Introduction to Machine Learning: Overview of Topics and Activities

1 Outline of Topics Covered and Learning Activities

(a) The Machine Learning (ML) topics will be covered in Weeks 4 and 5.

(b) The main text is "An Introduction to Statistical Learning with Applications in R" available here. There are also online copies available in our HWU library system.

(c) An R package `ISRL` is available on CRAN which provides the collection of data sets used in the book.

(d) ML topics covered:

(1) **Unit 1**:

- Introduction to Machine Learning.
- The Regression Setup.
- Training, Test and Validation Sets, and Related Errors.
- The Cross Validation (CV) Procedure.
- Bias vs. Variance Trade Off.

(2) **Unit 2**:

- Machine Learning Techniques for Classification.
- Logistic Regression.
- The CV procedure for Logistic Regression.

(3) **Unit 3**:

- Supervised vs. Unsupervised Learning.
- The k-means and Gaussian mixture model clustering algorithms
- Choosing the optimal number of clusters using the elbow criteria.

(e) I will also post a pdf file on knowledge transfer activities which will give you the sections of the Main Text that you should read, the lecture recordings you should follow and the exercises that you should do for the three Units.

(f) There will be a lab session on Week 7 (Monday, 22nd February, 2021) for you to do the exercises in all three Units that involve code development in R.

(g) There will be a ML class test worth 10% of your total grades for this course. More details on this later.