# Time series prediction

In this project, you will regress a neural network with sinusoidal activation functions to fit to some time-series data.

Instructions:

1. Download this data. (I scraped this data from the US department of labor statistics.)

2. Create a 3-layer neural network with a topology of 1 -> 101 -> 1. The first and last layers are linear. The hidden layer will be non-linear. The input to this network will be time. The output will be the unemployment rate indicated in the data. The hidden layer will use the sinusoid activation function in the first 100 units of the hidden layer. It will use the identity activation function (a.k.a. no activation function) for the last unit of the hidden layer. Initialize the weights in the first linear layer as follows:

   ```
                weight      bias
                ------      ----
       unit 1   1*2*pi      pi
       unit 2   2*2*pi      pi
       unit 3   3*2*pi      pi
       ...      ...         ...
       unit 50  50*2*pi     pi


       unit 51  1*2*pi      pi/2
       unit 52  2*2*pi      pi/2
       unit 53  3*2*pi      pi/2
       ...      ...         ...
       unit 100 50*2*pi     pi/2
       unit 101 0.01        0      (the linear unit)
   ```

   Initialize the weights in the last linear layer with small random values.

3. Train this network using the feature values

   ```
       0/256
       1/256
       2/256
       ...
       255/256
   ```

   and the first 256 rows from the labor statistics data as the corresponding labels.

4. Plot the data and corresponding predictions for values from 0/256 to 356/256. Label the axes. Indicate the point t=256/256 on your chart, pehaps with a vertical line. (This is the point where it begins predicting into the future, as far as it knows.)

5. Implement $L^1$ regularization. Regularize only the outbound weights from sinusoid units. (Don't regularize the bias terms or the weight from the one hidden linear unit.) Fiddle with the regularization term until you find one that makes a difference, but doesn't ruin the results. Add these results to the same chart.

6. Implement $L^2$ regularization. Regularize only the outbound weights from sinusoid units. (Don't regularize the bias terms or the weight from the one hidden linear unit.) Fiddle with the regularization term until you find one that makes a difference, but doesn't ruin the results. Add these results to the same chart. Now, you should have four curves/lines (the data, predictions made without regularization, predictions with L1 regularization, and predictions with L2 regularization). Label them, so it is clear which is which.

7. Submit an archive containing your code and chart here. There is no need to execute anything on the submission server.

## Hints:

- Here is some debug spew from a working implementation to help you debug a broken implementation. To use it, set all the bias values to 0 and all the weights in the output layer to 0.01. Also, present the training patterns in sequential order.

- **Q: Can you give us some simpler debug spew?**
  A: Ok, here is some...

  ```
      In this example, I used a 1->5->1 topology.
      The first four hidden units were sinusoidal, and the last hidden unit was linear.
  ```

I trained with the same training data, but I visited the patterns in sequential order.
(Note that for your final results, you should visit them in random order.)

Learning rate=0.01
Momentum=0.0
Weights: 3.1415926535898,3.1415926535898,1.5707963267949,1.5707963267949,0,
        6.2831853071796,12.566370614359,6.2831853071796,12.566370614359,0,
        0.01,0.01,0.01,0.01,0.01,0.01
Input: 0
Layer 0 activation: 3.1415926535898,3.1415926535898,1.5707963267949,1.5707963267949,0
Layer 1 activation: 1.2246467991474e-16,1.2246467991474e-16,1,1,0
Layer 2 activation: 0.03
Label: 3.4
Layer 2 blame: 3.37
Layer 1 blame: 0.0337,0.0337,0.0337,0.0337,0.0337
Layer 0 blame: -0.0337,-0.0337,2.0635298565633e-18,2.0635298565633e-18,0.0337
Gradient: -0.0337,-0.0337,2.0635298565633e-18,2.0635298565633e-18,0.0337,
        0,
        0,0,0,0,3.37,4.1270597131266e-16,4.1270597131266e-16,3.37,3.37,0
Weights: 3.1412556535898,3.1412556535898,1.5707963267949,1.5707963267949,0.000337,
        6.2831853071796,12.566370614359,6.2831853071796,12.566370614359,0,
        0.0437,
        0.01,0.01,0.0437,0.0437,0.01
Input: 0.00390625
Layer 0 activation: 3.165799346196,3.1903430388021,1.5953400194011,1.6198837120072,0.000337
Layer 1 activation: -0.024204328633827,-0.048731077478764,0.9996988186962,0.99879545620517,0.000337
Layer 2 activation: 0.13030821575206
Label: 3.8
Layer 2 blame: 3.6696917842479
Layer 1 blame: 0.036696917842479,0.036696917842479,0.16036553097163,0.16036553097163,0.036696917842479
Layer 0 blame: -0.036686166831694,-0.036653319529609,-0.003935567142773,-0.0078687636470596,0.036696917842479
Gradient: -0.036686166831694,-0.036653319529609,-0.003935567142773,-0.0078687636470596,0.036696917842479,
        -0.0001433053391863,-0.00014317702941253,-1.5373309151457e-05,-3.0737357996327e-05,0.00014334733532218,
        3.6696917842479,
        -0.088822425930792,-0.17882803466137,3.6685865416918,3.6652714797803,0.0012366861312916
Weights: 3.1408887919215,3.1408891203945,1.5707569711235,1.5707176391584,0.00070396917842479,
        6.2831838741262,12.566369182589,6.2831851534465,12.566370306986,1.4334733532218e-06,
        0.080396917842479,
        0.0091117757406921,0.0082117196533863,0.080385865416918,0.080352714797803,0.010012366861313
Input: 0.0078125
Layer 0 activation: 3.1899761659381,3.2390638796335,1.6198443551348,1.6688924071818,0.00070398037743537
Layer 1 activation: -0.048364637212142,-0.09731695950681,0.99879738658182,0.99519243656352,0.00070398037743537
Layer 2 activation: 0.23941974535589
Label: 4
Layer 2 blame: 3.7605802546441
Layer 1 blame: 0.034265563935192,0.030880830785197,0.30229749823934,0.30217283267567,0.037652309120906
Layer 0 blame: -0.034225464528331,-0.030734253062208,-0.014821152029072,-0.029594453358372,0.037652309120906
Gradient: -0.034225464528331,-0.030734253062208,-0.014821152029072,-0.029594453358372,0.037652309120906,
        -0.00026738644162759,-0.0002401113520485,-0.00011579025022712,-0.00023120666686228,0.00029415866500708,
        3.7605802546441,
        -0.18187909972301,-0.36596823636331,3.7560577303697,3.7425010265119,0.0026473747070403
Weights: 3.1405465372762,3.1405817778639,1.5706087596032,1.5704216946248,0.0010804922696339,
        6.2831812002618,12.566366781475,6.283183995544,12.566367994919,4.3750600032927e-06,
        0.11800272038892,
        0.007292984743462,0.0045520372897532,0.11794644272062,0.11777772506292,0.010038840608383
Input: 0.01171875
Layer 0 activation: 3.2141775669668,3.2878438885843,1.644239822051,1.7176838195653,0.0010805435398683
Layer 1 activation: -0.072521193719579,-0.14573042069629,0.99730423856202,0.9892314149972,0.0010805435398683
Layer 2 activation: 0.35095921439238
Label: 3.9
Layer 2 blame: 3.5490407856076
Layer 1 blame: 0.025883100303361,0.016155365998941,0.4185967357328,0.41799794988439,0.035628254759366
Layer 0 blame: -0.025814946775527,-0.015982896761636,-0.030715576956057,-0.06117811994514,0.035628254759366
Gradient: -0.025814946775527,-0.015982896761636,-0.030715576956057,-0.06117811994514,0.035628254759366,
        -0.0003025189075257,-0.00018729957142542,-0.00035994816745379,-0.00071693109310711,0.00041751861046133,
        3.5490407856076,
        -0.25738067433174,-0.51720320675489,3.5394734183159,3.5108226382294,0.0038348930936173
Weights: 3.1402883878084,3.1404219488963,1.5703016038336,1.5698099134254,0.0014367748172275,
        6.2831781750727,12.56636490848,6.2831803960623,12.566360825608,8.5502461079059e-06,
        0.153493128245,
        0.0047191780001446,-0.00061999477779568,0.15334117690378,0.15288595144522,0.010077189539319
Input: 0.015625
Layer 0 activation: 3.2384630467939,3.3367714005913,1.6684762975221,1.7661593013255,0.001436908414823
Layer 1 activation: -0.096718961026805,-0.19394189119524,0.99523310369769,0.98097727260728,0.001436908414823
Layer 2 activation: 0.4557592762934
Label: 3.5
Layer 2 blame: 3.0442407237066
Layer 1 blame: 0.014366313850461,-0.001887413351051,0.46680745535157,0.46542163947216,0.030677390776107
Layer 0 blame: -0.014298960711727,0.0018515770690307,-0.045525262292203,-0.090348866898991,0.030677390776107
Gradient: -0.014298960711727,0.0018515770690307,-0.045525262292203,-0.090348866898991,0.030677390776107,
        -0.00022342126112074,2.8930891703605e-05,-0.00071133222331568,-0.0014117010452967,0.00047933423087666,
        3.0442407237066,
        -0.29443579991239,-0.59040580320923,3.0297291438574,2.9863309623017,0.0043742951126407