

# Anisha Saha

✉ anisha0325@gmail.com

☎ 9647195300

🌐 My Webpage

in anisha0398

## Experiences

- Feb'23 – Present **Research Assistant**, Indian Institute of Technology, Patna (IITP).  
*Worked on utilizing multimodal data and multi-tasking frameworks for tasks such as clinical dialogue summarization, medical domain identification, dialogue generation.*
- Aug'22 – Present **Data Scientist**, Micron Technology.  
*Worked with semiconductor fabrication data, creating end-to-end prediction pipelines, utilizing tools like GCS, Snowflake, Pyspark and Airflow for automation and analysis.*
- Jan'22 – May'22 **Teaching Assistant, Linear Algebra and Applications**, CMI .  
*Designed and evaluated assignments, conducted tutorial sessions on crucial topics and was actively involved in the final evaluation.*
- Teaching Assistant, Computing and Data Science**, Sai University .  
*Guided students through a spectrum of topics, starting from fundamental Python programming and algorithms, to the basics and practical applications of Machine Learning.*
- Jan'22 – Apr'22 **Julia Statistics Developer**, XKDR Forum.  
*Contributed to the Surjey.jl package in Julia, required to study complex survey data.*
- Sep'21 – Dec'21 **Data Science Intern**, Delab Research Pvt. Ltd.  
*Curated datasets, analyzed papers, deployed classical ML and DL models for a smart energy monitoring system and day-ahead solar energy forecasting.*
- Jul'21 – Sep'21 **Summer Intern (Digital & AI)**, Legato Health Technologies.  
*Generated statistical quality control parameters for models processing audio transcripts, performed root cause analysis to address data anomalies and evaluate model performance.*

## Education






- 2020 – 2022 **M.Sc. Data Science, Chennai Mathematical Institute (CMI)**. CGPA : 9.44  
*Predictive Analytics, Machine Learning, Deep Learning, NLP, CV, Distributed Computing and Big Data, RDBMS, Data Structures, Financial Modelling, Time Series Analysis.*
- 2017 – 2020 **B.Sc. Mathematics, St. Xavier's College, Kolkata**. CGPA : 7.27  
*Real Analysis, Linear Algebra, Abstract Algebra, Numerical Analysis, Topology, Statistics. Electives : Physics and Chemistry*
- 2001 – 2017 **Mary Immaculate School** .  
ISC (XII): 96% - *Mathematics, Physics, Chemistry, Biology, English.*  
ICSE (X): 96.8% - *Science(PCB), Mathematics, Computer Science, English, HCG.*

## Publications

- 2023 **Experience and Evidence are the eyes of an excellent summarizer! Towards Knowledge Infused Multi-modal Clinical Conversation Summarization**,  
*ACM Conference on Information and Knowledge Management (CIKM)*
- 2021 **Guide to Distributed Representations in ML**  
*Analytics India Magazine*





## Projects

---

- 2022  **Diabetic Retinopathy Detection**  
Created an automated analysis system capable of classifying each retinal image to one of the 5 classes, corresponding to the severity of the disease.
- 2021  **Day Ahead Solar Photovoltaic Forecasting of Power Output : A Comparative Study**  
A comparative study of the prediction models, generated using the techniques of ARIMA, SVM, ANN and XGBoost, for forecasting day ahead power output of Photovoltaic systems. Evaluating ML models against traditional time series forecasting techniques. Identification of relevant features for day ahead energy production.
-  **Facial Recognition using Eigenfaces**  
Compared characteristics of a face image to that of face images of known individuals and non-face images. The approach treats the images as two-dimensional, rather than requiring three-dimensional geometry, by taking advantage of the fact that faces are generally upright. The method used is Eigenfaces, where we project the images into a low-dimensional feature space using Principal Component Analysis.
-  **Unreeling Netflix**  
Built a movie recommendation engine on Apache PySpark. Demonstrated how Netflix uses a recommendation system based on Big Data to recommend movies to its users. Focused on collaborative vs content-based filtering for recommendation.
- 2020  **Integrating Factors of 1st order Ordinary Differential Equations**  
This was my Bachelor's dissertation. Found a correlation between the integrating factors of a first order ordinary differential equation and infinitesimal Lie group transformations. Corresponding to each integrating factor of an Ordinary Differential Equation (ODE), a symmetry group can be found, under which the ODE is form-invariant.






## Skills

---

- Languages  Strong reading, writing and speaking competencies for English, Hindi and Bengali
- Coding  Python, R, SQL, Julia,  $\LaTeX$ .
- Tools  PyTorch, Tensorflow, PySpark, Git, Snowflake, Google Cloud Services, Airflow.
- Interpersonal  Academic research, Teaching, Content creation.





## Certifications

---

- DeepLearning.AI  Neural Networks and Deep Learning
-  Structuring Machine Learning Projects
-  Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization
- Google Cloud  Smart Analytics, Machine Learning and AI on GCP
-  Google Cloud Big Data and Machine Learning Fundamentals

## Achievements

---

- 2023  Volunteered at **The International Conference on Learning Representations (ICLR'23)**.
- 2020  **All India Rank 4**, in the entrance exam for MSQMS course of Indian Statistical Institute (ISI).
- 2017  **1st Rank**, in school in ISC (12th Boards).
- 2015  **3rd Rank**, in school in ICSE (10th Boards).