**Part 1**

**Q.1**: We have three identical six-sided dices. We role one dice first and the remaining two dices after that. What is the probability that the point obtained in the first roll is greater than sum of the points obtained in the second roll.

There are 20 cases satisfying the requirements.

|  |  |  |
| --- | --- | --- |
| Dice 1 - first roll | Dice 2 - second roll | Dice 3 - second roll |
| 3 | 1 | 1 |
| 4 | 1 | 1 |
| 4 | 1 | 2 |
| 4 | 2 | 1 |
| 5 | 1 | 1 |
| ... | ... | ... |

**Hence the probability of the event = 20 / (6 \* 6 \* 6) = 9.26%**

**Q.2**: Given a Bayesian Network as in Figure 1

The above Bayesian Network has the following probabilities P(A) = 0.6 P(B|A) = 0.7, P(B|~A) = 0.3 P(C|B) = 0.2, P(C|~B) = 0.8 P(D|B) = 0.1, P(D|~B) = 0.4

Calculate the probability of P(~C)

<https://www.slideshare.net/OrochiKrizalid/bayesian-networks-13646169>

<https://docs.google.com/spreadsheets/d/1ACp0yysMFy0dOPG1LatHIk7511DiGq1QLV7NCcJq2lg/edit#gid=0>

Probability tables:

|  |  |
| --- | --- |
| **A** | **P(A)** |
| TRUE | 0.6 |
| FALSE | 0.4 |

|  |  |  |
| --- | --- | --- |
| **A** | **B** | **P(B|A)** |
| TRUE | TRUE | 0.7 |
| TRUE | FALSE | 0.3 |
| FALSE | TRUE | 0.3 |
| FALSE | FALSE | 0.7 |

|  |  |  |
| --- | --- | --- |
| **B** | **C** | **P(C|B)** |
| TRUE | TRUE | 0.2 |
| TRUE | FALSE | 0.8 |
| FALSE | TRUE | 0.8 |
| FALSE | FALSE | 0.2 |

P(C) = P(A) \* P(B|A) \* P(~C|B)

+ P(A) \* P(~B|A) \* P(~C|~B)

+ P(~A) \* P(B|~A) \* P(~C|B)

+ P(~A) \* P(~B|~A) \* P(~C|~B)

= 0.6\*0.7\*0.8 + 0.6\*0.3\*0.2 + 0.4\*0.3\*0.8 + 0.4\*0.7\*0.2 = **0.524**

**Q.3:** Given a simple Hidden Markov Model with a hidden variable S and an

observation O as in Figure 2

Figure 2: a Hidden Markov Model

In the above Hidden Markov Model, S takes two possible states A and B while O has

four possible values X, Y, Z and T.

The initial state has the probabilities: P s0 (A) = 0.4, P s0 (B) = 0.6

The transition probabilities of the states are

o P T (A|A) = 0.2, P T (B|A) = 0.8

o P T (A|B) = 0.7, P T (B|B) = 0.3

The probabilities of getting observation values given states are

o P O (X|A) = 0.1, P O (Y|A) = 0.2, P O (Z|A) = 0.3, P O (T|A) = 0.4

o P O (X|B) = 0.2, P O (Y|B) = 0.3, P O (Z|B) = 0.4, P O (T|B) = 0.1

Calculate the probability of seeing the sequence of observations TZY at states 1, 2, 3.

|  |  |
| --- | --- |
| P(A) | 0.4 |
| P(B) | 0.6 |
| P(A|A) | 0.2 |
| P(B|A) | 0.8 |
| P(B|B) | 0.3 |
| P(A|B) | 0.7 |

|  |  |
| --- | --- |
| P(X|A) | 0.1 |
| P(X|B) | 0.2 |
| P(Y|A) | 0.2 |
| P(Y|B) | 0.3 |
| P(Z|A) | 0.3 |
| P(Z|B) | 0.4 |
| P(T|A) | 0.4 |
| P(T|B) | 0.1 |

|  |
| --- |
| 0.000384 |
| 0.002304 |
| 0.007168 |
| 0.004608 |
| 0.001764 |
| 0.001134 |
| 0.001008 |
| 0.000648 |

**P(T) = 0.22**

**P(T,Z) = 0.0806**

**P(T,Z,Y) = 0.019018**

**Part 2**

Q. 4. I use **Logistic Regression** and **Matrix Factorization**.

* **Matrix Factorization** to generate categories and product latent feature vectors, which can be used to create session vectors. These session vectors are used as advanced features.
* **Logistic Regression** since it performed well in my setting.

See **Data exploring.ipynb** in **notebook** folder for more detail.

Q. 5. See project folder attached to email.