Saumya Gaurang Shah

Fourth Year Undergraduate

: *(+41) 0766132181*
: *saumyagshah@gmail.com, saumya.shah@epfl.ch*
: *saumyagshah*
: *saumyagshah*
: *saumyagshah*

Education

Sep'19-Present **Semester Exchange, Computer Science**, *Ecole Polytechnique Fédérale de Lausanne*, Lausanne, Switzerland.

2016–2021 (Expected) **Bachelor of Technology - Electrical Engineering and Computer Science and Engineering (Double Major)**, *Indian Institute of Technology*, Kanpur, India, *CPI – 9.6/10*.

2016 **Grade XII**, *The New Tulip International School*, Ahmedabad, India, *Result – 94.4%*.
CBSE

2014 **Grade X**, *Delhi Public School*, Ahmedabad, India, *CGPA – 10/10*.
CBSE

Internships

May'19-Jul'19 **CHILI Lab, EPFL**, *Summer@EPFL Research Intern*.
Diagnosing Dysgraphia Using Handwriting Data Analysis, Prof. Pierre Dillenbourg
Report
Github
- Contributed to a project for **classifying dysgraphia within seconds** using data from a consumer tablet
- Explored the relation between **grasp type and legibility of handwriting** using the LeapMotion Controller
- Implemented an additional **60 new features** to improve prediction accuracies and interpretability for the diagnosis of dysgraphia

Feb'18-May'18 **Auquan Inc., Bangalore**, *Data Science Intern*.
Predicting Stock Prices to Develop Trading Strategies for different stock market indices
Github
- Developed **predictive models for stock prices** in **Python** using the fundamentals of quantitative finance research
- Designed, back-tested and optimized a data-driven **quantitative trading strategy on real-world data**.
- Developed an intra-day mean reversion strategy to give **>30% return on capital (RoC)** using **Hurst values and ARIMA models**.

Nov'17-Dec'17 **Kritsnam Technologies**, *Winter Intern*.
Analysis of Analog Signals from Atlas Scientific Sensors, Prof. Ketan Rajawat
- Explored the working and physical principles of Atlas Scientific pH, EC, DO and Temperature sensors
- **Integrated these sensors with Particle Electron** Embedded Platform via mobile data(2G) to retrieve their data using Atlas Scientific EZO Circuits
- Utilised this data to propose a **regression technique** that **reverse engineers** EZO circuits by establishing a relationship between the analog signals received from the sensors and the corresponding physical quantities

Projects

May'19-Aug'19 **Julia Seasons of Contributions**, *Student Developer*.
Model Zoo for Turing.jl, Open Source Contribution
Blog
Github
- Implemented the following **graph and time series** models in Julia using bayesian inference with the probabilistic programming language (PPL) Turing:
  - Mixed Membership Stochastic Blockmodel
  - Autoregressive (AR(p)) and Moving Average (MA(q)) models
  - Autoregressive Integrated Moving Average (ARIMA(p, d, q)) model
- Wrote blog posts, most of which are published in the Medium publication Towards Data Science, describing the work done in detail

Apr'19-Feb'19 **Scaling Up Gaussian Processes**.
Course Project for Topics in Probabilistic Modeling and Inference (CS698X), Prof. Piyush Rai, IIT Kanpur
Report
Github
Presentation
- Explored the approach of Scaling Gaussian Processes via Inducing Point Method
- Implemented **various approximations via Inducing points methods** and compared these methods
- Implemented the learning of Gaussian Process kernels from data using the **Spectral Mixture kernel**

Sep'18- Nov'18
Report
Github
Presentation

Explainable Machine Learning.
Course Project for Introduction to Machine Learning (CS771), Prof. Piyush Rai, IIT Kanpur
- Developed a web application to **explain the prediction** of any classifier on the user's dataset using LIME
- Implemented **feature visualisation** using matrix factorisation by generating adversarial examples using **BFGS** method
- Explored state of the art techniques for visualising CNNs using **Lucid and neuron group methods**

Aug'18- Nov'18
Report
Github

SLAM: Extraction of Backend Data and Comparison of Backend Solvers .
Undergraduate Project, Prof. Ketan Rajawat, IIT Kanpur
- Modified and implemented Google Cartographer to **retrieve backend data** for our own dataset in the form of raw nodes and edges before being published as visual marker arrays
- Implemented the solvers **Toro and g2o** on this retrieved backend data
- Compared error metrics for these solvers on the Manhattan3500 and city10000 datasets

May'18- Jul'18
Github

Multiple Sensors Dataset Repository.
Summer Project, Prof. Ketan Rajawat, IIT Kanpur
- Collaborated to **record and benchmark** a new dataset comprising of multiple sensors for research in short and long term SLAM
- Developed a **ROS Node** in **Python** to calibrate LIDAR and odometry parameters for this dataset by implementing CSM and calibration packages developed by Andrea Censi
  - Developed a **Bash + Python** script to automate the entire process
- Implemented and compared the results of Google Cartographer, Hector-SLAM and open-karto

Awards and Achievements

2019 Won the **Grand Prize** (team of 3 students) at Deloitte TechnoUtsav 2.0 - Cash Award of **INR 500,000** and a **job offer** from US-India Deloitte.

2016-2018 Received the **Academic Excellence Award** for 3 consecutive years at IIT Kanpur, awarded to **Top 10%** students

2016 Ranked in **National Top 1%** (amongst 1,200,000 candidates) in JEE Main

2016 Ranked in **National Top 4%** (amongst 150,000 candidates) in JEE Advanced (IIT-JEE)

2016-Present Awarded **A*** in 4 courses (**for outstanding performance**) including **Complex Analysis**

2015 Received Pradeeptam Award for securing 2nd position in academics in class XI

Technical Skills

Languages PYTHON, C/C++, Julia, SQL, Shell(bash), MATLAB/OCTAVE

Frameworks Pytorch, ROS, Particle Electron

Utilities Git, TravisCI, LaTeX, Scikit-Learn, Numpy, Pandas, Autodesk Fusion 360, Adobe Premier Pro

Relevant Coursework

Mathematics **Mathematics of Data: From Theory to Computation****, **Probability and Statistics**, **Linear Algebra**, Ordinary Differential Equations, Partial Differential Equations, Complex Analysis(A*), Introduction to Real Analysis

Computer Science **Introduction to Machine Learning**, **Probabilistic Modeling and Inference**, **Visual Recognition**, **Introduction to Natural Language Processing****, **Algorithms****, **Data Structures and Algorithms**, **Computer Organisation**, **Fundamentals of Programming**

Electrical and Electronics **Information Theory and Coding****, **Control Systems Analysis**, Signals and Systems, Communication Systems, Principles of Communication, Power Systems, Digital Electronics, Fundamentals of Electronics, Microelectronics-I, Introduction to Electrical Engineering

*** - Ongoing at EPFL*
A - Exceptional Grade*

Miscellaneous

2019 Standardized Scores- **GRE**: 332/340, **TOEFL**: 116/120

2017 Institute Table Tennis Team- Awarded Player of Summer Camp for exceptional performance in Table Tennis

2013 Secured 2nd position in the City Finals (Ahmedabad) of HDFC Life Spell Bee