Comparison for different learning rates:

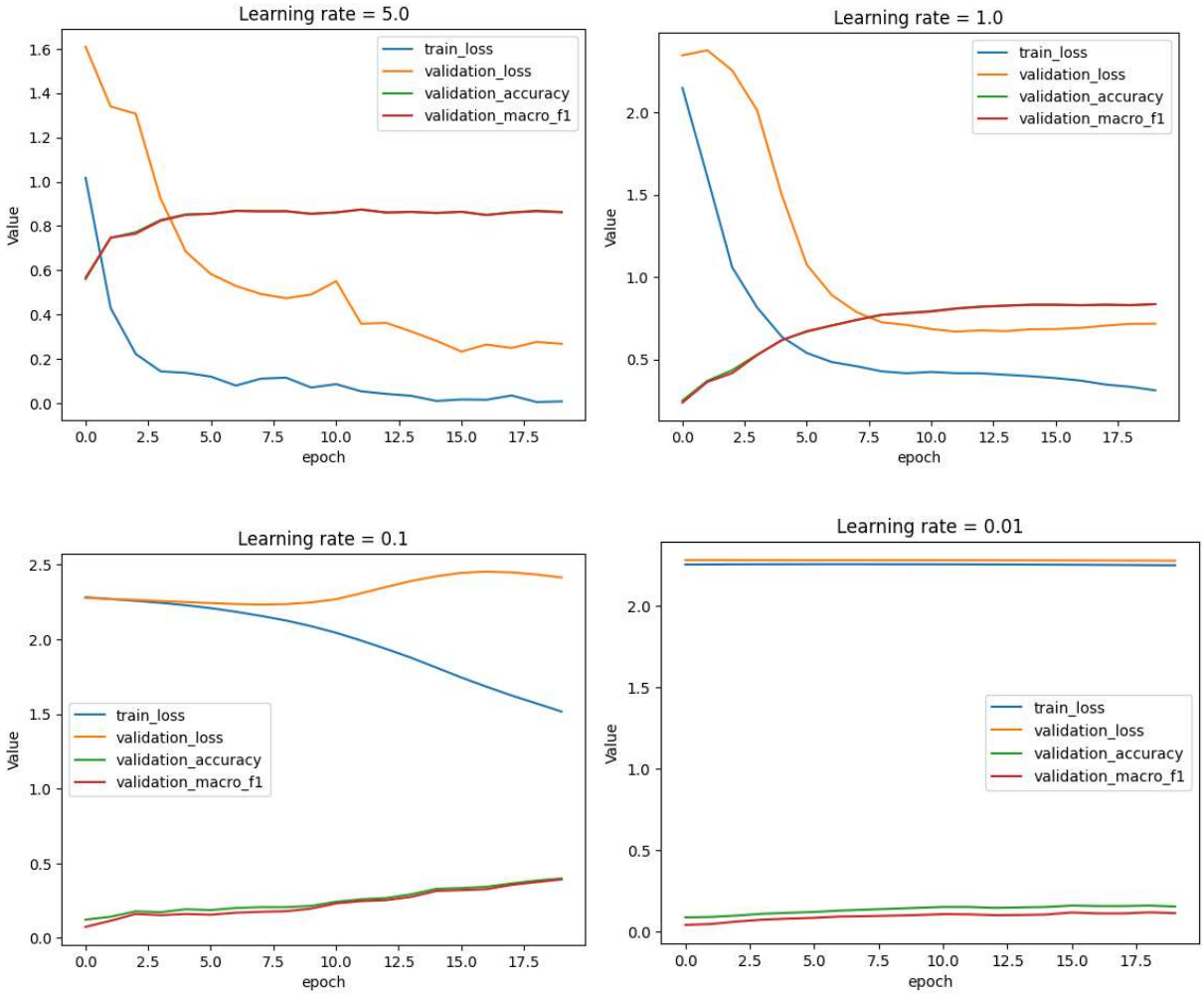


Fig 1- Train Loss, Validation Loss , Validation Accuracy and Validation Macro F1 for different learning rates.

For the best model with learning rate 5 , all the scores are shown in a table:

| epoch | train_loss | train_accuracy | train_macro_f1 | validation_loss | validation_accuracy | validation_macro_f1 |
|---|---|---|---|---|---|---|
| 0 | 1.017387 | 0.580626 | 0.554372 | 1.609324 | 0.559889 | 0.566688 |
| 1 | 0.429251 | 0.819885 | 0.818342 | 1.339466 | 0.746518 | 0.747152 |
| 2 | 0.221926 | 0.876405 | 0.875786 | 1.307097 | 0.771588 | 0.764882 |
| 3 | 0.143016 | 0.905208 | 0.904503 | 0.92134 | 0.827298 | 0.823851 |
| 4 | 0.136583 | 0.91789 | 0.917373 | 0.683975 | 0.852368 | 0.850043 |
| 5 | 0.11938 | 0.928633 | 0.928222 | 0.582714 | 0.855153 | 0.854788 |
| 6 | 0.078883 | 0.939716 | 0.939508 | 0.528911 | 0.869081 | 0.867481 |
| 7 | 0.110214 | 0.941932 | 0.941707 | 0.492372 | 0.866295 | 0.86605 |
| 8 | 0.114593 | 0.945953 | 0.945921 | 0.473502 | 0.866295 | 0.866683 |
| 9 | 0.070301 | 0.953218 | 0.953168 | 0.489912 | 0.855153 | 0.8547 |
| 10 | 0.08536 | 0.955068 | 0.955051 | 0.550777 | 0.860724 | 0.861312 |
| 11 | 0.053038 | 0.958585 | 0.958536 | 0.358125 | 0.874652 | 0.873996 |
| 12 | 0.041337 | 0.960111 | 0.960035 | 0.362197 | 0.860724 | 0.860302 |
| 13 | 0.0327 | 0.964965 | 0.964877 | 0.323074 | 0.86351 | 0.863637 |
| 14 | 0.009494 | 0.966204 | 0.966119 | 0.281346 | 0.857939 | 0.858775 |
| 15 | 0.016172 | 0.968199 | 0.968124 | 0.232639 | 0.86351 | 0.864225 |
| 16 | 0.014781 | 0.968578 | 0.968544 | 0.264132 | 0.849582 | 0.849509 |
| 17 | 0.03414 | 0.97169 | 0.971689 | 0.248655 | 0.860724 | 0.860586 |
| 18 | 0.004621 | 0.971662 | 0.971603 | 0.276118 | 0.869081 | 0.866089 |
| 19 | 0.007399 | 0.971551 | 0.971509 | 0.267808 | 0.86351 | 0.861408 |

Confusion matrix:  Lr = 5 ie the best model on test set

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1068. | 5. | 1. | 8. | 5. | 17. | 1. | 3. | 0. | 1. |
| 1 | 3. | 977. | 14. | 3. | 17. | 4. | 7. | 0. | 2. | 102. |
| 2 | 0. | 6. | 1023. | 1. | 5. | 8. | 1. | 1. | 2. | 1. |
| 3 | 11. | 7. | 3. | 1054. | 12. | 24. | 43. | 0. | 6. | 0. |
| 4 | 7. | 5. | 13. | 2. | 892. | 12. | 2. | 3. | 2. | 5. |
| 5 | 6. | 4. | 9. | 6. | 97. | 958. | 14. | 1. | 1. | 5. |
| 6 | 0. | 4. | 4. | 25. | 4. | 41. | 984. | 0. | 3. | 0. |
| 7 | 6. | 7. | 12. | 2. | 23. | 24. | 6. | 1061. | 0. | 25. |
| 8 | 2. | 0. | 12. | 4. | 24. | 13. | 6. | 2. | 1068. | 1. |
| 9 | 4. | 92. | 16. | 2. | 28. | 6. | 4. | 4. | 2. | 897. |

Test set Accuracsy: 0.9151081774844151

Test set macro-f1: 0.9145186790501025