

gradients/bert/encoder/layer\_0

output  
17.78%

attention  
63.47%

intermediate  
18.75%

```
graph TD; A["output<br/>17.78%"] --> B["intermediate<br/>18.75%"]; C["attention<br/>63.47%"]
```