UCLA CS 145 Homework #1

Kevin Shih 504946083

1. Linear Regression

1.1

1.1.1 (a)

$$X = \begin{bmatrix} 1 & 60 \\ 1 & 70 \\ 1 & 62 \\ 1 & 72 \\ 1 & 65 \end{bmatrix} \quad y = [130\ 155\ 125\ 162\ 150]^T$$

closed form $\Rightarrow \beta = (X^TX)^{-1}X^Ty$

$$\Rightarrow \beta = \begin{bmatrix} 5 & 329 \\ 329 & 21753 \end{bmatrix}^{-1} \begin{bmatrix} 722 \\ 47814 \end{bmatrix} = \begin{bmatrix} 41.5134 & -0.6279 \\ -0.6279 & 0.0095 \end{bmatrix} \begin{bmatrix} 722 \\ 47814 \end{bmatrix}$$

$$= \begin{bmatrix} -47.9771 \\ 2.9239 \end{bmatrix}$$

$$\Rightarrow y = 2.9239x - 47.9771$$

(b) predicted weight: (1) 127.4427
(2) 156.6794
(3) 133.2901
(4) 162.5267
(5) 142.0611

1.2 (Program's outputs are showed below)

(a)

They are different. Closed form solution is using derivative to find Beta that minimize the cost function. Batch and Stochastic gradient descent are methods that minimize the cost function by moving down in the steepest direction. The results of gradient descent methods are nearly impossible to be exactly same (could be very close) as the closed form solution. Since we have to move Beta with a specific step size in every loop, we might either pass or not yet reach the optimized point every time.

(b)

After applying Z-score, beta is changed and the MSE value is slightly bigger. I think this result tells that the features in the dataset might not need to be normalized and maybe those features with larger scale should just weight more. Before applying Z-score to any dataset, we should think if this dataset really need a normalization, and just try. It is no guarantee that applying normalization can make the prediction better.

Closed Form	Closed Form Without Normalization
	Beta:
	[-0.0862246 0.05340575 0.65803045 0.41731923 -0.01772481 0.30069864
	  1.02871152 0.48383363 0.26685697 0.04573456 0.31944742 1.14776959
	  0.29366213 0.41491543 0.85180482 -0.05950309 0.47235562 0.46198106
	  0.00497427 0.0205398 0.41310473 0.98508025 0.15573467 0.8618602
	  0.41974331 -0.06893699 0.33317496 0.27766637 -0.04184791 -0.23599504
	  0.15020297 0.37745027 0.80256455 0.16053288 0.2744667 0.63461071
	  0.74135259 0.56079776 0.94058723 -0.0432542 0.80803615 0.93967722
	  0.12225161 -0.19933624 0.09398732 0.11412993 0.35479619 0.78582876
	  0.38900433 0.11804526 0.67618837 0.70377377 0.05526258 -0.24919095
	  0.87339793 -0.01381723 0.83138416 0.90569236 0.39980648 0.25235308
	  0.69692397 -0.00949757 0.17676599 0.45822485 0.02743899 1.16718165
	  0.04176352 1.01993881 0.56015024 -0.29761224 0.3177761 0.55781578
	  1.1376088 0.55190283 0.4099807 0.91987238 1.34076835 0.53297825
	  0.63648277 0.22140583 0.21469531 -0.00609269 0.82898663 0.46891532
	 -0.25571565 0.1972989 1.38639797 0.87219453 0.65782257 0.54983464
	  1.11698567 0.94267463 0.79030138 0.30055848 0.53288973 0.22873689
	  0.86702876 0.98591924 0.08132528 0.30834368 0.70121488]
	
	MSE: 4.39609786082
		MSE: 4.396098

Batch Gradient	Batch Gradient Without Normalization
	Beta:
	[-0.08594345 0.05340518 0.65802942 0.41731402 -0.01772559 0.30069295
	  1.02870437 0.48382582 0.26685302 0.04572718 0.31944122 1.14776128
	  0.29365595 0.41490743 0.85179929 -0.05951067 0.47235067 0.46197321
	  0.00496676 0.02053584 0.41309724 0.98507597 0.15572622 0.86185534
	  0.41973952 -0.06894273 0.33316978 0.27766126 -0.0418592 -0.23600262
	  0.15020136 0.3774426 0.80255702 0.16052957 0.27446005 0.6346032
	  0.74134789 0.56079204 0.94057789 -0.04325919 0.80802963 0.93967125
	  0.12224943 -0.19934223 0.09398096 0.11412466 0.3547917 0.78582322
	  0.38900247 0.11803897 0.67618267 0.70376827 0.05525392 -0.24919547
	  0.87339385 -0.01382234 0.83137763 0.90568653 0.39980062 0.25234862
	  0.69692104 -0.00950351 0.1767622 0.45821941 0.02743001 1.16717395
	  0.0417592 1.01993327 0.56014208 -0.29761558 0.31777253 0.55781062
	  1.13760227 0.55189413 0.40998181 0.91986843 1.34076133 0.53297445
	  0.63647761 0.22140173 0.21469085 -0.00609649 0.82898084 0.46890806
	 -0.25572232 0.19729144 1.38639266 0.87219318 0.65781746 0.5498277
	  1.11697834 0.9426711 0.79029427 0.30055233 0.53288577 0.22872826
	  0.86702054 0.98591491 0.08132039 0.30833849 0.7012123]
	
	MSE: 4.39609846641
		MSE: (slightly different) 4.396098

Stochastic Gradient	Stochastic Gradient Without Normalization
	Beta:
	[0.21640291 0.05385335 0.6281821 0.40613939 -0.00283132 0.29650405
	  0.98628684 0.46874224 0.3000846 0.02208147 0.31312136 1.12196755
	  0.30442894 0.42872292 0.82981547 -0.07582267 0.47275074 0.44610326
	  0.01148563 0.02472412 0.41351757 0.99373931 0.17089601 0.84567714
	  0.42143774 -0.05960301 0.34174134 0.27879223 -0.04537258 -0.21998727
	  0.15782948 0.38437396 0.76692174 0.15670468 0.27219916 0.61517531
	  0.73090023 0.57276349 0.91817605 -0.04927191 0.79625943 0.92056157
	  0.10253085 -0.19492819 0.10594301 0.127431 0.3550384 0.76817413
	  0.39156388 0.12352769 0.67867606 0.68620512 0.06785142 -0.23254753
	  0.87612768 -0.02132111 0.8197791 0.89446426 0.39018232 0.23136953
	  0.67818345 -0.01871732 0.17514493 0.47286987 0.02958889 1.15843896
	  0.03363486 1.03392577 0.55303743 -0.28599406 0.31221876 0.55896211
	  1.11651168 0.53659549 0.40404441 0.89522891 1.30476557 0.52320146
	  0.63616424 0.21310028 0.20163266 -0.0031994 0.84555349 0.46842042
	 -0.27182358 0.19697907 1.3878699 0.85977204 0.66456278 0.56926253
	  1.10244568 0.92407039 0.779811 0.29942972 0.5026147 0.23073916
	  0.82587376 0.95211263 0.07432128 0.294332 0.66077334]
	
	MSE: 4.39731677199
		MSE: 4.397317

Closed Form With Norm	Closed Form With Normalization Beta: [2.27729720e+01 1.53267685e-01 1.85400036e-01 1.20001101e-01 -5.02894960e-03 8.91855522e-02 2.85477509e-01 1.40249729e-01 7.58001703e-02 1.29653087e-02 9.40114997e-02 3.31501951e-01 8.48405150e-02 1.19998020e-01 2.42101087e-01 -1.70904428e-02 1.37119556e-01 1.35350218e-01 1.41619004e-03 5.96423043e-03 1.15830867e-01 2.84837752e-01 4.40248244e-02 2.49185633e-01 1.20285952e-01 -1.97966211e-02 9.78939759e-02 8.05403060e-02 -1.21241111e-02 -6.77059821e-02 4.42940642e-02 1.07814670e-01 2.27982170e-01 4.72154203e-02 7.98729034e-02 1.82957097e-01 2.10609705e-01 1.62079663e-01 2.74455584e-01 -1.24456123e-02 2.32346197e-01 2.68821067e-01 3.49745502e-02 -5.73174263e-02 2.74558199e-02 3.22366923e-02 1.03219840e-01 2.23792899e-01 1.12445398e-01 3.34223468e-02 1.96611852e-01 2.04171370e-01 1.61259528e-02 -7.12316220e-02 2.51757075e-01 -3.88735810e-03 2.31055679e-01 2.65481860e-01 1.14239087e-01 7.19519080e-02 2.03225977e-01 -2.77922653e-03 5.10043840e-02 1.31478537e-01 7.74623329e-03 3.36781203e-01 1.19518825e-02 2.98145298e-01 1.64253970e-01 -8.57326109e-02 9.04810592e-02 1.57878654e-01 3.30578812e-01 1.58142457e-01 1.17519641e-01 2.66603450e-01 3.90619100e-01 1.54573813e-01 1.82230684e-01 6.25165215e-02 6.11873098e-02 -1.74345404e-03 2.34361003e-01 1.35158424e-01 -7.34879378e-02 5.72871764e-02 4.02966409e-01 2.50642329e-01 1.87572968e-01 1.57855445e-01 3.18225717e-01 2.66144412e-01 2.29911686e-01 8.53225833e-02 1.56706806e-01 6.57087894e-02 2.52781623e-01 2.90068806e-01 2.33284776e-02 9.01229905e-02 2.03998476e-01] MSE: 4.40454594906	MSE: 4.404546
Batch Gradient With Norm	Batch Gradient With Normalization Beta: [2.27729720e+01 1.53267739e-01 1.85399914e-01 1.20000886e-01 -5.02894366e-03 8.91854074e-02 2.85477680e-01 1.40250040e-01 7.58003097e-02 1.29653008e-02 9.40114470e-02 3.31502042e-01 8.48405681e-02 1.19998329e-01 2.42101125e-01 -1.70902481e-02 1.37119434e-01 1.35350247e-01 1.41623417e-03 5.96422142e-03 1.15830819e-01 2.84837842e-01 4.40250695e-02 2.49185581e-01 1.20285893e-01 -1.97965800e-02 9.78940153e-02 8.05404125e-02 -1.21237424e-02 -6.77056875e-02 4.42937347e-02 1.07814877e-01 2.27982264e-01 4.72152991e-02 7.98728987e-02 1.82957235e-01 2.10609573e-01 1.62079661e-01 2.74455648e-01 -1.24455720e-02 2.32346224e-01 2.68821004e-01 3.49745270e-02 -5.73173336e-02 2.74558620e-02 3.22367618e-02 1.03219809e-01 2.23792761e-01 1.12445300e-01 3.34224047e-02 1.96611882e-01 2.04171192e-01 1.61260396e-02 -7.12315575e-02 2.51757064e-01 -3.88730367e-03 2.31055835e-01 2.65482167e-01 1.14239152e-01 7.19519645e-02 2.03226008e-01 -2.77913637e-03 5.10042046e-02 1.31478558e-01 7.74644921e-03 3.36781392e-01 1.19520430e-02 2.98145403e-01 1.64254028e-01 -8.57324829e-02 9.04810535e-02 1.57878788e-01 3.30579127e-01 1.58142563e-01 1.17519481e-01 2.66603482e-01 3.90619304e-01 1.54573831e-01 1.82230606e-01 6.25164527e-02 6.11873081e-02 -1.74342182e-03 2.34361138e-01 1.35158638e-01 -7.34879734e-02 5.72872124e-02 4.02966351e-01 2.50642156e-01 1.87572946e-01 1.57855537e-01 3.18225596e-01 2.66144407e-01 2.29911630e-01 8.53224135e-02 1.56706756e-01 6.57089316e-02 2.52781599e-01 2.90068475e-01 2.33285484e-02 9.01229590e-02 2.03998339e-01] MSE: 4.40454569497	MSE: (slightly different) 4.404546

<table>
<tr>
<td>Stochastic Gradient With Norm</td>
<td><pre>
Stochastic Gradient With Normalization
Beta:
[2.27720127e+01   1.53872755e-01   1.83562882e-01   1.18619595e-01
  -5.15521208e-03   8.78790173e-02   2.85940656e-01   1.41687764e-01
   7.78379442e-02   1.29627580e-02   9.44404869e-02   3.32629543e-01
   8.61308812e-02   1.23567123e-01   2.42052567e-01 -1.62040103e-02
   1.37452158e-01   1.35726845e-01   3.61517950e-03   5.77649038e-03
   1.16648267e-01   2.85637895e-01   4.66129068e-02   2.48533499e-01
   1.19057912e-01 -1.87344161e-02   9.82311352e-02   8.28218219e-02
  -7.97507684e-03 -6.57658265e-02   4.35542099e-02   1.10764100e-01
   2.28422202e-01   4.58663923e-02   8.14041165e-02   1.83074992e-01
   2.09641770e-01   1.63105930e-01   2.75251487e-01 -1.27531983e-02
   2.34056852e-01   2.67489813e-01   3.29186533e-02 -5.72565221e-02
   2.81713059e-02   3.34824768e-02   1.03131480e-01   2.23093908e-01
   1.12449007e-01   3.59415418e-02   1.96768153e-01   2.03611135e-01
   1.73674765e-02 -7.07734698e-02   2.51828802e-01 -4.04079260e-03
   2.32734393e-01   2.66439279e-01   1.14351731e-01   7.14609223e-02
   2.01928976e-01 -3.20751988e-03   4.96974545e-02   1.32864631e-01
   1.08524544e-02   3.38902190e-01   1.36401894e-02   2.99465956e-01
   1.65270891e-01 -8.43715138e-02   9.16611310e-02   1.59308114e-01
   3.32530094e-01   1.58697611e-01   1.15659063e-01   2.66407018e-01
   3.91138302e-01   1.54052907e-01   1.81983592e-01   6.24051582e-02
   6.03888021e-02 -1.47862574e-03   2.35955691e-01   1.36626458e-01
  -7.33049947e-02   5.73694561e-02   4.04504134e-01   2.47696426e-01
   1.88714912e-01   1.59096641e-01   3.19252040e-01   2.65706869e-01
   2.30369997e-01   8.64933291e-02   1.55986982e-01   6.78409400e-02
   2.52658627e-01   2.87726558e-01   2.35753603e-02   9.06621874e-02
   2.01994991e-01]

MSE: 4.40412222937
</pre></td>
<td>MSE: 4.4041222</td>
</tr>
</table>

2. Logistic Regression and Model Selection

$$x = \begin{bmatrix} 1 & 60 & 155 \\ 1 & 64 & 135 \\ 1 & 73 & 170 \end{bmatrix},\ x^T = \begin{bmatrix} 1 & 1 & 1 \\ 60 & 64 & 73 \\ 155 & 135 & 170 \end{bmatrix},\ y = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

(a)

$$L = y_1x_{11}\beta_1 + y_1x_{12}\beta_2 + y_1x_{13}\beta_3 + y_2x_{21}\beta_1 + y_2x_{22}\beta_2 + y_2x_{23}\beta_3 + y_3x_{31}\beta_1 + y_3x_{32}\beta_2 + y_3x_{33}\beta_3 - \log_2(1 + e^{x_{11}\beta_1 + x_{12}\beta_2 + x_{13}\beta_3}) - \log_2(1 + e^{x_{21}\beta_1 + x_{22}\beta_2 + x_{23}\beta_3}) - \log_2(1 + e^{x_{31}\beta_1 + x_{32}\beta_2 + x_{33}\beta_3})$$

(b)

$$\beta_{1\,new} = \beta_1 + \eta\left(x_{11}y_1 + x_{21}y_2 + x_{31}y_3 - \frac{e^{\beta_1x_{11} + \beta_2x_{12} + \beta_3x_{13}}}{1 + e^{\beta_1x_{11} + \beta_2x_{12} + \beta_3x_{13}}} - \frac{e^{\beta_1x_{21} + \beta_2x_{22} + \beta_3x_{23}}}{1 + e^{\beta_1x_{21} + \beta_2x_{22} + \beta_3x_{23}}} - \frac{e^{\beta_1x_{31} + \beta_2x_{32} + \beta_3x_{33}}}{1 + e^{\beta_1x_{31} + \beta_2x_{32} + \beta_3x_{33}}}\right)$$

$$\beta_{2\,new} = \beta_2 + \eta(x_{12}y_1 + x_{22}y_2 + x_{32}y_3 - \frac{e^{\beta_1 x_{11}+\beta_2 x_{12}+\beta_3 x_{13}}}{1 + e^{\beta_1 x_{11}+\beta_2 x_{12}+\beta_3 x_{13}}}$$

$$- \frac{e^{\beta_1 x_{21}+\beta_2 x_{22}+\beta_3 x_{23}}}{1 + e^{\beta_1 x_{21}+\beta_2 x_{22}+\beta_3 x_{23}}} - \frac{e^{\beta_1 x_{31}+\beta_2 x_{32}+\beta_3 x_{33}}}{1 + e^{\beta_1 x_{31}+\beta_2 x_{32}+\beta_3 x_{33}}})$$

$$\beta_{3\,new} = \beta_3 + \eta(x_{13}y_1 + x_{23}y_2 + x_{33}y_3 - \frac{e^{\beta_1 x_{11}+\beta_2 x_{12}+\beta_3 x_{13}}}{1 + e^{\beta_1 x_{11}+\beta_2 x_{12}+\beta_3 x_{13}}}$$

$$- \frac{e^{\beta_1 x_{21}+\beta_2 x_{22}+\beta_3 x_{23}}}{1 + e^{\beta_1 x_{21}+\beta_2 x_{22}+\beta_3 x_{23}}} - \frac{e^{\beta_1 x_{31}+\beta_2 x_{32}+\beta_3 x_{33}}}{1 + e^{\beta_1 x_{31}+\beta_2 x_{32}+\beta_3 x_{33}}})$$

(c)

$$e_i = \frac{e^{\beta^T x_i}}{\left(1 + e^{\beta^T x_i}\right)} \frac{1}{\left(1 + e^{\beta^T x_i}\right)} = \frac{e^{\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3}}}{1 + e^{\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3}}} * \frac{1}{1 + e^{\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3}}}$$

$$H_{11} = -\left(\sum_{i=1}^{N} x_{i1} x_{i1} e_i\right) = -(x_{11}x_{11}e_1 + x_{21}x_{21}e_2 + x_{31}x_{31}e_3)$$

$$H_{12} = -\left(\sum_{i=1}^{N} x_{i1} x_{i2}\, e_i\right) = -(x_{11}x_{12}e_1 + x_{21}x_{22}e_2 + x_{31}x_{32}e_3)$$

$$H_{13} = -\left(\sum_{i=1}^{N} x_{i1} x_{i3}\, e_i\right) = -(x_{11}x_{13}e_1 + x_{21}x_{23}e_2 + x_{31}x_{33}e_3)$$

$$H_{21} = -\left(\sum_{i=1}^{N} x_{i2} x_{i1}\, e_i\right) = -(x_{12}x_{11}e_1 + x_{22}x_{21}e_2 + x_{32}x_{31}e_3)$$

$$H_{22} = -\left(\sum_{i=1}^{N} x_{i2} x_{i2}\, e_i\right) = -(x_{12}x_{12}e_1 + x_{22}x_{22}e_2 + x_{32}x_{32}e_3)$$

$$H_{23} = -\left(\sum_{i=1}^{N} x_{i2} x_{i3}\, e_i\right) = -(x_{12}x_{13}e_1 + x_{22}x_{23}e_2 + x_{32}x_{33}e_3)$$

$$H_{31} = -\left(\sum_{i=1}^{N} x_{i3} x_{i1}\, e_i\right) = -(x_{13}x_{11}e_1 + x_{23}x_{21}e_2 + x_{33}x_{31}e_3)$$

$$H_{32} = -\left(\sum_{i=1}^{N} x_{i3} x_{i2}\, e_i\right) = -(x_{13}x_{12}e_1 + x_{23}x_{22}e_2 + x_{33}x_{32}e_3)$$

$$H_{33} = -\left(\sum_{i=1}^{N} x_{i3} x_{i3}\, e_i\right) = -(x_{13}x_{13}e_1 + x_{23}x_{23}e_2 + x_{33}x_{33}e_3)$$

3. Decision Tree

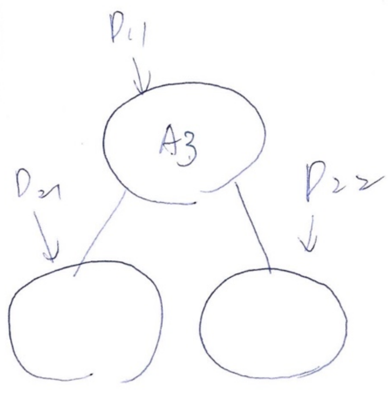
3.1

$$Info(D_{11}) = -\frac{1}{2}(\log_2 \frac{1}{2}) - \frac{1}{2}(\log_2 \frac{1}{2})$$
$$= \frac{1}{2} + \frac{1}{2} = 1$$

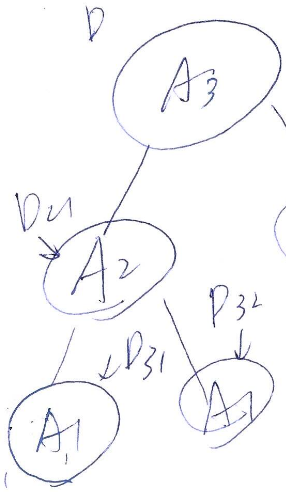
$$Info_{A1}(D_{11}) = \frac{8}{20} Info(D_{21}) + \frac{12}{20} Info(D_{22})$$
$$= \frac{8}{20}\left(-\frac{6}{8}(\log_2 \frac{3}{4}) - \frac{2}{8}(\log_2 \frac{1}{4})\right) + \frac{12}{20}\left(-\frac{4}{12}(\log_2 \frac{1}{3}) - \frac{8}{12}\log_2 \frac{2}{3}\right.$$
$$= 0.87549$$

$$Info_{A2}(D_{11}) = \frac{1}{2}\left(-\frac{3}{5}(\log_2 \frac{3}{5}) - \frac{2}{5}\log_2(\frac{2}{5})\right) + \frac{1}{2}\left(-\frac{2}{5}\log_2(\frac{2}{5}) - \frac{3}{5}\log_2(\frac{3}{5})\right)$$
$$= 0.97095$$

$$Info_{A3}(D_{11}) = \frac{11}{20}\left(-\frac{9}{11}\log_2(\frac{9}{11}) - \frac{2}{11}\log_2(\frac{2}{11})\right)$$
✓
$$+ \frac{9}{20}\left(-\frac{1}{9}\log_2(\frac{1}{9}) - \frac{8}{9}\log_2(\frac{8}{9})\right) = \underline{0.60269}$$



$$Info(D_{21}) = 0.68404$$
$$Info_{A2}(D_{21}) = \frac{4}{11}\left(-\frac{3}{4}\log_2(\frac{3}{4}) - \frac{1}{4}\log_2(\frac{1}{4})\right) +$$
$$\frac{7}{11}\left(-\frac{6}{7}\log(\frac{6}{7}) - \frac{1}{7}\log_2(\frac{1}{7})\right)$$
$$= \underline{0.67152}$$
$$Info_{A1}(D_{21}) = 0.67152$$



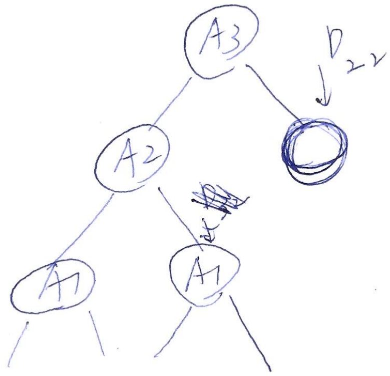
$$Info(D_{31}) = -\frac{3}{4}\left(\log\frac{3}{4}\right) - \frac{1}{4}\left(\log\frac{1}{4}\right) = 0.81128$$

$$Info_{A_1}(D_{31}) = \frac{1}{2}\left(-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2}\right) + \frac{1}{2}\left(\cancel{\log 1}\right)$$

$$= \underline{0.5}$$

$$Info(D_{32}) = -\frac{6}{7}\log\left(\frac{6}{7}\right) - \frac{1}{7}\log\left(\frac{1}{7}\right) = 0.59168$$

$$Info_{A_1}(D_{32}) = \frac{5}{7}\left(-\frac{4}{5}\log\left(\frac{4}{5}\right) - \frac{1}{5}\log\frac{1}{5}\right) + \frac{2}{7}\left(\cancel{\frac{2}{2}\log 1}\right) = \underline{0.51567}$$



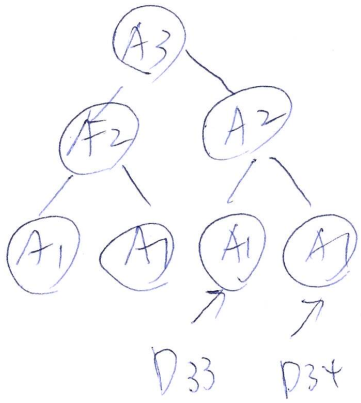
$$Info(D_{22}) = -\frac{1}{9}\log\left(\frac{1}{9}\right) - \frac{8}{9}\log\left(\frac{8}{9}\right) = 0.50326$$

$$Info_{A_1}(D_{22}) = \frac{1}{9}\left(-\frac{0}{1}\log\frac{0}{1} - \frac{1}{1}\log 1\right) + \frac{8}{9}\left(-\frac{1}{8}\log\frac{1}{8} - \frac{7}{8}\log\frac{7}{8}\right)$$

$$= 0.48317$$

$$Info_{A_2}(D_{22}) = \frac{6}{9}\left(-\frac{1}{6}\log\frac{1}{6} - \frac{5}{6}\log\frac{5}{6}\right) + \frac{3}{9} \times (0)$$

$$= \underline{0.43333}$$



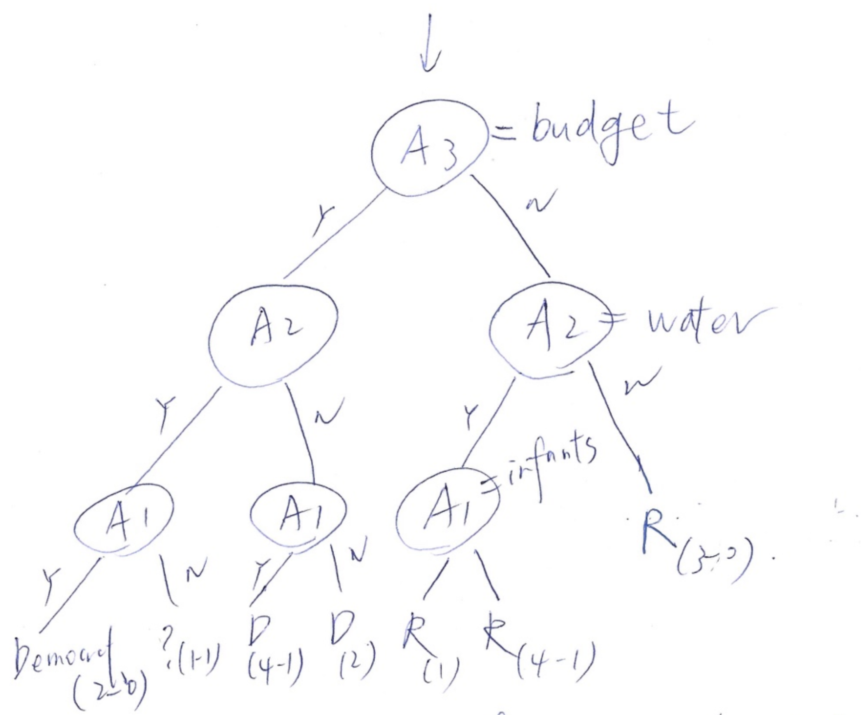
$$Info(D_{33}) = -\frac{1}{6}\log\frac{1}{6} - \frac{5}{6}\log\frac{5}{6} = 0.65000$$

$$Info_{A_1}(D_{33}) = \frac{1}{6} \times (0) + \frac{5}{6}\left(-\frac{1}{5}\log\left(\frac{1}{5}\right) - \frac{4}{5}\log\left(\frac{4}{5}\right)\right)$$

$$= 0.60161$$

$$Info(D_{34}) = \cancel{-\frac{0}{3}\log 0} - \frac{3}{3}\log 1 = \underline{0}$$

$$Info_{A_1}(D_{34}) = \frac{0}{3}(\text{——}) + \frac{3}{3}(\cancel{\log} 0) = \underline{0}$$



use majority when a leaf contains D and R.

if the amounts are equal, label the leaf as "?"

(b)

For house-votes-84, I pick the information gain method. Although there are "y, n, ?" in attributes, "?"s are defined as missing values, it would be ok if it makes the measure biased.(since the values already lost in the first place)

For tic-tac-toe, information gain measure might be biased since almost every attribute has 3 values, so use the gain ratio measure instead.