Homework # 2
MGT-302

Due to 21.05.2021

Please submit your answers by the end of 21.05.2021.

Problem 1

In this problem, you will see another formalization of the SVM. Prove that the following problems are equivalent

$$\begin{aligned}
&\mathbf{w}^*, b^*, r^* = \arg\max_{\mathbf{w},b,r} r \\
&\text{subject to } y_n(\mathbf{w}^T\mathbf{x}_n + b) \geq r,\ n = 1, ..., N, \\
&||\mathbf{w}||_2 = 1, \\
&r > 0.
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{w}_*, b_* = \arg\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||_2^2 \\
&\text{subject to } y_n(\mathbf{w}^T\mathbf{x}_n + b) \geq 1,\ n = 1, ..., N.
\end{aligned}$$

Moreover, find the relationship between $\mathbf{w}^*, b^*, r^*$ and $\mathbf{w}_*, b_*$.

Problem 2

Consider the DAG in figure 1.

- Write down the joint distribution for this graph.
- Find a subset of variables that d-separates X_i from X_j.
- Draw another DAG that is Markov equivalent to this graph.
- Find a valid adjustment set for the ordered pair (X_i, X_j), i.e., for computing $P(X_j|do(X_i))$.

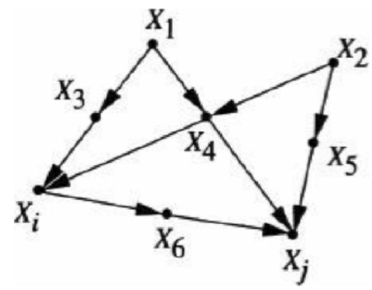

Figure 1: DAG of problem 2.

Problem 3
Prove the following statements,

- Chain rule for Entropy: $H(X_1, ..., X_m) = \sum_{i=1}^{m} H(X_i|X_1, ..., X_{i-1})$. Note that for $i = 1$, the term within summation is equal to $H(X_1)$.
- Using the non-negativity of KL-divergence, show that $I(X;Y|Z) \geq 0$,
- If the Bayesian network of X, Y, and Z is $X \to Y \to Z$, then show

$$I(X;Y|Z) \leq I(X;Y).$$