Monterrey Digital Hub - Data Science Immersive

week	topic(s)	delivery	capstone project
1	Python, math essential review	in person	Brainstroming
2	Data ingestion pipelines	remote	Brainstorming, get data
3	Statistics and EDA	remote	get data, proposal ideas
4	Linear models	remote	Proposals ideas discussed
5	Supervised learning	in person	Proposals discussed
6	NLP and unsupervised learning	remote	Proposals written
7	Big Data	remote	Proposal revisions
8		remote	Capstone ETL, EDA
9	Deep learning	in person	Capstone EDA deliverable
10	Production and specialized topics	remote	Capstone models
11	Capstones	remote	Capstone models
12	Capstones	in person	finish models and deliverable

Proposals for capstone projects are due by the end of week 4

Prework

Topic	Resource	Duration	Goals
Python	Python precourse	~40 hours	Exposure to Python fundamentals
Statistics	Statistics short-course	~4 hours	Exposure to classical statistics
Math Foundations	Linear algebra essentials	~4 hours	Exposure to linear algebra

Data Science Immersive

Week 1: Python and version control (IN PERSON)

Day	Main Topics
2019.09.23	Dev, environment, Git, Unix
2019.09.24	Python part I (primatives, loops)
2019.09.25	Python part II (functions, OOP)
2019.09.26	NumPy, pandas

Day	Main Topics
2019.09.27	CASE STUDY - pandas

Week 2: Data Ingestion (REMOTE)

Day	Main Topics
01	Data Visualization (matplotlib)
02	MongoDB
03	SQL
04	Python SQL
05	CASE-STUDY - data ingestion pipelines

Week 3: Statistics and EDA (REMOTE)

Day	Main Topics
01	Linear algebra
02	Probability distributions, statistical inference
03	Sampling and estimation
04	Null hypothesis testing
05	CASE STUDY - exploratory data analysis

Week 4: linear models (REMOTE)

Day	Main Topics
01	Evaluation metrics, KNN
02	Cross-validation
03	Linear and logistic regression
04	Regularized regression
05	CASE STUDY - linear models

Week 5: Supervised learning (IN PERSON)

Day	Main Topics
01	Decision trees, bagging
02	Random forests, gradient descent
03	Boosting
04	Neural networks (backpropagation, MLP)
05	CASE STUDY - supervised learning

Week 6: NLP and unsupervised learning (REMOTE)

Day	Main Topics
01	Natural language processing
02	Text classification, naive bayes
03	Topic models (LDA and NMF)
04	Clustering algorithms
05	CASE STUDY - NLP

Week 7: Big Data (REMOTE)

Day	Main Topics
01	Capstones, Troubleshoot Docker/Spark Install
02	Docker
03	Spark I (SQL and Dataframes)
04	Spark II (ML Pipelines)
05	CASE STUDY - Spark

Week 8: Special ML Topics (REMOTE)

Day	Main Topics
2019.11.11	Profit curves
2019.11.12	Support vector machines
2019.11.13	Principal component analysis
2019.11.14	Singular value decomposition
2019.11.15	Review followed by dedicated capstone time

Week 9: Deep Learning (IN PERSON)

Day	Main Topics
01	Time-series, recurrent neural nets
02	Convolutional neural networks
03	GANs and autoencoders
04	Deep learning case study
05	Capstones

Week 10: Case study Week (REMOTE)

Day	Main Topics
01	Capstones
02	Capstones
03	Capstones
04	HOLIDAY (NO CLASS)

Day	Main Topics
05	HOLIDAY (NO CLASS)

Week 11: Projects (REMOTE)

Day	Main Topics
01	Capstones
02	Capstones
03	Capstones
04	Capstones
05	Presentation practice

Week 12: Projects (IN PERSON)

Day	Main Topics
01	Capstones, final assessment
02	Code Freeze
03	Dress rehersal
04	Capstone Showcase