

# Jayson Alden Villena

Email: [Jayson.villena@gmail.com](mailto:Jayson.villena@gmail.com) | Cell: +1 (925) 285-2593 | Github: <https://github.com/jay28son>

Data scientist who enjoys exploring through data and developing models that predict future outcomes. My degree in Mechanical Engineering assisted me in developing a technical mindset which helped in breaking down technical complexities into simplified narratives. Looking to apply this mindset along with my passion for data to increase my knowledge and application of Machine Learning processes.

## Technical Skills

- Data Science: Machine Learning, Python, Pandas, Git, NLP, TensorFlow, Clustering
- Data Analysis: SQL, Excel
- Data Visualization : Tableau, Seaborns, Matplotlib

## Honors/Certificates

- [Google Data Analytics Certificate \(2022\)](#)
- Dean's Honor List (2016-2017, 2020-2021)

## Experience:

### General Assembly

November 2022- Present

#### Data Science Fellow

- **NFL Play Predictor**
  - Executed deep learning and decision tree algorithms to produce a model that can accurately classify 70% of the plays, 18% higher than our baseline model, and also achieved an average f1 score of 70% across the classes
  - Implemented data replacement techniques to balance the minority and majority classes
  - Generated and deployed a StreamLit that predicted a play via the given inputs
- **Covid 19 Classification**
  - Developed a Convolutional Neural Network Model that multi-classified radiology images that had a confidence interval of 90%
  - Implemented Data Augmentation techniques utilizing Tensorflow and Albumentations to help increase the models performance
  - Utilized Tensorflow to implement a Gradient-CAM visualizations to increase models interpretability
- **Reddit Analysis**
  - Utilized PRAW API to scrape posts from two SubReddits, collected 5000 posts from each SubReddit and made use of Natural Language Processing techniques to help classify the posts
  - Created XgBoost and Bagging Models and applied GridSearch for hyperparameter tuning, created confidence intervals of 71%, a 21% increase of our baseline model

### Tesla

November 2021-November 2022

#### Production Associate:

- Improved a battery manufacturing process in Excel by incorporating functions such as VLOOKUP, IF, COUNTIF and features such as Conditional Formatting and Sort/Filter, increased production time by 7%.
- Streamlined the manufacturing process to be more efficient such as 3D printing a tool for a manufacturing process

## Personal Projects:

### NFL Salary Analysis (<https://github.com/jay28son/NFL-Salary-Analysis-Project>)

August 2022 - September 2022

- Gathered data of the top one thousand paid players each year over the last 10 years by web scraping an external website
- Performed ETL on the csv files into a MySQL database for data analysis, such as discovering the player market increased by 79% from 2013-2022

## Education:

### California State University, Sacramento

B.S. - Mechanical Engineering

August 2016- June 2021