

Names to Nationality Hyper-Parameter Search:

Objective: Classify 123 classes with only their first and last names with 1000 examples per class. Run each task for 10 epoches.

Task 1: Get the best learning rate from [0.1, 0.01, 0.001, 0.0001, 0.00001]. Set momentum = 0, L2 = 0

- For learning rate 0.1:

Not started because we assume that since the learning rate for 0.01 is bad, then the learning rate for 0.1 would be worse.

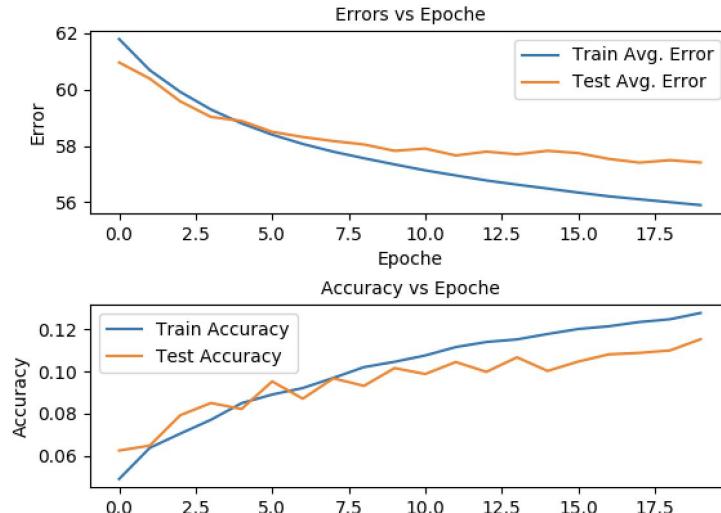
- For learning rate 0.01:

Terminated early because accuracy is decreasing:

```
RNN with learning rate: 0.01, momentum: 0, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epoche: 10
0 63.935244654407306 63.59686936500215 0.011864470754016683 0.012505871301080319 1.0 1571679966.8605103
1 64.75220591458638 63.59340961745712 0.012128684842918434 0.008953734147487083 1.0 1571681993.636366
2 65.24129120778454 64.08981004405547 0.010870522514814861 0.012740723344293095 1.0 1571684007.9601748
3 64.82251564838502 64.61231514039837 0.010203696480919969 0.011302254579614843 1.0 1571686010.365927
```

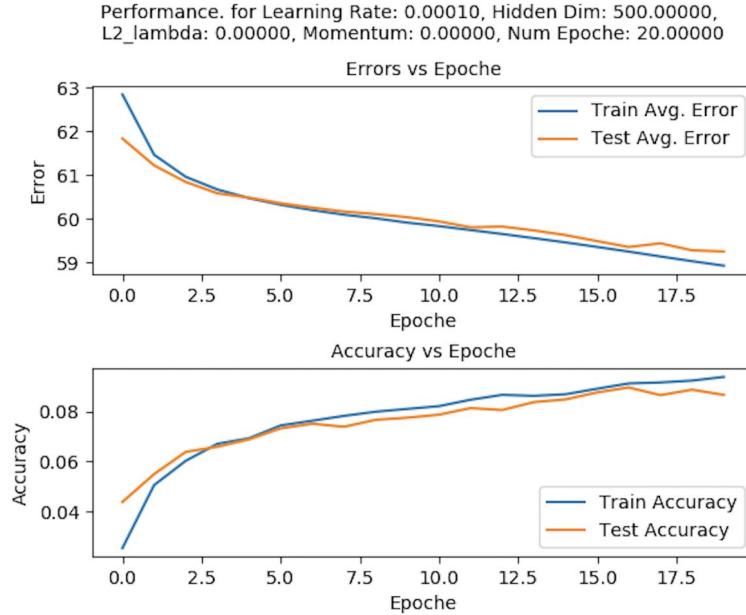
- For learning rate 0.001:

Performance. for Learning Rate: 0.00100, Hidden Dim: 500.000000, L2_lambda: 0.00000, Momentum: 0.00000, Num Epoche: 20.00000



```
1 60.68837796193696 60.38840474404318 0.0637861377166375 0.06487137319393868 1.0 1571731433.6165292
2 59.91464045531868 59.57847588563436 0.0705826085862082 0.07931986373781276 1.0 1571733453.387655
3 59.29105574964269 59.0287322010314 0.07713994436962733 0.08510513332550217 1.0 1571735492.5323923
4 58.7996441625606 58.88599456580997 0.08503140221262885 0.08225654880770586 1.0 1571737506.3469381
5 58.39772040224652 58.49842151626507 0.08913445684870301 0.0953541642194291 1.0 1571739511.7146425
6 58.068755412896515 58.3142296569843 0.09220545479717569 0.08716081287442734 1.0 1571741521.4828594
7 57.79375408705689 58.17865881526952 0.09702591469170453 0.09682250675437566 1.0 1571743545.102646
8 57.56087598699212 58.050783603805215 0.102148439582143 0.09323975096910607 1.0 1571745597.8143685
9 57.34184957617155 57.82528450383776 0.10469082350572036 0.10169748397039821 1.0 1571747624.8331826
10 57.13128030383463 57.904070247737046 0.10768630511119781 0.09887818630330084 1.0 1571749666.3846772
11 56.95028628613478 57.660211269711915 0.11167688523277912 0.1045753533889346 1.0 1571751687.2047694
12 56.772644628126486 57.79630778986574 0.11409260820862649 0.09990602607776342 1.0 1571753678.1702878
13 56.62471594169955 57.702181640418615 0.11532684181088191 0.10680723599201222 1.0 1571755709.8539927
14 56.48609810027512 57.82544569777177 0.11789359747271973 0.10034652883824739 1.0 1571757734.1273515
15 56.34367006649327 57.745055449804966 0.1202597762199836 0.10478092329378598 1.0 1571759747.123862
16 56.20971040258272 57.5380276956929 0.12155614910798838 0.1082168448255609 1.0 1571761744.934348
17 56.10455808643206 57.4062566886135 0.1235824953116937 0.10892164924233526 1.0 1571763756.3631785
18 56.00145615249203 57.49223744771872 0.12489144525694436 0.11006695641959356 1.0 1571765765.9596605
19 55.89820566718577 57.411907688278006 0.12787434088525595 0.11541172324679902 1.0 1571767797.5573351
Saved model to LX1-H-500-R-0-001-M-0-E-20-data.npz
```

- For learning rate 0.0001:



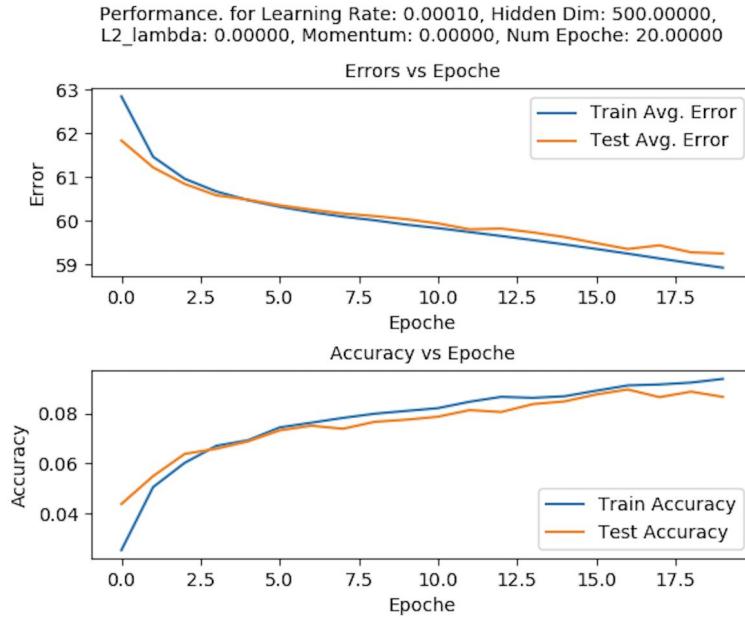
```
RNN with learning rate: 0.0001, momentum: 0, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epoche: 20
0 62.84911228494603 61.83849154595875 0.025260216731904397 0.04375660754140726 1.0 1571689348.782848
1 61.46716441572205 61.22269219926196 0.05050784740664292 0.05491601080700106 1.0 1571691365.3829112
2 60.96042078924245 60.84251508088694 0.06026204171819345 0.0638435334194761 1.0 1571693370.600126
3 60.6702391995622 60.58202287727405 0.06707109863693 0.06589921296840127 1.0 1571695380.0197427
4 60.47101331100456 60.48209057103728 0.0692988307552893 0.0688652648889933 1.0 1571697376.8949182
5 60.31910178844328 60.353633846197866 0.0744465281361398 0.07327029249383296 1.0 1571699356.605471
6 60.197286015062005 60.25299525666527 0.07634702276817741 0.07514977093856455 1.0 1571701352.1188824
7 60.092409146457456 60.16432235234786 0.07828527557172164 0.07391636320920944 1.0 1571703367.9184425
8 60.00846713470525 60.10557460524644 0.07995922117478257 0.0767862140256079 1.0 1571705396.333144
9 59.90916638853033 60.033476717932686 0.08107938026254516 0.0775578526958769 1.0 1571707379.8685918
10 59.83015811720988 59.938216216058116 0.08218695329314185 0.07876189357453307 1.0 1571709375.945313
11 59.739062085144255 59.802173883272964 0.08476709501214555 0.08140491013743686 1.0 1571711379.684177
12 59.648641328208384 59.819442446457614 0.08673051993002152 0.08064137201926465 1.0 1571713405.4628396
13 59.553148909780255 59.7310895264597 0.08632776610071363 0.08381299189474921 1.0 1571715391.2476563
14 59.45542257626364 59.62375483146922 0.0869578689590072 0.08489956537060965 1.0 1571717412.6164129
15 59.35161374531935 59.485636020352544 0.0891848010773665 0.08768941618700811 1.0 1571719429.1913457
16 59.24455326097998 59.35016115071834 0.0913118447383988 0.08968636203453542 1.0 1571721410.017823
17 59.13303246198649 59.43495770715283 0.09170201251054082 0.08660284271114765 1.0 1571723379.9344616
18 59.024597397326495 59.27352438267258 0.09244458988332725 0.08880535651356748 1.0 1571725396.8049989
19 58.922327642705866 59.24810904699405 0.09392974462890011 0.08672031011394338 1.0 1571727385.4490402
Saved model to LX1-H-500-R-0_0001-M-0-E-20-data.npz
```

- For learning rate 0.00001:

From the data above, we can see that lower learning rates will lead to slower convergence.

Task 2: Get the best momentum value from [0, 0.9, 0.98, 0.5]. Set the learning rate to 0.0001, and run it for 20 epoches:

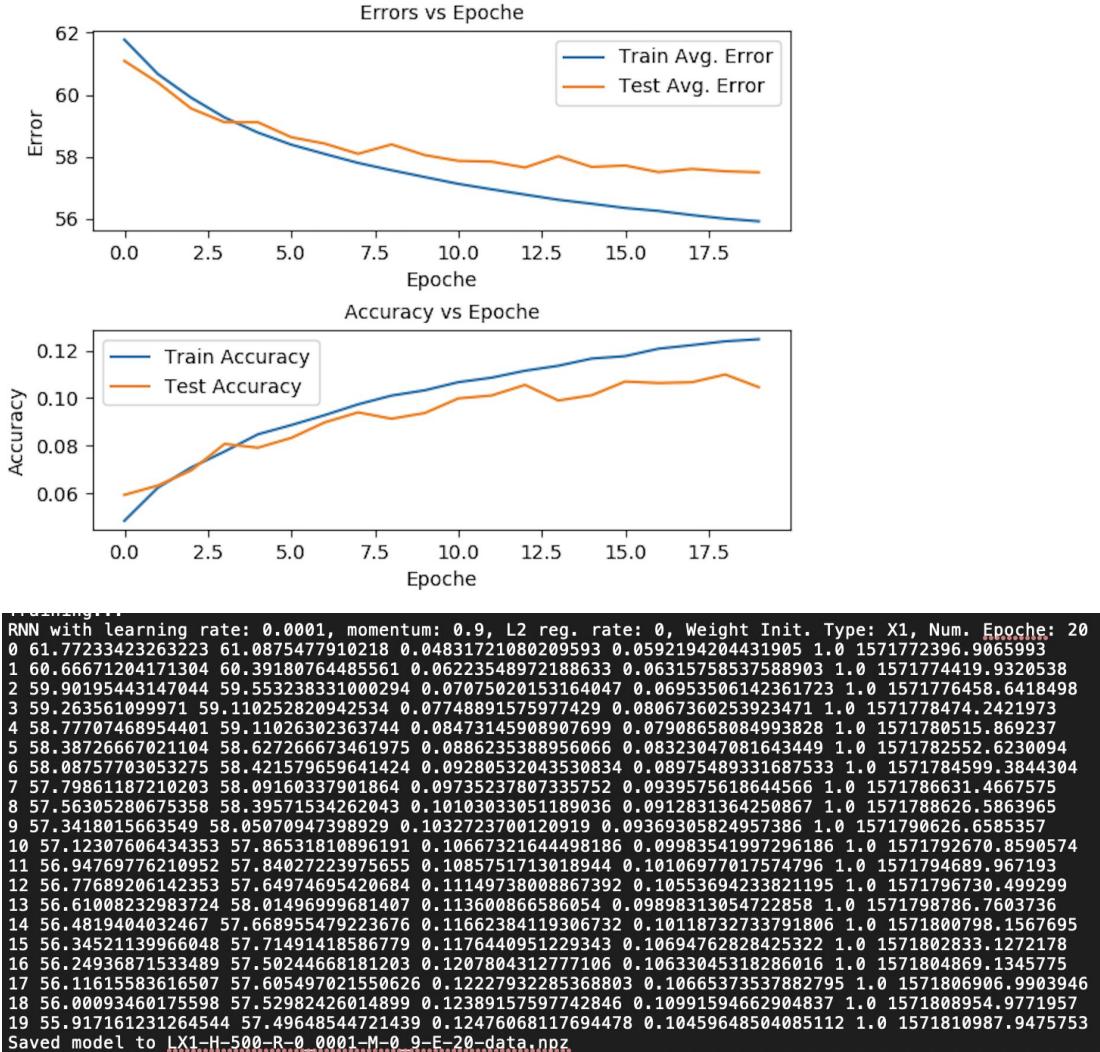
- For 0:



```
RNN with learning rate: 0.0001, momentum: 0, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epochs: 20
0 62.84911228494603 61.83849154595875 0.025260216731904397 0.04375660754140726 1.0 1571689348.782848
1 61.46716441572205 61.22269219926196 0.05050784740664292 0.05491601080700106 1.0 1571691365.3829112
2 60.96042070924245 60.84251508088694 0.06026204171019345 0.0638435334194761 1.0 1571693370.600126
3 60.6702391995622 60.582022287727405 0.06707109863693 0.06589921296840127 1.0 1571695380.0197427
4 60.47101331100456 60.48209057103728 0.0692988307552893 0.0688652648889933 1.0 1571697376.8949182
5 60.31910178844328 60.353633846197866 0.0744465281361308 0.07327029249383296 1.0 1571699356.605471
6 60.197286015062005 60.25299525666527 0.07634782276817741 0.07514977093856455 1.0 1571701352.1188824
7 60.092409146457456 60.16432235234786 0.07828527557172164 0.07391636320920944 1.0 1571703367.9184425
8 60.00846713470525 60.10557460524644 0.07995922117478257 0.0767062140256079 1.0 1571705396.333144
9 59.90916638853033 60.033476717932686 0.08107938026254516 0.0778578526958769 1.0 1571707379.8685918
10 59.83015811720988 59.938216216058116 0.08218695329314185 0.07876189357453307 1.0 1571709375.945313
11 59.739062085144255 59.802173883272964 0.08476709501214555 0.08140491013743686 1.0 1571711379.684177
12 59.648641328208384 59.819442446457614 0.08673051993002152 0.08064137201926465 1.0 1571713405.4628396
13 59.553148909780255 59.7310895264597 0.08632776610071363 0.08381299189474921 1.0 1571715391.2476563
14 59.45542257626364 59.62375483146922 0.08695706868950072 0.08489956537060965 1.0 1571717412.6164129
15 59.35161374531935 59.485636020352544 0.0891848010773665 0.08768941618700811 1.0 1571719429.1913457
16 59.24455326097998 59.35016115071834 0.0913118447383988 0.08968636203453542 1.0 1571721410.017823
17 59.13303246198649 59.43495770715283 0.09170201251054082 0.08660284271114765 1.0 1571723379.9344616
18 59.024597397326495 59.27352438267258 0.09244458988332725 0.08880535651356748 1.0 1571725396.8049989
19 58.922327642705866 59.24810904699405 0.09392974462890011 0.08672031011394338 1.0 1571727385.4490402
Saved model to LX1-H-500-R-0_0001-M-0-E-20-data.npz
```

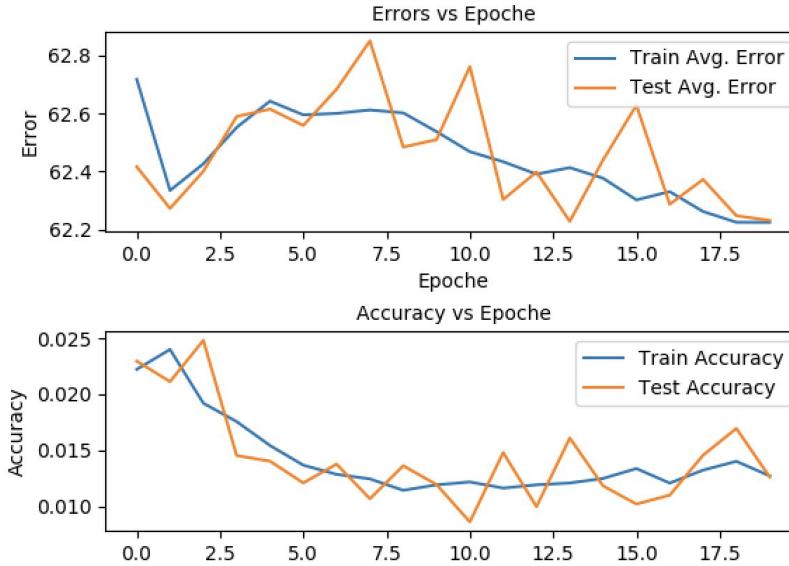
- For 0.9:

Performance. for Learning Rate: 0.00010, Hidden Dim: 500.00000,
L2_lambda: 0.00000, Momentum: 0.90000, Num Epochs: 20.00000



- For 0.98:

Performance. for Learning Rate: 0.00010, Hidden Dim: 500.00000,
L2_lambda: 0.00000, Momentum: 0.98000, Num Epoche: 20.00000

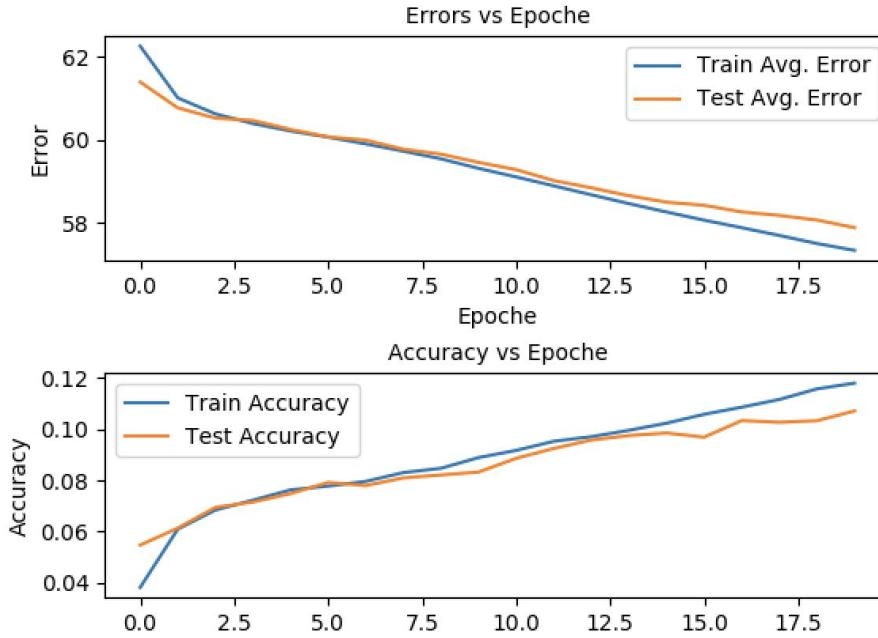


```
RNN with learning rate: 0.0001, momentum: 0.98, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epoche: 20
0 62.71823920849786 62.416573473490836 0.022218863361547762 0.022953035913713043 1.0 1571812983.2710261
1 62.33505944561285 62.273233048521675 0.023994860943168078 0.02113089990007641 1.0 1571814964.6539905
2 62.426789648972196 62.40102683002873 0.019208484482063685 0.0248045612178922 1.0 1571816963.4824932
3 62.5522719692508 62.58974495051689 0.017571039903264814 0.01454769881855052 1.0 1571819004.3803697
4 62.64282949475423 62.61540239294397 0.01544236195082628 0.014048080879327574 1.0 1571821003.80704
5 62.59548762906139 62.559041994409704 0.01369155582426441 0.012108387703520837 1.0 1571823032.0896478
6 62.600649671827234 62.68388124607012 0.012872833534864974 0.013783577264444836 1.0 1571825026.9642637
7 62.61232007631653 62.80506217371092 0.012469770253929867 0.010697701757479574 1.0 1571827017.2261984
8 62.602278674593094 62.484682483614165 0.011449516324062877 0.013636630811732204 1.0 1571829013.5626998
9 62.53781397675134 62.510004432835274 0.011940749697702539 0.011961441250808206 1.0 1571831038.012801
10 62.46856105032859 62.76167838465641 0.012192664248286982 0.008640451419502733 1.0 1571833065.3818665
11 62.43401196812122 62.30367547392853 0.01165104796453043 0.014812202433433256 1.0 1571835097.3722835
12 62.39116436278676 62.398766962417525 0.011940749697702539 0.009992358784458944 1.0 1571837119.31248
13 62.41366223315597 62.22866957090087 0.012104494155582426 0.016105331217304415 1.0 1571839117.268713
14 62.3771007854193 62.44069467582645 0.012507557436517533 0.011843884088638101 1.0 1571841111.0684266
15 62.30238619719256 62.62886140894305 0.01338925836356308 0.010227473108799153 1.0 1571843128.1119049
16 62.33099174680454 62.28756702954997 0.012091898428053204 0.011020983953447364 1.0 1571845144.4028282
17 62.262439099288834 62.37294566643085 0.013250705360741636 0.014577088109093047 1.0 1571847151.261321
18 62.22535706489181 62.24799800823317 0.014031640467553405 0.01695762064303768 1.0 1571849151.9331431
19 62.225034718366125 62.23173278974411 0.012734280532043531 0.01263739493328631 1.0 1571851158.0949655
Saved model to LX1-H-500-R-0_0001-M-0_98-E-20-data.npz
```

This means that it overshot the minima :(

- For 0.5:

Performance. for Learning Rate: 0.00010, Hidden Dim: 500.00000,
L2_lambda: 0.00000, Momentum: 0.50000, Num Epochs: 20.00000

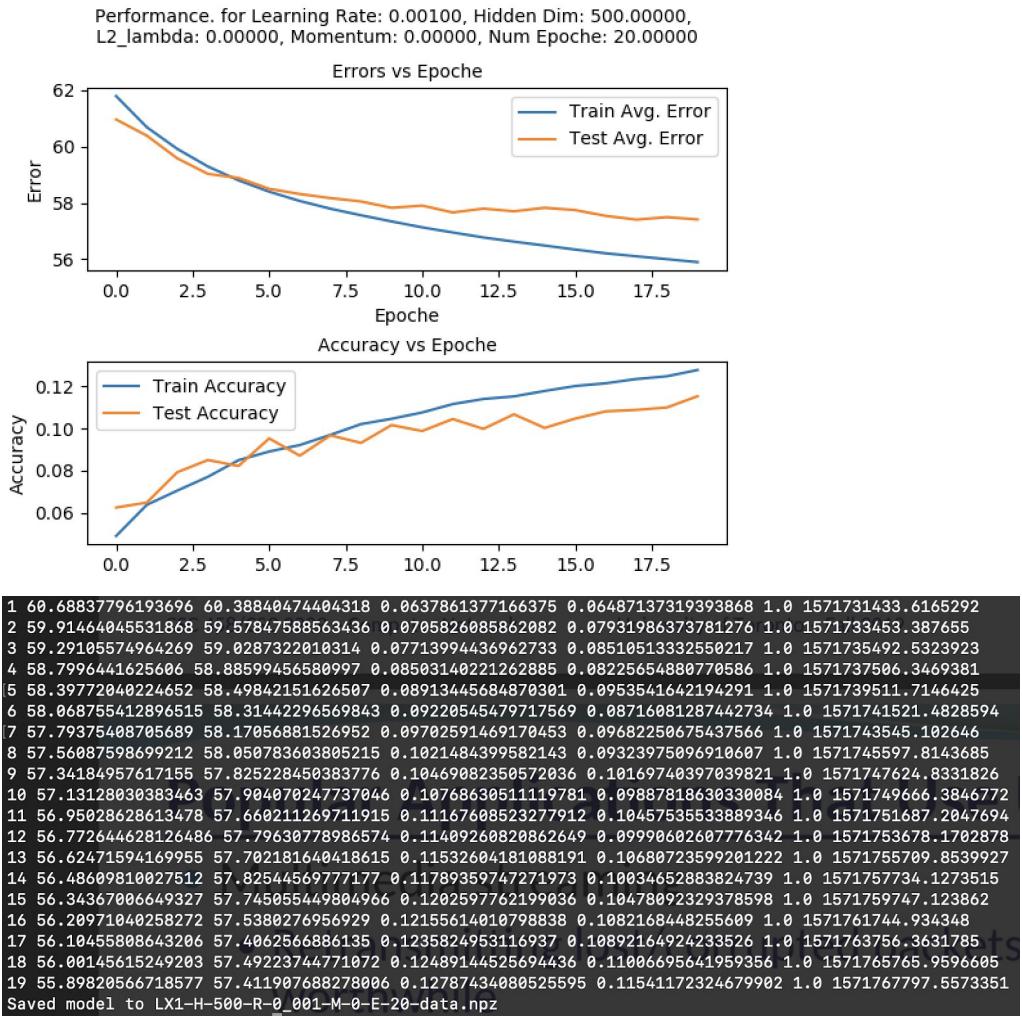


```
RNN with learning rate: 0.0001, momentum: 0.5, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epochs: 20
0 62.26656232696914 61.40124107666506 0.03811467150342604 0.054664080409098925 1.0 1571853170.4480286
1 61.01846892841385 60.77842027986427 0.060925534058847236 0.06130606007170987 1.0 1571855170.3411372
2 60.63425735911729 60.53242837968157 0.0683696090286175 0.06938811497090461 1.0 1571857198.3462367
3 60.40011945482965 60.4735727337157 0.07223649738008868 0.0715041438899665 1.0 1571859209.849016
4 60.22136163477922 60.25891596129002 0.07625453446191052 0.07479574443072944 1.0 1571861213.903078
5 60.07300661960927 60.079577085451156 0.07774083031035872 0.07914535943102334 1.0 1571863165.6627946
6 59.91076051885166 60.000812419798635 0.07957980652962515 0.07794039851877976 1.0 1571865187.3587427
7 59.73569658493712 59.78113487321671 0.08301844014510278 0.08093810615411744 1.0 1571867138.2295275
8 59.55100775745517 59.66414772109248 0.08469367190648931 0.08208428848527596 1.0 1571869141.5331087
9 59.32118683626806 59.465157664207 0.08888804917372027 0.08323047081643449 1.0 1571871126.086415
10 59.11482330378078 59.29084710647767 0.0916968964127368 0.08857932169517428 1.0 1571873103.329204
11 58.90069611865206 59.02698713929536 0.09526148730350666 0.09245870804678775 1.0 1571875079.122056
12 58.68469830866187 58.85271241247728 0.09706267634018541 0.0957503085875507 1.0 1571877080.2815812
13 58.47218183453562 58.66290105178981 0.09955663039097139 0.09751366602010228 1.0 1571879092.7150695
14 58.27058331591121 58.506747788976796 0.10234028617492946 0.09851290189854817 1.0 1571881126.2981164
15 58.073496063755165 58.43158952011448 0.10581670697299476 0.0968671016281667 1.0 1571883146.49481
16 57.892767124697166 58.272263391757235 0.10860036275695284 0.10336213483806501 1.0 1571885144.0633676
17 57.705528917048255 58.18767998770365 0.11161074163643692 0.1026861811555869 1.0 1571887179.0950093
18 57.511319229774166 58.077471030854106 0.11575473599355099 0.10330335625697995 1.0 1571889197.07243
19 57.347017818649554 57.893337880081724 0.11800937122128174 0.10712396402750837 1.0 1571891192.0071576
Saved model to LX1-H-500-R-0.0001-M-0.5-E-20-data.npz
```

- From this we can conclude that 0.98 is worse than 0.9 or 0.5 since it overshot the minima. But the performance between 0.9 and 0.5 is subtle.

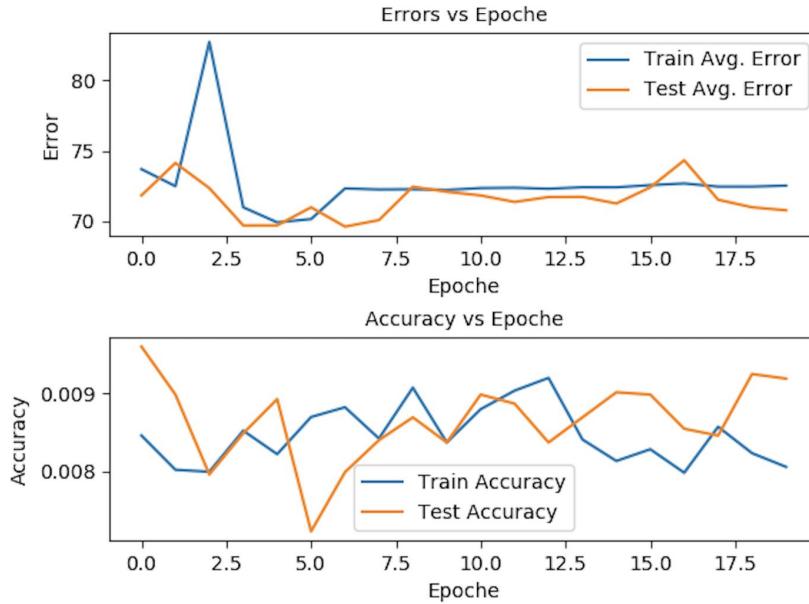
Task 3: Since the curve for the learning rates 0.001 and 0.0001 are very similar, we are going to do Task 2 again for learning rate 0.001. Get the best momentum value from [0.9, 0.98, 0.5]. Set the learning rate to 0.001, and run it for 20 epoches:

- For 0:



- For 0.98:

Performance. for Learning Rate: 0.00100, Hidden Dim: 500.00000,
L2_lambda: 0.00000, Momentum: 0.98000, Num Epoche: 20.00000



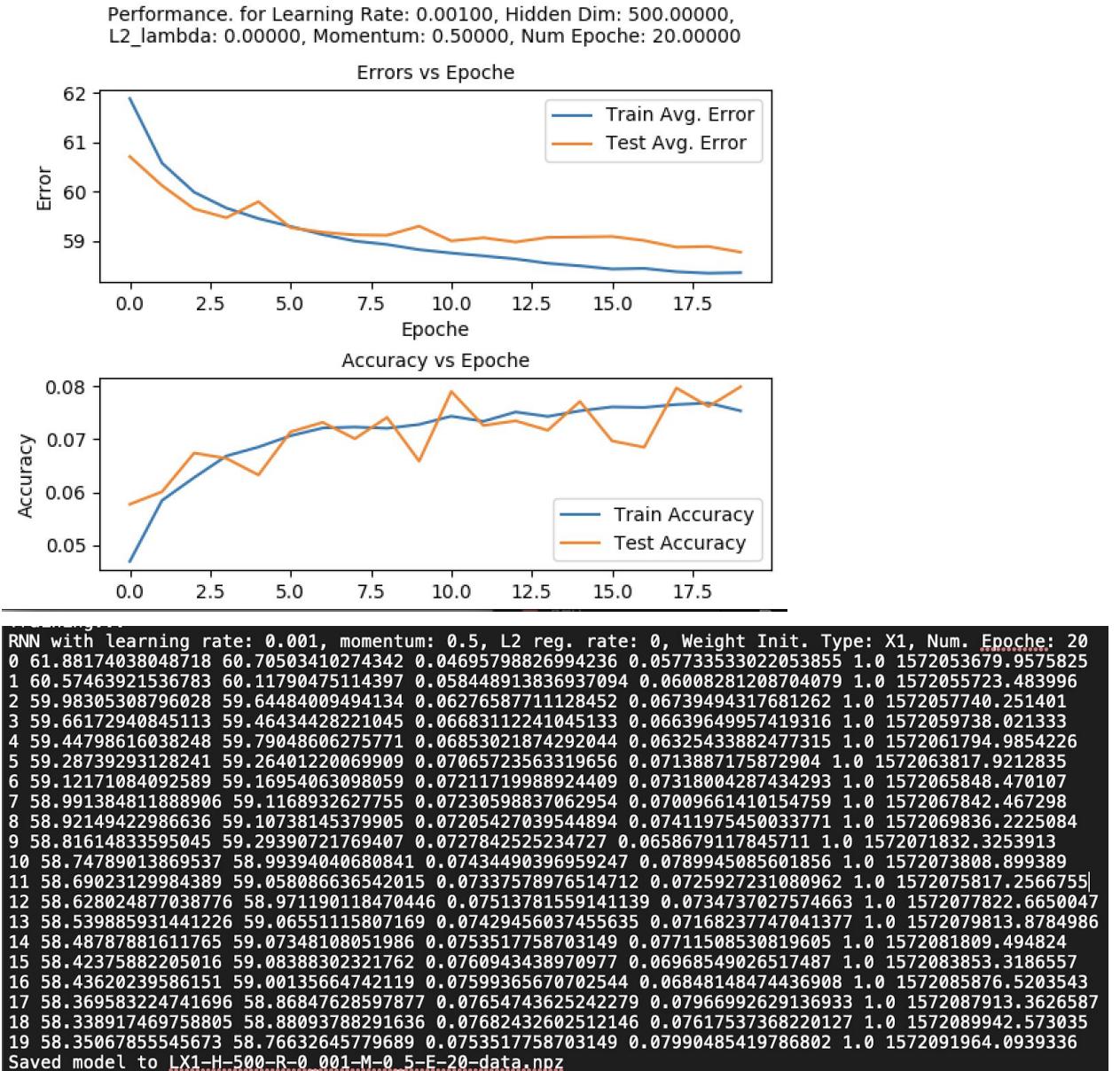
```
RNN with learning rate: 0.001, momentum: 0.98, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epoche: 20
0 73.696855628882 71.84188476412788 0.008457723966068417 0.009602678178134086 1.0 1572093948.4534814
1 72.4914289893714 74.13455244762743 0.008017217509502354 0.008985992423575015 1.0 1572095978.377177
2 82.72007220845929 72.35579706634634 0.00799204571984292 0.007958182832643233 1.0 1572098006.1053925
3 70.99544844058514 69.70844755171389 0.008520653459863569 0.008486770622265292 1.0 1572100149.6881049
4 69.92547045250132 69.71287709489816 0.00821859188964684 0.008927260446950342 1.0 1572102711.108196
5 70.16037737095536 70.99250505967062 0.008696856042489993 0.0072240331248348165 1.0 1572104934.4108472
6 72.33251920917884 69.62625033992568 0.008822715030080298 0.00798754882095557 1.0 1572106989.1775002
7 72.25950413802101 70.08679335637017 0.008419966269791326 0.008398672657328283 1.0 1572109042.134023
8 72.27476388349083 72.45404165290424 0.009074433005260906 0.00869233254045165 1.0 1572111063.0500536
9 72.22664803807221 72.09490449617088 0.008369622674755205 0.008369306669015945 1.0 1572113106.1427097
10 72.3578256407891 71.83660625232719 0.008797543232562237 0.008985992423575015 1.0 1572115160.1425745
11 72.38222446405479 71.37350516032072 0.009036675308983815 0.00886852847032567 1.0 1572117215.0334466
12 72.30966041333453 71.72346461869107 0.00920029199285121 0.008369306669015945 1.0 1572119242.536775
13 72.41605206149487 71.73255503096574 0.008407380371032296 0.00869233254045165 1.0 1572121250.6783316
14 72.41277138180457 71.26660009380424 0.008130490598333626 0.009015358411887352 1.0 1572123284.8152146
15 72.56834233112126 72.40875708009175 0.008281521383441992 0.008985992423575015 1.0 1572125310.2204554
16 72.68487791777052 74.32744511038082 0.007979459813225263 0.008545502598889965 1.0 1572127349.203146
17 72.46278362181397 71.53195988172034 0.008570997054899691 0.008457404633952956 1.0 1572129374.1363428
18 72.46171257196809 70.99615234004814 0.00823117778840587 0.009250286318386045 1.0 1572131430.9816108
19 72.53190154167733 70.78419168557618 0.008054975205779445 0.009191554341761372 1.0 1572133480.4219391
Saved model to LX1-H-500-R-0_001-M-0_98-E-20-data.npz
```

- For 0.90:



```
RNN with learning rate: 0.001, momentum: 0.9, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epoche: 20
0 63.707818990702236 63.526751173387225 0.009754071538248546 0.009543946201509413 1.0 1572013197.4304874
1 63.59368182932891 63.62146643223814 0.009452009968031817 0.008398672657328283 1.0 1572015191.8159947
2 63.58657415455723 63.412436620356424 0.008948574017670602 0.008985992423575015 1.0 1572017197.7903616
3 63.56528302570291 63.66464063206211 0.008747199637526116 0.008339940680703609 1.0 1572019230.1863885
4 63.59689320330199 63.66547263584945 0.00859616885241775 0.009279652306698381 1.0 1572021249.7931616
5 63.6542582678549 63.56026336738216 0.008797543232562237 0.008574868587202302 1.0 1572023291.8983753
6 63.894542903099094 63.98791644020782 0.008847886827598359 0.007928816844330895 1.0 1572025326.51113
7 64.10140067131182 63.88909805670404 0.009489767664308908 0.009074090388512025 1.0 1572027370.0266194
8 64.16666531037102 64.03722993359182 0.009162534296574118 0.011217807535312601 1.0 1572029377.6725495
9 63.98856510648519 63.7502609219925 0.009250635587887331 0.009220920330073708 1.0 1572031392.5191762
10 63.90839542547273 63.66697919078743 0.00921287789161024 0.009455848236572402 1.0 1572033429.923906
11 63.964656032592636 63.74497718417268 0.008709441941249025 0.007752620914456876 1.0 1572035457.409812
12 63.965359540547574 63.89050586514619 0.009401666372995696 0.010248729921005492 1.0 1572037500.9097292
13 64.03194303443578 63.75085387543601 0.009338736879200544 0.00886852847032567 1.0 1572039554.5880885
14 63.9923516903134 63.799422068676996 0.008508067561104539 0.009778874108008106 1.0 1572041558.4128127
15 63.945923164923265 63.73636626186217 0.009414252271754726 0.009632044166446422 1.0 1572043560.7363155
16 63.893988512298364 63.56950031284144 0.009628212550658242 0.010248729921005492 1.0 1572045605.525557
17 64.18846153711162 64.14334896870216 0.009527525360585999 0.010953513640501572 1.0 1572047634.917908
18 64.07548628080168 63.71603958263925 0.009401666372995696 0.009720142131383431 1.0 1572049666.3725605
19 63.991310746461544 63.71038770829188 0.009426838170513756 0.008897894458638006 1.0 1572051697.5121393
Saved model to LX1-H-500-R-0_001-M-0_9-E-20-data.npz
```

- For 0.50:

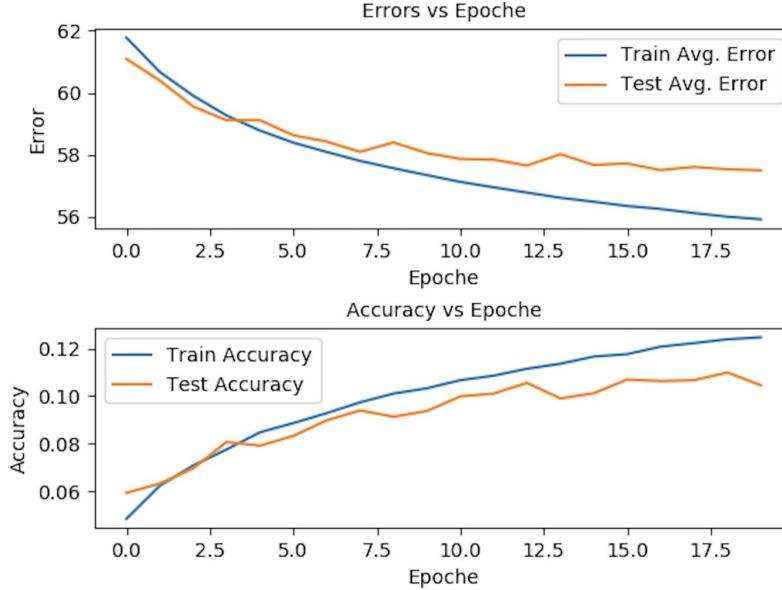


- As one can see, using 0.5 is better than using 0.9 or 0.98 simply because 0.9 and 0.98 caused the algorithm to overshoot its minima. Moreover, one can also see that using a learning rate of 0.01 (in general) is worse than using a learning rate of 0.0001. Thus, using 0.0001 as the learning rate seems to be the most sensible solution.

Task 4: Get the best L2 regularizer from [0.1, 0.01, 0.001, 0.0001]. Set learning rate to the best one from Task 2-3 (it is 0.0001), and set the momentum to the best one from Task 2-3 (it is 0.9):

- For L2 regularizer 0:

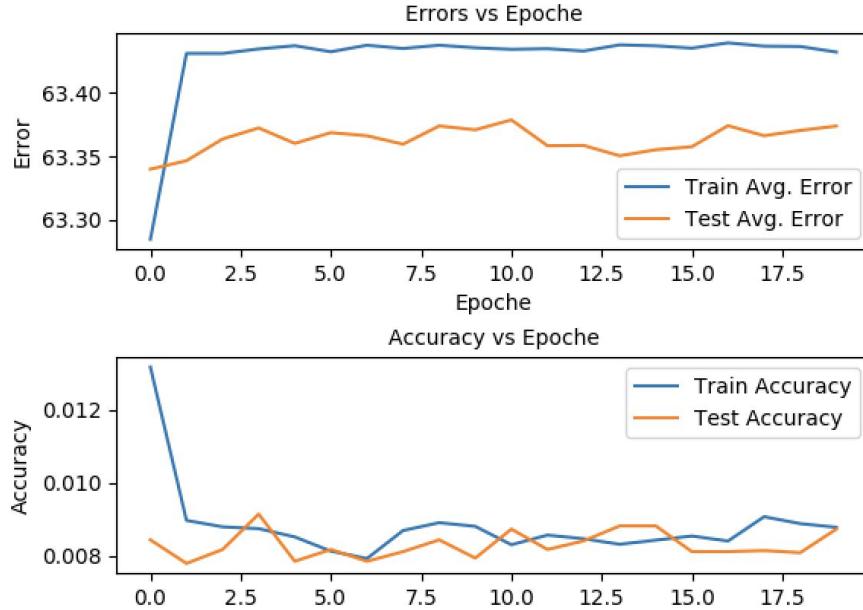
Performance. for Learning Rate: 0.00010, Hidden Dim: 500.00000,
L2_lambda: 0.00000, Momentum: 0.90000, Num Epochs: 20.00000



```
TRAINING.RNN
RNN with learning rate: 0.0001, momentum: 0.9, L2 reg. rate: 0, Weight Init. Type: X1, Num. Epochs: 20
0 61.77233423263223 61.0875477910218 0.04831721080209593 0.0592194204431905 1.0 1571772396.9065993
1 60.66671204171304 60.39180764485561 0.06223548972188633 0.06315758537588903 1.0 1571774419.9320538
2 59.90195443147044 59.553238331000294 0.07075020153164047 0.06953506142361723 1.0 1571776458.6418498
3 59.263561099971 59.110252820942534 0.07748891575977429 0.08067360253923471 1.0 1571778474.2421973
4 58.77707468954401 59.11026302363744 0.08473145908907699 0.07908658084993828 1.0 1571780515.869237
5 58.38726667021104 58.627266673461975 0.0886235388956066 0.08323047081643449 1.0 1571782552.6230094
6 58.08757703053275 58.421579659641424 0.09280532043530834 0.08975489331687533 1.0 1571784599.3844304
7 57.79861187210203 58.09160337901864 0.09735237807335752 0.0939575618644566 1.0 1571786631.4667575
8 57.56305280675358 58.39571534262043 0.10103033051189036 0.0912831364250867 1.0 1571788626.5863965
9 57.3418015663549 58.05070947398929 0.1032723700120919 0.09369305824957386 1.0 1571790626.6585357
10 57.12307606434353 57.86531810896191 0.10667321644498186 0.09983541997296186 1.0 1571792670.8590574
11 56.94769776210952 57.84027223975655 0.1085751713018944 0.10106977017574796 1.0 1571794689.967193
12 56.777689206142353 57.64974695420684 0.11149738008867392 0.10553694233821195 1.0 1571796730.499299
13 56.61008232983724 58.01496999681407 0.113600866586054 0.09898313054722858 1.0 1571798786.7603736
14 56.4819404032467 57.668955479223676 0.11662384119306732 0.10118732733791806 1.0 1571800798.1567695
15 56.3452113996048 57.71491418586779 0.1176440951229343 0.10694762828425322 1.0 1571802833.1272178
16 56.24936871533489 57.50244668181203 0.1207804312777106 0.10633045318286016 1.0 1571804869.1345775
17 56.11615583616507 57.605497021550626 0.12227932285368803 0.10665373537882795 1.0 1571806906.9903946
18 56.00093460175598 57.52982426014899 0.12389157597742846 0.10991594662904837 1.0 1571808954.9771957
19 55.917161231264544 57.49648544721439 0.12476068117694478 0.10459648504085112 1.0 1571810987.9475753
Saved model to LX1-H-500-R-0_0001-M-0_9-E-20-data.npz
```

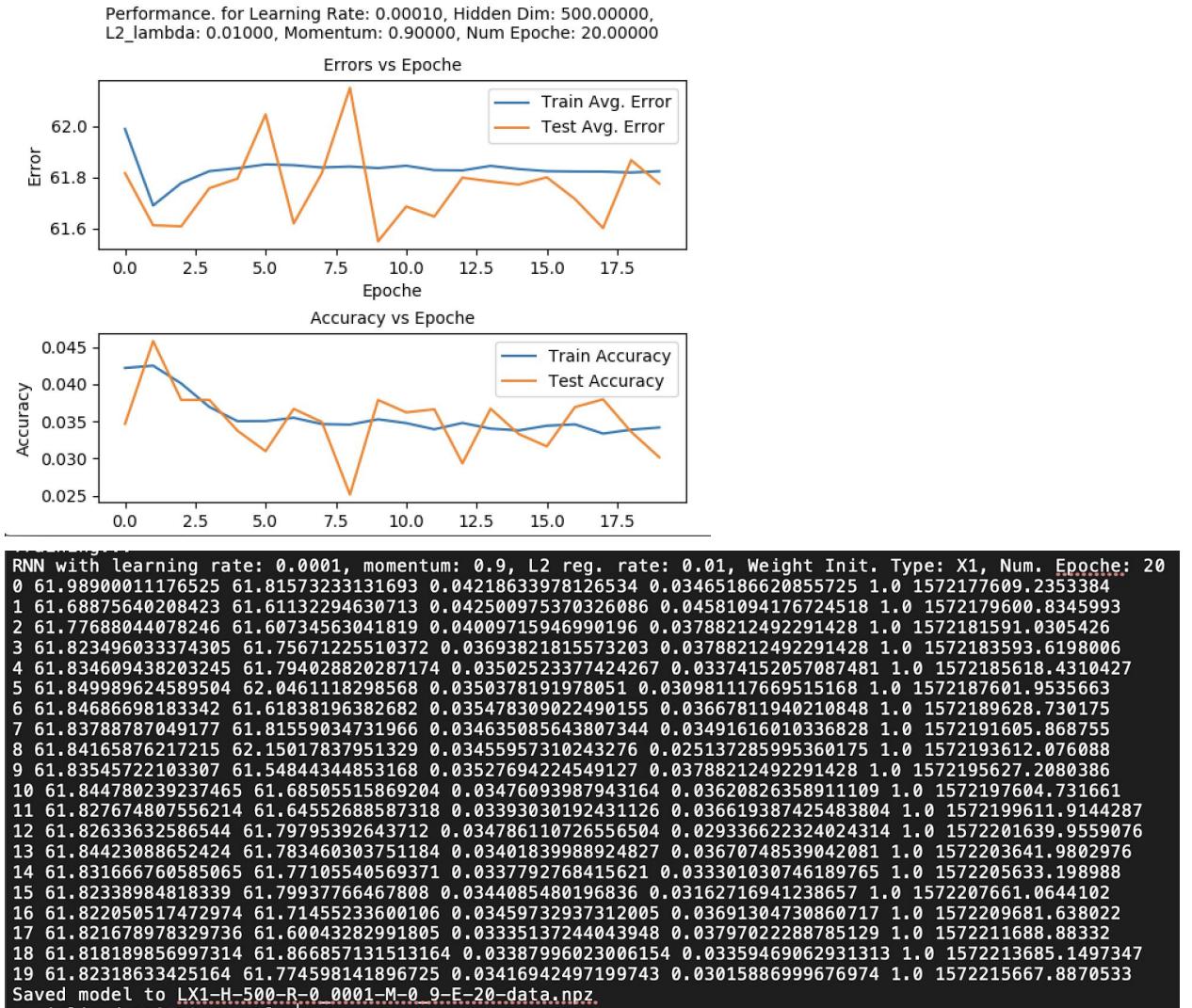
- For L2 regularizer 0.1:

Performance. for Learning Rate: 0.00010, Hidden Dim: 500.00000,
L2_lambda: 0.10000, Momentum: 0.90000, Num Epochs: 20.00000



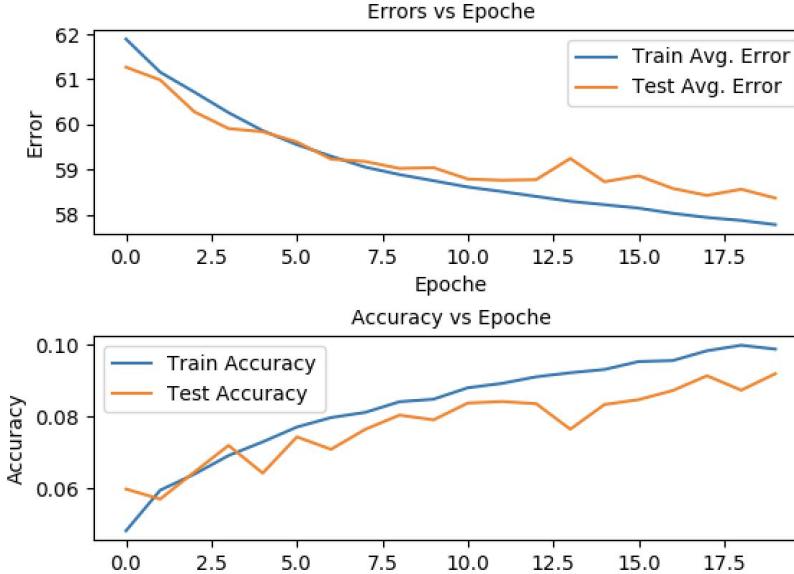
```
RNN with learning rate: 0.0001, momentum: 0.9, L2 reg. rate: 0.1, Weight Init. Type: X1, Num. Epochs: 20
0 63.28476120629908 63.33995473902535 0.013164353046301773 0.00842803864564062 1.0 1572137715.329302
1 63.43100110517935 63.346601437163166 0.008960821576450156 0.007781986902769213 1.0 1572139685.9906337
2 63.43098663450773 63.36367590358398 0.008784625646576135 0.00816374475082959 1.0 1572141674.308169
3 63.43455411682646 63.37237586064055 0.008734283952326416 0.0091328223651367 1.0 1572143656.2895672
4 63.43708384329996 63.36023494239466 0.00850774632802675 0.007840718879393886 1.0 1572145648.7144737
5 63.432357633586406 63.368686491429436 0.008117598197767346 0.00816374475082959 1.0 1572147619.1223493
6 63.437478380275095 63.366225612932375 0.007916231420768466 0.007840718879393886 1.0 1572149591.5327795
7 63.4349176996381 63.359605795475815 0.008683942258076695 0.008105012774204915 1.0 1572151598.726169
8 63.437482984642955 63.373944582868816 0.008897894458638006 0.00842803864564062 1.0 1572153580.6343298
9 63.43547885849712 63.37092399403168 0.008797211070138566 0.007928816844330895 1.0 1572155602.2044847
10 63.43424568058998 63.37873350902092 0.008293794127641366 0.008721698528763986 1.0 1572157609.1552577
11 63.434756710801345 63.35837498291155 0.008558088022452396 0.00816374475082959 1.0 1572159614.7995634
12 63.43290854370355 63.3586274198897 0.008457404633952956 0.008398672657328283 1.0 1572161590.669562
13 63.43777654013375 63.350477184537965 0.008306379551203795 0.008809796493700995 1.0 1572163605.5718033
14 63.437025341899 63.3552375309895 0.008419648363265666 0.008809796493700995 1.0 1572165609.427092
15 63.435186951147564 63.35763221595337 0.008532917175327536 0.008105012774204915 1.0 1572167619.8778334
16 63.43932522153243 63.374169967256464 0.008394477516140806 0.008105012774204915 1.0 1572169603.518981
17 63.4367627921906 63.36626042755542 0.009061504964949596 0.008134378762517252 1.0 1572171604.6569538
18 63.43645362775697 63.370432045594114 0.008872723611513145 0.008075646785892579 1.0 1572173615.2680094
19 63.43212012066944 63.37386319688442 0.008772040223013705 0.008721698528763986 1.0 1572175607.9251254
Saved model to LX1-H-500-R-0_0001-M-0_9-E-20-data.npz.
```

- For L2 regularizer 0.01:



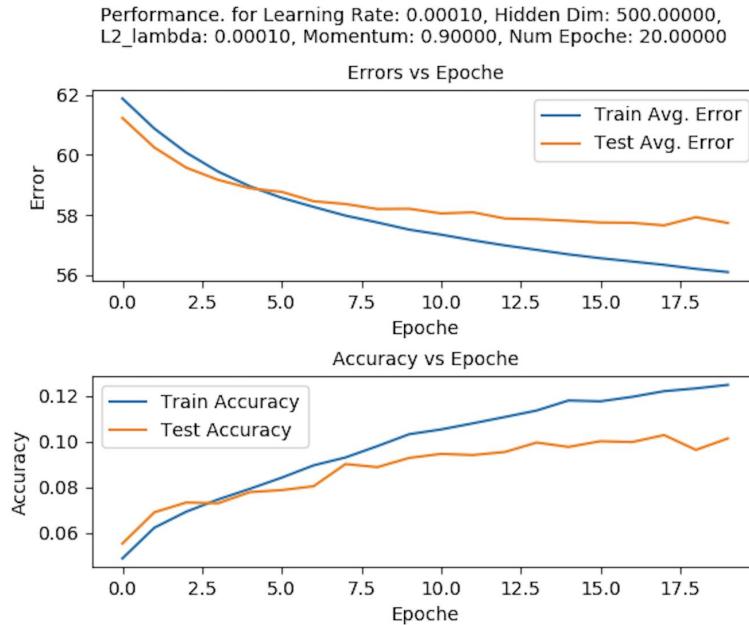
- For L2 regularizer 0.001:

Performance. for Learning Rate: 0.00010, Hidden Dim: 500.00000,
L2_lambda: 0.00100, Momentum: 0.90000, Num Epoche: 20.00000



```
RNN with learning rate: 0.0001, momentum: 0.9, L2 reg. rate: 0.001, Weight Init. Type: X1, Num. Epoche: 20
0 61.893630997698104 61.27158120879155 0.048177001396982014 0.05981851819222976 1.0 1572217674.7765021
1 61.1602357730953 60.987367757149336 0.05951646802673144 0.05702874930255778 1.0 1572219647.3361042
2 60.7182910029491 60.28128429197423 0.06398429339139408 0.0645758082988283 1.0 1572221669.02406
3 60.26235719882747 59.909171939534616 0.06918207332267767 0.07200540334184947 1.0 1572223701.0799544
4 59.86249380714823 59.839014721643764 0.07299545666209396 0.06428214841570493 1.0 1572225700.7833636
5 59.551726102898954 59.61285924868415 0.07713606101413344 0.07438404839514874 1.0 1572227736.889731
6 59.295627416409395 59.23118374620388 0.0797789996224373 0.07091886177429302 1.0 1572229757.5837061
7 59.0529680504055 59.17793829400396 0.08122632367192317 0.07649839955363698 1.0 1572231744.9257083
8 58.88792503231985 59.02450163024476 0.08419648363265665 0.08046280797580242 1.0 1572233786.0415905
9 58.75145142238078 59.040977042729736 0.08491385277571517 0.07914133850174727 1.0 1572235768.6667922
10 58.611595562942036 58.788692062081 0.08809796493700996 0.08383989663172114 1.0 1572237781.31913
11 58.51112071311184 58.76201123965325 0.08931875102256566 0.08425102046809385 1.0 1572239805.7632384
12 58.40354801305183 58.775138948120926 0.09115622286268045 0.08366370070184712 1.0 1572241791.4687839
13 58.294519214534034 59.24564066605723 0.09228891098329914 0.07652776554194932 1.0 1572243763.3539512
14 58.21969953461343 58.72863876423275 0.09320764690335653 0.08342877279534842 1.0 1572245767.1005306
15 58.14323004249963 58.860874283721124 0.09541009602678178 0.08477960825771591 1.0 1572247729.8192275
16 58.02750384748802 58.5800944193055 0.09569956076871768 0.0873344492408892 1.0 1572249754.814163
17 57.93606992932712 58.425530597338756 0.09843059768176499 0.09144568760461633 1.0 1572251733.077461
18 57.87008570324185 58.562476126056225 0.09996601935638144 0.0874225472058262 1.0 1572253750.0404885
19 57.776601786321606 58.36612900648306 0.09888367293001246 0.09203300737086306 1.0 1572255751.2414694
Saved model to LX1-H-500-R-0_0001-M-0_9-E-20-data.npz
```

- For L2 regularizer 0.0001:



```
RNN with learning rate: 0.0001, momentum: 0.9, L2 reg. rate: 0.0001, Weight Init. Type: X1, Num. Epochs: 20
0 61.878280739806456 61.22905881506498 0.048768516304416226 0.05529615599212991 1.0 1572257736.3026621
1 60.87967315386413 60.24845214804529 0.06222233409265389 0.0689807065456788 1.0 1572259763.7939742
2 60.07944205776427 59.57644030850247 0.0692324150169274 0.07326814083927995 1.0 1572261774.4328132
3 59.44677527305313 59.17121605201601 0.07461897630164743 0.07294511496784424 1.0 1572263806.1840177
4 58.95380480021372 58.88981409653359 0.0792816844330897 0.07784923501600446 1.0 1572265813.5889986
5 58.570043872609 58.76805912403653 0.08419648363265665 0.07873021466537457 1.0 1572267812.7227776
6 58.267886931870976 58.45723744440873 0.08958304491737669 0.08043344198749008 1.0 1572269796.2344644
7 57.978260780031 58.3646365034116 0.09304403639704495 0.09015358411887352 1.0 1572271828.5294795
8 57.75073742415655 58.198622108451566 0.09801527870420479 0.08880274865650603 1.0 1572273828.4084167
9 57.51203722836177 58.20673987519276 0.10225081490617567 0.09285525504360849 1.0 1572275836.6834888
10 57.344789721231045 58.05154959615587 0.10533999521753905 0.0946172143423487 1.0 1572277846.9863088
11 57.15684152591936 58.08831301519963 0.10802069043633664 0.09411799254103898 1.0 1572279867.8834386
12 56.98631351894655 57.88303358145964 0.11080206904363367 0.09543946201509412 1.0 1572281866.7945662
13 56.83774202604316 57.85722830100446 0.11363378934518041 0.09958006636713358 1.0 1572283909.9633968
14 56.68807622781035 57.80797329498641 0.11803868759203091 0.0976712771268317 1.0 1572285917.5362399
15 56.55942520727483 57.74854971872088 0.11767371030872044 0.10016738613338032 1.0 1572287919.5498219
16 56.44976582979693 57.73920386745836 0.11963703638445952 0.09981499427363227 1.0 1572289932.1798563
17 56.337540487141595 57.65280523647097 0.12212895024982066 0.10286905705811529 1.0 1572291961.8335776
18 56.20439442756442 57.927332466475384 0.12334973633537637 0.09637917364108889 1.0 1572293923.6226928
19 56.09778262062949 57.73311459900519 0.12489774343355525 0.10137139165418613 1.0 1572295947.3826632
Saved model to LX1-H-500-R-0_0001-M-0_9-E-20-data.npz
```