Abhay **Kumar**

Sr. Data Science Engineer · Data Science · Data Engineer · Leadership

☎ (+91) 8090470606 | ✉ kumarabhay.2016@iitkalumni.org | ⬡ kmrabhay | ✍ iamabhaykmr | in iamabhaykr | 🐦 @iamabhaykmr

Education

Indian Institute of Technology Kanpur
Kanpur

B.Tech in Computer Science and Engineering
Jul. 2011 - Jul. 2016

- **1.**Matrix Theory and Linear Estimation. **2.**Probability and Statistics. **3.**Principles of Database Systems. **4.**Advance Algorithm
- **5.**Machine Learning: Tools, Techniques, Applications. **6.**Data Structures and Algorithms. **7.**Randomized Algorithm. **8.**Game Theory

Anand Kumar's Super 30
Patna

JEE - Physics, Chemistry, Mathematics
Jun. 2010 - Jun. 2011

- Secured All India Rank 2356 in IIT-JEE 2011 among 450K aspirants
- Secured All India Rank 574 and State Rank 11 in AIEEE-2011

Work Experience

Trell Experiences Pvt. Ltd.
Bengaluru, Karnataka

Sr. Data Science Engineer
Nov.2019 - Present

- Built and deployed feed personalization algorithm using lightFM hybrid model for personalizing users based on historical and recent behaviour.
- Built and deployed near real time data migration ETL pipelines for migrating data from MySQL, MongoDB, Sheets, Big Query to ClickHouse, MySQL, Big Query and Redshift accordingly.
- Built and Deployed Multiclass Classification using BERT for predicting the category of videos based on textual information.
- Integrated and deployed Language Detection based on speech and transcription of speech using google and gnani.ai api.
- Implemented AI thumbnail generation, duplicate video detection and video padding detection using opencv.
- Built and Deployed NSFW Text Detection and Pre-trained NSFW Video detection and build backend for integration with internal QC dashboard for daily QC of detected nsfw content. Also Built model for intro/outro detection in the video based on transcription of the video
- Worked on Data and backend part for Creator Insight Dashboard(CID: `https://cid.trell.co/?userhandle=subhachugh`). Collected social media and google trends across location, language and category. Built vernacular content recommender models for CID.
- Built Custom object detection model using YoloV3 framework for competitors logo detection and masked detected logos using ffmpeg.
- Built & Published Trell-AI-Utils pip package on pypi for using common utils code in separate project repositories related to data science.
- Used Amazon MKS, Celery, Sagemaker, Uwsgi/Nginx, Docker and Kubernetes to deploy models for real time usage with scalability.
- Hired and built data science and data analytics team with full time employees and interns.
- **Tech Stacks:** Python, PySpark, ML, DL, CV, NLP, Flask, Docker, K8S, ELB, SageMaker, Uwsgi, Nginx, Amazon MSK, Celery, Redis, MySQL, BigQuery, ClickHouse, MongoDB, Databricks, AWS, GCP

Uniplatform Tech Pvt. Ltd
Delhi

(Lead) Machine Learning Engineer
Jul 2018- Nov 2019

- Built gradient boost models for predicting price and area with accuracy of **90.65**% and **85.67**% respectively
- Developed an android app and backend for uploading and processing real estate WhatsApp and Telegram groups messages concurrently in real time. Also developed various features in Multiply My Leads app and fixed bugs.
- Automated social media property posting for Facebook, Twitter, Linkedin, Instagram and Pinterest using APIs
- Implemented binary text classifier using CNN with accuracy of **95.96**%. Implemented ANN neuroNER model for extracting important information from classified relevant text with accuracy of **93.78**%
- Implemented captcha detection models using CNN with **99.52**% accuracy for various types of captchas
- Customized spacy sentence detector model for separating multiple real estate WhatsApp messages from a paragraph
- Hired and trained three data science interns for various responsibilities and led data science and android projects
- Developed property posting engine using selenium for posting properties on real estate websites and social media
- Wrote backend and android code to automate account creations and property postings on various real estate websites
- Wrote a scalable module called analyticrecorder to store insights and logs into influxdb and created visualization dashboards and alerts to monitor overall systems and products health using grafana
- **Tech Stacks:** Python, Django, Flask, Android, Celery, MySQL, MongoDB, InfluxDB, Grafana, Uwsgi, Nginx, OpenCV

RedSky Technologies
Mumbai

Senior Developer
Dec 2017- May 2018

- Built neural network model using RCNN and CNTK to identify objects in fridge with the accuracy of **78.54**%
- Developed retail genie website in Python using Django framework and deployed it on Microsoft Azure.
- Developed various features in POS streaming website using java, spring and hibernate and optimized existing code to load website faster.
- **Tech Stacks:** Python, OpenCV, Django, MSSQL, Java, Spring, Hibernate, PostgreSQL, JavaScript

Quantum Four Analytics - Artificial Intelligence Lab
Mumbai

Data Scientist & Programming Consultant
Feb 2017 - Nov 2017

- Wrote massive dynamics module for various data science phases in R along with various machine learning algorithms.
- Implemented ML and NLP algorithms for predicting posts like count and topic classification respectively.
- Developed local server, central server, desktop app and LibreOffice extension for First Vault project. Central server authenticates user via local server using OAuth2 spring security and upload files in encrypted format
- **Tech Stacks:** Java, R, Python, MySQL, Swing, Jersey, Spring, AWS

Bharat Matrimony App(Matrimony.com Limited)
Chennai

Technical Consultant
Jul 2015 - Aug 2016

- Developed block unblock, similar profiles, success stories, membership details, upgrade promos, and modified existing functionalities. Solved various bugs and optimized existing code.
- Got Accomplisher Award in recognition of valuable contributions.
- **Tech Stacks:**Java, Android

Blue Water Trade Winds Pvt. Ltd *Dehradun*

Machine Learning Intern: Project E-Therm *May 2015 - June 2015*

- Extracted data from a large database and applied mathematics and thermodynamics to build features out of extracted data. Built multivariate regression model to predict temperature
- Implemented projection plan for cargo heating using predicted temperature.
- Worked together with the dev team to develop interfaces to show above results.
- **Tech Stacks:** Python, scikit-learn, Django, MySQL

IIT Kanpur (Course: Machine learning tools and techniques) *Kanpur*

Number Plate Text Recognition *Apr 2016 - Apr 2016*

- It reads registration-number of four wheeler vehicles in a video in real time
- It extracts frames from video and use kNN classifier to recognize characters in the extracted license plate region
- Worked together with the dev team to develop interfaces to show above results.
- **Tech Stacks:** Python, scikit-learn, OpenCV

Extracurricular Activity

Savagtagami Association(IIT Kanpur) *Kanpur*

Android Coordinator and Community Volunteer *May 2013 – May 2015*

- Impacting lives of special kids through android apps on Aakash tablet and also visiting NGO on regular basis for tracking their progress and interacting with them

Techkriti (IIT Kanpur) *Kanpur*

Secretary *Mar 2012*

- Responsible for organizing and coordinating with various team participated in the event

Athletic Club(IIT Kanpur) *Kanpur*

Active Member *Jun 2011 - May 2015*

- Actively participating in running competitions in college
- Completed my first 10KM marathon in college in 2nd year

Programming Club(IIT Kanpur) *Kanpur*

Active Member *Jun 2011 - May 2015*

- Actively participating in programming competitions in college

Self Discipline & Fitness *Mumbai, Delhi, Bengaluru, Motihari*

Health & Fitness Enthusiast *Jun 2017 - Present*

- Actively participating in various marathon
- Participated in barefoot 10km marathon in coorg and awarded finisher medal by Milind Soman
- Completed 100 days out of 100 days of cycling challenge in Jul 2020 - Oct 2020
- Completed 97 days out of 100 days of running challenge in Jul 2020 - Oct 2020