Viktor Shaumann

vshaumann | http://viktorshaumann.com | viktorshaumann
(818) 455-7038 | viktor.shaumann@gmail.com | San Francisco, CA

SKILLS

LANGUAGES
Python • R • Bash • HTML • CSS

DATABASES
PostgreSQL • MySQL • MongoDB

BIG DATA
PySpark • Hive • Pig • Hadoop

DATA VISUALIZATION
D3.js • Matplotlib • ggplot2

OTHER
Excel • PowerPoint • LaTeX

EDUCATION

MS ANALYTICS
University of San Francisco
July 2016

BS ECONOMICS
University of California, San Diego
June 2013

COURSEWORK

GRADUATE
Machine Learning
Linear Regression
Relational and NoSQL Databases
Distributed Computing
Time Series Analysis
Web Analytics
Experimental Design
Data Visualization
Computation for Analytics
Exploratory Data Analysis
GIS

UNDERGRADUATE
Econometrics
Linear Algebra
Python Programming
Micro & Macro Economics
Business Strategy
Project Management
Product Marketing

EXPERIENCE

AT&T | Data Science Intern
Nov 2015 – Aug 2016 | San Francisco, CA

- Developed insights into churner's behavior by analyzing a social network graph.
- Improved existing churn models by engineering social features describing customer behavior.
- Processed and cleaned complex data sets with over 130 Billion records for the graph and churn analysis by using Pig, Hive and Python.
- Constructed a social graph and implemented a distributed community detection algorithm based on the academic paper.

BDO CONSULTING | Analyst
May 2013 – Jul 2015 | Los Angeles, CA

- Performed company and asset valuation, industry analysis and assistance through all stages of engagements.
- Designed integrated financial models with Excel to value entities, complex financial securities, intangible and fixed assets.
- Identified patterns and trends in financial and operational data, performed company, industry and economic research.
- Led an audit team of three in valuation of over 50 venture capital backed technology start-ups in the Silicon Valley.

PROJECTS

Amazon Review Usefulness Prediction
Feb 2016

Predicted usefulness measure based on the context of 4.6 million of reviews. Feature extraction with Word2Vec, modeling with Logistic Regression and Random Forest. Implemented with Apache Spark MLlib and Python. Deployed with Flask.

YouTube Trend Analysis
Nov 2015

Provided insights into trending global video consumption. Scraped metadata for YouTube videos trending worldwide using Python and Google API. Utilized Pandas and PostgreSQL to store and query the data. Visualized the data with Plotly to identify trends and relationships.

Restful Web Service
Nov 2015

Built an analytics web server (EC2) with a Flask backend and PostgreSQL database (RDS) that provided insight into Amazon beauty product data. Performed ETL of over 2 million Amazon product reviews.

Movie Review Sentiment Analysis
Aug 2015

Achieved 80% accuracy on a test set. Classified 2,000 movie reviews from polarity data set as positive or negative. Computed probabilities of each class using algorithm written in Python.