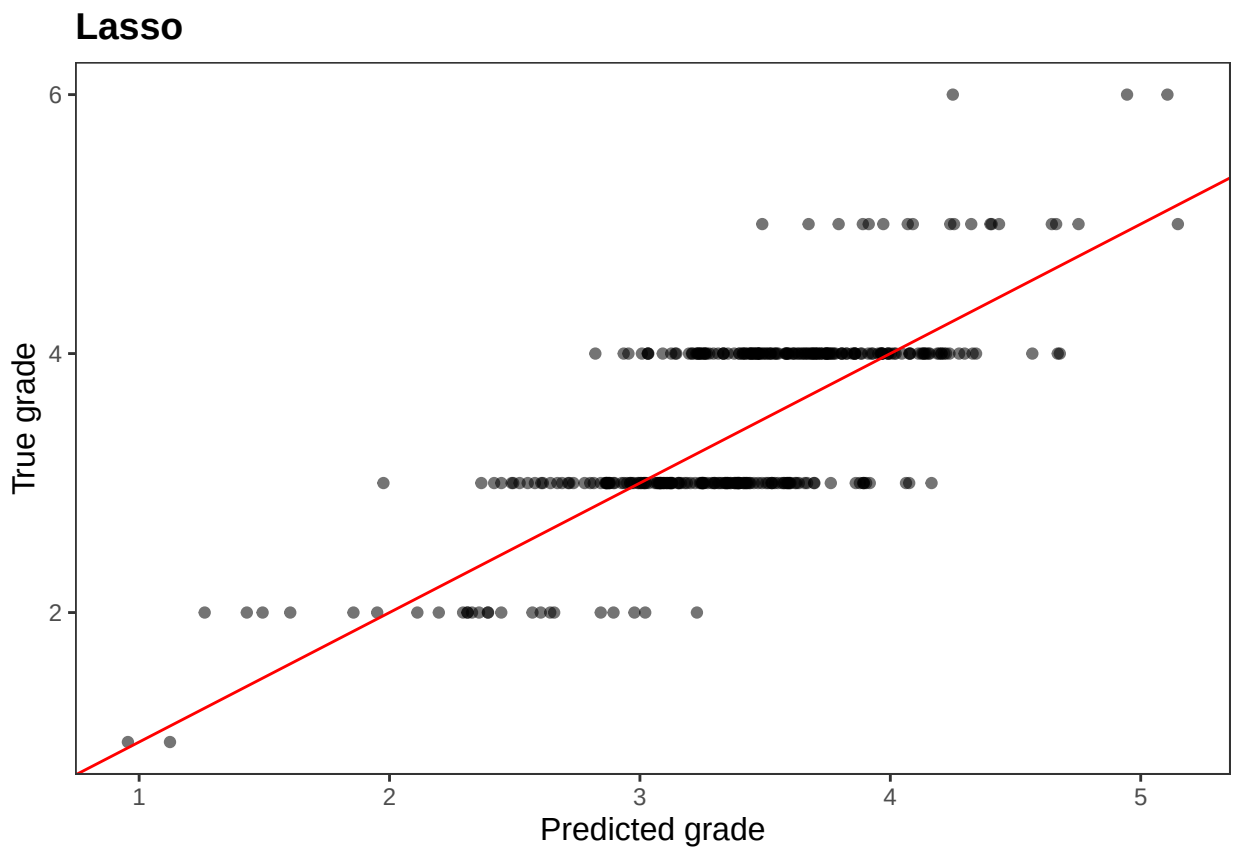
```
DS2A %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



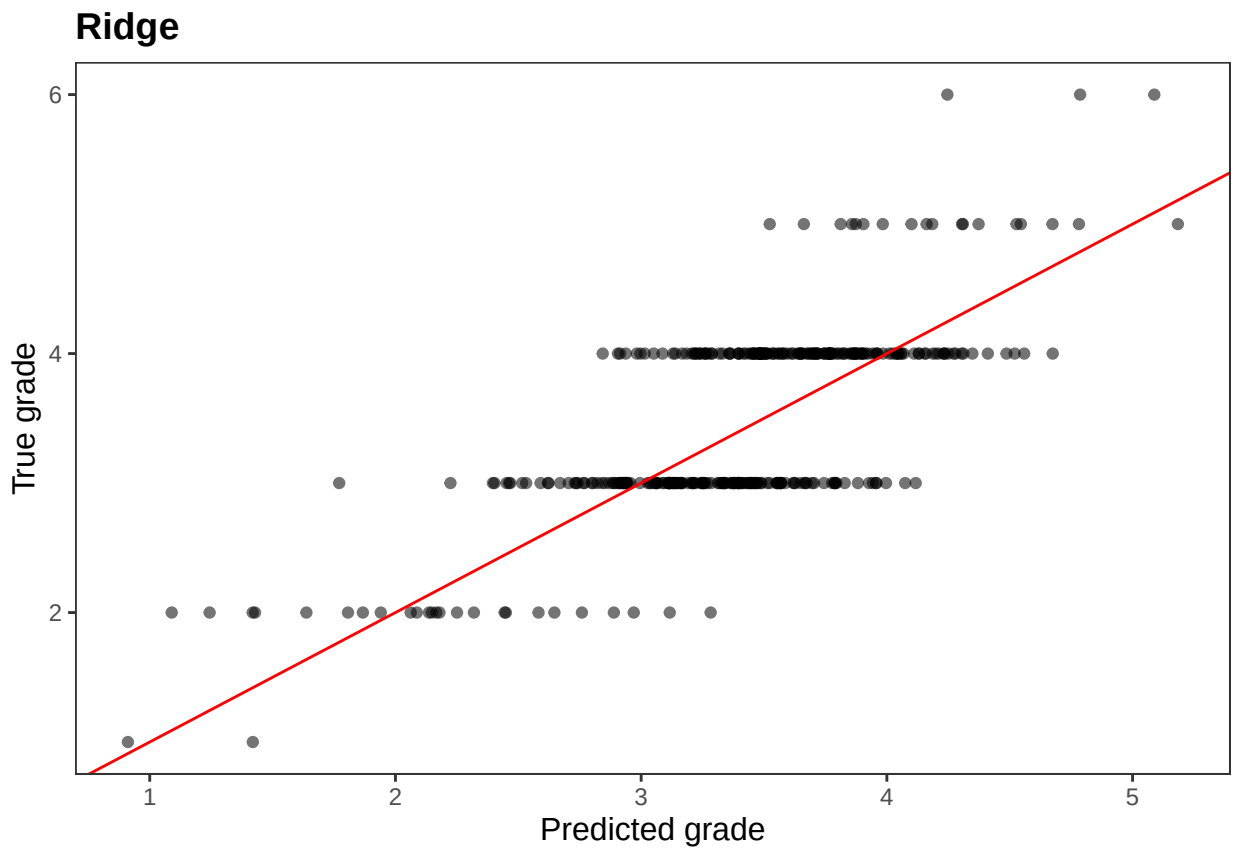

```
#ggsave("new_pred_plot/DS2A_true_vs_preds_NN.png")
```

```
DS2A %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



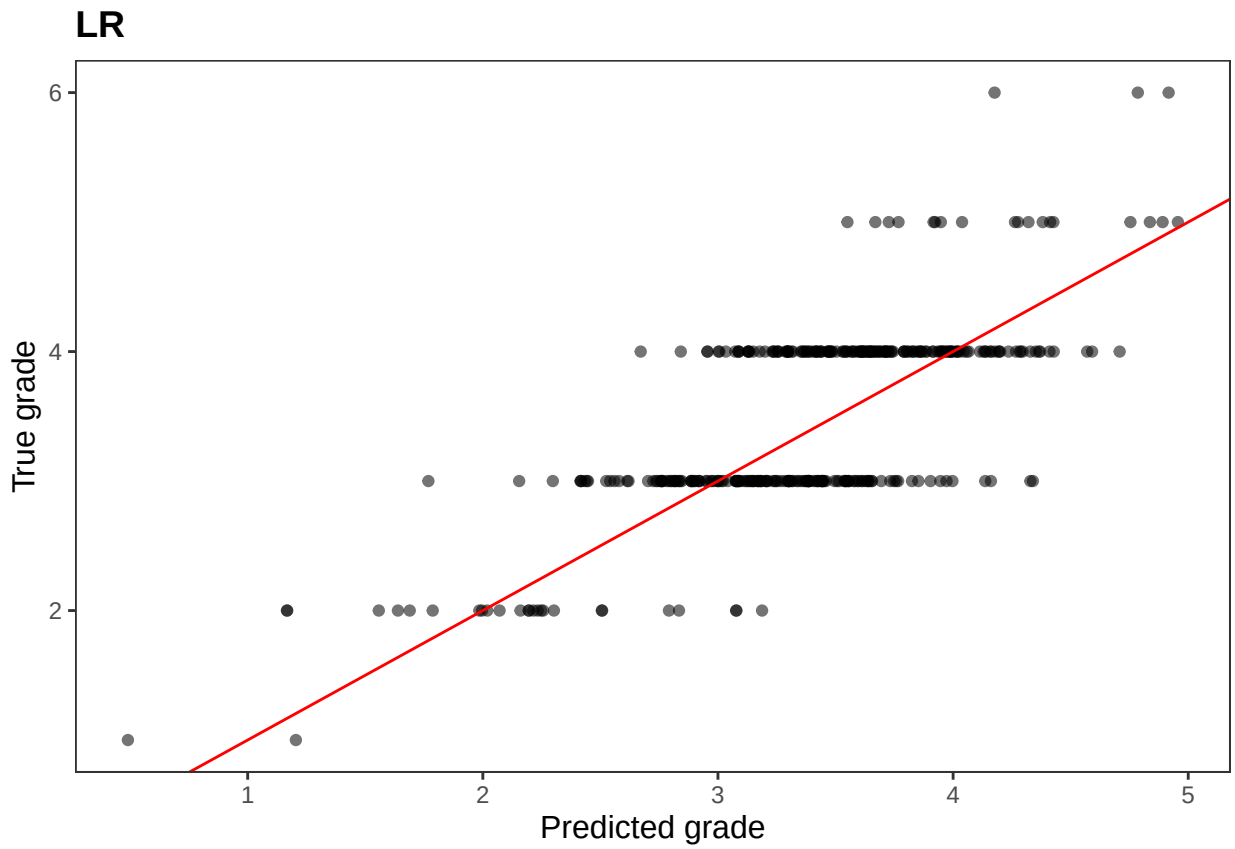

```
#ggsave("new_pred_plot/DS2A_true_vs_preds_Lasso.png")
```

```
DS2A %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



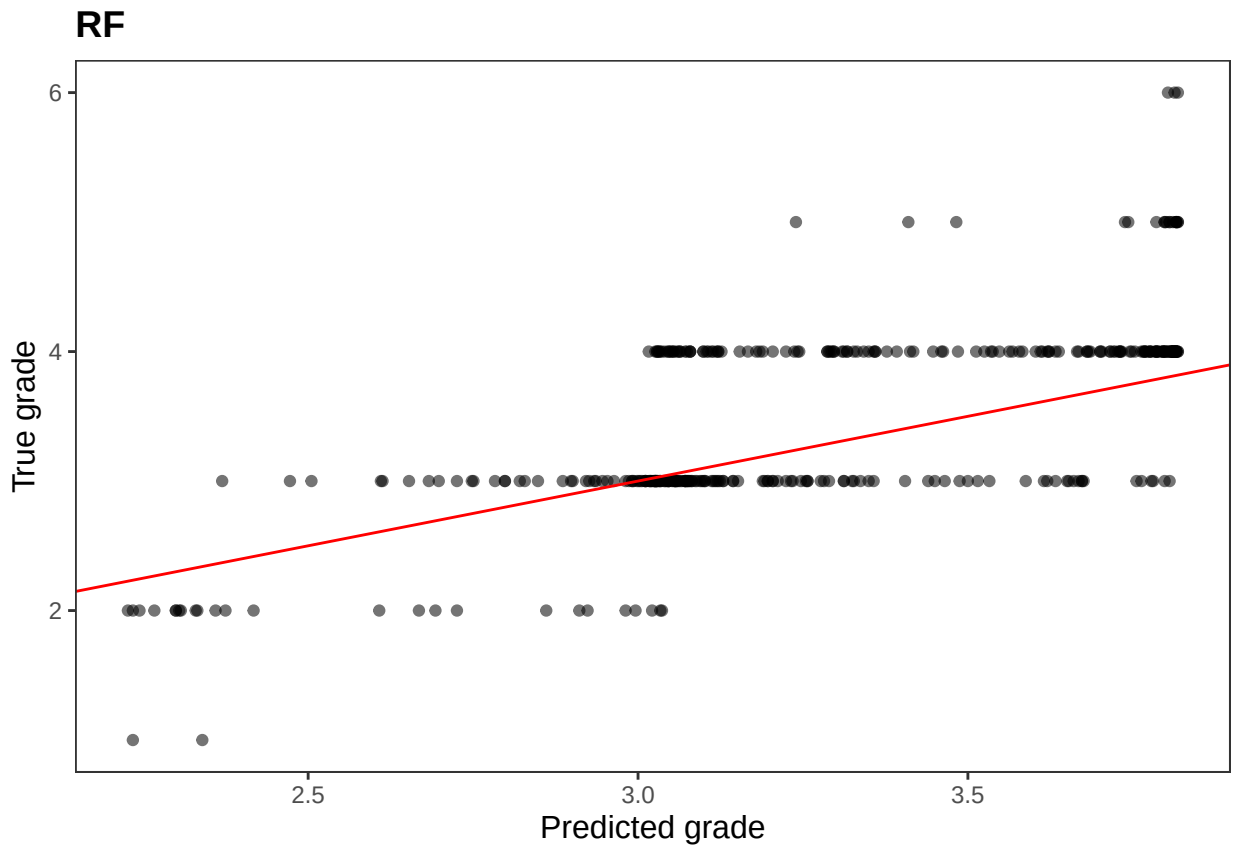

```
#ggsave("new_pred_plot/DS2A_true_vs_preds_Ridge.png")
```

```
DS2A %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



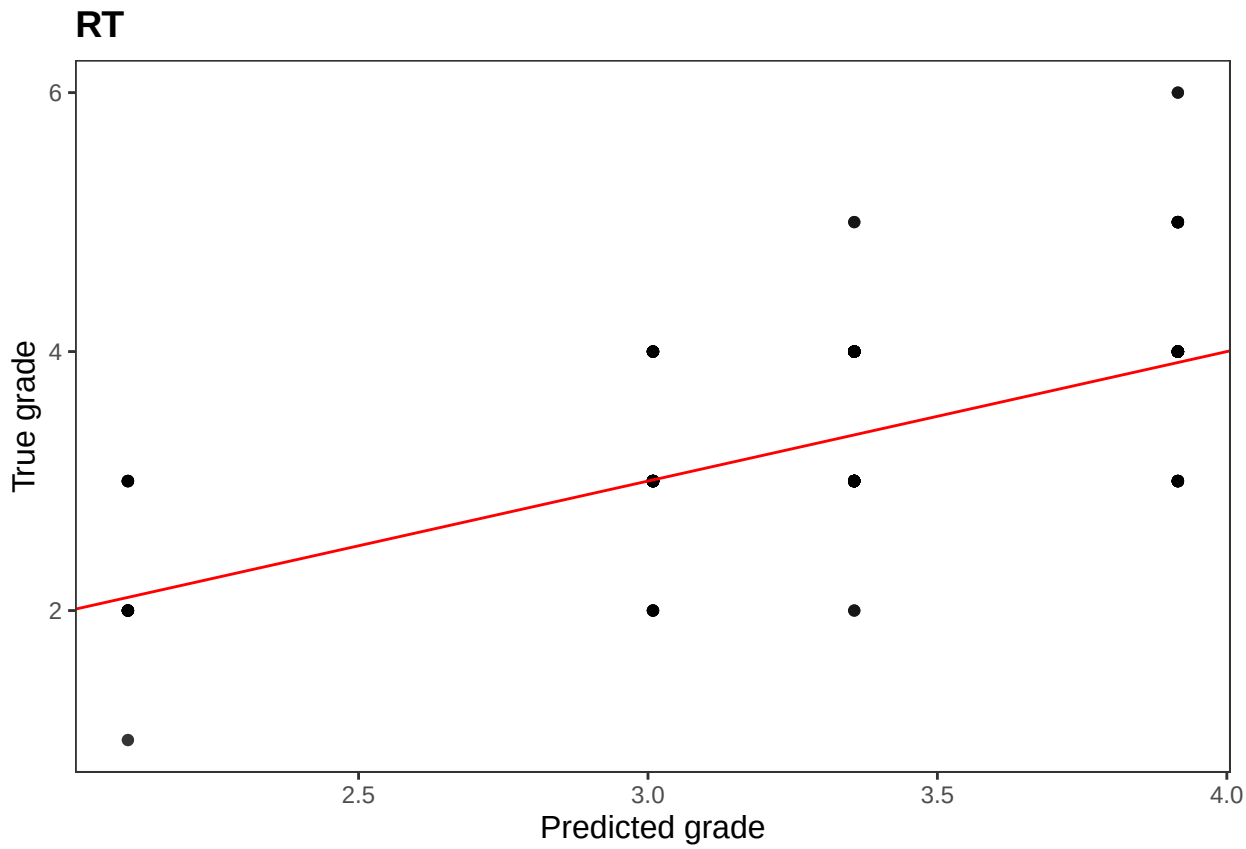

```
#ggsave("new_pred_plot/DS2A_true_vs_preds_LR.png")
```

```
DS2A %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



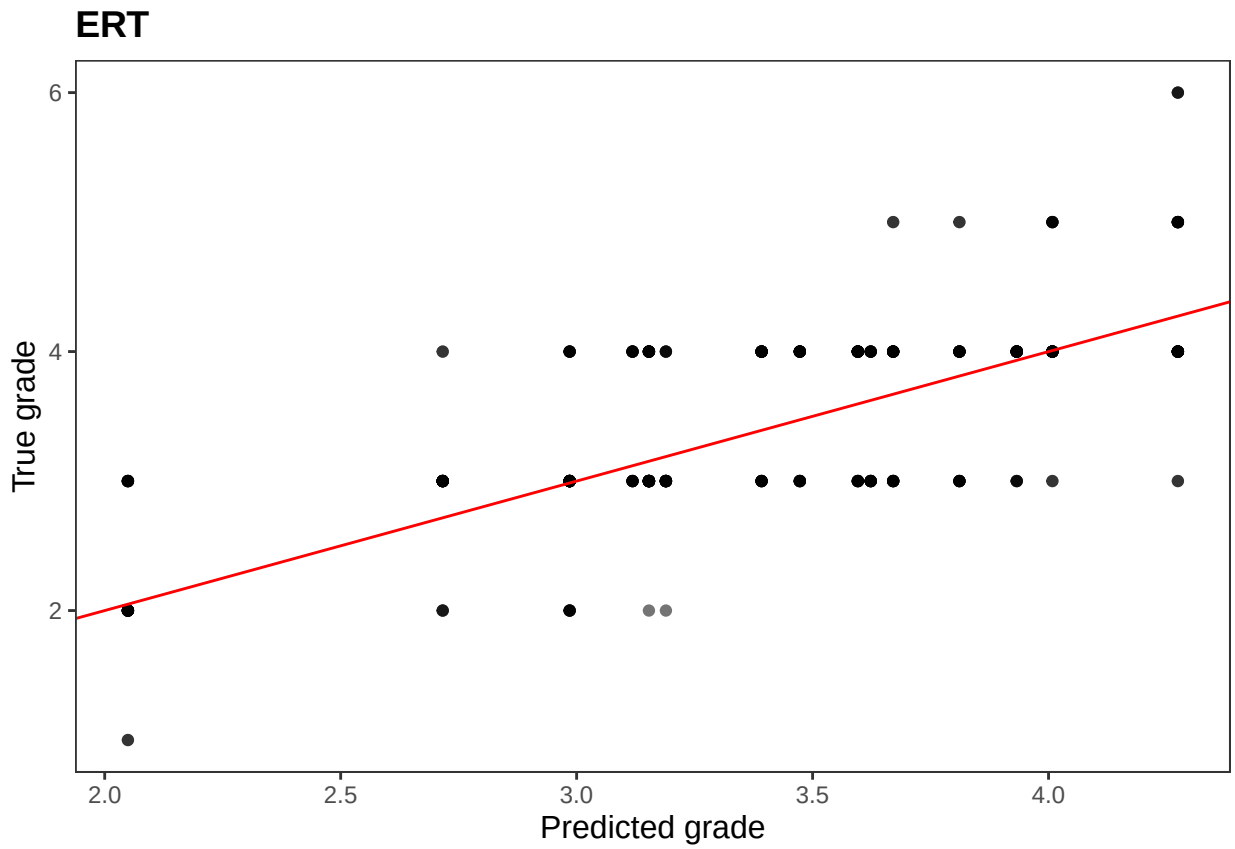

```
#ggsave("new_pred_plot/DS2A_true_vs_preds_RF.png")
```

```
DS2A %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS2A_true_vs_preds_RT.png")
```

```
DS2A %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS2A_true_vs_preds_ERT.png")
```

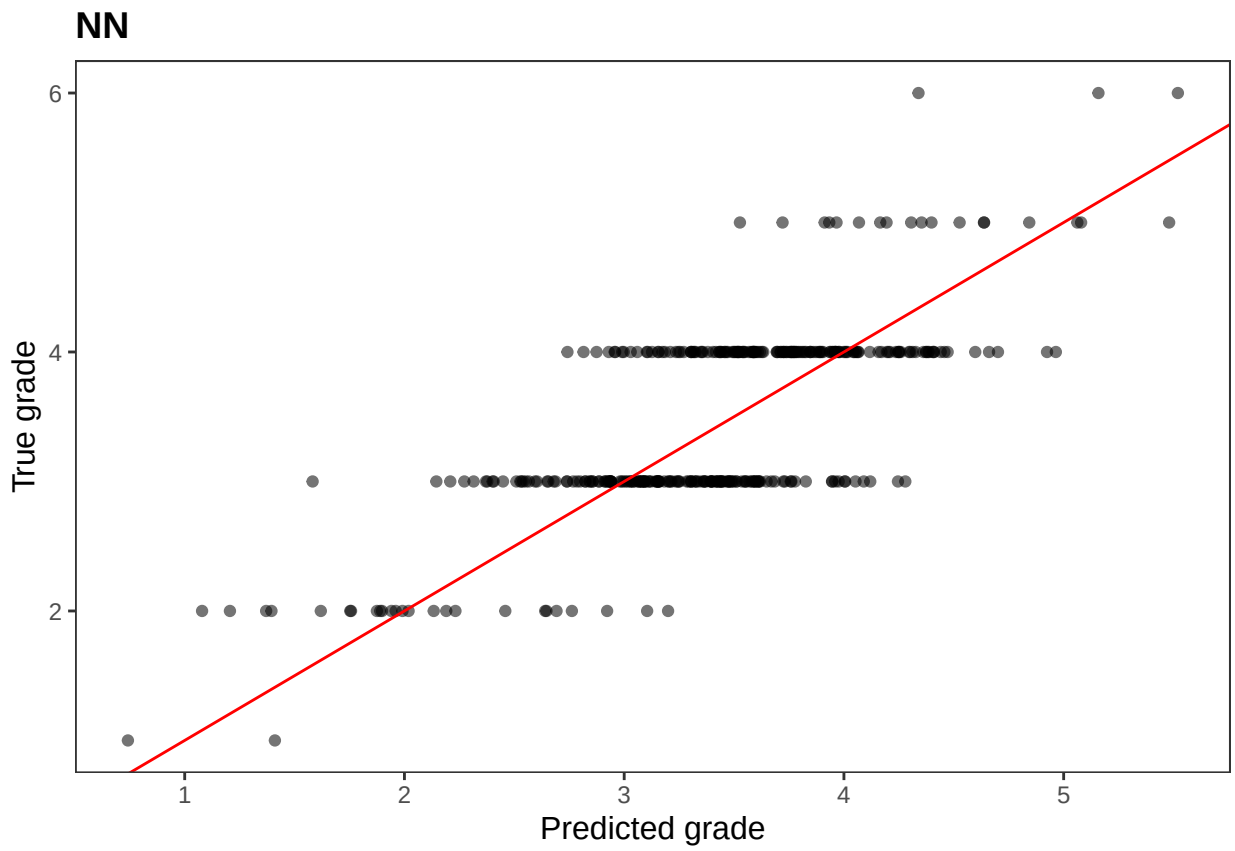
DS2B

```
DS2B <- DS2_2_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS2B)
```

```
##   y_test model    value
## 1      3    RF 3.204261
## 2      3    RF 3.667920
## 3      4    RF 3.795739
## 4      4    RF 3.309524
## 5      3    RF 3.487469
## 6      3    RF 3.672932
```

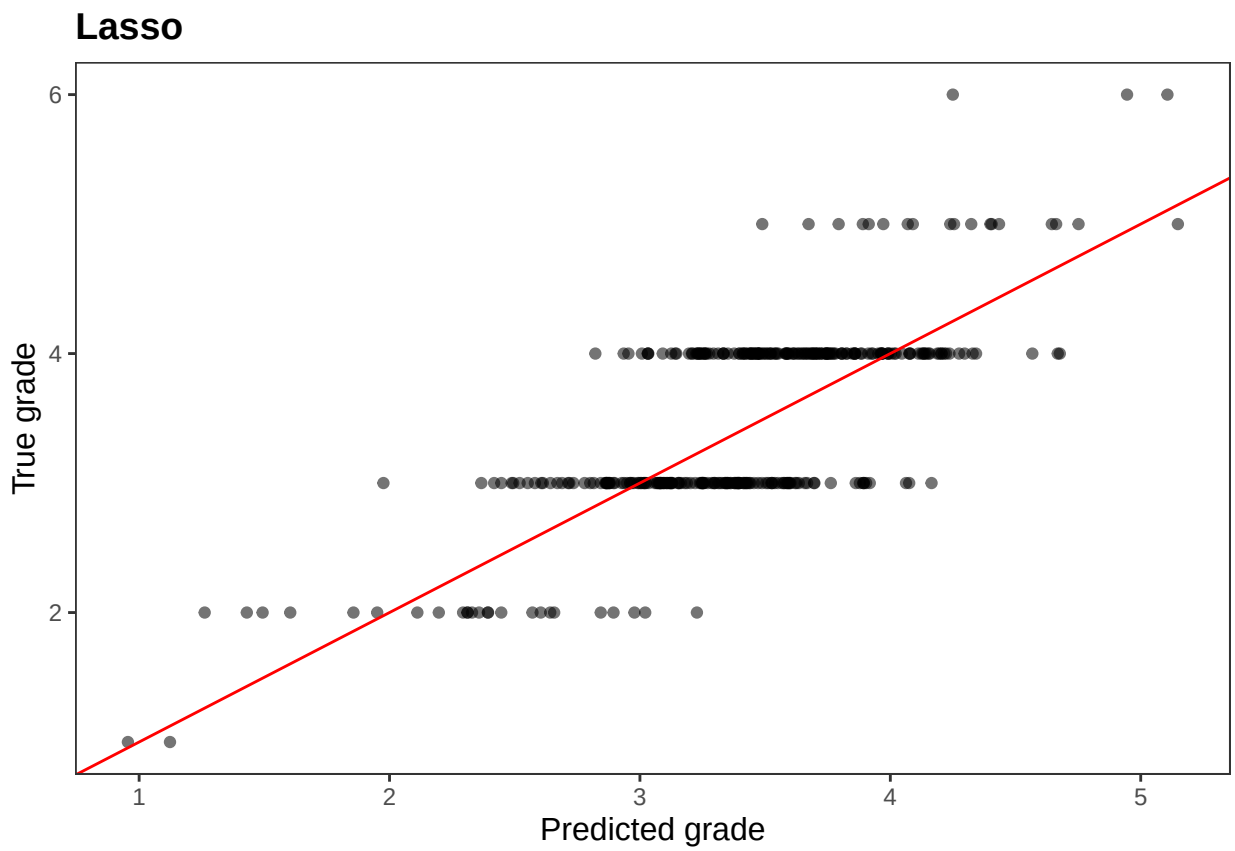
```
DS2B %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



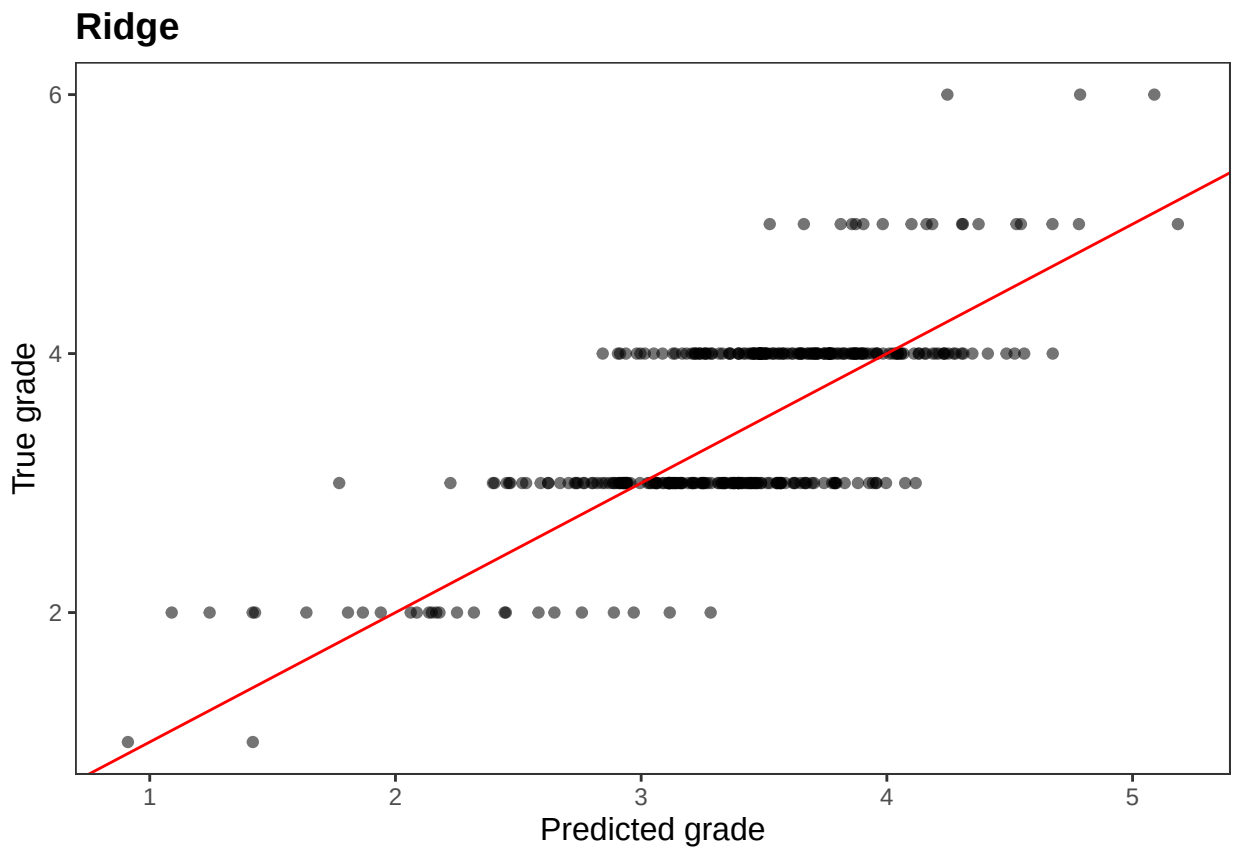

```
#ggsave("new_pred_plot/DS2B_true_vs_preds_NN.png")
```

```
DS2B %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



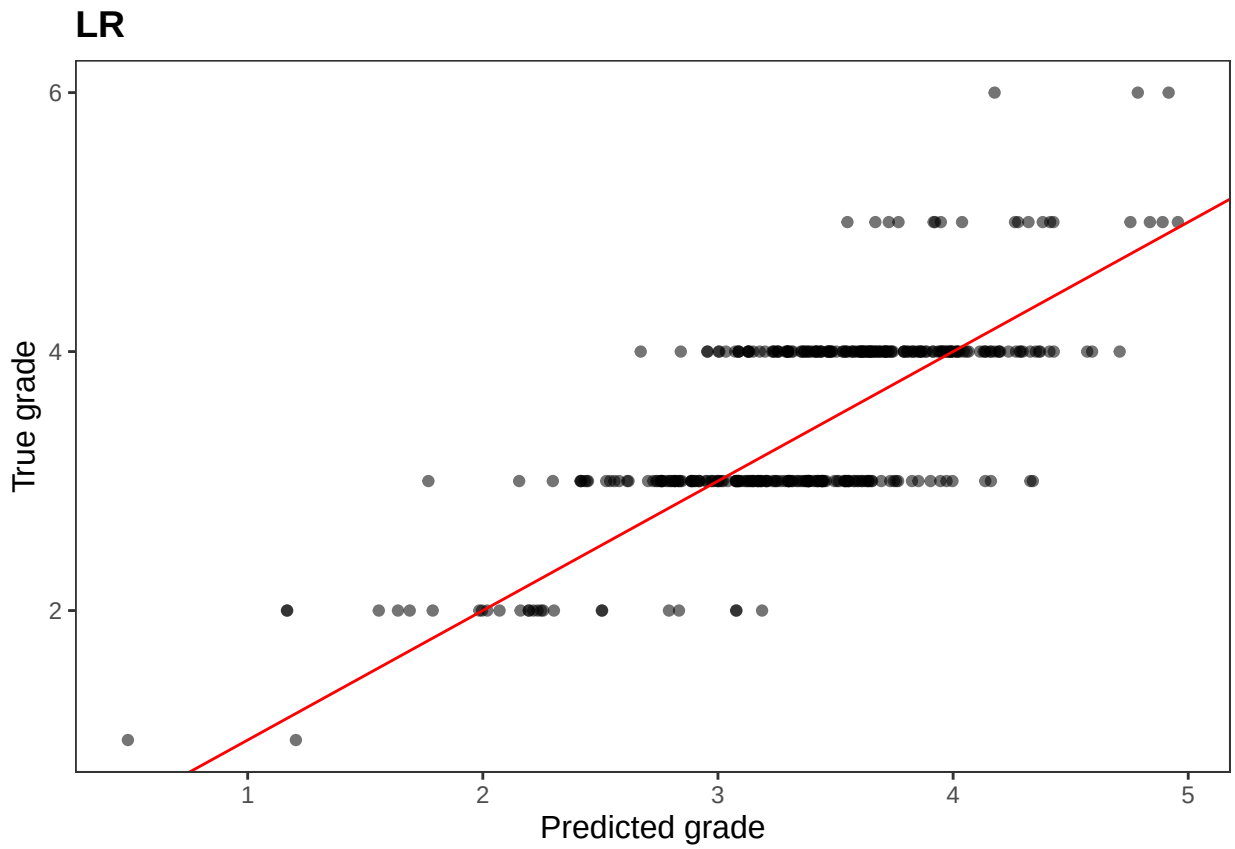

```
#ggsave("new_pred_plot/DS2B_true_vs_preds_Lasso.png")
```

```
DS2B %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



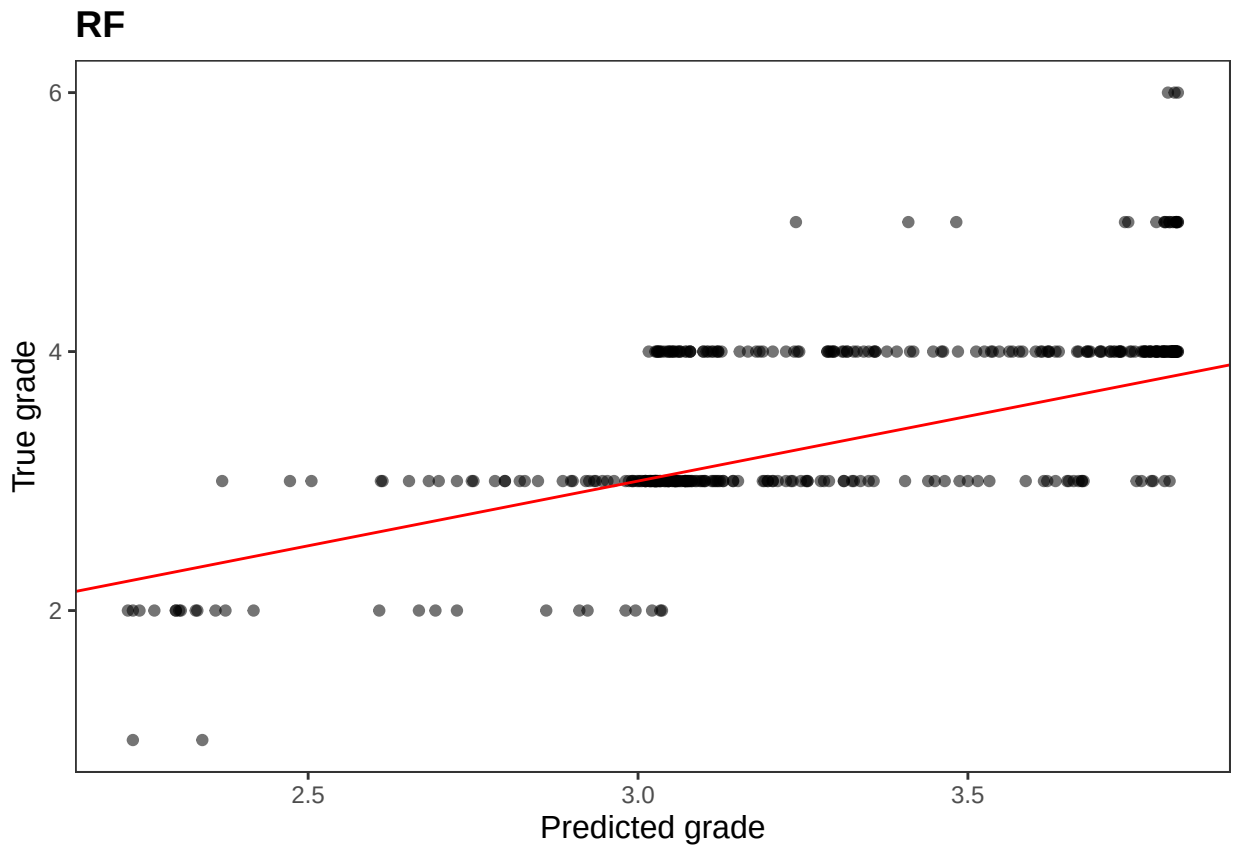

```
#ggsave("new_pred_plot/DS2B_true_vs_preds_Ridge.png")
```

```
DS2B %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



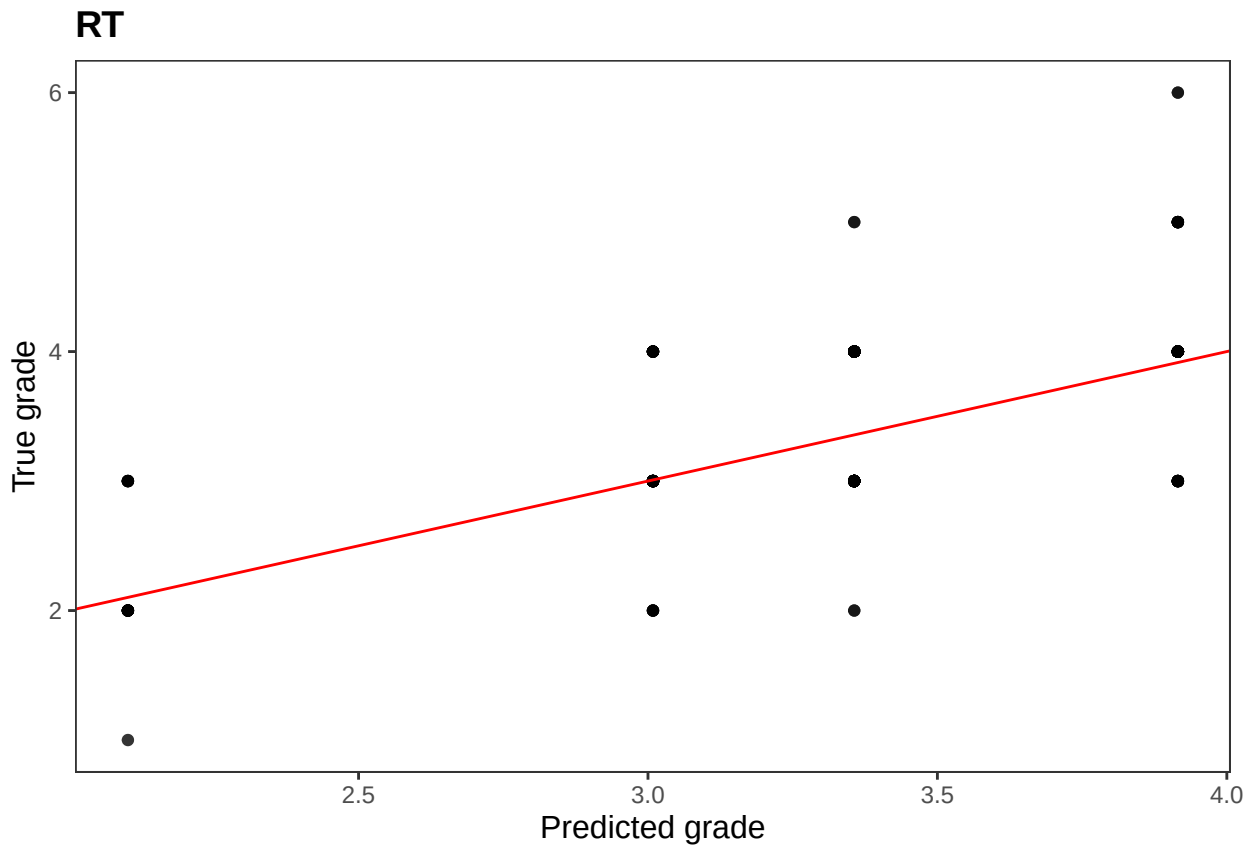

```
#ggsave("new_pred_plot/DS2B_true_vs_preds_LR.png")
```

```
DS2B %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



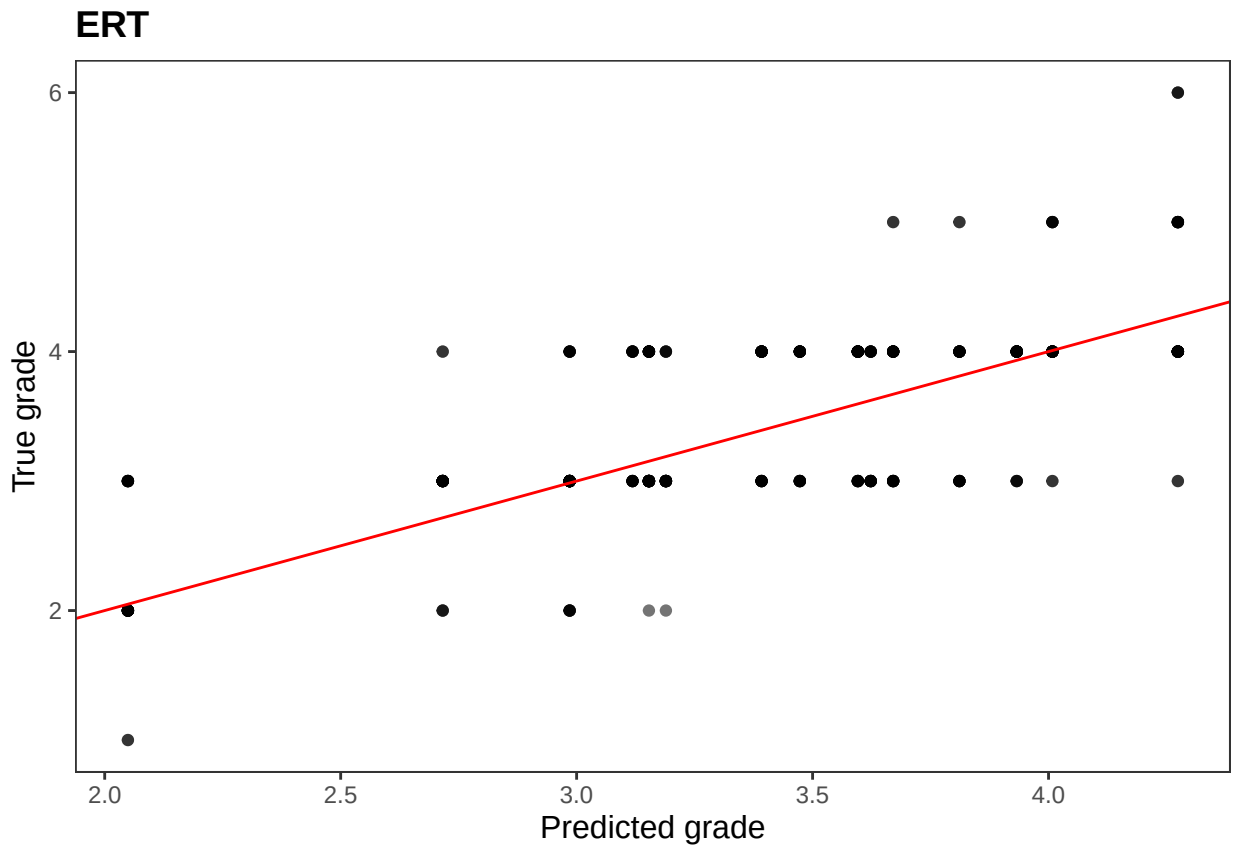

```
#ggsave("new_pred_plot/DS2B_true_vs_preds_RF.png")
```

```
DS2B %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS2B_true_vs_preds_RT.png")
```

```
DS2B %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS2B_true_vs_preds_ERT.png")
```

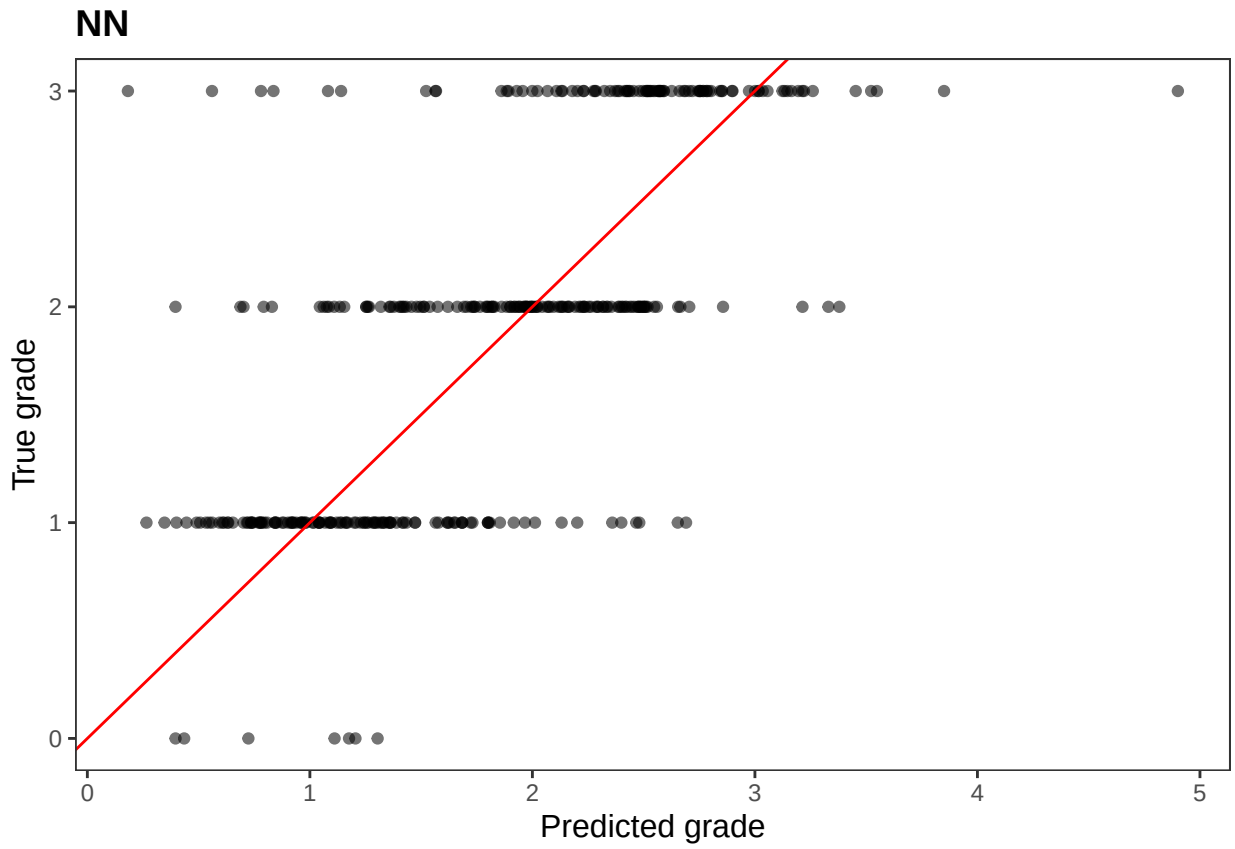
DS3

```
DS3 <- DS3_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS3)
```

```
##   y_test model    value
## 1      1    RF 1.220476
## 2      3    RF 2.530185
## 3      1    RF 1.132975
## 4      1    RF 1.143840
## 5      3    RF 2.363201
## 6      2    RF 2.210415
```

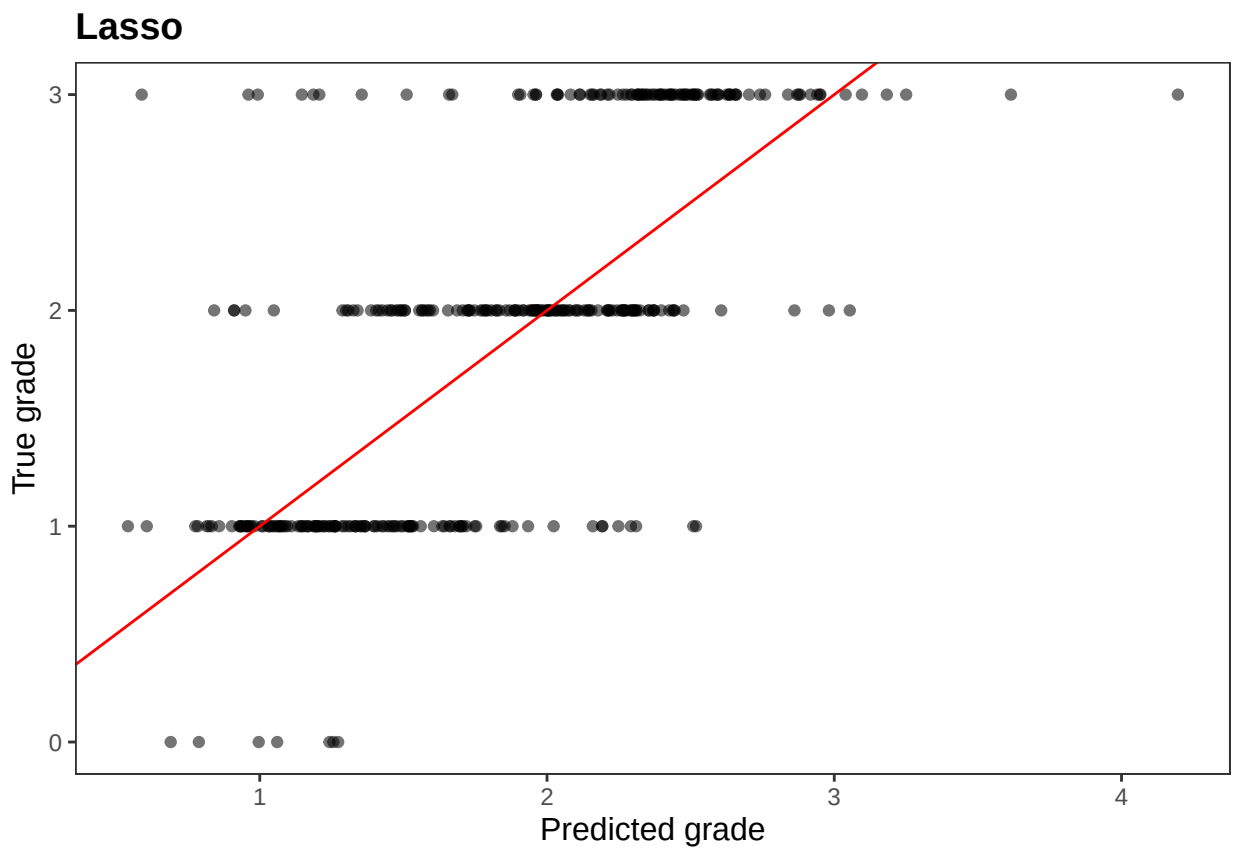
```
DS3 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



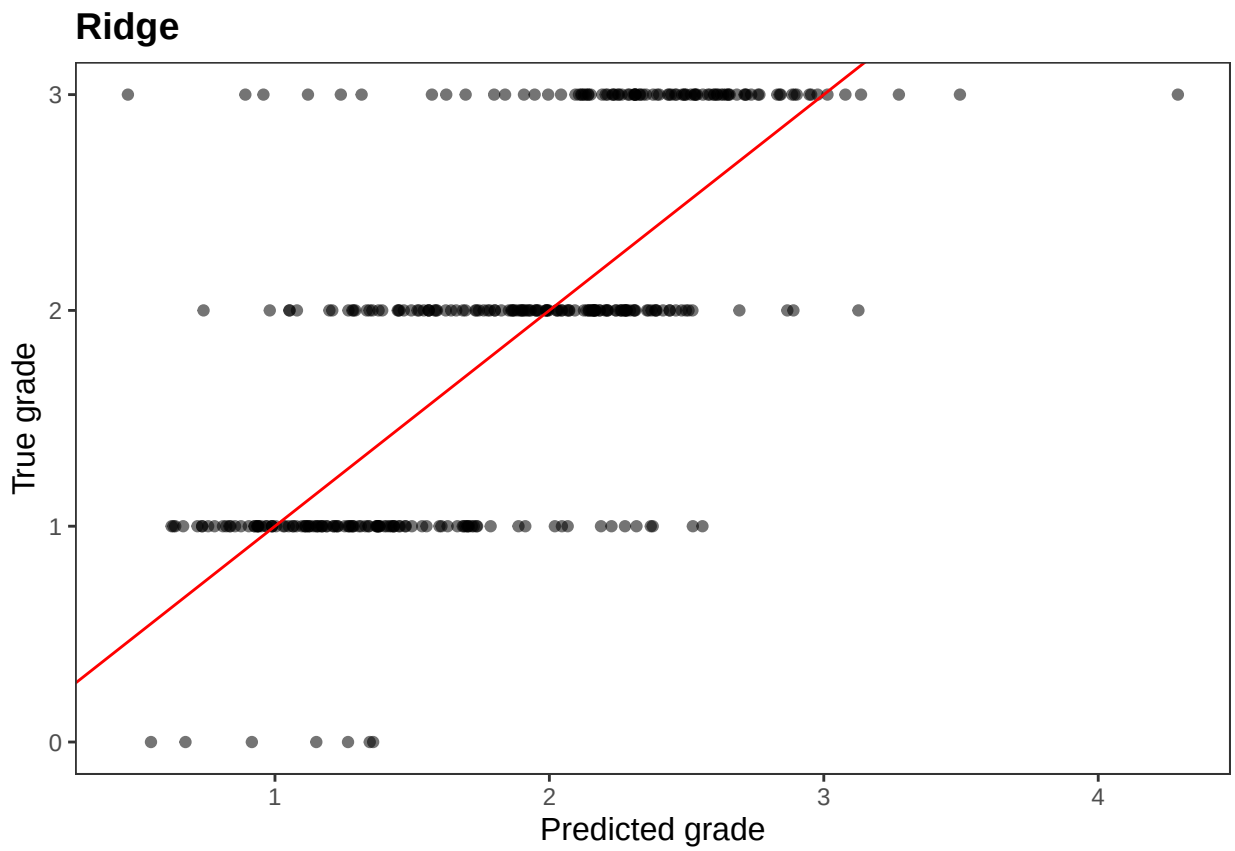

```
#ggsave("new_pred_plot/DS3_true_vs_preds_NN.png")
```

```
DS3 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



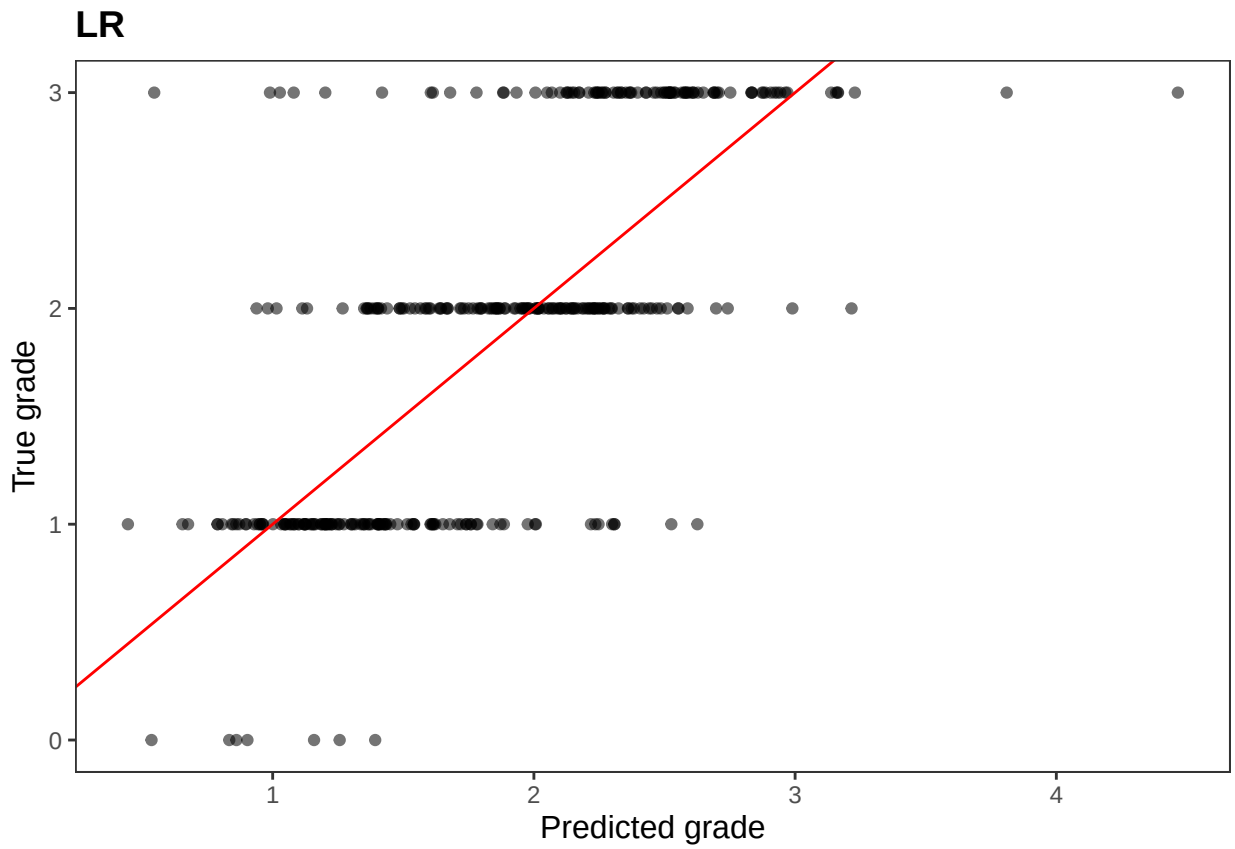

```
#ggsave("new_pred_plot/DS3_true_vs_preds_Lasso.png")
```

```
DS3 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



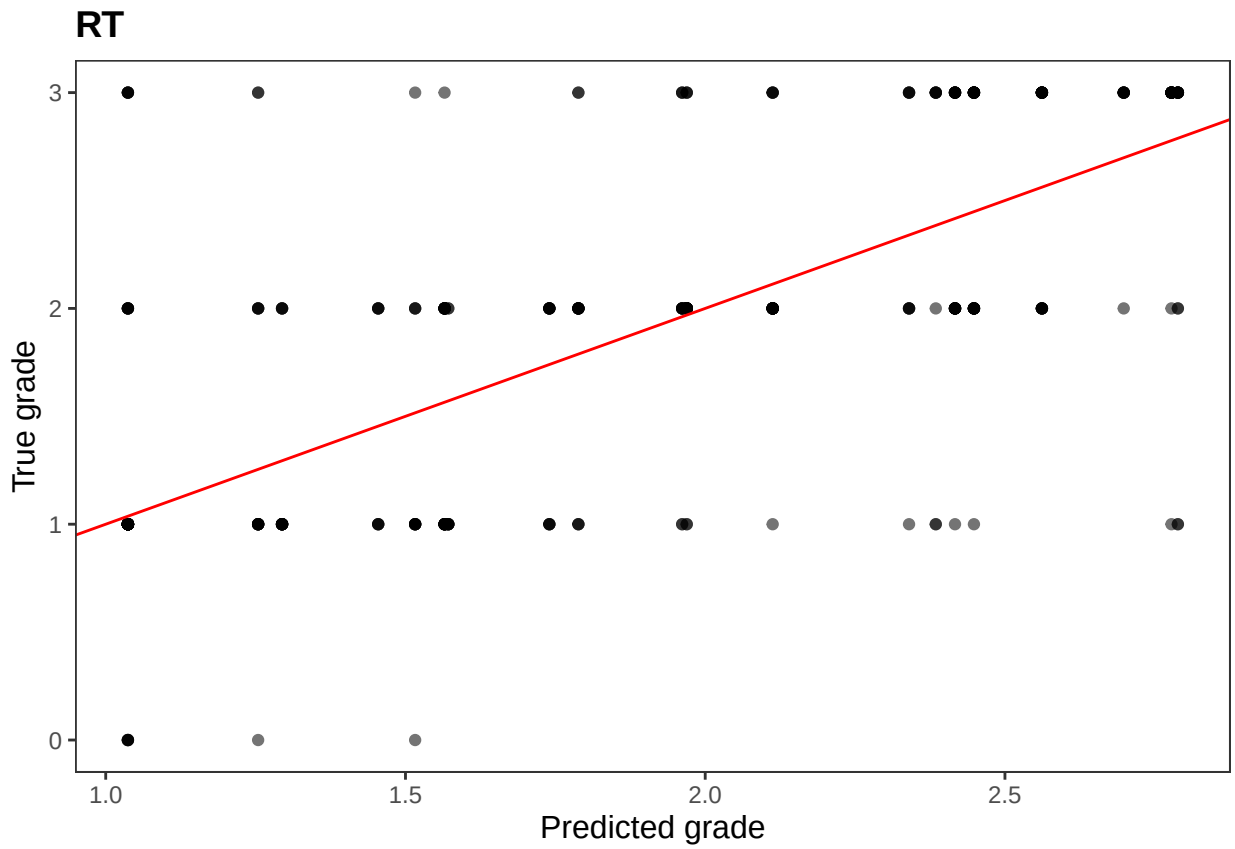

```
#ggsave("new_pred_plot/DS3_true_vs_preds_Ridge.png")
```

```
DS3 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



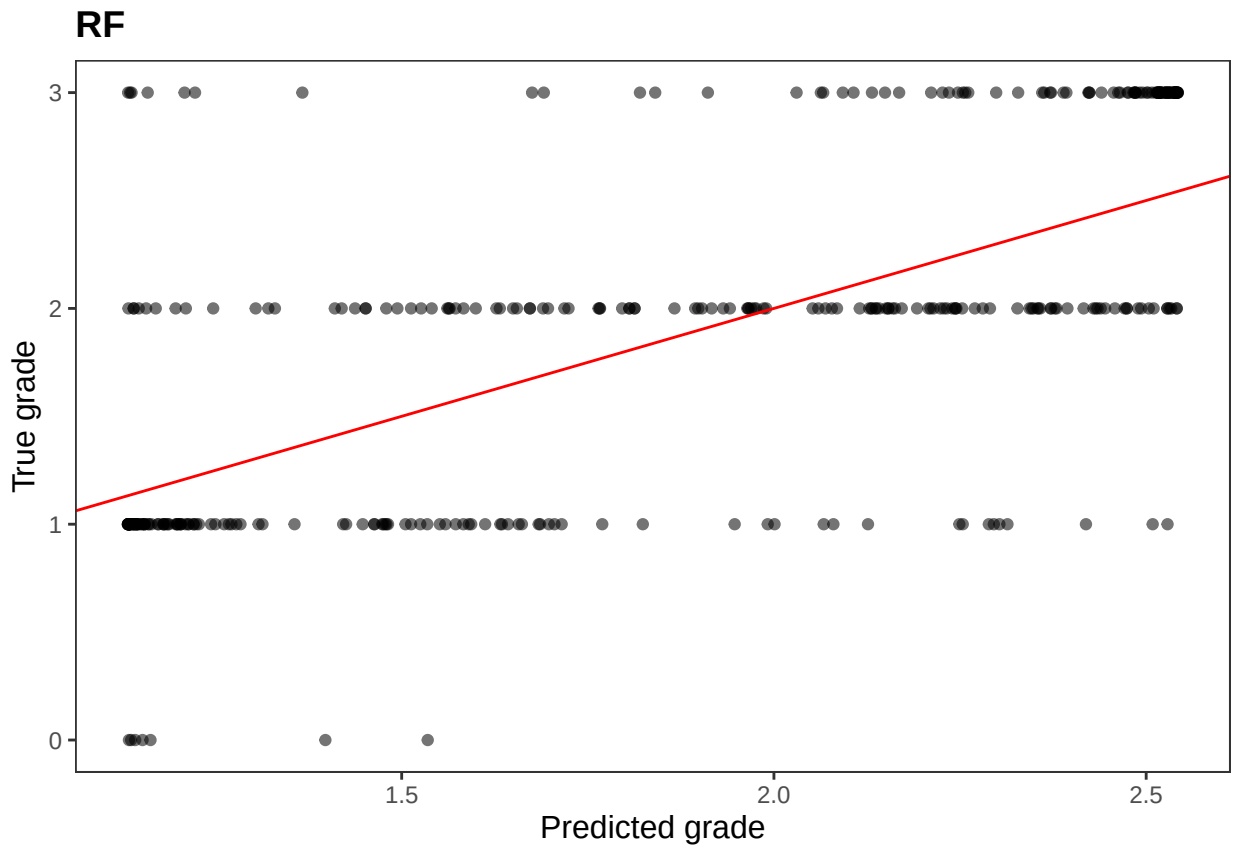

```
#ggsave("new_pred_plot/DS3_true_vs_preds_LR.png")
```

```
DS3 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



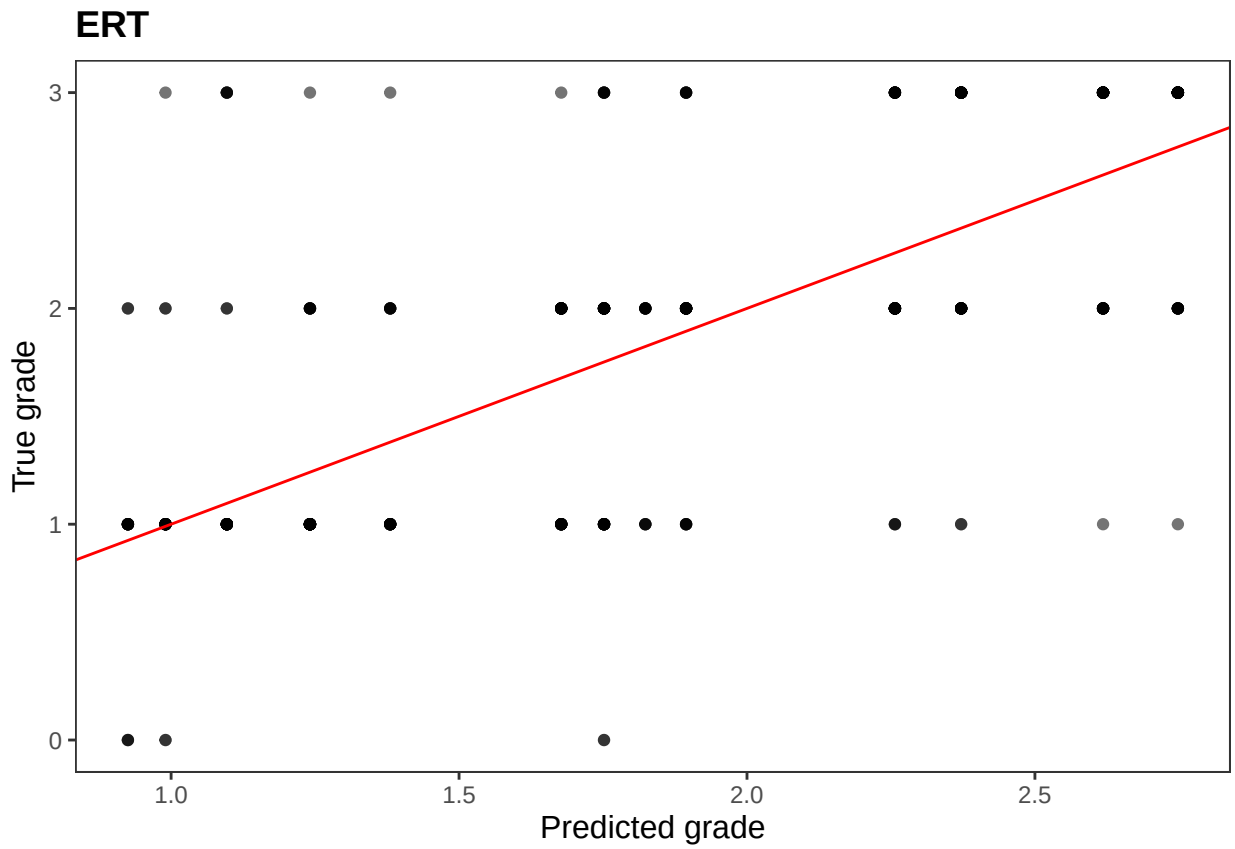

```
#ggsave("new_pred_plot/DS3_true_vs_preds_RT.png")
```

```
DS3 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS3_true_vs_preds_RF.png")
```

```
DS3 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS3_true_vs_preds_ERT.png")
```

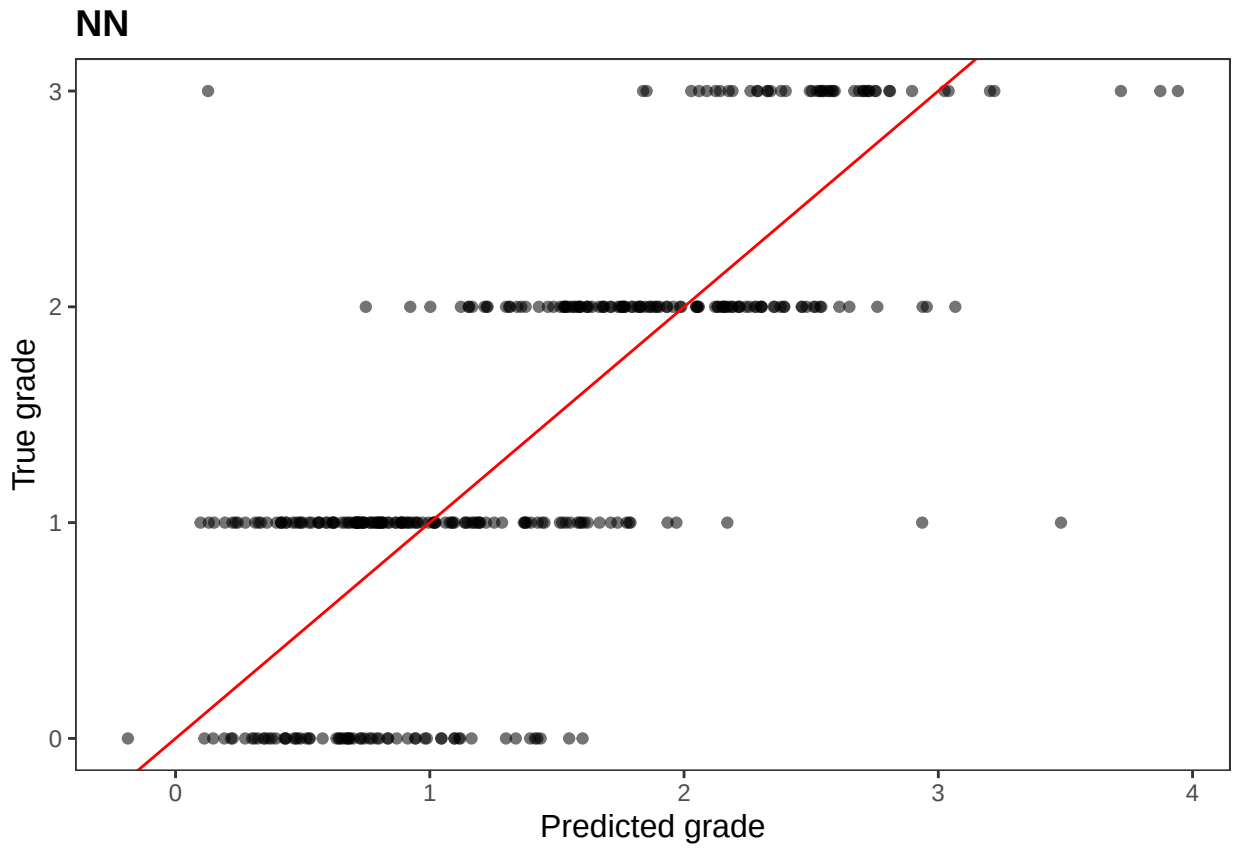
DS4

```
DS4 <- DS4_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS4)
```

```
##   y_test model    value
## 1      2    RF 1.589789
## 2      0    RF 1.790182
## 3      1    RF 1.709720
## 4      3    RF 2.325626
## 5      1    RF 1.549201
## 6      2    RF 1.713512
```

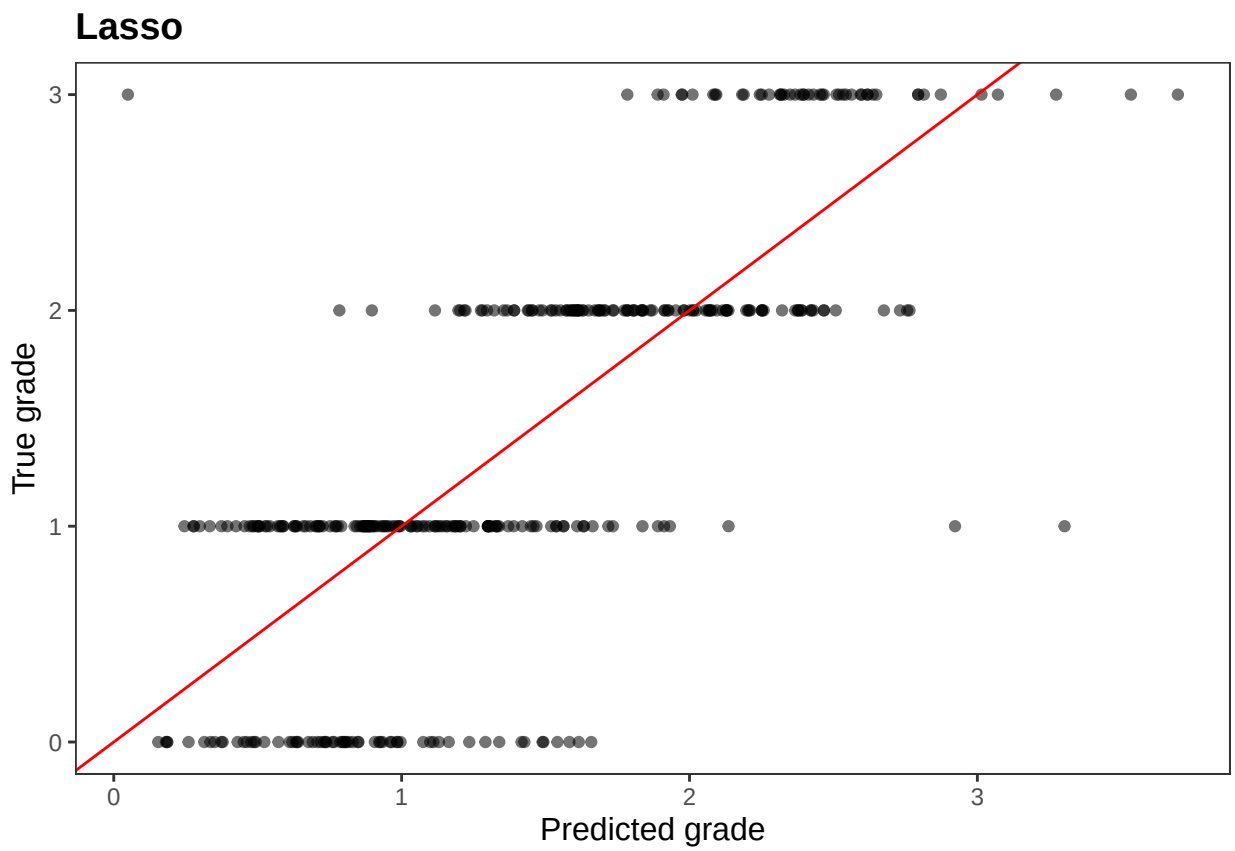
```
DS4 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



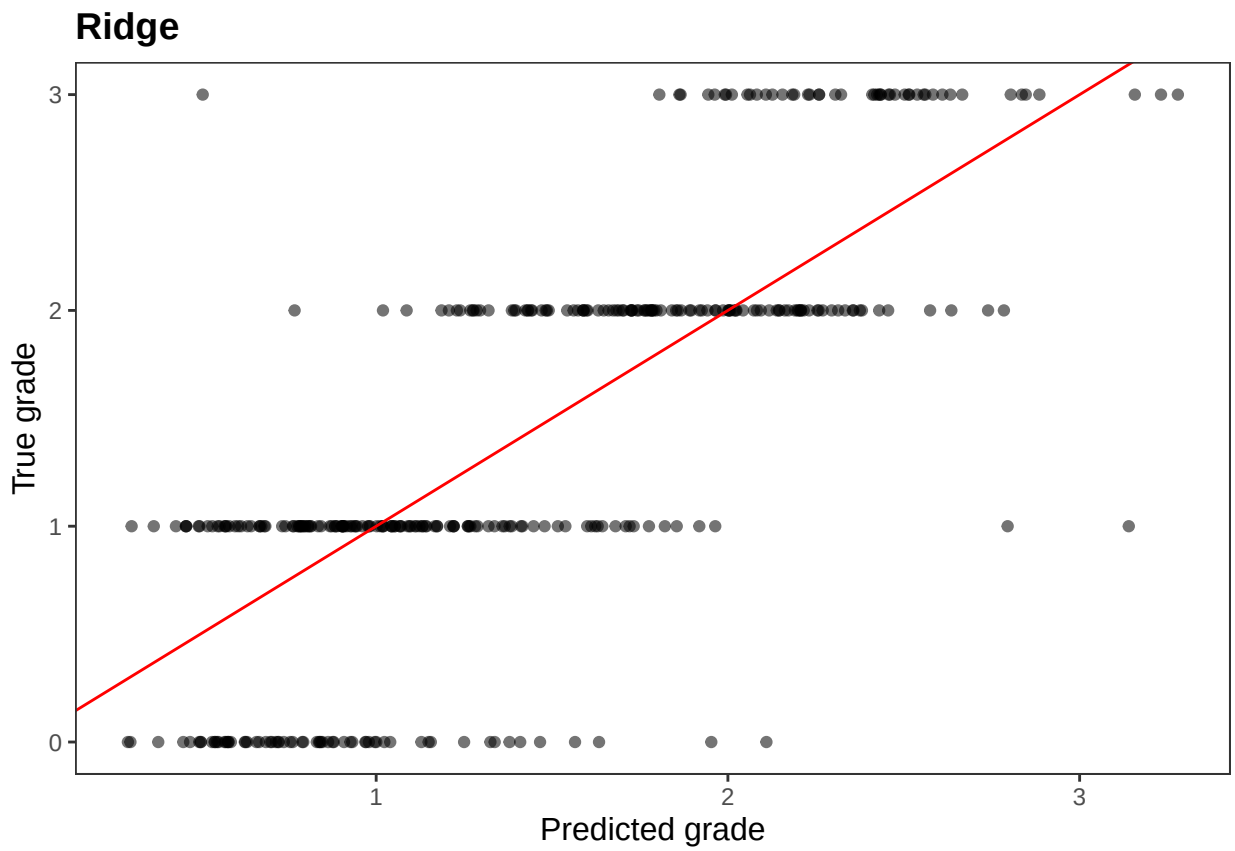

```
#ggsave("new_pred_plot/DS4_true_vs_preds_NN.png")
```

```
DS4 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



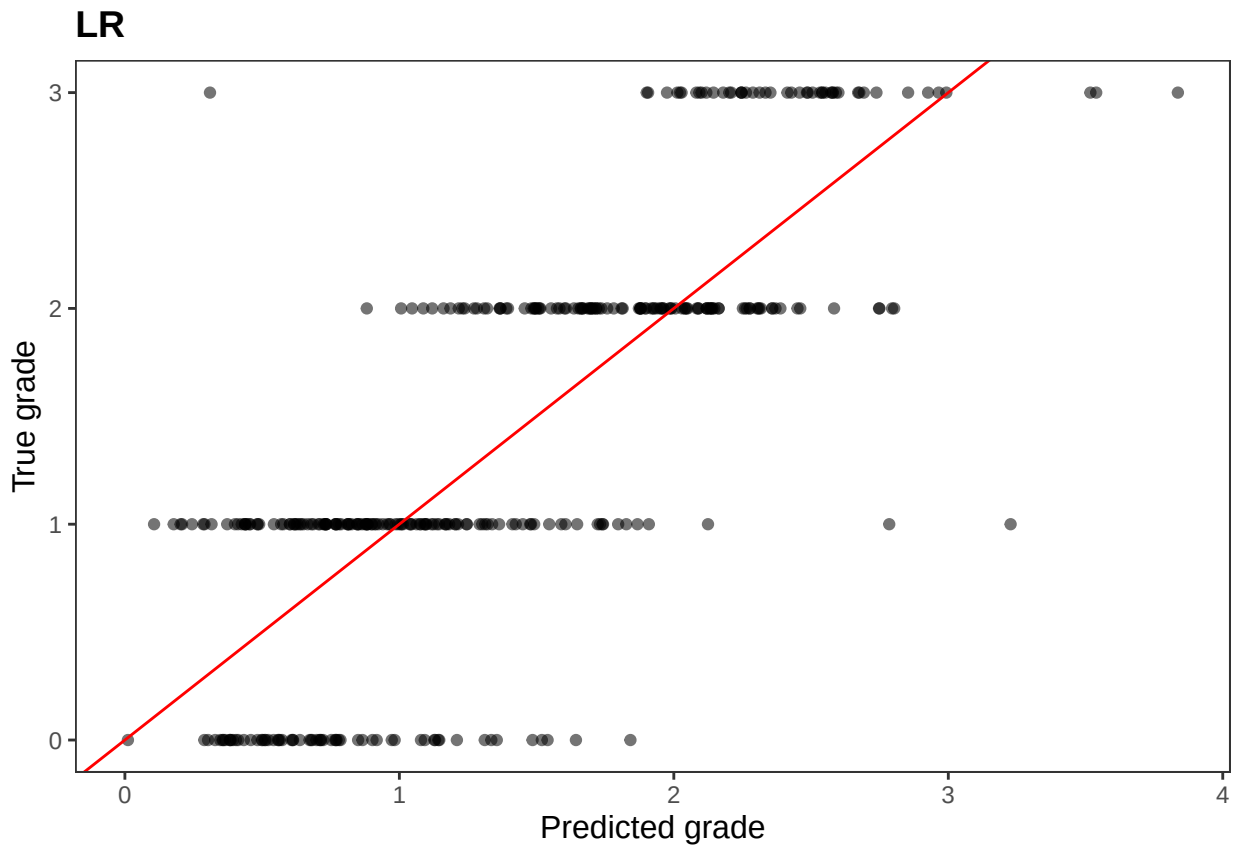

```
#ggsave("new_pred_plot/DS4_true_vs_preds_Lasso.png")
```

```
DS4 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



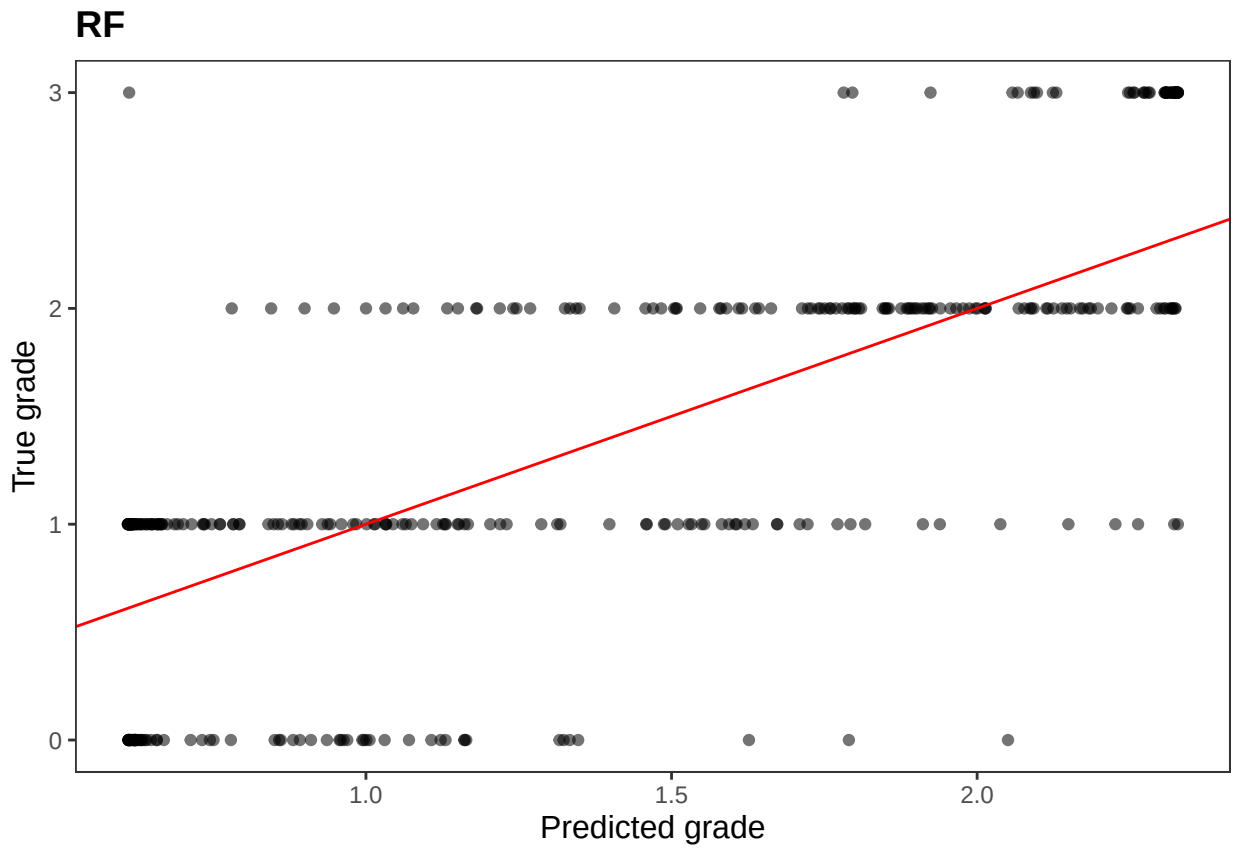

```
#ggsave("new_pred_plot/DS4_true_vs_preds_Ridge.png")
```

```
DS4 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



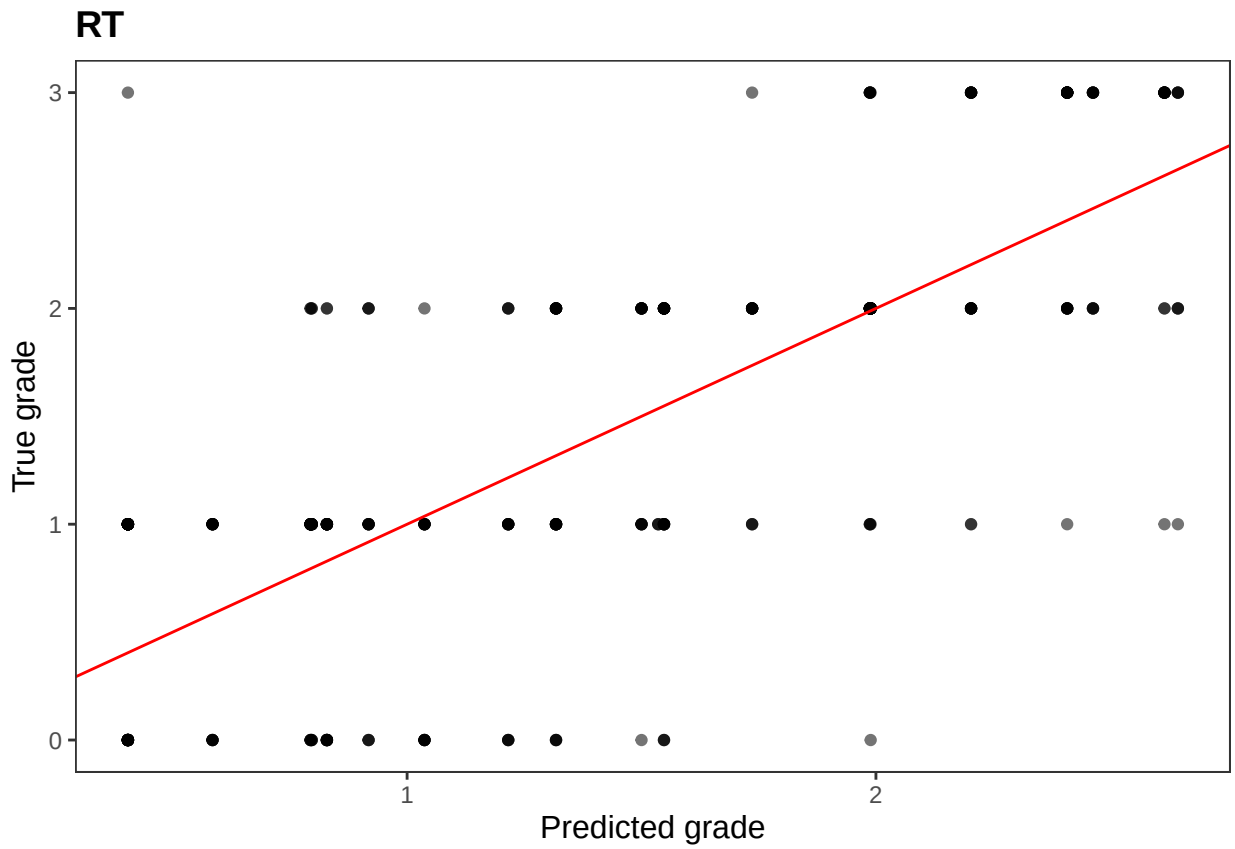

```
#ggsave("new_pred_plot/DS4_true_vs_preds_LR.png")
```

```
DS4 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



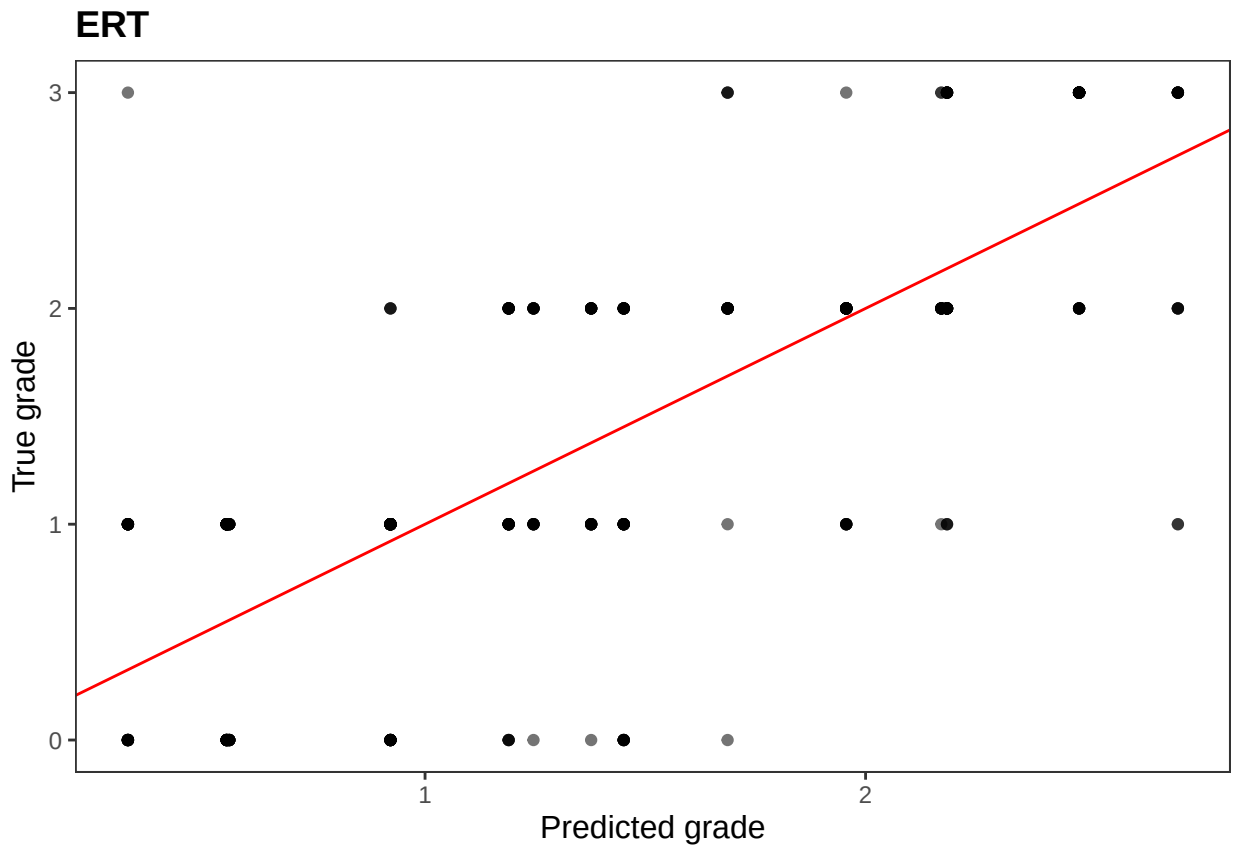

```
#ggsave("new_pred_plot/DS4_true_vs_preds_RF.png")
```

```
DS4 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS4_true_vs_preds_RT.png")
```

```
DS4 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS4_true_vs_preds_ERT.png")
```

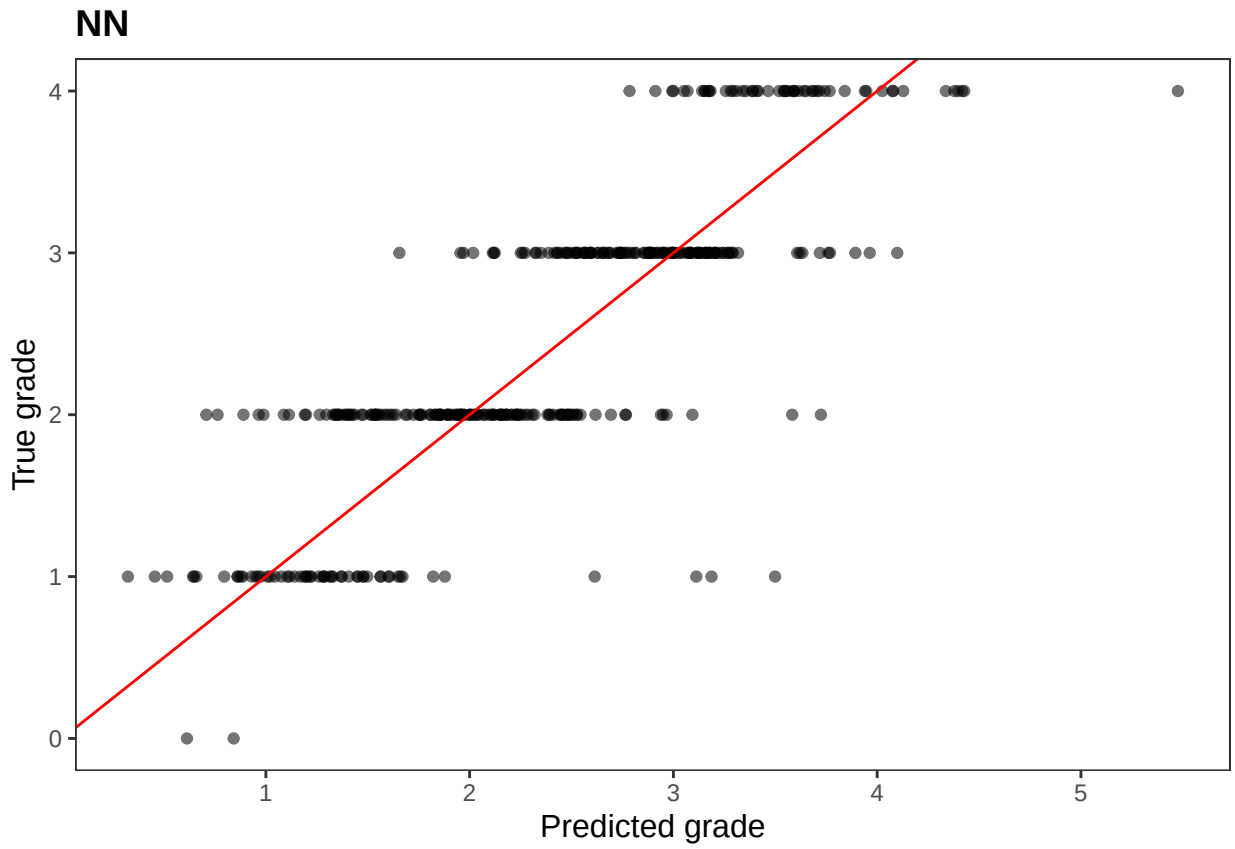
DS5

```
DS5 <- DS5_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS5)
```

```
##   y_test model    value
## 1      3    RF 2.251878
## 2      2    RF 2.022082
## 3      4    RF 3.406376
## 4      3    RF 3.402557
## 5      3    RF 3.071332
## 6      1    RF 1.949035
```

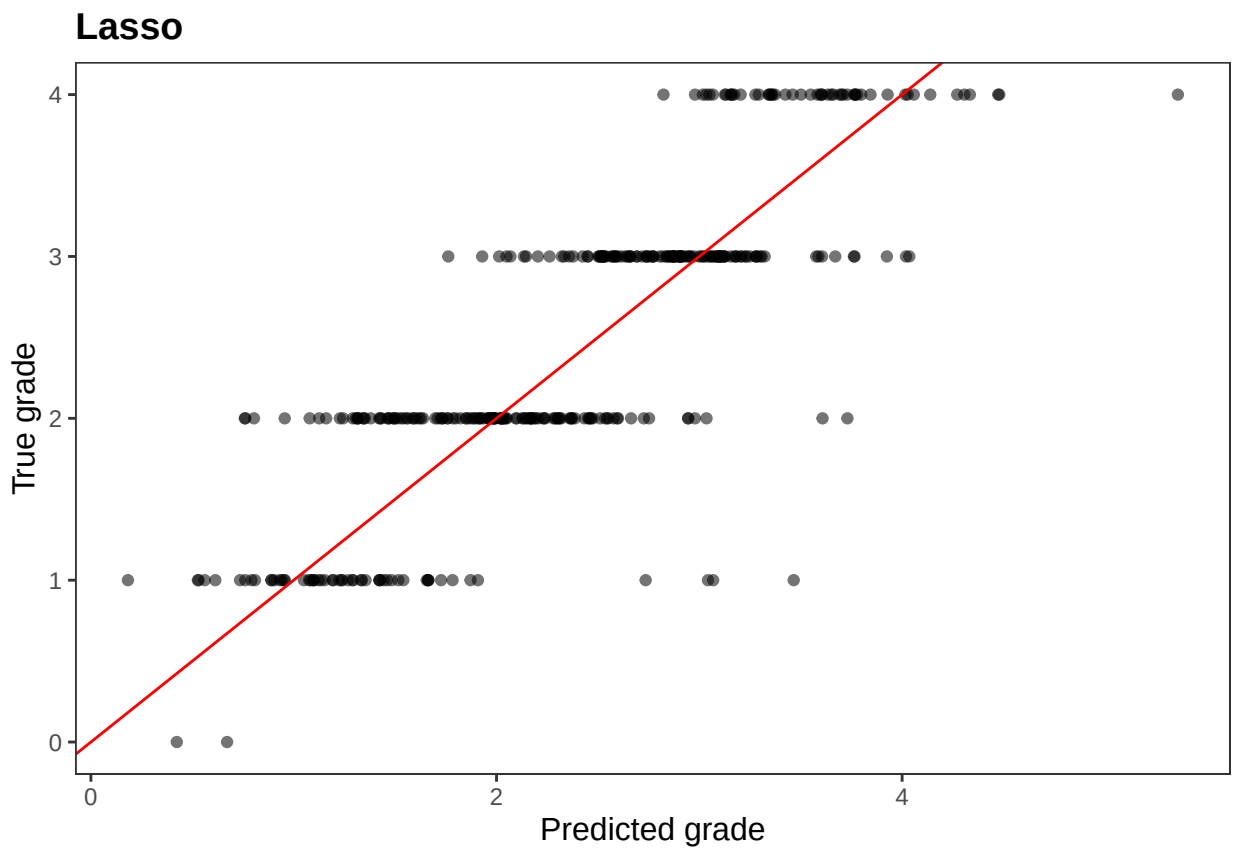
```
DS5 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



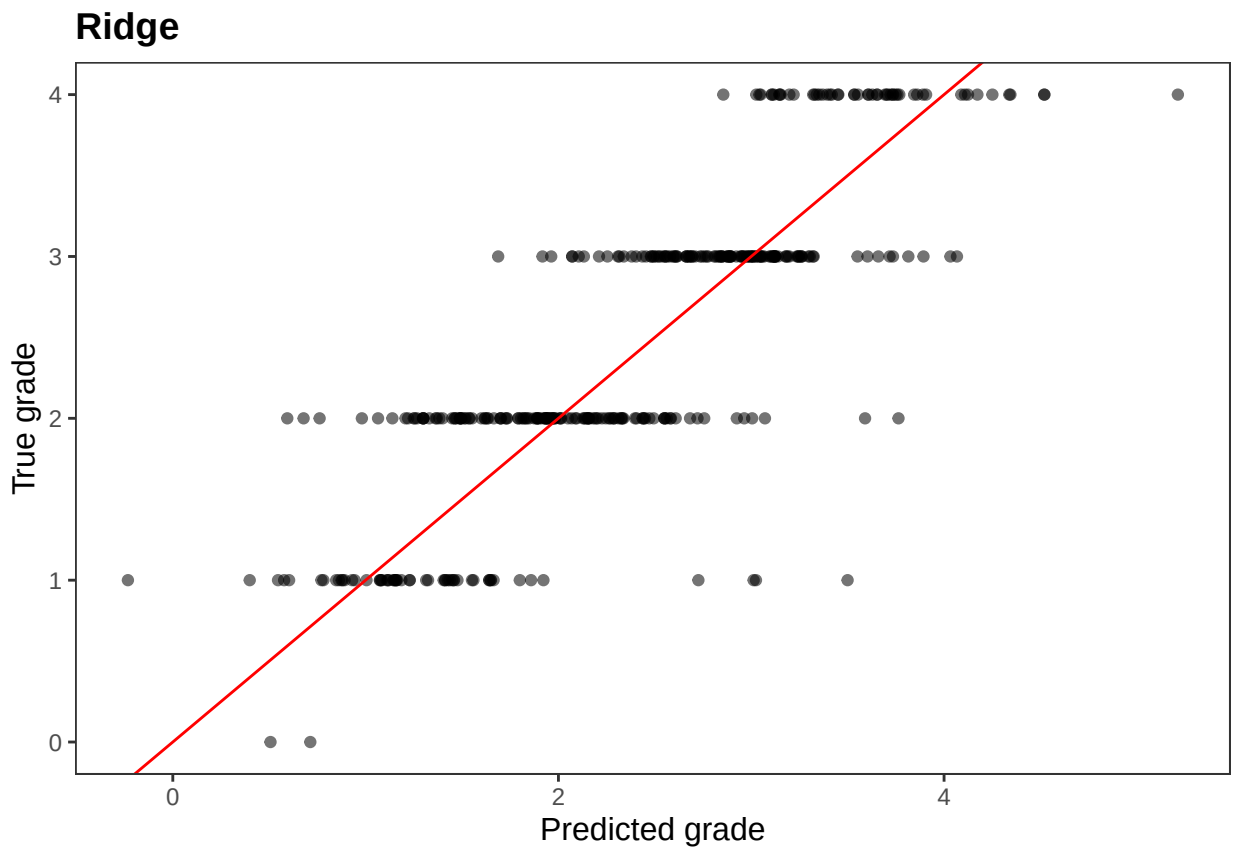

```
#ggsave("new_pred_plot/DS5_true_vs_preds_NN.png")
```

```
DS5 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



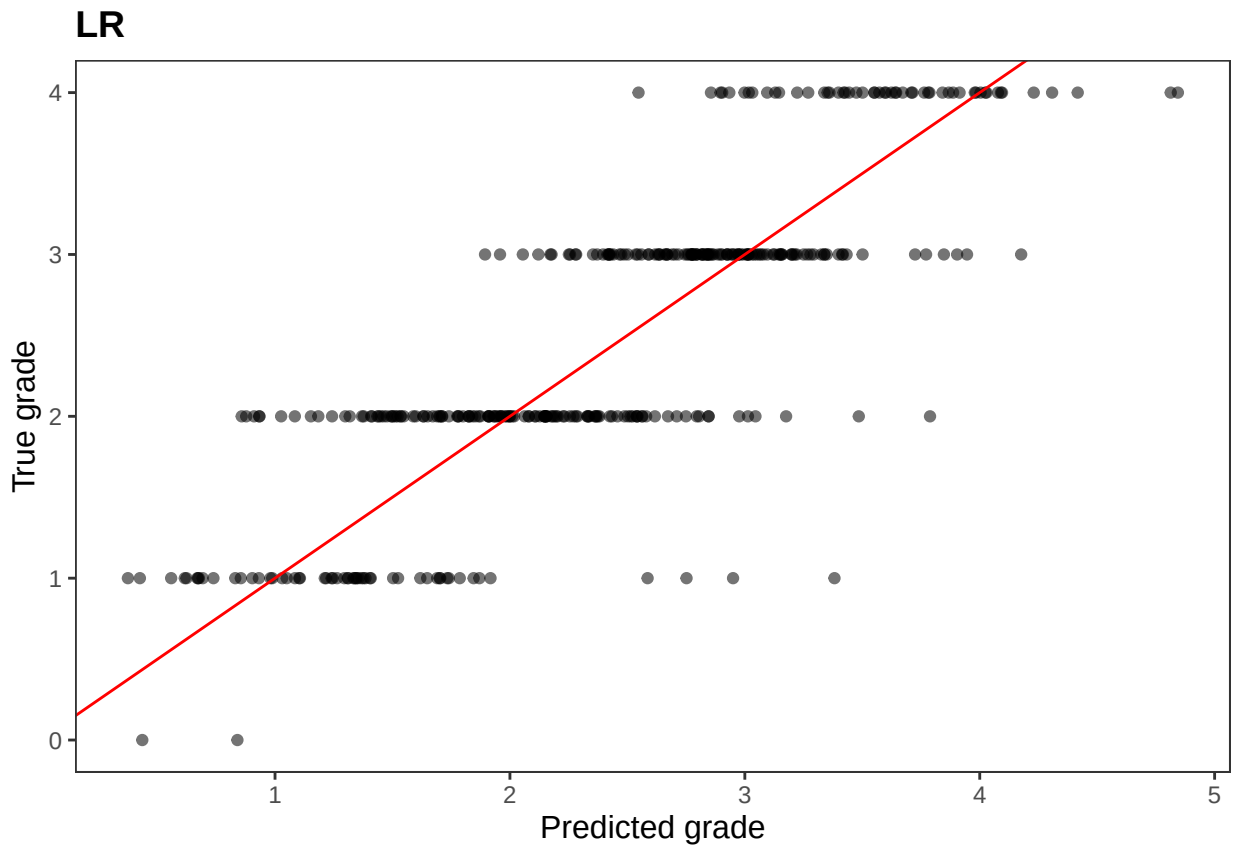

```
#ggsave("new_pred_plot/DS5_true_vs_preds_Lasso.png")
```

```
DS5 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



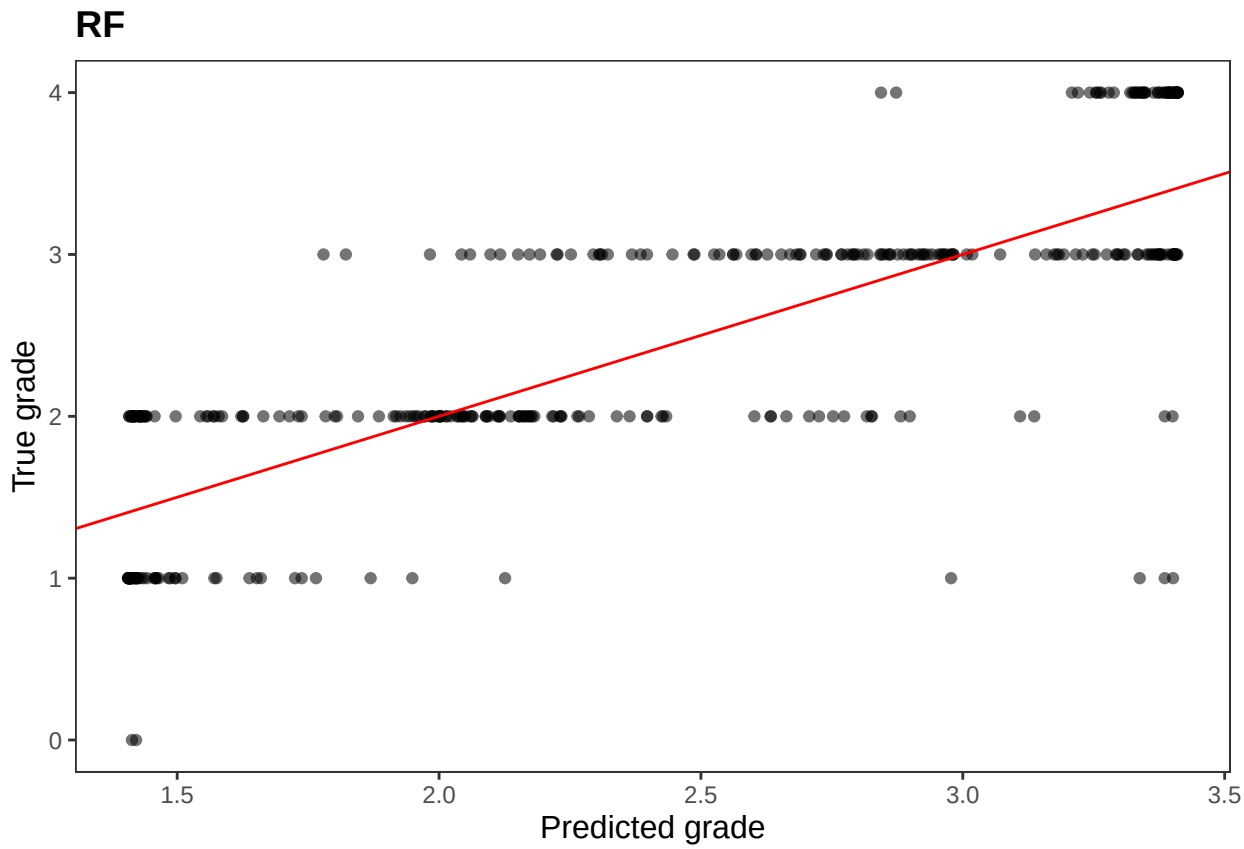

```
#ggsave("new_pred_plot/DS5_true_vs_preds_Ridge.png")
```

```
DS5 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



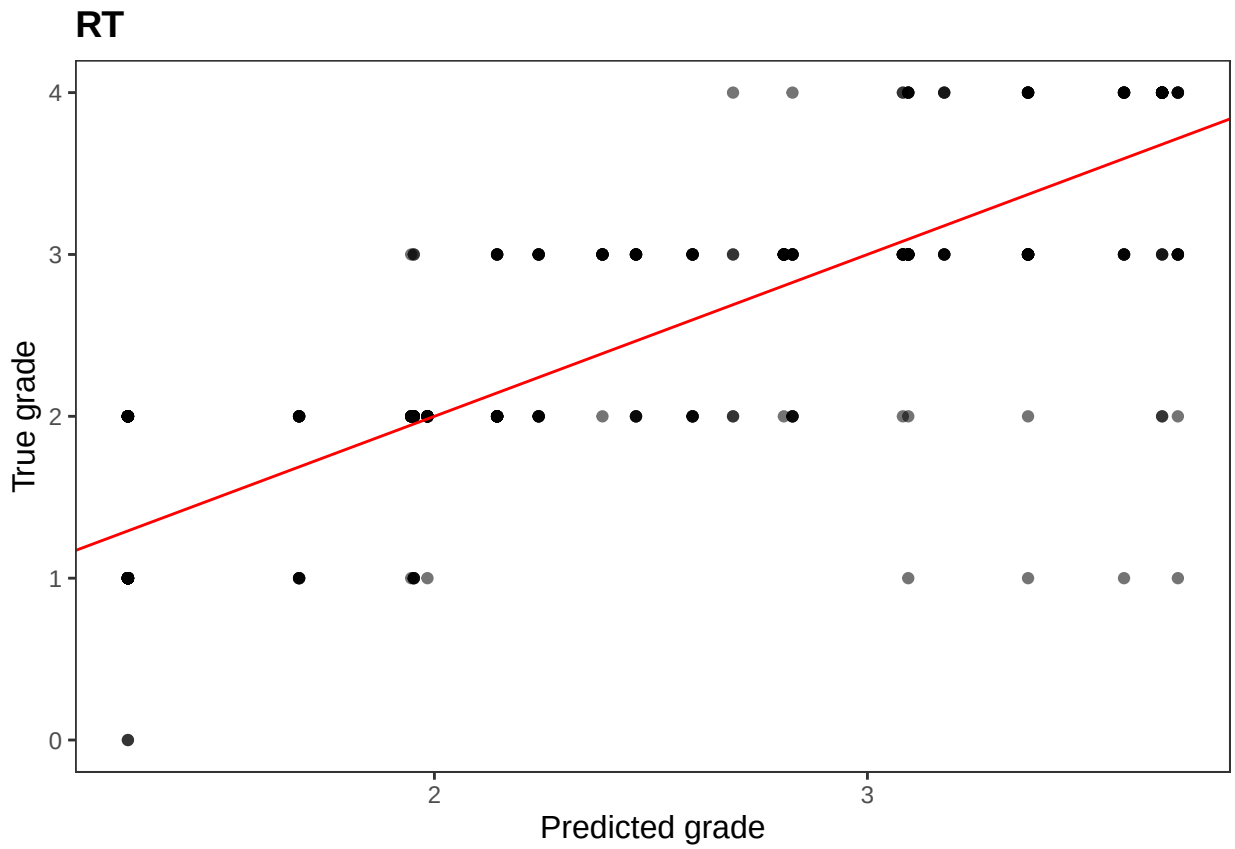

```
#ggsave("new_pred_plot/DS5_true_vs_preds_LR.png")
```

```
DS5 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



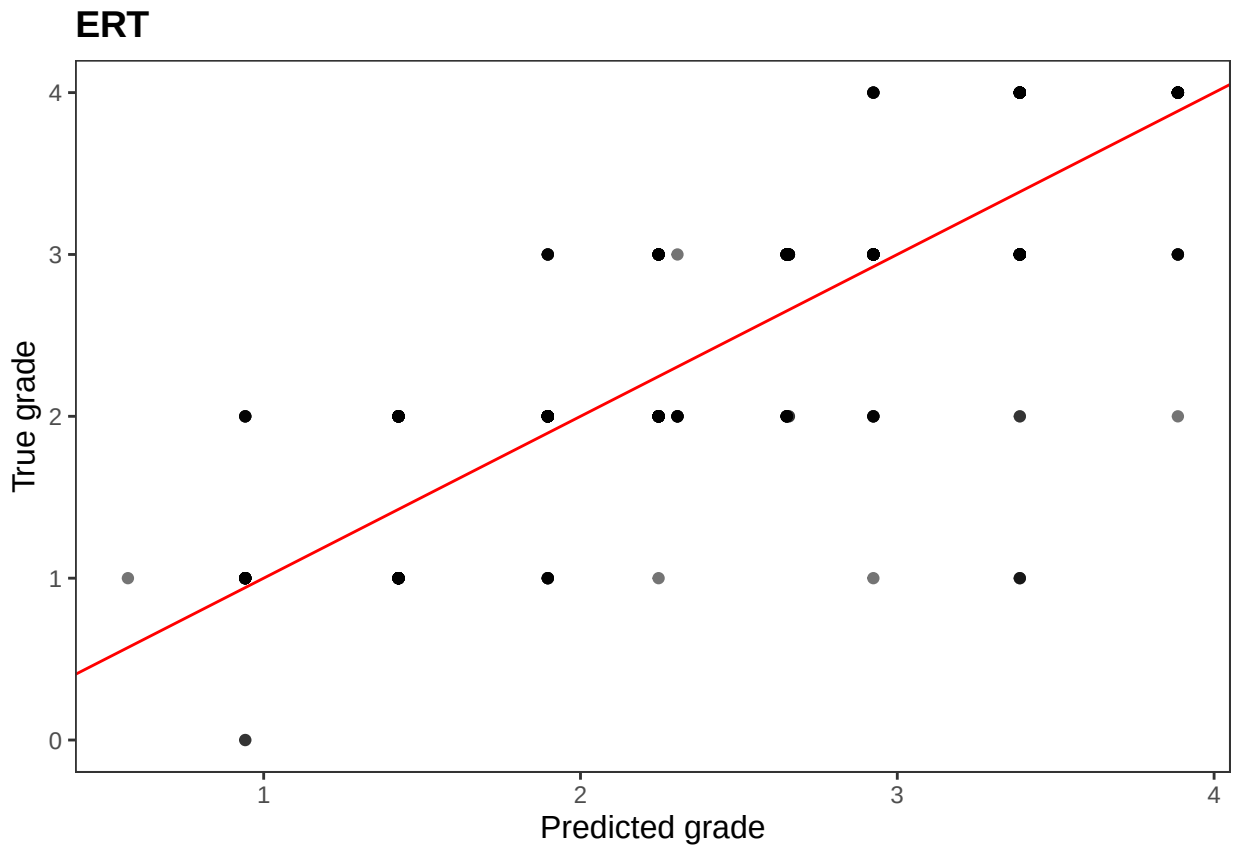

```
#ggsave("new_pred_plot/DS5_true_vs_preds_RF.png")
```

```
DS5 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS5_true_vs_preds_RT.png")
```

```
DS5 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS5_true_vs_preds_ERT.png")
```

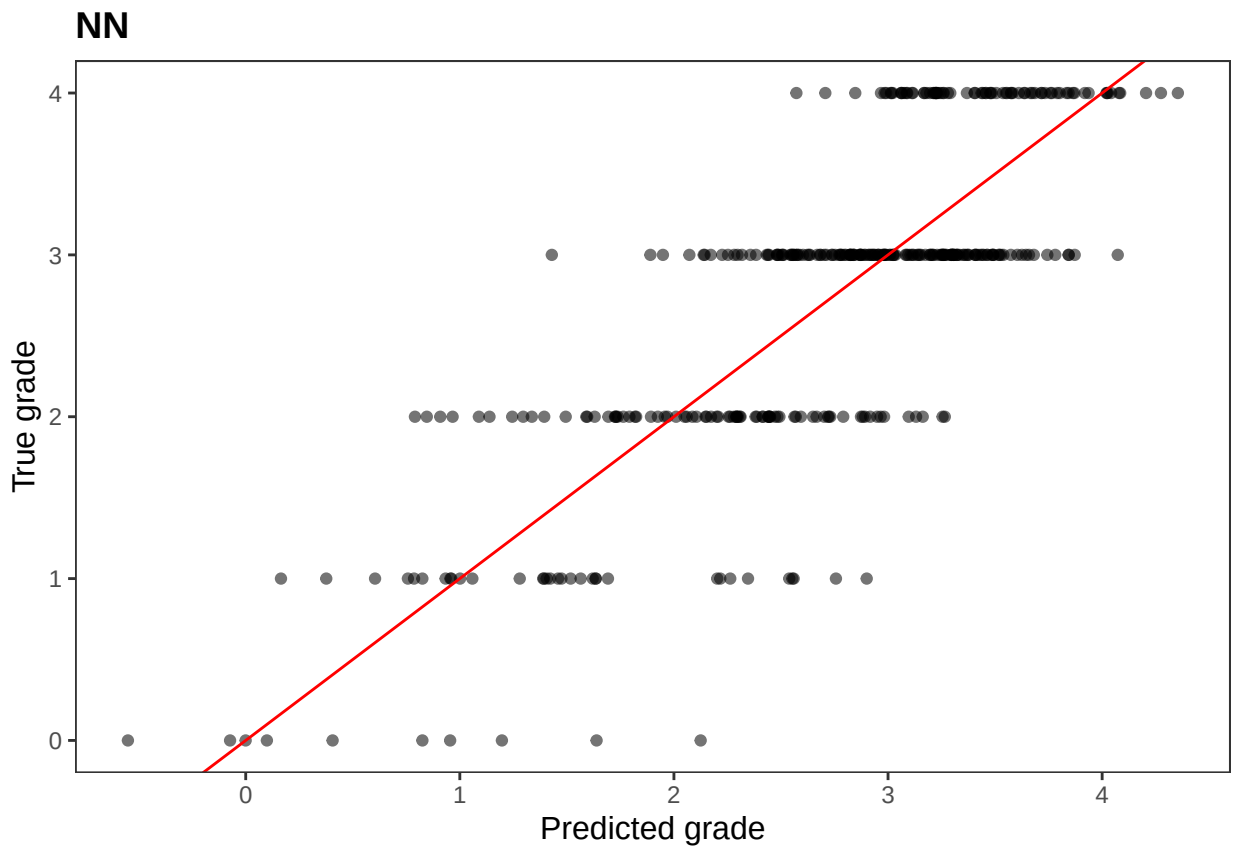
DS6

```
DS6 <- DS6_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS6)
```

```
##   y_test model    value
## 1      3    RF 3.012531
## 2      1    RF 1.814536
## 3      2    RF 1.942356
## 4      2    RF 2.676692
## 5      3    RF 2.967419
## 6      3    RF 3.020050
```

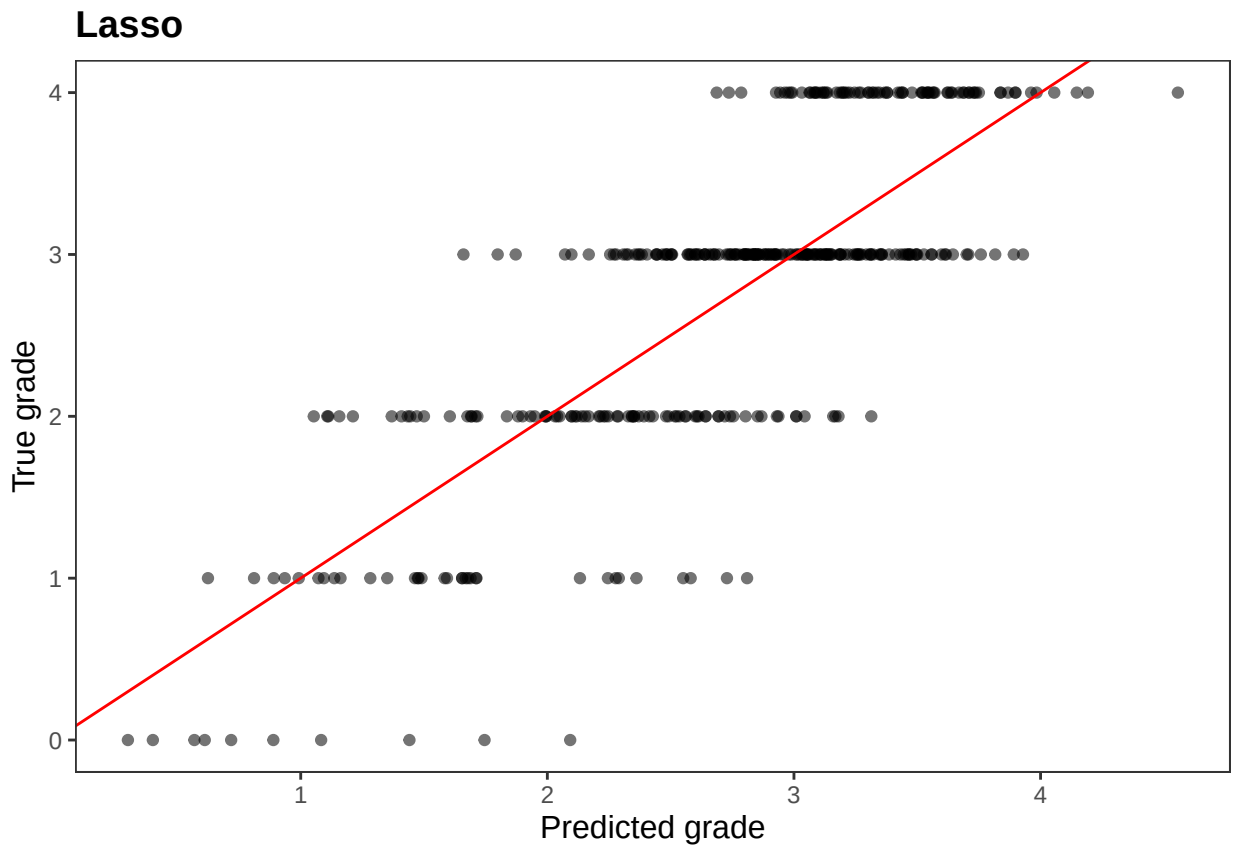
```
DS6 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



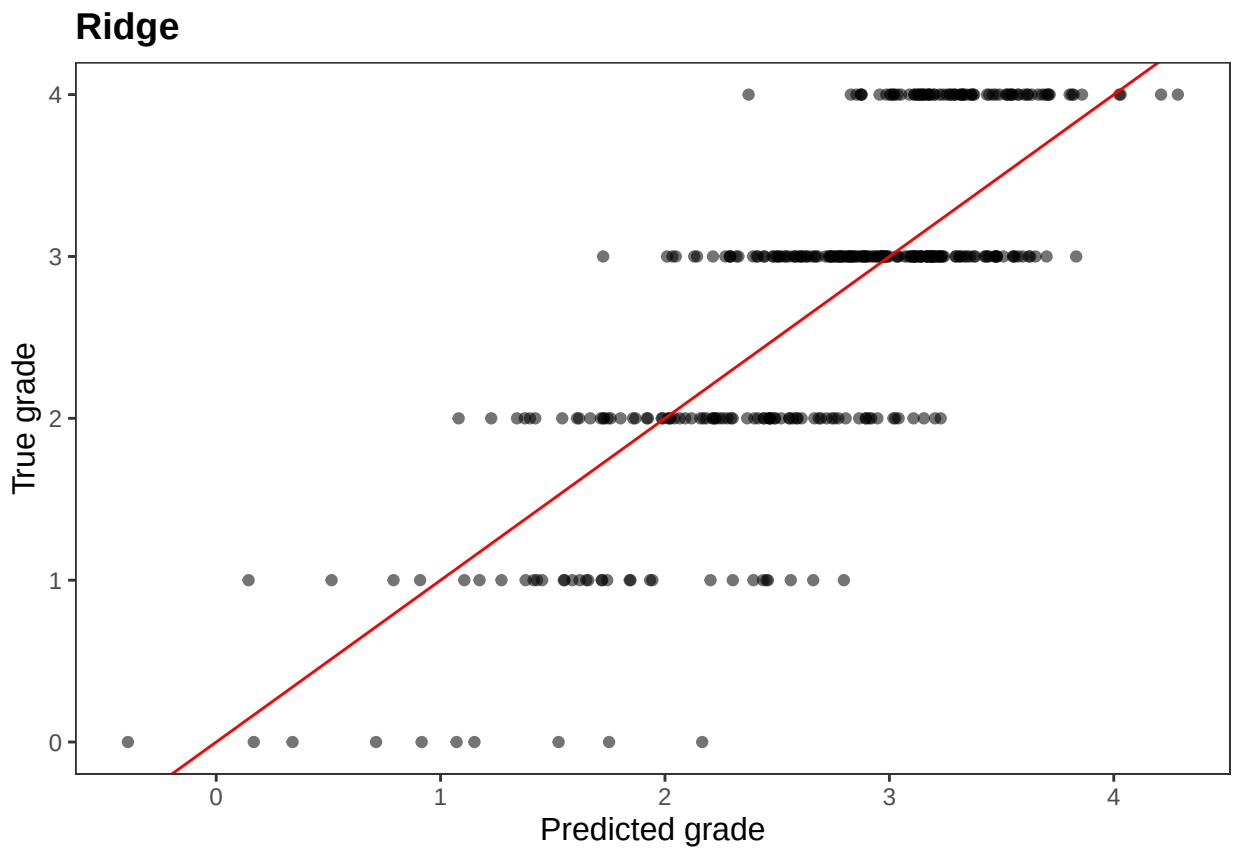

```
#ggsave("new_pred_plot/DS6_true_vs_preds_NN.png")
```

```
DS6 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



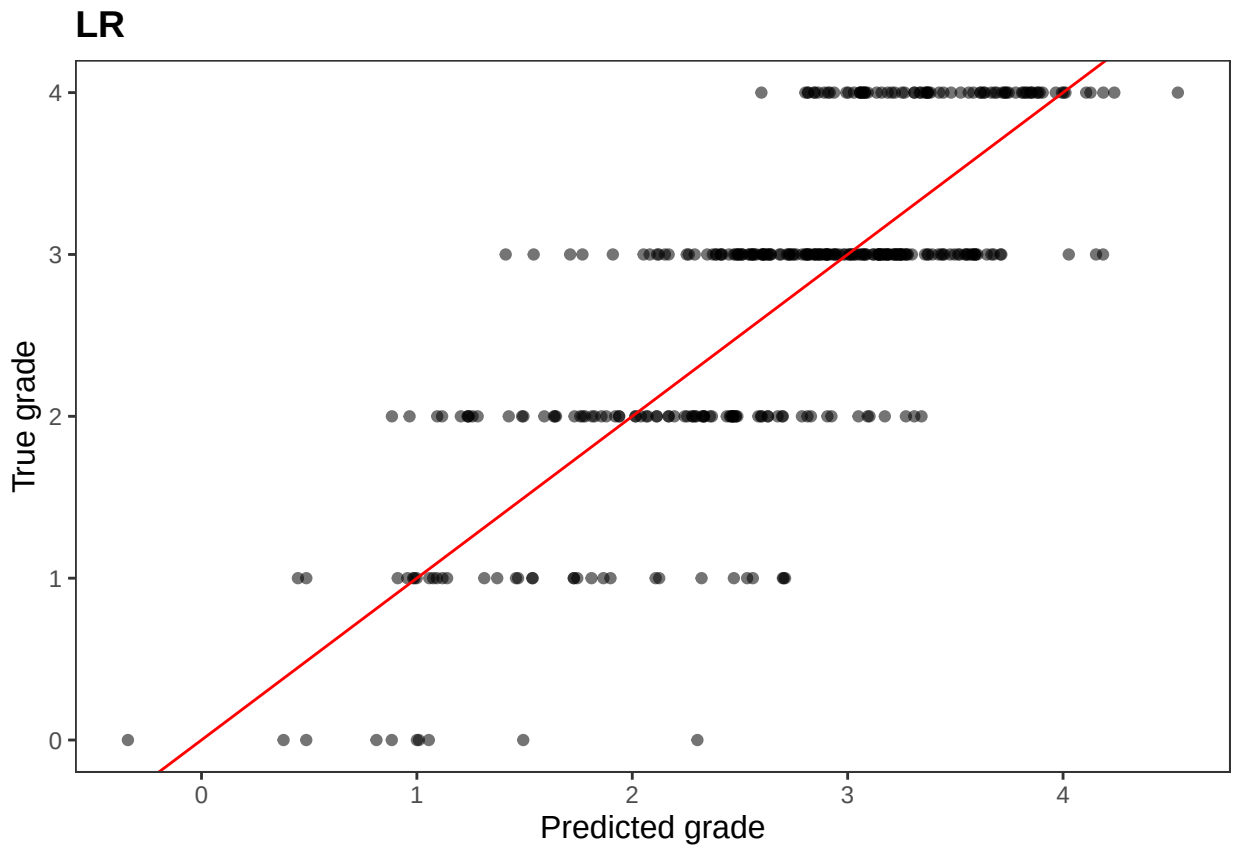

```
#ggsave("new_pred_plot/DS6_true_vs_preds_Lasso.png")
```

```
DS6 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



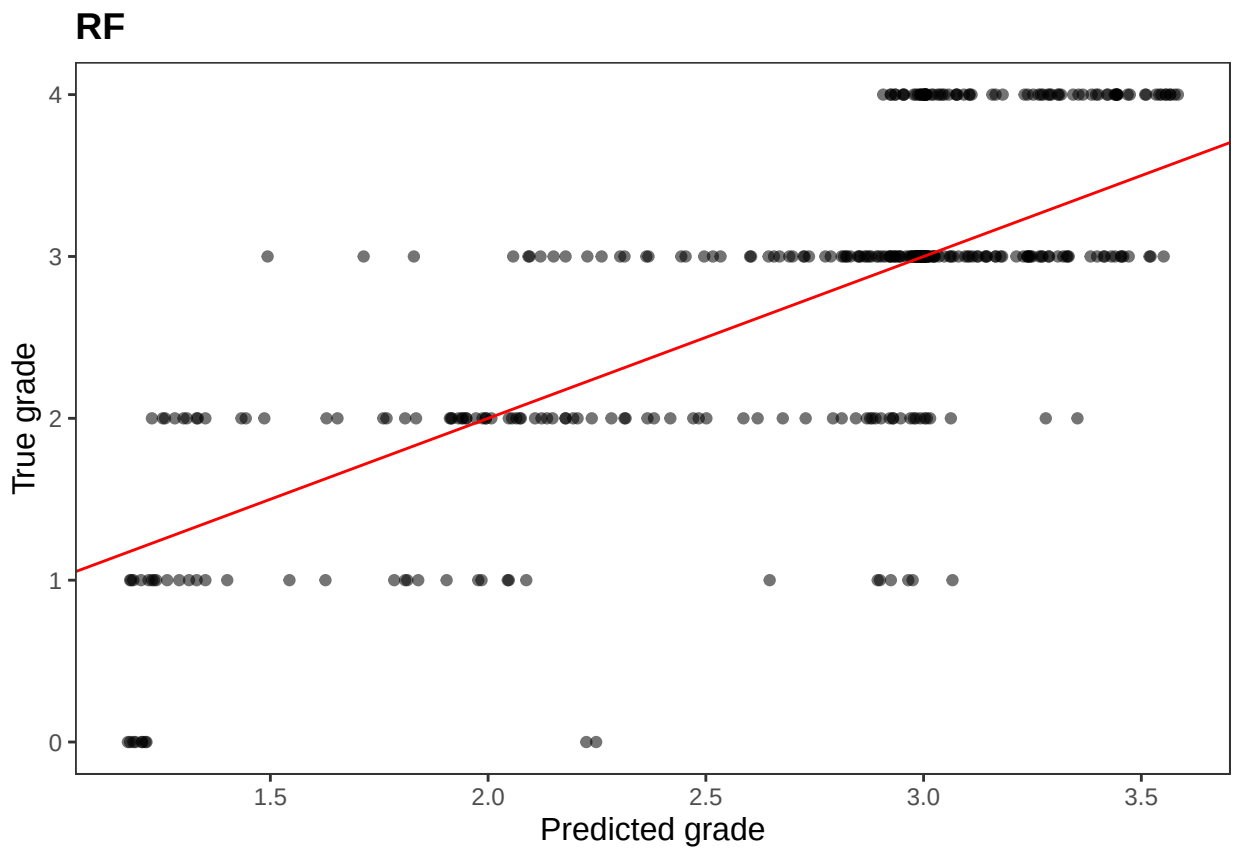

```
#ggsave("new_pred_plot/DS6_true_vs_preds_Ridge.png")
```

```
DS6 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



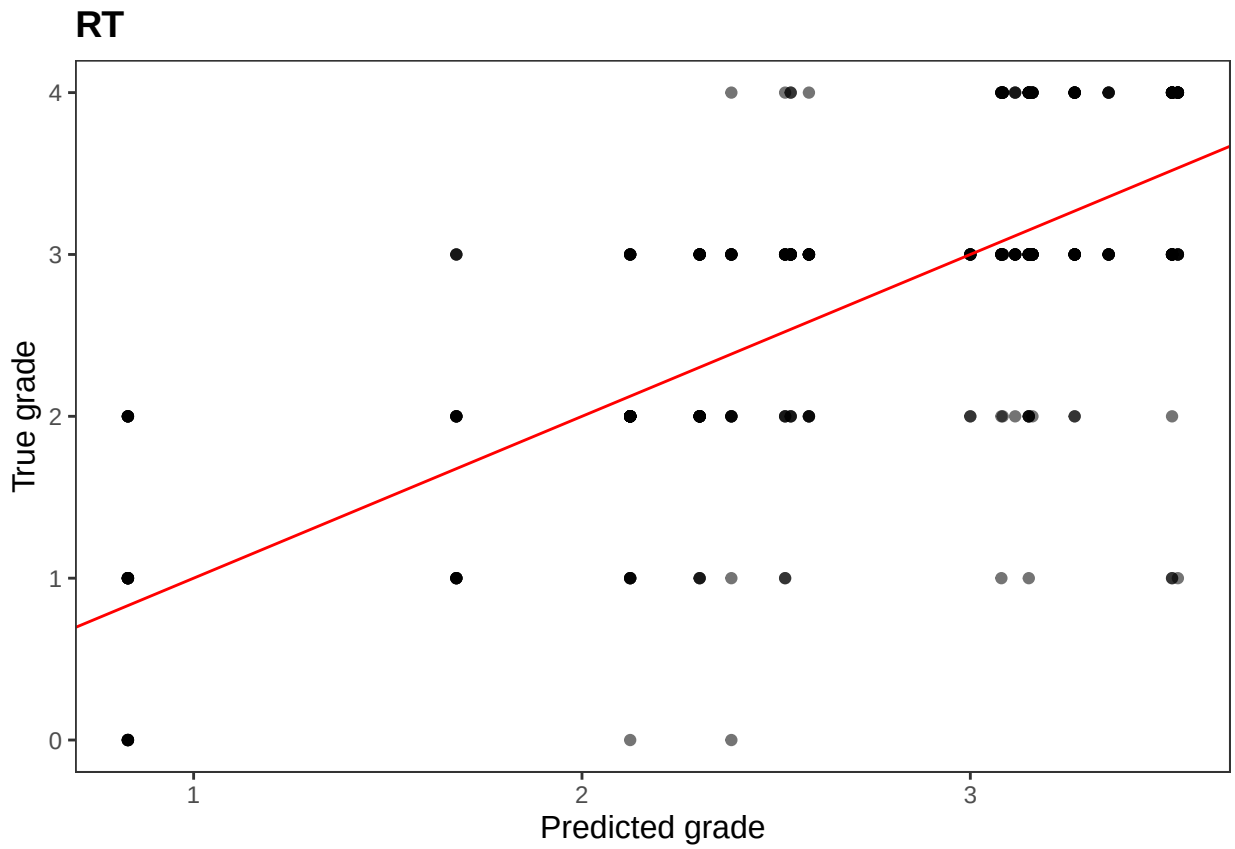

```
#ggsave("new_pred_plot/DS6_true_vs_preds_LR.png")
```

```
DS6 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



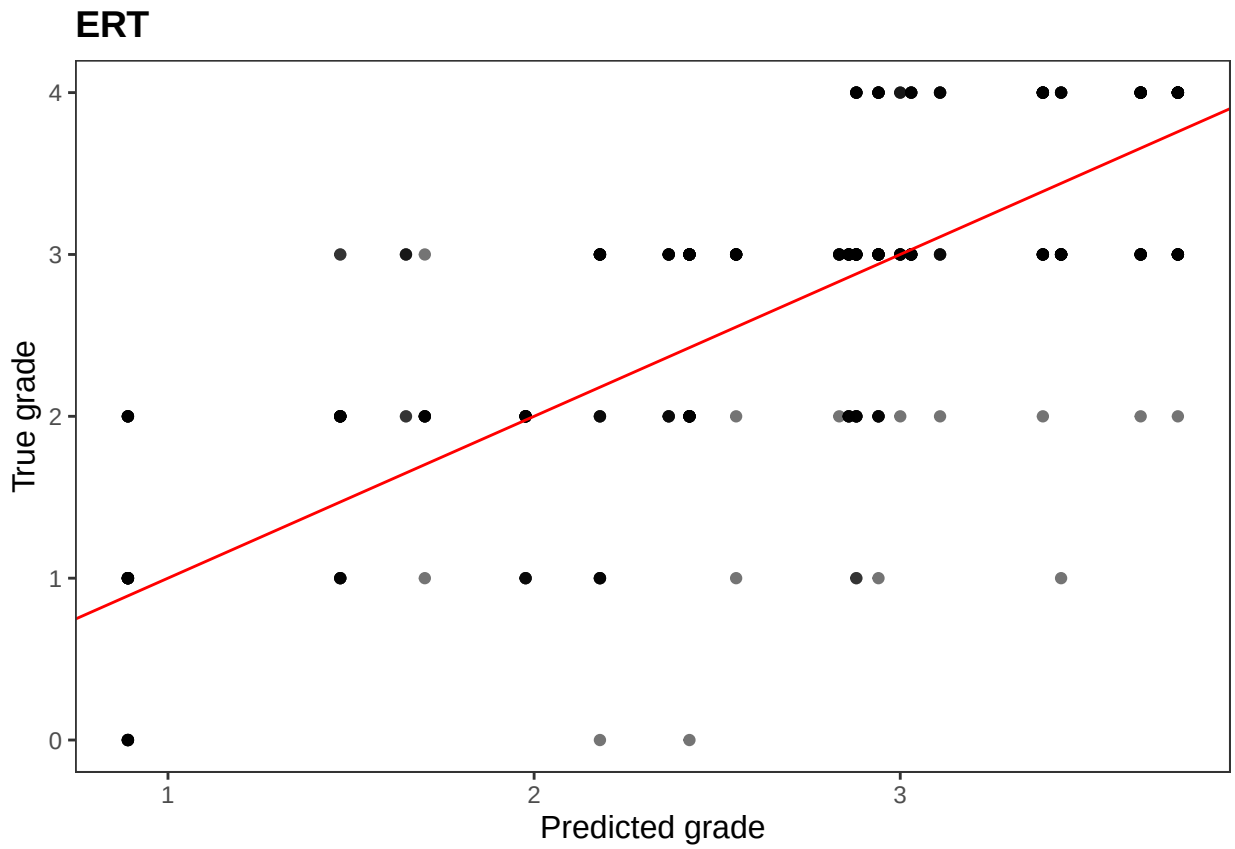

```
#ggsave("new_pred_plot/DS6_true_vs_preds_RF.png")
```

```
DS6 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS6_true_vs_preds_RT.png")
```

```
DS6 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS6_true_vs_preds_ERT.png")
```

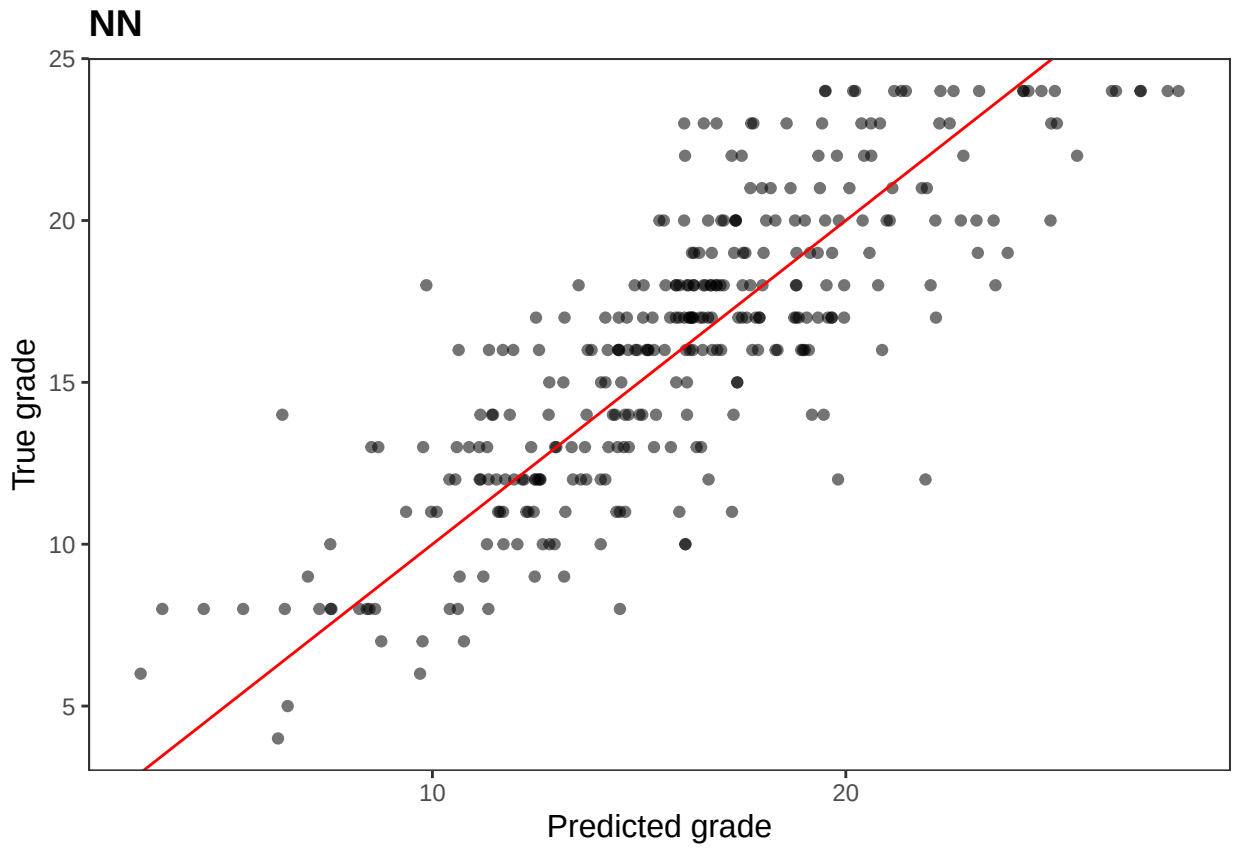
DS7

```
DS7 <- DS7_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS7)
```

```
##   y_test model     value
## 1      8    RF  9.419799
## 2     18    RF 20.750627
## 3     14    RF 13.706767
## 4     14    RF 19.827068
## 5     13    RF 14.952381
## 6     11    RF 11.175439
```

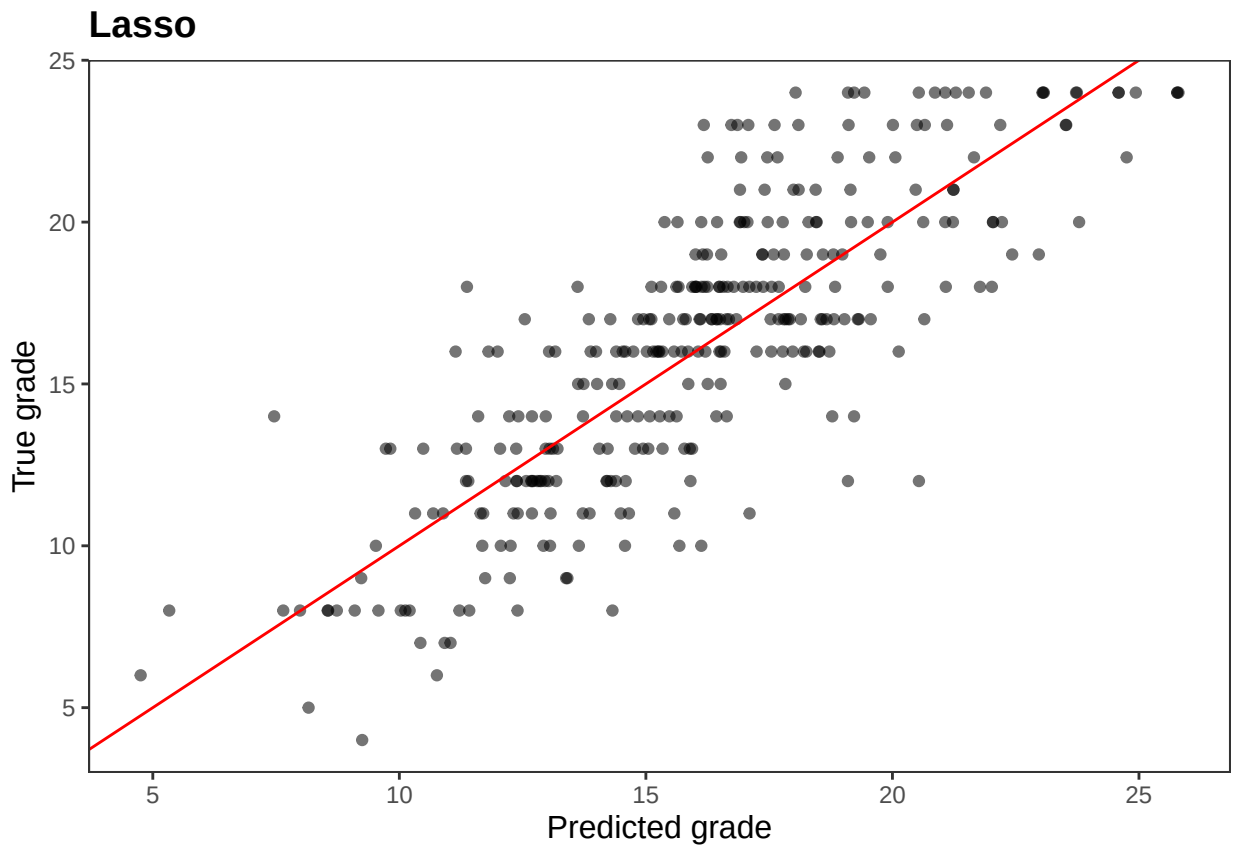
```
DS7 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



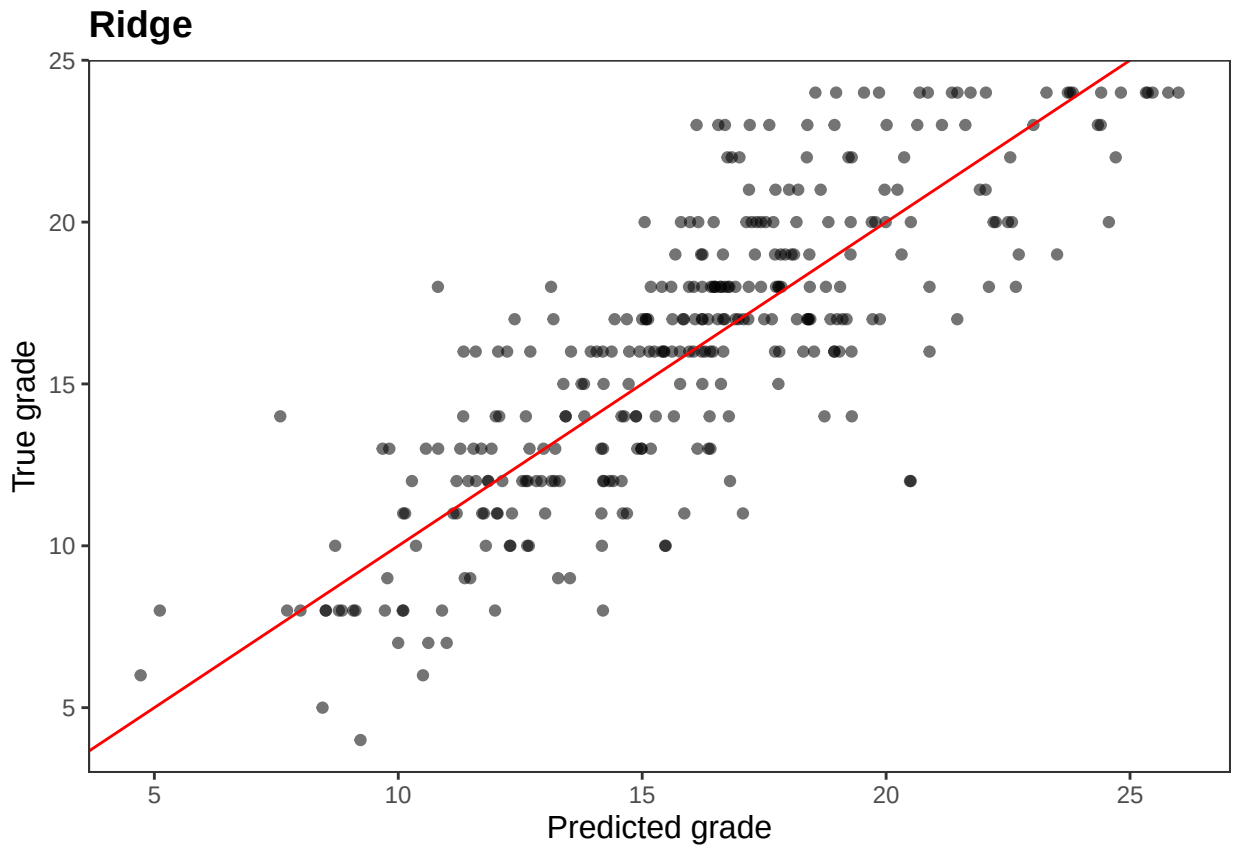

```
#ggsave("new_pred_plot/DS7_true_vs_preds_NN.png")
```

```
DS7 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



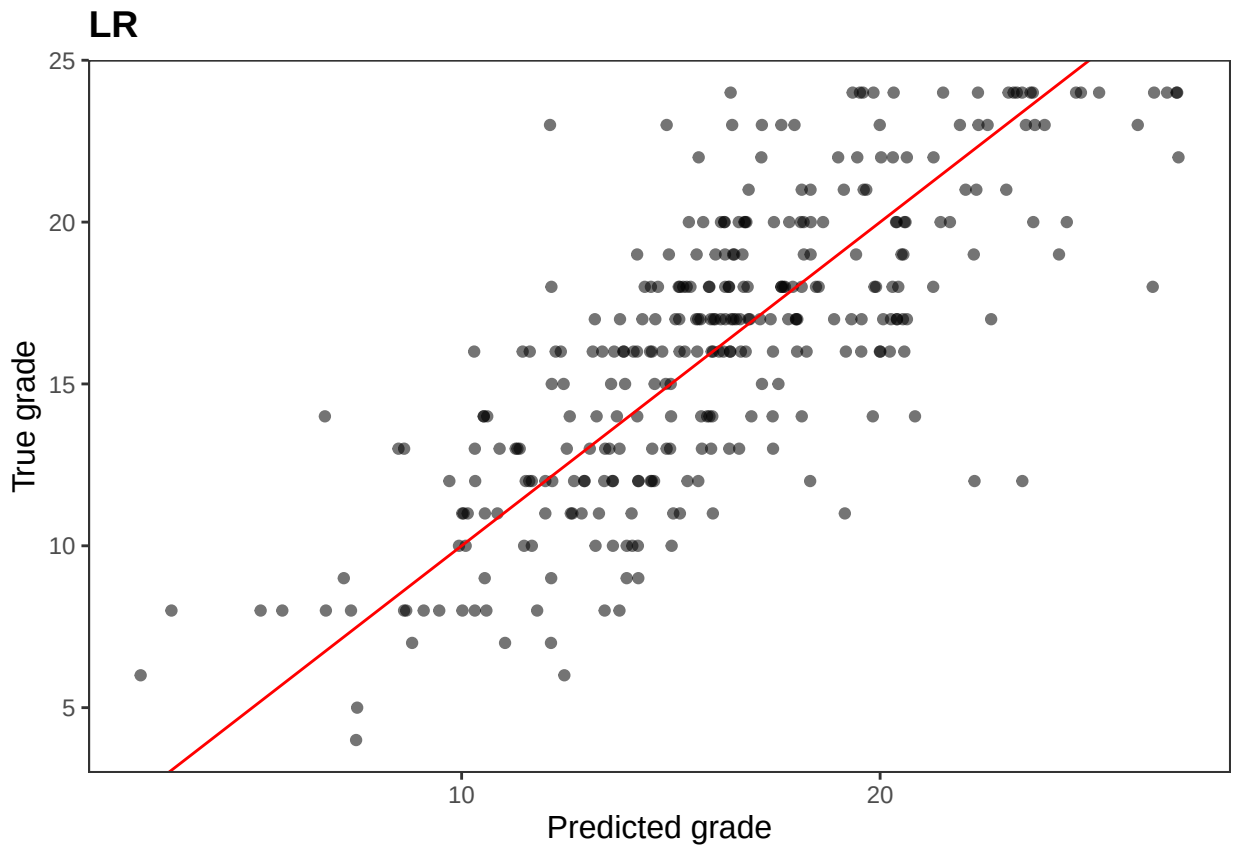

```
#ggsave("new_pred_plot/DS7_true_vs_preds_Lasso.png")
```

```
DS7 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



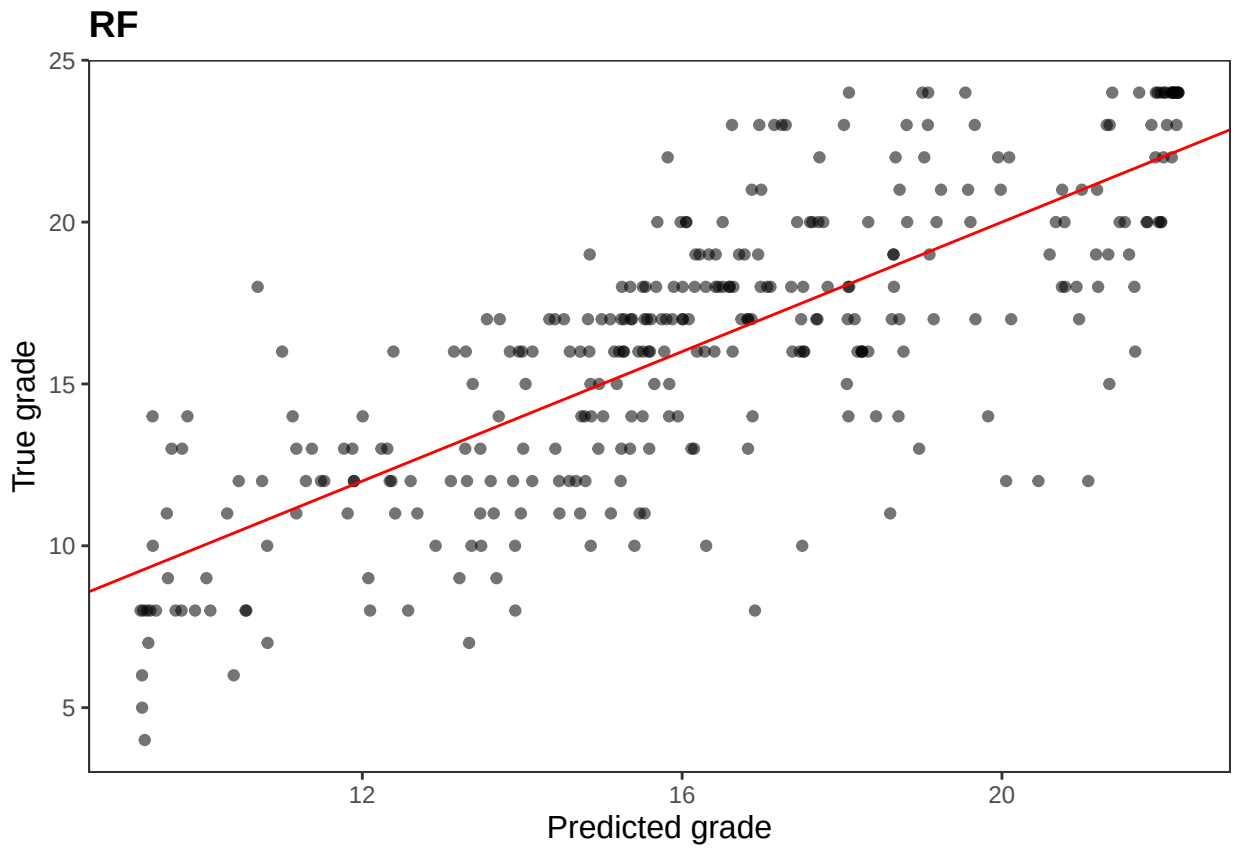

```
#ggsave("new_pred_plot/DS7_true_vs_preds_Ridge.png")
```

```
DS7 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



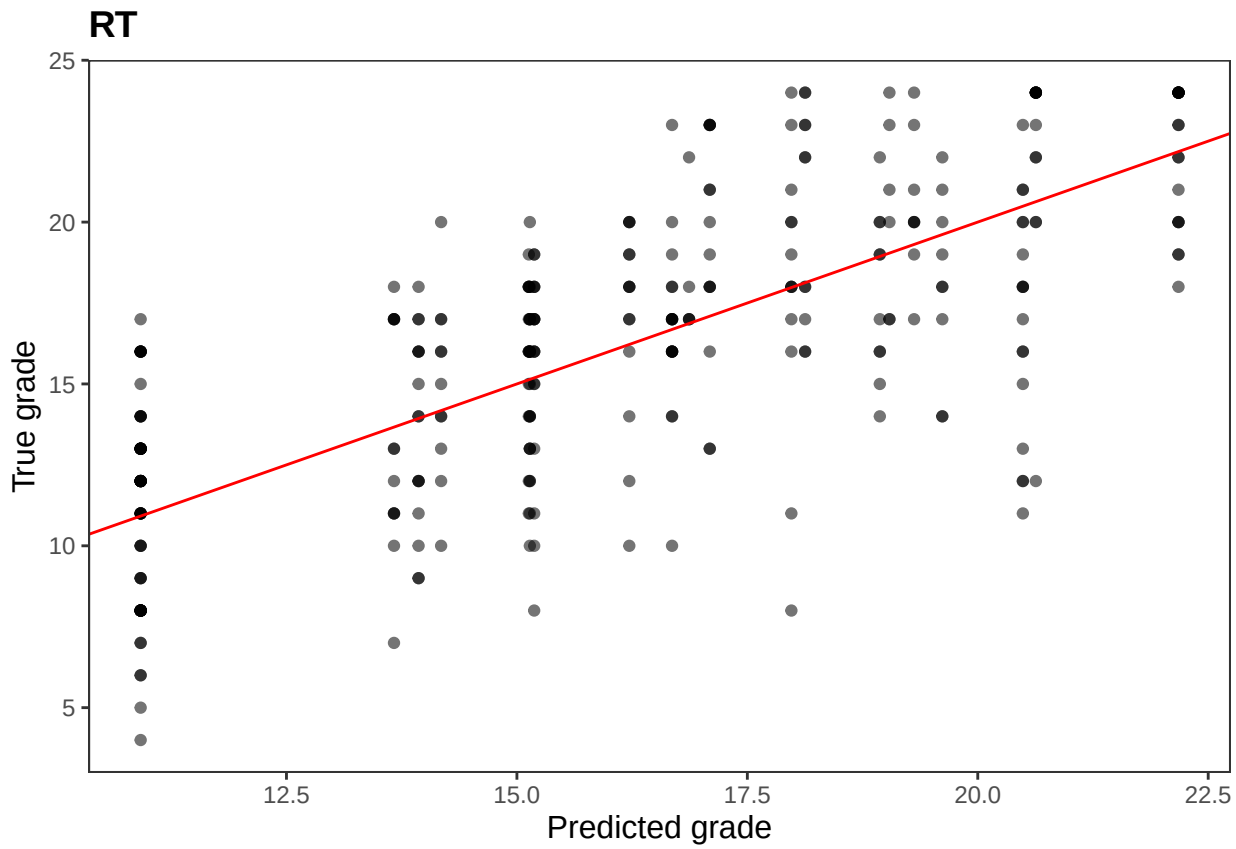

```
#ggsave("new_pred_plot/DS7_true_vs_preds_LR.png")
```

```
DS7 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



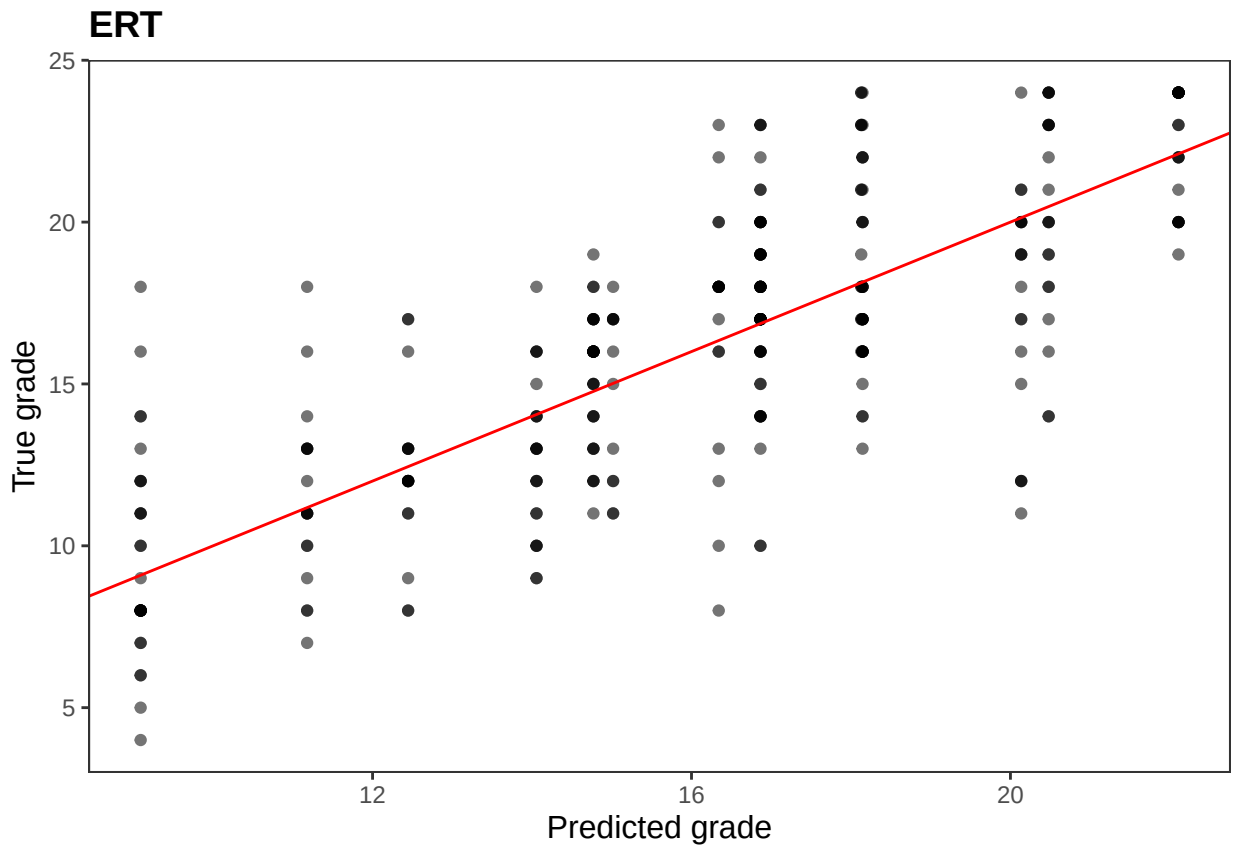

```
#ggsave("new_pred_plot/DS7_true_vs_preds_RF.png")
```

```
DS7 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS7_true_vs_preds_RT.png")
```

```
DS7 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS7_true_vs_preds_ERT.png")
```

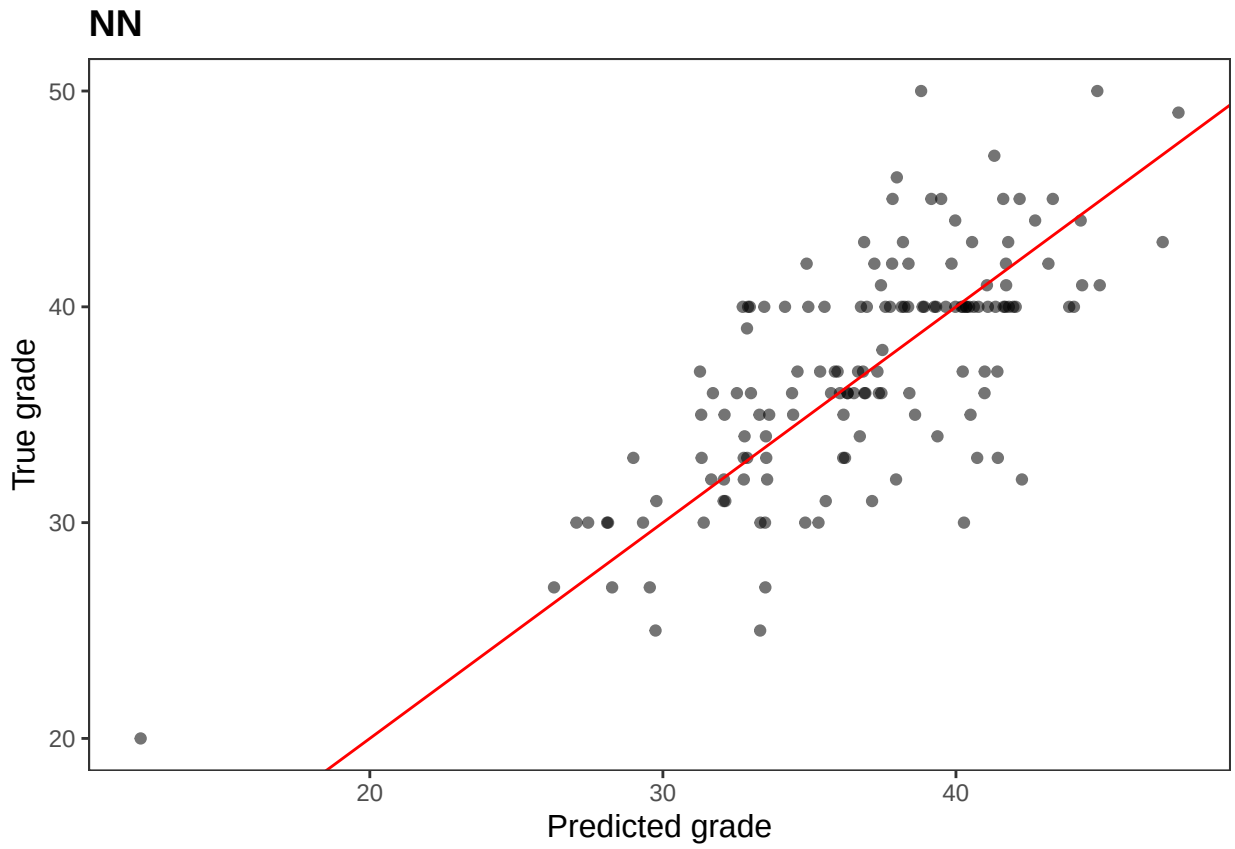
DS8

```
DS8 <- DS8_preds %>% select(y_test, ends_with("preds")) %>%
  rename(RF = "RF preds", RT = "DT preds", RF = "RF preds", NN = "NN preds",
         Lasso = "Lasso preds", Ridge = "Ridge preds", ERT = "ERT preds", LR = "LR preds") %>%
  melt(id.vars = 'y_test', variable.name = "model")
head(DS8)
```

```
##   y_test model    value
## 1     40    RF 38.21429
## 2     45    RF 39.08772
## 3     35    RF 34.61779
## 4     30    RF 39.08271
## 5     32    RF 39.62155
## 6     30    RF 34.96366
```

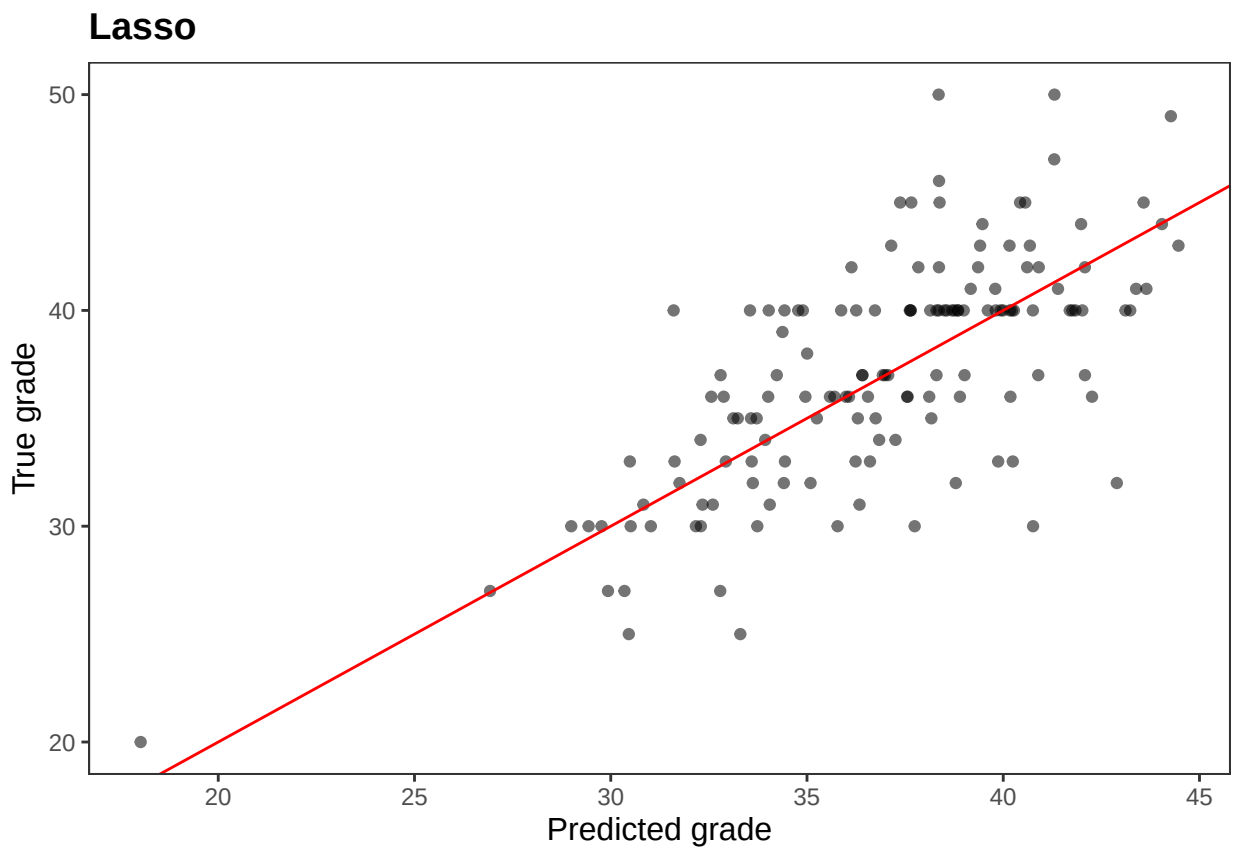
```
DS8 %>% filter(model %in% c("NN")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("NN") +
```

```
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



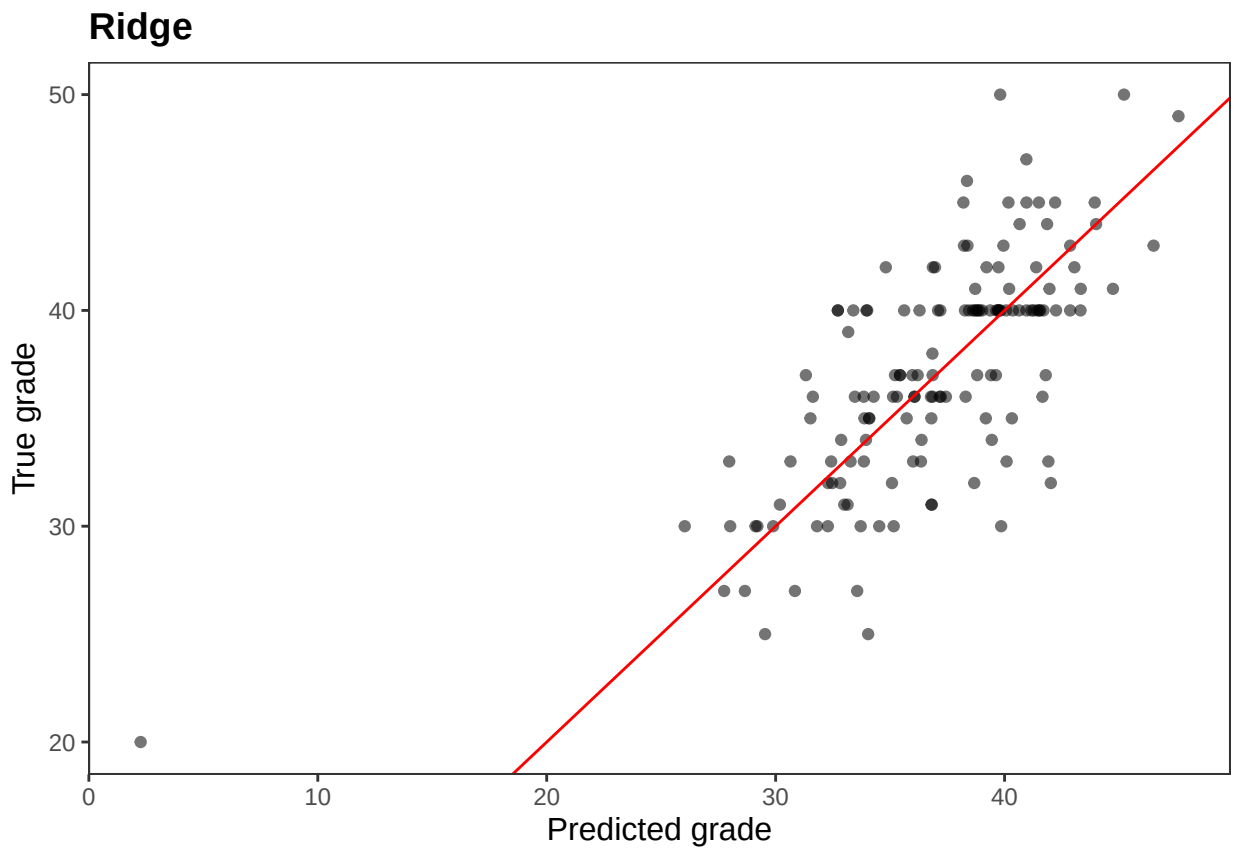

```
#ggsave("new_pred_plot/DS8_true_vs_preds_NN.png")
```

```
DS8 %>% filter(model %in% c("Lasso")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Lasso") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



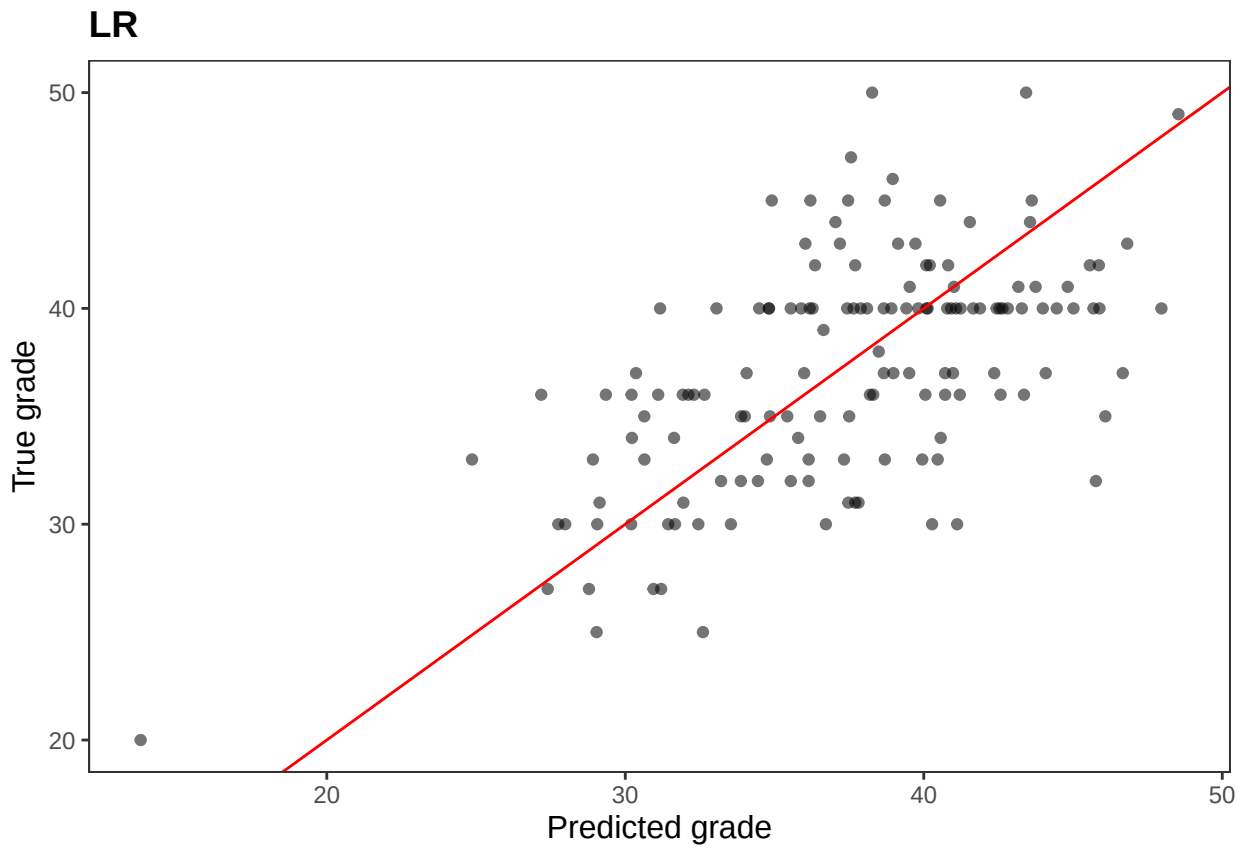

```
#ggsave("new_pred_plot/DS8_true_vs_preds_Lasso.png")
```

```
DS8 %>% filter(model %in% c("Ridge")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("Ridge") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



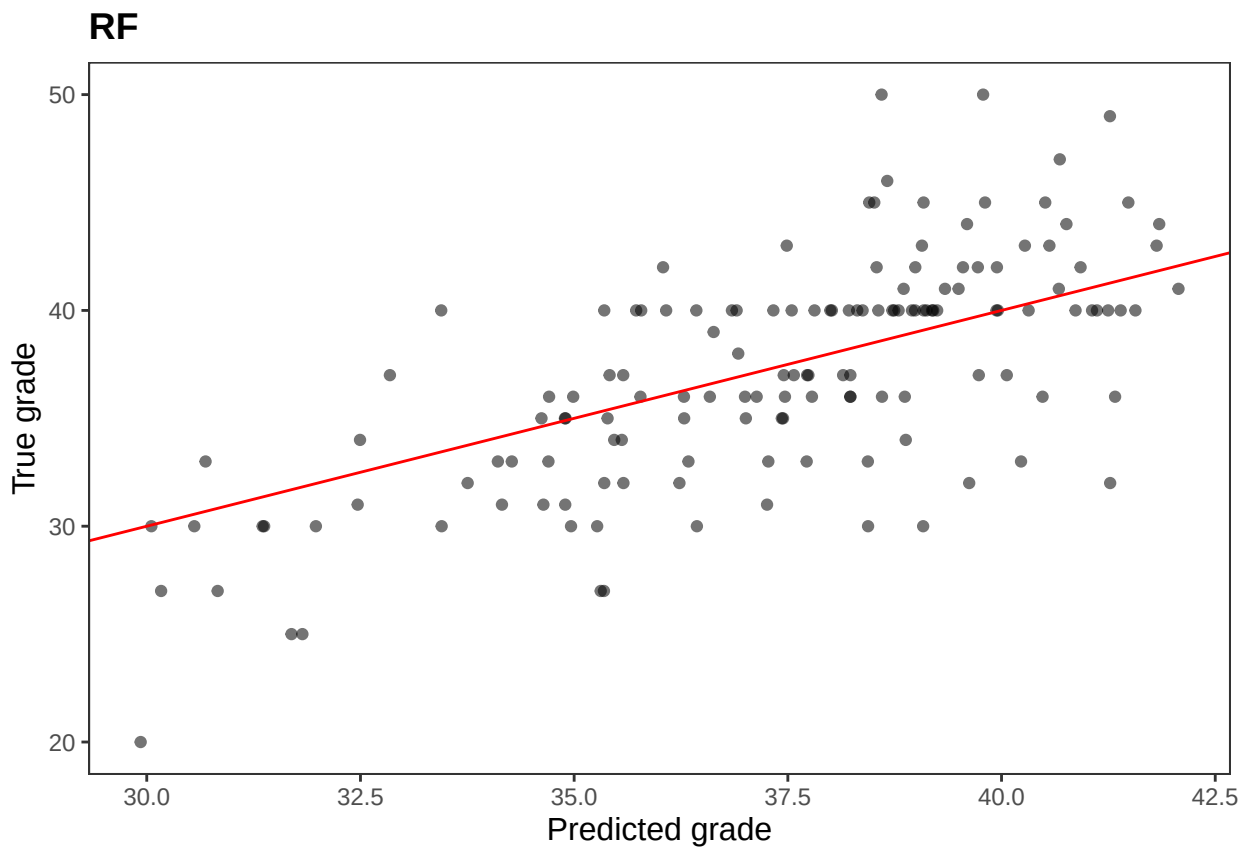

```
#ggsave("new_pred_plot/DS8_true_vs_preds_Ridge.png")
```

```
DS8 %>% filter(model %in% c("LR")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("LR") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



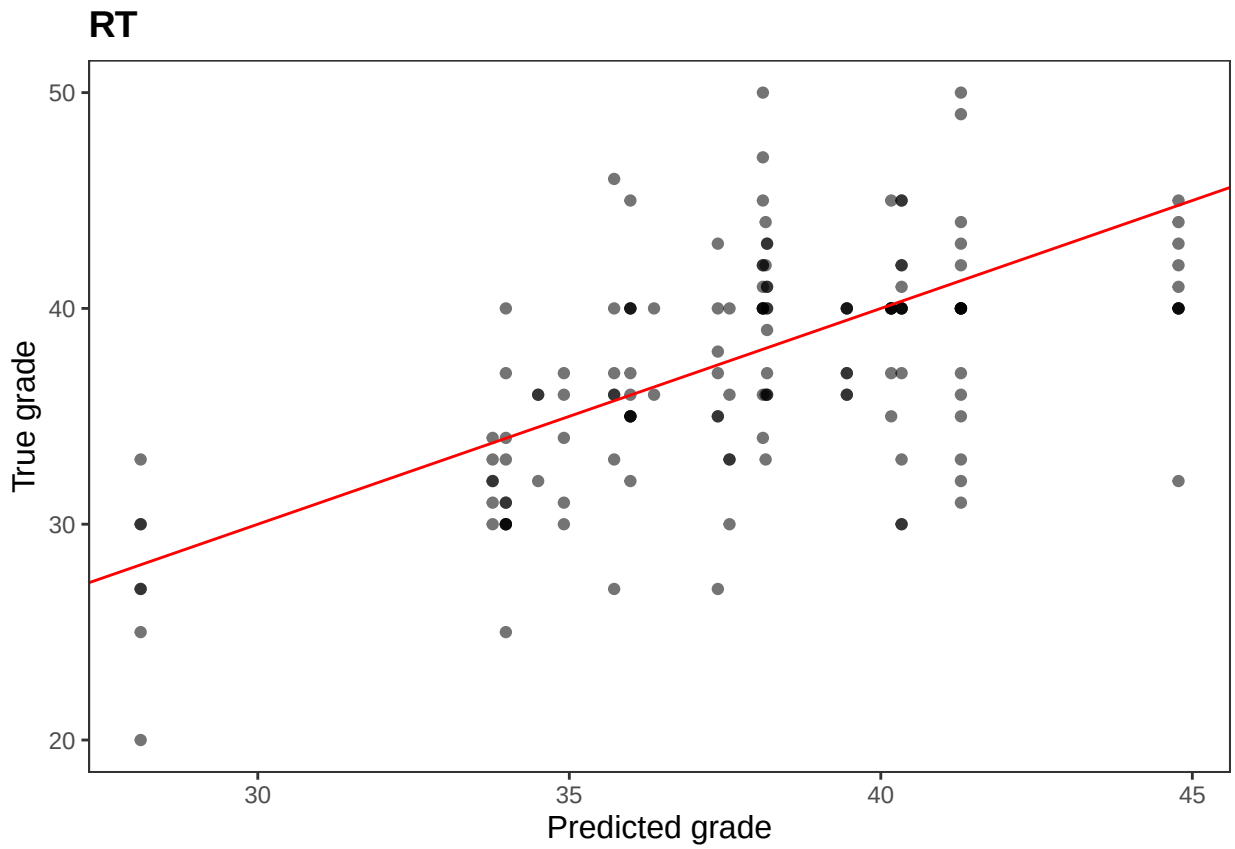

```
#ggsave("new_pred_plot/DS8_true_vs_preds_LR.png")
```

```
DS8 %>% filter(model %in% c("RF")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RF") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```



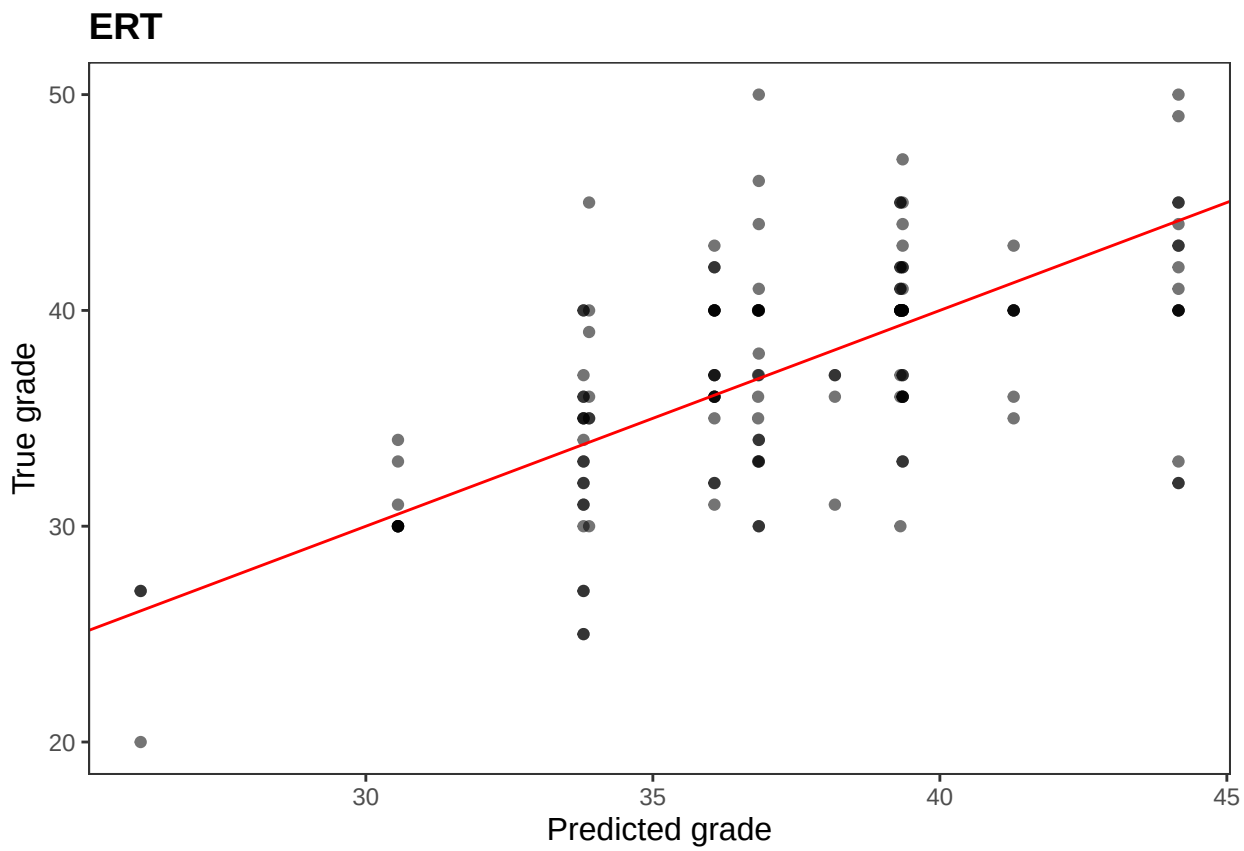

```
#ggsave("new_pred_plot/DS8_true_vs_preds_RF.png")
```

```
DS8 %>% filter(model %in% c("RT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("RT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS8_true_vs_preds_RT.png")
```

```
DS8 %>% filter(model %in% c("ERT")) %>%
  ggplot(aes(value, y_test, alpha = .5)) +
  geom_point() +
  geom_abline(col = "red") +
  ggtitle("ERT") +
  ylab("True grade") +
  xlab("Predicted grade") +
  theme_apa() +
  theme(legend.position = "none")
```




```
#ggsave("new_pred_plot/DS8_true_vs_preds_ERT.png")
```