Karan Singh

linkedin.com/in/karansingh
kboytron.dev

Email: kss4398@protonmail.com
Mobile: +1-217-402-6543
Iowa City, IA

Education

University of Iowa — Iowa City, IA
B.A. in Computer Science — *December 2024*

- **Coursework**: Artificial Intelligence, Computer Organization, Networks, Data Structures & Algorithms, Software Development
- **Extracurricular**: Association for Computing Machinery (ACM)

Certifications: AWS Certified Data Engineer (DEA-C01), AWS Certified Solutions Architect (SAA-C03), AWS Certified Developer (DVA-C02), AWS Certified Cloud Practitioner (CLF-C02), Google Data Analytics

Experience

Labelbox — Remote
Software Engineer - AI Training (Contract) — *June 2024 – Dec 2024*

- Evaluated data labeling accuracy of AI models during deployment by analyzing annotation discrepancies and model predictions, improving dataset quality by **18%**
- Contributed to the **Labelbox Python SDK** by extending API functionality and writing documentation, enhancing developer efficiency and integration speed
- Developed automation scripts using Python to validate and clean training data, reducing manual effort by **30%**
- Debugged and optimized **AI-assisted labeling workflows** by identifying edge cases in model-generated annotations, increasing labeling consistency and confidence scores

8AM App — Los Angeles, CA
Data Engineer Intern — *September 2023 – November 2023*

- Designed and implemented efficient ETL pipelines using **AWS Glue, Python, and SQL**, automating the ingestion and processing **6 TB** of customer records from S3, reducing processing time by **40%**
- Optimized query performance on large CSV-based financial datasets using **AWS Redshift Spectrum with S3**, leveraging SQL query tuning techniques to improve revenue forecasting accuracy by **30%**
- Analyzed user telemetry data using **Python**, tracking retention and churn trends; visualized insights in **Power BI** and presented findings to the CEO, guiding the transition from web to mobile
- Collaborated with an agile team to perform data cleansing using **Python** and **Pandas** on 4+ years of user data, reducing data inconsistencies by **35%**

Projects

SolarEnergyViz.tech *Python, Pandas, Javascript, Seaborn* — September 2023

- Led a team of **4** to develop a machine learning-powered solar analytics platform, visualizing **12+ years** of solar data
- Integrated solar energy APIs to retrieve irradiance data and compare solar array performance across locations
- Built time-series forecasting models with **Prophet** to predict missing solar irradiance data, restoring **23%** of data
- Automated data processing and deployment with **GitHub Actions** and **Netlify**, ensuring seamless CI/CD

SpeechServe *Python, TensorFlow, AWS (Sagemaker, Lambda), Docker* — Aug 2024 – Present

- Developed an end-to-end speech recognition system from scratch using **CTC** and a deep neural network for variable-length input/output mapping
- Built core components like acoustic modeling, forced alignment, and beam search decoding, enabling accurate speech transcription with a low error rate in performance evaluations
- Designing a modular architecture for real-time transcription with low latency, enabling scalable multi-instance execution on AWS Lambda to handle concurrent user requests
- Building a **FastAPI** backend to deploy the model, enabling real-time speech-to-text via API, targeting sub-second response times for **5+** concurrent streams

Awards & Achievements

- Best Data Visualization Award at Hack UIowa (SolarEnergyViz.tech)
- Built a SaaS platform generating **$250k+** revenue using Python, image recognition, and Wireshark

Skills

Languages: Python, Java, C++, JavaScript, React.js, Next.js, SQL, HTML/CSS

Technologies: TensorFlow, PyTorch, Scikit-Learn, Pandas, Prophet, AWS (S3, EC2, Lambda, DynamoDB, IAM, Redshift, CloudFront, CloudWatch), Azure, Docker, Kubernetes, Terraform, CI/CD (GitHub Actions, Jenkins), MongoDB, MySQL, REST APIs, Git, Unix, Visual Studio Code