

ANISH SHRESTHA

✉ sir.anishshrestha@gmail.com  [linkedin.com/in/anyesh](https://www.linkedin.com/in/anyesh)  github.com/Anyesh

Education

Lord Buddha Education Foundation (APU)

Bachelor of Science in Computer Science

2016 – 2019

Kathmandu, Nepal

Experience

AlayaCare

Software Engineer

May 2022 – Present

Sydney, Australia

- Developed and optimized backend features for Home Care Package (HCP) and the National Disability Insurance Scheme (NDIS), including invoicing, claiming, smart reconciliation, and budget management, addressing key pain points for customers and support teams.
- Implemented global features such as auto travel time in a cloud-based SAAS environment, enhancing functionality and user experience for international clients.
- Collaborated with a North American team to integrate regional and global functionalities, ensuring seamless operation and compliance with Australian standards.
- Founding member of a core product delivery team, providing strategic value for Australian clients while driving technical design and collaborating closely with North American Engineering teams.

Fusemachines Nepal

Machine Learning Engineer (TIME Magazine)

Feb 2020 – Nov 2021

Kathmandu, Nepal

- Collaborated with the best Data Team of the top US news magazine to streamline the data pipelines that provides brand audits, insights, and strategic recommendations that drive business outcomes at a fraction of the cost.
- Worked on various big data platforms of the news magazine ecosystem to ingest the scattered data in a single OLAP platform to analyze and get better insights.
- Collaborated with the Data VP and a Ph.D. to streamline the deployment of the Machine learning pipeline for viral content prediction.
- Collaborated with the senior product manager to prepare aggregated data reports that are used to get product insights and better strategy development.

Machine Learning Engineer (Hospital For Special Surgery)

- Collaborated with senior Data Scientists and Ph.D. of the US number 1 Orthopedics to build a solution for transforming their implant supply chain by determining a patient's implant needs and share it further upstream to manufacturers to optimize inventory and day-of operations..
- Using data from 15,000+ cases, we built an implant size prediction algorithm that utilizes manufacturers' implant data (e.g., product dimensions, design), patient demographics (i.e., sex at birth, height, weight, date of birth), and historical surgeon usage data to output a case-specific implant order.
- Machine learning models (Ridge Regression and Neural Network Classification) were built to predict implants required for a primary total knee replacement with an overall accuracy above 90%.

Machine Learning Engineer (Fuse AI)

- Collaborated with Machine Learning Ph.D. to re-structure and re-create the best machine learning and deep learning academic courses which are being taught in Fuse Classroom with tens of thousands of students worldwide.
- Worked on a core ML and NLP algorithm implementation at the Maths level to better explain the working mechanics of the algorithm to students.
- Collaborated with senior NLP Engineers to create the best NLP course focusing on the NLP starters.
- Involved in the core Machine Learning algorithm implementation and academic course development by analyzing the level of the students.

Xcellab

Frontend Engineer

Oct 2019 – Feb 2020

Lalitpur, Nepal

- Front end development of a crypto exchange platform.
- Worked with Devops and senior engineers to secure and manage the architecture for deployment of JavaScript backend and frontend services.
- Collaborated with the security teams on the best practices for user authentication and secured trading management.

Projects

Text-to-Image Generative AI with Advanced Image Analysis | *Python, GCP, Stable Diffusion*

- Developed an AI tool using Stable Diffusion for generating high-quality images from text, integrating custom models for age/gender detection and image captioning, attracting 1500 users within the first few months
- Designed and implemented an end-to-end architecture, optimizing performance through rigorous testing and fine-tuning, leading to hundreds of images generated daily and a 60% reduction in costs through efficient resource management
- Achieved high user satisfaction and financial sustainability with enough paid users to cover costs, continuously improving the product based on feedback and collaboration with cross-functional teams for a seamless user experience

Knee Transplant Size Estimation via Patient Demographics | *Python, sklearn, AWS*

- Researched and developed the best possible ML solution in collaborations with Ph.D. to predict implants required for a primary total knee replacement with an overall accuracy above 90% .

GrowBot AI (Viral content forecast) | *Python, GCP, Vertex AI*

- Built an AI model that gives a list of predictions of upcoming viral news given the previous history with an overall accuracy above 90%, so that sales/marketing team can focus more on those news.

Technical Skills

Languages: Python, JavaScript, TypeScript, SQL, NoSQL

Technologies/Frameworks: Linux, Git, GitHub, Docker, Figma, NextJS, React, Svelte, GraphQL, Django, Flask, Redis, Celery, Sklearn, Pytorch, Tensorflow/Keras, Pymc, VS Code, Google Cloud Platform, AWS, JIRA

Certifications

TensorFlow Developer Certification: <https://bit.ly/3i46gp3>

Rasa Developer Certification: <https://bit.ly/3iFJ7KP>

Rasa Advanced Deployment Certification: <https://bit.ly/3vmJaOk>

Sequence Models by deeplearning.ai: <https://bit.ly/3as2XSG>

Mathematics for Machin Learning: <https://bit.ly/3CJ3LC6>

Machine Learning by Stanford Online: <https://bit.ly/2wBuTFa>

Getting started with Machine Learning by AWS: <http://bit.ly/2xfCPw0>

Leadership / Extracurricular

Google I/O Extended

June 2024

Panelist

NSW Teachers Federation Conference Centre, Sydney

- Panel Discussion on What does the future of AI hold for us

Google Developer Student Clubs — Google Labs

May 2024

Speaker

Google HQ, Sydney

- Delivered a presentation at the GDSC Google labs event, sharing insights on leveraging open-source text-to-image generative AI technology for building practical applications.
- Discussed the high-level workings of text-to-image generative AI models and demonstrated an end-to-end side project with a user base of over 1500, deployed in production.
- Provided practical guidance on productionizing AI projects using Google Cloud Platform (GCP) resources, emphasizing cost-effective deployment strategies.

Google Developer Group — Google Cloud

April 2024

Speaker

Google Developer Group, Sydney

- Delivered a comprehensive talk at the GDG event, showcasing the Dreamery project – a generative AI platform for transforming selfies into artwork.
- Discussed the architectural decisions behind Dreamery, including leveraging a distributed service approach over a monolithic design, applying clean code architecture principles, and implementing domain-driven design.
- Demonstrated cost-effective deployment strategies by hosting backend services on Cloud Run and leveraging serverless GPU instances for Stable Diffusion XL, resulting in a 90% cost reduction.
- Highlighted additional cost optimization techniques, such as right-sizing resources, which led to a 60% reduction in overall GCP costs.
- Conducted a live demo, exhibiting Dreamery’s capabilities by transforming an example selfie into professional headshots, showcasing the power of generative AI.

Fuse AI Training

Nov 2020 – Dec 2020

AI Instructor

Fusemachines

- AI training provided to associate and trainee ML Engineers.
- Managed the whole schedule and curriculum by in collaboration with other instructors.
- Led the end-to-end ML pipeline development to deployment workshop.