Corbin Ulloa

442-226-8640 | corbinulloa619@gmail.com | linkedin.com/in/corbin-ulloa

EDUCATION

Bachelor of Science, Computer Science **Expected Graduation:** December 2023
California State University, San Bernardino
Relevant Coursework: Data Structures, Algorithm Analysis, AI & Machine Learning, Software Engineering

WORK EXPERIENCE

Data Science Intern September 2022 – Present
California State University, San Bernardino *San Bernardino, CA*
- Led the development of an ETL pipeline which processed 9,000,000+ student records using PySpark, transformed and cleaned data using SparkSQL with Pandas, and loaded data into Oracle Database using SQLAlchemy
- Extracted data from Oracle Database using SQL to perform data cleansing and aggregation
- Created new metrics and dashboards using data aggregation methods within Tableau and Excel
- Performed EDA using Python libraries to identify and visualize relationships or anomalous events for several ad-hoc requests

Virtual Reality Developer October 2021 – September 2022
California State University, San Bernardino – xREAL Lab *San Bernardino, CA*
- Optimized VR gameplay through developing and debugging features using C++, reducing development time by 30%
- Articulated highly technical methods to diverse teams to shape decision-making and integrate agile software development
- Collaborated with stakeholders to update on project features, give demos, and provide project timelines

Data Science Intern June 2021 – August 2021
University of California, Santa Barbara - CalCOFI *Santa Barbara, CA*
- Modeled larval fish abundances in relation to pH of the water as an indicator of climate change using Linear Regression and Interpolation (Kriging) to estimate water values for locations where data was not collected
- Combined larval fish and water characteristic datasets from 1904 to present, identifying correlations between larval fish abundance, pH, and other water characteristics
- Applied K-means clustering to detect fish abundance patterns and grouped data into coastal and oceanic groups
- Improved the Linear Regression model's accuracy to 96% providing further observations on the effects of water characteristics towards predicting future states of marine life

Research Analyst Intern June 2020 – August 2021
California State University, San Bernardino *San Bernardino, CA*
- Conducted feature engineering, data wrangling, and merging 3+ datasets to increase data-driven decision-making
- Developed visuals with Tableau and used MS Excel to organize and analyze statistical student data
- Led a time series analysis on student performance rates to improve instructional support and uphold student success outcomes

PROJECTS

Real Estate Price Prediction Model: github.com/cdulloa/Real-Estate-Price-Prediction-Model
- Built a model using Scikit-Learn and Linear Regression to approximate the price for houses using a house dataset
- Achieved an accuracy of 88% in predicting the real-estate price with the Linear Regression model

Deep Learning Facial Recognition System: github.com/cdulloa/Deep-Learning-Facial-Recognition-System
- Developed a Siamese Neural Network model using Keras for a one-shot image recognition to authenticate a user
- Built a data pipeline to load samples from preprocessed images that returns an accuracy greater than 90% for prediction
- Evaluated the effectiveness of the model's performance with Precision and Recall returning a F1-score of 0.962

SKILLS

Programming Languages: Python (PySpark, TensorFlow, Pandas, Matplotlib, Scikit-Learn, NumPy), SQL, C++, JavaScript
Data Science & Technologies: Apache Spark, MS Excel, AWS, Git, Agile Scrum Methodology, Tableau, Oracle, Linux CLI
Certifications: Google Data Analytics, Accelerated Data Science