

gradients/bert/encoder/layer_0

output
20.46%

attention
50.80%

intermediate
28.75%

```
graph TD; A["output<br/>20.46%"] --> B["intermediate<br/>28.75%"]; C["attention<br/>50.80%"]
```