C:\Users\whitn\Anaconda2\envs\ml4qs2\python.exe C:/Users/whitn/OneDrive/Documenten/Groupwork\_TommyErik/ML4QS/ML4QS-master/PythonCode/crowdsignals\_ch8\_REGRESSION\_question2.py

Training set length is: 541

Test set length is: 112

Training set length is: 541

Test set length is: 112

#basic features: 8

#PCA features: 7

#time features: 320

#frequency features: 432

#cluster features: 1

These were the selected\_feautures returned by Forward Selection:

['pca\_2', 'pca\_7', 'pca\_1', 'mag\_phone\_z\_freq\_0.3\_Hz\_ws\_10', 'acc\_phone\_y\_temp\_mean\_ws\_30\_freq\_0.5\_Hz\_ws\_10', 'pca\_5\_temp\_std\_ws\_30\_freq\_0.2\_Hz\_ws\_10', 'pca\_2\_freq\_0.0\_Hz\_ws\_10', 'pca\_3\_temp\_std\_ws\_30\_freq\_weighted', 'mag\_phone\_z\_freq\_0.2\_Hz\_ws\_10', 'acc\_phone\_z\_temp\_mean\_ws\_30\_freq\_0.0\_Hz\_ws\_10']

These were the ordered\_feautures returned by Forward Selection:

['pca\_2', 'pca\_7', 'pca\_1', 'mag\_phone\_z\_freq\_0.3\_Hz\_ws\_10', 'acc\_phone\_y\_temp\_mean\_ws\_30\_freq\_0.5\_Hz\_ws\_10', 'pca\_5\_temp\_std\_ws\_30\_freq\_0.2\_Hz\_ws\_10', 'pca\_2\_freq\_0.0\_Hz\_ws\_10', 'pca\_3\_temp\_std\_ws\_30\_freq\_weighted', 'mag\_phone\_z\_freq\_0.2\_Hz\_ws\_10', 'acc\_phone\_z\_temp\_mean\_ws\_30\_freq\_0.0\_Hz\_ws\_10']

These were the ordered\_scores returned by Forward Selection:

[0.25376797649577815, 0.14046827929593075, 0.019228018154642035, 0.01149144714815113, 0.0096296810974533225, 0.0080865726621685667, 0.0072709282856087539, 0.0063059250361838237, 0.0059458059208417294, 0.005616124749405549]

\_\_\_\_\_\_\_\_

We will go ahead with this top 10 of features, which we define as selected\_features in the remainder of this script

**['pca\_2', 'pca\_7', 'pca\_1', 'mag\_phone\_z\_freq\_0.3\_Hz\_ws\_10', 'acc\_phone\_y\_temp\_mean\_ws\_30\_freq\_0.5\_Hz\_ws\_10', 'pca\_5\_temp\_std\_ws\_30\_freq\_0.2\_Hz\_ws\_10', 'pca\_2\_freq\_0.0\_Hz\_ws\_10', 'pca\_3\_temp\_std\_ws\_30\_freq\_weighted', 'mag\_phone\_z\_freq\_0.2\_Hz\_ws\_10', 'acc\_phone\_z\_temp\_mean\_ws\_30\_freq\_0.0\_Hz\_ws\_10']**

**Rerunning with edited timestamps gave these: ['pca\_2', 'pca\_7', 'pca\_1', 'mag\_phone\_x\_freq\_0.5\_Hz\_ws\_10', 'acc\_phone\_z\_freq\_0.2\_Hz\_ws\_10', 'pca\_5\_temp\_std\_ws\_30\_freq\_0.3\_Hz\_ws\_10', 'pca\_5\_temp\_mean\_ws\_30\_freq\_0.0\_Hz\_ws\_10', 'pca\_5\_max\_freq', 'acc\_phone\_y\_temp\_std\_ws\_30\_freq\_0.3\_Hz\_ws\_10', 'acc\_phone\_y\_freq\_0.0\_Hz\_ws\_10']**

(-13.931505502561807, 5.0796508380405437e-26, 8L, 644L, {'5%': -2.866038402294262, '1%': -3.440544963888275, '10%': -2.569165592859072}, 1485.5683836336104)

---- 0

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[100, True, 250]

-------

---- 1

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[100, True, 500]

-------

---- 2

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[100, True, 250]

-------

---- 3

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[50, True, 250]

-------

---- 4

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[100, True, 500]

-------

[0, 0, 1]

Normal ARIMAX(0,0,0)

======================================================= ==================================================

Dependent Variable: gyr\_phone\_x Method: MLE

Start Date: 1970-01-01 00:00:00 Log Likelihood: -462.833

End Date: 1970-01-01 00:06:00 AIC: 945.666

Number of observations: 378 BIC: 985.015

==========================================================================================================

Latent Variable Estimate Std Error z P>|z| 95% C.I.

======================================== ========== ========== ======== ======== =========================

Beta 1 0.3858 0.2757 1.3997 0.1616 (-0.1544 | 0.9261)

Beta acc\_phone\_x 0.0549 0.0215 2.5568 0.0106 (0.0128 | 0.097)

Beta acc\_phone\_y -0.0153 0.0127 -1.2073 0.2273 (-0.0402 | 0.0096)

Beta acc\_phone\_z 0.131 0.0193 6.7714 0.0 (0.0931 | 0.1689)

Beta gyr\_phone\_y 0.7845 0.0611 12.8389 0.0 (0.6647 | 0.9042)

Beta gyr\_phone\_z 0.4105 0.1856 2.2116 0.027 (0.0467 | 0.7742)

Beta mag\_phone\_x 0.0042 0.0037 1.1393 0.2546 (-0.003 | 0.0115)

Beta mag\_phone\_y 0.0113 0.0066 1.7101 0.0872 (-0.0016 | 0.0242)

Beta mag\_phone\_z -0.0096 0.003 -3.2078 0.0013 (-0.0155 | -0.0037)

Normal Scale 0.8232

==========================================================================================================

[0, 5, 1]

Normal ARIMAX(0,0,5)

======================================================= ==================================================

Dependent Variable: gyr\_phone\_x Method: MLE

Start Date: 1970-01-01 00:00:00 Log Likelihood: -378.5108

End Date: 1970-01-01 00:06:00 AIC: 787.0215

Number of observations: 373 BIC: 845.8452

==========================================================================================================

Latent Variable Estimate Std Error z P>|z| 95% C.I.

======================================== ========== ========== ======== ======== =========================

MA(1) 0.2986 0.0502 5.9459 0.0 (0.2002 | 0.397)

MA(2) -0.0504 0.0431 -1.1699 0.2421 (-0.1348 | 0.034)

MA(3) -0.4461 0.0436 -10.2321 0.0 (-0.5316 | -0.3606)

MA(4) -0.5081 0.0453 -11.2166 0.0 (-0.5969 | -0.4193)

MA(5) -0.2958 0.046 -6.4261 0.0 (-0.386 | -0.2056)

Beta 1 0.0962 0.1441 0.668 0.5042 (-0.1861 | 0.3786)

Beta acc\_phone\_x -0.0134 0.006 -2.2368 0.0253 (-0.0252 | -0.0017)

Beta acc\_phone\_y 0.0109 0.0077 1.4193 0.1558 (-0.0042 | 0.026)

Beta acc\_phone\_z 0.044 0.0141 3.1159 0.0018 (0.0163 | 0.0717)

Beta gyr\_phone\_y 0.7892 0.0423 18.6356 0.0 (0.7062 | 0.8722)

Beta gyr\_phone\_z 0.0353 0.1375 0.257 0.7972 (-0.2342 | 0.3048)

Beta mag\_phone\_x 0.0004 0.0009 0.4258 0.6702 (-0.0013 | 0.0021)

Beta mag\_phone\_y 0.0069 0.0028 2.4197 0.0155 (0.0013 | 0.0124)

Beta mag\_phone\_z -0.0007 0.0004 -1.7723 0.0763 (-0.0016 | 0.0001)

Normal Scale 0.6395

==========================================================================================================

[5, 0, 1]

Normal ARIMAX(5,0,0)

======================================================= ==================================================

Dependent Variable: gyr\_phone\_x Method: MLE

Start Date: 1970-01-01 00:00:00 Log Likelihood: -349.4968

End Date: 1970-01-01 00:06:00 AIC: 728.9936

Number of observations: 373 BIC: 787.8172

==========================================================================================================

Latent Variable Estimate Std Error z P>|z| 95% C.I.

======================================== ========== ========== ======== ======== =========================

AR(1) 0.366 0.0375 9.7548 0.0 (0.2925 | 0.4395)

AR(2) 0.0751 0.0352 2.1352 0.0327 (0.0062 | 0.1441)

AR(3) -0.0653 0.0352 -1.8577 0.0632 (-0.1342 | 0.0036)

AR(4) -0.144 0.0356 -4.0478 0.0001 (-0.2137 | -0.0742)

AR(5) -0.0263 0.0355 -0.7405 0.459 (-0.0959 | 0.0433)

Beta 1 0.0393 0.2174 0.1809 0.8564 (-0.3867 | 0.4653)

Beta acc\_phone\_x -0.0125 0.017 -0.7342 0.4628 (-0.0458 | 0.0209)

Beta acc\_phone\_y 0.0009 0.0097 0.095 0.9243 (-0.0181 | 0.0199)

Beta acc\_phone\_z 0.0706 0.0152 4.6519 0.0 (0.0409 | 0.1004)

Beta gyr\_phone\_y 0.8648 0.0504 17.1733 0.0 (0.7661 | 0.9635)

Beta gyr\_phone\_z -0.1316 0.1542 -0.8534 0.3934 (-0.4339 | 0.1707)

Beta mag\_phone\_x 0.0015 0.0029 0.5244 0.6 (-0.0042 | 0.0073)

Beta mag\_phone\_y 0.0039 0.0051 0.7596 0.4475 (-0.0061 | 0.0139)

Beta mag\_phone\_z 0.0004 0.0025 0.1456 0.8842 (-0.0045 | 0.0052)

Normal Scale 0.6176

==========================================================================================================

[5, 5, 1]

Normal ARIMAX(5,0,5)

======================================================= ==================================================

Dependent Variable: gyr\_phone\_x Method: MLE

Start Date: 1970-01-01 00:00:00 Log Likelihood: -333.0099

End Date: 1970-01-01 00:06:00 AIC: 706.0198

Number of observations: 373 BIC: 784.4513

==========================================================================================================

Latent Variable Estimate Std Error z P>|z| 95% C.I.

======================================== ========== ========== ======== ======== =========================

AR(1) 0.3796 0.0488 7.7838 0.0 (0.284 | 0.4751)

AR(2) 0.1072 0.0495 2.1644 0.0304 (0.0101 | 0.2043)

AR(3) -0.0284 0.0497 -0.5719 0.5674 (-0.1258 | 0.069)

AR(4) -0.1598 0.0513 -3.1168 0.0018 (-0.2602 | -0.0593)

AR(5) 0.0073 0.0558 0.1315 0.8954 (-0.1019 | 0.1166)

MA(1) -0.1529 0.0763 -2.0047 0.045 (-0.3024 | -0.0034)

MA(2) -0.1907 0.076 -2.5095 0.0121 (-0.3397 | -0.0418)

MA(3) -0.2185 0.0741 -2.9471 0.0032 (-0.3638 | -0.0732)

MA(4) -0.0464 0.0854 -0.5426 0.5874 (-0.2138 | 0.1211)

MA(5) -0.0286 0.0706 -0.4053 0.6853 (-0.1671 | 0.1098)

Beta 1 0.0667 0.1478 0.4511 0.6519 (-0.223 | 0.3563)

Beta acc\_phone\_x -0.0092 0.0106 -0.8662 0.3864 (-0.03 | 0.0116)

Beta acc\_phone\_y -0.0048 0.0086 -0.561 0.5748 (-0.0217 | 0.0121)

Beta acc\_phone\_z 0.0571 0.015 3.8071 0.0001 (0.0277 | 0.0865)

Beta gyr\_phone\_y 0.8444 0.0494 17.1092 0.0 (0.7477 | 0.9412)

Beta gyr\_phone\_z -0.0911 0.1493 -0.6101 0.5418 (-0.3838 | 0.2016)

Beta mag\_phone\_x 0.0009 0.0014 0.6483 0.5168 (-0.0018 | 0.0036)

Beta mag\_phone\_y 0.003 0.003 1.0004 0.3171 (-0.0029 | 0.009)

Beta mag\_phone\_z 0.0003 0.001 0.2872 0.7739 (-0.0016 | 0.0022)

Normal Scale 0.5907

==========================================================================================================

-------

[5, 0, 1]

-------

Normal ARIMAX(5,0,0)

======================================================= ==================================================

Dependent Variable: gyr\_phone\_x Method: MLE

Start Date: 1970-01-01 00:00:00 Log Likelihood: -474.0454

End Date: 1970-01-01 00:08:00 AIC: 978.0907

Number of observations: 536 BIC: 1042.3527

==========================================================================================================

Latent Variable Estimate Std Error z P>|z| 95% C.I.

======================================== ========== ========== ======== ======== =========================

AR(1) 0.3739 0.032 11.6714 0.0 (0.3111 | 0.4367)

AR(2) 0.0694 0.03 2.3166 0.0205 (0.0107 | 0.1281)

AR(3) -0.0885 0.0294 -3.009 0.0026 (-0.1461 | -0.0308)

AR(4) -0.16 0.0297 -5.395 0.0 (-0.2182 | -0.1019)

AR(5) -0.0322 0.0299 -1.075 0.2824 (-0.0909 | 0.0265)

Beta 1 -0.0543 0.1649 -0.3292 0.742 (-0.3775 | 0.2689)

Beta acc\_phone\_x 0.0044 0.0147 0.2971 0.7664 (-0.0244 | 0.0331)

Beta acc\_phone\_y -0.0007 0.0078 -0.0859 0.9315 (-0.016 | 0.0146)

Beta acc\_phone\_z 0.052 0.0127 4.0921 0.0 (0.0271 | 0.0769)

Beta gyr\_phone\_y 0.8581 0.0427 20.1017 0.0 (0.7744 | 0.9417)

Beta gyr\_phone\_z -0.1831 0.127 -1.4416 0.1494 (-0.4321 | 0.0658)

Beta mag\_phone\_x 0.0007 0.0021 0.3305 0.741 (-0.0035 | 0.0049)

Beta mag\_phone\_y 0.0011 0.004 0.2637 0.792 (-0.0068 | 0.0089)

Beta mag\_phone\_z 0.0014 0.002 0.7164 0.4737 (-0.0024 | 0.0052)

Normal Scale 0.586

==========================================================================================================

[(4.8701219922226091, 5.5840919129142623, 0.1969092978660269, 0.84027512025379303), (2.4199459172528508, 3.1202339068404354, 0.15714961167793176, 0.83586124338948031)]

initial set & 4.8701 \emph{( 5.5841 )} & 0.1969 \emph{( 0.8403 )} & 2.4199 \emph{( 3.1202 )} & 0.1571 \emph{( 0.8359 )} \\\hline

---- 0

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[50, True, 500]

-------

---- 1

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[50, True, 250]

-------

---- 2

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[100, True, 500]

-------

---- 3

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[50, True, 500]

-------

---- 4

[50, True, 250]

[50, True, 500]

[100, True, 250]

[100, True, 500]

-------

[50, True, 500]

-------

[(5.5040128053538435, 5.971016595839421, 0.11526396697357214, 0.66463316839815367), (0, 0, 0, 0)]

Chapter 3 & 5.5040 \emph{( 5.9710 )} & 0.1153 \emph{( 0.6646 )} & 0.0000 \emph{( 0.0000 )} & 0.0000 \emph{( 0.0000 )} \\\hline

---- 0

[50, True, 250]

