Fine Tuning Process

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **lcf model** | model 1 | model 2 | model 3 | model 4 | model 5 | model 6 | model 7 |
| batch\_size | 16 | 16 | 16 | 16 | 20 | 20 | 20 |
| lr | 5e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 |
| num\_epoch | 2 | 2 |  |  | 2 | 3 | 3 |
| valset\_ratio | 0.29 | 0.29 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| local context focus | cdw | cdw | cdm | cdw | cdw | cdw | cdw |
| SRD | 3 | 3 |  | 64 | 64 | 128 | 256 |
| optimizer | adam | adam |  |  |  |  |  |
| max seq len |  |  |  |  |  |  | 256 |
| test accuracy | 0.6105 | 0.6514 | 0.6361 | 0.6548 | 0.6497 | 0.6684 | 0.6514 |
| test f1 | 0.5816 | 0.6347 | 0.5860 | 0.6414 | 0.6266 | 0.6554 | 0.6382 |
| validation accuracy | 0.6472 | 0.66 | 0.6021 | 0.6515 | 0.6495 | 0.6330 | 0.6495 |
| validation f1 | 0.6240 | 0.6429 | 0.5959 | 0.6377 | 0.6217 | 0.6289 | 0.6257 |

|  |  |
| --- | --- |
| lcf model | model 8 |
| batch\_size | 23 |
| lr | 2e-5 |
| num\_epoch | 2 |
| valset\_ratio | 0.1 |
| local context focus | cdw |
| SRD | 3 |
| optimizer | adam |
| max seq len | 256 |
| drop out | 0 |
| test accuracy | 0.6837 |
| test f1 | 0.6817 |
| validation accuracy | 0.6198 |
| validation f1 | 0.6157 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **bert\_spc** |  |  |  |  |  |  |  |
| batch\_size | 16 | 32 | 64 | 64 | 32 | 16 |  |
| lr | 5e-5 | 5e-5 | 5e-5 | 2e-5 | 2e-5 | 2e-5 |  |
| num\_epoch | 2 | 2 | 2 | 2 | 2 | 2 |  |
| valset\_ratio | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 | 0.29 |  |
| local context focus |  |  |  |  |  |  |  |
| SRD | 3 | 3 | 3 | 3 | 3 | 3 |  |
| optimizer | adam | adam | adam | adam | adam | adam |  |
| test accuracy | 0.5289 | 0.5374 | 0.6054 | 0.5459 | 0.5782 | 0.5714 |  |
| test f1 | 0.3659 | 0.4798 | 0.5516 | 0.4135 | 0.5380 | 0.5332 |  |
| validation accuracy | 0.5149 | 0.5263 | 0.5676 | 0.5277 | 0.5448 | 0.5633 |  |
| validation f1 | 0.3501 | 0.4632 | 0.5113 | 0.3875 | 0.4996 | 0.5187 |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| bert\_spc |  |  |  |  |  |  |  |
| batch\_size | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| lr | 2e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 |
| num\_epoch | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| valset\_ratio | 0.2 | 0.1 | 0.01 | 0.05 | 0.8 | 0.11 | 0.09 |
| local context focus |  |  |  |  |  |  |  |
| SRD | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| optimizer | adam | adam | adam | adam | adam | adam | adam |
| test accuracy | 0.6327 | 0.6412 | 0.5017 | 0.5289 | 0.5442 | 0.6514 | 0.6480 |
| test f1 | 0.5947 | 0.6180 | 0.4686 | 0.3659 | 0.4503 | 0.6354 | 0.6165 |
| validation accuracy | 0.6041 | 0.6322 | 0.5000 | 0.5455 | 0.5278 | 0.6504 | 0.6193 |
| validation f1 | 0.5563 | 0.5916 | 0.4713 | 0.3909 | 0.4328 | 0.6252 | 0.5752 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| bert\_spc |  |  |  |  |  |  |  |
| batch\_size | 16 | 16 | 64 | 64 | 64 | 64 | 64 |
| lr | 2e-5 | 2e-5 | 2e-5 | 5e-5 | 5e-5 | 5e-5 | 5e-5 |
| num\_epoch | 3 | 5 | 5 | 5 | 2 | 10 | 6 |
| valset\_ratio | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 |
| local context focus | cdw | cdw | cdw | cdw | cdw | cdw | cdw |
| SRD | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| optimizer | adam | adam | adam | adam | adam | adam | adam |
| test accuracy | 0.6514 | 0.6514 | 0.6463 | 0.6548 | 0.5901 | 0.6514 | 0.6514 |
| test f1 | 0.6354 | 0.6354 | 0.6474 | 0.6544 | 0.5315 | 0.6477 | 0.6477 |
| validation accuracy | 0.6165 | 0.5639 | 0.6278 | 0.6391 | 0.5564 | 0.6241 | 0.6466 |
| validation f1 | 0.6167 | 0.5696 | 0.6213 | 0.6345 | 0.4827 | 0.6197 | 0.6403 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| bert\_spc |  |  |  |  |  |  |  |
| batch\_size | 64 | 64 | 64 |  |  |  |  |
| lr | 5e-5 | 5e-5 | 5e-5 |  |  |  |  |
| num\_epoch | 4 | 5 | 5 |  |  |  |  |
| valset\_ratio | 0.11 | 0.11 | 0.11 |  |  |  |  |
| local context focus | cdw | cdw | cdw |  |  |  |  |
| SRD | 3 | 3 | 3 |  |  |  |  |
| optimizer | adam | adam | adam |  |  |  |  |
| test accuracy | 0.6463 | 0.6548 |  |  |  |  |  |
| test f1 | 0.6413 | 0.6544 |  |  |  |  |  |
| validation accuracy | 0.6128 | 0.6391 |  |  |  |  |  |
| validation f1 | 0.5989 | 0.6345 |  |  |  |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| lcf model | model 1 | model 2 |  |  |  |  |  |
| batch\_size | 16 | 16 | 16 | 16 | 20 |  |  |
| lr | 5e-5 | 2e-5 | 2e-5 | 2e-5 | 2e-5 |  |  |
| num\_epoch | 2 | 2 |  |  | 2 | 3 |  |
| valset\_ratio | 0.29 | 0.29 | 0.2 | 0.2 | 0.2 |  |  |
| local context focus | cdw | cdw | cdm | cdw | cdw |  |  |
| SRD | 3 | 3 |  | 64 | 64 | 128 | 256 |
| optimizer | adam | adam |  |  |  |  |  |
| test accuracy | 0.6105 | 0.6514 | 0.6361 | 0.6548 | 0.6497 | 0.6684 | 0.6514 |
| test f1 | 0.5816 | 0.6347 | 0.5860 | 0.6414 | 0.6266 | 0.6554 | 0.6382 |
| validation accuracy | 0.6472 | 0.66 | 0.6021 | 0.6515 | 0.6495 | 0.6330 | 0.6495 |
| validation f1 | 0.6240 | 0.6429 | 0.5959 | 0.6377 | 0.6217 | 0.6289 | 0.6257 |