

Analyze shortcut effects

Zining Zhu

12/17/2021

Load data

(Loading the bigram_shift as example. Will repeat for other SentEval tasks later.)

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(lmerTest)
```

```
## Warning: package 'lmerTest' was built under R version 4.1.2
```

```
##  
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':  
##  
##      lmer
```

```
## The following object is masked from 'package:stats':  
##  
##      step
```

```
df <- rbind(  
  read.csv("../reports/report_bigram_shift.roberta.csv"),  
  read.csv("../reports/report_coordination_inversion.roberta.csv"),  
  read.csv("../reports/report_obj_number.roberta.csv"),  
  read.csv("../reports/report_odd_man_out.roberta.csv"),  
  read.csv("../reports/report_past_present.roberta.csv"),  
  read.csv("../reports/report_subj_number.roberta.csv"))  
head(df)
```

```
##   train_acc train_loss val_acc  val_loss test_acc test_loss      model rs
## 1      0.5  0.6931472    0.5 0.6931472    0.5 0.6931472    LogReg  0
## 2      0.5  0.6940289    0.5 0.6940289    0.5 0.6940289    MLP-10  0
## 3      0.5  0.6931588    0.5 0.6931588    0.5 0.6931588    MLP-20  0
## 4      0.5  0.6931473    0.5 0.6931473    0.5 0.6931473    RF-100  0
## 5      0.5  0.6932686    0.5 0.6932686    0.5 0.6932686    RF-10   0
## 6      0.5  0.6931472    0.5 0.6931472    0.5 0.6931472 DecisionTree 0
##   config train_size_per_class      task nclasses
## 1   Full                1200 bigram_shift.roberta_layer_0      2
## 2   Full                1200 bigram_shift.roberta_layer_0      2
## 3   Full                1200 bigram_shift.roberta_layer_0      2
## 4   Full                1200 bigram_shift.roberta_layer_0      2
## 5   Full                1200 bigram_shift.roberta_layer_0      2
## 6   Full                1200 bigram_shift.roberta_layer_0      2
```

```
df_fvz = df[(df$config=='Full') | (df$config=='ZeroMI'),]
df_nvz = df[(df$config=='Nonzero') | (df$config=='ZeroMI'), ]
df_fvn = df[(df$config=='Full') | (df$config=='Nonzero'),]
```

Linear mixture model.

```
model_fvz <- lm(test_acc ~ task + model + config, data=df_fvz)
anova(model_fvz)
```

```
## Analysis of Variance Table
##
## Response: test_acc
##           Df Sum Sq Mean Sq F value    Pr(>F)
## task       77 25.9882  0.33751   176.99 < 2.2e-16 ***
## model        6  2.8071  0.46785    245.34 < 2.2e-16 ***
## config       1  0.2996  0.29965    157.14 < 2.2e-16 ***
## Residuals 5375 10.2496  0.00191
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model_nvz <- lm(test_acc ~ task + model + config, data=df_nvz)
anova(model_nvz)
```

```
## Analysis of Variance Table
##
## Response: test_acc
##           Df Sum Sq Mean Sq F value    Pr(>F)
## task       77 26.1858  0.34008   183.24 < 2.2e-16 ***
## model        6  2.4941  0.41568    223.99 < 2.2e-16 ***
## config       1  0.3007  0.30071    162.03 < 2.2e-16 ***
## Residuals 5375  9.9752  0.00186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model_fvn <- lm(test_acc ~ task + model + config, data=df_fvn)
anova(model_fvn)
```

```
## Analysis of Variance Table
##
## Response: test_acc
##           Df Sum Sq Mean Sq  F value Pr(>F)
## task       77 30.8118  0.40015  230.6586 <2e-16 ***
## model        6   3.8915  0.64859  373.8626 <2e-16 ***
## config       1   0.0000  0.00000    0.0005  0.9814
## Residuals 5375   9.3247  0.00173
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In both Full vs ZeroMI (`fvz`) and Nonzero vs ZeroMI (`nvz`) settings, the configuration has significant effects on the test accuracy.

LMM with random effects

```
model_fvn_re <- lmer(test_acc ~ task + model + config + (1+config|rs), data=df_fvn)
```

```
## boundary (singular) fit: see ?isSingular
```

```
anova(model_fvn_re)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF  F value Pr(>F)
## task    30.8118  0.40015     77   5375  230.6586 <2e-16 ***
## model     3.8915  0.64859      6   5375  373.8626 <2e-16 ***
## config    0.0000  0.00000      1   5375    0.0005  0.9814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

When considering the random effects of the seeds, the config still has no effects.