

Yogendra Verma

+91 639-288-8879 | yogendraverma.ec@gmail.com | [Portfolio](#) | [Github](#) | [LinkedIn](#)

CAREER OBJECTIVE

Data Scientist with 5+ years of experience. Always seeking to work with reputed organization where I could get an opportunity to grow my skills and improve and maintain the team's productivity using my Skills & Knowledge.

EDUCATION

BPUT University

Master of Technology in Computer Science and Engineering CGPA: 8.12

Orissa

Aug 2016 – June 2018

ICFAI University

Bachelor of Technology in Electronics and Communications Engineering CGPA: 6.20

Dehradun

Aug 2008 – June 2012

EXPERIENCE

Sr. Data Scientist

Gramener

March 2023 – Present

Delhi

- On-site Data Scientist at National Health Authority, tasked with automating medical claims processing for Pradhan Mantri Jan Arogya Yojana Scheme.
- Developed advanced AI algorithms for automated claim processing using NLP, CV, and MLOps.
- Integrated MLOps methodologies to ensure scalability, achieving the processing capability of 100,000 documents per hour
- Streamlined medical claims processing, significantly reducing time, increasing operational efficiency, and ensuring a 94% accuracy rate in document classification for the National Health Authority.
- Tools: AIML, Deep Learning, Computer Vision, Tesseract, Python, API Design, Flask, Docker, K8 Cluster, AWS Cloud, Computer Vision, OpenCV, Documents Classification

Sr. Data Scientist

Thinkbridge

June 2022 – Feb 2023

Pune

- Design and develop an e-billing assistant machine learning model which can automatically identify or flag the bills that have the potential to be rejected.
- Design and develop an automation bill remediation system for the rejected lines Items In the e-bills to quantify the current rejection rate and help to reduce the rejection ratio.
- Building an AIML based model to identifying the fraud while processing the bills by the legal firms to its customers. Also it is removing maximum human effort.
- Tools: Python, OpenCV, Keras, Tensorflow, PyTorch, BERT, Word Embeddings, Text Classification, AWS textract

NLP Platform Engineer

Quattro

Dec 2020 – Dec 2021

Gurugram

- Designing and constructing a document classification system to automate the loan underwriting process.
- Clustering documents from comprehensive loan packages, comprising over 1500 pages and spanning 30+ categories concurrently.
- Extracting data from scanned PDF bank documents, including tables and images, to expedite the loan application underwriting process.
- Tools: Python, Layout Detection, Documents Classification, BERT, Random forest, CNN, RNN, RCNN, LSTM, Pandas, NumPy, Image Processing, OpenCV, Computer Vision, Camelot, AWS textract

Data Scientist

Edu-Fabrica

Aug 2018 – Nov 2020

Delhi

- Design and building an Automation system based on AIML to Support Process and Hotel customer Reviews Sentiment Analysis
- Tools: Machine Learning, Python, Pandas, Sklearn, Natural language processing, NLTK, RegEx, Stemming, Lemmatization, CountVectorizer, TF-IDF, WordCloud, Selenium web scraping, Classification Models

- Design and develop IOT Based Smart home Automation To control appliances at fingertips from anywhere in the world
- Tools: IOT, Raspberry Pi, NodMCU AWS, MQTT, Node Red, Pub Sub, http, html,php, webhook

PROJECTS

Retina Disease Classification | *Supervised Learning, Python, Computer Vision, CNN, Exploratory Data Analysis (EDA), Acc*

- Used Retina based diseases dataset containing 10000 images to classify the diseases of the patient.
- Implemented a Deep Learning based classification model for identifying potentially infected or damaged retina.
- Utilized Computer Vision and other tools for achieving this.

Classification Solution for Scanned Documents | *Python, Documents Classification, BERT, Random forest, CNN, RNN, R*

- Designing and constructing a document classification system to automate the loan underwriting process.
- Clustering documents from comprehensive loan packages, comprising over 1500 pages and spanning 30+ categories concurrently.
- Tested the model on the real scanned documents and got 90% overall accuracy on the evaluation metric.

TECHNICAL SKILLS

Programming Languages:Python, Embedded C, SQL, MySQL

Frameworks: Tensorflow, PyTorch, Streamlit, Gramex, Keras, Matlab, Flask

Developer Tools: Git, VS Code, Anaconda, Azure CLI, AWS CLI, Heroku, Docker, Kubernetes

Libraries: NumPy, Matplotlib, Pandas, scikit-learn, Seaborn, OpenCV, Huggingface, Plotly

EXTRA CURRICULAR

Embedded System Trainer Conducted multiple trainings for pan India institutes like CDAC and NTPC.

Technical Head of Collegiate Students Welfare Committee