Sean Marcus Pereira

Long Beach, CA | Email: sean.marcus.pereira@gmail.com **LinkedIn** **GitHub** **Mobile**: (858) 254-9831

SUMMARY

MS (Computer Science); Experience in Full Stack Web Development **(React.js, Node.js, Express, MongoDB, PostgreSQL),** Back End Development **(Python, Django, Flask);** Currently working on a **REST API** using **Python**. Experience in **Docker, AWS,** and **Git. Research Driven** with **4** published research papers with publishers like **IEEE**.

EDUCATION

California State University Long Beach, U.S.A. **Aug 2019 - May 2021**
Master's in Computer Science CGPA (3.3/4.0)

TECHNICAL SKILLS

- **Languages**: Python, C, Java, C++, C#, R, SQL.
- **Databases:** MongoDB, PostgreSQL, MySQL, Firebase.
- **Web Stack**: HTML5, CSS3, JavaScript, React.js, Node.js, D3.js, Bootstrap, Flask, Django, Express.js, AJAX, jQuery, JSON, Angular.js.
- **Web Services:** RESTful, XML, SOAP, Spring MVC, Hibernate.
- **Web Servers:** Digital Ocean, Heroku, Cloudflare, Amazon AWS EC2.
- **Other Technologies**: Tableau, Git, SoapUI, Postman, Web Scraping, ML, Babel, Linux, Docker, C1/CD, GitLab, Jenkins, Apache Groovy, Hadoop, Apache Spark.

WORK EXPERIENCE

Teaching and Research Assistant - California State University Long Beach **Aug 2020 - Current**

- Instructing and assisting 60-70 students with Python, MySQL, and Tableau technologies, include helping and reviewing lesson plans and conducting doubt sessions for students over Zoom calls.
- Conducting and grading more than 100 quizzes, assignments, and assisting Professor with Research work on numerous projects, involves coordination with Graduate Assistants of various Universities.
- **Technology: Python, Selenium, Beautiful Soup, Twitch API, NLP, Machine Learning, Amazon AWS EC2.**

Project Intern- Tech Mahindra Ltd, Mumbai India **June 2018 - July 2018**

- Drafted 7 use cases to research on Application of Artificial Intelligence in Business Operational Systems (B/OSS) of Communication Providers.
- Delivered presentation of 7 use cases to senior stakeholders within the company; one use case was shortlisted.
- Developed technical architecture for "Call Drop & Network Fault Prediction and Management" was selected for Tech Mahindra's Telecom customer development.

Industrial Intern- Honeywell Automation India Ltd, Pune India **June 2017 - July 2017**

- Evaluated code for Process History Database (P.H.D) that records data from a Historian in a tabulated format using Excel Companion and worked on a research project on Industrial Internet of Things (IIoT) preparation a PowerPoint Presentation and presented to Stakeholders.

PROJECTS

Twitch API Analytics **Dec 2020- Present**

- Extracting data using Twitch Developer Console to gain authorization to the Twitch server for data retrieval.
- Analyzed Data obtained via Web Scraping of 10,000 users through the connection made to get each user's information.
- Executed the process by creating 100 parallel sub-processes and firing them simultaneously to reduce latency.
- Performed segregation of each data type in separate CSV files obtained when each sub-process was fired; each file will act as an input for performing sentimental analysis using NLP.
- **Technology: Python.**

Web Scraping Dow Jones Website for George Mason University **Oct 2020- Present**

- Spearheaded and improved a Web scraping code for George Mason University in coordination with the university's research assistant.

- Extracted and converted information of 10000 companies from Dow Jones website to a CSV file, further programmed sentimental analysis using N.L.P from the data obtained.
- **Technology: Python (NumPy, pandas, sklearn, Selenium, Beautiful Soup).**

Multi-Layer Perceptron from scratch **Nov 2020 - Dec 2020**

- Developed a Multi-Layer Perceptron (M.L.P.) back-propagation network style of artificial neural network classifier. A Single M.L.P. was constructed with one hidden layer and one multi-class output layer. ANN/MLP libraries were not utilized for the project, i.e., code was written from scratch. The accuracy obtained was 75%.
- **Technology: Python.**

Interactive visualization from eye gaze dataset using D3.js **Nov 2020 – Dec 2020**

- Built, designed, and engineered an interactive visualization using the given Dataset that captures eye gaze recorded during a human-computer interaction session. JavaScript D3 library was utilized. Project aims to provide interactive visualization support to users in examining whether a particular trend/pattern is present.
- **Technology: JavaScript, HTML, CSS, D3.js, Java.**

Building Decision Tree from scratch and Ensemble the Tree **Oct 2020- Nov 2020**

- Created and shaped an automated Ensemble Decision Tree Builder and wrote a binary pattern recognition Decision Tree Ensemble using Builder using KRK training dataset.
- Built Two Decision Trees for the Ensemble, first with an initial selection of vectors and second with the same number but of boosted feature vectors.
- Streamlined an ensemble, structured as a weighted vote of two Decision Trees based on both the model's accuracy resulting in a boosted accuracy of 88%.
- Executed crucial tasks in all phases of Software Development Life Cycle (SDLC) following Agile Methodology.
- **Technology: Python.**

Interactive Word Cloud Generator **Sept 2020- Oct 2020**

- Implemented an interactive tag cloud to visualize text of 300 words. The input was a free text extracted from a Wikipedia page.
- Developed a placement algorithm using Python, with the team's help, to place the word on the canvas in various orientations (0 or 90) and colors.
- Programmed Interactive Features, clicking on the word, directs to the Wikipedia page, and numerous animations were integrated on canvas using JavaScript, HTML, and CSS.
- Connected Python API with JavaScript Frontend using Flask.
- **Technology: JavaScript, HTML, CSS, Python (Flask, PIL).**

Sudden Cardiac Death (S.C.D.) Prediction using E.C.G. Machine **Aug 2018 – Dec 2019**

- Engineered a Sudden Cardiac Death prediction model using Python, where a database was formulated based on a combination of E.C.G. Data generated from an E.C.G. Machine and an online Medical database (PhysioNet).
- Managed testing and analysis of various Machine Learning algorithms on Dataset, Classification, and Regression Tree Model (CART), which resulted in the highest accuracy of 83% compared to the other models.
- Signal Processing was done on signal extracted from the E.C.G machine, using LabView and MATLAB.
- **Technology: Python, LabView, MATLAB.**

Railway Crowd Management System for Suburban Railway using Image Processing **Jan 2018 -- May 2018**

- Co-led a team to develop a passenger safety APP where Crowd details of Railway Compartments and Foot Over Bridge (F.O.B.) were displayed.
- Collected data with cameras installed at various Compartments, F.O.B., and Booking counters, data was processed and analyzed using Image Processing algorithms and tools. An accuracy of 90% was obtained of the number of passengers detected in each frame by filtering each snapshot's noise.
- **Technology: Python, Embedded C, Java.**

PUBLICATIONS

- Prediction of Sudden Cardiac Death using Classification and Regression Tree Model
- A.I. Use Cases in Operational Support System and Business Support System
- C-Indicator: Crowd Management System for Suburban Railway using Image Processing
- STEM EDUCATION: why math is so important?