Isha Patel

Irvine, CA | +1 (949) 337-2324 | ishamp@uci.edu | linkedin.com/in/ishapatel1108 | github.com/Isha1108 | isha1108.github.io/

EDUCATION

- **University of California, Irvine | Irvine, CA** — Sep 2023 - Dec 2024
  Master of Computer Science
  Coursework: Graphical Models and Statistical Learning, Artificial Intelligence, Database and Data Management
- **K. J. Somaiya Institute of Technology | Mumbai, India** — Aug 2019 - May 2023
  Bachelor of Technology - Computer Engineering
  Coursework: Software Engineering, Data Structures and Algorithms, Applied Data Science, Machine Learning, NLP

TECHNICAL SKILLS

Programming Languages: Python, C++/C, Java, R, HTML, CSS, JavaScript
Frameworks and Libraries: Flask, NumPy, Pandas, Keras, TensorFlow, PyTorch, Sci-Kit Learn, Matplotlib
Databases and Tools: SQL, Snowflake, GitHub, Azure, MS Fabric, Power BI, Tableau, Google Cloud Platform

PROFESSIONAL EXPERIENCE

Data Science Intern | Engage3 — Jul 2024 - Present
- Preparing data tables using the large company database on the Snowflake platform, ensuring reliable and scalable data for analysis.
- Creating an AI agent using Snowflake Cortex Analyst to generate a market data analysis report in natural human language.
- Performing data preparation techniques on the system manual PDFs with textual and image data for efficient model training.
- Spearheading the development of a Client Support AI agent by leveraging Google Cloud Vertex AI and DialogFlow CX API, utilizing advanced NLP and NLU techniques, to accurately interpret and respond to client queries regarding the company tool.
- Integrating the AI Agent into the company tool to deliver real-time, accurate answers to client questions.

Machine Learning Intern | Arkision.AI, India — Feb 2022 - Jun 2022
- Applied image preprocessing techniques, such as resizing for standardized input dimensions, noise reduction to upgrade image quality, and segmentation to isolate pertinent objects, on real-time surveillance footage.
- Developed an image captioning model achieving 87% accuracy by integrating VGG16 for feature extraction, connecting to LSTM for sequential caption generation, leading to efficient monitoring of the video footage.

Python Developer Intern | Connect Club, India — Jun 2021 - Jul 2021
- Created a robust web application making use of Flask, combined with a MySQL database managed through XAMPP, to provide a startup ecosystem to investors and startups.
- Utilized predictive analytics on historical startup data, empowering investors with insights to make informed decisions and optimize investment strategies based on growth trends.

PROJECTS

ELI - *Explain it to me Like I am*
- Fine-tuned a Gemini Pro 1.0 large language model (LLM) using Vertex AI to power ELI, specifically enhancing its ability to simplify text across varying comprehension levels.
- Orchestrated the data preparation and training phases, utilizing a comprehensive custom dataset to cover a broad spectrum of text styles and complexities.
- Integrated the fine-tuned LLM into a React-developed Chrome extension, enabling real-time, dynamic text simplification.

CancerCare AI | Participated in HackTogether 2024, organized by Microsoft | Github
- Built a Cancer Care Solution utilizing Microsoft Fabric's tools: Synapse Data Science, Data Warehouse, and Power BI.
- Deployed a chatbot using Azure OpenAI to boost patient interaction within the Cancer Care AI platform.
- Applied data augmentation methodologies, including CTGAN, to enrich dataset quality and diversity, enhancing the efficacy of deep learning model training and evaluation.
- Managed the database infrastructure adopting Fabric's Lakehouse architecture, ensuring efficient data storage and retrieval within the Cancer Care Solution ecosystem.

Sparkle | Winner at Smart India Hackathon, 2022 | GitHub
- Constructed a learning difficulty screening system with 88% accuracy in identifying learning difficulties among children, training an XGBoost model to provide insights into potential types of learning difficulties.
- Leveraged machine learning techniques, including speech recognition and head pose estimation integrated with OpenCV, to innovate the design of screening tests, strengthening accuracy in assessing user responses.
- Elevated learning outcomes and generated detailed progress reports through descriptive data analytics, contributing to a significant improvement in monitoring and lifting each child's intellectual journey.

puRED | Second Runner up at Irada 2022 Hackathon, organized by L D COE, Ahmedabad | GitHub
- Generated a Random Forest classification model with an accuracy of 95% to diagnose PCOS/PCOD based on the user's symptoms.
- Conceptualized and implemented an intent-based chatbot, using NLP and Tensorflow, to address and alleviate a broad spectrum of menstruation-related queries and concerns.
- Incorporated period cycle monitoring using FullCalendar javascript library to help users track their menstrual health and identify patterns, resulting in an increase in user adoption of these tools.