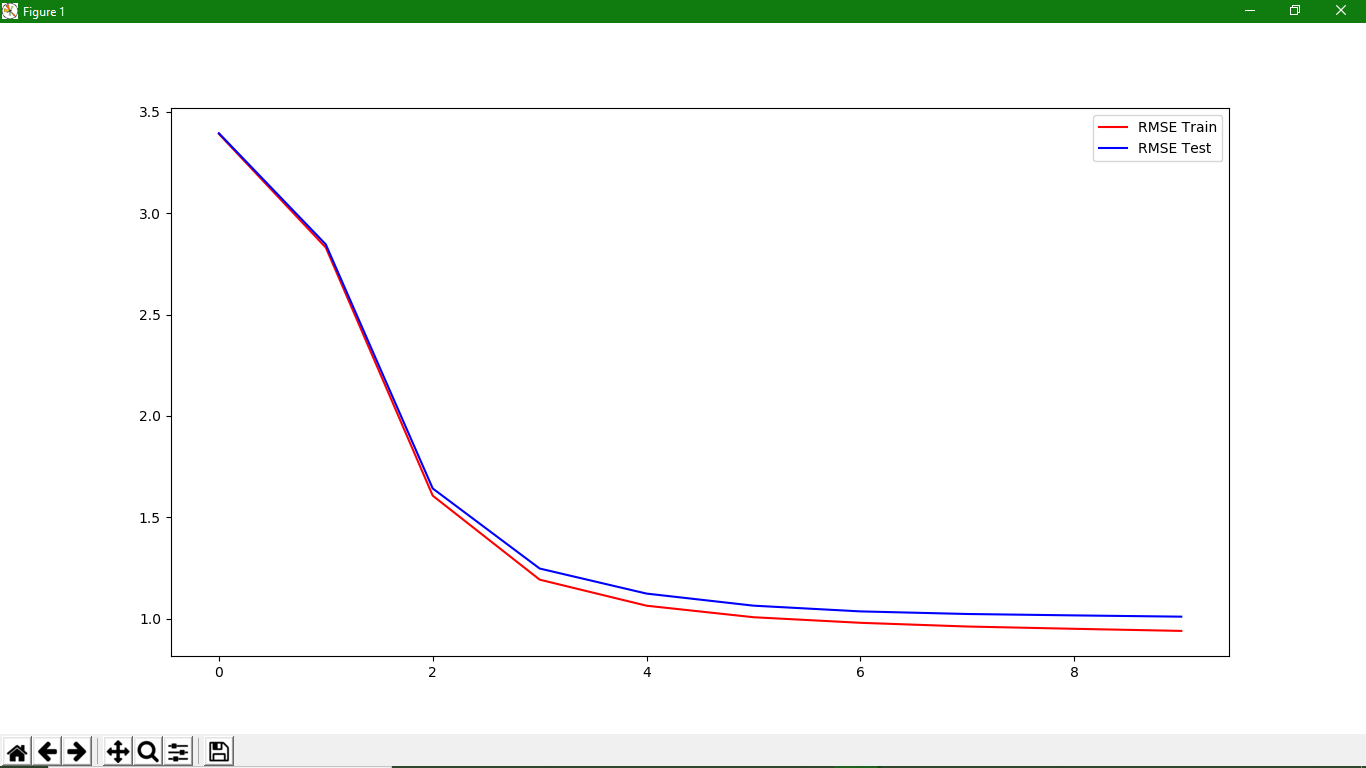
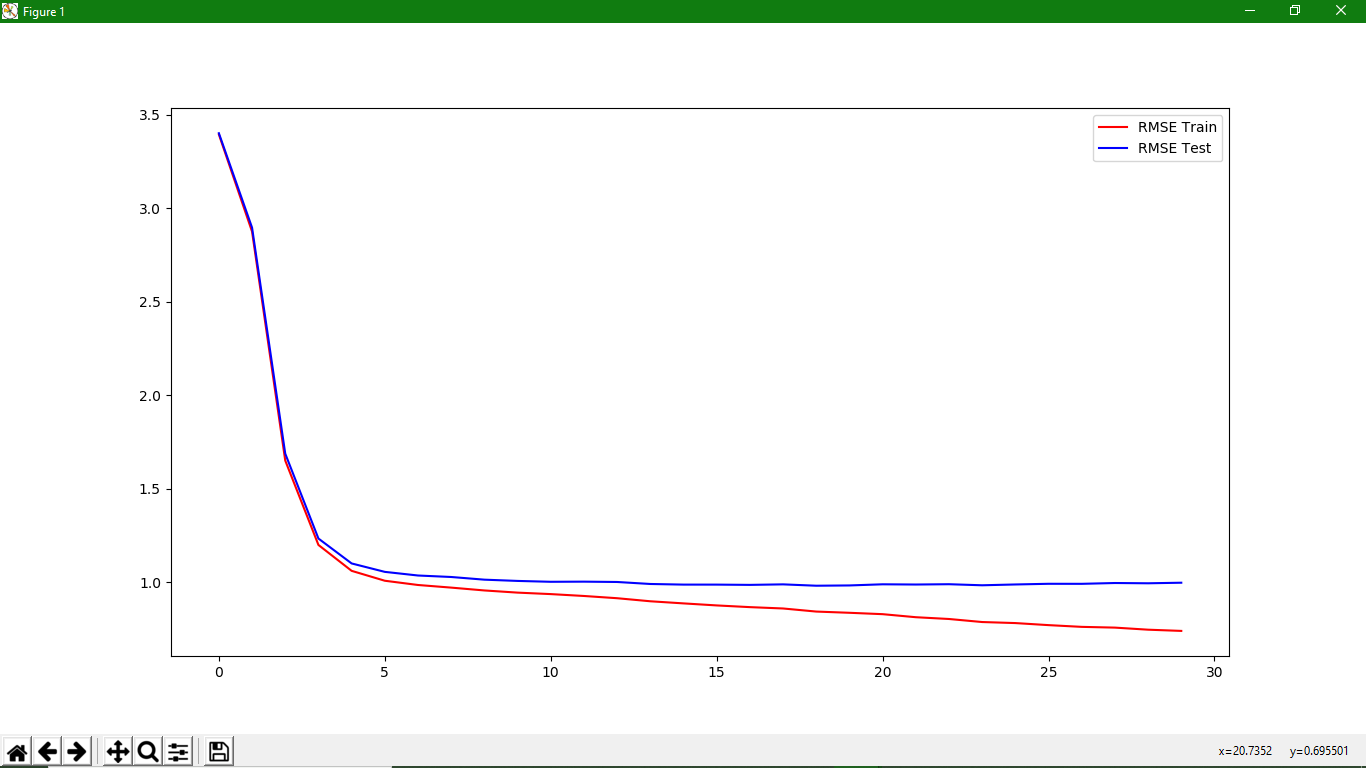
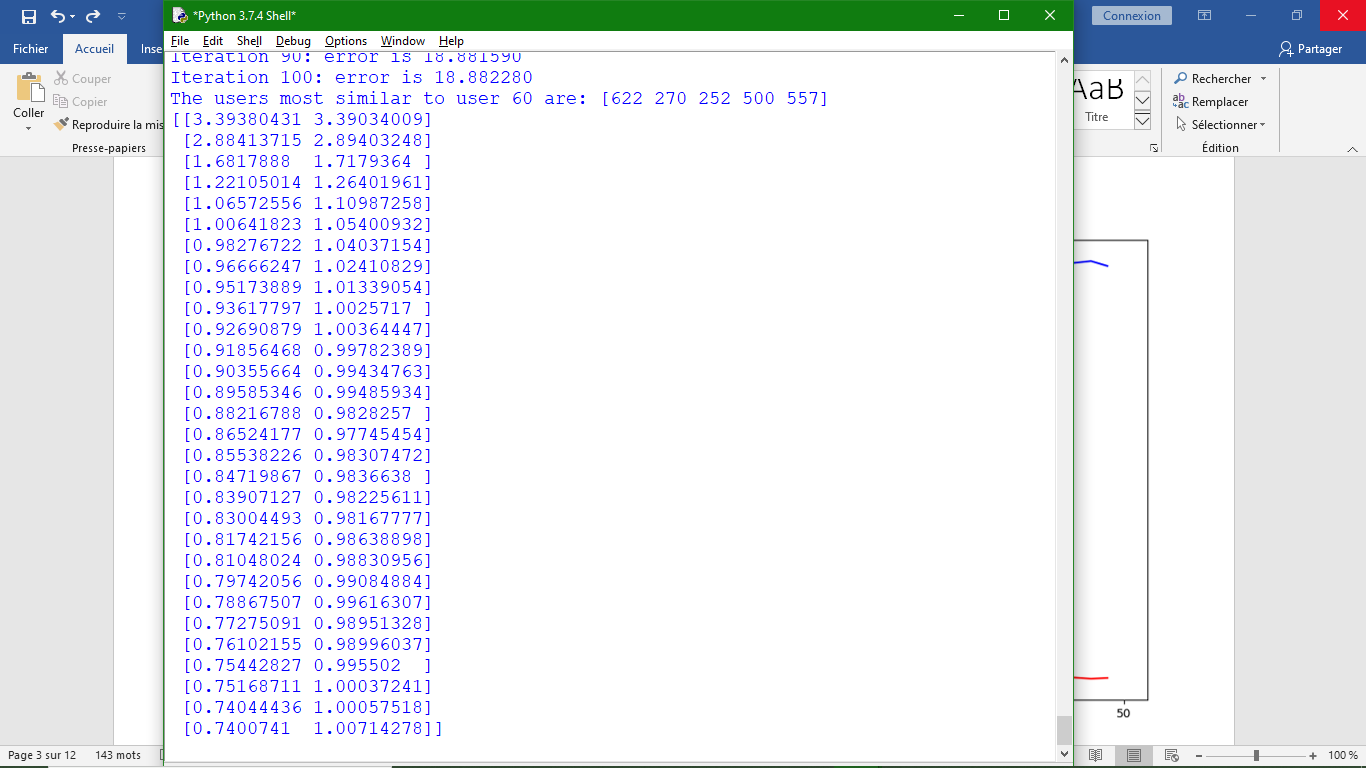
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k² | 1e-2 (default) | 10 x 1 Step | 80/20 |  |

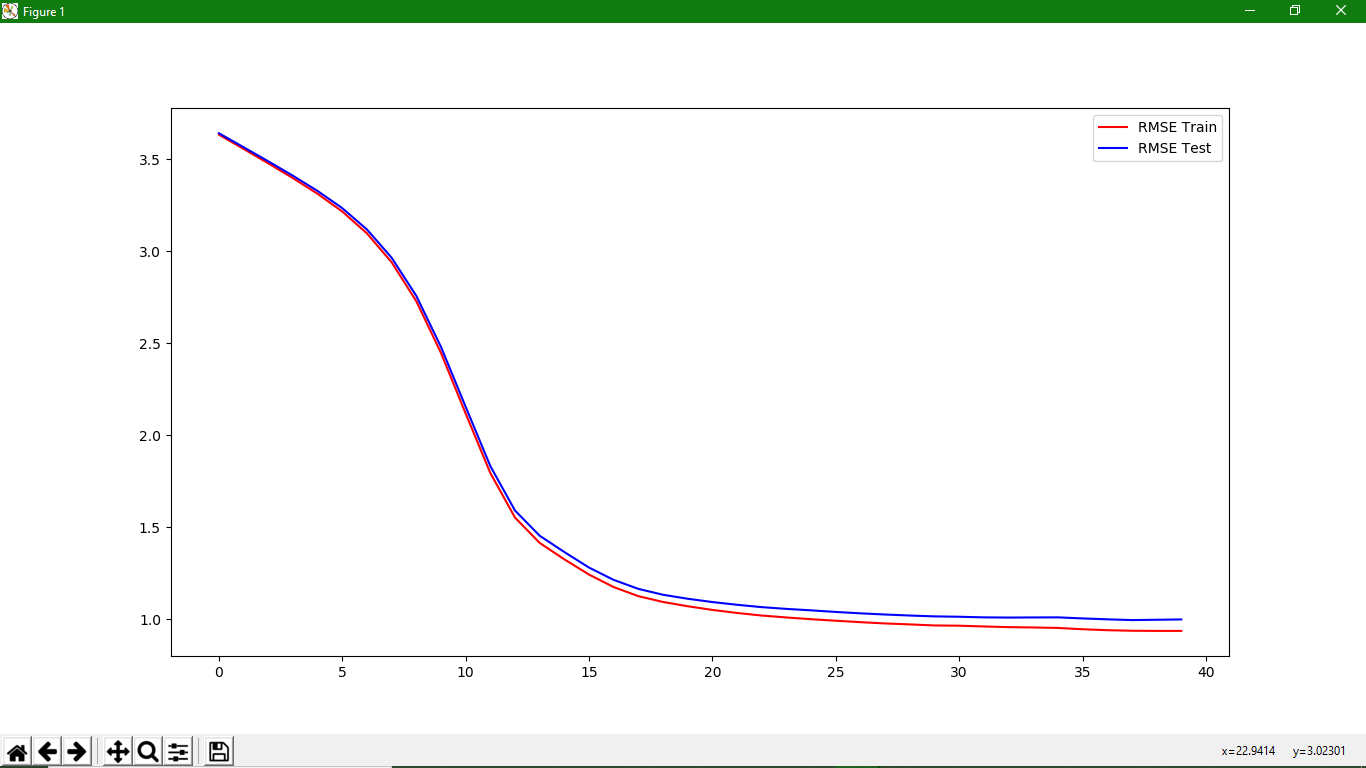


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 10 x 3 Step | 80/20 |  |

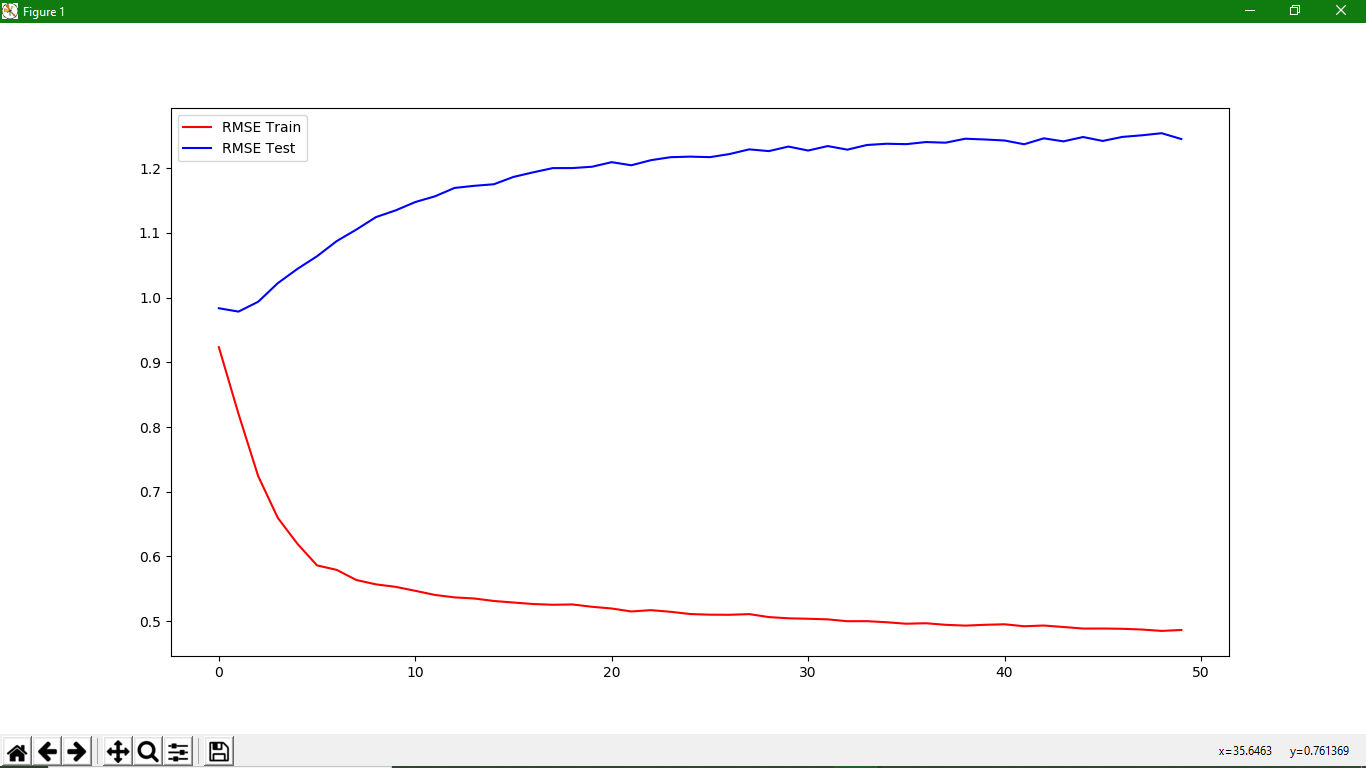




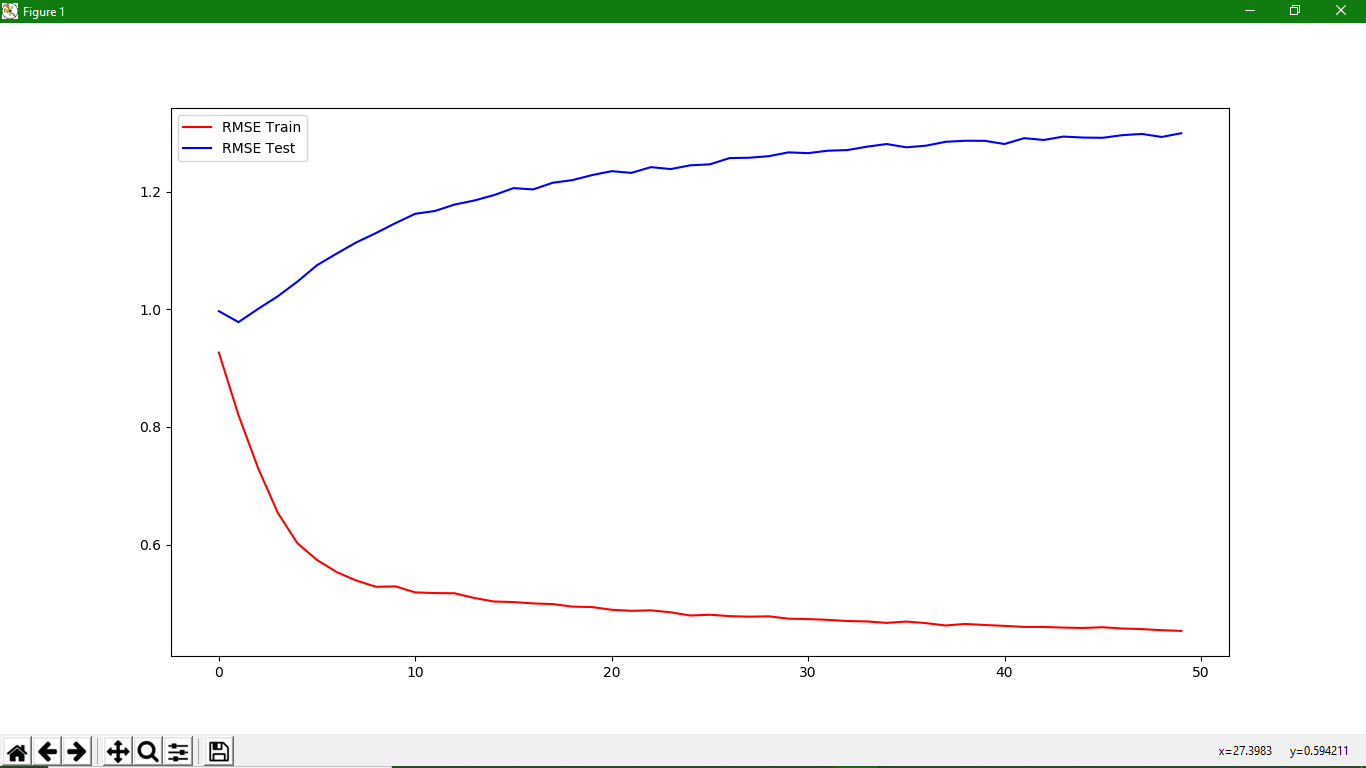
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 40 x 1 Step | 80/20 |  |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 90/10 |  |

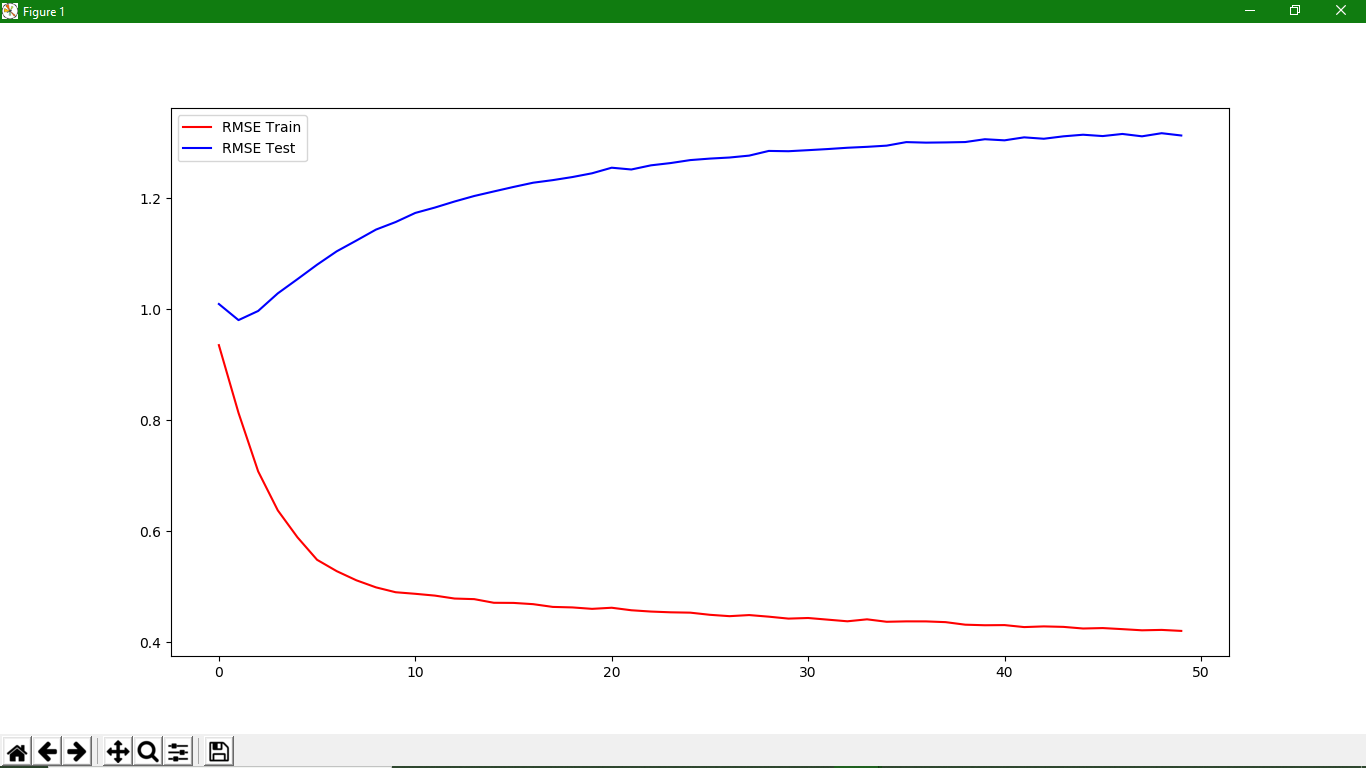


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 80/20 |  |

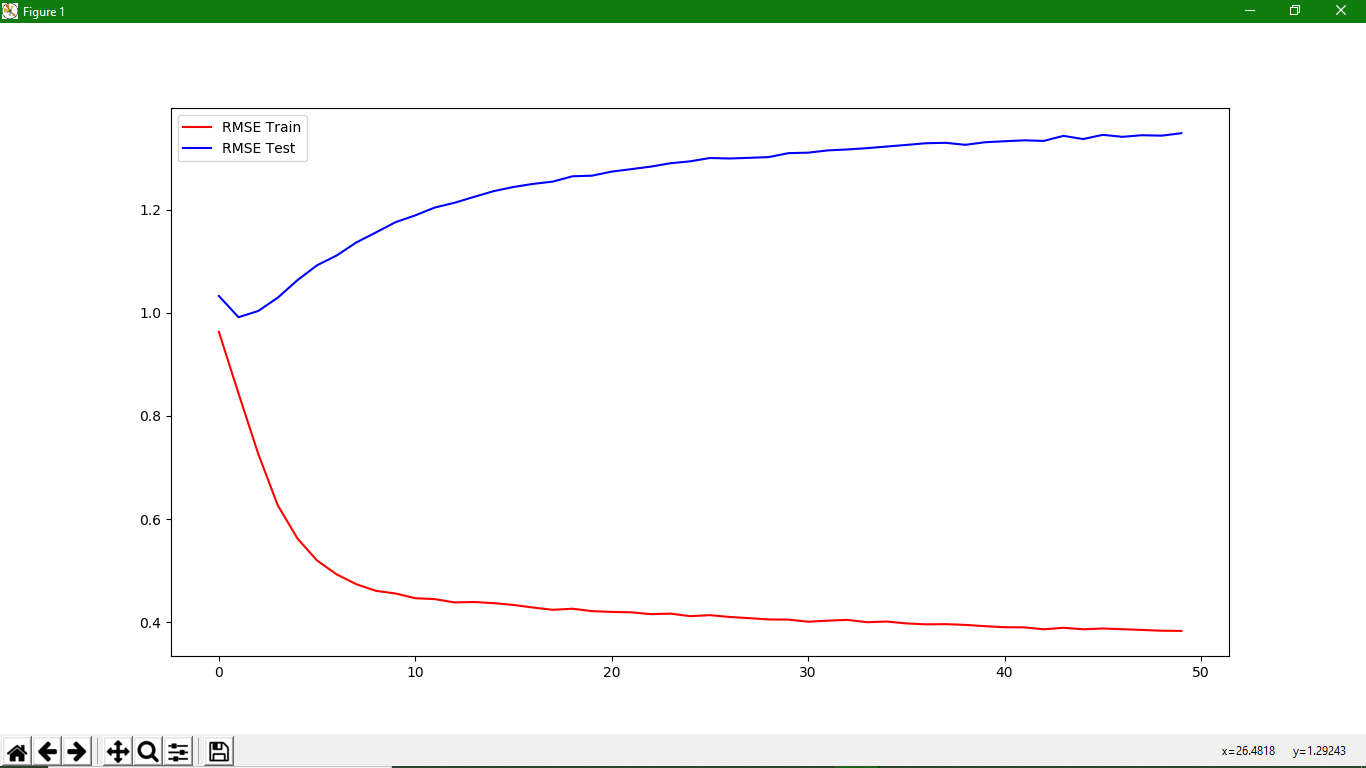


|  |  |  |
| --- | --- | --- |
| **Iteration** | **Train RMSE** | **Test RMSE** |
| 1 | 0.92048186 | 0.99910051 |
| 2 | 0.81162608 | 0.98670751 |
| 3 | 0.71692789 | 1.00611413 |
| 4 | 0.65404791 | 1.0309608 |
| … | … | … |
| 48 | 0.46033046 | 1.27686942 |
| 49 | 0.45619506 | 1.27536404 |
| 50 | 0.45576793 | 1.27951813 |

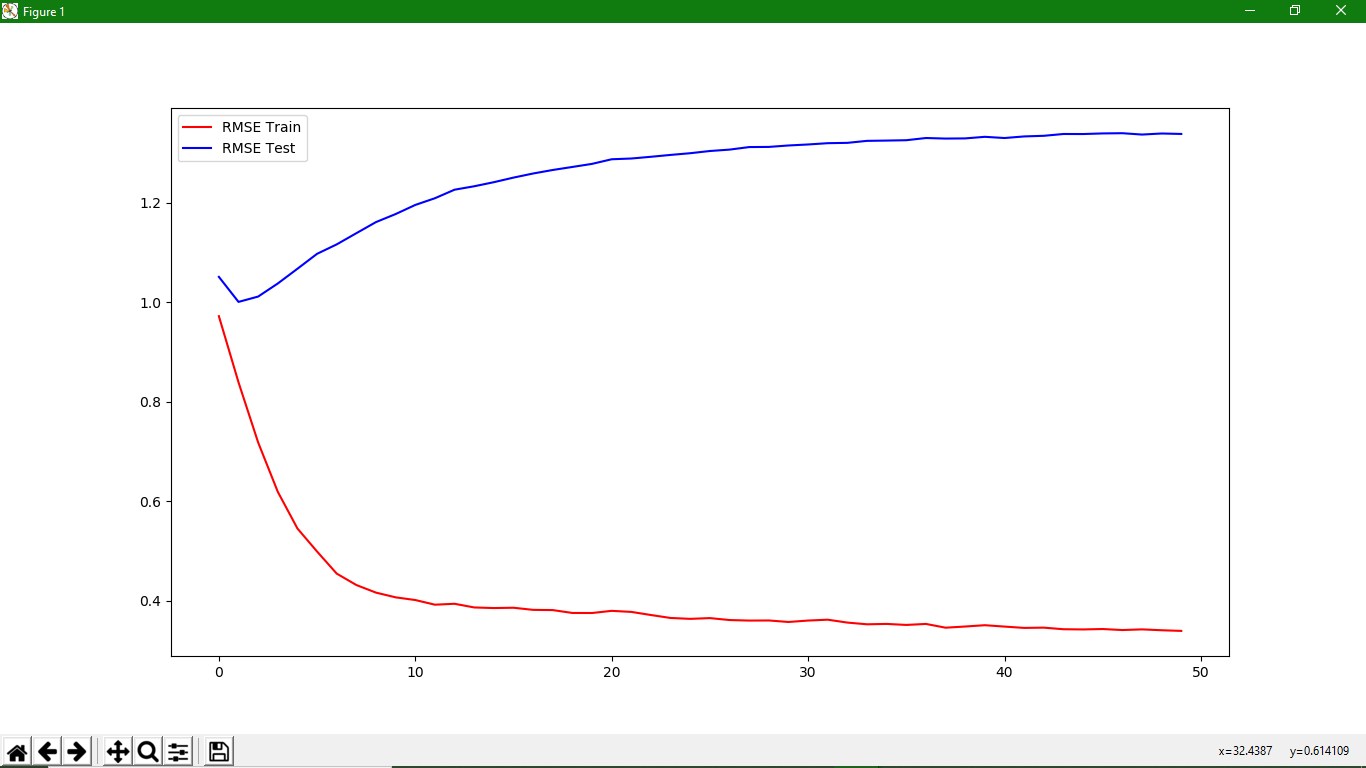
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 70/30 |  |



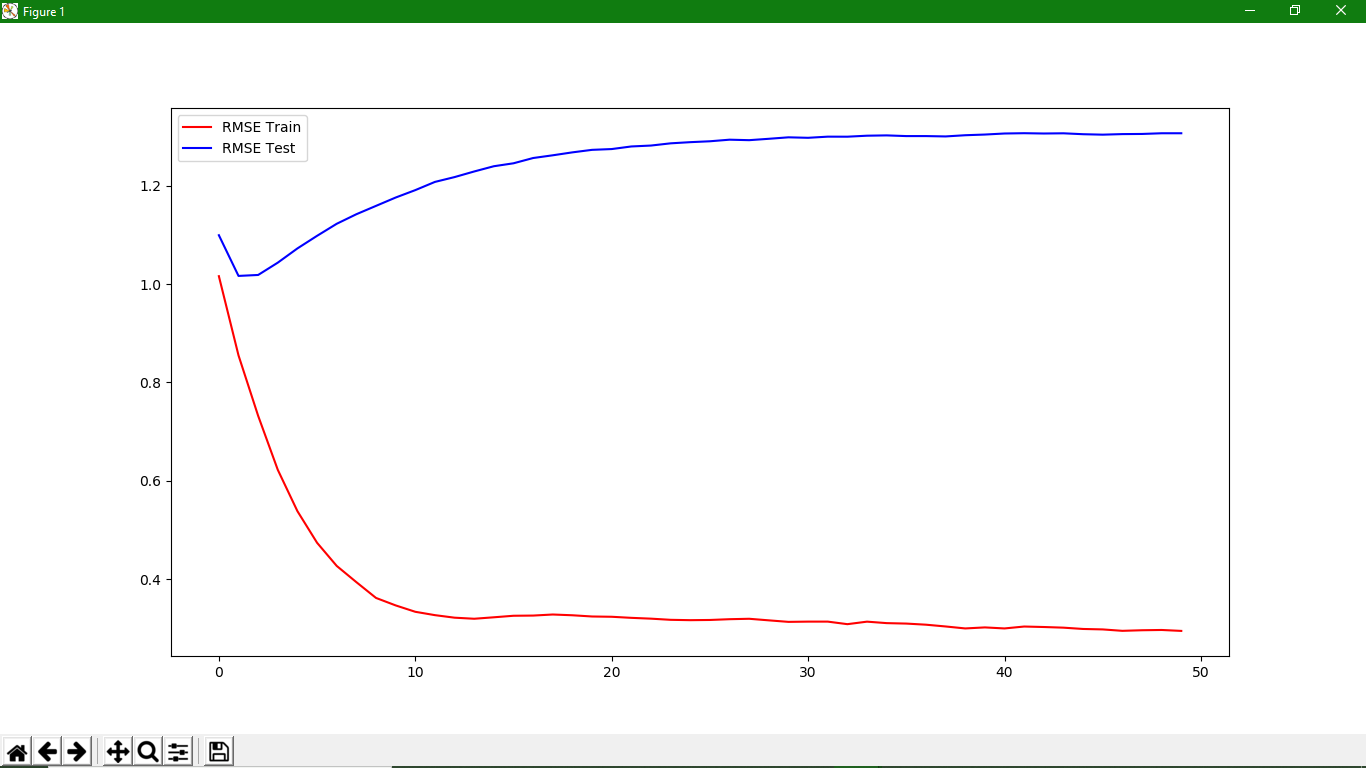
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 60/40 |  |



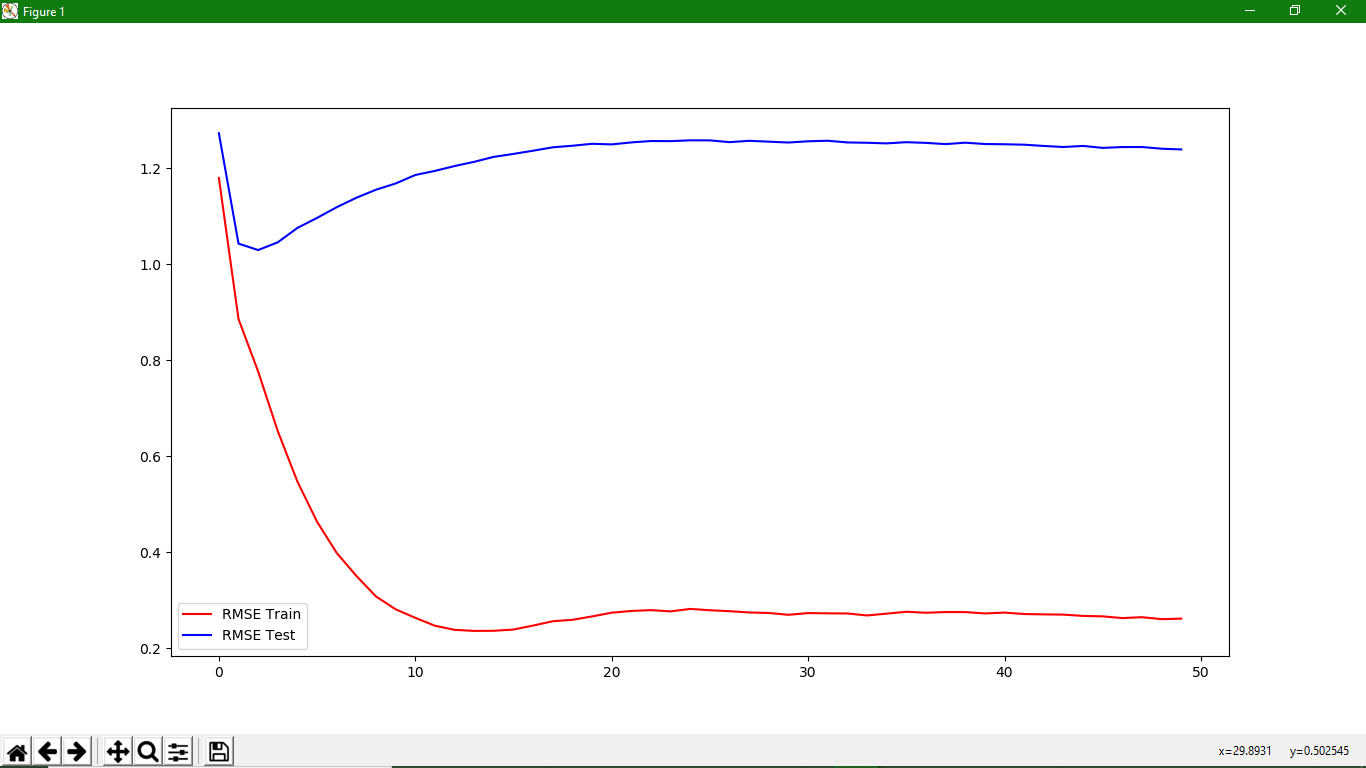
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 50/50 |  |



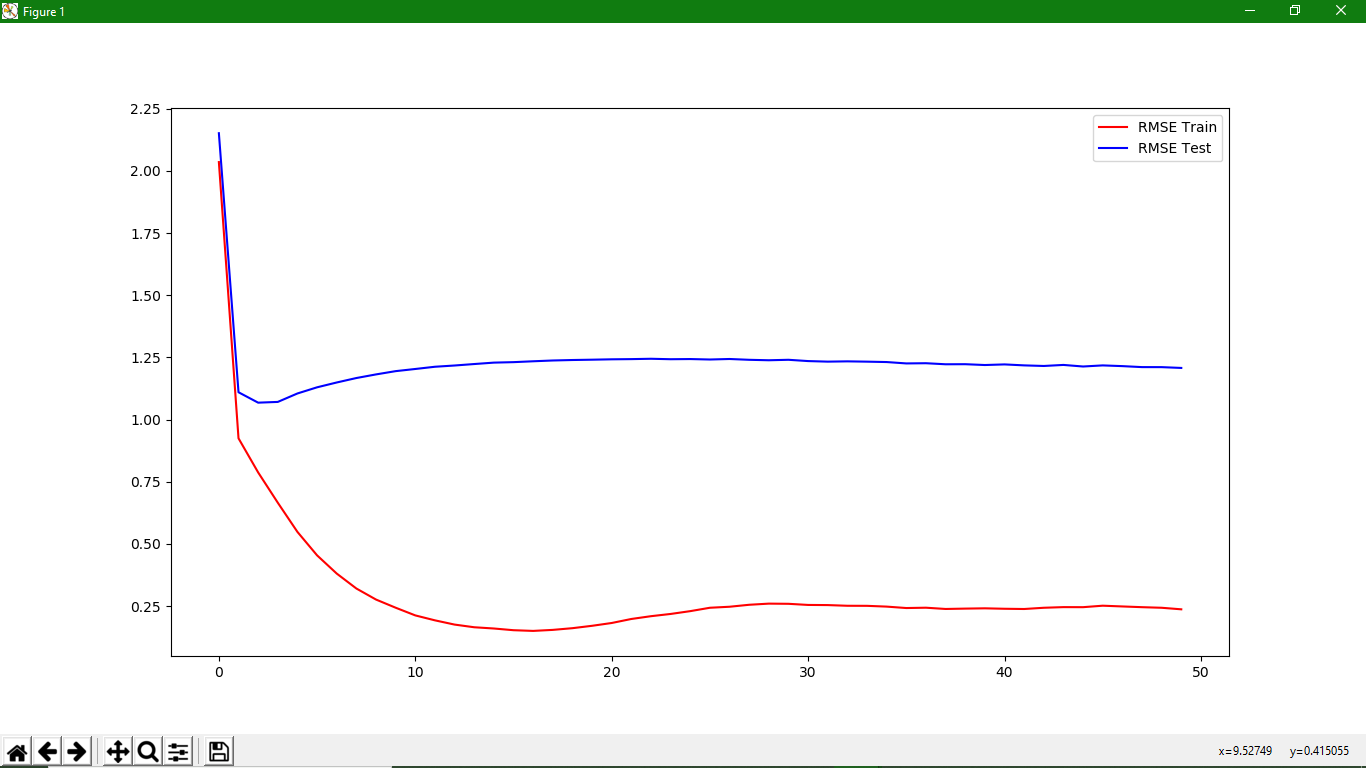
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 40/60 |  |



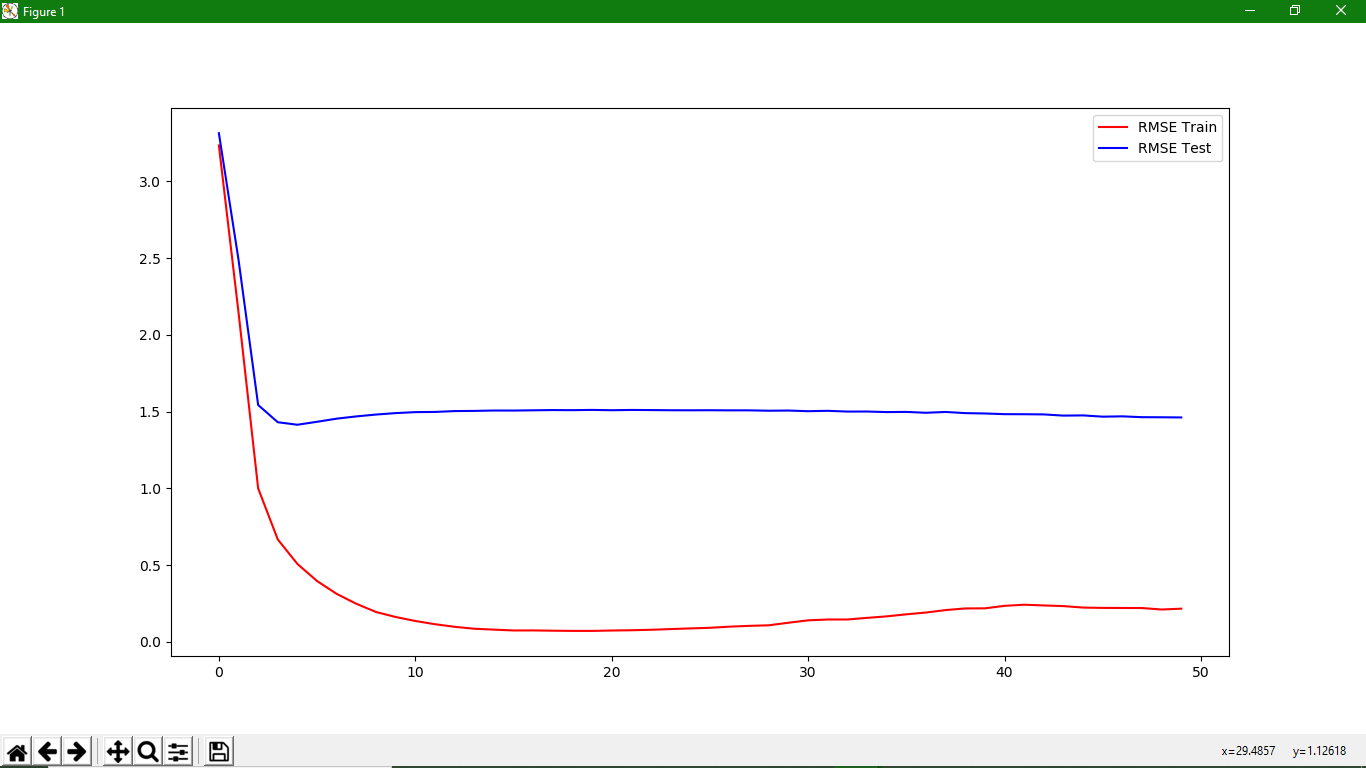
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 30/70 |  |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 20/80 |  |

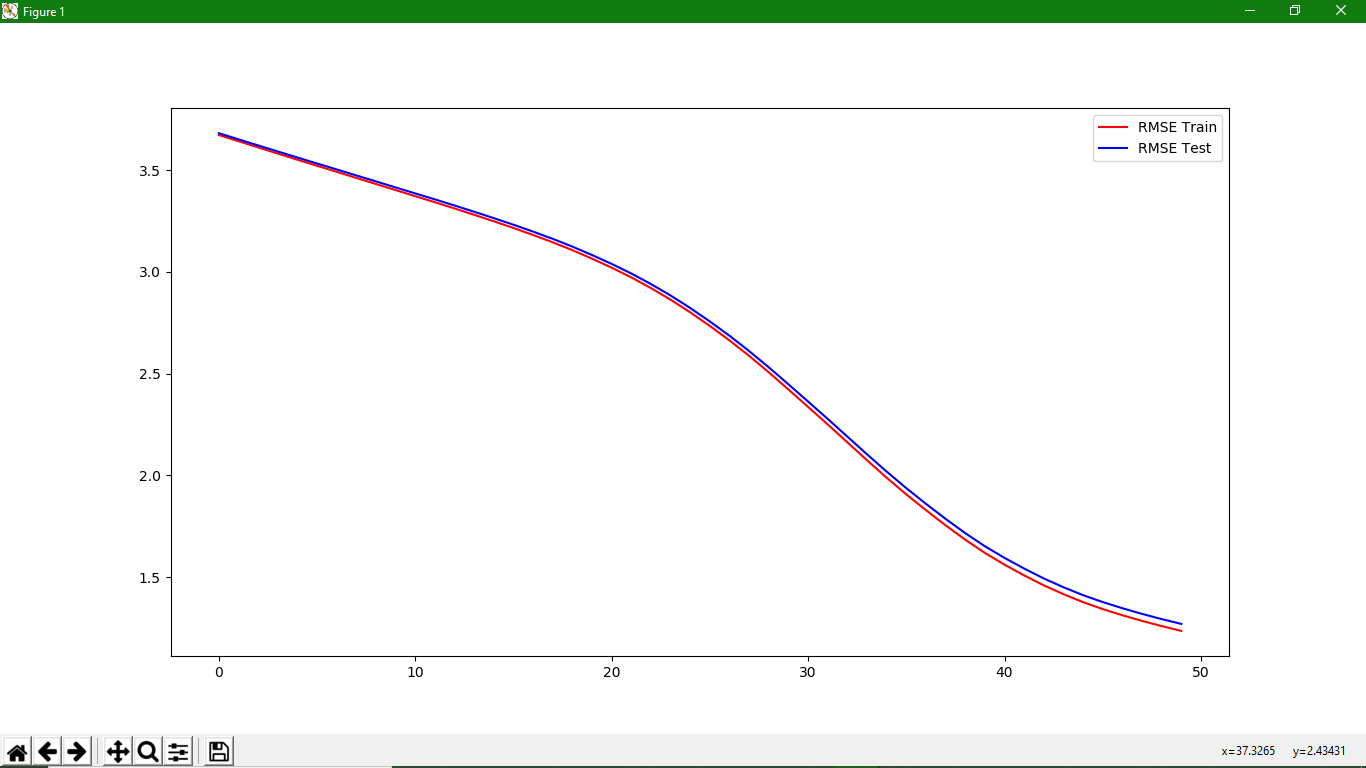


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-2 (default) | 50 times entire model | 10/90 |  |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-3 | 10 x 5 Step | 80/20 |  |

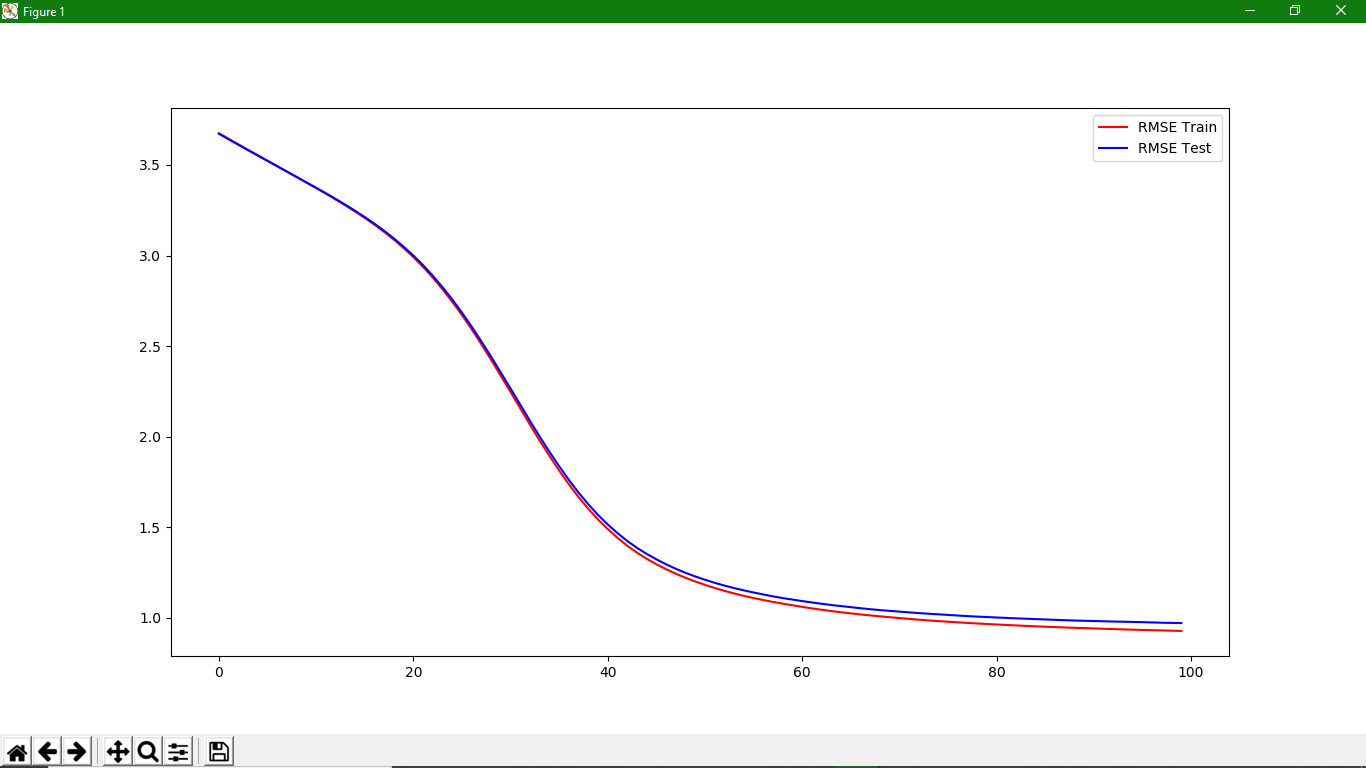
Also visualisations of neighbours have sudden shape changes (see Image Saves). (no convergence).



|  |  |  |  |
| --- | --- | --- | --- |
| [[3.67280149 3.68166804]  [3.64223552 3.6515553 ]  [3.61179423 3.62159419]  [3.58152413 3.59187222]  [3.551265 3.56204343]  [3.52143693 3.53271723]  [3.4917047 3.50345445]  [3.46192527 3.47410774]  [3.43191552 3.44457817]  [3.40197635 3.41504097]  [3.37221026 3.38571644]  [3.34233236 3.35632205]  [3.3118403 3.3262825 ]  [3.28065872 3.29563427]  [3.24872494 3.26414633]  [3.21585798 3.23181176] | [3.1815443 3.19798899]  [3.14536357 3.16233397]  [3.10674214 3.12425995]  [3.06498337 3.08293819]  [3.02072477 3.03922534]  [2.97290349 2.99204874]  [2.9204073 2.9401505 ]  [2.86337948 2.88399291]  [2.80111337 2.82240677]  [2.73415279 2.75634122]  [2.66309357 2.68601274] | [2.58708596 2.61075759]  [2.50623655 2.53080225]  [2.42201638 2.44737935]  [2.3364594 2.36274385]  [2.25011659 2.27730536]  [2.16240764 2.19045305]  [2.07430458 2.10345626]  [1.98849702 2.01834798]  [1.90659428 1.93718147]  [1.82902718 1.86011946]  [1.75449395 1.78625619] | [1.6839999 1.71612489]  [1.6189425 1.65158987]  [1.56107962 1.59437311]  [1.50832999 1.542032 ]  [1.4594928 1.49331832]  [1.41602123 1.4502852 ]  [1.37692976 1.41114438]  [1.34294343 1.37731171]  [1.31245685 1.34696436]  [1.28481686 1.31927264]  [1.25916493 1.29357958]  [1.2355516 1.27004468]] |

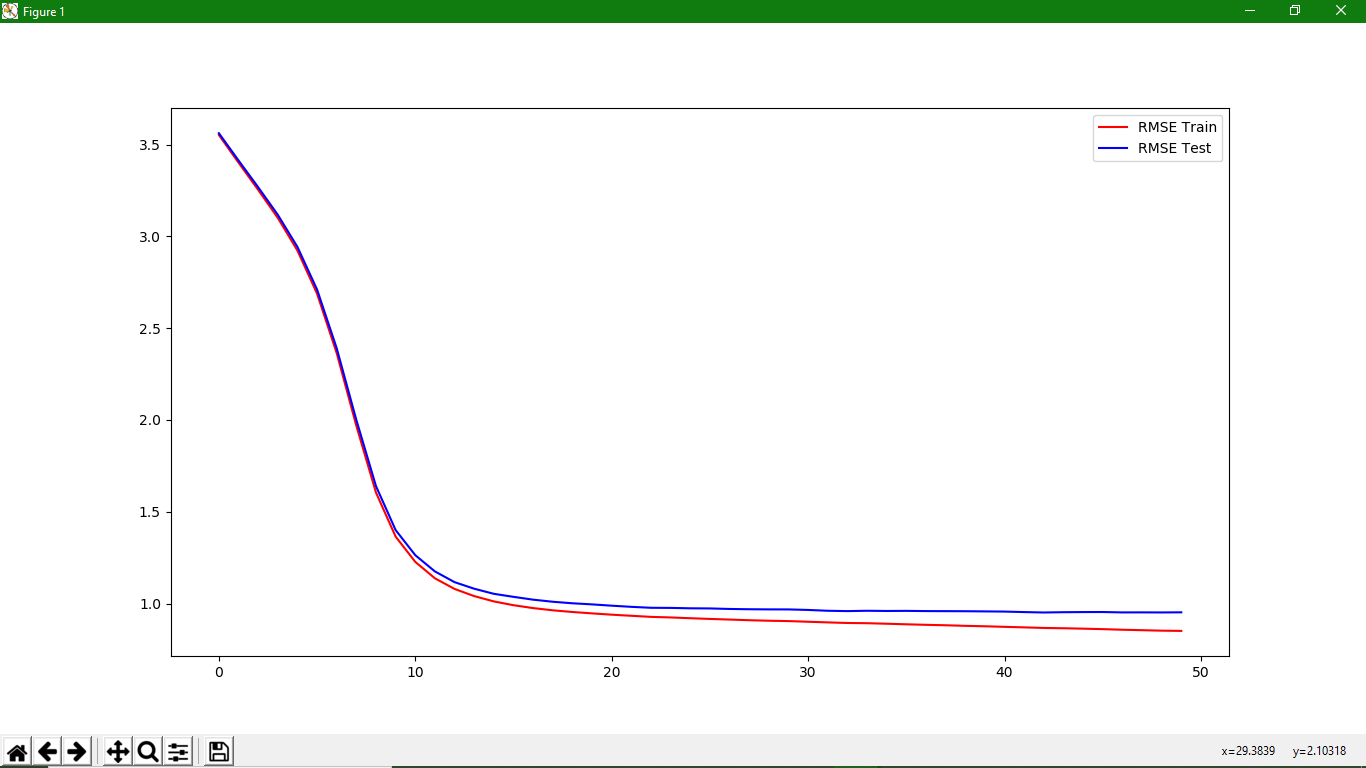
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 1e-3 | 10 x 10 Step | 80/20 |  |

(no convergence): Visualisations vary quite a lot (line and bat shapes).



|  |  |  |  |
| --- | --- | --- | --- |
| [[3.67530084 3.67222428]  [3.64444876 3.64178705]  [3.61380911 3.61161494]  [3.58338451 3.58174014]  [3.55324316 3.55214667]  [3.52328873 3.52269506]  [3.49323344 3.49314594]  [3.46304417 3.46347809]  [3.43290997 3.43383527]  [3.4026227 3.40408134]  [3.37244391 3.37446761]  [3.34152579 3.34404707]  [3.3098309 3.31289554]  [3.27710843 3.28075671]  [3.24319673 3.24749899]  [3.20787454 3.21273589]  [3.17057562 3.17606115]  [3.13052225 3.13671994]  [3.0877409 3.09463406]  [3.04135633 3.04899287]  [2.99194026 3.00043416] | [2.93759537 2.94685698]  [2.87850332 2.88858414]  [2.81435251 2.82542872]  [2.74564385 2.75798011]  [2.67205977 2.68537569]  [2.5935173 2.60793829]  [2.51065135 2.52612805]  [2.42485189 2.44161725]  [2.33623028 2.35421658]  [2.24719071 2.2662642 ]  [2.15694523 2.17699862]  [2.0665915 2.08739829]  [1.9785105 2.00050545]  [1.89398909 1.91722691]  [1.81345093 1.83722711]  [1.73754776 1.76176548]  [1.6667639 1.69143355]  [1.6021204 1.62714195]  [1.54297757 1.56832719]  [1.49014187 1.51573443]  [1.4422406 1.46803892]  [1.39873362 1.42477155]  [1.36038387 1.38688993]  [1.32637632 1.35340655]  [1.2961427 1.32323503] | [1.26877356 1.29587865]  [1.24375582 1.27082467]  [1.22123384 1.2485764 ]  [1.20082629 1.22849739]  [1.18241334 1.21049213]  [1.16494787 1.19352996]  [1.14912629 1.1782757 ]  [1.13473177 1.1643672 ]  [1.12155855 1.15189886]  [1.10937393 1.14007056]  [1.09846485 1.12923849]  [1.08787334 1.11862254]  [1.07836235 1.1092658 ]  [1.06947684 1.100968 ]  [1.06078589 1.09261394]  [1.05251372 1.08506644]  [1.04481661 1.07783043]  [1.0374738 1.07102275]  [1.03082132 1.06484735]  [1.02470291 1.05911434]  [1.01902282 1.05314362]  [1.01356673 1.04779792]  [1.00833738 1.04285228]  [1.00358844 1.03869593]  [0.9989779 1.03443921]  [0.9943735 1.03019238]  [0.99004191 1.02639008] | [0.98586708 1.02276599]  [0.98215318 1.01953769]  [0.97859782 1.01633894]  [0.97531748 1.01287484]  [0.97206593 1.00950277]  [0.96906972 1.00693059]  [0.96636391 1.00466359]  [0.96351987 1.00197756]  [0.96090358 0.99971777]  [0.95822597 0.99742407]  [0.95574945 0.99540025]  [0.9533776 0.99323344]  [0.95113081 0.99135268]  [0.94909209 0.98927999]  [0.94709134 0.98721439]  [0.94510388 0.98535484]  [0.94332016 0.98395103]  [0.94155169 0.98247755]  [0.93979764 0.98112422]  [0.93814224 0.97981477]  [0.93643498 0.97872955]  [0.93485308 0.97750258]  [0.9333002 0.97624731]  [0.93196189 0.97470987]  [0.93059087 0.97316104]  [0.9292475 0.97214186]  [0.92801327 0.9713468 ]] |

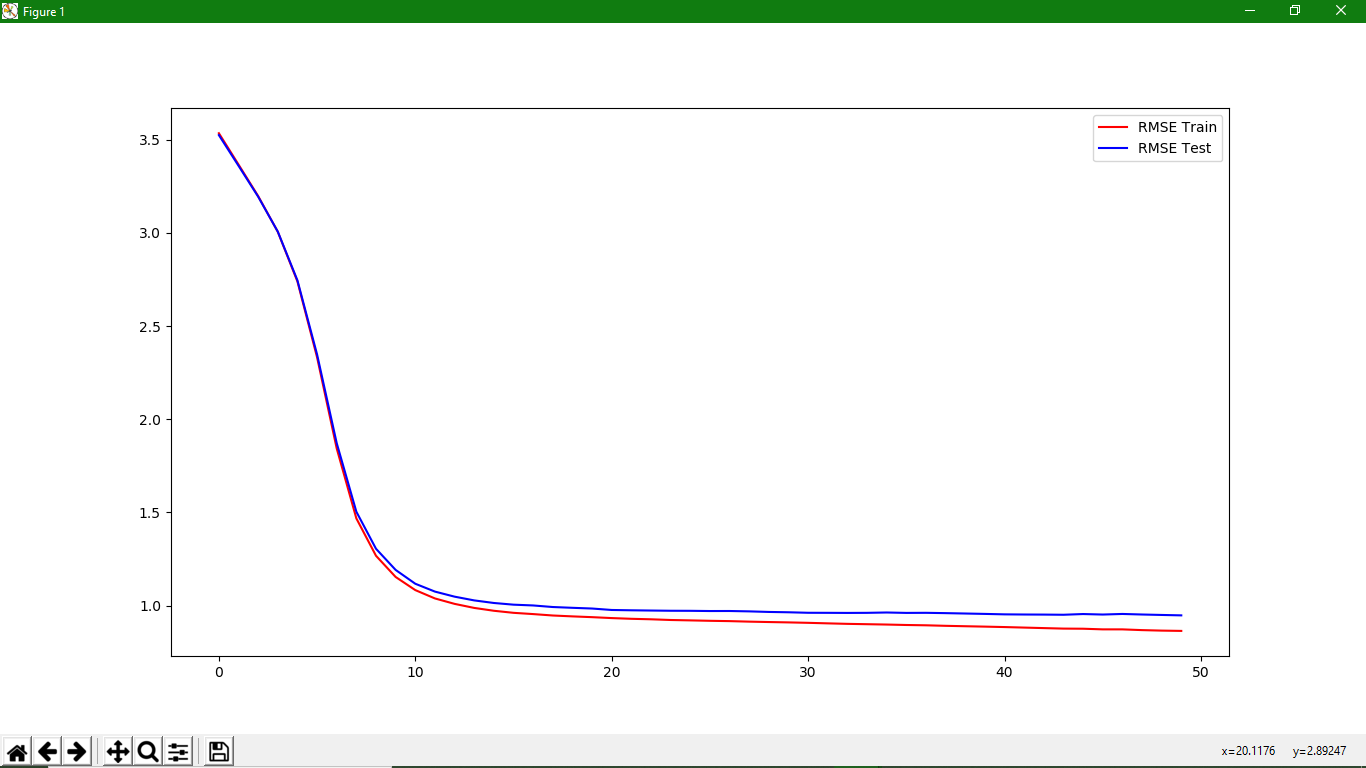
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 5e-3 | 10 x 5 Step | 80/20 |  |

(converged more or less). Visualisations have some odd variations (line shapes).

|  |  |  |  |
| --- | --- | --- | --- |
| [[3.55220413 3.56165338]  [3.40142965 3.4134202 ]  [3.25301361 3.26765585]  [3.100317 3.11733437]  [2.92257953 2.94240355]  [2.68733549 2.71015358]  [2.36345816 2.39019775]  [1.96795774 1.99974227]  [1.60423231 1.63947272]  [1.36455357 1.4001894 ]  [1.22698271 1.26348662]  [1.13752317 1.17498124]  [1.07957625 1.11661887] | [1.0406214 1.08082449]  [1.01171267 1.0530442 ]  [0.99103129 1.03658557]  [0.97536695 1.02118635]  [0.96324778 1.00985956]  [0.95407891 1.0015285 ]  [0.94648975 0.99554497]  [0.93947381 0.988428 ]  [0.9334262 0.98215675]  [0.9275136 0.97695613]  [0.92433482 0.97633922]  [0.92016882 0.9741354 ]  [0.9162631 0.97339487]  [0.9128679 0.9706018 ]  [0.90926224 0.96904188]  [0.90648603 0.96826744]  [0.90453678 0.96809208] | [0.9009161 0.96522284]  [0.89743775 0.96074742]  [0.89417142 0.95890927]  [0.89298797 0.96074426]  [0.89013922 0.95964688]  [0.88704509 0.96017855]  [0.88421047 0.9590112 ]  [0.88177884 0.95863074]  [0.87862176 0.95825535]  [0.87628835 0.95731753]  [0.87318385 0.95632559]  [0.87027937 0.95358497]  [0.8671726 0.95134377]  [0.86552918 0.95290506] | [0.86328673 0.95364803]  [0.8607114 0.95390242]  [0.85743845 0.95176101]  [0.8549391 0.95196176]  [0.85238934 0.95161027]  [0.85107285 0.952241 ]] |

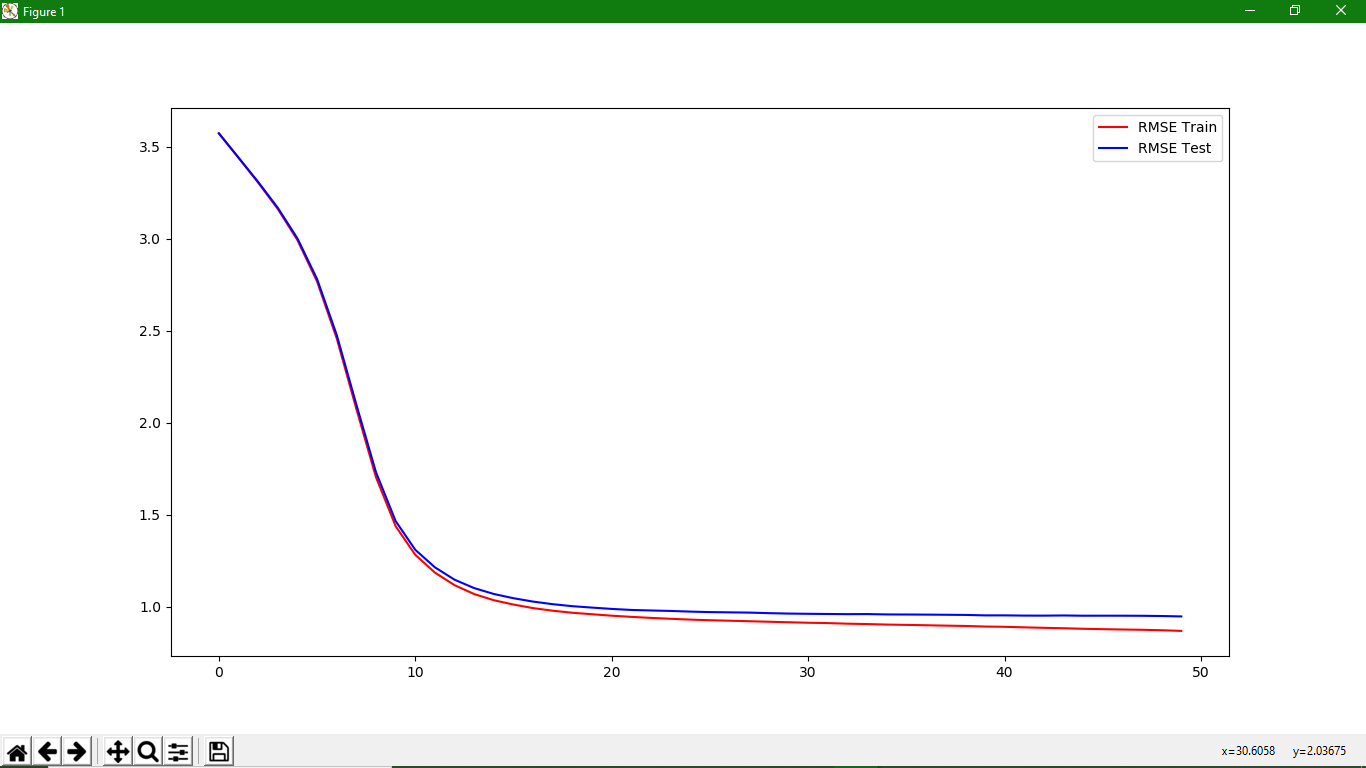
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 5e-3 | 10 x 5 Step | 90/10 |  |

(converged more or less): Visualisations have some odd variations (line shapes).



|  |  |  |  |
| --- | --- | --- | --- |
| [[3.5355885 3.52541542]  [3.36737752 3.36084819]  [3.1981039 3.19492936]  [3.00631547 3.00719643]  [2.73963284 2.74544907]  [2.33251452 2.34789038]  [1.84327996 1.87166464]  [1.46763492 1.50355506]  [1.26703179 1.30467057]  [1.15320945 1.1910522 ]  [1.08279133 1.11692107]  [1.03796804 1.07472384]  [1.00883961 1.04756725]  [0.98708153 1.0274533 ]  [0.9717952 1.01357603] | [0.96074384 1.00454473]  [0.95399708 1.0003233 ]  [0.94636869 0.99191272]  [0.94164127 0.98766094]  [0.93749279 0.98380625]  [0.93253219 0.97605109]  [0.92865032 0.97433227]  [0.9260335 0.9730345 ]  [0.92223388 0.97185171]  [0.92014134 0.97153306]  [0.91789359 0.97035688]  [0.91621453 0.97053868]  [0.91344047 0.96856421]  [0.91126597 0.96548092]  [0.90935558 0.96370077] | [0.90703005 0.96098334]  [0.90420038 0.96070999]  [0.90152168 0.96022642]  [0.8996762 0.96067041]  [0.8978098 0.96252817]  [0.89542162 0.96032405]  [0.89377815 0.96077585]  [0.89096928 0.95909858]  [0.88877624 0.95702559]  [0.88668072 0.95490968]  [0.88431531 0.95254129] | [0.88153309 0.9517076 ]  [0.87874717 0.95127141]  [0.87605977 0.95035827]  [0.8753863 0.95423216]  [0.87195837 0.95154917]  [0.87189734 0.95467186]  [0.86792159 0.951729 ]  [0.8652547 0.94933897]  [0.86368757 0.9470973 ]] |

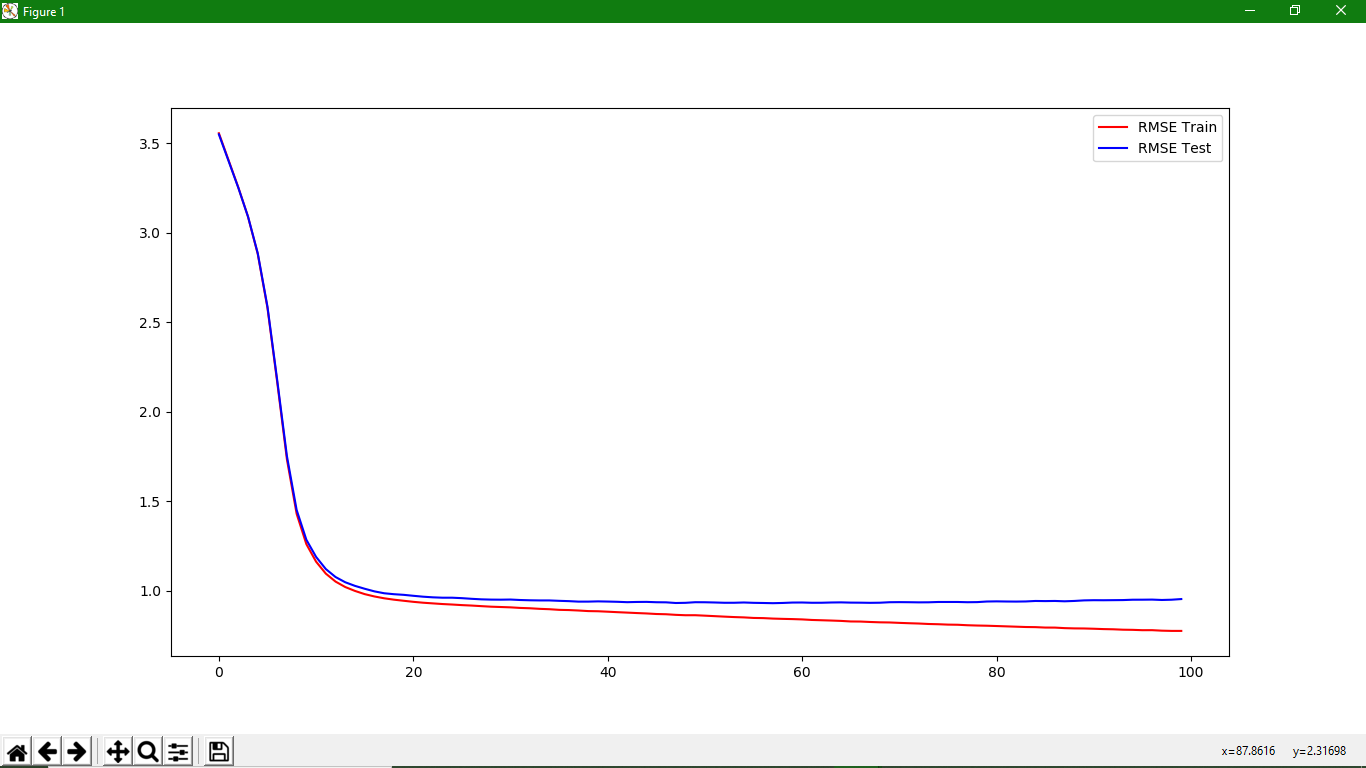
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 5e-3 | 10 x 5 Step | 70/30 |  |



|  |  |  |  |
| --- | --- | --- | --- |
| [[3.57361102 3.57224679]  [3.44000602 3.44106865]  [3.30507898 3.3087604 ]  [3.1618638 3.16808653]  [2.99265671 3.00184631]  [2.76772523 2.7810216 ]  [2.45867157 2.47710919]  [2.07549787 2.09885192]  [1.70389068 1.73089588]  [1.43812907 1.46547759]  [1.28221703 1.30953729]  [1.18562424 1.21433711] | [1.1172719 1.14709628]  [1.06910992 1.10143363]  [1.03550959 1.06946564]  [1.01203835 1.04647458]  [0.99270535 1.0280565 ]  [0.97857553 1.01405334]  [0.96735966 1.00312746]  [0.95947689 0.99552649]  [0.95149791 0.98835826]  [0.94515347 0.98275882]  [0.93949896 0.97991073]  [0.93473172 0.97733271]  [0.93021518 0.97369504]  [0.92687988 0.97107834] | [0.92422754 0.96967977]  [0.9215011 0.96835786]  [0.91859424 0.96534723]  [0.91595078 0.96302277]  [0.91318631 0.96178037]  [0.91132236 0.96078962]  [0.90806037 0.95991343]  [0.90605873 0.96036083]  [0.90328163 0.95836163]  [0.90152222 0.95799106]  [0.89961767 0.95740908]  [0.89748031 0.95657265]  [0.89574528 0.95571709] | [0.89252853 0.95320892]  [0.89111674 0.95326155]  [0.88817143 0.9520368 ]  [0.88563126 0.95176488]  [0.88333035 0.95278233]  [0.88037843 0.95135128]  [0.87830025 0.95154023]  [0.87645012 0.95140362]  [0.87495416 0.95099223]  [0.87245631 0.94974422]  [0.86884838 0.94745862]] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 5e-3 | 10 x 10 Step | 80/20 |  |

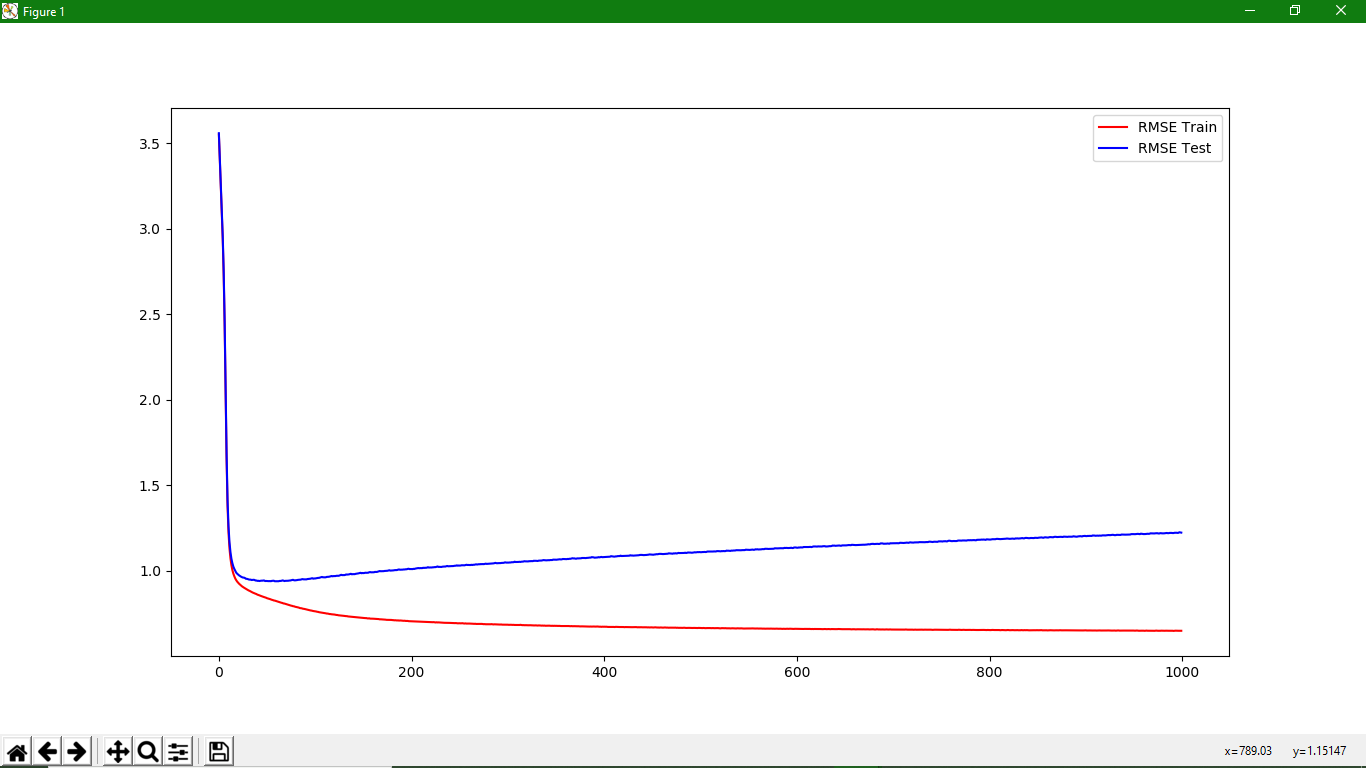
(converged): Visualisations are very irregular

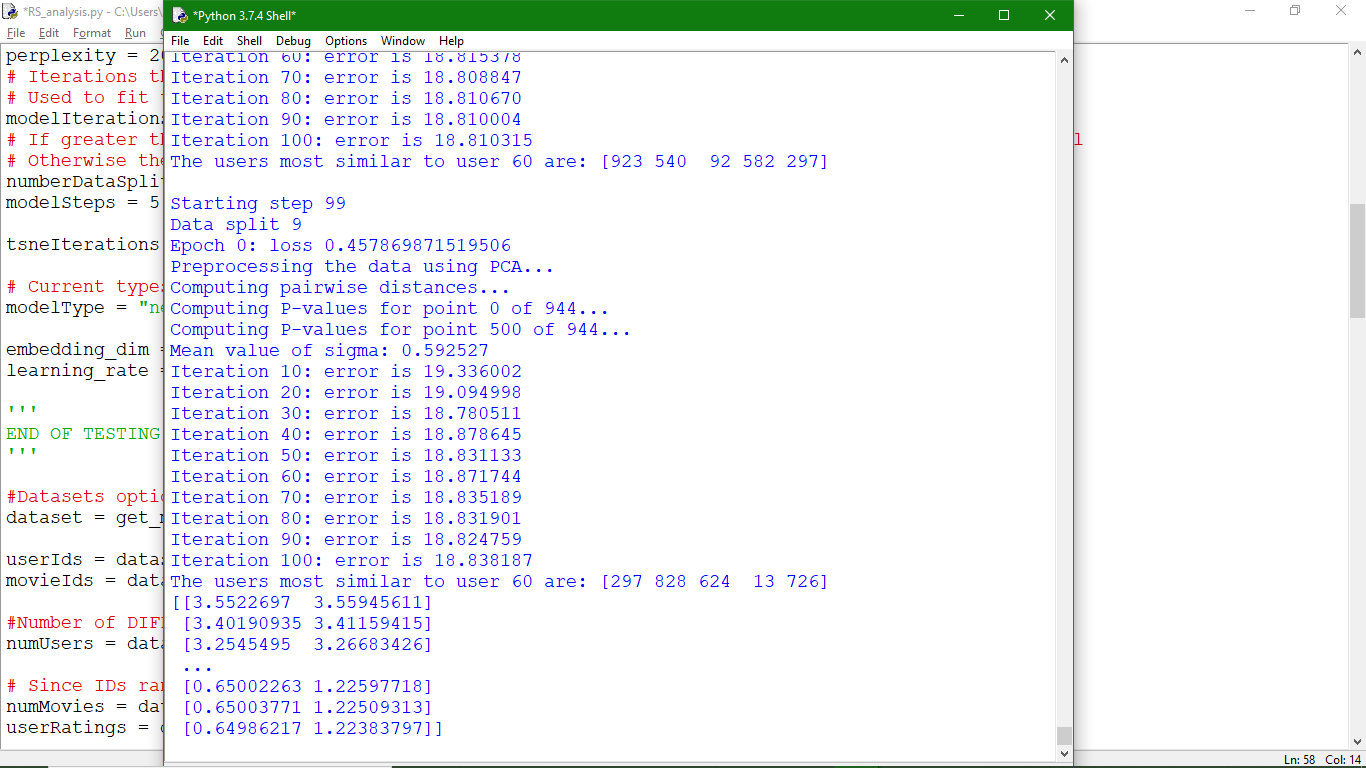


|  |  |  |  |
| --- | --- | --- | --- |
| [[3.55517769 3.54890108]  [3.40407491 3.39980745]  [3.25258565 3.25073195]  [3.08833218 3.08927321]  [2.88076925 2.88484669]  [2.57954574 2.58660841]  [2.16573238 2.17675281]  [1.73447418 1.75262654]  [1.43033552 1.4518342 ]  [1.2611084 1.28549743]  [1.1632576 1.18985415]  [1.09654832 1.12200832]  [1.05210876 1.07762146]  [1.02162242 1.04827034]  [0.99989283 1.02811909]  [0.98204768 1.01198578]  [0.9690128 0.99743658]  [0.95881927 0.98690349]  [0.95111537 0.98162419]  [0.94498938 0.9781155 ]  [0.93905205 0.97275716]  [0.9341144 0.96774101]  [0.93038625 0.96416885]  [0.92669773 0.96205729]  [0.92396098 0.96205622] | [0.92070925 0.95952863]  [0.91821349 0.95587379]  [0.91493779 0.95290905]  [0.91190511 0.95144379]  [0.91009784 0.95104712]  [0.9082076 0.9517988 ]  [0.90495068 0.9491967 ]  [0.9030835 0.94764787]  [0.89986473 0.94683868]  [0.89781034 0.94680429]  [0.89426374 0.94463682]  [0.89305985 0.94287729]  [0.89055526 0.94015074]  [0.88734549 0.94023943]  [0.88639754 0.94158077]  [0.88393527 0.94054157]  [0.88126588 0.93920314]  [0.87885624 0.93708181]  [0.87634939 0.9383074 ]  [0.87415886 0.93874758]  [0.87107432 0.93677598]  [0.86971146 0.93620616]  [0.86634427 0.93249404]  [0.86420351 0.93347704]  [0.86430347 0.93689144]  [0.86178756 0.93662119]  [0.8590923 0.93537623] | [0.85640174 0.93375921]  [0.85394681 0.93381935]  [0.85215044 0.93535691]  [0.84906173 0.93346816]  [0.84813476 0.93255997]  [0.84546983 0.93146157]  [0.84390032 0.93275779]  [0.8426801 0.93495965]  [0.84090173 0.93528295]  [0.8378188 0.9337787 ]  [0.83618432 0.93391323]  [0.83449495 0.93533438]  [0.83254462 0.93600774]  [0.82943231 0.93464619]  [0.82892305 0.93427467]  [0.82653773 0.93344569]  [0.82464617 0.93408287]  [0.82381296 0.93681824]  [0.82155752 0.93732077]  [0.81957614 0.93689615]  [0.81792313 0.93608564]  [0.81534564 0.93640989]  [0.81398219 0.93818069]  [0.81182218 0.938196 ]  [0.8110342 0.93821228]  [0.80833787 0.93686426]  [0.80669808 0.93746132]  [0.80564451 0.9407959 ] | [0.80389225 0.94160634]  [0.80207568 0.94083303]  [0.80028409 0.94040823]  [0.79846203 0.94113332]  [0.79775143 0.94384801]  [0.79535574 0.94313705]  [0.79516566 0.94387436]  [0.79219753 0.94205028]  [0.79073 0.94394416]  [0.7902776 0.94710141]  [0.78877777 0.94814068]  [0.78684247 0.9478721 ]  [0.78574449 0.94836169]  [0.78345436 0.94883388]  [0.78263229 0.9507907 ]  [0.78071094 0.95095879]  [0.78082281 0.95150191]  [0.77802736 0.94937348]  [0.7767511 0.95081896]  [0.77671212 0.95450383]] |

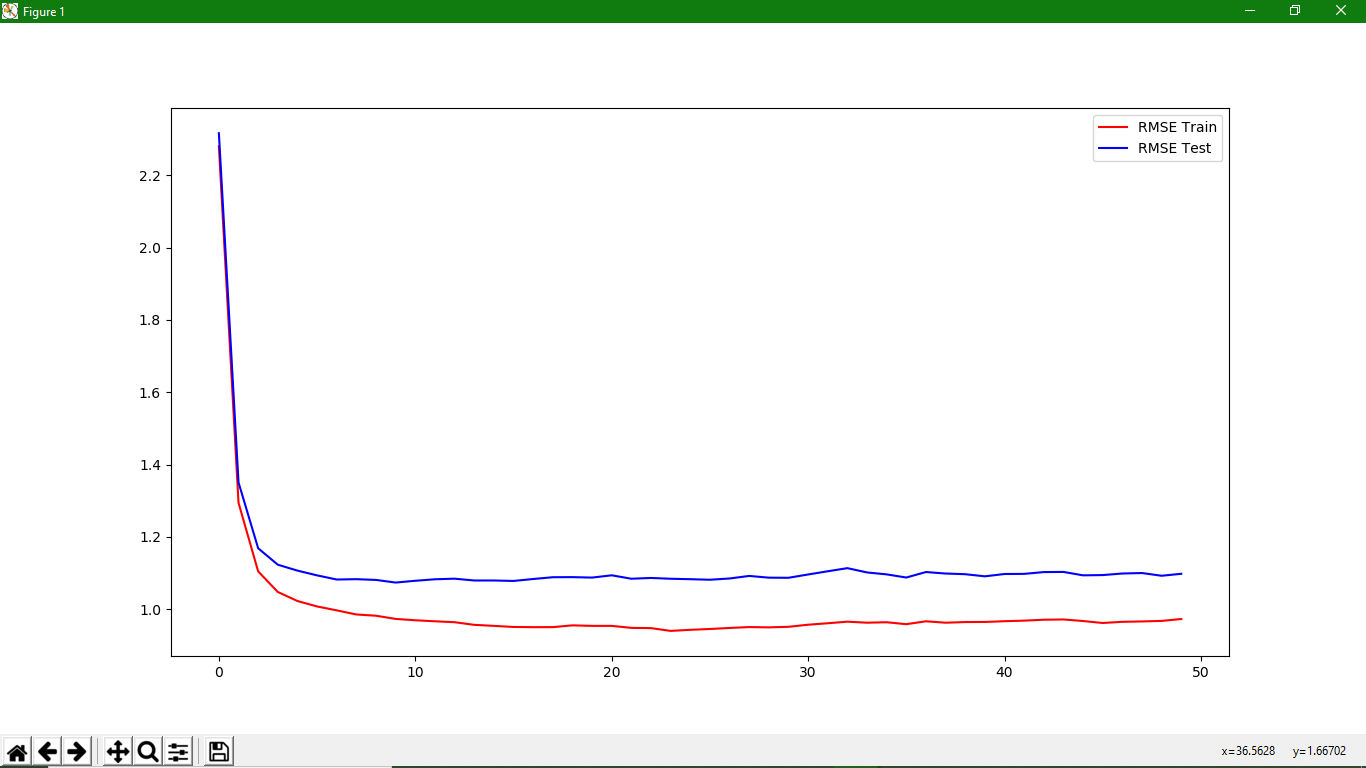
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 5e-3 | 10 x 100 Step | 80/20 |  |

(converged): Visualisations are very irregular



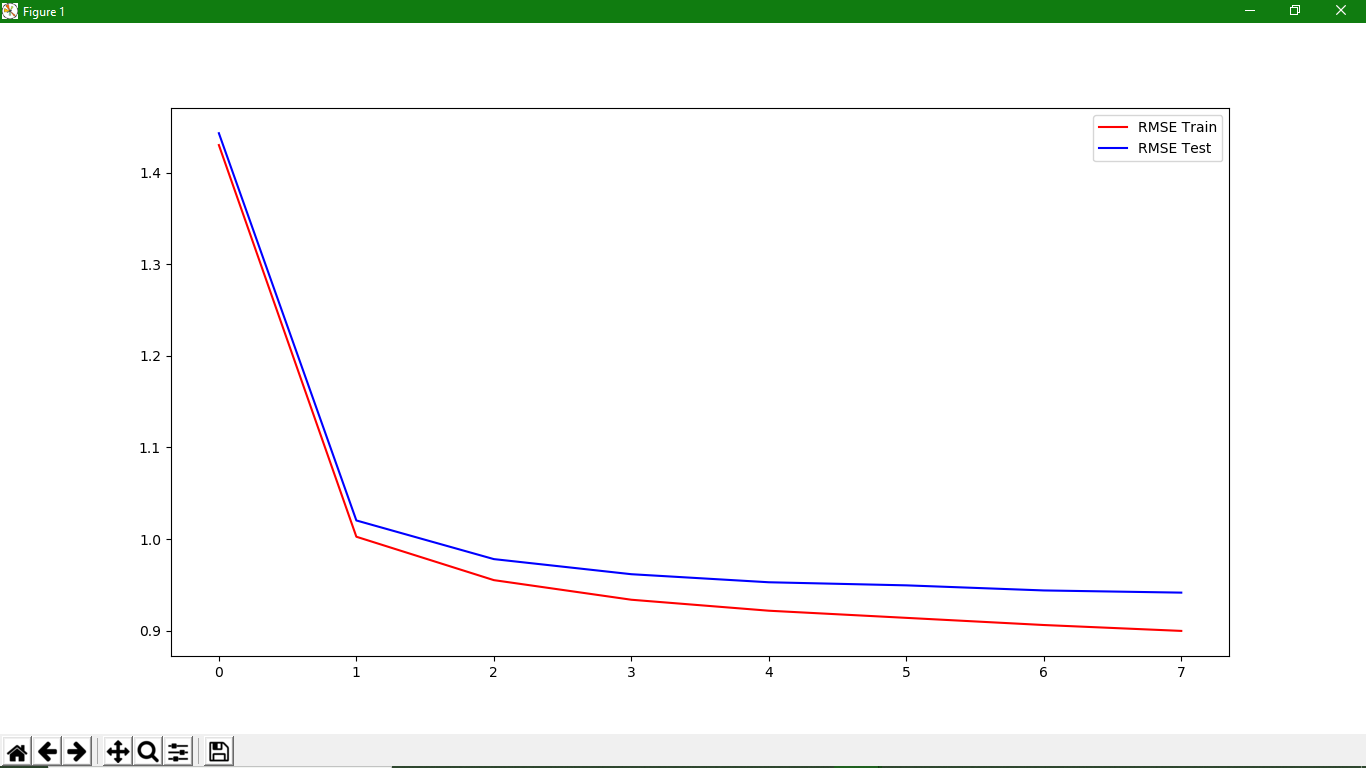


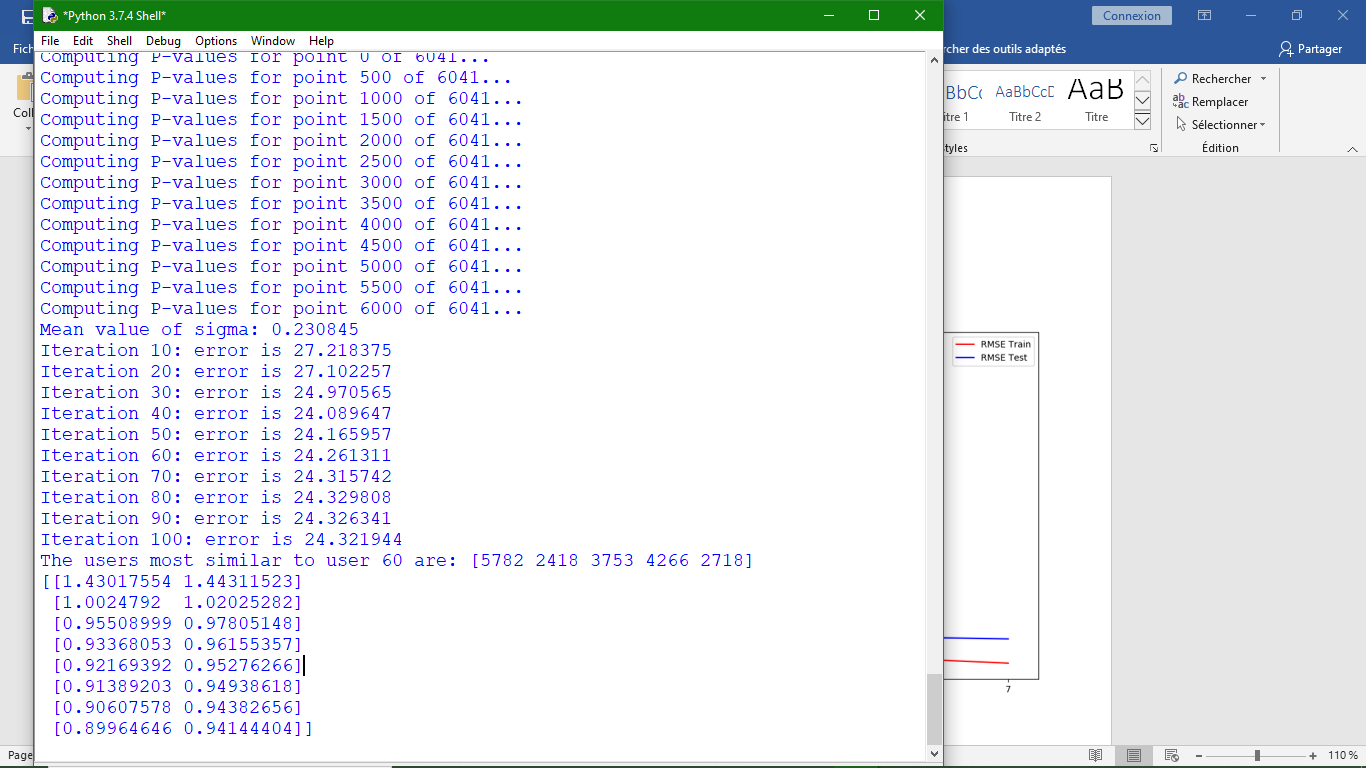
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split | Train/Test split |  |
| 100k | 5e-2 | 10 x 5 Step | 80/20 |  |

(converged): Visualisations look correct, some outliers at certain steps condense the scatterplots.

|  |  |  |  |
| --- | --- | --- | --- |
| [[2.28014445 2.31601143]  [1.29447746 1.35081387]  [1.10486305 1.1687423 ]  [1.0479635 1.12335014]  [1.02295375 1.10697579]  [1.00795066 1.09380984]  [0.99709517 1.08235633]  [0.98573983 1.08330584]  [0.98228896 1.08123517]  [0.97344482 1.07394564]  [0.9697625 1.07889593]  [0.96693462 1.08304441]  [0.96437323 1.08468258] | [0.95709997 1.07969701]  [0.95425445 1.07967663]  [0.9512195 1.07842302]  [0.95065516 1.08375704]  [0.95075327 1.08870649]  [0.95566702 1.08896255]  [0.95421094 1.08769286]  [0.95426351 1.09401655]  [0.94871056 1.08455777]  [0.94803154 1.08666277]  [0.94033045 1.08447242]  [0.94335628 1.08325112]  [0.94562441 1.08174503]  [0.94843787 1.0851934 ] | [0.95094228 1.09226751]  [0.95010692 1.08748281]  [0.95166576 1.08714521]  [0.95742905 1.09628761]  [0.96148753 1.10514975]  [0.9658193 1.11378741]  [0.96301848 1.10192883]  [0.9642669 1.09655416]  [0.95889109 1.08788288]  [0.96687907 1.10305917]  [0.96298224 1.09893501]  [0.9647944 1.09704363]  [0.96509296 1.09118259]  [0.96711099 1.0976156 ] | [0.96870625 1.09797537]  [0.97123915 1.10294044]  [0.97198784 1.10331452]  [0.9676367 1.09395075]  [0.96220422 1.0946486 ]  [0.96542436 1.09897959]  [0.96650684 1.10029876]  [0.96796894 1.09275341]  [0.97323388 1.09819949]] |

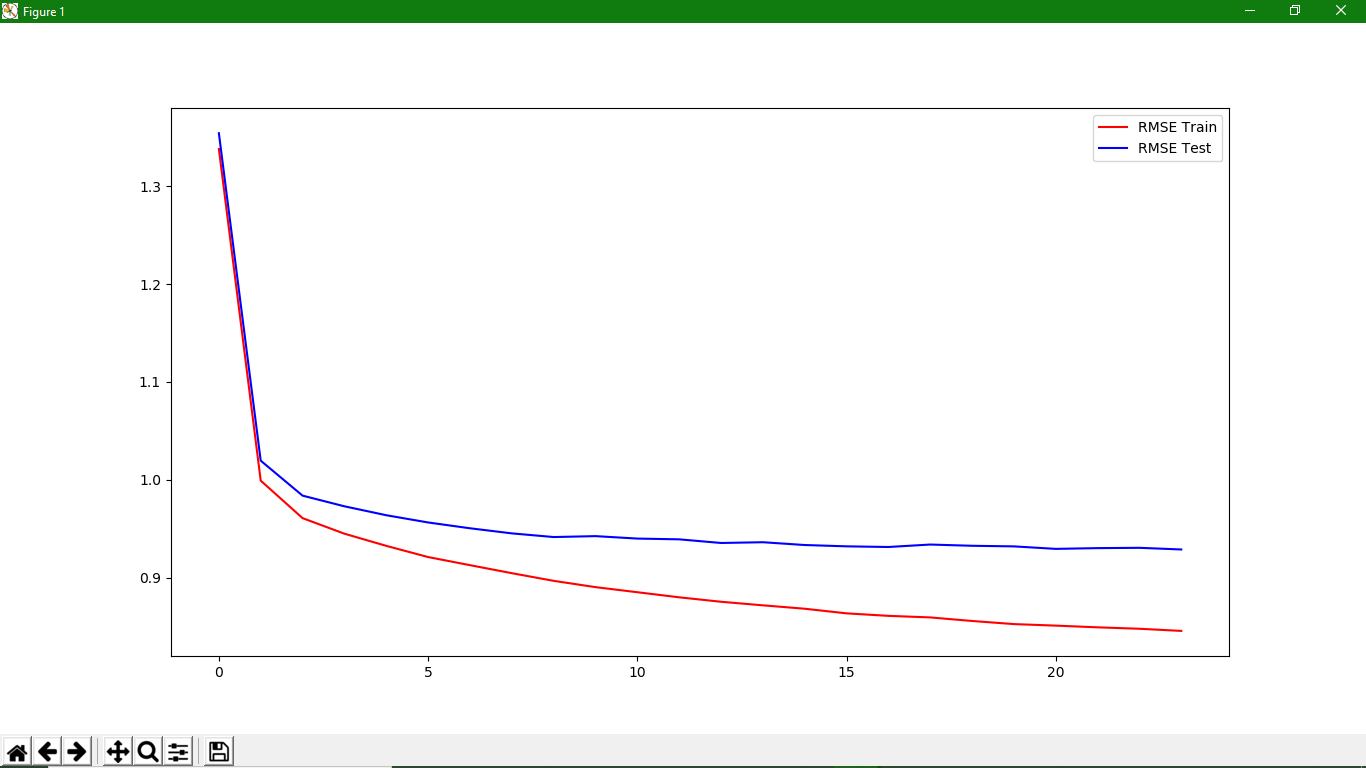
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split |  |  |
| 1M | 1e-2 | 8 x 1 Step |  |  |

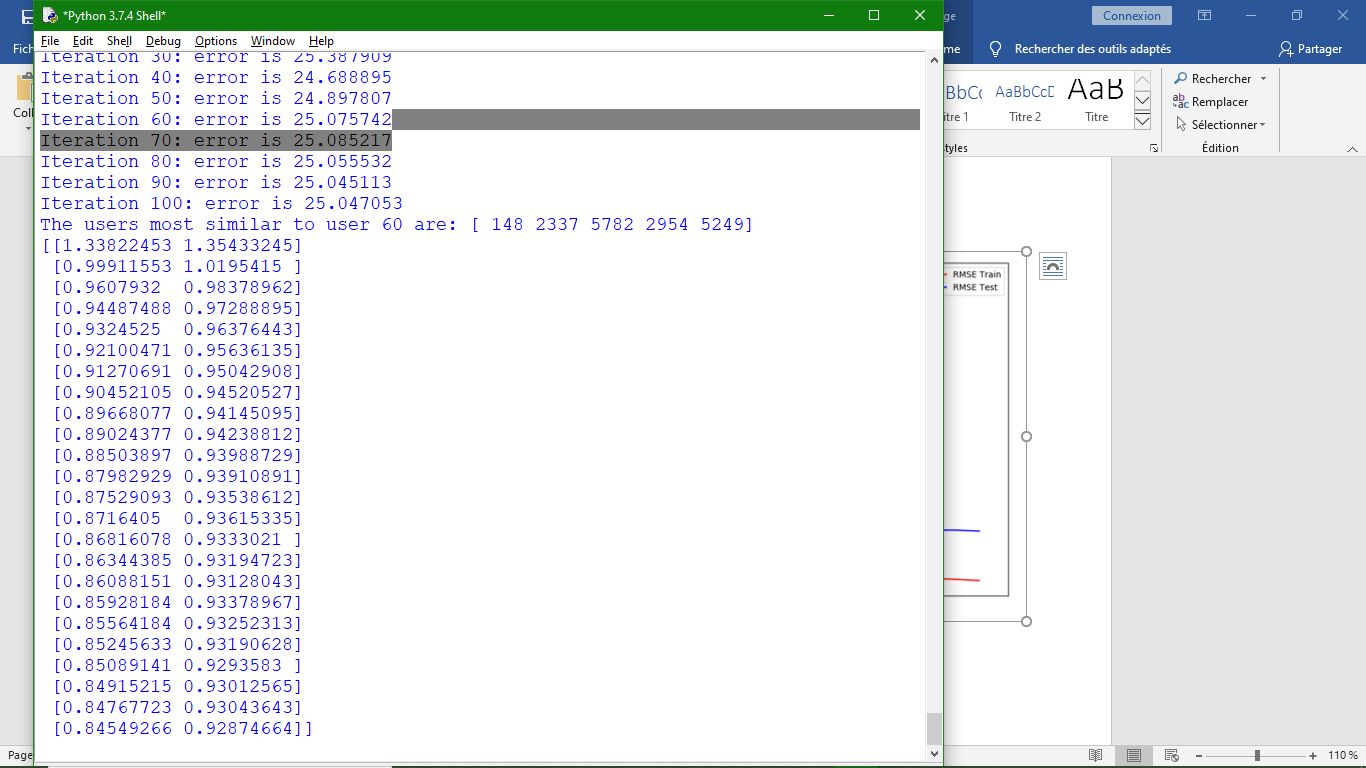




|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split |  |  |
| 1M | 1e-2 | 8 x 3 Step |  |  |

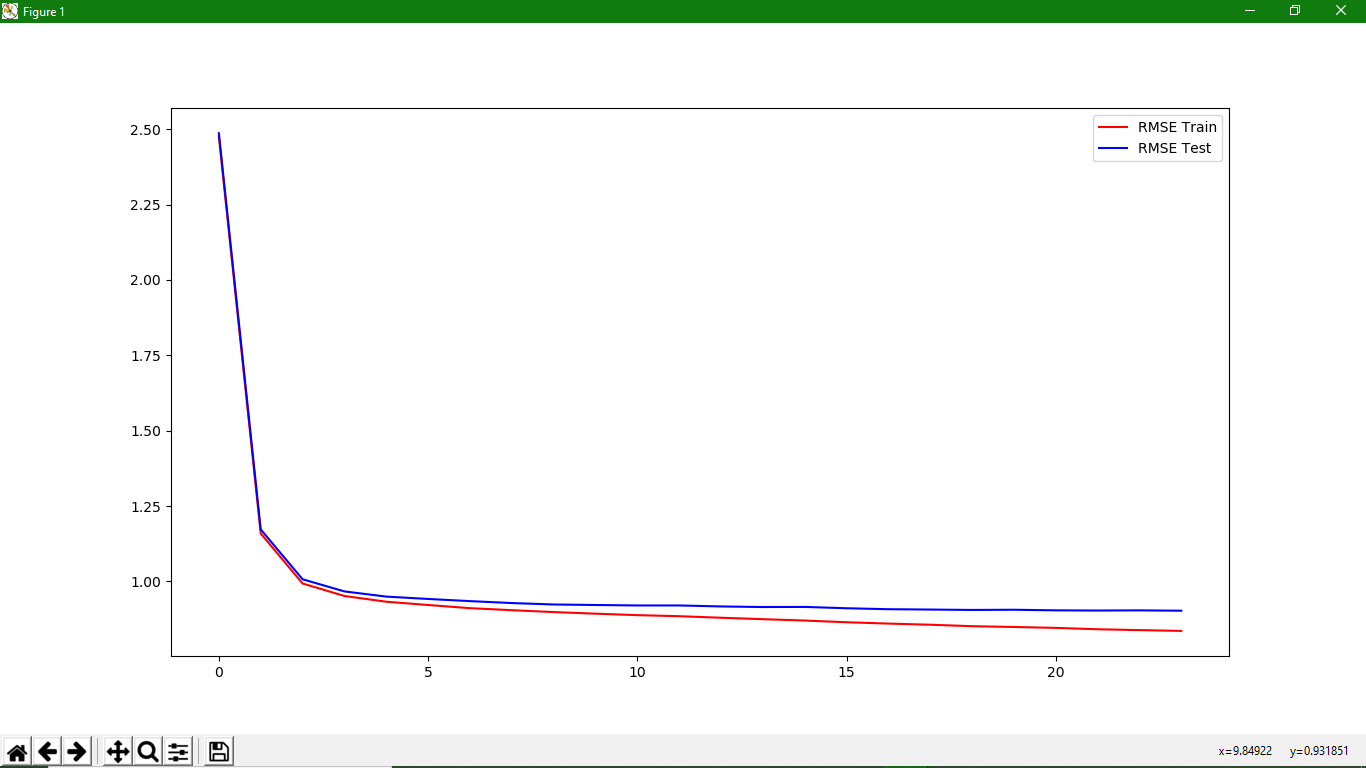
Visualisations seem consistent (images) but too dense.

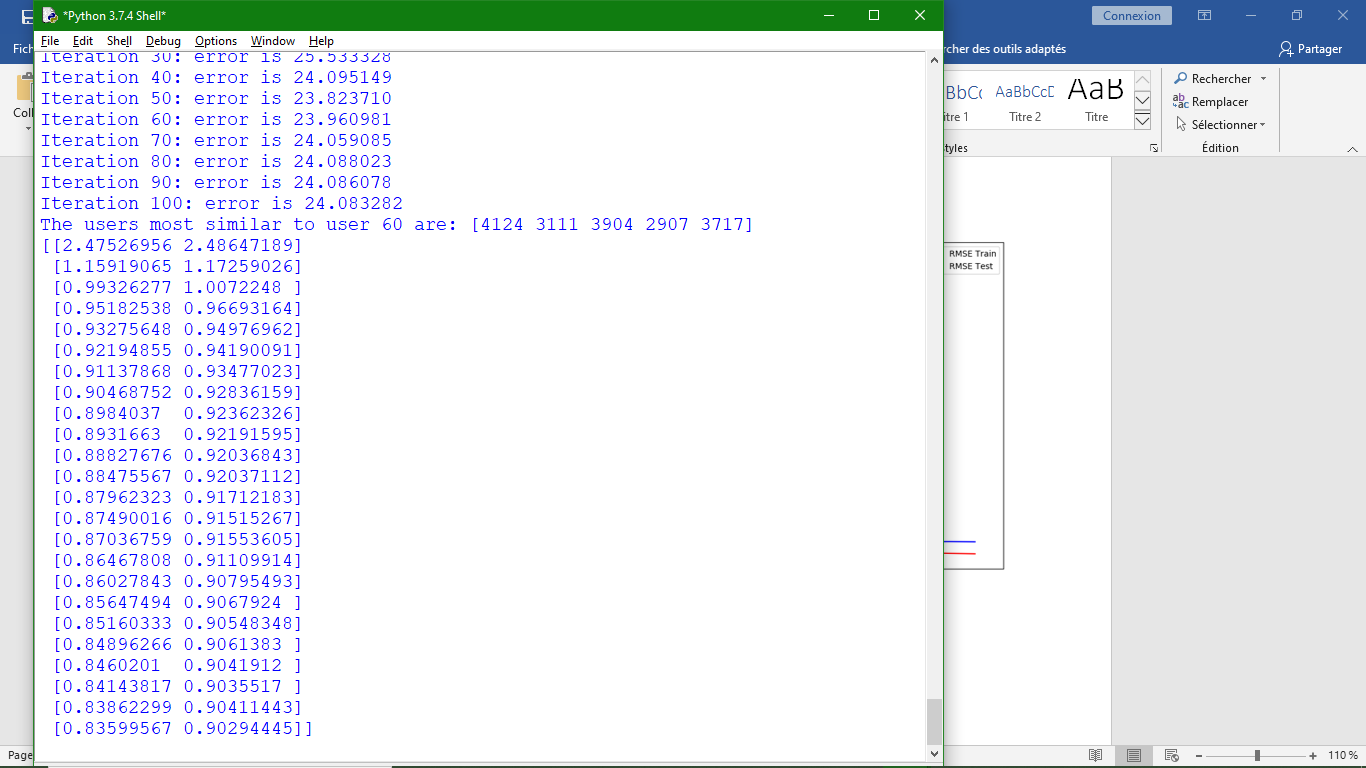




|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split |  |  |
| 1M | 5e-3 | 8 x 3 Step |  |  |

Visualisations seem consistent (images) but too dense. (perplexity 20)





|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Learning Rate | Data Split |  |  |
| 1M | 5e-3 | 8 x 2 Step |  |  |

Visualisation shapes have odd shapes. (perplexity 5)

