

gradients/bert/encoder/layer_1

output
17.97%

attention
62.84%

intermediate
19.19%

```
graph TD; A["output<br/>17.97%"] --> B["intermediate<br/>19.19%"]; C["attention<br/>62.84%"]
```