Naman Lal

+91 8601251608
namanlal.lal92@gmail.com
Bengaluru, Karnataka, IN

SUMMARY

Data Science professional with 3+ years of experience in data analysis, visualization and modeling for key business problems and 2 years of experience in software development. Proficient in collaborating with teams to design, develop and enhance ML solutions to revamp business processes and products. Hand-on experience in RL, Generative-AI, Computer Vision and Predictive modeling across diverse industry problems in a fast-paced startup environment.

PUBLICATIONS

TC-OCR: TableCraft OCR for Efficient Detection & Recognition of Table Structure & Content, **ACM MMIR Workshop 2023**

RanLayNet: A Dataset for Document Layout Detection used for Domain Adaptation and Generalization, **ACM MM Asia 2023**

GEC-DCL: Grammatical Error Correction Model with Dynamic Context Learning for Paragraphs & Scholarly Papers, **BDA 2023**

Context-Enhanced Language Models for Generating Multi-Paper Citations, **BDA 2023**

KG-CTG: Citation Generation through Knowledge Graph-guided Large Language Models, **BDA 2023**

PhyCoTQA: Answering Physics Questions Using Chain-of-Thought Prompting With Multiple Images, **IJCAI 2024** (Submitted)

ACHIEVEMENTS

Kaggle Competition Expert

Winner of Garden Nerd Hackerearth competition.

Runner-up in Affine Analytics ML challenge.

ACM ICPC National Rank:
Amritapuri-72/255 and Kharagpur- 26/67

SKILLS

Generative-AI NLP EDA Data Visualization Time Series Machine Learning MLOps Predictive modeling System Design User Behavior Recommendation Dataset Creation

WORK

Github - https://github.com/Altaireon

Kaggle - https://kaggle.com/Altaireon

Hackerearth - https://hackerearth.com/@Altaireon

Enhancing Student Engagement Prediction: Leveraging Image Data and Attention Mechanisms for Efficient Engagement Prediction, **IJCAI 2024** (Submitted)

Advancing Citation Text Generation: Leveraging Multi-Source Seq2Seq Models and Large Language Models for Improved Citation Text Generation, **IJCAI 2024** (Submitted)

EXPERIENCE

Games24x7 Pvt. Ltd., Bengaluru - Applied Scientist - 1
Jun 2020-Present

Asset Generation using Stable Diffusion - Fine tuning and deploying Stable Diffusion on in-house assets developed across 4 games launched by the U-Games Studio. We captured the style of the assets and used the model to generate food items for Cafe Rescue. We also created a plugin for integrating the model capabilities like sample generation, in-painting and out-painting based on the prompt given by the Artist, which successfully increased the efficiency of the art creation process by ~40% which was observed on 10 assets in production.

Leaderboard Point System and Reward Restructuring - Simulating games based on priors calculated from previous games to optimize the point system to increase user engagement on the platform. We also restructured the reward system so that users are willing to compete even when the chances of moving up in the rankings are low. We also tweaked the point system to minimize rewards distributed to users who got in the top rankings purely by chance. We achieved +2% in overall d30 revenue and +5% d30 revenue in HVP players.

Procedural Content Generation, Bot Testing and Level Recommendation- Applying Conditional GANs and MRF-based approach to generate game content and testing the playability of the content via MCTS and SAC-based bot. We customized GAN to control multiple metrics of the game to generate varying difficulty levels. Reduced game level creation time by more than 50%. Analyzing user gameplay to recommend a hard level for every skill-based user cohort with 72% recall.

User Personalization - Utilizing the MAB framework to optimize revenue and D1 conversion for acquired users. Evaluating campaigns containing different users' specific product features and recommending personalized features for certain sets of cohorts. Average 2% lift in D1 conversion on specific cohorts of users. Also deploying a Pick-Discard skill evaluation ML model on K8 cluster with Grafana for monitoring, Sumo logic/Kibana for logging and creating a

TECHNICAL EXPERTISE

Tools - Pandas, scikit-learn, Pytorch, Spark, Python, AWS, Kubernetes, Opencv-python, SQL

ML Modelling - Ensembling, Deep Learning, Decision Trees, Clustering, Linear and Logistic Regression/Classification

Jenkins pipeline to auto-deploy updated models.

3D Reconstruction and LOD Generation (Research Project) - Generating low poly 3D mesh from a single sketch. We used sketches to extract object geometry information using Resnet-18 and used a transformer-based "Polygen" to generate vertices and faces using separate models. We curated an in-house dataset of sketches and 3D models and merged it with public datasets. We covered 20 different classes with 2 assets from the "Vass" class going into production. We also generated LODs of 3D models by mesh decimation based on deep learning models which optimized our game performance.

Research Collaboration with MiDAS labs, IIIT Delhi

Jun 2021 - Present

Text Generation - Discuss ideas related to problems researchers face to generate citation text. We used a combination of heterogeneous Graph embeddings, Stylistic and Keyphrase embedding with BART transformer model to generate citation text and outperformed current SOTA models. We have compared our models with LLMs' like Vicuna, LLama and Alpaca.

Grammar Error Correction(GEC) for documents - We tackled the problem of GEC for documents by preserving dynamic relevant context as compared to the current SOTA context-aware model that uses a fixed no of sentences as context. We developed a DCA algorithm to get the context and use GeCTOR to perform corrections.

Oracle India Pvt. Ltd., Bengaluru — *SDE I*

Jul 2018 - Aug 2019

Inter-Working Function(IWF) - Design and Development of a Networking Function in a 5G Network. Build from scratch, implemented in Spring Boot J2EE. Deployed as a microservice using Kubernetes and helm.

Quittung Labs Pvt. Ltd., Noida — *Software Developer*

May 2017 - Nov 2017

Developed an iOS app for "BillFree-Your Own Bill Wallet". Involved in the design, development and implementation of bill scraping tool using various text analysis techniques (Stopwords treatment, stemming/lemmatization, Word Embedding) on unstructured emails and classifying them using Machine Learning Techniques (Naive Bayes, FastText, LDA)

CERTIFICATION

Deep Learning Nanodegree - *Udacity*
May 2021

Machine Learning for Data Analysis - *Coursera*
Aug 2019

Machine Learning - *Coursera*
Aug 2019

PROJECTS

Demand Prediction based on Feedback Analysis, Noida
Nov 2017 - Nov 2017

Using customer ratings and comments to enhance product and purchase experience. Built a base model to determine customer satisfaction, using LDA, Deep Learning and Naive Bayes thus providing a score-based feedback mechanism.

Detecting Defects in Steel sheets, Kaggle
Dec 2019

Developed a model for localizing and classifying surface defects on a steel sheet. Used ensemble of UNet architecture with se-resnext50, resnet34 and mobilenet2 as encoders to localize the defect and resnet34, se-resnext50 and se-resnet50 to classify the defect.

Understanding Clouds, Kaggle
Dec 2019

Developed a model for identifying (segmentation and classification) clouds in satellite images. Used UNet architecture with efficientnet-b2 as encoders to segment (create mask) different clouds and threshold on mask size to classify them.

Convolutional Neural Network for Steady Flow Approximation
Jan 2018 - Apr 2018

Implemented the research paper with Tensorflow's CNN.

EDUCATION

Indian Institute of Information Technology, Jabalpur — *B-Tech Computer Science and Engineering*
2014-2018

CGPA - 7.2/10