



- Different Architecture
- Different Context Length
- Different Optimizer
- Different Size