Miguel Niblock

Data Scientist/ Analyst

Summary

Experienced in accounting and the media industry.
Do **not** require visa sponsorship and willing to relocate.

Personal Info

miguelniblock@posteo.net
+1(415)-867-6356
Los Angeles, CA
github.com/miguelniblock
linkedin.com/in/miguelniblock

Skills

Machine Learning:
- Supervised and Unsupervised Machine Learning with Python in Jupyter Notebooks.
- Classification and Regression.
- Hyper-parameter tuning.

Natural Language Processing:
- Latent Semantic Analysis: TF-IDFVectorizer Feature Extraction and Singular Value Decomposition.
- Tokenization, Lemmatization, Regex String Methods, TextRank, Word2Vec, Sense2Vec.

Python Packages:
- Sci-kit Learn, Pandas, Numpy, Seaborn, Matplotlib, Keras Neural Networks, ImbLearn.
- Spacy, NLTK, Gensim, Polyglot.

Business Intelligence Tools:
- Tableau, Google Analytics, Google Data Studio, Adobe Reports, Adobe Analysis Workspace, dashboards.

Oral/ Written Communication:
- Product-focused presentations of technical material in plain-language.
- Talented exercise of political diplomacy in corporate culture.

Data Cleaning/ Processing:
- Missing data, Class-Imbalance, Binarization, Discretization, Feature Engineering / Selection.

Experimental Design:
- A/B Tests, Research Proposals, Defining Key Metrics, Evaluation and Rollout Plans.

Other:
- PostgreSQL and SQLite.
- Git, Github version-tracking.
- Linux OS & Command Line.
- Advanced Microsoft Excel.
- Spanish- 100% Written & Verbal

Recent Projects

Deep NLP: Predicting Ad Demand. <Python>
github.com/MiguelNiblock/Deep-NLP_Predicting-Ad-Demand
- Predictive model for estimating the deal probability of online classified ads.
- Used PLSR Cross-Decomposition on IDF matrix with Ngrams.
- Big Data: Implemented a memory-management script for successful batch cross-decomposition of a 64k-feature CSR matrix with 1.5Million datapoints.

Imbalanced Classification: Detecting Credit Card Fraud. <Python>
github.com/MiguelNiblock/Supervised-Learning_Credit-Card-Fraud
- Identifying fraudulent transactions in a highly unbalanced dataset.
- Implemented SMOTE: Synthetic Minority Over-Sampling Technique.
- 100% Recall and optimization of specificity via custom scoring functions.

Research Proposal: Mental Health in Tech. <Python>
github.com/MiguelNiblock/Research-Proposal_Mental-Health-in-Tech
- Investigative exploration of causes for mental illness in tech employees.
- Identification of problem and solution, key metrics, evaluation, roll-out plan, scenario planning.

Recent Experience

THINKFUL — Dec 18'- Jun 19'
Fellow Data Scientist
- Developed Machine Learning models under the mentorship of an industry professional.
- Produced Jupyter Notebooks of high quality Python code.
- Presented project and outcomes to a live audience in a product-focused manner, effectively conveying highly-technical information in plain language.
- Wrote complex SQL queries, involving new 'Window' functions.
- Set up a PostreSQL database in Linux.
- Skillfully manipulated large and noisy datasets in preparation for modeling.

VIACOM — Dec 17'- Dec 18'
Accounting Data Manager
- Designed and deployed a Digital Time Card automated system; this eliminated manual errors and decreased turnaround times by 300% in payroll processing.
- Integrated the reporting of daily labor costs ("Hot Sheets") into above-mentioned system, reducing Hot Sheets processing time by 50% and achieving data consistency across reports.
- Maintained construction budgeting system by routinely updating labor rates for all necessary labor unions, successfully preventing errors when scaling labor sources.
- Played key role in managing the department's relationship with key figures inside the company, contributing to extension of the team's contract.

Education

CALIFORNIA STATE UNIVERSITY, NORTHRIDGE — May 12'- May 17'
Bachelor Degree: Electronic Media Management
- Major Coursework: Audience Analysis, Business Law, Media Regulations, Post Production Management.
- CS/ Math Coursework: Statistics, Algorithms, Calculus, Linear Algebra, Data Structures.