Denis Leang

dsl2179@columbia.edu linkedin.com/in/denis-leang github.com/YeriAddict yeriaddict.github.io/

Education

Columbia University Sep 2023 - Dec 2024
M.S. in Data Science *New York, NY*
- **Relevant Coursework:** Machine Learning, Artificial Intelligence, Computer Vision, Statistical Inference, Exploratory Data Analysis & Visualization, Computer Systems

Telecom Saint-Etienne Sep 2020 - Aug 2023
B.S./M.S. in Computer Science & Mathematics (Diplome d'ingenieurs) *Saint-Etienne, FRA*
- Prepared for the French national entrance examinations to Grandes Ecoles at Lycee du Parc *(2017–2020 — Lyon, FRA)*
- **Relevant Coursework:** Data Structures & Algorithms, Database and Information Systems, Operating Systems, Distributed Computing, Cloud Computing, Big Data, Computer Graphics, Probability & Statistics, Optimization & Estimation methods

Skills

Technical: Proficient in Python, Java, TypeScript, React, HTML, CSS, SQL, pandas, scikit-learn, Git, Agile. Experience in Dart, Flutter, R, polars, OpenCV, Tensorflow, PyTorch, NLP, Deep Learning, Azure, Spring Boot, Vue.js, Next.js, Tailwind, Docker, Spark, Design Patterns. Used Kubernetes, AWS, GCP, Kafka, MongoDB, Redux, Leaflet, Jira, Asana.
Spoken Languages: French, German, Khmer

Experience

Analysis Group Jun 2024 - Aug 2024
Data Scientist Intern *Boston, MA*
- Revamped a CRUD patient survey management app built with C# (.NET) and React.js, enabling more efficient distribution of clinical drug testing surveys by honing workflows and strengthening vendor engagement
- Reduced user navigation time by 50% by redesigning the application's structure using scalable React components and optimizing API calls, improving overall user efficiency
- Executed Python scripts leveraging the Etherscan API to analyze cryptocurrency transactions across multiple clusters, tracking billions of dollars on Binance and facilitating investigations into suspicious wallet activity

Bionomous SA Feb 2023 - Aug 2023
Software Engineer Intern *Villaz Saint-Pierre, CHE*
- Developed a CRUD client update and support portal written in Java Spring Boot and Vue.js to automate software updates for in-house products, eliminating manual engineer involvement and improving customer experience and release management
- Implemented a secure authentication system with Microsoft Azure Active Directory and JWT tokens, boosting login efficiency and strengthening security for users and administrators
- Upgraded company products (Java for Android, Flutter/Dart for Android/iOS) by incorporating client-requested functionalities, increasing usability and customer satisfaction

Projects

Portfolio Website
- Designed a personal portfolio in React.js using the App Router paradigm and deployed it as a static website using GitHub Pages

Spotify Dashboard
- Built a Spotify listening habits dashboard in React.js with real-time, dynamically fetched data from the Spotify and Last.fm APIs, visualizing trends through interactive charts
- Refined data fetching for real-time music analytics by executing an efficient caching strategy and a CI/CD pipeline with GitHub Actions, reducing memory usage and query time by 4×

Iris Recognition Model
- Engineered an iris recognition package with OpenCV and Python, incorporating preprocessing, feature extraction, and LDA-based matching to support one-to-one and one-to-many biometric recognition modes
- Achieved a consistent 95% Correct Recognition Rate (CRR) by implementing a Gabor-based multichannel filtering technique, enhancing feature extraction across L1, L2, and Cosine similarity metrics

Space Simulation Game
- Optimized memory efficiency in C++/OpenGL by dynamically rendering space elements (asteroids, planets, starship) to prevent memory leaks and improve performance
- Integrated real-time starship navigation by leveraging OpenCV for hand movement detection, enhancing interactivity and gameplay fluidity

Electronic Health Record (EHR) Study on Interstitial Lung Disease with Columbia Medical School
- Conducted data wrangling in Python to harmonize large-scale EHR data across multiple health systems
- Fine-tuned Named Entity Recognition (NER) and BERT-based models on medical text to analyze chest CT reports, achieving an 30% improvement in accuracy for identifying drug-disease associations, contributing to enhanced clinical treatment guidelines