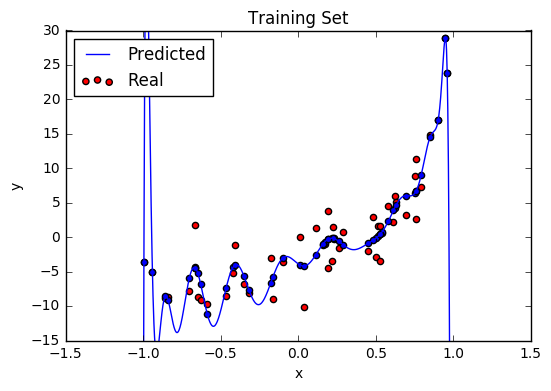
**COMP-551 : Applied Machine Learning**

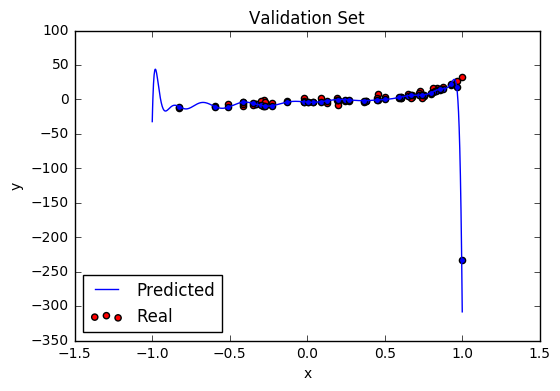
*Rony Azrak – 260606812 Due on : Jan 26, 5pm*

Programming Assignment #1

Question 1:

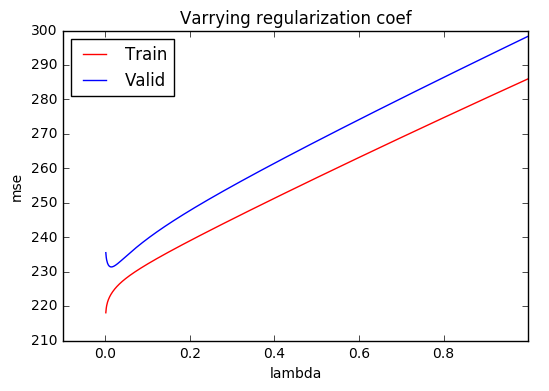
1) Training MSE = 6.4747033381. Validation MSE: 1421.94624006





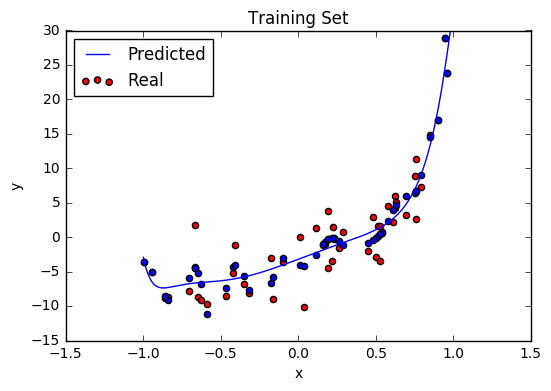
The MSE for the training set is not bad, but the one for the validation set is too high. This is explained by the fact that the validation set has data points with an x value closer to 1.0. We can see that the predicted 20 degree polynomial goes way out of the real data set values for x values closer to 1.0. This fit is not optimal because even though the training set does not have data points with as large x values as the validation set, those values do exist, so the fit must be closer to the data points.

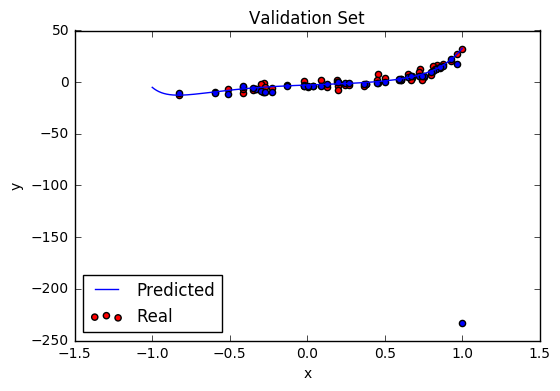
2)



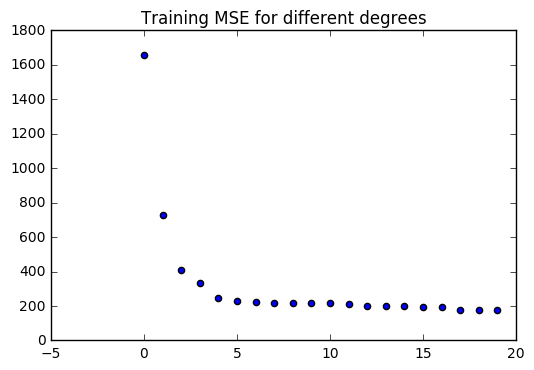
The minimum validation MSE has a value of 231.353 and the lambda that gives that value is 0.014.

The following graphs show the fit for the training data and validation data respectively.



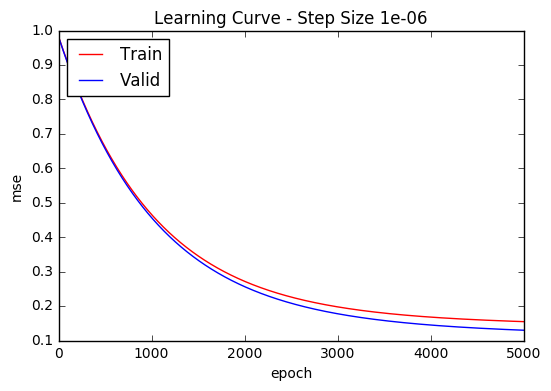


We can see that we have the same issue with the validation set where data points close to x = 1 are far from their predictive value. This tells us that a 20 degree polynomial is too high for this data.

3) It cannot be inferred exactly which degree the source polynomial is from the previous questions, but looking at the data points, it does not seem like there are more than 4-5 highs and lows. In the following graph, we plotted the training MSE for degrees 0 to 19 and found that the MSE plateaus at degree 4. Therefore, degree 4 is probably the source polynomial.

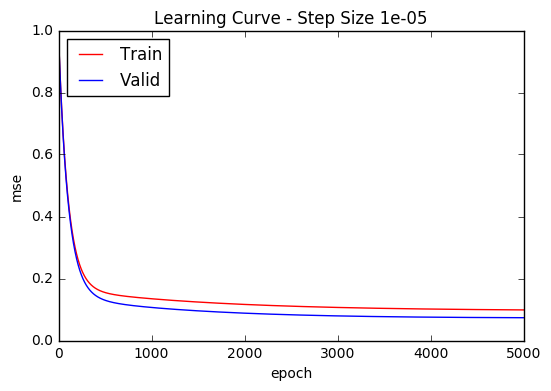
Question 2:

1) The following graph is a learning curve which shows the evolution of the training MSE and validation MSE for epochs 1 to 5000 with a step size of 1e-6.

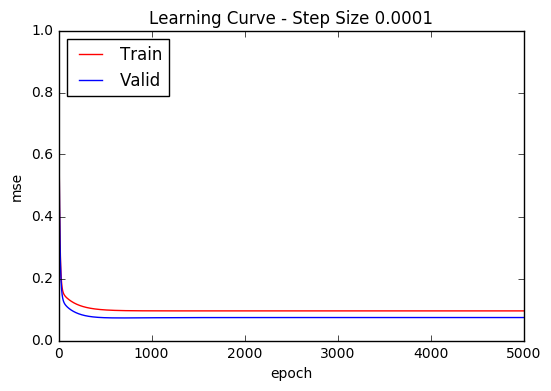


With a step size of 1e-6, the validation MSE does not totally converge in 5000 epochs.

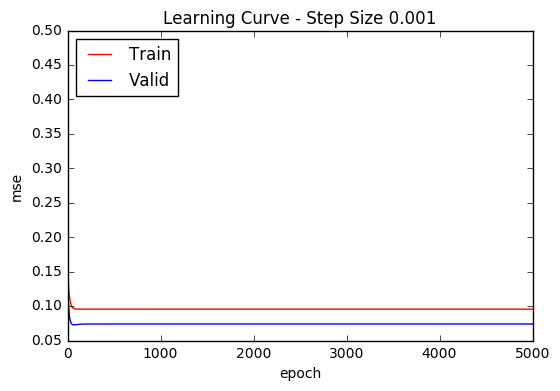
2) We will vary the step size from 1e-6, 1e-5, 1e-4 … 1e-1 and visualize the learning curve for all of them. We will report the epoch at which the validation MSE converges, with a convergence rate of 1e-5



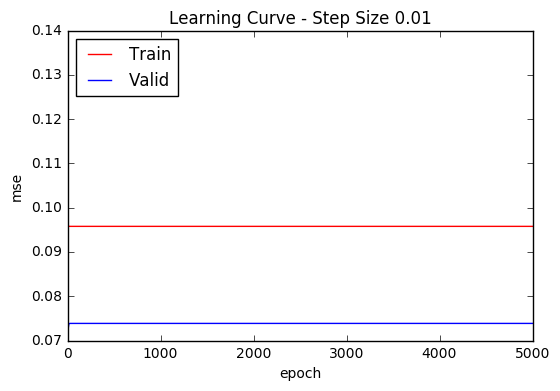
With a step size of 1e-5, the validation MSE converges after 2302 epochs.



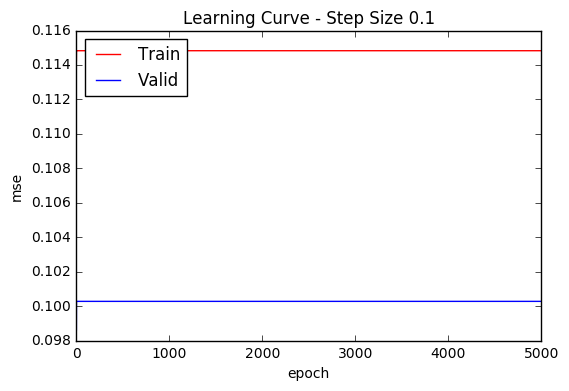
With a step size of 1e-4, the validation MSE converges after 503 epochs.



With a step size of 1e-3, the validation MSE converges after 64 epochs.



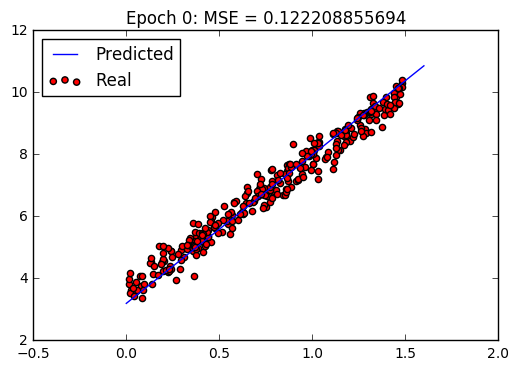
With a step size of 1e-2, the validation MSE converges after 21 epochs.

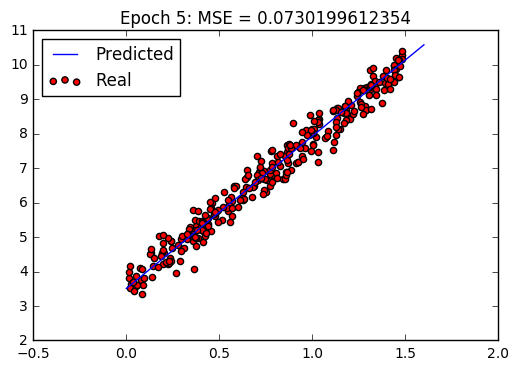


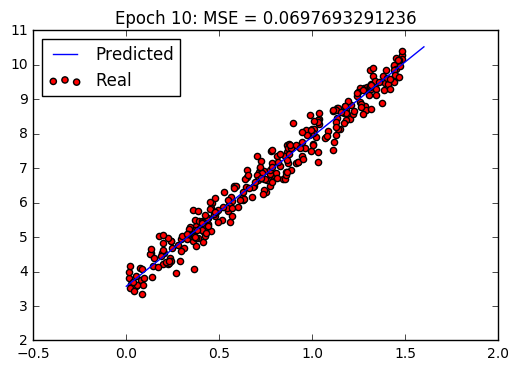
With a step size of 1e-1, the validation MSE converges after 3 epochs. However, the MSE does not converge to the same value as the ones displayed above, so this does not give weights as optimal as if we had a smaller step size.

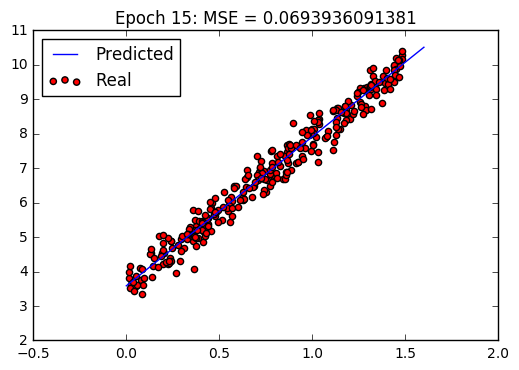
Therefore, the best step size is 1e-2. With that step size, the testing MSE = 0.069, and it gives the following values for the weight parameters [3.589232765108918, 4.3247691957865344] (w0 and w1).

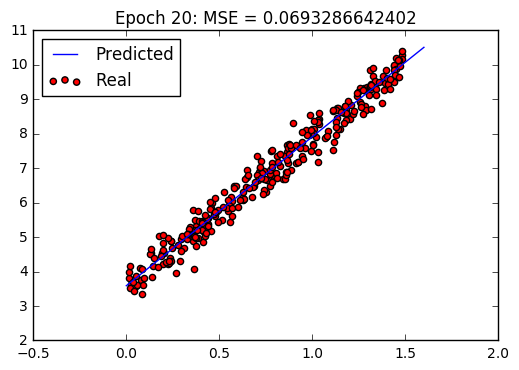
3) With a step size of 1e-2, the validation MSE converges after 21 epochs, so we will visualize the evolution of the fit between epochs 0 and 21.











Question 3:

1) Instead of filling all the missing values with the mean of the corresponding column, we chose to fill them with random values between (mean-std, mean+std). That way, there wouldn’t be many duplicate values and it would create less skew. The complete data csv file ‘Dataset\_3\_complete.csv’ can be found in the Datasets folder.

2) All splits have been saved in csv files under the Datasets folder. They are named ‘Dataset\_3\_test\_split0.csv’ and ‘Dataset\_3\_train\_split0.csv’ for the test set and training set. The following data will consist of the MSE of the 5 different splits, from split0 to split5, along with the weight parameters learnt. All data has been generated with a step size of 1e-9 as we found it gave us the smallest average MSE, and any bigger step size than 1e-6 would make the weights diverge. The final value shows the average MSE.

Test MSE: 511.382143253

Weight params: [-0.033370573412133203, 0.99913848987471543, 0.99994713334560281, 0.999596004699422, 0.9998386883756315, 0.99936095512451417, 0.99985989956579813, 0.99986965297619324, 0.99963622343690306, 0.99957540034443904, 0.99971139253352903, 0.99963786607938865, 0.99994118035316804, 0.99938974764438826, 0.99968260066958992, 0.99951557697973203, 0.99974864136331631, 0.99957147555930559, 0.99959731354528469, 0.99972542090675498, 0.99958724227265527, 0.99967183044787511, 0.99969366559209127, 0.99967718648060944, 0.99974214993284261, 0.99982199829236629, 0.9997137118608963, 0.99975043083167914, 0.99966706095137947, 0.99994906102637882, 0.99974210506521943, 0.99972640217232822, 0.99966975778664569, 0.99968523485720984, 0.99968508961284963, 0.99956599727508155, 0.99966305968328495, 0.99961971758253576, 0.99966963542222576, 0.99961587135601015, 0.99960653258741905, 0.99962428762690969, 0.99957987904408707, 0.99957607392350423, 0.99957566261073094, 0.9994734145025822, 0.99946595152084117, 0.99942491007221002, 0.99949823643742797, 0.99956923776810058, 0.99954621825787637, 0.99996510736896782, 0.99978156061967915, 0.99996956702940365, 0.99972903524449608, 0.99969170011766162, 0.9996553649829889, 0.99962899646903813, 0.99983590047864024, 0.99983394881854226, 0.9998305596288809, 0.99983133197993834, 0.99933162700353673, 0.99986072278289229, 0.99976390713690377, 0.99977689798076574, 0.99959571869917907, 0.99956886419060298, 0.99964597031755664, 0.99951543070003579, 0.99983195888833942, 0.99957466734330147, 0.99972820247753524, 0.99993043618571198, 0.99938075367713597, 0.99952687468690915, 0.99981896774882673, 0.99963025768388936, 0.99957087069784512, 0.99977392797908682, 0.9997876249141191, 0.99976474629887757, 0.9997650567727745, 0.99975997172686393, 0.99969303886102212, 0.99967136577012794, 0.99962508291562069, 0.9996605882778089, 0.99957173674733568, 0.99960622184112746, 0.99965505129144017, 0.99997316488410148, 0.99997905887033844, 0.99980230065677222, 0.99947941875756319, 0.99953936624158124, 0.99946150574145742, 0.99943835968094197, 0.99991444479414537, 0.9998133677620874, 0.99920659586687899, 0.99978326777006266, 0.99988984084034371, 0.99981503430381313, 0.99970796708204057, 0.99980927964771704, 0.99940473811499719, 0.99937447933425561, 0.9998047457543533, 0.99985865343370939, 0.99985775588157821, 0.99977798594068501, 0.99991590698648802, 0.99952435225580905, 0.99973255788458082, 0.99994228341429992, 0.99979435062975086, 0.99985555595503772, 0.99983879691941213, 0.99990884208464492, 0.9993980963904977, 0.99963164520706238, 0.99991864682889309, 0.99982975541960717]

Test MSE: 482.104994092

Weight params: [-0.033931005265575374, 0.99913193621569896, 0.99994551117105601, 0.99959067725951511, 0.99984021313271387, 0.99935362251877113, 0.99985720760167418, 0.99987008354147477, 0.9996337039126405, 0.99957211272423374, 0.99971076339745868, 0.99964313897119228, 0.99993918648062974, 0.99937522836184289, 0.99967011333109668, 0.99950309800680792, 0.99974975456137005, 0.99956473073191876, 0.99960192286083682, 0.99973256515802877, 0.99958702667437016, 0.99965984234047256, 0.99968460726194031, 0.99966825674339843, 0.99974416410240741, 0.99982265079102084, 0.9997174925305492, 0.9997421878469771, 0.99965800578876207, 0.99994943979500406, 0.9997483375238938, 0.99972974053624719, 0.99967276905007141, 0.99967685061777667, 0.99969320275607554, 0.99955382211258947, 0.99965927990687875, 0.99962094840321014, 0.99966924272198965, 0.99960863158530233, 0.99960480175202326, 0.99962363747504535, 0.9995789696930536, 0.99957474201501406, 0.99956984376561364, 0.99946326586916012, 0.99945609418215997, 0.9994162755784145, 0.99948977571479547, 0.99956570531964284, 0.99954431223719453, 0.99996532357248635, 0.99977969637237329, 0.99997109355162372, 0.99972206626812621, 0.99968621049414597, 0.99965059082315966, 0.9996234333879167, 0.99983230657861488, 0.99983195307701089, 0.99982865318499115, 0.99982982182144509, 0.99932285782193686, 0.99986339260147672, 0.99976273618429179, 0.99977562760072514, 0.99958927726372038, 0.9995629305129422, 0.99964180085069043, 0.99950483790183253, 0.99983355313479716, 0.99957810772495559, 0.99971829989366268, 0.99993018819000457, 0.99937008484643952, 0.99951666206485912, 0.99981724076483769, 0.99962944808847476, 0.99956061577707167, 0.99977734668454699, 0.99979035401793104, 0.99975845524419649, 0.99975930372148358, 0.99975497645659828, 0.99968499631595276, 0.99966258578243017, 0.99961584438300155, 0.99965068896119214, 0.99957278924404547, 0.99960303613100354, 0.99964951103402222, 0.99997343808719164, 0.99997893554361295, 0.99980275478128156, 0.99948160344208137, 0.99953664477588977, 0.99946070627467942, 0.99943829810083573, 0.9999148443835838, 0.99981428164385755, 0.99919649451029213, 0.99978806363052097, 0.99989064282416196, 0.99981666446603223, 0.99969994944257357, 0.99981685092965922, 0.99940018023219013, 0.99935977350510174, 0.99980091888121569, 0.99986007387042752, 0.99985734920427616, 0.9997719531973559, 0.99991648909706909, 0.99951466085792151, 0.99973376394332658, 0.99993740060520886, 0.9997976021425975, 0.99985143603404147, 0.9998396895517323, 0.99991029924909014, 0.99938871434537002, 0.99960899750416199, 0.99991287350143876, 0.99983164521659607]

Test MSE: 562.599696989

Weight params: [-0.03328088221640868, 0.99913922808107147, 0.99994924999750689, 0.99959160981422557, 0.99984279453310188, 0.99935810419404125, 0.999858611902538, 0.99986900049072036, 0.99963171781292626, 0.99957304828470217, 0.999709782825841, 0.9996485661660629, 0.9999433952056056, 0.99938876207226512, 0.99967760787587445, 0.99950937304418919, 0.99975198717617886, 0.99957029956378229, 0.99960536939504852, 0.99972762268319915, 0.99959337240398394, 0.99966692988717765, 0.99969097225260506, 0.99967515567663134, 0.99974282843015816, 0.99982066117263058, 0.99971623605431226, 0.99975080643995617, 0.99966255215775501, 0.99995171989921228, 0.99974249866333242, 0.99973186133415703, 0.99967609148577619, 0.99967935551473375, 0.9996911733084245, 0.99955990444355247, 0.99966302434810106, 0.99961934455378343, 0.99967397513012257, 0.99961229624939452, 0.99960575753408043, 0.99962312450955981, 0.99958049760642564, 0.99957621289706067, 0.99957261911358875, 0.99946987630426798, 0.9994617350089614, 0.99942310516147415, 0.99949651937870454, 0.99957492730886677, 0.99955179144623563, 0.99996762969353681, 0.99978431834492676, 0.99997269216797147, 0.99972701229945526, 0.99969038836443092, 0.99965828182936489, 0.99963214825697633, 0.99984030634062038, 0.99983818094477994, 0.99983581380997077, 0.9998374636638977, 0.99933201904306612, 0.99986518908227906, 0.99976303626722662, 0.99977601257537163, 0.99959351696959053, 0.99956646072500943, 0.99964513948137834, 0.99951403902410041, 0.99983511202038056, 0.99958017250158604, 0.9997212847548862, 0.99993402438802215, 0.99937876193307007, 0.99952572377770033, 0.99981538597918651, 0.99962874926738432, 0.9995673201920483, 0.99977691803539481, 0.99978838565217387, 0.99976004680337327, 0.99976039912797088, 0.99975550771206589, 0.99968964060018528, 0.99966746224323877, 0.9996221380993936, 0.99965696035049967, 0.99957737269327718, 0.99960508525212444, 0.9996519625245629, 0.99997310360921121, 0.99997730785570471, 0.9998085720981178, 0.99948335759191687, 0.99954316990707992, 0.99946484087735799, 0.99944388803611661, 0.99991815863978861, 0.999812622169582, 0.99919941467896234, 0.99978491469798303, 0.99989162890075545, 0.99981244451338747, 0.99970228184730947, 0.99980865009237974, 0.99941464347496678, 0.99937233724607799, 0.9998050310109351, 0.99986377579310137, 0.99985226156295293, 0.99977805456653668, 0.99992100309166365, 0.999514929393023, 0.99973961067080652, 0.99994102337419277, 0.99980008125332631, 0.99985700110048115, 0.99984369811090301, 0.99991197823724876, 0.99940717236668808, 0.99962457525970505, 0.99991141150706264, 0.99982916015109724]

Test MSE: 484.815304488

Weight params: [-0.033409744481980898, 0.99912765262359826, 0.99995136168668086, 0.99958982143903485, 0.99984523713138951, 0.99934073086606734, 0.99986315760783073, 0.9998746298292186, 0.99962963843063024, 0.99956912146851074, 0.99970711389889233, 0.99963822339929076, 0.99994539318325404, 0.99937869780729582, 0.99967402976083508, 0.99950460304573041, 0.99974422895472903, 0.99956026823433852, 0.99959569542285565, 0.99972922252842766, 0.99958724445746394, 0.99966253473157474, 0.99968677754895563, 0.99967204053411429, 0.99974095133633489, 0.99981833647063834, 0.99971893029909209, 0.99974818502531371, 0.99965722621746245, 0.99995311405337051, 0.99974218826304651, 0.99972676798051419, 0.99966997299984806, 0.99967786901537958, 0.99968899509705555, 0.99955615884435711, 0.99965909415649989, 0.99961501060463021, 0.99966864135531497, 0.99960881292738224, 0.9996069170009515, 0.99962148846585031, 0.99958218683817035, 0.99957753957976125, 0.99956905016120445, 0.99945830608130382, 0.99945034223448548, 0.99941031135646585, 0.9994850522191504, 0.99956327006583023, 0.99954414989149543, 0.99996880165126345, 0.9997869598243353, 0.99997440397624915, 0.99971990998532445, 0.99968453176525918, 0.99965021945341337, 0.9996256333109268, 0.99984025954831401, 0.99983946564536363, 0.99983706132700756, 0.99983819675356511, 0.99931569708296764, 0.99986881745285605, 0.9997648754766586, 0.99977788440387638, 0.99959006784512983, 0.99956182487996692, 0.99964384864232414, 0.9995026364750319, 0.99983771969094082, 0.99957553822757628, 0.99971845144584115, 0.99993448350491509, 0.99937119955403042, 0.99951508632327479, 0.9998210915503829, 0.99962219914670536, 0.99956486947378886, 0.99977546472583945, 0.99979129702191771, 0.99976046348394298, 0.99976114152280471, 0.99975637315476518, 0.99968762319374627, 0.99966548591515036, 0.99962041043241656, 0.99965515026944185, 0.99957472837284722, 0.99960204789660312, 0.99964223856780376, 0.9999749327201829, 0.99998045291822568, 0.99980896294356003, 0.99946567664161279, 0.99953048142120382, 0.99945154531138081, 0.99942882618581164, 0.99991515520535224, 0.9998093750127427, 0.99919324950159227, 0.99978798539870162, 0.99989354485959514, 0.99981201932009922, 0.99970122791666693, 0.99980942392602223, 0.99939443351998858, 0.99936600642964468, 0.99980055417997726, 0.99986593413909386, 0.99985481595140158, 0.99977669584455386, 0.99991906161625999, 0.9995152934677366, 0.99973326184954425, 0.99994294081573443, 0.99980065074945945, 0.99985616816341494, 0.9998403768822588, 0.99991269897202073, 0.99940050674602865, 0.9996118667859103, 0.99992265991166118, 0.99983069762093413]

Test MSE: 524.04306682

Weight params: [-0.03305300851933754, 0.9991340067639336, 0.99995058803498971, 0.99958993151252529, 0.99984420787262607, 0.99935441952975568, 0.99985831515661139, 0.99986632653225438, 0.99962851465497937, 0.99956902620400168, 0.99970548858387764, 0.999639353489848, 0.99994493617644131, 0.99939676282322065, 0.99968114423245524, 0.99951179997729722, 0.99974324072123422, 0.99957059771991386, 0.99959726555400441, 0.99972500508089668, 0.99958760445591377, 0.99966990158694502, 0.99969361783039046, 0.99967750778984921, 0.9997452743678128, 0.99982362750228726, 0.99971405809814629, 0.99975212992703977, 0.99966247202357883, 0.99995364449603508, 0.99974064845928468, 0.99972188011335084, 0.99966574802502228, 0.9996837394338921, 0.99968548465911289, 0.9995622078997618, 0.99965421379950614, 0.99961533627726828, 0.99966072805227468, 0.99961567602006263, 0.99960557225986035, 0.99961799603149193, 0.99958179391700874, 0.99957672266486952, 0.99956982120097837, 0.99946862457838692, 0.99946105984873224, 0.99942251884512201, 0.99949453557060508, 0.99956492316860579, 0.99954437648377625, 0.9999696617552789, 0.99978070427280308, 0.99997308362565074, 0.99972432465562455, 0.99968793148536361, 0.99965246383802453, 0.99962824296131592, 0.99983239997961748, 0.99983108901923945, 0.99982793376190848, 0.99982992150939998, 0.99933020725984056, 0.99985843164306809, 0.99975797152724177, 0.99977153893108295, 0.99959232033291812, 0.99956548237916676, 0.99964295921938617, 0.99951298147713885, 0.99983093935602352, 0.99957055461162314, 0.99972949916932985, 0.99993748709926444, 0.99937316801963938, 0.9995246084560373, 0.99981838491653252, 0.99962549829896563, 0.99957108646510673, 0.99977506319644782, 0.99979207783669866, 0.99976016904615739, 0.99975989051691772, 0.99975519564752191, 0.99969343366481411, 0.9996703629586895, 0.99962243417389463, 0.99966038830196158, 0.9995698578219292, 0.99959923721668964, 0.99964646451340677, 0.99997635358073067, 0.99998176161126218, 0.99980048064988158, 0.99947384273572282, 0.99953678577557714, 0.99945871214352744, 0.99943281419770813, 0.9999126528816259, 0.9998112191097549, 0.99919956785289421, 0.99978806346296845, 0.99989377692155113, 0.99981357438247742, 0.99969826848571919, 0.99981286745056608, 0.99940117266295259, 0.99937190401873, 0.99980195835105368, 0.99986103016131311, 0.99986129921328737, 0.99978081667667351, 0.99991834565400695, 0.99951737499242754, 0.99973847912654568, 0.99994302578331795, 0.99979147971331872, 0.9998556916551522, 0.99984070354653265, 0.99991174401213978, 0.99939067004141025, 0.99962033954383289, 0.99991793452420907, 0.99982934808487089]

Average MSE: 512.989041128

3) Not Completed