



# Abhishek Kumar

Experienced and dedicated Data Scientist with several years of experience identifying efficiencies and problem areas within AI and NLP, while communicating needs for projects.

## GET IN CONTACT

Mobile: +91-6204196506

Email: abhishek.kr.patel99@gmail.com

## PERSONAL DETAILS

- Total Experience 3 Years 2 Months
- Date of Birth Sep 25, 1999
- Gender Male

## SKILLS

- Python
- MongoDB
- Elastic Search
- Pytorch
- Transformers
- Tensorflow
- Computer Vision
- NLP
- Data Structures
- Algorithms
- Machine Learning
- Deep Learning
- BERT
- GCP
- RegEx
- Text Mining
- Pandas
- LLM
- Fast API
- Flask Web Framework
- Django
- React.Js
- Docker
- GIT
- OCR
- Beautiful Soup
- Selenium

## TECHNICAL SKILLS

- Data Preparation
- Data Cleaning
- Transfer Learning

## EDUCATION HISTORY

### Graduation

Course	B.Tech/B.E.( Metallurgy )
College	Indian Institute of Technology (BHU) Varanasi (IIT BHU)
Year of Passing	2020
Grade	7.2/10

## WORK EXPERIENCE

Aug 2022 to Present

### Senior Data Scientist at BAJAJ FINSERVE

- Work on multiple projects on AI and NLP.
- Focused on pdf extraction and cleaning of various bank statements having different formats on each bank.
- Work on OCR project where build deep learning model for bank logo detection using YOLO V7 model, accuracy of the model is around 98 percent.
- Work on bank demographics extraction from pdf and images inside bank statement.
- Build dashboard for Perfios banking index using Django and React js.

Dec 2021 to Jul 2022

### Software Engineer at Carpl.ai

- Work on building the dashboard for x ray and diacom images.
- Fixed the bugs and developed the API for the CARPL dashboard using Django,nginx,docker and react js.
- Maintain and reduced the api call time using Fast api in some cases.

Aug 2021 to Dec 2022

### Associate Data Engineer at Innoplexus Consulting Services

- Focused on data crawling and data extraction from different types of htmls and pdfs.
- Port and clean the crawled data into the CLS of different asset classes  
like:-Publications(PubMed), Congresses and News.
- Using Elasticsearch query fetch the data from different asset classes  
for product development and analytics.
- Extract the unstructured pdfs text data from different sources using a  
pdf to text conversion and after this string

- Problem Solving
- Model Tuning
- Model Training
- Generative AI
- Competetive Programming
- Data Augmentation
- Data Crawling
- Data Extraction
- PDF Extraction

---

## LANGUAGES KNOWN

- English
- Hindi

---

## COURSES & CERTIFICATIONS

- A-Z Machine Learning Course
- Deep Learning Specialization By Andrew Ng

---

## SOCIAL LINKS

- <https://leetcode.com/abhi1999/>
- <https://www.linkedin.com/in/abhishek-kumar-040a9a153>

manipulation for all sources.

- Worked on hospital wise crawling of physicians data from different sources in the Curia App development.
- Tools:-Beautifulsoup,Selenium,Elasticsearch, Mongoddb and BERT.

May 2019 to Jul 2019

### Machine Learning Engineer(Internship) at Quantiphi Analytics Solutions

- Focused on research project to reduce the training time of machine learning classification model(ResNet50) on image datasets using clusters of multiple GPUs.
- Using Distributed Deep Learning with Horovod technique after benchmarking of ML mode ltraining time reduced by two third.  
Benchmarking and accuracy measured by F1 score and confusion matrix.
- Tools &Technologies:-Google Cloud Platform(GCP), Keras, Tensorflow,Python, Ring all reduce algorithm, Nvidia Tesla k-80,GPU's,Horovod,Supervised Learning,Distributed Deep Learning,ResNet-50.
- Worked on OCR and BERT sentence transformer model for detect the text from the TV data images and convert it into english language.

Jan 2019 to Mar 2019

### Undergraduate Project at Indian Institute of Technology BHU (IIT BHU)

- Sentimental Analysis of Varanasi based hotel reviews using supervised learning methods.
- Using NLTK and ML trained a model which gives the positive or negative sentiment forthe given reviews and ratings.  
It can rate the Hotels, PG, Resturants etc by using customer reviews.
- Used classifier algorithms are:- NaiveBayes, SVM, DecisionTree and Logistic regression. Further accuracy increased by Ensemble Learning.  
Tools:-Python, NLTK,Confusion Matrix,Sklearn.