Graduate School of Management
University of California -- Davis

BAX 493B: Topics in Business Analytics - Implementing Machine Learning on the Cloud

Instructor:	Mak Ahmad, Adjunct Professor
Phone:	(415) 935-3188
Email:	shahmad@ucdavis.edu
Office Hours:	By appt (over VC)

Course Description:

Machine learning and cloud computing are complementary technologies. Machine learning takes advantage of larger data sets and elastic computing to create highly accurate predictions at scale. A newer trend is the democratization of AI/ML. This democratization involves creating new abstractions for ML, including pre-trained models, and higher level ML Services. This course teaches students about how to leverage AutoML to quickly develop prediction machines using the best abstraction for the task at hand. Restricted to students enrolled in the MSBA program

Course Objectives:

This course aims to bring hands-on experience to the Machine Learning fundamentals you have already studied. AutoML (Automated Machine Learning) is the process of automating the process of applying machine learning to real-world problems. The theory behind the course is agnostic to any specific cloud provider however the hands-on exercises focus on Google Cloud Platform and Salesforce. The topics include:

- Introduction
- Machine Learning review
- Feature Engineering
- Art & Science of ML
- End-to-end on structured data ML
- Production ML Systems
- Intro to Google Cloud Platform's Tensorflow & AutoML Tables
- Intro to Salesforce's TransmogrifAI & Einstein Prediction Builder

Class Rules

Academic Honor Code All students are expected to adhere to the University of California, Davis' Code of Conduct as noted here: http://sja.ucdavis.edu/files/cac.pdf. Please also note that

if “academic misconduct is admitted or is determined by adjudication to have occurred,” per Regulation 550 the student could potentially receive a grade of “F” not only for the assignment or project in question, but also for the *entire course*.

Use of Electronics in Class You are allowed to use your laptop/tablet in class – however this use is conditional. You cannot use your laptops for any other activity other than those pertaining to the class. Hence, surfing, emailing, chatting, facebook visits and other related activities are not allowed. If such activity is observed in class, then laptop privileges will be revoked – *for the entire class*. Smartphone/phone use is not allowed in class – you are welcome to step out of class, with my permission, if you need to answer a call. Failure to comply with this policy will result in the creation of ‘laptop zones’ (e.g., only the first two rows will be allowed to use laptops).

Course Material

There is no required textbook: all materials will be available on the class website. However the following are recommended readings.

1. Prediction Machines: The Simple Economics of Artificial Intelligence by Ajay Agrawal, Joshua Gans and Avi Goldfarb
2. Google Cloud Platform in Action 1st Edition by JJ Geewax

Grading

- 50% Final Project
- 40% Assignments
- 10% Class Participation

Late assignments, exams, or projects, will not be accepted.

Lecture Outline

Lecture #1

Theory

- Machine Learning review
- Feature Engineering
- Salesforce & Google Cloud's AutoML

Practice

- Salesforce
    - TransmogrifAI
    - Einstein Prediction Builder

        - Einstein Next Best Action (if time allows)

Guest Speaker, Salesforce (15 min)

- Marco Casalaina, SVP Products, Salesforce Einstein: how ML can solve real business problems w/ Salesforce Einstein

Reading / Homework

- See Canvas

Lecture # 2

Theory

- Art & Science of ML
- End-to-end on structured data ML

Practice

- Google Cloud Platform
    - Tensorflow
    - AutoML Tables

Guest Speaker, Google (15 min)

- Scott Penberthy, Director, Applied AI at Google: how ML can solve real business problems w/ Google Cloud Platform

Reading / Homework

- See Canvas

Lecture # 3

Theory

- Production ML Systems

Reading / Homework

- See Canvas

Practice

- Final group presentation practice

The Team Project

You have just joined an AI startup and have been tasked to launch your MVP in 6 months. Your first task is to quickly create a POC for the head of product.

Prompt:

- Determine what your POC will be
- Gather and clean required data for your POC
- Train a model using any ML technology (Google, Salesforce or any vendor)
- Build an end to end ML application pitch deck and incorporate the ML findings in your deck.
    - Note: a ML API could also be a product

Project Topic 1 - COVID-19
Pick any project to forecast something impactful to society or economy related to COVID-19.

For example:

- Create your own health score similar to the one here: https://covidly.com/
- Forecasting business segments that would move from brick and mortar to pure digital business

The following links can be helpful to brainstorm an idea:

- https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge/tasks
- https://www.kaggle.com/c/covid19-local-us-ca-forecasting-week-1/
- https://www.kaggle.com/c/covid19-global-forecasting-week-1

Potential data:

- https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series
- Data from Tableau here that I found by going to this site and clicking “Access the Data” and selecting CSV

Project Topic 2 - Climate Change
Predict economic impact due to climate change.

Potential Data:

- https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data

Project Topic 3 - Twitter post sentiment analysis
Retrieve data from Twitter and then use GCP’s Natural Language API for sentiment prediction.

Potential Data:

- https://towardsdatascience.com/creating-the-twitter-sentiment-analysis-program-in-python-with-naive-bayes-classification-672e5589a7ed

Project Topic X
Topic of your choice (approval required)

Submission
Your team will deliver a 5-10 minute presentation based on your findings over VC. I will provide the date/times that works for me and the team leaders can select any open slot that works for the group.

Turn in using Canvas:
- Pitch deck
- Cleaned up version of your data used
- Any code or screenshots required to create your ML model