

Saransh Gupta



[Contact](#) [Email](#) [LinkedIn](#) [Github](#) [Website](#)

ACADEMIC QUALIFICATIONS

| Year | Degree | Institute | CGPA/Marks |
|------|--------------------------------------|---------------|------------|
| 2022 | M.Tech in Engineering Product Design | IIT Kharagpur | 9.26 / 10 |
| 2022 | B.Tech in Engineering Product Design | IIT Kharagpur | 8.07 / 10 |

PUBLICATIONS

- (Gupta et al.) An integrative machine learning and Bayesian modeling approach highlights the crucial roles of the Rho-GDI signaling pathway in the progression of non-small cell lung cancer (NSCLC); implications for drug target discovery, (under review), **IEEE Journal of Biomedical and Health Informatics (JBHI-01644-2021)**, 2021
- (Gupta et al.) Development of a virtual reality-based fire training simulator and machine learning-based path guidance system (working paper), **IHIET-AI**, 2020, **Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland**

RESEARCH EXPERIENCE

Osaka University, Japan

[Jan'20 - Sep '20]

Guide: [Dr. Kenji Mizuguchi](#), Mizuguchi Lab, NIBIOHN, Osaka, Japan

Project: Predict the Non-Small Cell Lung Cancer (NSCLC) using Machine Learning, identify its potential drug targets

- Extracted **412** essential genes out of **10,077** by applying **Boruta** Feature selection on their gene expression dataset
- Obtained **F-1 score** of **1.0** on validation and **0.98** on test dataset by using the **XGBoost** model to predict NSCLC
- Predicted drug targets for the NSCLC by simulating a **Bayesian Network Model** on the Rho-GDI signaling pathway
- Discovered methodology leads to an accurate treatment of the disease impacting **85%** of the lung cancer patients

Tools and Software: Python, TargetMine, scikit-learn, smote, NetworkX, NumPy, pandas, Plotly, joblib

INTERNSHIP EXPERIENCE

ZS Associates Inc. | Data Science Associate – Intern

[Jan '21 – June' 21]

Project - 1: Extract biomedical text dataset, identify entities, and classify if there exists a relation between entities

- Created a pipeline to extract texts from PubMed database, identifying the entities using **Selenium** and **PubTator**
- Implemented **Binary Classification rules**, devised **four** labeling functions using bio-verbs, co-occurrence of entities
- Generated a training dataset utilizing the four labeling functions in **Snorkel** by applying the **Weak Supervision**
- Achieved F1 score of 0.88 on the gold-standard dataset in relation-classification by training **RoBERTa base** model

Project - 2: Identify the type of relationship between two entities if it exists from the results of the Project-1

- Created a new set of **three** labelling functions for **relation-type identification** by using the results of the project-1
- Attained **F1 score** of **0.83** on the gold-standard dataset using **XGBoost** Model followed by feature engineering

Tools and Software: Python, TensorFlow, Transfer Learning, Medline-Plus API, PubTator, Selenium, Snorkel

Sapio Analytics India PVT. LTD. | Data Science Intern

[May '20 – July '20]

Project: Build a data-driven **Decision Support System** for the prediction of COVID-19 at the hyper-local level

- Fabricated **SEIRD** model using Migration, Lockdown conditions, reduced RMSE up to **4.8%** by tuning parameters
- **The Govt. Of Telangana** has used this application to mitigate COVID-19 while maximizing the economic activities

Tools and Software: Python, Jupyter Notebook, scikit-learn, SciPy, AWS Server, NumPy, pandas, Plotly, joblib

- Worked towards automating the **bounding box tracking** in different conditions of Day light and road semantics
 - Programmed **SuperPoint** algorithm to extract the features of an image for Point tracking and Visual Odometry
- Tools and Software:** Python, pandas, numpy, tensorflow, pytorch, OpenCv, matplotlib, Jupyter Notebook

SKILLS

Programming & Applications: Python, C++, R, C, Image Processing, Computer Vision, Machine Learning

Software: Tableau, MATLAB, ANSYS, Solid works, Unity3D, R-Studio, SOLIDWORKS, Visual Studio, MS-EXCEL

ACHIEVEMENTS

- **Academic Achievements:**
 - Featured as one of the **Top 30 Undergraduate Achievers** of IIT Kharagpur in the UG Achievers Directory 2020
 - Conferred merit-based scholarship of **2200 €** by The A*Midex Foundation of **Aix-Marseille University, France**
 - Selected among **Top 5%** out of all for the summer fellowship at **The Institute of Science & Technology Austria**
 - Got featured in the ISE Newsletter Autumn-2020 under the Department Spotlight of **ISE fights COVID-19**
- **Internship Achievements:**
 - Awarded as **Intern of The Month** for my contribution as a Data Analyst at Sapio Analytics by the CEO
 - Offered a position of Data Scientist as Pre-Placement Offer from Sapio Analytics PVT LTD

COURSEWORK INFORMATION

- **Department Courses:** Probability and Statistics, Programming and Data Structures, Machine Learning Foundations
- **MOOC Courses:** Machine Learning, Data Analytics, Algorithms in programming, Deep Learning, Virtual Reality

COMPETITIONS / CONFERENCES

Young Data Scientists annual meetup at Kaggle - days, Dubai World Trade Centre [Mar 2020]

International Conference on Human Interaction & Emerging Technologies: Future Applications [Aug 2020]

- Virtually Presented the research paper on "Development of a virtual reality based fire training simulator and machine learning based path guidance system: A case of hospital fire breakouts" at Lausanne, Switzerland

ML Summer School 2021 | Amazon | Virtual [July 2021]

- Passed a competitive process consisting of coding and MCQ rounds to check understanding of Probability
- Got a chance to interact and learn from some of the Subject Matter Experts and Data Scientists at Amazon

Jumpstart | Publicis Sapient | Virtual [July 2021]

- Passed the aptitude round (**Level-1**) and Coding Round (**Level-2**) to be shortlisted for the on-campus conference

Winner| Data Analytics and Business Case Study Event| Shaastra, IIT Madras [Jan 2020]

- Recognition for most accurate Machine Learning Model, presented at IIT Madras, representing IIT Kharagpur

Gold Winner| Inter-Hall Data-Analytics competition| IIT Kharagpur [Feb 2019]

- Worked in the Visualization of data-set and then making interpretation of trends followed by different variables
- Applied **XGBOOST** followed by feature engineering and obtained **82.55%** on training and **80.66%** on the test data
- Awarded best in the category of best F-1 Score Model, Overall stood second in terms of Business Case Study

EXTRA-CURRICULAR ACTIVITIES

- Part of the **Gold-Winning** water-polo team of Azad Hall of Residence at Inter-Hall General Championship, 2019
- Lead presenter of the **Gold-Winning** team of Azad Hall of Residence at Inter-Hall Gardening Competition, 2019
- Raised a sponsorship of **51,000/- INR** at KHOJ, the treasure hunt event organized by Azad Hall of Residence, 2019
- Recognised by the Dean, Student Affairs for mentoring six juniors under the Student Mentorship Programme
- Managed a technical workshop on SolidWorks at Jadavpur University organized by **teamKART**, IIT Kharagpur