Harshkumar Patel

+1 (437) 421 3950 | Theharsh1@outlook.com | Toronto, Ontario

Portfolio Website | LinkedIn Profile | GitHub Profile

Education & Certifications

- Postgraduate Diploma – **Artificial Intelligence & Machine Learning** | Lambton College, Toronto, ON | Jan 2022 – Sept 2023
- Bachelor of Engineering – **Computer Engineering** | Gujarat Technological University, India | Aug 2016 – Aug 2020
- **Machine Learning Specialization by Stanford,** Coursera
- **Deep Learning Specialization by Stanford,** Coursera

Skills

- **Programming Languages**: SQL, Python, R, JAVA, C++
- **Libraries & Frameworks:** TensorFlow, PyTorch, Keras, Pandas, NumPy, OpenCV, Matplotlib, Seaborn, SciPy, Scikit-Learn, XGBoost, NLTK, Flask, Django, Selenium, PyAutoGUI, Scrapy, BeautifulSoup
- **Machine Learning & AI:** Supervised, Unsupervised, Semi-Supervised Learning, Predictive Modeling, NLP, Image Processing, Audio Processing, Deep Learning, Model Development, Data Visualization
- **Cloud & Big Data:** Azure, AWS, Apache Spark, Docker, MLflow
- **Databases & Visualization:** MySQL, MariaDB, Oracle, SQL Server; Tableau, Power BI, Excel
- **Methodologies:** SOLID Principles, Programming Design Patterns, Agile, Test-Driven Development
- **Operating Systems**: Linux, Unix, macOS, Windows
- **Professional Skills:** Data Storytelling, Problem Solving, Critical Thinking, Detail-Oriented

Experience

Data Scientist | Dataannotation.tech | Remote | Jan 2024 – Present

- Analyzed and Optimized large **AI** training **datasets** by **20%**, improving **model accuracy** and **performance** by **40%**.
- Executed data **cleaning**, **feature validation**, and **EDA**, enhancing **data precision** and reducing **anomalies** by 35%.
- Monitored performance of **10+ AI models**, tracking **training/testing** accuracy **metrics** to improve **model** reliability by **15%.**
- Enhanced **model** efficiency by **20%** through optimized **hyperparameter tuning** and selection of the effective **ML** algorithms.
- Delivered data insights via **visualizations**, driving **25%** improvement in **annotation** workflow **efficiency** and d**ecision-making** accuracy.

Automated Speech Research Scientist Intern | Microsoft (Nuance), Montreal | Mar 2023 – Aug 2023

- Implemented **Speech Completion** in MS Teams & PowerPoint, boosting r**eal-time recognition** by **80%** and integrating **GPT-4**.
- Leveraged **Python, ReactJS, Gradio, Azure APIs, GPT-4**, and Cognitive Speech Service to develop and deploy advanced **AI solutions**.
- Designed and deployed **CI/CD** pipeline with **GitHub Actions**, cutting **deployment failures** by **40%** and accelerating release cycles.
- **Awarded Regional Winner at Microsoft Global Hackathon 2023** for developing a **real-time speech summarization system**, improving accessibility and user experience.

Data/Business Analyst | Welltech Minerals, India | May 2020 – Sept 2021

- Leveraged **SQL**, and **Python** to extract, cleanse, and dissect datasets; pinpointed **5 vital KPIs**, elevating business alignment by **20%.**
- Boosted the **data processing time by 40%** through the extensive use of **python data manipulation libraries** and **SQL queries**.
- Boosted reporting **efficiency 25%** by **automating data extraction** and transformation **pipelines** for faster, reliable business **insights**.
- Developed **10+ interactive Tableau** dashboards visualizing **KPIs**, **trends**, and **finances**, boosting decision-making by **30%** .

Deep Learning Scientist | SoftVan, India | May 2019 – Apr 2020

- Trained and deployed 8+ **Deep Learning** models, boosting **accuracy** from **70% to 92%** for enhanced predictions.
- Automated **monitoring scripts** that reduced downtime in image and text **processing pipelines by 25%**.
- Conducted **50+ code reviews**, fixed **100+ bugs**, **optimized code** using Agile, boosting **team efficiency 15%**.
- Built **full-stack** website using **Flask, Django, Jinja2, OpenCV** backend; **HTML/CSS/JS** frontend, handling **5K+** daily users.

Projects

Smart Product Pricing Amazon Dataset

- Developed a **DL** model for the **Smart Product Pricing Challenge**, analyzing **75K** products' **text, images**, and **attributes** using CLIP model with text/image transformers to get feature embedding vectors to predict optimal prices.
- Processed **catalog content, images,** and pack quantities, **boosting** model **accuracy** from baseline to 92% for e-commerce pricing optimization.
- Generated CSV **predictions** for **75K test** set, improving marketplace efficiency and customer satisfaction through precise, data-driven insights.