Justin Lai

justinlslai@gmail.com | linkedin.com/in/lai-justin | github.com/Lai-Justin | lai-justin.github.io

Education

University of California, Santa Barbara September 2019 – June 2023
Bachelor of Science in Computer Science *GPA: 3.61*

Experience

Machine Learning Engineer Intern June 2022 – August 2022
Sermonis.ai *Berlin, Germany*

- Diversified training data by developing a series of text scrapers for a variety of data sources
- Accelerated data processing speeds by automating data cleaning, labeling, and lemmatization processes
- Tokenized text data using the BERT transformer for improved natural language processing model training
- Deployed ElasticSearch to index and search tokenized data for efficient querying and retrieval

Frontend Web Developer Intern June 2018 – August 2018
University of California, San Francisco *San Francisco, California*

- Improved usability and user productivity by developing frontend features such as task lists, reminders, and toolbars
- Enhanced accessibility, viewing experience, and content delivery speed by revising backend elements
- Automated medical document processing for metadata extraction and redaction for confidentiality and compliance
- Performed rigorous unit and integration testing with JUnit, Selenium, and Mockito

Technical Skills

Languages: Python, C++, C, JavaScript, TypeScript, Java, SQL, R, OCaml, C#, HTML5, CSS
Tools: Git, Unix/Linux, Postman, PostgreSQL, Docker, ElasticSearch, Kibana, AWS EC2, Kanban
Frameworks: React, Node, MongoDB, Express, Flask, JUnit, RESTful API, Spring Boot, CodeCov, Storybook
Processes: Agile, CI/CD, Scrum, Object-Oriented Programming, Waterfall Development, Design Patterns
Libraries: Numpy, Pandas, Seaborn, RSQLite, Scikit-Learn, TensorFlow, PyTorch, HuggingFace, Keras, OpenCV, Spark

Projects

CIFAR-10 Convolutional Neural Network Image Classifier | *Python, TensorFlow, Keras, Scikit-Learn, Matplotlib*

- Developed a deep learning model for classifying 32x32 colored images of objects, based on the CIFAR-10 dataset
- Implemented a multi-layer CNN with batch normalization, max pooling, and dropout layers
- Monitored training and validation convergence using EarlyStopping
- Visualized model training and outputs to analyze CNN feature extraction and model learning

Four Year Schedule Planner | *JavaScript, Python, React.js, Node.js, Express.js, Mongoose*

- Developed a MERN-based course scheduling web app for UC Santa Barbara students to plan four-year schedules
- Implemented a web scraper to gather up-to-date course IDs, titles, and grade distributions
- Utilized Mongoose and React.js for efficient data storage and dynamic, responsive user experience
- Ensured relevance and usability by surveying professors and peers to improve and iterate over application features

Ecommerce Linear Regression and Data Analysis | *Python, Pandas, Seaborn, Matplotlib, Scikit-Learn*

- Analyzed customer purchase data to optimize online storefronts and increase annual sales for a clothing retailer
- Conducted exploratory data analysis on key purchasing metrics such as customer application usage
- Identified factors influencing annual customer spending by training a Linear Regression model
- Evaluated model performance and accuracy using metrics such as Root Mean Squared Error
- Provided actionable insights to improve customer experience and increase annual sales

Temperature-based Automated Window Control Software | *Python, GPIOZero, PostgreSQL, Flask, Request*

- Developed an automated window control software using a Raspberry Pi, temperature sensors, and motors
- Integrated geolocation and NationalWeatherService APIs to fetch outdoor temperature and other weather data
- Utilized PostgreSQL to store indoor temperature data for machine learning analysis
- Created a Flask server to display real-time indoor/outdoor temperatures and window status