Andrew Nicola 801136465 HW3

Problem 1)

Inspired by the course example, train and validate rnn.RNN, rnn.LSTM and rnn.GRU for learning the above sequence. Use sequence lengths of 10, 20, and 30 for your training. Feel free to adjust other network parameters. Report and compare training loss, validation accuracy, execution time for training, and computational and mode size complexities across the three models over various lengths of sequence.

RNN sequence of 10:

```
Epoch 10, Loss: 2.2318365573883057, Validation Loss: 2.293134927749634, Validation Accuracy: 0.3886554539203644
Epoch 20, Loss: 1.7558807134628296, Validation Loss: 2.0356361865997314, Validation Accuracy: 0.44117647409439087
Epoch 30, Loss: 1.4005903005599976, Validation Loss: 1.90595543384552, Validation Accuracy: 0.48949578404426575
Epoch 40, Loss: 1.0829108953475952, Validation Loss: 1.8602923154830933, Validation Accuracy: 0.49369746446609497
Epoch 50, Loss: 0.7962321639060974, Validation Loss: 1.8867394924163818, Validation Accuracy: 0.5252100825309753
Epoch 60, Loss: 0.556976318359375, Validation Loss: 1.9297168254852295, Validation Accuracy: 0.5273109078407288
Epoch 70, Loss: 0.3610989451408386, Validation Loss: 2.0324878692626953, Validation Accuracy: 0.5231092572212219
Epoch 80, Loss: 0.2205822616815567, Validation Loss: 2.1859323978424072, Validation Accuracy: 0.5315126180648804
Epoch 90, Loss: 0.13359877467155457, Validation Loss: 2.3336892127990723, Validation Accuracy: 0.5105041861534119
Epoch 100, Loss: 0.09143994003534317, Validation Loss: 2.465543746948242, Validation Accuracy: 0.5210084319114685
Total execution time for training: 1.5855510234832764 seconds
```

RNN sequence of 20:

```
Epoch 10, Loss: 2.2617180347442627, Validation Loss: 2.3123133182525635, Validation Accuracy: 0.40084388852119446
Epoch 20, Loss: 1.8056467771530151, Validation Loss: 2.075916290283203, Validation Accuracy: 0.45780590176582336
Epoch 30, Loss: 1.456538438796997, Validation Loss: 1.953408122062683, Validation Accuracy: 0.4831223487854004
Epoch 40, Loss: 1.1540477275848389, Validation Loss: 1.9044122695922852, Validation Accuracy: 0.5105485320091248
Epoch 50, Loss: 0.8750730752944946, Validation Loss: 1.8882423639297485, Validation Accuracy: 0.5189873576164246
Epoch 60, Loss: 0.6367104649543762, Validation Loss: 1.969519019126892, Validation Accuracy: 0.5358649492263794
Epoch 70, Loss: 0.43332192301750183, Validation Loss: 2.079019546508789, Validation Accuracy: 0.5316455960273743
Epoch 80, Loss: 0.2846674919128418, Validation Loss: 2.216718912124634, Validation Accuracy: 0.5379746556282043
Epoch 90, Loss: 0.1785351186990738, Validation Loss: 2.376279592514038, Validation Accuracy: 0.5253164768218994
Epoch 100, Loss: 0.11596552282571793, Validation Loss: 2.5036566257476807, Validation Accuracy: 0.5189873576164246
Total execution time for training: 2.9763951301574707 seconds
```

RNN sequence of 30:

```
Epoch 10, Loss: 2.270785331726074, Validation Loss: 2.406099557876587, Validation Accuracy: 0.33474576473236084
Epoch 20, Loss: 1.7989249229431152, Validation Loss: 2.157938003540039, Validation Accuracy: 0.40466102957725525
Epoch 30, Loss: 1.4406944513320923, Validation Loss: 2.058704137802124, Validation Accuracy: 0.4682203531265259
Epoch 40, Loss: 1.124627947807312, Validation Loss: 2.046135425567627, Validation Accuracy: 0.4894067943096161
Epoch 50, Loss: 0.8385301828384399, Validation Loss: 2.0404891967773438, Validation Accuracy: 0.4957627058029175
Epoch 60, Loss: 0.60858154296875, Validation Loss: 2.113710641860962, Validation Accuracy: 0.508474588394165
Epoch 70, Loss: 0.4128665030002594, Validation Loss: 2.223323345184326, Validation Accuracy: 0.49152541160583496
Epoch 80, Loss: 0.26966893672943115, Validation Loss: 2.350189685821533, Validation Accuracy: 0.5021186470985413
Epoch 90, Loss: 0.16926011443138123, Validation Loss: 2.481628179550171, Validation Accuracy: 0.48728814721107483
Epoch 100, Loss: 0.11020784825086594, Validation Loss: 2.6183078289031982, Validation Accuracy: 0.4894067943096161
Total execution time for training: 4.459903955459595 seconds
```

The loss value and training accuracy were the best for a RNN sequence of 20.

LTSM sequence of 10:

Andrew Nicola 801136465 HW3

```
Epoch 10, Loss: 2.5319554805755615, Validation Loss: 2.4772162437438965, Validation Accuracy: 0.3025210201740265
Epoch 20, Loss: 2.010073184967041, Validation Loss: 2.1620285511016846, Validation Accuracy: 0.424369752407074
Epoch 30, Loss: 1.6024867296218872, Validation Loss: 1.999237298965454, Validation Accuracy: 0.462184876203537
Epoch 40, Loss: 1.2371381521224976, Validation Loss: 1.9405391216278076, Validation Accuracy: 0.4957983195781708
Epoch 50, Loss: 0.9121354222297668, Validation Loss: 1.950217366218567, Validation Accuracy: 0.5042017102241516
Epoch 60, Loss: 0.6232479214668274, Validation Loss: 2.0126564502716064, Validation Accuracy: 0.5042017102241516
Epoch 70, Loss: 0.3906773328781128, Validation Loss: 2.12318754196167, Validation Accuracy: 0.4957983195781708
Epoch 80, Loss: 0.23331937193870544, Validation Loss: 2.261284112930298, Validation Accuracy: 0.5021008253097534
Epoch 90, Loss: 0.14367955923080444, Validation Loss: 2.3871896266937256, Validation Accuracy: 0.48949578404426575
Epoch 100, Loss: 0.09821955114603043, Validation Loss: 2.501943588256836, Validation Accuracy: 0.49159663915634155
Total execution time for training: 3.554847478866577 seconds
```

LTSM sequence of 20:

```
Epoch 10, Loss: 2.566255569458008, Validation Loss: 2.523092031478882, Validation Accuracy: 0.3206751048564911
Epoch 20, Loss: 2.080193281173706, Validation Loss: 2.173811674118042, Validation Accuracy: 0.3987341821193695
Epoch 30, Loss: 1.6896708011627197, Validation Loss: 1.9862499237060547, Validation Accuracy: 0.4873417615890503
Epoch 40, Loss: 1.329180359840393, Validation Loss: 1.890060305595398, Validation Accuracy: 0.502109706401825
Epoch 50, Loss: 1.030712604522705, Validation Loss: 1.8771021366119385, Validation Accuracy: 0.5295358896255493
Epoch 60, Loss: 0.7546561360359192, Validation Loss: 1.909353256225586, Validation Accuracy: 0.5316455960273743
Epoch 70, Loss: 0.527492105960846, Validation Loss: 1.962944507598877, Validation Accuracy: 0.5189873576164246
Epoch 80, Loss: 0.37476521730422974, Validation Loss: 2.0272555351257324, Validation Accuracy: 0.5232067704200745
Epoch 90, Loss: 0.24019590020179749, Validation Loss: 2.146002769470215, Validation Accuracy: 0.5147679448127747
Epoch 100, Loss: 0.16719377040863037, Validation Loss: 2.2577321529388428, Validation Accuracy: 0.502109706401825
Total execution time for training: 6.711322546005249 seconds
```

LTSM sequence of 30:

```
Epoch 10, Loss: 2.5331737995147705, Validation Loss: 2.580303907394409, Validation Accuracy: 0.29025423526763916
Epoch 20, Loss: 2.050171375274658, Validation Loss: 2.2532997131347656, Validation Accuracy: 0.3792372941970825
Epoch 30, Loss: 1.6833680868148804, Validation Loss: 2.076892614364624, Validation Accuracy: 0.4173728823661804
Epoch 40, Loss: 1.345648169517517, Validation Loss: 1.9826481342315674, Validation Accuracy: 0.4661017060279846
Epoch 50, Loss: 1.0471264123916626, Validation Loss: 1.9437787532806396, Validation Accuracy: 0.4830508530139923
Epoch 60, Loss: 0.7629194855690002, Validation Loss: 1.9505035877227783, Validation Accuracy: 0.485169500112533357
Epoch 70, Loss: 0.5238338112831116, Validation Loss: 2.002650022506714, Validation Accuracy: 0.5105932354927063
Epoch 80, Loss: 0.35704299807548523, Validation Loss: 2.0903637409210205, Validation Accuracy: 0.508474588394165
Epoch 90, Loss: 0.22252334654331207, Validation Loss: 2.1816632747650146, Validation Accuracy: 0.4936440587043762
Epoch 100, Loss: 0.13908858597278595, Validation Loss: 2.2727761268615723, Validation Accuracy: 0.4957627058029175
Total execution time for training: 10.852977275848389 seconds
```

The sequence of 20 had the best accuracy, the validation loss was best for the sequence of 30.

GRU sequence of 10:

```
Epoch 10, Loss: 2.362579107284546, Validation Loss: 2.3552424907684326, Validation Accuracy: 0.38235294818878174
Epoch 20, Loss: 1.8574811220169067, Validation Loss: 2.045283317565918, Validation Accuracy: 0.4432772994041443
Epoch 30, Loss: 1.4390254020690918, Validation Loss: 1.9021484851837158, Validation Accuracy: 0.48949578404426575
Epoch 40, Loss: 1.0665242671966553, Validation Loss: 1.854416012763977, Validation Accuracy: 0.4957983195781708
Epoch 50, Loss: 0.7337533831596375, Validation Loss: 1.8989617824554443, Validation Accuracy: 0.5252100825309753
Epoch 60, Loss: 0.4612691104412079, Validation Loss: 1.996089220046997, Validation Accuracy: 0.5378151535987854
Epoch 70, Loss: 0.26790571212768555, Validation Loss: 2.1543712615966797, Validation Accuracy: 0.529411792755127
Epoch 80, Loss: 0.1528775990009308, Validation Loss: 2.3064486980438232, Validation Accuracy: 0.5168067216873169
Epoch 90, Loss: 0.0957653820514679, Validation Loss: 2.4412593841552734, Validation Accuracy: 0.5105041861534119
Epoch 100, Loss: 0.07029607892036438, Validation Loss: 2.5433695316314697, Validation Accuracy: 0.5231092572212219
Total execution time for training: 4.140954494476318 seconds
```

Andrew Nicola 801136465 HW3

GRU sequence of 20:

```
Epoch 10, Loss: 2.364863634109497, Validation Loss: 2.369717836380005, Validation Accuracy: 0.3164556920528412
Epoch 20, Loss: 1.8670744895935059, Validation Loss: 2.0857629776000977, Validation Accuracy: 0.4556961953639984
Epoch 30, Loss: 1.4557466506958008, Validation Loss: 1.9398282766342163, Validation Accuracy: 0.5063291192054749
Epoch 40, Loss: 1.0812913179397583, Validation Loss: 1.8598074913024902, Validation Accuracy: 0.5168776512145996
Epoch 50, Loss: 0.7500610947608948, Validation Loss: 1.8427746295928955, Validation Accuracy: 0.5569620132446289
Epoch 60, Loss: 0.478264719247818, Validation Loss: 1.9129105806350708, Validation Accuracy: 0.5632911324501038
Epoch 70, Loss: 0.28179556131362915, Validation Loss: 2.0088882446289062, Validation Accuracy: 0.5590717196464539
Epoch 80, Loss: 0.15955805778503418, Validation Loss: 2.141486644744873, Validation Accuracy: 0.5569620132446289
Epoch 90, Loss: 0.0911194235086441, Validation Loss: 2.2721376419067383, Validation Accuracy: 0.5485231876373291
Epoch 100, Loss: 0.05830315873026848, Validation Loss: 2.365845203399658, Validation Accuracy: 0.5295358896255493
Total execution time for training: 8.543624639511108 seconds
```

GRU sequence of 30:

```
Epoch 10, Loss: 2.377859592437744, Validation Loss: 2.493075132369995, Validation Accuracy: 0.2817796468734741
Epoch 20, Loss: 1.881993293762207, Validation Loss: 2.1907496452331543, Validation Accuracy: 0.40042373538017273
Epoch 30, Loss: 1.4610310792922974, Validation Loss: 2.002131938934326, Validation Accuracy: 0.4512711763381958
Epoch 40, Loss: 1.0819740295410156, Validation Loss: 1.9371222257614136, Validation Accuracy: 0.4957627058029175
Epoch 50, Loss: 0.7442370653152466, Validation Loss: 1.965140700340271, Validation Accuracy: 0.5254237055778503
Epoch 60, Loss: 0.4645291864871979, Validation Loss: 2.0598089694976807, Validation Accuracy: 0.5338982939720154
Epoch 70, Loss: 0.26403042674064636, Validation Loss: 2.2407562732696533, Validation Accuracy: 0.5254237055778503
Epoch 80, Loss: 0.1450941413640976, Validation Loss: 2.430475950241089, Validation Accuracy: 0.5105932354927063
Epoch 90, Loss: 0.08272408694028854, Validation Loss: 2.57120418548584, Validation Accuracy: 0.5148305296897888
Epoch 100, Loss: 0.051943961530923384, Validation Loss: 2.6827750205993652, Validation Accuracy: 0.5042372941970825
Total execution time for training: 12.375067234039307 seconds
```

For Problem 1, it seemed like the training length increased as the sequence length increased. And for all three models the best accuracy was seen for a sequence length of 20.

Problem 2:

Build the model for.LSTM and rnn.GRU for the tiny Shakespeare dataset, the data loader code is already provided.

1. Train the models for the sequence of 20 and 30, report and compare training loss, validation accuracy, execution time for training, and computational and mode size complexities across the two models.

2. Adjust the hyperparameters (fully connected network, number of hidden layers, and the number of hidden states) and compare your results (training and validation loss, computation complexity, model size, training and inference time, and the output sequence). Analyze their influence on accuracy, running time, and computational perplexity.

3. What if we increase the sequence length to 50. Perform the training and report the accuracy and model complexity results.

Andrew Nicola 801136465 HW3

LSTM sequence of 20:

```
Epoch 1/20, Training Loss: 1.711473316227781
Epoch 2/20, Training Loss: 1.485150503568518
Epoch 3/20, Training Loss: 1.4233778502858376
Epoch 4/20, Training Loss: 1.3878972135992893
Epoch 5/20, Training Loss: 1.3628450302354782
Epoch 6/20, Training Loss: 1.3431480968573006
Epoch 7/20, Training Loss: 1.3279951620960262
Epoch 8/20, Training Loss: 1.3146173407292134
Epoch 9/20, Training Loss: 1.3034009359909486
Epoch 10/20, Training Loss: 1.294143872883767
Epoch 11/20, Training Loss: 1.2862671143062758
Epoch 12/20, Training Loss: 1.2784827889909842
Epoch 13/20, Training Loss: 1.2721875276175028
Epoch 14/20, Training Loss: 1.266045561926979
Epoch 15/20, Training Loss: 1.2608564193571565
Epoch 16/20, Training Loss: 1.255879172915299
Epoch 17/20, Training Loss: 1.251721989147566
Epoch 18/20, Training Loss: 1.2477851554674453
Epoch 19/20, Training Loss: 1.244573937858307
Epoch 20/20, Training Loss: 1.2407389667464341
Total execution time for training: 606.2889075279236 seconds
Accuracy on test set: 58.054465986775746%
```

LSTM sequence of 30:

```
Epoch 1/20, Training Loss: 1.697342387942904
Epoch 2/20, Training Loss: 1.4726576827257034
Epoch 3/20, Training Loss: 1.4113048192549644
Epoch 4/20, Training Loss: 1.375155462476794
Epoch 5/20, Training Loss: 1.3505117829527857
Epoch 6/20, Training Loss: 1.3311707029569033
Epoch 7/20, Training Loss: 1.3150749606162104
Epoch 8/20, Training Loss: 1.3021008289807567
Epoch 9/20, Training Loss: 1.291757515019818
Epoch 10/20, Training Loss: 1.2821564125778753
Epoch 11/20, Training Loss: 1.2734858036400325
Epoch 12/20, Training Loss: 1.2663496789820086
Epoch 13/20, Training Loss: 1.2603301568511431
Epoch 14/20, Training Loss: 1.254051627652165
Epoch 15/20, Training Loss: 1.2488141292668729
Epoch 16/20, Training Loss: 1.2438960185576908
Epoch 17/20, Training Loss: 1.2387259014450949
Epoch 18/20, Training Loss: 1.2359358922835404
Epoch 19/20, Training Loss: 1.2318230640406043
Epoch 20/20, Training Loss: 1.2287620891802487
Total execution time for training: 748.9822533130646 seconds
Accuracy on test set: 58.84441416038696%
```

LSTM sequence of 50:

```
Epoch 1/20, Training Loss: 1.7001969526324001
Epoch 2/20, Training Loss: 1.4711108197271183
Epoch 3/20, Training Loss: 1.4074978094295216
Epoch 4/20, Training Loss: 1.3704794130662543
Epoch 5/20, Training Loss: 1.3441443178835326
Epoch 6/20, Training Loss: 1.3238139850151724
Epoch 7/20, Training Loss: 1.3077308766606923
Epoch 8/20, Training Loss: 1.294043606399139
Epoch 9/20, Training Loss: 1.2827759207945235
Epoch 10/20, Training Loss: 1.2727265480669032
Epoch 11/20, Training Loss: 1.2645177987739904
Epoch 12/20, Training Loss: 1.2569114672215935
Epoch 13/20, Training Loss: 1.2498197430450046
Epoch 14/20, Training Loss: 1.2443295110932486
Epoch 15/20, Training Loss: 1.2378445891255954
Epoch 16/20, Training Loss: 1.2333719689848432
Epoch 17/20, Training Loss: 1.228789914083556
Epoch 18/20, Training Loss: 1.2248058697360265
Epoch 19/20, Training Loss: 1.221132945431269
Epoch 20/20, Training Loss: 1.2182156594082438
Total execution time for training: 870.7704834938049 seconds
Accuracy on test set: 59.194688638941315%
```

The best accuracy and training loss was seen on a sequence length of 50, however it took the longest to train.

GRU sequence of 20:

```
Epoch 1/20, Training Loss: 1.6957282759954515
Epoch 2/20, Training Loss: 1.4996592183942024
Epoch 3/20, Training Loss: 1.4534013345880339
Epoch 4/20, Training Loss: 1.4285255410995419
Epoch 5/20, Training Loss: 1.4115607911016637
Epoch 6/20, Training Loss: 1.399496109439119
Epoch 7/20, Training Loss: 1.3905963871886657
Epoch 8/20, Training Loss: 1.38464664220297
Epoch 9/20, Training Loss: 1.379856461775556
Epoch 10/20, Training Loss: 1.3759207847179247
Epoch 11/20, Training Loss: 1.374044992769168
Epoch 12/20, Training Loss: 1.37172578532533
Epoch 13/20, Training Loss: 1.369672107888034
Epoch 14/20, Training Loss: 1.3691545740086255
Epoch 15/20, Training Loss: 1.3696721814192374
Epoch 16/20, Training Loss: 1.3690125433641538
Epoch 17/20, Training Loss: 1.3704669377411225
Epoch 18/20, Training Loss: 1.3722592879832647
Epoch 19/20, Training Loss: 1.373418925542703
Epoch 20/20, Training Loss: 1.3763837425601393
Total execution time for training: 474.571594953537 seconds
Accuracy on test set: 56.455900481900706%
```

GRU sequence of 30:

```
Epoch 1/20, Training Loss: 1.6861268352723573
Epoch 2/20, Training Loss: 1.4944427352541831
Epoch 3/20, Training Loss: 1.4488336394828965
Epoch 4/20, Training Loss: 1.4227918446388863
Epoch 5/20, Training Loss: 1.407423982812424
Epoch 6/20, Training Loss: 1.396357479455549
Epoch 7/20, Training Loss: 1.3875296961898715
Epoch 8/20, Training Loss: 1.3806749197108883
Epoch 9/20, Training Loss: 1.3746316732087493
Epoch 10/20, Training Loss: 1.3712931449855803
Epoch 11/20, Training Loss: 1.3683031995834734
Epoch 12/20, Training Loss: 1.3650096261973226
Epoch 13/20, Training Loss: 1.3632348598297888
Epoch 14/20, Training Loss: 1.3652427787419645
Epoch 15/20, Training Loss: 1.366415409532622
Epoch 16/20, Training Loss: 1.3651240237283762
Epoch 17/20, Training Loss: 1.3639760853501481
Epoch 18/20, Training Loss: 1.3656101293489977
Epoch 19/20, Training Loss: 1.3670186329541094
Epoch 20/20, Training Loss: 1.3685405236325316
Total execution time for training: 390.2198209762573 seconds
Accuracy on test set: 56.78051579527778%
```

GRU sequence of 50:

```
Epoch 1/20, Training Loss: 1.683142794476505
Epoch 2/20, Training Loss: 1.4841036172221644
Epoch 3/20, Training Loss: 1.4351261619720546
Epoch 4/20, Training Loss: 1.4081026340485165
Epoch 5/20, Training Loss: 1.3900087800605314
Epoch 6/20, Training Loss: 1.3779923070060216
Epoch 7/20, Training Loss: 1.3691461610753026
Epoch 8/20, Training Loss: 1.3626731731357746
Epoch 9/20, Training Loss: 1.3562186070672582
Epoch 10/20, Training Loss: 1.3521321563138593
Epoch 11/20, Training Loss: 1.3495454395831425
Epoch 12/20, Training Loss: 1.3487541041728535
Epoch 13/20, Training Loss: 1.3455588955342932
Epoch 14/20, Training Loss: 1.3471375236875818
Epoch 15/20, Training Loss: 1.3462522669102537
Epoch 16/20, Training Loss: 1.3446821476000315
Epoch 17/20, Training Loss: 1.3478725592403633
Epoch 18/20, Training Loss: 1.3475939247550384
Epoch 19/20, Training Loss: 1.350850970634616
Epoch 20/20, Training Loss: 1.353002951667087
Total execution time for training: 493.34515857696533 seconds
Accuracy on test set: 56.89315861908199%
```

Andrew Nicola 801136465 HW3

The best accuracy and training loss was seen on a sequence length of 50, however it took the longest to train. It appears that for both models this time the best sequence length was 50, the only issue is the long run time in comparison to the other model length.

GitHub Link

P1: https://github.com/Eskdagoat/4106/blob/main/Homework3P1.ipynb

P2: https://github.com/Eskdagoat/4106/blob/main/Homework3P2.ipynb