بسم الله الرحمن الرحيم

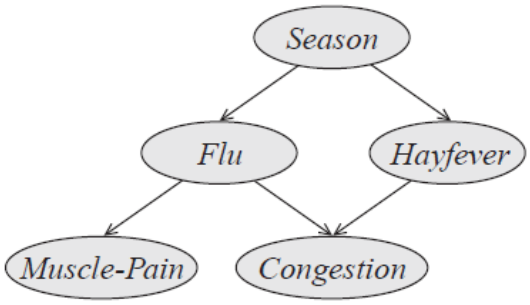
NOVEMBER 23, 2017 *In the name of God, the Beneficent the Merciful*

ASSIGNMENT 3

INSTRUCTOR: M.A. KEYVANRAD | TEACHING ASSISTANT: E. KHADEMI | COURSE: DEEP LEARNING

PGM

Consider the following Bayes net with five discrete variables. All variables except "Season" are binary valued and the "Season" variable have four values.



1. Complete the network using local probabilistic models (here CPTs) using your information and Wikipedia's about "Flu", "Hayfever" and other variables. You should provide four CPTs to complete the model. The priors over the "Season" variable is provided for you.

Season	Spring	Summer	Autumn	Winter
Probability	0.25	0.25	0.25	0.25

The values which are selected for CPT cells must make sense. Moreover, you should provide a brief reason for selecting such a values. An example for $P(H|S)$ CPT is provided here with the following reasons:

Hayfever / Season	H^0	H^1
Spring	0.83	0.17
Summer	0.93	0.07
Autumn	0.87	0.13
Winter	0.97	0.03

- Hay fever is the seasonal Allergic rhinitis triggered by the pollens of specific seasonal plants.
- About 10% of adult population were diagnosed with the hay fever.

2. Inspect the three types of reasoning patterns over your model: causal reasoning, evidential reasoning and inter-causal reasoning. Justify your inference results. You may find it useful to use numpy multi-dimensional arrays and their broadcasting

feature for constructing the full joint distribution then conditional-marginalize the joint to complete your inference. For this purpose, element-wise multiplication, array reshaping and summation over specific axis are useful.

DBN (Deep Belief Network)

1. Using Equation (1), show that $P(H|V)$, the distribution of the hidden units conditioned on all of the visible units, can be factorized as

$$p(H|V) = \prod_j p(H_j|V)$$

Where

$$p(H_j = 1|V = v) = \sigma\left(b_j + \sum_i w_{ij} v_i\right)$$

and $\sigma(s) = \frac{e^s}{1+e^s}$ is the sigmoid function.

Note that: $p(\boldsymbol{H}_j = 0|\boldsymbol{V} = v) = 1 - p(\boldsymbol{H}_j = 1|\boldsymbol{V} = v)$

$$p(V = v, H = h) = \frac{1}{Z} e^{-\boldsymbol{E}(v,h)} \qquad (1)$$

2. First, train an RBM using following fake data and this RBM code (an RBM with 6 visible units and 2 hidden units).

Ali: (The Kingdom of Solomon=1, Gold and Copper=1, Ekhrajiha=1, Kolah Ghermezi=0, Be Roh-e Pedaram=0, The Messiah=0), Top Importance to Fajr International Film Festival Winner.

Reza: (The Kingdom of Solomon=1, Gold and Copper=0, Ekhrajiha=1, Kolah Ghermezi=0, Be Roh-e Pedaram=0, The Messiah=0), Top Importance to Fajr International Film Festival Winner but uninterested in Gold and Copper.

Mohammad: (The Kingdom of Solomon=1, Gold and Copper=1, Ekhrajiha=1, Kolah Ghermezi=0, Be Roh-e Pedaram=0, The Messiah=0), Top Importance to Fajr International Film Festival Winner.

Javad: (The Kingdom of Solomon=0, Gold and Copper=0, Ekhrajiha=1, Kolah Ghermezi=1, Be Roh-e Pedaram=1, The Messiah=0), Top Importance to comic film.

Hosein: (The Kingdom of Solomon=0, Gold and Copper=0, Ekhrajiha=1, Kolah Ghermezi=1, Be Roh-e Pedaram=0, The Messiah=0), Top Importance to comic film but uninterested in Be Roh-e Pedaram.

Ahmad: (The Kingdom of Solomon=0, Gold and Copper=0, Ekhrajiha=1, Kolah Ghermezi=1, Be Roh-e Pedaram=1, The Messiah=0), Top Importance to comic film.

- Plot the network learned weights and explain what each hidden neuron learned?
- What happens if we give the RBM a new user, George, who has (The Kingdom of Solomon=0, Gold and Copper=0, Ekhrajiha=0, Kolah Ghermezi=1, Be Roh-e Pedaram=1, The Messiah=0) as his preferences?
- What happens if we activate only the Fajr Winner unit, and run the RBM a bunch of different times?

3. Create a DBN (according to structure of this paper) for classifying MNIST dataset, using a DBN toolbox (such as DeeBNet, Tensorflow, Theano, …). Compare your results with the paper.