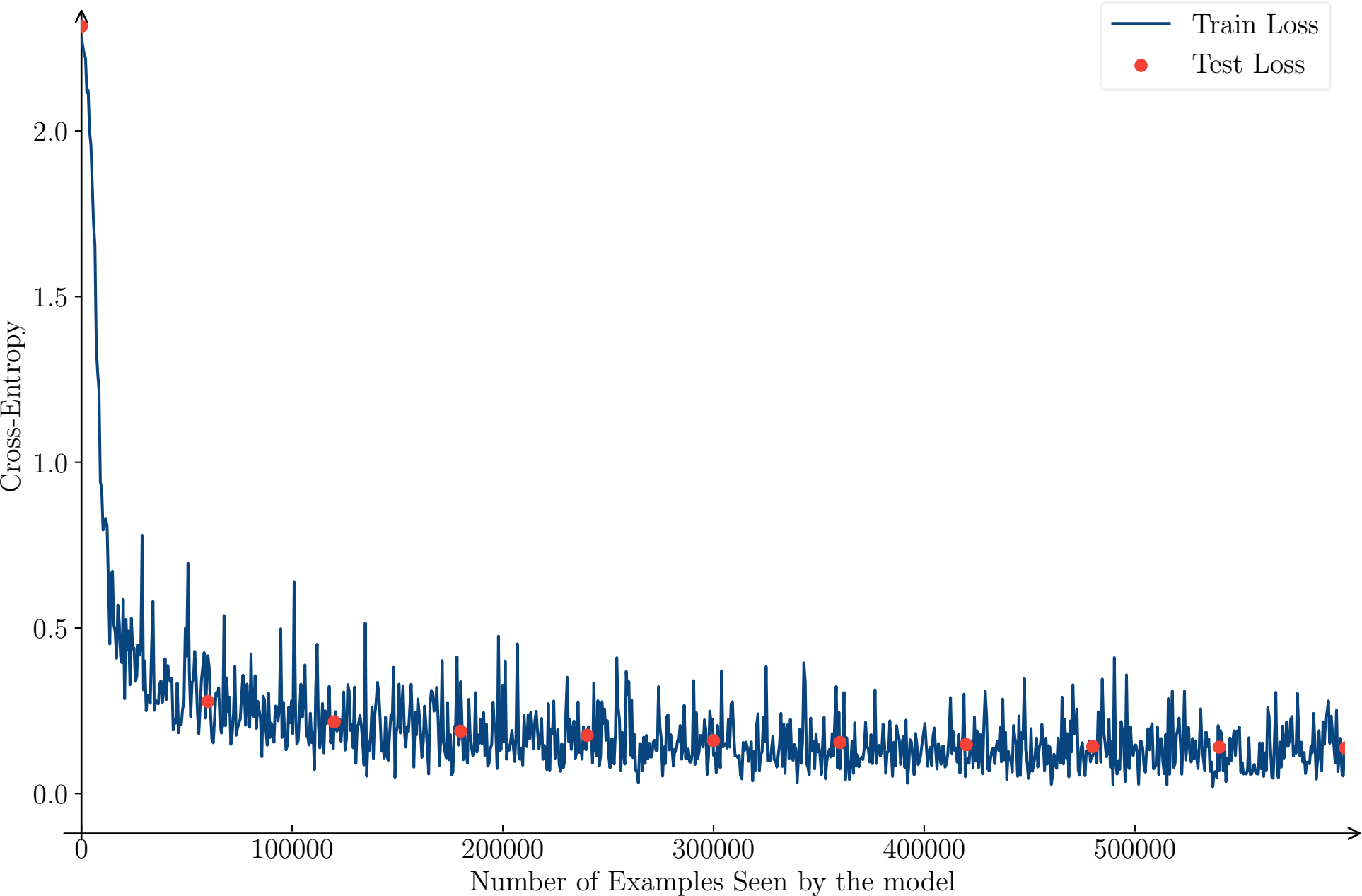
Train Loss
Test Loss
Cross-Entropy
2.0
1.5
1.0
0.5
0.0
0
100000
200000
300000
400000
500000
Number of Examples Seen by the model