NATHAN DIEKEMA

(425) 879-0907 | nathandiek@icloud.com | LinkedIn: nathandiekema

Summary

Excellent at asking the right questions to derive actionable insights in a fast-paced environment. Proficient problem-solver, out-of-the box thinker, and dependable team member with strong technical skills. Enthusiastic about learning new skills and growing as a professional.

Key Competencies

- Programming | Python, R, SQL, C, Java, HTML
- Machine Learning | Sklearn, Tensorflow, Keras
- Data Processing | NumPy, SciPy, Pandas, tidyverse
- Web Scraping | BeautifulSoup, rvest, NLP
- Statistics | Econometrics, Forecasting, Excel, R
- Data Analysis | EDA, Visualization, Communication
- Data Management | MySQL, Snowflake, Mongodb
- Visualization | Tableau, Matplotlib, ggplot2
- Market Research | A/B testing, Segmentation
- Collaboration | Storytelling, Teamwork, Git, Palantir

Education

M.S. Data Analytics - California Polytechnic State University, San Luis Obispo   [June 2022]
- Relevant Coursework – Econometrics, Cloud Computing, Data Management, Data Mining, Machine Learning
- Cumulative GPA: 4.0; Academic Honors (top 10% of graduating class)
- Certifications: AWS Cloud Practitioner Badge, AWS Machine Learning Course Certificate

B.S. Electrical Engineering - California Polytechnic State University, San Luis Obispo   [June 2021]
- Minor in **Computer Science**
- Relevant Coursework – Data Structures, Continuous & Discrete Signals, Computer Vision, Neural Networks
- Major GPA: 3.75; 6-time Dean's list recipient
- Involvement: Electrical Engineering Freshman Mentor, Design Lead for Engineers Without Borders

Work Experience

Data Science Intern | Netflix   [Mar. 2022 – June 2022]
- Web-scraped external data on over 3.4 million apps from both the App Store and Google Play Store
- Trained a linear regression model to help identify important trends in the mobile gaming market
- Cleaned and derived a table of aggregate variables from 7.6 million rows of internal customer data using SQL and then utilized PCA and k-means to form distinct customer segments
- Combined external and internal findings to formulate actionable recommendations

Data Science Intern | AT&T   [Jan. 2022 - Mar. 2022]
- Utilized PySpark and SQL to clean, wrangle, and analyze hundreds of GB worth of customer data
- Developed models, examined trends, and extracted key insights to map the typical customer journey
- Segmented over 250 million customers into 7 distinct segments using K-modes, resulting in the ability to derive more personalized retention strategies, potentially reducing churn for "at-risk" customers
- Derived a customer-centric segmentation strategy and synthesized final results and business recommendations into a concise report (5 pages, 42-page appendix) and presentation (10 min)

University Projects

ECG Arrhythmia Classification using FNN | Capstone Project   [Sept. 2020 – June 2021]
- Developed an arrhythmia classification software in python utilizing the discrete wavelet transform for feature extraction and a feed-forward neural network for classification
- Achieved an accuracy of 97.87% on six of the most common arrhythmias

Face Recognition System using CNN   [April 2021 – June 2021]
- Developed a facial recognition software in python using TensorFlow and OpenCV
- The CNN was trained, tested, and improved over a couple months using a labeled dataset
- The final iteration achieved an adjusted accuracy of 85% on the open-source LFW dataset