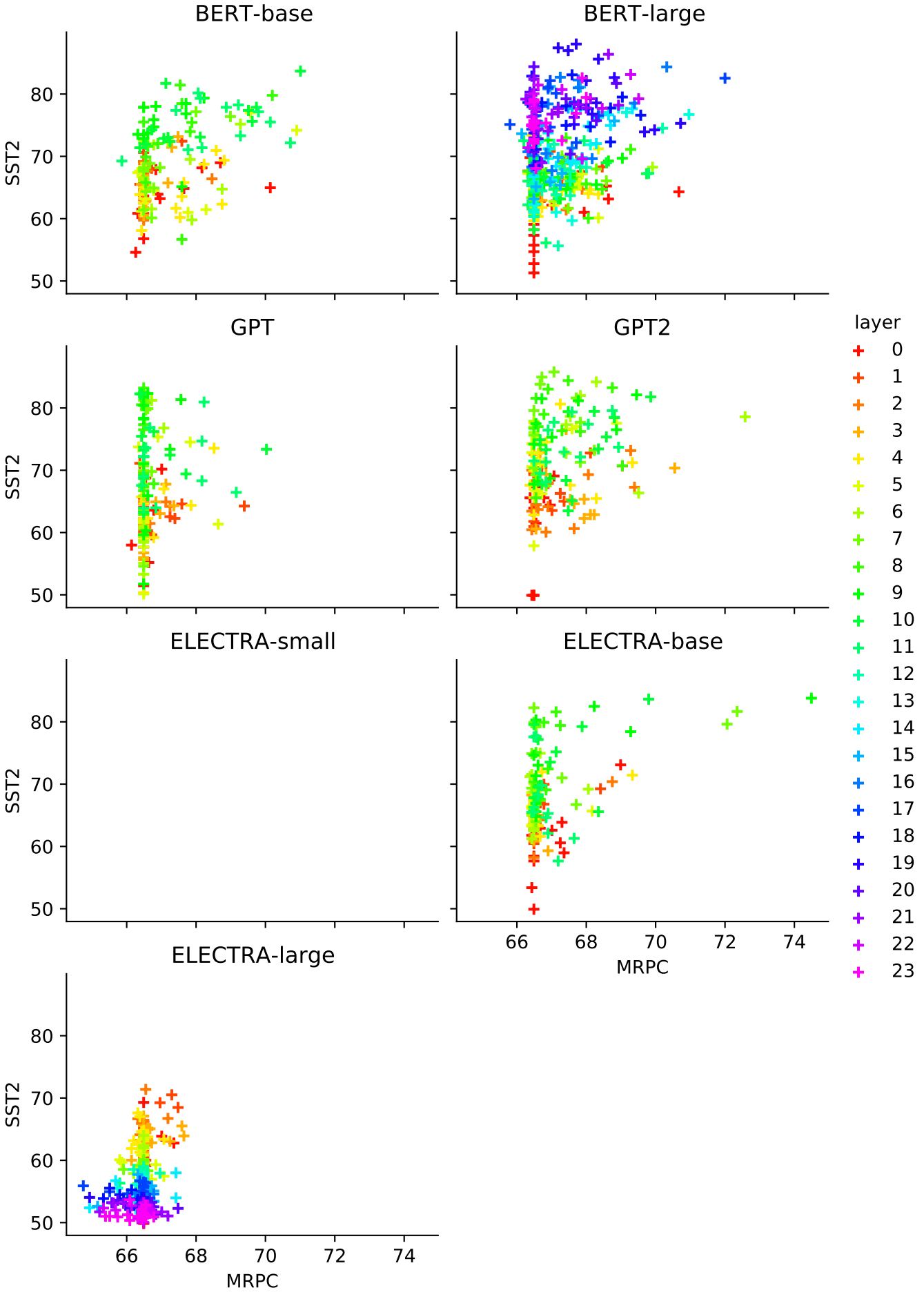
BERT-base
BERT-large
GPT
GPT2
ELECTRA-small
ELECTRA-base
ELECTRA-large
SST2
MRPC
50
60
70
80
66
68
70
72
74
layer
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23