# CS60021: Scalable Data Mining

## Practice Questions: DL Frameworks and Subset Selection

State whether following statements are true or false with max 2-sentences of explanations:

1. Batchnorm layers do not introduce any new parameters to the neural network.
2. Each node of a Pytorch computational graph is a Tensor or a Variable.
3. Submodular set functions are monotone, in the sense that adding new elements to a set increase the function value.
4. Let $v(x)$ be the value of an element $x$. The function $f(s) = \sum_{x \in S} v(x)$ is a submodular function.
5.

**Long-answer questions:**

Q1. Which component (class) of Pytorch is responsible for the automatic differentiation property? Consider the following code:

```
x =
torch.autograd.Variable(torch.tensor([10.0,10.0]),require
s_grad=True)
y =
torch.autograd.Variable(torch.tensor([5.0,5.0]),requires_
grad=True)
z = x+y
u = z.mean()
u.backward()
print(x.grad)
print(y.grad)
```

Write the output of the above program. Draw the forward and backward graphs for the above program with the values of appropriate variables at each node.

Q2. Define Submodular Functions, mathematically. What are monotone submodular functions? Consider the function $f$ for a set $S$, $f(S) = \sum_{x \in S} g(x)$, where $x \in X$ are elements that constitute the sets, and $g$: $X \rightarrow [0, \infty)$ are positive real functions. Show that $f$ is a monotone submodular function.

Q3. Write the greedy approximation algorithm for maximization for monotone submodular functions. State and derive the approximation error bound for the above algorithm.

Q4. What are the 2 main features of a Pytorch tensor over a Numpy array ? Explain how they are used.

Q5. Write a Pytorch code fragment to compute the output of following function where all the matrix and vector operations are performed on GPU, but the comparison happens on CPU:

$$y = \begin{cases} \sigma(W_2 x) & if. \quad w_1^T x > 0 \\ \sigma(W_3 x) & otherwise \end{cases}$$

Q6. What is the 2 main advantages of using Pytorch Modules. Write the code for creating a new module M3, assuming that Modules M2 and M3 are already defined, which performs the following computation:

$$M3(x, y) = M2(M1(x), y)$$

where $x$ and $y$ are input tensors.

Q7. Write the facility location objective function for selection of a subset of datapoints from a dataset, given the pairwise similarities between the datapoints. Define a submodular function. Show that the facility location objective function is indeed submodular.

Q8. Write the convex relaxation of the above mentioned facility location objective function. Clearly describe the variables, and explain why the two are equivalent.

Q9. Write the randomized greedy algorithm for maximization of a submodular function, subject to an upper bound on the cardinality of the selected set. Under what assumption is this algorithm guaranteed to provide an approximate solution?