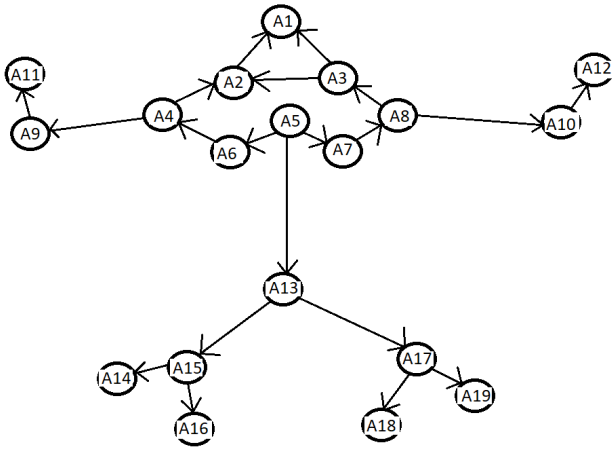
Final exam: practice

Introduction to Machine Learning
Fall 2018
Instructor: Anna Choromanska

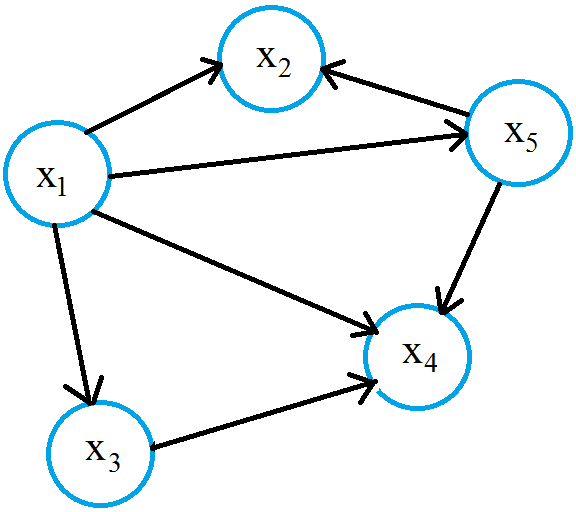
Problem 1

Eve is looking for WallE using her cameras but can't find WallE. Eve has small circuits for performing the junction-tree algorithm. Help her out by designing a junction-tree from the graph below which Eve has in her mind for WallE.



Problem 2

Consider the Bayesian network below with binary variables $x_1, x_2, \ldots, x_5$.



Write out the factorization of the probability distribution $p(x_1, ..., x_5)$ implied by this directed graph. (10 points) Then, using the Bayes ball algorithm, indicate for each statement below if it is True or False and justify your answers (100 points)

- x_2 and x_4 are independent.
- x_2 and x_4 are conditionally independent given x_1, x_3, and x_5.
- x_2 and x_4 are conditionally independent given x_1 and x_3.
- x_5 and x_3 are conditionally independent given x_4.
- x_5 and x_3 are conditionally independent given x_1, x_2, and x_4.
- x_1 and x_3 are conditionally independent given x_5.
- x_1 and x_3 are conditionally independent given x_2.
- x_2 and x_3 are independent.
- x_2 and x_3 are conditionally independent given x_5.
- x_2 and x_3 are conditionally independent given x_5 and x_4.

Problem 3

Consider the fragment of the convolutional architecture given below:

- Input image: $1 \times x \times y$
- Convolutional layer: $\underbrace{1 \to 4}_{\text{number of input and output channels}}$, $\underbrace{3 \times 4}_{\text{filter size}}$, $\underbrace{2 \times 2}_{\text{stride}}$
- ReLU
- MaxPooling: $\underbrace{2 \times 2}_{\text{region size}}$, $\underbrace{2 \times 2}_{\text{stride}}$
- Convolutional layer: $4 \to 6, 3 \times 3, 2 \times 2$
- ReLU
- MaxPooling: $2 \times 2, 2 \times 2$
- Flattening (3D to 1D): $\underbrace{6 \times 9 \times 6}_{\text{number of feature maps } \times \text{ size of the feature map } (9 \times 6)} \to 324$

What is the size of the input (in other words what is x and y)?

Problem 4

Explain overfitting and underfitting on an example.

Problem 5

HMM Example

You are given the parameters of a 2-state HMM. You observed the input sequence AB (from a 2-symbol alphabet A or B). In other words, you observe two symbols from your finite state machine, A and then B. Using the junction tree algorithm, evaluate the likelihood of this data p(y) given your HMM and its parameters. Also compute (for decoding) the individual marginals of the states after the evidence from this sequence is observed: $p(q_0|y)$ and $p(q_1|y)$. The parameters for the HMM are provided below. They are the initial state prior $p(q_0)$, the state transition matrix given by $p(q_t| q_{t-1})$, and the emission matrix $p(y_t| q_t)$, respectively.

$$\pi = p(q_0) = \begin{matrix} 1 & 2 \end{matrix} \\ \begin{bmatrix} 1/3 & 2/3 \end{bmatrix}$$

$$a^T = p(q_t \mid q_{t-1}) = \begin{matrix} & 1 & 2 \\ 1 & 3/4 & 1/2 \\ 2 & 1/4 & 1/2 \end{matrix}$$

$$\eta^T = p(y_t \mid q_t) = \begin{matrix} & 1 & 2 \\ A & 1/2 & 1/3 \\ B & 1/2 & 2/3 \end{matrix}$$

Problem 6

What is the VC dimension of the hypothesis space consisting of triangles in the 2D plane (justify your answer)? Points inside the triangle are classified as positive examples.