

Snehal Bhartiya

snehal.bhartiya@gmail.com — LinkedIn — Webpage

EDUCATION

Indian Institute of Science Education and Research, Pune BSMS

Pune, Maharashtra

Aug 2015- June 2020

- **DST-Inspire Fellow**: Granted **INR 100,000/annum** by the Department of Science and Technology during studies at IISER Pune
- Co-founded Entrepreneurship and Innovation Cell of IISER Pune, and raised funding of **INR 10 Cr.** for establishing AIC-SEED under Atal Innovation Mission.
- Achieved an **A.I.R 110** in **GATE** and **A.I.R 85** in **CSIR-NET**.
- Relevant coursework: Single Variable Calculus, Multi-Variable Calculus, Linear Algebra, Probability and Statistics, Mathematical Methods, Biostatistics, Data Science, Satellite Data Analysis and Image Processing etc

Master's Thesis: Diagnosing Cervical Cancer with Deep Learning

Pune, Maharashtra

PI: Dr. Pranay Goel

June 2019- June 2020

- The primary objective of the thesis was to develop a deep learning-based system for diagnosing cervical cancer more accurately, reliably, and efficiently compared to traditional methods
- We used publicly available cervical cancer data from Herlev Hospital to train and evaluate deep networks
- We compared Multi-layer Perceptrons, Transfer Learning models, and Optimizers(ADAM, RMSprop, and SGD with momentum) using **TensorFlow**
- RMSprop and SGD achieved similar validation accuracy, but RMSprop required less computational time
- **InceptionResnetV2**, the best-performing model, achieved an accuracy of 96.25% on the Binary Classification problem and 66.78% in the Seven-Class Classification problem

WORK EXPERIENCE

IIT Kanpur

Kanpur, Uttar Pradesh

Senior Project Associate(Dept. of Computer Science and Engineering)

Jan 2024- July 2024

- Developing models to identify fraudulent behaviour of users using unsupervised learning techniques for Prutor.ai.

Auronova Consulting

Bangalore, Karnataka

Associate Consultant(Model Risk Management)

Mar 2023- Sep 2023

- Worked with the model-validation team of a UAE based client.
- Leveraged the RAG framework for monitoring a comprehensive mortgage loan portfolio worth **17 Bn. AED**
- Compiled the user data and model output for the previous 4 quarters and pre-processed it to calculate key-performance metrics.
- Created a data waterfall for excluding accounts based on the exceptions provided by the model development team.
- Contributed towards writing the **Python** code for calculating the KPIs such as **GINI index, CSI, PSI, and MAPE**.
- Reconciled discrepancies across various data sources being used to train the model, and resolved differences in accounts that affected the overall loan balance by 1.37 Bn AED, and reported the same in a **MIS reconciliation report**.

Dono Consulting

Remote

Junior Data Scientist(Model Risk Management)

Mar 2022- Oct 2022

Worked on POC for building a P2P lending platform for enabling loans to high-risk class individuals, and providing a new asset class for retail investors using openly available data by LendingClub.

- Worked towards identifying the risk associated with a loan, allotting a Loan Grade based on the A-Score and classifying a loan as good or bad (based on the probability of default), enabling investors to make better choices about accepting/rejecting a loan
- In the case of predicting default, we achieved a **91% Precision** and **96% Recall** rate in the classification problem (good vs bad loans) using a **L2 -logistic regression**
- Leveraged metrics such as **AUC** and **GINI index** to estimate the discriminatory power of the model.

SEC data extraction: Built a pipeline using Layout Parser, and **Camelot** library for extracting unstructured data (pdf format) in the U S Securities and Exchange Commission filings into a structured tabular format (csv format)

Feedback survey of state-level reforms BRAP 2020: Created an end-to-end data processing workflow using Pandas for 301-point Business Reform Action Plan (2020), which sought to promote and create sector-specific business environment enablers

Indian School of Business

Mohali, Punjab

Junior Data Engineer(Bhartiya Institute of Public Policy)

Jan 2021- Mar 2022

- Responsible for curating the data repository for **India Data Portal(IDP)**, funded by the Bill and Melinda Gates Foundation to create a one-stop portal for all openly available socio-economic data in India
- Created data processing pipeline for **30+ datasets** using **Python** libraries like **Requests, BS4, Selenium, Pandas and Apache Airflow**
- Single-handedly managed India Pulse, a sub-division of IDP for tracking high-frequency socio-economic activity in India
- Built a script using **Google's OAuth 2.0 API** for automatically uploading the data on Google Drive saving **3 hours/month**
- Contributed to the development of the IDP repository for scrapers and data processing codes using **Git**

SKILLS

Technical: Python, SQL, LaTeX, Git, TensorFlow, Tableau, Apache Airflow, Microsoft Azure

Graphics and Video Editing: DaVinci Resolve, Adobe Indesign, Inkscape