SHUBHAM LOHIYA

279, Hostel 6, Indian Institute of Technology Bombay, Powai, Mumbai - 400076

Research Interests

Knowledge Graphs, Representation Learning, Natural Language Processing, Reinforcement Learning

Education

Indian Institute of Technology Bombay — Mumbai, India
Bachelor of Technology in Mechanical Engineering — *July 2018 – Present*

- Cumulative GPA: **9.51/10.0**
- Minor Degree: Dual Minor in Computer Science & Engineering, Machine Intelligence & Data Science

Scholastic Achievements

- Two-time awardee of the **Institute Academic Prize** (top **1%**) for outstanding academic performance (2021, 2020)
- Awarded **AP grades** for exceptional performance in Economics (**1/256**) and Thermodynamics (**1/58**) (2019)
- Secured **All India Rank 888** out of 0.16 million candidates in **JEE Advanced** entrance exam (2018)
- Achieved **99.78 percentile** in **JEE-Main** entrance examination amongst 1.2 million candidates (2018)

Research Experience

Entity Alignment in Multilingual Knowledge Graphs [RnD Project] — *Jan 2021 - Apr 2021*
Guide: Prof. Soumen Chakrabarti, Department of Computer Science and Engineering — *IIT Bombay*

- **Introduction:** *Entity Alignment (EA) plays a vital role in automatically integrating multiple knowledge bases. Has BERT-INT, which gives almost saturation results on DBP15K, solved the Entity Alignment problem?*
- Conducted detailed EA performance analysis of **mBERT** and **BERT-INT** on WikiData and DBPedia KGs
- Analyzed **entity multiplicity** in the WikiData KG and experimented on novel datasets with higher ambiguity
- Examined the performance of the Interaction module on EA for low-resource language pairs like Bengali-Hindi
- Developed a web-application using Django for quick insights and comparison between several EA techniques

Online Reinforcement Learning for Lane Following [CS 748 Project] — *Mar 2021 - Apr 2021*
Guide: Shivaram Kalyanakrishnan, Department of Computer Science and Engineering — *IIT Bombay*

- **Introduction:** *Lane following is a crucial part of the Autonomous driving problem. Is it possible to learn lane following in an online and efficient fashion using a very low-dimensional discretized state-space?*
- Extracted left & right distance features from dashboard camera feed using Semantic Segmentation and masking
- Employed **Tile Coding** to encode continuous state-variables like velocity, steer and throttle in discrete form
- Handcrafted a multi-objective reward function for our online **Q-learning** agent, which learned a policy on our surprisingly small **5-dimensional** state-space, and achieved collision-free lane following on the 300m test-track

Professional Experience

Data Scientist | Anheuser-Busch InBev | Growth Analytics Center — *May 2021 – July 2021*
AB-InBev aims to embed Analytics in Business Operations to streamline and optimise tasks

- Engineered a Machine Learning Framework to forecast Accounts receivables for the Mexico region, achieving a mean dispersion of $<$ **5 %** over **5**+ categories and delivering over **$40 million** in working capital benefits
- Developed trend and seasonal features from historical data, and data from Sales, Collections and the Economy
- Trained ensembles of models like **ARIMAX**, **XGBoost**, **Random Forest**, and **MLP** for accurate forecasting
- De-constructed the ML blackbox to analyze the decision making and feature impact using **Shapley Values**
- Awarded the **Pre-Placement Offer (PPO)** for outstanding work during the internship period

Python Developer | Avrio Energy — *May 2020 – Sep 2020*
The firm is developing AI and IoT powered technology to improve the energy efficiency of businesses

- Designed Schema, Models, and APIs in **Django** for version 1 of Avrio Energy's Outlet Manager android app
- Worked with raw time series data in **InfluxDB** from **1300**+ appliances to perform feature-extraction for **ML**

Winter Intern | Unacademy — *Dec 2020*
*Unacademy is an Indian online education technology company with **6**+ million active users*

- Planned course map and created content for a **Data Structures and Algorithms** course for GATE aspirants

Key Projects and Hackathons

Winner – Prospect 100 Global Tech Challenge | *COVID-19 Hackathon* *Summer 2020*
- Brainstormed a solution for a globally feasible, logistical solution for **vaccine** storage and distribution
- Pitched the idea live to head judge **Steve Wozniak** – the co-founder of **Apple** alongside other **top-5** teams

Automatic Raga Recognition in Indian Classical Music | *IE643: Deep Learning* *Fall 2020*
- Leveraged **tonic** normalized **pitch-tracked** frequencies of a music sample as features for raga classification
- Trained a model based on **LSTM**s with **attention** on random subsequences from the Carnatic Music Dataset
- Achieved a testset accuracy of **96.67%** with 60% majority voting, and **100%** with 50% majority voting

Mastering Atari Games using Deep Reinforcement Learning | *CS419: Intro to ML* *Spring 2021*
- Trained a Deep Reinforcement Learning agent capable of surpassing human performance on classic Atari games like Pong, Breakout and Boxing using **high-level sensory information** in the form of game screen pixels
- Compared the performance of **off-policy** frameworks like **Deep Q-Network (DQN)** and **Double-DQN**
- Leveraged techniques like frame stacking for incorporating temporal information, frame skipping for speed and efficiency, and **experience replay** for allowing the Deep RL agent to learn from memory

Camouflaged Object Detection | *GNR638: Machine Learning for Remote Sensing* *Fall 2020*
- Built the **SINet** deep learning framework to precisely detect and localize camouflaged objects in nature images
- Incorporated **Receptive Field** modules and **Partial Decoder Components** on top of a ResNet50 backbone
- Obtained output masks precisely localizing the camouflaged body if present, with a testset MAE of **0.091**

Image Quilting for Texture Synthesis and Transfer | *CS663: Digital Image Processing* *Fall 2020*
- Implemented a **patch-based algorithm** to synthesize a texture of any desired size from the given sample
- Used a modified quilting algorithm to transfer any given texture to any target image and obtained good results

Shortest Path in a Maze | *CS747: Foundations of Intelligent and Learning Agents* *Fall 2020*
- Modelled given 2D mazes as **MDP**s with appropriate states, actions, rewards and transition probabilities
- Compared Howard's Policy Iteration, Value Iteration and Linear Programming algorithms to find shortest path

Other Projects

Facial Expression Recognition | *Self Project* *Summer 2020*
- Constructed a deep **CNN** trained on the FER kaggle dataset to recognize facial expressions from **7 categories**
- Trained on **30,000+** images in **Keras** to achieve **70%** training-set accuracy and **65%** test-set accuracy
- Deployed the trained model to a web interface using **Flask** and **OpenCV** and applied it to **live video** streams

Twitter Sentiment Analysis | *Self Project* *Summer 2020*
- Created a pipeline to remove punctuation and stopwords from tweets and to perform **Count Vectorization**
- Accomplished an F1-score of **0.71** and **96%** accuracy on the test set by training a **Multi-layered Perceptron**

Star-Shooter! Game | *Self Project* *Spring 2020*
- Used the **PyGame** package to develop the game environment and render graphics, animation, and sound
- Maintained an **OOP** structure for game assets and tracked game files and versions with **Git** version control

Technical Skills

Programming Languages: C++, Python, R, Go, MATLAB, Bash, SQL
Machine Learning: PyTorch, TensorFlow, Keras, OpenCV, Numpy, Pandas, Seaborn
Web Development: HTML, CSS, JavaScript, Node.js
Software: Docker, LaTeX, Git, PowerBI, AutoCAD, SolidWorks, Arduino IDE

Key Coursework

Computer Science: Data Structures and Algorithms, Design and Analysis of Algorithms, Operating Systems[1], Digital Image Processing, Programming for Data Science, Introduction to Machine Learning, Machine Learning for Remote Sensing, Deep Learning, Intelligent and Learning Agents (I & II), Information Retrieval and Web Mining, Learning with Graphs, Organizing Web Information[1] , Statistical Machine Learning and Data Mining, Advanced Machine Learning[1]

Robotics: Microprocessors and Automatic Control, Robotics, Kinematics and Dynamics of Machines, Machine Design

Mathematics: Differential Equations, Calculus, Linear Algebra, Numerical Analysis

[1]To be taken in Spring 2022

Certifications: 6.86x - Machine Learning with Python (MIT), 6.431x - Probability: The Science of Uncertainty and Data (MIT), 18.6501x - Fundamentals of Statistics (MIT), Deep Learning Specialization (Deeplearning.ai), Fundamentals of Reinforcement Learning (University of Alberta), Game Theory (University of Tokyo)

MOOCs: CS229 - Machine Learning (Stanford), CS231n - Convolutional Neural Networks for Visual Recognition (Stanford), CS285 - Deep Reinforcement Learning (UC Berkeley)

Key Mentoring and Leadership Roles

Institute Secretary Technical Affairs | Institute Technical Council *Apr 2020 – Mar 2021*

Head of the Electronics and Robotics Club and part of a 23-member core team catering to 5000+ students

- Elected to lead and manage a team of **15+ members** to organize **20+** events, competitions and hackathons and mentor 1200+ electronics and robotics enthusiasts with an annual budget of over **INR 300,000**
- Coordinated the Institute Technical Summer Project program with **70% y-o-y** increase in completed projects
- Initiated the development of **'ERC Wiki'** - a repository of easily accessible resources for enthusiastic learners

Institute Student Mentor and Department Academic Mentor *May 2021 – Present*

- Selected based on overall performance in a rigorous process comprising of interviews, SOP and **peer reviews**
- Mentoring a group of incoming **freshmen** and **7 sophomores** with their academics and life at IIT Bombay

Teaching Assistantships | IIT Bombay

Facilitating smooth course organization, grading papers, mentoring students, conducting help sessions, etc.

- **IE 643 - Deep Learning**, Prof. P. Balamurugan, IEOR Department *Fall 2021*
- **MA 106 - Linear Algebra**, Prof. Sudhir Ghorpade, Department of Mathematics *Spring 2021*
- **MA 108 - Differential Equations**, Prof. Prachi Mahajan, Department of Mathematics *Spring 2021*
- **PH 107 - Quantum Physics**, Prof. Shankaranarayanan S, Department of Physics *Fall 2020*

Extra Curricular Activities and Other Achievements

Achievements	• Selected among the four **delegates from India** to the 5-day virtual "Humanizing Digital 2021" **AI and Data Science** conference at **Chulalongkorn University, Thailand** • Ranked **4** in India's Best Student Contest 2015 organized by RaoIIT amongst **0.3** million participants • Selected among **top 30** students in a Nationwide Aptitude Test 'DEXTER' conducted by VNIT
Mentorship	• Conducted a summer course for **Practical Python Programming** with **1000+ enrollments** • Mentored **9 freshmen** on a project to create an AI agent for mastering the snake game using **RL** • Guided 6 students with reading projects on **Deep Learning** and **Reinforcement Learning**
Technical	• **Analyzed COVID-19 data** using Python to find if there is a correlation between a country's infection rate and its score on various life factors according to the World Happiness Report 2020 • Led a team of **4** to build a radio-controlled trainer aircraft capable of dropping payloads • Constructed an all-terrain obstacle manoeuvring bot controlled using a mobile application • Participated in the **Boeing Aeromodelling Competition** at Techfest 2019, IIT Bombay • Completed a reading project on the use of Deep Learning in **Computer Vision** under SoS 2019
E-Cell, IIT Bombay	• Led a team of **5** organizers during the **E-Summit 2020**, to successfully execute **20+** talks, interviews, and lectures as a **Corporate Relations Coordinator**
Miscellaneous	• Certified in Introduction to **Public Speaking** MOOC by the University of Washington • Undertook year-long training in **Swimming** under the National Sports Organization (NSO) • Completed a two-week-long **Finance Bootcamp** conducted by the Finance Club, IIT Bombay