IFI 8430 – Advanced Analytics Experience
Course Syllabus – Spring 2021

(Draft – Subject to Change)

Instructor:	TBD
Class Schedule:	TBD
Classroom:	TBD
Office Hours:	TBD
Course Description:	We introduce the procedure of framing and solving business problems with machine learning and predictive analytics. Several real world business problems such as pricing, analyzing customer review data, market segmentation, customer discovery, and process improvement will be explored and tackled with analytics methods. Students will learn how to frame a business problem, collect and clean data, handle missing values, examine descriptive statistics, build different models and choose the most accurate and interpretable one, as well as present insights from the model output. This course is project-based with hands-on application of analytics tools and available packages, specifically in R and Python, to business data and problems to efficiently extract patterns and insights to create value.
Course Objectives:	At the end of the course, students will be able to proceed through the four stages of an analytics workflow: • Business problem framing • Data preparation and exploration • Analytics methodologies • Presentation and visualization
Required Texts:	R and R Studio (downloaded and installed) Online version; text version also available for purchase) Gareth James, Daniela Witten, Trevor Hastie & Robert Tibshirani, An Introduction to Statistical Learning: with Applications in R, available at http://faculty.marshall.usc.edu/gareth-james/ISL/
Contributing Texts:	Roger D. Peng, Exploratory Data Analysis with R, available at https://leanpub.com/exdata Garrett Grolemund & Hadley Wickham, R for Data Science, available at https://r4ds.had.co.nz/

Nicola Sturaro, Rabbit: Introduction to R, available at http://www.quantide.com/rabbit-introduction-to-r/

Class Schedule and Activities:

CLASS	TOPIC AND COURSE CONTENT
Week 1	Problem Framing – Business example: Pricing Problem What is a good business problem?; Benchmarking, identifying pain points; Theory and hypothesis formation; Measurement questions; Falsifiability
Week 2	Team Formation and Team Business Problem Selection
Week 3	Data Ethics and Bias – Business example: Use and limits of self-reported consumer data Selection effects; Privacy and security; Replication and transparency
Week 4	Data Preparation – Business example: Incomplete or missing sales or customer review data Data quality, sufficiency, and availability; Identification of external data sources; Cleaning and pre-processing; Problems of missing data; Features engineering
Week 5	
Week 6	Data Exploration – Business example: Internal data audit Descriptive statistics: mean, median, mode, min, max, standard deviation; Visualization: scatter plots, histograms, box plots, maps; Correlations and time series; Trends, seasonality
Week 7	Statistical Methodologies – Business example: Assessment of marketing efficacy Ordinary Least Squares; Logistic regression; Model selection
Week 8	Machine Learning Methodologies: Supervised – Business example: Customer discovery Logistic regression; Random forest; Linear Discriminant Analysis (LDA); Support Vector Machines (SVM)
Week 9	Machine Learning Methodologies: Unsupervised – Business example: Customer discovery (continued) Clustering; Principal Component Analysis (PCA)
Week 10	Text Analytics Methodologies – Business example: Extracting meaning from consumer reviews Bag of Words, frequencies; Term Frequency-Inverse Document Frequency (TF-IDF); Sentiment analysis; Word clouds; Topic modeling: Latent Dirichlet Allocation (LDA), Structural Topic Models (STM)
Week 11	
Week 12	Deep Learning – Business example: Process improvement

	Deep Learning; Convolutional Neural Networks (CNN); Generative methodologies, e.g. Generative Adversarial Networks (GANS)
Week 13	Visualization, Reporting, and Debriefing
Week 14	Final team presentations