

Please answer the following questions related to Machine Learning concepts:

1. [18 points] Explain the key properties of the following concepts:

- 1) supervised learning,
- 2) unsupervised learning,
- 3) online learning,
- 4) batch learning,
- 5) model-based learning,
- 6) instance-based learning.

- 1) Supervised Learning : Model should be trained on labelled data. Labelled data basically means that some input data is tagged with the resulting/correct output data. Models will train and learn from these tagged/labelled and create a formula which will be able to predict the output for unseen data. Human supervision is necessary for supervised learning models.
- 2) Unsupervised Learning : Here model uses unlabelled data. There is no human intervention. In Unsupervised learning the model itself has to find hidden patterns and underlying structure of data. Also unsupervised is very important as most of the real world data is unlabelled and expensive to label.
- 3) Online Learning : In situations where model has to be updated frequently there we use online learning. Here system can be trained sequentially on real time data, for ex - prediction of stock prices. Also can be used in places where datasets are really huge and the memory cannot handle that. Here then the data is passed sequentially or in mini-batches
- 4) Batch Learning : Also known as offline learning, here the system cannot learn sequentially or incrementally. It must be trained at once, using all the available data. And if new data points come in which have to be trained then the entire process has to be run again. Basically used in models where data is not real time.
- 5) Model Based training : In model based learning, we develop models, train them on existing data so that the algorithm can learn from this data and then give output or predictions. Better for larger amounts of data and complex problems.
- 6) Instance Based Training : In instance based, the system **compares** new examples/data with existing labelled data and tries to find a pattern which can be generalised on the new examples/ data. Better for smaller amounts of data and simpler problems.