# Derek Albosta

derek.albosta@gmail.com | (760) 710-9932 | linkedin.com/in/derek-albosta

## Summary

Data Scientist and Engineer with expertise in developing machine learning models, analytical solutions, and end-to-end data pipelines to drive business and product success. Skilled in collaborating with stakeholders to deliver data-driven insights and effectively communicating technical concepts to non-technical audiences.

## Technologies

**Languages**: Python, Java, R, SQL, HTML, CSS, C
**Frameworks**: AWS, Flask, Django, REST, GraphQL
**Tools**: Jupyter, Tableau, Git, Jenkins, Datadog
**Databases**: MongoDB, PostgreSQL, PrestoDB, mySQL, SQLite
**Data Science:** NumPy, Pandas, SciPy, Scikit-Learn, SciPy, RStudio, A/B testing
**Deep Learning:** Tensorflow, Keras, Pytorch
**Operating Systems**: Ubuntu, MacOS, Windows

## Experience

**Data Scientist & Engineer**
**Indeed.com, Austin, TX | May 2021 - May 2024**

- Designed and optimized the employer pricing model and pricing APIs for SimplyHired.com, a website used by **1.2 million monthly users**, leading to improved pricing accuracy and increased revenue.
  - **Improved pricing model accuracy (MAE) by 15% by** utilizing global-mean-shrunk encoding for high level categorical variables.
  - **Reduced model prediction latency by over 70%** by writing a script to prepopulate model predictions with a NoSQL database (MongoDB), significantly enhancing response time and scalability.
  - Evaluated the impact of the new pricing model using a rigorous **A/B testing** methodology, estimating the required sample size and time needed for a statistically significant result.
- Optimized ad campaign efficiency by reallocating **~$5M+ (5% of budget)** using a **K-means clustering model** to identify underperforming job markets on Indeed.com.
- Partnered with three core product teams at Indeed.com to **encourage data driven decision making,** analyzing key data insights, presenting findings and **leading stakeholders to actionable solutions.**
- Created/maintained **real-time monitoring dashboards in Tableau and Datadog**, enabling marketing and product leaders to quickly respond to changes in campaign performance and optimize strategy
- Delivered reliable and scalable engineering solutions by following rigorous testing, clean coding standards, and **agile development practices**, supporting the long-term stability of high-impact applications.

**Machine Learning Intern**
**Juva Health, San Diego, CA | Dec 2020 - May 2021**

- Researched and implemented a **video de-identification pipeline** using a deep learning auto-encoder architecture with Keras/TensorFlow, **safeguarding user privacy in alignment with HIPAA standards.**
- Built Tableau dashboards for analyzing user behavior, providing insights to improve mobile app engagement and retention.

## Education

B.S. in Computer Science, Minor in Mathematics
University of Puget Sound, Tacoma, WA | 2016 – 2020

## Projects

**ChirpGAN: Generative Modeling for Bird Call Synthesis**
A GAN-based tool for synthesizing artificial bird calls using a custom data pipeline, progressive GAN architecture, and wavelet scalograms.