Joel Hanson

joelhanson025@gmail.com
LinkedIn +91-815-6978-004
Portfolio GitHub

Skills

- Python, JAVA, GoLang, JavaScript, TypeScript, SQL, Bash, Django, Flask, Pandas, Numpy, Pytorch, SQLAlchemy, FastAPI, Celery, NodeJS, React, NextJS, jQuery, REST APIs, FastAPI
- Git/Github, Github Actions, Cloud Computing, CI/CD, Unit Testing, Docker, Kubernetes, Openshift, PostgreSQL, Jenkins, Kubernetes Operators, Kafka, Kafka Connectors (sink/source), Strimzi Kafka Operator, Microservices, Functional Programming, Head First Design Patterns, Clean Code, CircleCI, Agile methodology
- Machine Learning Model Deployment, ML Model Training Pipeline, Scikit-learn, TensorFlow, Keras, Natural Language Processing (NLP), Computer Vision, Deep Learning, Linux, Google Cloud Platform (GCP), MLFlow, PyCaret, IBM Cloud, AWS EC2, ECS, CDK, SQS, SNS, Glue, Fargate, Redshift, Lambda functions, RDS, S3, Cloudformation, Cloud Development Kit, IBM Message Queue, IBM MQ Sink/Source Kafka Connectors, Neo4j, Elasticsearch, Redis, RabbitMQ – *All professional proficiency or above*

Experience

- **IBM - International Business Machines**
  Software Engineer *August 2021 - Present*
  - **IBM Event Streams**: • Architected cutting-edge cloud infrastructure to effortlessly handle billions of events, implementing robust information security measures. • Collaborated closely with the open-source project Strimzi Apache Kafka on Kubernetes, enabling seamless integration and optimal performance. • Achieved an impressive uptime of 99.9% for the IBM Event Streams platform, delivering reliable and uninterrupted event processing. • Improved the deployment process of Kafka clusters, significantly reducing deployment time by 90%.
  - **Kafka Connectors**: • Developed and meticulously maintained official open-source IBM Kafka Connectors, facilitating smooth integration with both IBM Cloud services and third-party services. • Ensured compatibility with the latest MQ features and implemented exactly-once semantics in IBM MQ Sink and Source Connectors, allowing reliable and guaranteed message delivery to end consumers. • Streamlined data transfer between IBM Cloud services and third-party services, resulting in an impressive 40% reduction in data integration time. • Obtained 95% test coverage for Kafka Connectors, ensuring high-quality and reliable data integration.
  - **Kubectl CLI Plugin**: • Enhanced the user experience by 20%, resulting in a 15% increase in user adoption and satisfaction. • Optimized API interactions, eliminating manual effort by 90% and saving an average of 5 hours per week. Allowing customers to use the CLI with the CNCF Kubernetes cluster
  - **Jenkins Nightly Pipeline**: • Designed and configured over 15 Jenkins components to run as a nightly routine. • Established the maintenance of a healthy codebase through regular code checks and tests.
  - **Mentoring And Leadership**: • Educated and mentored 6 candidates, fostering their development as future professionals in collaboration with two academic institutions.
- **Impress.ai** StartUp
  AI Engineer / Software Engineer *September 2018 - August 2021*
  - **Automated Candidate Evaluation - Chatbot-based pipeline for applicant assessment**: • Led the design and implementation of a project utilizing machine learning to evaluate candidates. • Reduced the time needed to assess over 4,000 applications from 470 hours to only 2 hours for a renowned Singapore school. • This project resulted in a 95% reduction in the time required to assess applicants, and a 10% increase in the number of qualified applicants. • Employed innovative chatbot technology to automate the candidate evaluation process, improving efficiency and accuracy.
  - **Automated Data Migration**: • Applied ETL methods to migrate client assets from the user acceptance testing platform to production. • Reduced the time needed for releasing the product by 90% by automating the process of moving required data from one server to another. • This project resulted in a 20% increase in the number of products released per month.
  - **Refactoring And Unit Testing**: • Reorganized the platform, delivering substantial improvements in performance and functionality. • Implemented unit tests to assure proper platform critical area coverage. • Increased test coverage from 4% to 29%, enhancing the platform's stability and reliability.
- **Travidux Technologies** StartUp
  Software Engineer *October 2017 - September 2018*
  - **Booking Platform**: • Collaborated with the Kerala Tourism Development Cooperation (KTDC) to create an online travel booking portal, resulting in a 20% increase in bookings and revenue for the organization. • Led the development of the platform, maintaining a smooth user experience and incorporating essential features for efficient booking. • Successfully integrated a secure payment gateway, resulting in enhanced transaction security, increased user trust, and generating a profit of INR 500,000.

- **Source Code Management And CI/CD Pipelines**: • Restructured three projects to implement version control using Git, improving better collaboration and code management. • Introduced CI/CD pipelines to enable automated website deployment. • Leveraged locally hosted GitLab to streamline the deployment process and ensure continuous integration. • Established code style checks to maintain code quality and consistency.

Education

- **Calicut University** **(Sahrdaya College of Engineering)** KL, India
  Bachelor of Technology (B.Tech) in Computer Science and Engineering *June 2013 – June 2017*

Projects

- **Speech-To-Text Microservice Demo Application *(April 2023)***: ◦ Designed and put together a Speech-To-Text demo application using a microservice architecture with a REST API. ◦ Put into effect functionalities that allowed users to upload audio files and receive accurate text output on the frontend. ◦ Leveraged technologies such as Django, Celery, Redis, RabbitMQ, and PostgresSQL to ensure flawless integration and efficient processing of audio data.
- **ML Experiment Tracker using Django and Celery *(May 2023)***: ◦ Constructed an ML experiment tracking tool using Django, Celery, and Django Rest Framework. ◦ Built a comprehensive backend system that enabled the efficient monitoring and tracking of machine learning experiments.. ◦ Utilized Django Celery Results and monitored experiment progress using Flower. ◦ Authored a blog post on Medium to share insights and knowledge about the ML experiment tracker, reaching a wide audience.
- **Webservice for Pycaret AutoML *(January 2022)***: ◦ Developed a webservice for Pycaret AutoML, a RESTful API web application that performs classification and regression on structured data. ◦ Created advanced features, including a central ML model store with versioning, allowing collaborative management of the full lifecycle of ML models. ◦ Utilized Kubernetes API to enable efficient deployment of ML models.
- **Large Scale ETL Pipeline (Python, SQL, AWS, Redshift) *(July 2021)***: ◦ Built and launched a large scale ETL pipeline, enabling the extraction, transformation, and loading of data from massive files to Redshift. ◦ Formulated and optimized the pipeline to handle files exceeding 10GB in size, achieving efficient processing and storage. ◦ Collaborated with cross-functional teams to determine data requirements and to guarantee seamless integration with existing systems and data sources. ◦ Achieved a significant increase in data processing efficiency, reducing processing time by 30% and facilitating timely insights and analytics.
- **Autonomize.ai (Healthcare data insights at scale) *(August 2021)***: ◦ Contributed to the creation of a trusted AI platform that empowers healthcare organizations to leverage data effectively, enhancing human-centered experiences with the right oversight and accountability. ◦ Collaborated closely with a team dedicated to making the world's biomedical knowledge accessible for improved health outcomes. ◦ Engineered a scalable architecture that processed millions of files, simplifying data extraction and search capabilities. ◦ Throughout the project lifecycle, played a crucial role in providing the accuracy, quality, and security of the data. ◦ Employed Elasticsearch, Python, and AWS technologies such as Lambda, SQS, SNS, AWS Glue, Opensearch (Elasticsearch, Kibana), Typescript, Cloudformation, and CDK.
- **Health Informatics (OCR, NLP, Data Analysis with Neo4j and Elasticsearch) *(November 2021)***: ◦ Created a health informatics project incorporating OCR and NLP techniques for entity recognition and relation extraction. ◦ Integrated the project with Neo4j and Elasticsearch to analyze patient journeys, treatment patterns, and disease patterns based on healthcare data. ◦ Put in place a scalable and efficient architecture using microservices, ensuring robust data analysis capabilities.

Publications

- **"Survey On Image Processing Based Plant Leaf Disease Detection" (2016)**: Conducted a survey on some of the existing image processing methods for leaf disease detection and examined common machine learning methods.
- **"Plant Leaf Disease Detection Using Deep Learning And Convolutional Neural Network" (2017)**: A solution to solve this using transfer learning. We came up with a production-ready web application to detect diseases from plant leaf images.