CNT4153 - Applied Machine Learning
Three Credits, Four and a half hours, Engineering Topic.

**Instructor:** Dr. Himanshu Upadhyay.

Textbook: Hands-On Machine Learning with Scikit-Learn by Aurelien Geron, O’Reilly Media Inc

**Specific Course Information:**
This course will focus on the application of traditional machine learning algorithms and popular framework to simple data sets. Students will learn about supervised/unsupervised learning, classification / regression algorithms, clustering, model building and prediction, model optimization & dimensionality reduction.

**Specific Goals for the Course**

- **a. Specific outcomes of instruction**
  
  Upon successful completion of this course, the student will:

  1. Understand fundamentals of AI/ML/DL
  2. Perform preprocessing of large data sets
  3. Knowledge of machine learning algorithms and applications
  4. Identify and apply classification & regression algorithms
  5. Develop clusters using large data sets
  6. Understand error metrics, dimensionality reduction & feature engineering
  7. Perform analysis with multiple algorithms and operationalize the models for prediction
  8. Display analytics results and reporting with data visualization tools
  9. Able to store and analyze results in the persistent data store
  10. Understand data science life cycle

- **b. Explicitly indicate which of the student outcomes listed in Criterion 3 or any other outcomes are addressed by the course.**

  In this course the student will have to show

  (a) an ability to apply knowledge of mathematics, science, and engineering (N/A)
  (b) an ability to design and conduct experiments (simulations), as well as to analyze, interpret data (N/A)
  (c) an ability to design a system, component, or process to meet desired needs (N/A)
  (d) an ability to function in multi-disciplinary teams (N/A)
Brief list of the topics to be covered

1. Introduction to Artificial Intelligence, Machine Learning & Deep Learning
2. Python data structures and packages- Pandas, NumPy and ScikitLearn
3. Data visualization with MatPlotLib and Seaborn
4. Supervised / Unsupervised Learning
5. Pre-processing techniques for sample datasets
6. Data analysis with classification algorithms
7. Application of linear regression algorithms
8. Classification Error Metrics
9. Regularization & Gradient Descent
10. KNN, Logistic Regression, Naïve Bayes, Decision Trees and SVM/Kernels
11. Bagging & Random Forest
12. Clustering application for sample datasets
13. K-Mean Clustering
14. Dimensionality Reduction/PCA
15. Data Science Life Cycle and applications

GRADING:

<table>
<thead>
<tr>
<th>Course Requirements</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Assignments    10%
Quiz 1      10%
Midterm    25%
Quiz 2     10%
Final     25%
Project   20%
Overall Grade 100%

**Conversion of Numerical Grade to Letter Grad**

<table>
<thead>
<tr>
<th>Numerical Grade</th>
<th>Letter Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 &lt;= A &lt;=100</td>
<td>A</td>
</tr>
<tr>
<td>90 &lt;= A- &lt; 94</td>
<td>A-</td>
</tr>
<tr>
<td>86 &lt;= B+ &lt; 89</td>
<td>B+</td>
</tr>
<tr>
<td>82 &lt;= B&lt; 85</td>
<td>B</td>
</tr>
<tr>
<td>78&lt;= B-&lt; 81</td>
<td>B-</td>
</tr>
<tr>
<td>74 &lt;= C+ &lt; 77</td>
<td>C+</td>
</tr>
<tr>
<td>70 &lt;= C &lt; 73</td>
<td>C</td>
</tr>
<tr>
<td>60 &lt;= D &lt; 69</td>
<td>D</td>
</tr>
<tr>
<td>F: Below 60</td>
<td></td>
</tr>
</tbody>
</table>