Advanced Methods in ML 2017 - Exercise 1

1. Consider a distribution q over three random variables X_1, X_2, X_3 defined as:

$$q(x_1, x_2, x_3) = \begin{cases} 1/12 & x_1 \oplus x_2 \oplus x_3 = 0 \\ 1/6 & x_1 \oplus x_2 \oplus x_3 = 1 \end{cases} \tag{1}$$

   (a) What is $I(q)$ (namely the set of all correct conditional independence statements) in this case?

   (b) Is there a DAG G where $I_{LM}(G) = I(q)$?

   (c) Is there an undirected graph G such that $I_{sep}(G) = I(q)$

2. Consider four random variables W, X, Y, Z where the distribution $p(w, x, y, z)$ is positive (i.e., not zero for any assignment). Assume that the two following properties are known:

$$(X \perp Y|Z, W) \ , \ (X \perp W|Z, Y) \tag{2}$$

   Show that $X \perp Y, W|Z$.

3. Consider random variables $X_1, \ldots, X_n$. The Markov Blanket of X_i is the minimal subset $S \subset \{1, \ldots, n\}$ such that $X_i \perp X_{\bar{S} \setminus i}|X_S$ (here $\bar{S}$ is the complement of S). In other words, conditioned the Markov blanket, variables X_S the variable X_i is independent of all the other variables. Given a DAG G and variable X_i, find a subset S that is the Markov blanket of X_i for any Bayesian network on G. The blanket should be described in terms of graph properties such as children, parents, non-descendents etc.

4. Given a distribution $p(x)$, we say that an undirected graph G is a *minimal I-map* for p if it satisfies $I_{sep}(G) \subseteq I(p)$, and any edge removed from G will make this false. Given a *positive* distribution p, construct a graph as follows: if $(X_i \perp X_j|X_{1,\ldots,n} \setminus i, j) \notin I(p)$ add the edge (i, j) to G.

   (a) Show that the G constructed above satisfies $I_{sep}(G) \subseteq I(p)$.

   (b) Show that this G is a minimal I-map for p.

5. (No need to submit) Familiarize yourself with the TensorFlow library. Read the MNIST basic tutorial, the Deep MNIST tutorial, and the CIFAR10 CNN tutorial.

6. Here you will show that there exist distributions that satisfy $I_{sep}(G)$ but are not Markov networks with respect to G. Consider the distribution $p(x_1, x_2, x_3, x_4)$ which has probability $1/8$ for each of the assignments $(0,0,0,0), (1,0,0,0), (1,1,0,0), (1,1,1,0), (0,0,0,1), (0,0,1,1), (0,1,1,1), (1,1,1,1)$, and probability zero for all others. Show that $I(p) = I_{sep}(G)$ where G is a square graph. But that p is not a Markov network with respect to this graph.