Neelambuj Chaturvedi

Neelambuj.nitj@gmail.com | LinkedIn | GitHub | Google Scholar | Kaggle | +91-9855672717

EXPERIENCE

Walmart Global Tech
July 2022 – Present
Senior Data Scientist, Marketplace science team
Bengaluru, India

- **Cross Border HTS Classification:** Working on a Vision + NLP based model to classify HTS codes for various products being imported exported, currently deployed on Walmart Marketplace

GoKwik
Oct 2022 – July 2025
Applied Scientist 2, ML Science Department
Bengaluru, India

- **Return to Origin (RTO) Model:** Built and deployed three end-to-end versions of GoKwik's flagship **XGBoost-based RTO prediction model** on the GoKwik Checkout, serving over **2 million real-time requests/day** using both batch and streaming features via an in-house Feature Store. Reduced RTO by **40%**, saving **50M annually**.
- **Model Retraining and Monitoring Pipeline:** Designed an end-to-end pipeline to monitor feature distribution drift and trigger retraining. The updated model is deployed in **shadow mode** for validation before production rollout.
- **pLTV Prediction:** Developed an end-to-end system to predict the pLTV value of a customer event in real time using **gradient boosting models** on session-level and batch features. Serving **5M+ daily requests**. A/B testing resulted in a **10–15% increase in ROAS** via targeted retargeting.
- **Customer Profiling (KwikAudiences):** Built an intelligent system to segment GoKwik customers based on cross-brand buying behavior. Led POCs with D2C merchants to run marketing campaigns using these segments, leading to a **25% ROAS improvement**.
- **Payment Method Prediction:** Developed an intelligent, **rule-based approach** to predict the preferred payment method for a user on the GoKwik checkout page. A/B testing showed a **2% increase in conversion rate (CR)** for D2C merchants.
- **Stakeholder Management:** Collaborated with product and tech teams to ensure timely delivery of data science models. Defined phase-wise timelines to improve transparency during development.
- **Mentoring:** Leading a team of 2 Applied Scientists (AS1s), delivering impact through enhancements to the RTO and pLTV models.

Accern AI
Jan 2022 – Sep 2022
Senior Data Scientist
Bengaluru, India

- **Dynamic Taxonomy Labeller:** Developed an **active learning engine** (BADGE and uncertainty estimation) combined with **BERT** to automate taxonomy labeling of internal documents. Enabled faster training of in-house sentiment models and reduced labeling costs by **40%**.

ZS Associates
Jul 2019 – Dec 2021
Data Science Associate Consultant, Advanced Data Science Group
Bengaluru, India

- **Rare Disease Patient Identification:** Built a **gradient boosting model** using **Positive-Unlabeled (PU) learning** to identify patients with rare diseases via healthcare claims. Devised go-to-market (GTM) strategy for reps using prescribing patterns, increasing patient capture by **46%** in the US Southeast region.
- **NASH Patient Identification and Progression:** Developed ML models to detect NASH patients using EHR and NIDDK datasets. Published findings in Nature Scientific Reports (NASHMap). Built a progression model to predict cirrhosis using Optum EHR data, published in Gastroenterology and Hepatology.
- **RA2 Dream Challenge:** Represented ZS in a computer vision challenge predicting erosion and narrowing scores for rheumatoid arthritis patients from X-rays. Built a multi-stage pipeline using **RetinaNet** for joint detection and **EfficientNet + attention** for scoring. Secured **4th and 8th place** among 1,000+ participants (Leaderboard), presented at RSG Dream Conference.
- **Unaddressed Opportunities:** Built a multi-stage **adaptive PU learning** system to enrich claims data. Created targeting strategies for breast cancer therapies based on model scores, increasing coverage by **10%**.
- **Achievements & Leadership:** Got a fast track promotion within 1.5 years to DSAC, Mentored 3 DSAs to formulate, design, and deploy ML solutions while deepening domain expertise.

PUBLISHED RESEARCH: *Google Scholar*

- **Ranking and Clustering of Drosophila Olfactory Receptors using Mathematical Morphology**
  Published in *Genomics Journal*, Elsevier (Impact Factor: 3.6)

- **DeepRA: Predicting Rheumatoid Arthritis Progression from Joint-Level Radiographs**
  Preprint published on *arXiv*

- **A crowdsourcing approach to develop machine learning models to quantify radiographic joint damage in rheumatoid arthritis**
  Published in *JAMA Network Open*, 2022

Kaggle:

- **MICCAI Brain Tumor Radiogenomic Classification** Won a bronze medal in the competition (Top 8% rank globally), built a Image Segmentation model coupled with a CNN - LSTM model to predict the regions for tumor

Education

Birla Institute of Technology and Science Pilani (WILP)   Pilani, India
MTech, Artificial Intelligence and Machine Learning   *April 2023 - April 2025*

National Institute of Technology Jalandhar   Jalandhar, Punjab
B.Tech - Industrial and Production Engineering - CGPA 7.86/10   *July 2015 – May 2019*

Technical Skills

- **Programming Languages & Tools:** Python, PySpark, SQL, Git, LaTeX, MS Excel

- **Machine Learning:** Gradient Boosting Models (XGB), Tree based Models, PU Learning

- **Deep Learning & NLP:** PyTorch, Transformers, BERT, CNN's, Object Detection, Active Learning

- **MLOps & Deployment:** Docker, Kubeflow, Feature Store, Shadow Deployment, GitHub Actions

- **Cloud & Workflow:** Google Cloud Platform (GCP), Vertex AI, Airflow