



Bibhash Chandra Mitra

A dynamic professional, targeting challenging & rewarding opportunities in Data Science/Machine Learning / **Artificial Intelligence** with an organization of high repute in **Medical Consultancy/Aviation Industry** preferably in **Pune, Bangalore, Kolkata**

LinkedIn: <https://www.linkedin.com/in/bibhashmitra220896/>

Kaggle: <https://www.kaggle.com/bibhash>

Github: <https://github.com/Bibyutatsu>



Profile Summary

- ❖ A focused and goal-oriented professional with **11 months of industrial exposure in Data Science, Machine Learning-supervised/unsupervised, Artificial Intelligence and Algorithms**
- ❖ Alumni of **IIT, Kharagpur**; graduated with a **Major in Aerospace Engineering and a Minor in Computer Science Engineering**
- ❖ Currently associated with **Innoplexus Consulting Services Pvt. Ltd.**; working on critical projects like **Novel Drug Discovery with AI and Clustering Graph Networks for Entity Normalisation**
- ❖ Received **rating of 5/5 employee for "Outstanding" performance for two quarters (Jun'19 to Dec'19)**
- ❖ **Expertise in** OCR Engines, Deep Learning, Data Exploration and Visualization, Predictive Modelling and Optimization
- ❖ Proficiency in using various **AI techniques such as RCNN, VAE, GAN and RL**
- ❖ **Worked on projects** such as 'Table Detection and Extraction using FRCNN and Image processing' & 'Hierarchy using Graphs
- ❖ Hands-on experience in **Docker, Virtual Environments, Anaconda, DGX-1, Tesla V100**
- ❖ **Modeling:** Designing and implementing statistical/predictive models and cutting edge algorithms by utilizing diverse sources of data to predict demand, risk and price elasticity; analyzing large simulation datasets
- ❖ **Experienced in analyzing data by using statistical techniques** like Regression / Factor/ Cluster Analysis, PCA, SVD, LDA etc.
- ❖ **Exposure in** designing procedures to ensure timely availability of business information to facilitate the critical decision-making process
- ❖ Steering the project for converting PDF into a categorically normalized structured Data Frame and built products and proofs-of-concept in the areas of Graph Clustering for Entity Normalisation
- ❖ Skills in analysing & interpreting unique problems, with combination of and logical & analytical thinking to find the right solutions



Education

- ❖ **Dual Degree from IIT, Kharagpur** in 2019
 - **Major in Aerospace Engineering** with CGPA 8.45/10
 - **Minor in Computer Science Engineering** with CGPA 8.20/10
- ❖ **12th** from DAV Public School, Kota, AISSCE in 2014 with 92.4%
- ❖ **10th** from Kendriya Vidyalaya, Burdwan, SSCE in 2012 with CGPA 10



Academic Project

Title: Optimal Simultaneous Rendezvous of Fixed-Wing UAVs in 3D

Mentor/Co-Author: Dr. Sikha Hota **Period:** Jul'17-Apr'19

Highlight: Selected at AIAA SciTech Forum 2019, San Diego, California, USA

Description: The project was to find an ideal route in an area filled with static as well as dynamic obstacles that keep a safe distance from the obstacle and take into account the UAV turning constraints. Used circular approximations to approximate the path when turning, such that the direction changes smoothly and the total distance shift is small. Employed Fortune's algorithm for creating Voronoi diagrams around different real-time boundaries that shift along with the moving obstacles



Soft Skills

- Team Player
- Communicator
- Innovator
- Thinker
- Analytical



Core Competencies

Machine Learning	★★★★★
Deep Learning / Neural Networks	★★★★★
Computer Vision / Image Processing	★★★★★
Probability and Statistics	★★★★★
NLP / Text analysis	★★★★★



IT/Technical Skills

- ❖ **Programming Languages:** Python, C#, C, C++
- ❖ **Deep Learning:** Faster RCNN, YOLO, Inception and ResNet Architecture, GCN, VAE, GAN
- ❖ **OCR Engines:** Tesseract OCR Engine, Asprise OCR
- ❖ **Image Processing:** OpenCV, Skimage
- ❖ **NLP:** Stanford Core NLP, Spacy, NLTK
- ❖ **Libraries:** Keras, Tensorflow, Pytorch
- ❖ **Modules:** TensorRT, Rapids, Kubernetes



Extracurricular Activities

- ❖ Voted for the position of General Secretary Technology for LLR Hall of Residence, Jul'16-Apr'17
 - Responsible for our Hall's performance and participation in 10 different competitions in General Championship
 - Personally managed all the teams consisting of more than 150 participants and supporting personnel



Areas Of Interest



Book Reading



Cycling



Travel



Work Experience

Since Jun'19 with Innoplexus Consulting Services, Pvt. Ltd., Pune as Associate Data Scientist

Project : Drug Discovery using AI

Period: Sep'19-Present

Key Result Areas:

- ❖ Creating new compounds which show similar properties as the target ligands but are novel by utilizing different Variational Autoencoder Networks (VAE, GVAE and JTVAE) with Reinforcement Learning
- ❖ Utilizing new Junction Tree Models for encoding the molecules
- ❖ Developing custom reward functions based on Self Organising Maps (SOMs); reinforcing the model to sample from specific parts of the latent space of the VAE
- ❖ Integrating custom tokenizers for encoding the SMILES structure
- ❖ Benchmarking models using Zinc dataset and MOSES benchmarking tool

Project: Hierarchy using Graphs

Period: Sep'19-Nov'19

Key Result Areas:

- ❖ Generated different Proof of Concepts to find the hierarchy of different nodes using Graphs
- ❖ Utilized:
 - Neo4j to find the hierarchy of nodes in a graph using a custom function built including Louvain Modularity and Centrality Algorithms such as Page Rank
 - Graphs to create clusters of similar nodes and NLP to find the relationship between two nodes which do not share a Graph edge directly but are related to each other
 - LDA to discover keywords and essential nodes for the graph

Project: PDF Extraction

Period: Jun'19-Sep'19

Key Result Areas:

- ❖ Performed segmentation of paragraphs in a scanned PDF by using rule-based Image Processing Techniques
- ❖ Employed Tesseract OCR on these segments to get the contents and separated different sections based on the type of content
- ❖ Utilized PDF encodings to convert a digitally generated PDF to HTML
- ❖ Documented, recorded style information and segregated different contents in PDF based on measurable style change
- ❖ Originated different information like headers, footers and page numbers to refine the final output of the PDF
- ❖ Analyzed the structured PDF data and used Machine Learning to find information like Author, Title, Affiliation and so on to map PDF and the author
- ❖ Executed Differential Analytics on top of the extraction for information which shows the consistency of the contents of the PDFs (mainly for PDFs which are prescriptions of patients)
- ❖ Produced NLP based post-processing scripts for refinement of the output; merged unnecessarily separated contents

Project: Table Extraction

Period: Jun'19-Aug'19

Key Result Areas:

- ❖ Performed detection of a table by using both Rule-based and Deep Learning Modules
- ❖ Worked on IOU approach to maximize the coverage and accuracy of table detection
- ❖ Developed a custom Table-Maker function that converts any table (different background colors, no grids and so on) into a standard table which can be used for extracting the contents of the table to generalize the table extraction module
- ❖ Employed Tesseract-OCR to fetch textual information in the generalized table and mapped it onto a preferred file type



Internships

Organization: Innoplexus Consulting Services, Pvt. Ltd., Pune

Period: May'18-Jul'18

Role: Utilised Machine Learning and Image Processing methods to segment an image of a document into Title, Author, Affiliation and Abstract and map it onto a CSV file. Performed detection of a table from an image of a document and converted that into CSV format using deep learning and rule-based mechanisms. Modified wrappers for Asprise OCR Engine and Tabula thus contributing to Tesseract OCR Engine, a 20-year long project of HP and Google.

Organization: Invention Labs, Chennai

Period: May'17- Jul'17

Role: Administered Machine Learning along with Decision Trees based algorithms for processing text. Assisted in creating a language-specific engine capable enough to process sentences into a tree structure and in turn be converted into Universal Natural Language.



Academic Achievements/Competitions

- ❖ Inter IIT Tech Meet, 2017- UAV Design
 - Designed a UAV with 58 seconds glide time
 - Secured the fourth position in the event and achieved an overall silver at inter-IIT
- ❖ Intelligent Game Design-Artificial Intelligence
 - Gained knowledge of Unity 3D and incorporating AI in different elements of a game like NPC & environment
 - Developed a game called 'Beyond Infinity', about space and physics simulator



Personal Details

Date of Birth: 22nd August 1996

Languages Known: English, Hindi & Bengali

Address: Ashok Meadows, Hinjewadi Phase-I, Pune-411057, MH