AI 3603 Artificial Intelligence: Principles and Techniques

By: Kaiwen Zhu 520030910178

HW#: 3

December 22, 2022

1 Problem 1

The size of the network is $131072 = 8 \times 2 \times 2 \times 2 \times 2 \times 4 \times 4 \times 2 \times 2 \times 2 \times 2 \times 4$, which is the product of the numbers of values every variable can take respectively. The program result is shown in Figure 1.

```
The size of this network is 131072.
```

Figure 1: Program result of problem 1

2 Problem 2

For the four health outcomes (diabetes, stroke, heart attack, angina), the four subtables of Table 1 list the probabilities given bad habits, good habits, poor health or good health. The program result is shown in Figure 2.

Table 1: Probabilities of health outcomes given conditions of habits or health

(a) Probability of diabetes

outcome level	bad habits	good habits	poor health	good health
1	0.179597	0.075195	0.115423	0.057710
2	0.008754	0.009409	0.007662	0.009543
3	0.791160	0.903426	0.860873	0.922194
4	0.020489	0.011970	0.016043	0.010553

(b) Probability of stroke

outcome level	bad habits	good habits	poor health	good health
1	0.053214	0.029202	0.082686	0.01446
2	0.946786	0.970798	0.917314	0.98554

(c) Probability of attack

outcome level	bad habits	good habits	poor health	good health
1	0.085704	0.036655	0.140784	0.016161
2	0.914296	0.963345	0.859216	0.983839

(d) Probability of angina

outcome level	bad habits	good habits	poor health	good health
1	0.09542	0.03551	0.161608	0.013326
2	0.90458	0.96449	0.838392	0.986674

```
Probabilities of diabetes given conditions of habits or health:
   diabetes  bad habits  good habits  poor health  good health
0         1    0.179597     0.075195     0.115423     0.057710
1         2    0.008754     0.009409     0.007662     0.009543
2         3    0.791160     0.903426     0.860873     0.922194
3         4    0.020489     0.011970     0.016043     0.010553

Probabilities of stroke given conditions of habits or health:
   stroke  bad habits  good habits  poor health  good health
0       1    0.053214     0.029202     0.082686      0.01446
1       2    0.946786     0.970798     0.917314      0.98554

Probabilities of attack given conditions of habits or health:
   attack  bad habits  good habits  poor health  good health
0       1    0.085704     0.036655     0.140784     0.016161
1       2    0.914296     0.963345     0.859216     0.983839

Probabilities of angina given conditions of habits or health:
   angina  bad habits  good habits  poor health  good health
0       1     0.09542      0.03551     0.161608     0.013326
1       2     0.90458      0.96449     0.838392     0.986674
```

Figure 2: Program result of problem 2

3 Problem 3

For the four health outcomes (diabetes, stroke, heart attack, angina), Figure 3 shows the probabilities given different income statuses. The program result is shown in Figure 4.

Clearly, the probabilities of health outcomes decrease almost linearly with increasing income status in general. It is easy to understand: a person with more income usually lives a better life, including superior healthcare, less exposure to health risks (like heavy work and harsh environment) and stronger health consciousness.

The only abnormality is that people at income status 2 (whose annual income is between $10000 and $15000) have a higher risk of stroke, attack or angina than those at income status 1. One possible explanation is that these people may be pillars of less affluent families, who laboriously earn not-too-little money, with hard work damaging their health.

4 Problem 4

The assumption is that there are no direct relationships between these habits and health outcomes. That is, if these three health conditions (BMI, blood pressure, and level of choles-

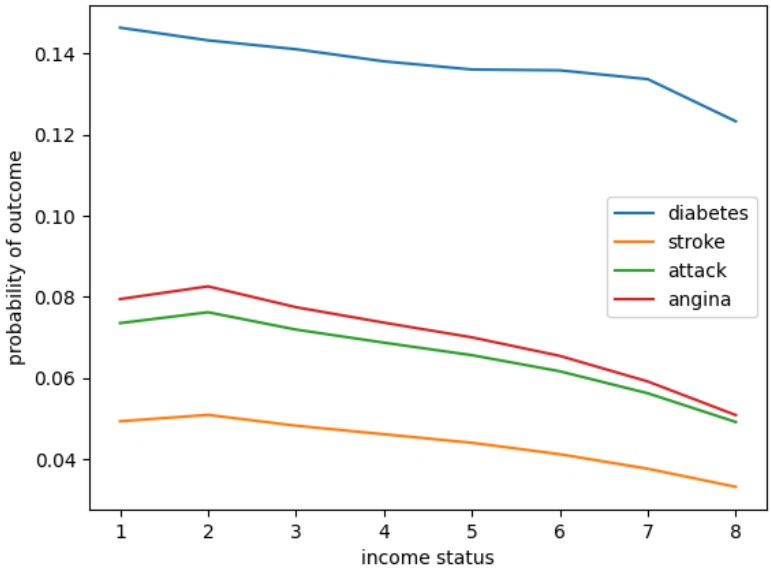


Figure 3: Relations between income and health outcomes

```
Probability of diabetes given different incomes:
   income     probs
0       1  0.146381
1       2  0.143259
2       3  0.141068
3       4  0.138079
4       5  0.136066
5       6  0.135862
6       7  0.133681
7       8  0.123280

Probability of stroke given different incomes:
   income     probs
0       1  0.049305
1       2  0.050885
2       3  0.048218
3       4  0.046113
4       5  0.044010
5       6  0.041183
6       7  0.037597
7       8  0.033132
```

```
Probability of attack given different incomes:
   income     probs
0       1  0.073515
1       2  0.076203
2       3  0.071924
3       4  0.068700
4       5  0.065610
5       6  0.061636
6       7  0.056209
7       8  0.049129

Probability of angina given different incomes:
   income     probs
0       1  0.079443
1       2  0.082579
2       3  0.077446
3       4  0.073632
4       5  0.070007
5       6  0.065431
6       7  0.059120
7       8  0.050829
```

Figure 4: Program result of problem 3

terol) are given, then the probability of outcomes is independent of habits. Mathematically,

$$\Pr(outcome \mid bmi, bp, cholesterol) = \Pr(outcome \mid bmi, bp, cholesterol, smoke, exercise).$$

Now we test the validity of this assumption. Having added edges from smoking and exercise to the four outcomes, Table 2 lists the probabilities of health outcomes given habits or health conditions. The program result is shown in Figure 5.

Figure 6 shows the difference between results before and after adding these edges, which is highly conspicuous when habits are given: assuming direct relationships between the two habits and the outcomes, the probabilities of outcomes are much higher than previously believed given bad habits and lower given good habits. This strongly demonstrates that bad habits of smoking and exercising less will directly induce these health problems, and vice versa. Note that there is almost no difference when health conditions are given because variables `smoke` and `exercise` have been marginalized.

Therefore, the assumption that there is no direct relationships between these habits and health outcomes is not valid.

Table 2: Probabilities of health outcomes given conditions of habits or health

(a) Probability of diabetes

outcome level	bad habits	good habits	poor health	good health
1	0.245992	0.056227	0.121241	0.055937
2	0.006928	0.010160	0.007492	0.009697
3	0.723721	0.923710	0.854769	0.924042
4	0.023359	0.009903	0.016498	0.010323

(b) Probability of stroke

outcome level	bad habits	good habits	poor health	good health
1	0.080488	0.019464	0.082697	0.014544
2	0.919512	0.980536	0.917303	0.985456

(c) Probability of attack

outcome level	bad habits	good habits	poor health	good health
1	0.135301	0.021213	0.140083	0.016183
2	0.864699	0.978787	0.859917	0.983817

(d) Probability of angina

outcome level	bad habits	good habits	poor health	good health
1	0.138072	0.023948	0.161096	0.013328
2	0.861928	0.976052	0.838904	0.986672

```
Probabilities of diabetes given conditions of habits or health:
   diabetes  bad habits  good habits  poor health  good health
0         1    0.245992     0.056227     0.121241     0.055937
1         2    0.006928     0.010160     0.007492     0.009697
2         3    0.723721     0.923710     0.854769     0.924042
3         4    0.023359     0.009903     0.016498     0.010323

Probabilities of stroke given conditions of habits or health:
   stroke  bad habits  good habits  poor health  good health
0       1    0.080488     0.019464     0.082697     0.014544
1       2    0.919512     0.980536     0.917303     0.985456

Probabilities of attack given conditions of habits or health:
   attack  bad habits  good habits  poor health  good health
0       1    0.135301     0.021213     0.140083     0.016183
1       2    0.864699     0.978787     0.859917     0.983817

Probabilities of angina given conditions of habits or health:
   angina  bad habits  good habits  poor health  good health
0       1    0.138072     0.023948     0.161096     0.013328
1       2    0.861928     0.976052     0.838904     0.986672
```

Figure 5: Program result of problem 4

5 Problem 5

The assumption is that there are no direct relationships among health outcomes. That is, if these habits and health conditions are given, then the probability of one outcome is independent of another outcome.

Now we test the validity of this assumption. Having added edges from diabetes to stroke, Figure 7 shows values of

$$\Pr(stroke = 1 \mid diabetes = 1), \Pr(stroke = 1 \mid diabetes = 3)$$

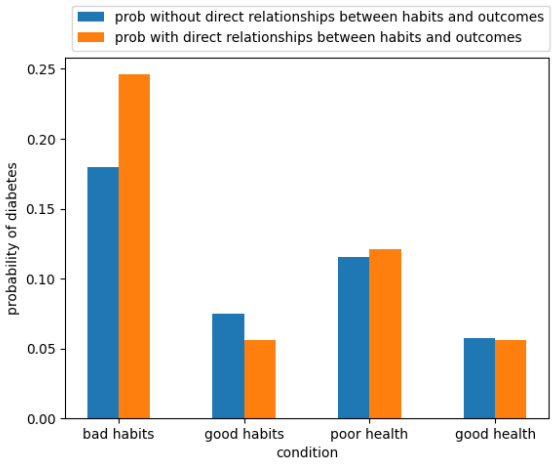
before and after adding these edges. The program result is shown in Figure 8.

Again, the difference is remarkable: assuming direct relationships between stroke and diabetes, the probability of stroke is much higher than previously believed with diabetes and lower without diabetes. This demonstrates that diabetes will directly induce stroke and vice versa.

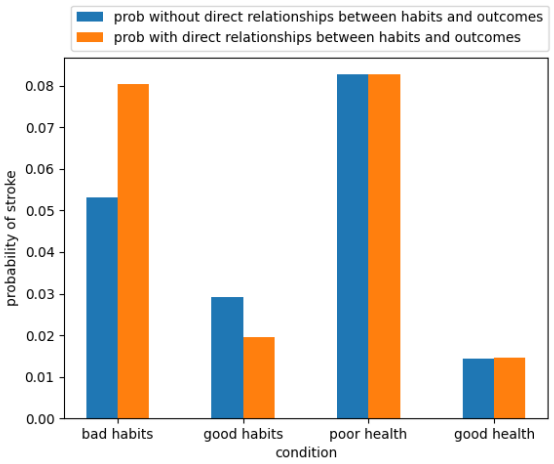
Therefore, the assumption that there are no direct relationships among health outcomes is not valid.

6 Problem 6

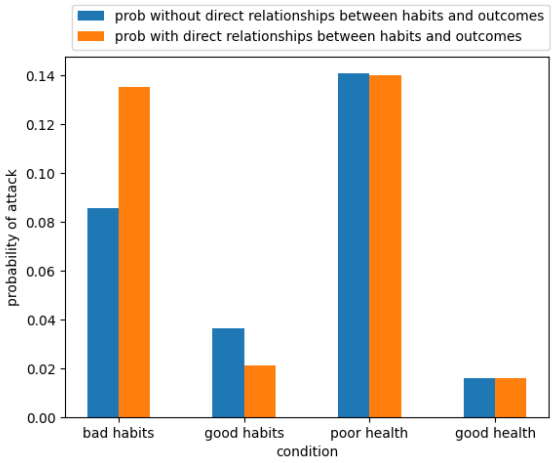
The results of provided examples are shown in Figure 9.



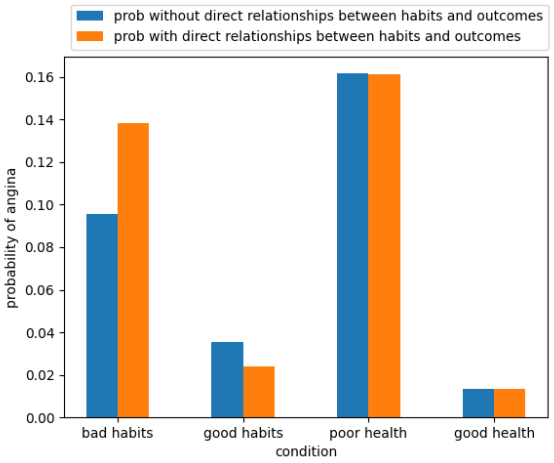
(a) diabetes



(b) stroke



(c) attack



(d) angina

Figure 6: Comparison between results before and after adding edges

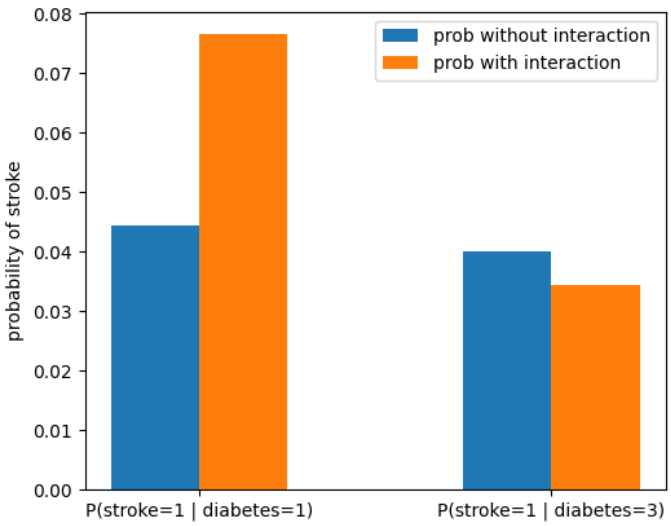


Figure 7: Comparison between results before and after adding edges

```
P(stroke=1 | diabetes=1) = 0.044 in network in problem 4.
P(stroke=1 | diabetes=3) = 0.040 in network in problem 4.
P(stroke=1 | diabetes=1) = 0.077 in network in problem 5.
P(stroke=1 | diabetes=3) = 0.034 in network in problem 5.
```

Figure 8: Program result of problem 5

```
inference starts
   gauge  probs
0      1  0.685
1      0  0.315
   fuel  gauge  probs
0     0      1   0.19
1     0      0   0.81
   fuel  gauge     probs
0     1      0  0.742857
1     0      0  0.257143
   battery  fuel  gauge     probs
0        0     1      0  0.888889
1        0     0      0  0.111111
inference ends
```

```
income dataframe is
      probs  income
0  0.050848       1
1  0.059429       2
2  0.074042       3
3  0.094414       4
4  0.116356       5
5  0.150725       6
6  0.164430       7
7  0.289755       8
   diabetes  exercise  long_sit  smoke     probs
0         1         2         1      1  0.136815
1         2         2         1      1  0.008916
2         3         2         1      1  0.837218
3         4         2         1      1  0.017052
```

Figure 9: Program result of examples