

## Assignment 2

Advanced Machine Learning ( CS566 )

Department of CSE, IIT Patna

(Read all the instructions carefully and adhere to them.)

Date: 13-Feb-2020

Deadline:- 02-Mar-2020

### Instructions:

1. Markings will be based on the correctness and soundness of the outputs.
2. Marks will be deducted in case of plagiarism.
3. Proper indentation and appropriate comments (if necessary) are mandatory.
4. You should zip all the required files and name the zip file as:  
***roll\_no\_of\_all\_group\_members .zip , eg. 1501cs11\_1201cs03\_1621cs05.zip.***
5. Upload your assignment ( the zip file ) in the following link:

<https://www.dropbox.com/request/13FbfkouQxzyI8ynkcUW>

***For any queries regarding this assignment contact:***

Chanchal Suman([chanchaliitp@gmail.com](mailto:chanchaliitp@gmail.com))

### Question:

Intrusion detection is the process of identifying the Intruders from a network. This is generally done by analyzing the header of the data packets.

1. Design an intrusion detection system using feature selection.
2. Use Multi-objective optimization for selecting relevant features from the data, and decision tree for the classification.
3. Objective functions:

The pairwise [mutual information](#) should be low.

The length of the feature subset should be high.

### Dataset:

[NSL-KDD](#) is the standard dataset for The IDS.

Use '20 Percent Training Set.csv' for training the System, and 'KDDTest-21.txt' for the testing purpose.

The Input is the raw header, having 41 features.

Class Labels: The output is the 42nd feature. (Please ignore the last one).

**Evaluation:**

- 1) Use NSGA-II for finding the relevant feature subset.
- 2) Use Decision Tree on the feature subset to evaluate the performance.