

Arpan Ghoshal

contact@arpanghoshal.com | +1 (332) 699-1429 | linkedin.com/in/arpanghoshal | arpanghoshal.com

EXPERIENCE

Machine Learning Engineer

Sep. 2023 – Present

Giisual Inc.

Philadelphia, PA

- Developed the world's first commercial ML model for power outage detection with a 96% accuracy rate, overseeing the full lifecycle from data extraction to deployment in production using traditional ML/DL techniques
- Proposed and implemented "Down Detector for Power Outages" which extracts content from social media and predicts power outage probabilities with LLM and GenAI tools
- Led cross-functional teams to execute strategic goals for US and international AI model deployment, achieving successful market expansion

Machine Learning Engineer

Sep. 2021 – Jun. 2022

Hitch Works (Acquired by ServiceNow)

Remote, Seattle, WA

- Established an ML-data pipeline using AWS services, enabling continuous automated 24/7 data crawling, ingestion, structuring, preprocessing, executing ML algorithms, and generating a real-time database for job market analysis
- Developed LLM-based algorithm that boosted skills contextual mapping from job descriptions & resumes by 170%

Machine Learning Engineer

Jan. 2021 – Aug. 2021

Brainentech Neuroscience

Bangalore, India

- Trained and tuned LLM-based NLP-audio models to generate 20+ insights for various text-audio use cases
- Created a real-time AI-powered audio API for deployment to over 150k users, enhancing insights extraction, and detecting emotions, engagement, and other attributes from audio data
- Research: Canonical Correlation Analysis to find the engaging score between brain EEG signals and audio stimuli

Data Engineer Intern

May 2020 – Jul. 2020

NayaOne

Remote, London, UK

- Created the company's first data infrastructure and Data APIs used by top FinTech companies in UK

EDUCATION

New York University

New York, NY

Master of Science in Computing, Entrepreneurship, and Innovation; Courant & Stern

Jul. 2022 – Jul. 2023

PES University

Bangalore, India

Bachelor of Technology in Computer Science and Engineering; Data Science Specialization

Aug. 2017 – May. 2021

PROJECTS

Entrepreneurship: Created a *GenAI SaaS Survey Product* currently used by more than 10 companies

Semi-Supervised Learning in Video Data: Predicted the 27th frame from the first 13 frames of a video

VibeMatch Algorithm: Matched venues' songs with users' moods using advanced NLP techniques

Clinical Context: Improved the embeddings of clinical notes and proposed LLM-embeddings evaluation strategies

GMM: Implemented an Incremental Learning System for Clustering Confidential Finance Data using GMM-EM algo

Research Paper: An Enhanced Context-based Emotion Detection Model using RoBERTa; IEEE-CONECCT, 2022.

The paper has been *cited 40+ times* and the model has been *downloaded for 10+ million times* in huggingface as *arpanghoshal/emorRoBERTa*

Research Paper: CATAQ: Concise Answer to Any Question; IEEE-ICSES, 2021 (NLP, RAG, LLMs and GenAI)

Hackathon: Developed an Instagram Hashtag Generator that analyzes images, captions, and real-time trends to generate hashtags, increasing post virality and visibility by 75%

LangChain Contributions: Contributed to LangChain by solving code bugs, enhancing its usability for the users

Entrepreneurship: Co-founded a startup at NYU that integrates GenAI into business solutions and developed a mental health platform that allows users to express their emotions, regulated by AI

TECHNICAL SKILLS

Skills: ML, MLOps, NLP/LLM, LLMops, Data Science, Deep Learning, Cloud Computing, Data Engineering, Statistics

Tools: Python, PyTorch/Cuda, Langchain, Sklearn/Numpy/Pandas/XGBoost, Spark, SQL/NoSQL, Flask/Docker, git

Certifications: AWS Solutions Architect Associate