INFO 526 Applied Machine Learning

Table of Contents

INFO 526 Applied Machine Learning .................................................. 1
Course Intro ....................................................................................... 1
Course Syllabus ................................................................................ 2
Grading Scale .................................................................................... 3
Course Infrastructure ......................................................................... 3
References ........................................................................................ 4
Instructor Bio ..................................................................................... 4

Course Intro

Just like electricity, the automobile, the Internet, and mobile phones transformed the 20th century, machine learning is transforming the 21st century, changing how people perceive and interact with technology, enabling machines perform a wider range of tasks, in many cases doing a better job than humans. These applications include: voice assistants on our smartphones, product recommendation engines, self-driving cars, high frequency stock market trading, applications for social good (combating crime), playing games (from Go to Atari), preventing credit card fraud, filtering out spam from our email inboxes, detecting and diagnosing medical diseases, the list goes on and on. Large companies, such as Google, Facebook, Apple, Amazon, and Hooli Con (GAFAH), and venture capitalists alike are investing heavily in machine learning research and applications.

If you want to become a machine learning practitioner, a better problem solver, or maybe even consider a career in machine learning research, then this course is for you. However, for a novice, the theoretical concepts behind machine learning can be quite overwhelming. This course focuses on: introducing theoretical concepts and algorithms in a step-by-step manner, while infusing them with intuition, examples and python jupyter notebooks. In this spirit you will code up core ML algorithms from scratch, while also working through numerous example applications of machine learning. Concrete examples help illustrate the broader concepts by putting the learned material directly into action. This combination of theory and hand-ons will help you master core ML concepts and algorithms that are used, not only in Silicon Valley but, throughout the world, while also offering intuitive yet informative explanations of how machine
Learning algorithms work, how to use them, and most importantly, how to avoid the most common pitfalls.

Course Syllabus

PART 1: Introduction and Infrastructure [3-4 hours]
- L1: Introduction to course; course logistics; survey of ML algorithms from KNN to deep learning to graph algos
- L1: Sample Practical Session
  - Install Docker container (install, pull, github basics, homework submissions)
  - Notebook for sample regression and classification problems (high level overview)
  - HW1 review (based on L1 and L2); part 1: using Scikit Learn
    - Basic ML pipeline using SciKit Learn on Titanic Data, Kaggle submission

PART 2: Supervised ML
- L2: KNN: classification, regression + EDA + ML pipelines
- L3: Optimization theory: gradient descent, SGD
- L4: Linear Regression from basics to regularization
- L5: Classification: logistic/softmax regression
- L6: Perceptron + SVM
- L7: Non-gradient-based approaches: decision trees, random forests
- L8: Mid-term

PART 3: Unsupervised learning
- L9: Clustering: K-means, Hierarchical clustering, DBSCAN
- L10: Dimensionality reduction: PCA, tSNE
- L11: Embeddings: text classification

PART 4: Advanced/Recent topics
- L12: Neural networks: backprop, core concepts, deep learning
- L13 Recommender systems: popularity-based, item-to-item, matrix factorizations
- L14: Reinforcement learning: multi-armed bandit, Bellman equation, Q-learning
- L15: end of semester exam

**Course Prerequisites**

- Intermediate programming capabilities in python and other object-oriented languages.
- Enough knowledge of calculus to be able to differentiate simple functions.
- Enough knowledge of linear algebra to understand simple equations involving vectors and matrices.
- Enough knowledge of probability theory to understand what a probability density is.

**Grading Scale**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm exam</td>
<td>30%</td>
</tr>
<tr>
<td>Class participation</td>
<td>20%</td>
</tr>
<tr>
<td>End of semester exam</td>
<td>30%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final Grade</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A “Excellent”</td>
<td>&gt;= 85%</td>
</tr>
<tr>
<td>B “Good”</td>
<td>from 75% to 84%</td>
</tr>
<tr>
<td>C “Satisfactory”</td>
<td>from 60% to 74%</td>
</tr>
<tr>
<td>D “Poor”</td>
<td>from 50% to 59%</td>
</tr>
<tr>
<td>F “Unacceptable”</td>
<td>&lt; 50%</td>
</tr>
</tbody>
</table>

**Course Infrastructure**

- GitHub code repository
• Docker Container (for local machine and on the cloud for CPU)
• Canvas

References

• Overview of Gradient Descent variations for machine Learning
  o http://sebastianruder.com/optimizing-gradient-descent

Instructor Bio

Professor Dr. James G. Shanahan (UC Berkeley and Church and Duncan Group Inc)

Professor Dr. James G. Shanahan has spent the past 30 years developing and researching cutting-edge artificial intelligence systems splitting his time between industry and academia. He has (co) founded several companies including: Church and Duncan Group Inc. (2007), a boutique consultancy in large scale AI which he runs in San Francisco; Xona Partners (2013) an ICT consultancy with technical and business depth; RTBFast (2012), a real-time bidding engine infrastructure play for digital advertising systems; and Document Souls (1999), a document-centric anticipatory information system. In 2012 he went in-house as the SVP of Data Science and Chief Scientist at NativeX, a mobile ad network that got acquired by MobVista in early 2016. In addition, he has held appointments at AT&T (Executive Director of Research), Turn Inc. (founding chief scientist; acquired by Amobee), Xerox Research, Mitsubishi Research, and at Clairvoyance Corp (a spinoff research lab from CMU). He also advises several high-tech startups including Aylien (deep learning for natural language processing), InferSystems (ad technology acquired by Kochava), ChartBoost (mobile ad network), DigitalBank, VoxEdu, and others.

Dr. Shanahan has been affiliated with the University of California at Berkeley and at Santa Cruz since 2008 where he teaches graduate courses on big data analytics, artificial intelligence, deep learning, and stochastic optimization. In addition, he is currently visiting professor of data science at the University of Ghent, Belgium. He has published six books, more than 50 research publications, and over 20 patents in the areas of machine learning and information processing. Dr. Shanahan received his PhD in engineering mathematics from the University of Bristol, U. K., and holds a Bachelor of Science degree from the University of Limerick, Ireland. He is a EU Marie Curie fellow. In 2011 he was selected as a member of the Silicon Valley 50 (Top 50 Irish Americans in Technology).
Training set

Training folds

Test fold

1\textsuperscript{st} iteration

2\textsuperscript{nd} iteration

3\textsuperscript{rd} iteration

\ldots

10\textsuperscript{th} iteration

\[ E = \frac{1}{10} \sum_{i=1}^{10} E_i \]