# Machine Learning and Deep Learning II

Homework 1

September 21, 2022

**Due date is 11:59 PM, October 5**.

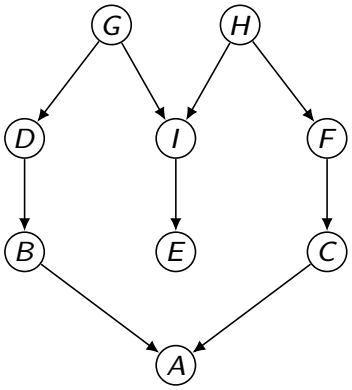
1. [7 pts] Given a Bayesian network over variables $\mathbf{V}$, let $X \in \mathbf{V}$ be a node in the graph. A *Markov Blanket* of $X$, denoted as $\mathrm{MB}(X)$, is a minimal set of variables which, when given, makes the rest of the variables $(\mathbf{V} \setminus \{X\} \setminus \mathrm{MB}(X))$ conditionally independent to $X$. That is,

$$X \perp\!\!\!\perp \mathbf{V} \setminus \{\mathbf{X}\} \setminus \mathrm{MB}(X) \mid \mathrm{MB}(X).$$

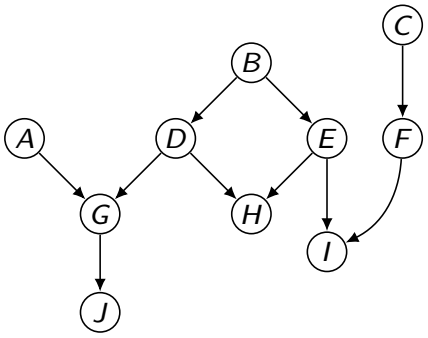
   Categorize variables in $\mathrm{MB}(X)$ into three groups, and explain why those three groups should be part of $\mathrm{MB}(X)$. (Informally, what kind of variables relative to $X$ should be blocked to be independent with the rest?)

2. [8 pts] For each graph below containing $A$ and a set of variables $\mathbf{S}$ (specified for each problem), determine which variables among $\mathbf{V} \setminus (\mathbf{S} \cup \{A\})$ are d-connected to $A$ given $\mathbf{S}$.

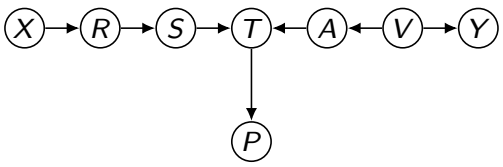
**(a)** $\mathbf{S} = \{B\}$



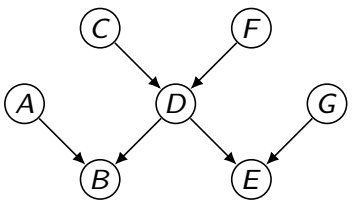

**(b)** $\mathbf{S} = \{J\}$




**(c)** $\mathbf{S} = \{R, P\}$




**(d)** $\mathbf{S} = \{B, D\}$



3. [20 pts] Implement expectation maximization algorithm for Gaussian Mixture Model in Python. Only for this problem, you can collaborate up to two other MLDL2 students. Please make sure to write down any collaborating students' names in the code. You *may* use the provided file as a template.

4. [15 pts] Implement Viterbi algorithm in Python. Make sure to return a sequence of hidden states *not* the max value. You *may* use the provided file as a template.

Please submit a `20xx-xxxxx-HW1.zip` file with `gmm.py`, `hmm.py`, and `HW1.pdf` for the rest of the questions below.