

gradients/bert/encoder/layer_1

output
21.98%

attention
42.87%

intermediate
35.15%

```
graph TD; A["output<br/>21.98%"] --> B["intermediate<br/>35.15%"]; C["attention<br/>42.87%"]
```