

gradients/bert/encoder/layer_5

output
17.90%

attention
62.61%

intermediate
19.48%

```
graph TD; A["output<br/>17.90%"] --> B["intermediate<br/>19.48%"]; C["attention<br/>62.61%"]
```