Models Comparison

Model	Architecture	Bleu score	Bleu score with smoothing
BERT	Dropout=0.3	0.572	0.00033
BERT	Dropout=0.7	0.573	0.00070
MobileBERT	Dropout=0.3	0.622	0.00050
MobileBERT	Dropout=0.7	0.619	0.00045
Roberta	Dropout=0.3	0.565	0.00025

- Rouge Scores of each model are in Notebook
- Smoothing is done using method 2 to avoid zero division error

Charts

Add panel

Training loss
— Roberta-0.3 ● — mobileBERT-0.7 — mobileBERT-0.3
— bert-0.7 — bert-0.3

0.8
0.6
0.4
0.2
0
Step
0   5   10   15

train_loss
Roberta-0.3
bert-0.7
bert-0.3
0.0   0.5   1.0   1.5   2.0

eval_loss
Roberta-0.3
bert-0.7
bert-0.3
-6   -4   -2   0

incorrect
Roberta-0.3
mobileBERT-0.7
mobileBERT-0.3
bert-0.7
bert-0.3
0   50   100   150   200

lr
— Roberta-0.3 ● — mobileBERT-0.7 — mobileBERT-0.3
— bert-0.7 — bert-0.3

4e-5
3e-5
2e-5
1e-5
0
Step
0   5   10   15

similar
Roberta-0.3
mobileBERT-0.7
mobileBERT-0.3
bert-0.7
bert-0.3
0   1,000   2,000   3,000   4,000

correct
Roberta-0.3
mobileBERT-0.7
mobileBERT-0.3
bert-0.7
bert-0.3
0   200   400   600

global_step
— Roberta-0.3 ● — mobileBERT-0.7 — mobileBERT-0.3
— bert-0.7 — bert-0.3

1000
800
600
400
200
Step
0   5   10   15   20