DARRELL L. NELSON II

Home: 11734 Pierce St, Lake View Terrace CA 91342 | **Mobile:** (818) 912 – 9877| **Email:** darrellnelson2@outlook.com |
LinkedIn: www.linkedin.com/in/darrell-l-nelson-ii | **GitHub:** https://github.com/darrellnelson2/Data_Science_Projects

Syracuse University
M.S. Applied Data Science **Fall 2018 – December 2019**

Washington University in St. Louis
B.S. Chemical Engineering **Fall 2014 – Spring 2016**
Research: Machine Learning Approaches to Design Catalysts for C1 Chemistry

Lewis & Clark College
B.A. Chemistry **Fall 2011 – Spring 2014**
Capstone: Experimental & Computational Investigation of the Synthesis of 1,3-Dienes by Au-Catalyzed Claisen Rearrangements of Allenyl Vinyl Ethers

AWARDS

Gustav Kurt Mesmer Scholar Fall 2014 – Spring 2016
Leadership & Service Award Fall 2011 – Spring 2014
Miller Science Scholar Fall 2011 – Spring 2014

ONGOING AND RECENT PROJECTS

Syracuse University

What Makes a Super Bowl Winner? (Sports Statistician)

- Used descriptive and predictive analytics to calculate probabilities of team finishes in NFL based on regular season performance
- Valuated Machine Learning Models based on test accuracy via hold-out method
    - Models: Neural Network (NN), Gradient Boost Classifier (GBC), Support Vector Machine (SVM), and Random Forest (RF)

Achieved test accuracy ~71% w/ GBC when classifications were constrained to real world outputs

Database Management at Community Tutors, LLC. (Database Owner/Manager)

- Developed management system on SQL Server for my tutoring business
    - Built Entity-Relational, Relational, and Logical Model Diagrams
    - Able to identify high-value problems and potential bottlenecks in business operations with SQL Queries

Wing Measurement Analysis (Aerospace Structure Analyst)

- Evaluated Triumph Group aerospace structure data
- Created correlation matrix to identify interdependencies between structures
    - Trained Linear Regression and NN models on interdependent parts to predict new out-of-spec relationships
        - Produced model with ~ 74% training accuracy
- Identified strongest and weakest processes in build-flow
- Identified outliers with 3-sigma quality control

Improving College Athletics without All the Bribery (Twitter Analyst)

- Utilized text mining to evaluate probability of social media influence on college football recruiting and pre-season rankings
- Obtained Twitter API data from Python library Tweepy for all college football teams in Pacific-12 Conference
- Developed novel features based on metadata to rank model performance
- Demonstrated that the number of retweets per tweet on Twitter can influence the number and quality of recruits a team receives

Analyzing Job Market & Predicting Salary in New York City (Salary Forecasting and Job Market Assessment)

- Analyzed NYC job listings to determine demand for job skills, positions, and predict salary
- Generated NB and SVM models to predict salary range based on minimum qualifications

- Achieved training accuracy of ~70% with SVM

NOTABLE CLASSES/SKILLS

Database Administration Concepts & Database Management - *SQL*

- Examined data structures, file organizations, concepts, and principles of database management systems (DBMS)
- Executed database design, modeling, management, and implementation
- Data analysis and normalization
- Created hierarchical, network, and relational data models
- Used Microsoft's Access and SQL Server DBMS as implementation vehicles for data warehouse
- Transaction management and concurrency control, distributed databases, multitier client/server architectures, web-based database applications, NoSQL.
- Entity & Logical Relationship Diagrams w/ Draw.io & Microsoft Access
- Connected statistical programming software (R) to database for descriptive analytics

Introduction to Data Science - *R*

- Structured, semi-structured, and unstructured data collection, processing, transformation, management, and analysis
- Leveraged applied statistics, information visualization (including Geolocation), text mining and machine learning for problem solving

Scripting for Data Analysis - *Python*

- Scripted for data science pipeline
- Acquired, accessed, and transformed all 3 types of data (structured, semi-structured, and unstructured data)

Data Analytics - *R*

- Acquired and munged data
- Utilized prescriptive analytics, and Machine Learning
  - Machine Learning: Association Rule Mining, k-means Clustering, Classification (Decision Tree, Naïve-Bayes, k-Neural Network, Support Vector Machines, Random Forest), and Model Evaluation
- Preliminary Text Mining processes

Big Data Analytics - *Python*

- Acquired and munged data
- Utilized prescriptive analytics, and Machine Learning
  - Machine Learning: Association Rule Mining, k-means Clustering, Classification (Decision Tree, Naïve-Bayes, k-Neural Network, Support Vector Machines, Random Forest), and Model Evaluation
- Preliminary Text Mining processes

Text Mining - *Python*

- Corpora acquisition, document representation, information extraction, text classification and clustering, opinion mining, and topic modeling analysis
- Commercial and open-source text analysis and visualizations for pattern extraction

Data Warehouse - *SQL*

- Utilized database constructs: Operational Data Store (ODS), Data Warehouse, and Data Mart
- Understand and implement project management guidelines/techniques, requirement gathering, and dimensional modeling
- Extract Transform and Load (ETL) architecture, specification and data loading, master and reference data management, integration approaches (ETL, EII, EAI), and analytical reporting from database

Marketing Analytics – *XLStat, Excel*

- Market segmentation, market response, customer profitability, product recommendation, churn predictions, media attribution, and resource allocation models

Scripting for Data Analysis - *Python*

- Built Python scripts for data science pipeline
- Acquired, accessed, and transformed data in the forms of structured, semi-structured, and unstructured data

Business Analytics - *Excel*

- Utilized demand modeling, price optimization, customer choice (Logit, Probit, Neural Networks), and Google Analytics for business websites

Natural Language Processing - *Python*

- Processed text files for linguistic analysis, tokenization, word-level semantics, part-of-speech tagging, syntax, and semantics
- Utilized NLP techniques in the NLTK package in Python for: information retrieval, summarization, question answering, sentiment analysis, and summarization
- Used regular expressions, morphology, supervised machine learning (k-Neural Networks & Naive-Bayes), context-free grammars, parsing, semantics, and sentiment analysis

PROGRAMMING LANGUAGES

Python, R, SQL, MATLAB, and GNU Octave

MEMBERSHIPS

STEMs for Youth	2014
National Society of Black Engineers	2014
ACTS (Association of Christian Truth Seekers)	2014
Multicultural Union	2011