Scaling up the final Approach

- full d2v2 and bert
- 2 shared layers
- lower lr for d2v2 and bert
- 15 epochs