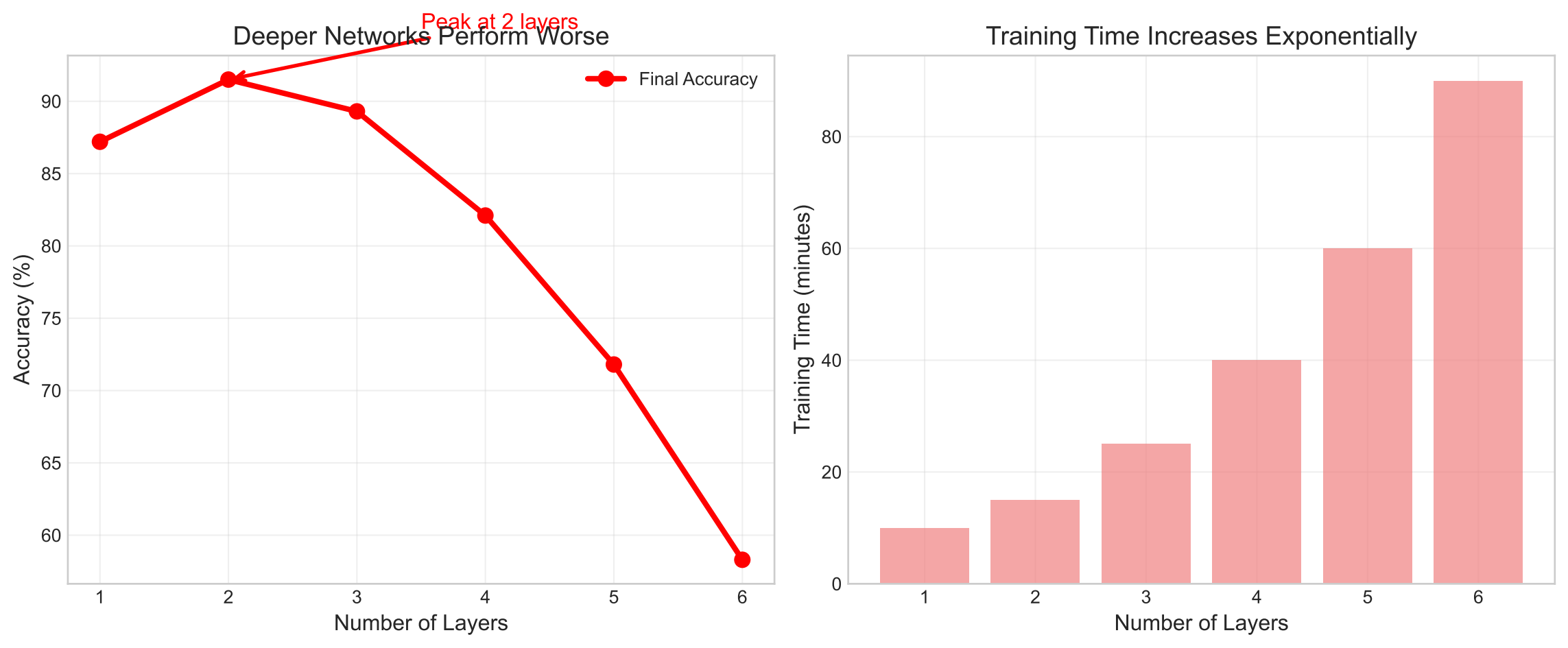
Deeper Networks Perform Worse
Peak at 2 layers
Final Accuracy
Accuracy (%)
90
85
80
75
70
65
60
1
2
3
4
5
6
Number of Layers
Training Time Increases Exponentially
Training Time (minutes)
80
60
40
20
0
1
2
3
4
5
6
Number of Layers