Phase 3

April 5, 2020

```python
[22]: import pandas as pd
      import numpy as np
      import matplotlib.pyplot as plt
      import seaborn as sns
      from sklearn.model_selection import train_test_split
      from sklearn.metrics import accuracy_score
      from sklearn.ensemble import RandomForestClassifier
      from sklearn.naive_bayes import GaussianNB
      from sklearn.neural_network import MLPClassifier
      from sklearn.metrics import confusion_matrix
```

```python
[2]: demo_data = pd.read_csv('nhanes_data//demographic.csv')
     ques_data = pd.read_csv('nhanes_data//questionnaire.csv')
     exam_data = pd.read_csv('nhanes_data//examination.csv')
     lab_data = pd.read_csv('nhanes_data//labs.csv')
     diet_data = pd.read_csv('nhanes_data//diet.csv')
```

```python
[3]: ques_cols = list(ques_data.corr().columns[1:])
     lab_cols = list(lab_data.corr().columns[1:])
```

```python
[4]: df = pd.merge(ques_data, lab_data, on='SEQN')
     df = df.fillna(0)

     train, test = train_test_split(df)
```

```python
[5]: from sklearn.ensemble import RandomForestRegressor
```

```python
[6]: # Cholesterol

     rfr = RandomForestRegressor(n_estimators=101)
     rfr.fit(train[ques_cols],train['LBXSCH'])
```

```
[6]: RandomForestRegressor(bootstrap=True, ccp_alpha=0.0, criterion='mse',
                           max_depth=None, max_features='auto', max_leaf_nodes=None,
                           max_samples=None, min_impurity_decrease=0.0,
                           min_impurity_split=None, min_samples_leaf=1,
                           min_samples_split=2, min_weight_fraction_leaf=0.0,
```

```
            n_estimators=101, n_jobs=None, oob_score=False,
            random_state=None, verbose=0, warm_start=False)
```

[7]: `rfr.score(test[ques_cols], test['LBXSCH'])`

[7]: 0.7572845403672199

[8]: `#dict(zip(rfr.predict(test[ques_cols]), test['LBXSCH']))`

[9]:
```python
#Total Protein Content

rfr2 = RandomForestRegressor(n_estimators=101)
rfr2.fit(train[ques_cols],train['LBXSTP'])
rfr2.score(test[ques_cols], test['LBXSTP'])
```

[9]: 0.8010691329809412

[10]: `#dict(zip(rfr.predict(test[ques_cols]), test['LBXSTP']))`

[10]:
```
{213.64356435643563: 7.3,
 165.62376237623764: 7.4,
 131.25742574257427: 0.0,
 182.25742574257427: 0.0,
 0.0: 0.0,
 143.0891089108911: 6.6,
 198.4851485148515: 7.8,
 184.44554455445544: 7.1,
 209.66336633663366: 7.0,
 200.86138613861385: 7.5,
 170.17821782178217: 7.5,
 187.4851485148515: 6.6,
 159.74257425742573: 0.0,
 203.9108910891089: 6.3,
 144.1980198019802: 6.6,
 194.45544554455446: 6.6,
 187.64356435643563: 6.5,
 197.98019801980197: 6.3,
 185.4851485148515: 6.9,
 185.98019801980197: 7.0,
 153.12871287128712: 0.0,
 178.35643564356437: 7.2,
 178.68316831683168: 7.6,
 185.9108910891089: 7.8,
 176.5148514851485: 6.9,
 170.69306930693068: 7.2,
 191.31683168316832: 7.9,
 217.9108910891089: 7.5,
```

176.22772277227722: 6.9,
164.59405940594058: 6.7,
182.5049504950495: 7.6,
214.23762376237624: 7.6,
167.86138613861385: 7.2,
183.0891089108911: 7.0,
148.33663366336634: 7.5,
96.63366336633663: 7.7,
145.26732673267327: 6.9,
158.11881188118812: 6.5,
1.4455445544554455: 0.0,
164.9108910891089: 0.0,
185.16831683168317: 7.1,
208.37623762376236: 6.4,
178.40594059405942: 0.0,
150.83168316831683: 6.9,
184.46534653465346: 7.6,
152.31683168316832: 0.0,
154.78217821782178: 6.1,
147.0891089108911: 7.4,
207.16831683168317: 6.9,
171.59405940594058: 7.7,
193.33663366336634: 6.7,
144.40594059405942: 7.1,
192.74257425742573: 6.8,
183.30693069306932: 6.7,
209.78217821782178: 9.0,
193.69306930693068: 6.8,
140.0891089108911: 6.5,
178.65346534653466: 7.0,
175.47524752475246: 7.3,
217.6138613861386: 6.4,
138.32673267326732: 6.3,
106.51485148514851: 0.0,
161.05940594059405: 7.4,
173.94059405940595: 6.9,
183.8019801980198: 7.4,
139.13861386138615: 7.1,
175.66336633663366: 7.3,
197.37623762376236: 6.6,
185.66336633663366: 7.3,
205.11881188118812: 7.0,
194.93069306930693: 7.2,
173.84158415841586: 6.2,
174.3960396039604: 6.6,
188.0990099009901: 7.2,
176.65346534653466: 6.8,

196.34653465346534: 6.9,
142.59405940594058: 8.2,
146.02970297029702: 7.2,
186.78217821782178: 7.4,
217.07920792079207: 7.4,
169.53465346534654: 6.9,
137.79207920792078: 6.8,
187.31683168316832: 7.3,
164.06930693069307: 7.0,
72.60396039603961: 7.8,
175.8019801980198: 7.5,
194.3960396039604: 7.5,
201.78217821782178: 6.3,
202.56435643564356: 7.3,
159.52475247524754: 7.7,
142.01980198019803: 8.3,
168.53465346534654: 8.0,
174.01980198019803: 8.2,
156.78217821782178: 7.9,
179.11881188118812: 7.2,
216.24752475247524: 7.1,
166.3960396039604: 7.1,
168.36633663366337: 6.7,
149.990099009901: 6.5,
175.35643564356437: 7.3,
104.98019801980197: 0.0,
147.46534653465346: 7.1,
171.53465346534654: 7.9,
145.3861386138614: 7.1,
191.46534653465346: 6.7,
214.16831683168317: 7.0,
184.3861386138614: 7.3,
172.4851485148515: 6.9,
173.52475247524754: 6.8,
167.44554455445544: 6.8,
154.5148514851485: 7.2,
147.78217821782178: 7.3,
200.8019801980198: 8.0,
180.1089108910891: 7.5,
163.75247524752476: 5.9,
174.46534653465346: 6.8,
191.21782178217822: 7.0,
170.3861386138614: 6.9,
182.990099009901: 8.3,
164.02970297029702: 7.6,
187.1881188118812: 6.1,
162.69306930693068: 6.7,

```
176.54455445544554: 0.0,
172.82178217821783: 7.5,
196.11881188118812: 7.3,
186.30693069306932: 0.0,
190.33663366336634: 0.0,
193.37623762376236: 7.1,
199.35643564356437: 7.8,
151.6039603960396: 6.8,
208.67326732673268: 7.5,
108.57425742574257: 0.0,
187.78217821782178: 7.7,
123.64356435643565: 6.8,
166.54455445544554: 6.9,
216.75247524752476: 7.4,
163.20792079207922: 6.9,
158.78217821782178: 6.4,
177.13861386138615: 7.2,
168.8910891089109: 6.5,
156.86138613861385: 6.3,
150.59405940594058: 7.6,
178.12871287128712: 6.7,
186.3960396039604: 7.7,
210.97029702970298: 7.2,
183.32673267326732: 7.7,
160.95049504950495: 6.5,
165.85148514851485: 8.1,
129.05940594059405: 6.8,
199.41584158415841: 7.1,
221.23762376237624: 6.6,
182.63366336633663: 7.2,
195.15841584158414: 7.0,
175.04950495049505: 6.6,
173.96039603960395: 7.1,
174.96039603960395: 7.9,
174.8118811881188: 7.5,
211.78217821782178: 6.6,
153.36633663366337: 0.0,
205.2871287128713: 6.9,
201.33663366336634: 6.6,
223.6138613861386: 7.0,
176.84158415841586: 7.2,
193.9009900990099: 0.0,
144.78217821782178: 6.4,
181.75247524752476: 7.6,
159.86138613861385: 7.8,
184.04950495049505: 6.3,
127.92079207920793: 7.8,
```

175.17821782178217: 6.7,
175.24752475247524: 0.0,
168.12871287128712: 6.4,
166.95049504950495: 6.4,
173.86138613861385: 6.5,
143.67326732673268: 7.4,
125.54455445544555: 8.0,
189.63366336633663: 7.7,
204.4950495049505: 7.2,
173.35643564356437: 6.9,
141.01980198019803: 8.0,
115.0990099009901: 0.0,
181.03960396039605: 7.1,
195.64356435643563: 0.0,
181.68316831683168: 6.9,
127.13861386138613: 7.6,
170.2970297029703: 6.8,
207.36633663366337: 6.9,
151.12871287128712: 0.0,
190.87128712871288: 6.9,
185.27722772277227: 7.2,
162.93069306930693: 7.6,
199.64356435643563: 7.5,
190.17821782178217: 7.5,
187.23762376237624: 6.8,
184.72277227722773: 7.1,
168.30693069306932: 7.5,
171.56435643564356: 6.9,
217.009900990099: 6.3,
154.4851485148515: 6.5,
230.63366336633663: 7.2,
180.25742574257427: 6.7,
176.33663366336634: 0.0,
162.94059405940595: 7.3,
174.85148514851485: 7.6,
155.40594059405942: 7.3,
217.55445544554456: 7.1,
227.7029702970297: 6.8,
187.0: 6.8,
208.92079207920793: 7.0,
187.5049504950495: 6.6,
183.1881188118812: 7.1,
170.15841584158414: 7.0,
163.96039603960395: 7.1,
191.69306930693068: 7.1,
170.16831683168317: 6.5,
167.34653465346534: 6.9,

```
104.61386138613861: 7.3,
150.62376237623764: 7.2,
168.98019801980197: 7.3,
188.30693069306932: 7.5,
186.45544554455446: 7.9,
176.03960396039605: 7.7,
172.21782178217822: 7.9,
157.82178217821783: 6.5,
214.46534653465346: 6.7,
171.92079207920793: 7.5,
125.85148514851485: 7.1,
187.9009900990099: 7.7,
175.94059405940595: 7.0,
156.7029702970297: 6.1,
87.06930693069307: 7.3,
196.5742574257426: 7.7,
173.65346534653466: 7.9,
154.07920792079207: 8.0,
171.1089108910891: 7.8,
187.88118811881188: 7.8,
169.07920792079207: 7.0,
190.4257425742574: 6.9,
165.16831683168317: 6.0,
167.52475247524754: 6.9,
199.21782178217822: 7.9,
204.77227722772278: 6.8,
135.990099009901: 7.7,
177.04950495049505: 6.6,
166.13861386138615: 6.6,
115.95049504950495: 0.0,
195.6138613861386: 7.0,
154.87128712871288: 0.0,
148.68316831683168: 6.9,
180.92079207920793: 7.3,
187.67326732673268: 6.8,
193.8019801980198: 7.3,
210.64356435643563: 7.4,
130.56435643564356: 7.2,
177.1881188118812: 7.6,
185.8118811881188: 6.7,
206.94059405940595: 6.5,
170.36633663366337: 7.6,
173.5742574257426: 7.0,
151.14851485148515: 0.0,
113.85148514851485: 6.5,
175.86138613861385: 7.2,
177.14851485148515: 7.2,
```

```
140.7128712871287: 6.8,
184.98019801980197: 7.5,
179.01980198019803: 6.2,
95.36633663366337: 0.0,
157.44554455445544: 8.6,
178.1881188118812: 6.3,
123.29702970297029: 7.6,
212.5148514851485: 7.3,
92.1980198019802: 0.0,
171.66336633663366: 7.6,
191.0990099009901: 7.2,
183.9108910891089: 7.3,
165.4257425742574: 6.2,
188.31683168316832: 7.1,
189.27722772277227: 6.6,
189.73267326732673: 6.7,
214.990099009901: 0.0,
181.30693069306932: 7.2,
183.7029702970297: 7.3,
178.74257425742573: 6.2,
216.6138613861386: 7.4,
218.86138613861385: 6.5,
194.8019801980198: 7.1,
182.4257425742574: 7.5,
197.02970297029702: 7.1,
129.21782178217822: 0.0,
151.11881188118812: 6.8,
207.93069306930693: 7.1,
211.30693069306932: 7.2,
162.06930693069307: 7.0,
204.97029702970298: 7.2,
238.25742574257427: 6.9,
176.87128712871288: 7.0,
181.04950495049505: 7.0,
151.1980198019802: 7.1,
203.6138613861386: 6.7,
156.8910891089109: 7.1,
93.15841584158416: 7.0,
132.07920792079207: 8.0,
192.82178217821783: 7.1,
206.87128712871288: 7.3,
148.0891089108911: 7.8,
180.4257425742574: 7.0,
182.62376237623764: 6.5,
154.85148514851485: 7.3,
215.92079207920793: 7.5,
161.0: 7.4,
```

```
191.77227722772278: 6.9,
133.36633663366337: 7.1,
196.98019801980197: 6.8,
145.04950495049505: 7.3,
201.52475247524754: 7.1,
152.009900990099: 0.0,
171.5049504950495: 6.4,
184.1089108910891: 7.5,
113.86138613861387: 7.7,
133.06930693069307: 7.0,
175.0891089108911: 7.9,
186.14851485148515: 6.4,
150.1881188118812: 7.1,
96.18811881188118: 7.7,
138.4851485148515: 7.3,
132.94059405940595: 7.3,
196.7128712871287: 6.8,
176.21782178217822: 6.5,
195.9108910891089: 6.5,
180.06930693069307: 7.1,
186.36633663366337: 6.6,
177.97029702970298: 7.0,
207.31683168316832: 7.3,
176.59405940594058: 7.1,
119.36633663366337: 7.9,
167.69306930693068: 7.7,
173.5148514851485: 7.2,
221.78217821782178: 0.0,
171.26732673267327: 7.4,
169.93069306930693: 7.3,
158.4950495049505: 7.2,
158.8118811881188: 7.9,
152.69306930693068: 7.5,
172.35643564356437: 6.9,
133.4851485148515: 7.4,
176.25742574257427: 7.3,
155.64356435643563: 7.6,
198.53465346534654: 6.8,
108.15841584158416: 7.4,
164.75247524752476: 0.0,
186.8910891089109: 7.2,
165.0891089108911: 7.8,
184.17821782178217: 0.0,
177.5049504950495: 7.3,
137.009900990099: 7.6,
127.43564356435644: 7.4,
177.4950495049505: 7.2,
```

```
209.88118811881188: 7.0,
176.1881188118812: 7.5,
184.30693069306932: 6.6,
155.1881188118812: 7.1,
180.2871287128713: 6.6,
166.94059405940595: 7.1,
199.990099009901: 7.3,
213.009900990099: 7.1,
141.0990099009901: 8.0,
162.2871287128713: 7.1,
190.40594059405942: 7.4,
167.21782178217822: 7.2,
146.83168316831683: 7.5,
163.98019801980197: 6.8,
130.83168316831683: 6.8,
130.17821782178217: 7.6,
185.65346534653466: 7.6,
156.73267326732673: 7.6,
184.78217821782178: 6.8,
147.86138613861385: 7.5,
175.82178217821783: 6.6,
184.37623762376236: 7.1,
206.36633663366337: 7.1,
173.44554455445544: 7.5,
148.07920792079207: 7.7,
183.35643564356437: 7.3,
185.84158415841586: 7.6,
148.47524752475246: 7.5,
127.2871287128713: 8.2,
127.8019801980198: 0.0,
136.95049504950495: 7.9,
155.72277227722773: 7.4,
1.5148514851485149: 0.0,
173.15841584158414: 6.5,
194.54455445544554: 6.7,
199.56435643564356: 7.2,
203.0: 7.5,
185.96039603960395: 7.0,
164.65346534653466: 7.0,
122.57425742574257: 6.5,
163.47524752475246: 6.9,
188.59405940594058: 6.4,
132.3861386138614: 6.6,
167.77227722772278: 7.5,
138.35643564356437: 6.7,
148.4950495049505: 5.5,
179.69306930693068: 7.8,
```

```
150.21782178217822: 7.2,
235.93069306930693: 7.3,
166.14851485148515: 7.9,
184.87128712871288: 7.2,
198.43564356435644: 7.5,
189.1881188118812: 7.6,
185.20792079207922: 6.4,
146.64356435643563: 6.9,
177.009900990099: 0.0,
169.3861386138614: 7.2,
173.55445544554456: 7.0,
229.20792079207922: 7.2,
180.54455445544554: 6.4,
189.3960396039604: 7.8,
177.92079207920793: 7.0,
185.94059405940595: 6.6,
191.35643564356437: 7.0,
192.25742574257427: 7.9,
135.74257425742573: 0.0,
185.009900990099: 6.3,
188.55445544554456: 6.9,
206.43564356435644: 6.8,
189.76237623762376: 7.0,
190.11881188118812: 7.2,
188.23762376237624: 6.0,
170.33663366336634: 8.0,
198.64356435643563: 7.5,
161.17821782178217: 0.0,
138.1980198019802: 7.3,
182.86138613861385: 7.7,
157.05940594059405: 7.6,
146.7128712871287: 6.8,
196.86138613861385: 7.0,
166.01980198019803: 7.8,
204.06930693069307: 6.6,
200.0990099009901: 7.2,
114.57425742574257: 7.5,
203.04950495049505: 7.4,
186.17821782178217: 7.1,
198.6138613861386: 6.8,
142.1881188118812: 7.2,
172.1980198019802: 7.1,
180.4950495049505: 7.4,
114.08910891089108: 0.0,
154.1881188118812: 7.3,
175.0990099009901: 6.3,
154.05940594059405: 6.5,
```

160.83168316831683: 6.8,
174.1881188118812: 7.6,
157.98019801980197: 7.6,
168.63366336633663: 7.6,
185.32673267326732: 0.0,
139.87128712871288: 6.8,
147.77227722772278: 7.4,
174.65346534653466: 7.3,
212.63366336633663: 6.9,
152.0990099009901: 7.2,
138.79207920792078: 6.4,
167.63366336633663: 7.0,
170.82178217821783: 6.6,
167.35643564356437: 7.0,
184.24752475247524: 7.9,
172.33663366336634: 7.4,
177.12871287128712: 7.8,
190.5049504950495: 8.3,
178.66336633663366: 7.4,
192.16831683168317: 6.1,
156.0891089108911: 6.5,
188.9108910891089: 6.7,
149.76237623762376: 7.1,
136.03960396039605: 7.1,
185.67326732673268: 6.9,
160.66336633663366: 7.1,
180.76237623762376: 6.5,
151.95049504950495: 7.3,
200.97029702970298: 6.5,
210.58415841584159: 6.7,
220.7029702970297: 7.2,
130.1089108910891: 7.2,
179.1881188118812: 7.2,
170.46534653465346: 7.6,
200.11881188118812: 7.6,
145.02970297029702: 0.0,
221.24752475247524: 6.8,
149.0891089108911: 7.1,
141.62376237623764: 7.2,
150.85148514851485: 7.0,
165.76237623762376: 7.2,
166.35643564356437: 6.4,
177.16831683168317: 7.0,
203.07920792079207: 6.8,
206.68316831683168: 6.8,
184.7029702970297: 7.2,
173.82178217821783: 6.7,

187.68316831683168: 7.4,
174.82178217821783: 7.5,
133.82178217821783: 6.8,
215.82178217821783: 6.8,
172.46534653465346: 7.1,
194.68316831683168: 7.1,
157.34653465346534: 6.7,
138.02970297029702: 0.0,
131.0990099009901: 7.1,
137.6138613861386: 7.3,
197.73267326732673: 7.4,
160.0990099009901: 6.5,
108.4059405940594: 0.0,
179.05940594059405: 6.9,
239.46534653465346: 7.4,
182.66336633663366: 7.7,
116.15841584158416: 8.0,
167.20792079207922: 6.3,
182.23762376237624: 7.0,
147.37623762376236: 6.9,
187.9108910891089: 7.2,
147.53465346534654: 6.7,
146.05940594059405: 7.2,
167.14851485148515: 6.6,
200.8118811881188: 0.0,
128.94059405940595: 7.1,
172.74257425742573: 7.4,
184.93069306930693: 6.4,
159.41584158415841: 7.5,
199.68316831683168: 6.8,
208.4851485148515: 7.3,
190.67326732673268: 7.8,
182.03960396039605: 6.7,
185.4257425742574: 7.3,
183.23762376237624: 7.7,
174.43564356435644: 7.3,
184.83168316831683: 7.4,
166.0: 7.2,
209.24752475247524: 7.0,
142.21782178217822: 7.2,
152.83168316831683: 7.3,
158.5742574257426: 7.4,
194.32673267326732: 6.9,
175.59405940594058: 6.5,
186.27722772277227: 7.0,
142.58415841584159: 7.2,
172.01980198019803: 8.2,

```
138.98019801980197: 7.2,
227.4851485148515: 7.1,
184.5049504950495: 7.7,
186.95049504950495: 6.6,
171.9108910891089: 7.5,
191.53465346534654: 7.4,
185.78217821782178: 0.0,
142.35643564356437: 6.5,
197.3960396039604: 6.6,
159.4851485148515: 6.1,
174.33663366336634: 6.7,
208.16831683168317: 7.3,
182.64356435643563: 7.2,
220.12871287128712: 7.5,
191.86138613861385: 6.7,
178.6138613861386: 7.0,
172.2871287128713: 6.4,
151.58415841584159: 7.1,
139.4257425742574: 7.4,
234.03960396039605: 7.2,
150.65346534653466: 7.3,
229.53465346534654: 7.5,
150.7029702970297: 7.1,
205.990099009901: 8.0,
213.73267326732673: 8.2,
177.6138613861386: 7.3,
130.15841584158414: 7.0,
170.77227722772278: 0.0,
137.07920792079207: 6.8,
170.1980198019802: 7.1,
150.98019801980197: 7.8,
168.2871287128713: 6.5,
129.82178217821783: 7.7,
209.02970297029702: 7.8,
196.7029702970297: 6.7,
193.53465346534654: 6.8,
140.45544554455446: 7.0,
212.6138613861386: 7.0,
126.58415841584159: 6.6,
147.07920792079207: 7.3,
184.2970297029703: 6.7,
173.75247524752476: 6.9,
176.35643564356437: 7.3,
148.01980198019803: 6.7,
181.0: 7.5,
173.62376237623764: 7.3,
140.79207920792078: 7.3,
```

```
191.14851485148515: 7.4,
216.84158415841586: 6.4,
156.6039603960396: 8.6,
173.93069306930693: 6.3,
133.2871287128713: 7.9,
186.22772277227722: 7.1,
131.36633663366337: 6.8,
189.04950495049505: 6.8,
164.8118811881188: 6.5,
159.95049504950495: 7.8,
194.14851485148515: 6.6,
185.35643564356437: 7.8,
154.22772277227722: 6.1,
159.53465346534654: 6.9,
175.5148514851485: 6.7,
127.52475247524752: 7.4,
186.21782178217822: 7.2,
179.2970297029703: 7.3,
182.59405940594058: 0.0,
158.35643564356437: 7.2,
197.72277227722773: 7.2,
135.36633663366337: 7.4,
164.37623762376236: 7.1,
166.009900990099: 6.9,
153.7029702970297: 6.9,
156.03960396039605: 6.5,
159.06930693069307: 6.5,
176.27722772277227: 7.5,
188.0891089108911: 7.1,
162.5049504950495: 7.0,
168.32673267326732: 6.8,
153.96039603960395: 7.6,
193.2970297029703: 6.6,
229.11881188118812: 7.6,
181.47524752475246: 6.5,
188.02970297029702: 8.1,
168.02970297029702: 6.5,
148.66336633663366: 6.9,
189.82178217821783: 6.8,
156.2970297029703: 7.2,
194.0: 6.8,
214.68316831683168: 7.2,
116.91089108910892: 7.1,
175.2970297029703: 0.0,
188.40594059405942: 6.8,
170.72277227722773: 7.4,
150.88118811881188: 7.2,
```

138.82178217821783: 7.3,
193.1089108910891: 6.4,
148.53465346534654: 6.1,
131.8910891089109: 7.3,
195.26732673267327: 7.1,
156.5148514851485: 6.6,
183.1089108910891: 7.2,
162.44554455445544: 7.5,
161.87128712871288: 6.6,
146.79207920792078: 6.8,
155.41584158415841: 6.9,
179.56435643564356: 6.9,
186.26732673267327: 7.1,
196.3960396039604: 7.5,
217.97029702970298: 7.3,
160.92079207920793: 7.4,
153.1089108910891: 7.5,
149.15841584158414: 7.2,
112.98019801980197: 0.0,
172.86138613861385: 6.8,
187.24752475247524: 7.1,
172.78217821782178: 6.9,
148.62376237623764: 7.2,
179.64356435643563: 7.0,
190.85148514851485: 6.9,
176.8019801980198: 6.4,
183.52475247524754: 7.4,
99.63366336633663: 7.0,
173.67326732673268: 7.0,
138.05940594059405: 6.9,
140.17821782178217: 7.2,
187.4950495049505: 6.8,
173.59405940594058: 6.9,
131.97029702970298: 7.0,
189.15841584158414: 7.3,
165.21782178217822: 6.9,
178.25742574257427: 6.7,
185.02970297029702: 6.8,
179.65346534653466: 7.3,
160.22772277227722: 6.2,
208.13861386138615: 7.3,
213.02970297029702: 0.0,
211.8019801980198: 0.0,
86.08910891089108: 0.0,
176.37623762376236: 7.3,
139.93069306930693: 7.8,
196.04950495049505: 7.2,

171.8910891089109: 6.9,
112.77227722772277: 7.5,
188.009900990099: 7.8,
189.36633663366337: 8.4,
177.43564356435644: 6.4,
181.3861386138614: 6.2,
179.23762376237624: 7.4,
147.5148514851485: 7.5,
124.85148514851485: 6.7,
194.5049504950495: 6.4,
167.0: 7.0,
177.8910891089109: 6.1,
186.56435643564356: 7.3,
146.21782178217822: 6.9,
164.64356435643563: 7.1,
188.95049504950495: 6.2,
163.32673267326732: 6.5,
212.88118811881188: 6.5,
175.72277227722773: 7.0,
161.54455445544554: 7.2,
176.20792079207922: 6.9,
175.30693069306932: 7.4,
198.54455445544554: 6.9,
171.67326732673268: 6.9,
148.45544554455446: 7.4,
193.25742574257427: 7.3,
200.65346534653466: 6.5,
207.6039603960396: 6.3,
171.73267326732673: 6.9,
188.67326732673268: 7.0,
178.16831683168317: 7.0,
143.41584158415841: 7.4,
169.0990099009901: 7.5,
193.32673267326732: 7.3,
138.31683168316832: 6.8,
208.72277227722773: 7.6,
180.59405940594058: 6.6,
145.56435643564356: 7.8,
211.43564356435644: 7.1,
175.4257425742574: 6.8,
136.24752475247524: 7.5,
151.53465346534654: 6.9,
162.25742574257427: 5.9,
166.11881188118812: 7.0,
207.06930693069307: 7.3,
183.03960396039605: 6.9,
190.97029702970298: 7.5,

```
185.73267326732673: 6.7,
185.40594059405942: 6.6,
179.3960396039604: 7.3,
160.2970297029703: 6.4,
152.93069306930693: 7.5,
230.88118811881188: 7.3,
220.95049504950495: 7.4,
178.14851485148515: 0.0,
147.88118811881188: 7.0,
129.02970297029702: 6.8,
203.72277227722773: 7.0,
144.26732673267327: 7.1,
181.59405940594058: 7.2,
145.45544554455446: 7.8,
150.0: 6.3,
228.4257425742574: 6.5,
169.87128712871288: 6.9,
156.0: 7.2,
195.85148514851485: 7.8,
184.40594059405942: 7.3,
168.47524752475246: 7.1,
176.88118811881188: 7.4,
181.83168316831683: 6.5,
179.88118811881188: 7.3,
180.8910891089109: 7.0,
149.44554455445544: 7.3,
176.96039603960395: 6.9,
176.1980198019802: 8.3,
227.84158415841586: 6.7,
128.87128712871288: 7.4,
230.14851485148515: 7.0,
160.44554455445544: 6.8,
169.21782178217822: 7.6,
178.26732673267327: 7.0,
191.25742574257427: 6.8,
188.6138613861386: 7.4,
179.4257425742574: 8.0,
192.990099009901: 6.1,
153.69306930693068: 7.1,
163.77227722772278: 6.6,
193.4950495049505: 6.6,
150.87128712871288: 7.0,
184.01980198019803: 8.0,
158.32673267326732: 7.6,
157.4851485148515: 7.1,
137.58415841584159: 8.2,
166.97029702970298: 7.1,
```

```
194.69306930693068: 7.6,
190.16831683168317: 0.0,
187.53465346534654: 7.1,
183.98019801980197: 8.2,
168.93069306930693: 7.6,
181.97029702970298: 7.2,
141.7128712871287: 6.7,
154.16831683168317: 7.0,
154.4257425742574: 7.3,
148.04950495049505: 7.1,
118.9009900990099: 7.3,
147.06930693069307: 7.2,
152.32673267326732: 6.5,
164.95049504950495: 7.0,
166.30693069306932: 6.9,
168.85148514851485: 7.6,
188.34653465346534: 7.4,
121.04950495049505: 6.3,
182.5148514851485: 6.6,
148.97029702970298: 7.0,
150.40594059405942: 6.3,
214.21782178217822: 7.3,
235.6039603960396: 7.0,
186.9108910891089: 7.0,
133.67326732673268: 6.9,
179.44554455445544: 6.6,
199.34653465346534: 8.3,
174.5049504950495: 6.5,
185.59405940594058: 7.7,
178.69306930693068: 7.0,
174.30693069306932: 0.0,
111.48514851485149: 0.0,
176.3861386138614: 7.1,
119.29702970297029: 7.0,
171.47524752475246: 7.6,
118.34653465346534: 6.9,
159.85148514851485: 7.0,
137.59405940594058: 7.0,
183.97029702970298: 7.5,
145.2871287128713: 7.6,
172.07920792079207: 6.9,
199.72277227722773: 6.6,
150.30693069306932: 6.2,
171.65346534653466: 7.4,
202.4257425742574: 7.2,
177.17821782178217: 0.0,
189.40594059405942: 7.2,
```

```
154.25742574257427: 7.1,
176.0891089108911: 6.9,
150.22772277227722: 7.2,
223.13861386138615: 7.3,
224.07920792079207: 7.5,
196.8910891089109: 8.5,
143.79207920792078: 0.0,
146.45544554455446: 7.1,
140.87128712871288: 7.5,
204.62376237623764: 6.8,
178.03960396039605: 6.4,
180.79207920792078: 8.2,
174.79207920792078: 6.4,
200.24752475247524: 6.6,
176.45544554455446: 7.9,
193.35643564356437: 6.5,
151.76237623762376: 5.8,
209.31683168316832: 6.7,
195.4257425742574: 7.9,
167.85148514851485: 7.0,
162.24752475247524: 7.1,
180.36633663366337: 0.0,
173.9009900990099: 0.0,
210.30693069306932: 6.7,
194.35643564356437: 5.9,
148.2871287128713: 5.7,
149.66336633663366: 7.2,
173.92079207920793: 7.3,
183.69306930693068: 6.6,
129.52475247524754: 7.4,
135.59405940594058: 7.6,
199.32673267326732: 6.8,
181.36633663366337: 8.0,
174.009900990099: 6.3,
173.23762376237624: 6.8,
147.26732673267327: 7.3,
158.54455445544554: 6.9,
112.03960396039604: 8.1,
185.93069306930693: 7.0,
182.20792079207922: 7.6,
153.2970297029703: 6.9,
186.3861386138614: 6.9,
165.11881188118812: 7.5,
187.87128712871288: 7.4,
172.96039603960395: 7.1,
150.78217821782178: 0.0,
179.75247524752476: 7.0,
```

```
174.78217821782178: 7.5,
188.65346534653466: 7.1,
175.64356435643563: 7.3,
190.14851485148515: 7.0,
164.0891089108911: 7.9,
184.27722772277227: 7.2,
73.61386138613861: 7.8,
231.03960396039605: 7.3,
172.11881188118812: 7.8,
129.59405940594058: 6.9,
160.93069306930693: 6.6,
191.85148514851485: 7.0,
177.53465346534654: 6.7,
148.05940594059405: 7.8,
219.32673267326732: 7.1,
175.5049504950495: 7.8,
179.45544554455446: 6.9,
114.21782178217822: 7.6,
153.64356435643563: 7.0,
208.94059405940595: 6.9,
175.78217821782178: 7.5,
193.30693069306932: 6.3,
195.55445544554456: 7.6,
145.36633663366337: 6.9,
145.1980198019802: 7.4,
208.8910891089109: 7.2,
124.20792079207921: 6.5,
189.5148514851485: 6.9,
182.27722772277227: 7.3,
139.27722772277227: 7.1,
190.2871287128713: 7.5,
142.9009900990099: 7.4,
145.54455445544554: 6.9,
167.990099009901: 6.8,
200.8910891089109: 7.2,
138.990099009901: 7.1,
184.56435643564356: 7.1,
194.43564356435644: 7.8,
178.92079207920793: 7.0,
199.45544554455446: 6.8,
176.05940594059405: 7.8,
178.20792079207922: 6.7,
185.68316831683168: 7.1,
149.93069306930693: 7.9,
184.67326732673268: 7.1,
135.2871287128713: 7.7,
188.3960396039604: 6.6,
```

```
125.67326732673267: 7.1,
172.69306930693068: 6.9,
182.33663366336634: 7.0,
198.35643564356437: 6.8,
180.95049504950495: 7.0,
184.34653465346534: 7.1,
143.8118811881188: 6.6,
157.30693069306932: 7.6,
119.99009900990099: 6.6,
187.06930693069307: 6.2,
133.58415841584159: 7.6,
153.0891089108911: 7.1,
156.68316831683168: 7.4,
188.68316831683168: 6.8,
135.9009900990099: 6.5,
181.46534653465346: 6.2,
180.26732673267327: 7.0,
214.7128712871287: 7.4,
235.0891089108911: 7.1,
137.0: 7.3,
189.92079207920793: 7.2,
168.55445544554456: 6.3,
177.47524752475246: 7.2,
191.40594059405942: 7.9,
193.55445544554456: 6.8,
173.16831683168317: 6.4,
153.3960396039604: 6.3,
131.17821782178217: 7.4,
141.30693069306932: 7.3,
223.1881188118812: 7.5,
186.54455445544554: 7.3,
134.66336633663366: 6.2,
176.94059405940595: 7.5,
152.06930693069307: 6.7,
211.5148514851485: 7.4,
174.84158415841586: 6.4,
182.8019801980198: 7.1,
194.55445544554456: 6.7,
203.8118811881188: 7.2,
171.52475247524754: 6.2,
139.47524752475246: 7.6,
210.74257425742573: 0.0,
177.78217821782178: 7.1,
190.3960396039604: 6.5,
181.65346534653466: 6.9,
186.43564356435644: 6.7,
187.0990099009901: 7.2,
```

```
       157.6039603960396: 7.4,
       150.0891089108911: 6.9,
       142.06930693069307: 7.9,
       160.4851485148515: 6.9,
       203.03960396039605: 6.9,
       181.63366336633663: 7.2,
       162.02970297029702: 7.7,
       146.0: 6.2,
       160.4950495049505: 6.7,
       191.07920792079207: 8.0,
       205.6039603960396: 7.2,
       172.45544554455446: 7.4,
       192.69306930693068: 7.0,
       184.9009900990099: 7.8,
       208.4950495049505: 7.6,
       141.63366336633663: 6.9,
       159.990099009901: 6.5,
       159.14851485148515: 7.0,
       172.32673267326732: 7.6,
       152.76237623762376: 6.4,
       147.3861386138614: 6.4,
       195.68316831683168: 0.0,
       166.32673267326732: 8.5,
       165.86138613861385: 7.1,
       157.12871287128712: 7.5,
       183.8910891089109: 8.0,
       132.84158415841586: 7.5,
       172.24752475247524: 6.4,
       205.3861386138614: 5.8,
       203.35643564356437: 6.9,
       182.1089108910891: 0.0,
       172.8910891089109: 6.4,
       …}
```

[11]:
```python
diet_cols = list(diet_data.corr().columns[1:])
```

[12]:
```python
df2 = pd.merge(diet_data, lab_data, on='SEQN')
df2 = df2.fillna(0)

train2, test2 = train_test_split(df2)
```

[13]:
```python
df3 = pd.merge(diet_data, ques_data, on='SEQN')
df3 = df3.fillna(0)

train3, test3 = train_test_split(df3)
```

```
[14]: # Hepatitis A

     print('Hepatitis A, RandomForest Classifier \n')
     rfr2 = RandomForestClassifier(n_estimators=100)
     rfr2.fit(train.loc[train['LBXHA']!=2,ques_cols],train.loc[train['LBXHA']!
      ↪=2,'LBXHA'])
     print('Train Score :')
     print(rfr2.score(train.loc[train['LBXHA']!=2,ques_cols],train.
      ↪loc[train['LBXHA']!=2,'LBXHA']))
     print('Test Score :')
     print(rfr2.score(test.loc[test['LBXHA']!=2,ques_cols], test.loc[test['LBXHA']!
      ↪=2,'LBXHA']))
```

Hepatitis A, RandomForest Classifier

Train Score :
1.0
Test Score :
0.8667079975201488

```
[15]: # Hepatitis A

     print('Hepatitis A, MLP Classifier \n')
     rfr2 = MLPClassifier()
     rfr2.fit(train.loc[train['LBXHA']!=2,ques_cols],train.loc[train['LBXHA']!
      ↪=2,'LBXHA'])
     print('Train Score :')
     print(rfr2.score(train.loc[train['LBXHA']!=2,ques_cols],train.
      ↪loc[train['LBXHA']!=2,'LBXHA']))
     print('Test Score :')
     print(rfr2.score(test.loc[test['LBXHA']!=2,ques_cols], test.loc[test['LBXHA']!
      ↪=2,'LBXHA']))
```

Hepatitis A, MLP Classifier

Train Score :
0.9521803244154202
Test Score :
0.7792932424054557

C:\Users\udit2\anaconda3\lib\site-
packages\sklearn\neural_network_multilayer_perceptron.py:571:
ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and
the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)

```
[16]: # Hepatitis B

      print('Hepatitis B, RandomForest Classifier \n')
      rfr2 = RandomForestClassifier(n_estimators=100)
      rfr2.fit(train.loc[train['LBXHBC']!=2,ques_cols],train.loc[train['LBXHBC']!
       ↪=2,'LBXHBC'])
      print('Train Score :')
      print(rfr2.score(train.loc[train['LBXHBC']!=2,ques_cols],train.
       ↪loc[train['LBXHBC']!=2,'LBXHBC']))
      print('Test Score :')
      print(rfr2.score(test.loc[test['LBXHBC']!=2,ques_cols], test.loc[test['LBXHBC']!
       ↪=2,'LBXHBC']))
```

Hepatitis B, RandomForest Classifier

Train Score :
1.0
Test Score :
0.9319620253164557

```
[17]: # Hepatitis B

      print('Hepatitis B, GaussianNB \n')
      rfr2 = GaussianNB()
      rfr2.fit(train.loc[train['LBXHBC']!=2,ques_cols],train.loc[train['LBXHBC']!
       ↪=2,'LBXHBC'])
      print('Train Score :')
      print(rfr2.score(train.loc[train['LBXHBC']!=2,ques_cols],train.
       ↪loc[train['LBXHBC']!=2,'LBXHBC']))
      print('Test Score :')
      print(rfr2.score(test.loc[test['LBXHBC']!=2,ques_cols], test.loc[test['LBXHBC']!
       ↪=2,'LBXHBC']))
```

Hepatitis B, GaussianNB

Train Score :
0.8975535168195719
Test Score :
0.8829113924050633

```
[28]: # Hepatitis B

      print('Hepatitis B, MLP Classifier\n')
      rfr2 = MLPClassifier()
      rfr2.fit(train.loc[train['LBXHBC']!=2,ques_cols],train.loc[train['LBXHBC']!
       ↪=2,'LBXHBC'])
      print('Train Score :')
```

```
print(rfr2.score(train.loc[train['LBXHBC']!=2,ques_cols],train.
 ↪loc[train['LBXHBC']!=2,'LBXHBC']))
print('Test Score :')
print(rfr2.score(test.loc[test['LBXHBC']!=2,ques_cols], test.loc[test['LBXHBC']!
 ↪=2,'LBXHBC']))
```

Hepatitis B, MLP Classifier

Train Score :
0.9709480122324159
Test Score :
0.9018987341772152

C:\Users\udit2\anaconda3\lib\site-
packages\sklearn\neural_network_multilayer_perceptron.py:571:
ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and
the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)

[30]:
```
print(confusion_matrix(test.loc[test['LBXHBC']!=2,'LBXHBC'],rfr2.predict(test.
 ↪loc[test['LBXHBC']!=2,ques_cols])))
```

[[484 36]
 [26 86]]

[19]:
```
# Hepatitis D

print('Hepatitis D, RandomForestClassifier\n')
rfr2 = RandomForestClassifier(n_estimators=100)
rfr2.fit(train.loc[train['LBDHD']!=2,ques_cols],train.loc[train['LBDHD']!
 ↪=2,'LBDHD'])
print('Train Score :')
print(rfr2.score(train.loc[train['LBDHD']!=2,ques_cols],train.
 ↪loc[train['LBDHD']!=2,'LBDHD']))
print('Test Score :')
print(rfr2.score(test.loc[test['LBDHD']!=2,ques_cols], test.loc[test['LBDHD']!
 ↪=2,'LBDHD']))
```

Hepatitis D, RandomForestClassifier

Train Score :
1.0
Test Score :
0.9961685823754789

[20]:
```
# Hepatitis D

print('Hepatitis D,  GaussianNB\n')
```

```
rfr2 = GaussianNB()
rfr2.fit(train.loc[train['LBDHD']!=2,ques_cols],train.loc[train['LBDHD']!
 ↪=2,'LBDHD'])
print('Train Score :')
print(rfr2.score(train.loc[train['LBDHD']!=2,ques_cols],train.
 ↪loc[train['LBDHD']!=2,'LBDHD']))
print('Test Score :')
print(rfr2.score(test.loc[test['LBDHD']!=2,ques_cols], test.loc[test['LBDHD']!
 ↪=2,'LBDHD']))
```

Hepatitis D, GaussianNB

Train Score :
0.8184559710494572
Test Score :
0.8295019157088123

[21]:
```
# Hepatitis D

print('Hepatitis D,  MLP Classifier\n')
rfr2 = MLPClassifier()
rfr2.fit(train.loc[train['LBDHD']!=2,ques_cols],train.loc[train['LBDHD']!
 ↪=2,'LBDHD'])
print('Train Score :')
print(rfr2.score(train.loc[train['LBDHD']!=2,ques_cols],train.
 ↪loc[train['LBDHD']!=2,'LBDHD']))
print('Test Score :')
print(rfr2.score(test.loc[test['LBDHD']!=2,ques_cols], test.loc[test['LBDHD']!
 ↪=2,'LBDHD']))
```

Hepatitis D, MLP Classifier

Train Score :
0.9993968636911942
Test Score :
0.9885057471264368

[27]:
```
confusion_matrix(test.loc[test['LBDHD']!=2,'LBDHD'],rfr2.predict(test.
 ↪loc[test['LBDHD']!=2,ques_cols]))
```

[27]:
```
array([[516,    4],
       [  2,    0]], dtype=int64)
```

[ ]: