Snehal Bhartiya

snehal.bhartiya@gmailcom — (+91)-8840753551 — LinkedIn

EDUCATION

Indian Institute of Science Education and Research, Pune   **Pune, Maharashtra**
BSMS   Aug 2015- June 2020
- Relevant coursework: Single Variable Calculus, Multi-Variable Calculus, Linear Algebra, Probability and Statistics, Biostatistics, Data Science, Satellite Data Analysis and Image Processing, etc
- Achieved an **A.I.R 110** in **GATE(Ecology and Evolution)** and **A.I.R 85** in **CSIR-NET(Life Sciences)**

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)
- 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(Credit Risk)   Mar 2023- Sep 2023
- Responsbile for validating a comprehensive mortgage loan portfolio model, enhancing accuracy by 30% through rigorous model validation and monitoring techniques.
- Reconciled discrepancies across multiple data sources, identifying and resolving differences in accounts that affected the overall loan balance by $1.37 Bn AED, and reported the same in an MIS reconciliation report.
- Detected calibration errors in the model, leading to an estimated reduction in expected credit loss (ECL) provisions by 15%, which resulted in significant cost savings for the organization.

Dono Consulting   **Remote**
Junior Data Scientist(Credit Risk)   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
- Built models for 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
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**
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 data for agriculture and financial inclusion 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 20 API for automatically uploading the data on Google Drive saving 90 min/month

SKILLS

Technical: Python, SQL, LaTeX, MS Office
Libraries and Framework: Tensorflow, Pandas, Numpy, Sklearn, Camelot, BS4, Selenium, Git, Tableau
Research: Experiment Design, A/B testing, Literature Review
Graphics and Video Editing: DaVinci Resolve, Adobe Indesign, Inkscape

EXTRA-CURRICULAR ACTIVITIES

- Co-founded E-Cell of IISER Pune, and raised funding of INR 10 Crore for establishing AIC-SEED under Atal Innovation Mission
- Represented the IISER Pune Athletics team in the Inter IISER Sports Meet 2017

- Represented the IISER Pune basketball team and secured a bronze medal at the Inter IISER Sports Meet 2018
- Represented the IISER Pune basketball team and secured a gold medal at Inter IISER Sports Meet 2019
- Represented the ISB Mavericks and secured a IInd position in 100-meter and 400-meter races in ISB Sports League 2022
- Produced a short documentary on 'The Science of Sprinting' as a part of coursework(Audio-Video Communication of Science) and received an A grade