Q2(a)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | Hdim | Lookback | Learning rate | Best loss |
| 1 | 25 | 0 | 0.1 | 5.21897271803343 |
| 2 | 25 | 0 | 0.5 | 5.016975861837507 |
| 3 | 25 | 0 | 0.05 | 5.407677518765296 |
| 4 | 25 | 2 | 0.1 | 5.214505668687516 |
| 5 | 25 | 2 | 0.5 | 5.035141274765579 |
| 6 | 25 | 2 | 0.05 | 5.403824569896471 |
| 7 | 25 | 5 | 0.1 | 5.214640703207686 |
| 8 | 25 | 5 | 0.5 | 5.018008955421851 |
| 9 | 25 | 5 | 0.05 | 5.403781106364606 |
| 10 | 50 | 0 | 0.1 | 5.132931888175017 |
| 11 | 50 | 0 | 0.5 | 4.973139445402921 |
| 12 | 50 | 0 | 0.05 | 5.312957608309724 |
| 13 | 50 | 2 | 0.1 | 5.136560027192997 |
| 14 | 50 | 2 | 0.5 | 5.029222280349153 |
| 18 | 50 | 2 | 0.05 | 5.314183768404833 |
| 16 | 50 | 5 | 0.1 | 5.136824058303708 |
| 17 | 50 | 5 | 0.5 | 5.021098622112128 |
| 18 | 50 | 5 | 0.05 | 5.314223521752088 |

We choose number 11 with Hdim=50, lookback=0 and learning rate=0.5, for its best performance among our experiments.

Q(2)b

On dev set

Dev mean loss 4.415768044290772

On test set

Test mean loss 4.426598071706132

Unadjusted: 83.646

Adjusted for missing vocab: 112.925

Q3 (b)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Number | Hdim | Lookback | Learning rate | Best loss | Acc | At which epoch |
| 1 | 25 | 0 | 0.1 | 0.6951596223673758 | 0.659 | 10 |
| 2 | 25 | 0 | 0.5 | 0.6452609985233952 | 0.658 | 7 |
| 3 | 25 | 0 | 0.05 | 0.8281694991228595 | 0.659 | 10 |
| 4 | 25 | 2 | 0.1 | 0.692399833363384 | 0.659 | 10 |
| 5 | 25 | 2 | 0.5 | 0.6448560079592647 | 0.658 | 7 |
| 6 | 25 | 2 | 0.05 | 0.8142654185920801 | 0.659 | 10 |
| 7 | 25 | 5 | 0.1 | 0.6923229610881881 | 0.659 | 10 |
| 8 | 25 | 5 | 0.5 | 0.6449029214637166 | 0.658 | 7 |
| 9 | 25 | 5 | 0.05 | 0.8137990304504382 | 0.659 | 10 |
| 10 | 50 | 0 | 0.1 | 0.6789812774832906 | 0.669 | 10 |
| 11 | 50 | 0 | 0.5 | 0.6447408478359821 | 0.669 | 7 |
| 12 | 50 | 0 | 0.05 | 0.7193314891305294 | 0.666 | 10 |
| 13 | 50 | 2 | 0.1 | 0.67800530242373 | 0.669 | 10 |
| 14 | 50 | 2 | 0.5 | 0.6414452303429862 | 0.669 | 7 |
| 18 | 50 | 2 | 0.05 | 0.7169518959050305 | 0.666 | 10 |
| 16 | 50 | 5 | 0.1 | 0.6778664329549344 | 0.669 | 10 |
| 17 | 50 | 5 | 0.5 | 0.6412843774790113 | 0.669 | 7 |
| 18 | 50 | 5 | 0.05 | 0.7167688771994244 | 0.666 | 10 |

Q 4

Number prediction accuracy on dev set: 0.675  
Number prediction accuracy on test set: 0.6505