Rudresh **Veerkhare**

Senior Analyst, Deutsche Bank

rudresh.net | veerkharerudresh@gmail.com | RudreshVeerkhare | rudresh-veerkhare | Google Scholar

Research Interests

I am interested in creating robust and dependable **Document Understanding** systems that facilitate data extraction and analysis from multi-modal documents, further supporting data-driven machine learning.

- **Multimodal Document Processing**: Developing techniques for structured data extraction that encompass textual, visual, and spatial information.
- **Document Analysis utilizing Generative AI**: Employing Multimodal Document Processing to empower Generative AI, like Large Language Models, in building robust and reliable Data Analysis Pipelines.

Education

May 2022 | **Sardar Patel Institute of Technolgy** — **Mumbai, India**
Aug 2018 | B.Tech, Computer Engineering | **CGPA**: 9.63/10

Publications

[1] **HRescue: A Modern ML approach for Employee Attrition Prediction** [🔗]
Rudresh Veerkhare*, Parshwa Shah*, Jiten Sidhpura*, Sudhir Dhage (* = Equal Contribution)
Springer Proceedings in Mathematics & Statistics, vol 401. Springer, Cham. [**ICMLBDA 2022**]

[2] **FedSpam: Privacy Preserving SMS Spam Prediction** [🔗]
Jiten Sidhpura*, Parshwa Shah*, Rudresh Veerkhare*, Anand Godbole (* = Equal Contribution)
Communications in Computer and Information Science, vol 1793. Springer, Singapore. [**ICONIP 2022**]

[3] **Face To BMI: A Deep Learning Based Approach for Computing BMI from Face** [🔗]
Jiten Sidhpura*, Rudresh Veerkhare*, Parshwa Shah*, Surekha Dholay (* = Equal Contribution)
In Proceedings of ICITIIT, Kottayam, India, 2022, pp. 1-6. IEEE [**ICITIIT 2022**]

Experience

Present / **Jul 2022** | **Deutsche Bank | Chief Technology Office, TDI** — **Pune, India**
Senior Analyst
I'm working on OCR and **Document Understanding** to streamline document processing. My work spanned **Custom Document Entity Extraction**, **Smart Document Splitting**, **LLM-aided Contract Drafting**, and innovative software automations like **Automated Vulnerability Scanning**. I have also worked on the adaptation of **Document AI** within the organization and have experience in **Software Governance**.

Jul 2021 / **Jun 2021** | **Deutsche Bank | Chief Technology Office, TDI** — **Pune, India**
Technology Analyst Intern
As part of the research project, I worked on Optical Character Recognition, which encompassed tasks such as **Table Detection**, **Table Structure Recognition** from scanned documents, **Signature Extraction**, **Custom Named Entity Recognition**, and **Intelligent Character Recognition**.

Apr 2021 / **Jan 2021** | **Sardar Patel Institute of Technology** — **Mumbai, India**
Research Intern | Advisor: **Prof. Pramod Bide**
I conducted research in the field of **Cross Event Detection** and **Topic Evolution Mining** from Social Media Posts, developing innovative algorithms for Cross Event Detection and **Sub-Topic Evolution** through the application of statistical NLP techniques.

Awards

- **India Excellence Award (Deutsche Bank, Feb 2023)** Award granted to only **280 out of 17000** (1.65%).
- **Best Graduation Project Exhibition Award (SPIT Mumbai, Apr 2022)** Awarded **1st prize** for project **FedSpam**.

Research Projects

FedSpam: Privacy Preserving SMS Spam Detection

Federated Learning, NLP, Edge Computing

› FedSpam is an edge computing application which leverages **Federated ML** to **preserve the user data privacy** while using the advanced data-driven ML solutions for spam detection.

CustomXGBoost: XGBoost Implementation with Optimizers [in |]

Gradient Boosting, Optimizers

› XGBoost **Implementation from scratch** where I've modified the **gradient boosting** to utilize optimizers such as **ADAM** and **RMSProp**.

Arbit: A Decentralized Crypto Exchange Arbitrage System

Linear Algebra & Graphs, Blockchain Smart Contracts

› Developed an efficient algorithm with **O(1)** time complexity for detecting **Nth order arbitrage** opportunity in Decentralized Crypto Exchanges (DeX).

› Utilized graphs and linear algebra in the derivation process.

› Implemented the system on the Cloud with optimal regions for low network latency, enabling continuous real-time blockchain monitoring and swift execution of profitable arbitrage transactions.

› Created a profitable personal project, generating approximately **$10,000** in cryptocurrency in early 2022.

HRescue: A Modern ML approach for Employee Attrition Prediction

Gradient Boosting, Explainable AI, Data Augmentations

› Numerous attrition prediction methods have been developed in the past. However, This approach focuses on **interpretability** of ML models for sensitive employee attrition decisions and **outperforms prior methods** while addressing **data imbalance**.

Face To BMI: A Deep Learning Based Approach for Computing BMI from Face

Computer Vision, Transfer Learning, Discriminative Learning

› Developed a BMI prediction model from facial images using **transfer learning** on deep convolution networks, implementing **discriminative learning** to train the last few layers with varying learning rates for further model fine-tuning.

› Used **Tensor Processing Unit** for training Deep Learning models.

Software Projects

Catalyst [| | in]

Open Source, Node js, VSCode Extension

› Catalyst is a VS code Extension to accelerate the process of solving problems on Codeforces. It has **3000+ installed user and 12000+ downloads**

ReactPy [| in]

Open Source, Web Python, Algorithm

› ReactPy is a implementation of React in Python using Brython. It's a **from scratch implementation** of **React Fiber**, along with **diffing and Virtual DOM** in Python 3.

Numras [ ]

Open Source, Numpy, Neural Networks

› A **mini-framework** completely **implemented from scratch** using Numpy. Its api is similar to **Keras**. All of the operations like forward pass, backward pass and optimizations are carried mathematically from scratch.

Recruitment Assisting Platform [ ]

Data Mining, Data Visualization

› Used **Latent Dirichlet Allocation (LDA)** for grouping candidates based on the Resume.

Elliptical Curve Diffie Hellman [ ]

Cryptography, Algorithms

› Implemented **Elliptical Curve Diffie Hellman Key Exchange algorithm**, from scratch.

Technical Skill

Programming	Python 3, Java, C, C++, Java, JavaScript, Typescript, Solidity, PowerShell, Shell Script
Frameworks & Libraries	PyTorch, Numpy, Keras, Tensorflow, Scikit-Learn, Huggingface, LangChain

Volunteering

CSR, Deutsche Bank **2022 - Present**

- Volunteered for School Kit Assembly and Distribution for underprivileged children.
- Volunteered for crafting of environment friendly paper bags.

Scikit Learn Open Source Contribution [|] **Nov 2022**

- Implemented an Enhancements Proposal for allowing Minkowski distance with $0 < p < 1$.

SEVA, Mumbai **2018 - 2020**

- Volunteered for teaching Maths and Science to high school underprivileged children.
- Volunteered for Mumbai's Beach Cleaning.