ABHISHEK SAINI

206-887-5357 | work.abhisheksaini@gmail.com
linkedin.com/in/saini-abhishek | github.com/abhishekiitm | abhishekiitm.github.io

Experience

University of Washington May '23 – Current
Research Assistant, EMIT Lab, Dept. of ECE *Seattle, WA*

- Developed a novel framework to train cascade classifiers using Genetic Algorithms and LightGBM for feature selection that reduced the training time by 50x.
- Extended existing open-source electronic design automation tool to generate SKILL code from an input netlist.

Bosch Research Jun '22 – Sep '22
Machine Learning Intern *Sunnyvale, CA*

- **Reduced unexpected downtime of factory lines by 35%** through research and development of a real-time fault identification predictive maintenance framework using ML. **Patent application** for this framework to be filed by Bosch.
- As part of this framework, pipeline for feature engg., evaluation and cross-validation, training (XGBoost), hyperparameter optimization (hyperopt library), explainable inference (shap library), drift detection (alibi-detect library) were developed, leading to an F-score of 74%.

OpsMx Dec '20 – Sep '21
Software Engineer - Data Science platform *Bangalore*

- Member of data science team that is responsible for the R&D of tools to automate software deployment.
- Reduced software error **reidentification time from minutes to seconds with >95% precision** by designing an algorithm to identify repeated error patterns; feature in use by 3 Fortune 500 clients.
- Improved F-measure of software log clustering to 90% by modification of Spell and Drain clustering algorithms.
- Enhanced application to handle large log files by creating an S3-compatible object storage interface with Boto3 and MinIO.
- Improved microservice scalability by redesigning compute-intensive log analysis APIs to function as a distributed task queue using Celery and RabbitMQ.

Palpx Mar '20 – Dec '20
Machine Learning Engineer *Bangalore*

- Automated remote proctoring of online exams by developing a browser-side Javascript solution that performed facial recognition and activity detection using Tensorflow.js, Faceapi.js.
- Eliminated the cost of data labeling by generating simulated training data using Unity's ML-Agents SDK. Achieved 95% accuracy across 3 classes to classify industrial fasteners using Keras by finetuning a ResNet-50 model trained on synthetic data.
- Engineered a Python and Selenium-based web scraping tool to gather comprehensive movie metadata and Google search trends data to forecast movie revenue.

Xiaomi Aug '16 – Mar '20
Business Analyst *Bangalore*

- **Reduced e-commerce product return rates by 50%** by estimating the probability of product return using a Logistic regression model on hand-engineered features.
- Forecasted monthly smartphone sales with >95% accuracy, set targets for Online and Retail teams
- Increased retail business market share by 8 percentage points via data-driven process improvements, for which I earned the best employee award in Q1 2019.
- Applied connected components analysis to identify and restrict sales originating from shopkeepers, thereby optimizing the availability of fast-moving smartphones leading to a notable improvement in product accessibility and CSAT scores.

Education

University of Washington - Seattle Sep '21 - Mar '23
Master of Science in Data Science, GPA: 3.94/4.0

- **Course Projects**: Reduced deep learning inference time by a factor of 10 through benchmarking PySpark performance on Azure Databricks.
  Reduced neural network training time by more than 3x through benchmarking of distributed training using PyTorch's DistributedDataParallel module; model trained across multiple GCP spot instances.
- **Industrial Project**: Generated synthetic handwritten data using a diffusion model to train an OCR for the handwritten medical prescription image digitization reducing WER by 43%. Project sponsored by Flipkart.

Indian Institute of Technology - Madras Aug '11 - Jul '16
B.Tech. + M.Tech. in Electrical Engineering, Minor in Physics

Skills and Certifications

Languages & Skills: Python, C++, SQL, PyTorch, scikit-learn, PySpark, Docker, Kubernetes, PowerBI, AWS, CI/CD, TDD
MOOCs: Deep Learning Specialization, Natural Language Processing on Coursera, Practical Deep Learning on fast.ai