

gradients/bert/encoder/layer\_0

output  
18.03%

attention  
62.85%

intermediate  
19.12%

```
graph TD; A["output<br/>18.03%"] --> B["intermediate<br/>19.12%"]; C["attention<br/>62.85%"]
```