**Assignment 1**

1. What type of Machine Learning Algorithms would be used to allow your bipedal bot to walk in various terrains?

Ans: Reinforcement Machine Learning algorithm. As this will help bot to learn and adopt by interaction with the surroundings.

1. What is the difference between online learning and out-of-the-core learning?

Ans: Online learning is the method in which model learns continuously with the live stream of data. Used when dataset size is huge. And model is complex.

In Out-of-the core learning, model is trained on a fixed set of dataset. Used when dataset is small.

1. What is the importance of a train-dev set in machine learning?

Ans:

Train-dev set is useful for identifying Data Mismatch. This happens when the model is trained on dataset which is ‘Non-representative’ of the actual production data.

This Train-dev set contains data/images taken from web. If the testing accuracy is high on this Train-dev test, then it can be concluded that there is a Data Mismatch.

1. If suppose your model is poorly generalizing the unseen data, what possibilities are there, and how you will solve those situations?

Ans:

This means model is having low test accuracy which means model is overfitted.

Resolve Overfitting:

* + Reduce / Optimize the features to be considered
  + Reduce complexity of the model(ex, using lower degree polynomial)
  + Constrain the model using Regularization (Increasing Lambda value)

1. What is data leakage and what can go wrong if you tune your hyperparameters on the test set?

Ans:

Data Leakage: Data leakage is when information from outside the training dataset is used to create the model. This additional information can allow the model to learn or know something that it otherwise would not know and in turn invalidate the estimated performance of the mode being constructed.

Using test set for hyper parameter tuning will cause this Data leakage. And will result in an abnormally high model accuracy which may not be true.