Srijan Chattopadhyay

Curriculum Vitae

3 Vivekananda Nagar
Belgharia, Kolkata-700056
+91 9163885964
srijanisi26@gmail.com

Education

2021–present **Bachelor of Statistics (B.Stat), *Indian Statistical Institute (ISI)*, Kolkata, WB, India**, *Marksheet Link upto 2nd Year*.

- First Year, 1st Sem: 87.8%
- First Year, 2nd Sem: 91.4%
- Second Year, 1st Sem: 92.2%
- Second Year, 2nd Sem: 93.4%

2019-2021 **Higher Secondary Examination**, *Ramakrishna Mission Boys' Home High School, Rahara*.
GRADE: 484/500(O) with 99 in Mathematics and Statistics

2008-2019 **Secondary Examination**, *Ramakrishna Mission Boys' Home High School, Rahara*.
GRADE: 671/700(AA) with 100 in Mathematics

Areas of Interest

- Big data Handling & Machine Learning.
- Biostatistics
- Conformal Inference
- Change Point Detection
- Astro Statistics
- Model Selection, Missing data problem
- Spatial and Time Series
- Bayesian, Causal

Publications and Pre-Prints

1. **Chattopadhyay, S.**, Bhattacharyya, S., Basu, S., Analysis of Pleiotropy for Testosterone and Lipid Profiles in Males and Females (2023).
[medRxiv][arXiv:2312.16241]
2. **Chattopadhyay, S.**, Bhattacharyya, S., A Statistical approach to ecological modeling by a new similarity index (2022).
[arXiv:2304.01944][Under journal review]

Academic achievements

2021-23 **Indian Statistical Institute**, Awarded Prize Money for Excellent performance in 1st,2nd,3rd and 4th Semesters.

2020 **Mathematics Teachers' Association**, 2nd State Rank in Indian Olympiad Qualifier of Mathematics.

2020 **HBCSE,Tata Institute of Fundamental Research**, Selected for **INMO Training Camp**.

2021 **Chennai Mathematical Institute**, Qualified with Sri Rama Scholarship.

2021 **Joint Entrance Examinations**, **WBJEE** rank 193, **JEE MAIN** 99.45 percentile, **KIIT** 190 AIR(Selected with Scholarship), **VITEE** 810 AIR.

2020 **SUM IT, Integration, ISIK**, Selected in top 25 candidates.

2019 **JBNSTS**, Selected for junior scholarship.

2015-18 **Brilliant International Olympiad of Mathematics**, ranks:17,33,130,374(arranged by **FCES**).

2016-18 **International Olympiad of Mathematics**, ranks: 12,102,115(arranged by **Silver Zone Foundation**).

Research Projects

- **Inference problems on spectral analysis in spatial datasets** (2023)
  - **Guide: Dr. Soudeep Deb, Decision Science, IIMB**
  - **Guide: Dr. Sayar Karmakar, Department of Statistics, University of Florida**
  - **Project Partners: Swapnaneel Bhattacharyya**
  - Overview of the ongoing work available here
- **Joint modeling of multivariate categorical longitudinal outcomes and time-to-event data** (2023)
  - **Guide: Dr. Kiranmoy Das, Beijing Institute of Mathematical Sciences and Applications**
  - Data provided by **Tata Medical Center, Kolkata**
  - **Project Partners: Swapnaneel Bhattacharyya, Sevantee Basu**
  - Acute lymphoblastic leukemia (ALL) is the most frequent childhood malignancy. Nine studies were included (one population-based and the other eight hospital-based), which covered variable time periods (range 1985–2011). Together, they constituted 3761 children with ALL with some overlap (two studies reported patients from the Cancer Institute, Chennai and two studies reported patients from the All India Institute of Medical Sciences). But the first time, any full study of medicine and observation of a total of 5 years was conducted by TATA Medical Center, Kolkata in the very recent past. We are using that data to do some survival analysis.
- **A Machine Learning Based approach to model Himalayan Biogeography** (2023)
  - **Guide: Dr. Kumar Mainali, Data Science Lead, Chesapeake Conservancy**
  - **Project Partners: Swapnaneel Bhattacharyya**
- **Temporal Changepoint detection in spatiotemporal data using conformal inference** (2023)
  - **Guide: Dr. Soumendu Sundar Mukherjee, SMU, ISIK**
  - **Project Partners: Swapnaneel Bhattacharyya**

- **An ANOVA based Cluster Improvement method** (2023 +)
  - **Project Partners: Swapnaneel Bhattacharyya, Dr. Kiranmoy Das**
- **An ANOVA-based approach for sensitivity analysis, changepoint detection, and model selection** (2023+)
  - **Project Partners: Swapnaneel Bhattacharyya, Dr. Kiranmoy Das**

Projects/Internships/Study

Fall 2023 **Application of Stochastic Gradient Descent algorithms in spatial modeling (Ongoing)**, *Study*.
- **Guide: Dr. Debashis Paul, ASU, ISIK, & Department of Statistics, University of California, Davis**
- In spatial parametric modeling, often the parameters are estimated in likelihood-based approaches where obtaining the exact zeroes of the score function is pretty difficult. We are applying stochastic gradient descent-based algorithms and trying to obtain the zeroes of the score function by stochastic approximations.

Summer 2023 **National Institute of Biomedical Genomics**, Research Internship.
- **Guide: Dr. Analabha Basu, NIBMG**
- **Co-Guide: Dr. Anasuya Chakrabarty**
- **Co-Guide: Dr. Diptarup Nandi**
- **Project Partners: Swapnaneel Bhattacharyya, Sevantee Basu**
- We measured hormonal pleiotropy from UK Biobank raw dataset. After a bunch of data cleaning, the data was munged and then it was ready for analysis. We used the traits testosterone, triglycerides, low-density lipoprotein (LDL), high-density lipoprotein (HDL), and cholesterol. We then measured the genetic correlation between testosterone and the other traits for both males and females using LDSC score regression.

Winter 2022 **Reading Project, ISIK**, An Introduction to Non Linear Regression Methods.
- **Guide:Dr. Shyamal Krishna Dey, ASU, ISIK**
- Learned some notions of non-linearity, the asymptotic study of the estimators, and application to some real and simulated datasets.

Winter 2022 **Analysis of Long Covid data**, analyzing raw data from the districts of Gurugram, Jhajjar, Rohtak, a group project with Rudrashis Bardhan and Sevantee Basu .
- **Guide: Dr. Malay Bhattacharyya, MIU, ISIK**
- The dataset was from Gurugram, Jhajjar, and Rohtak, which was collected by means of surveying for around 6 months. We reanalyzed the data but didn't find any significant post covid syndrome.

Fall 2022 **The Artificial Intelligence Institute Of UofSC (AIISC)**, Applied a newly developed method of Joint Embedding by AIISC to memotion dataset and related analysis.

Summer 2022 **Reading Project, ISIK**, An Introduction to Bayesian Learning and Bayesian Computing.
- **Guide: Dr. Kiranmoy Das, Beijing Institute of Mathematical Sciences and Applications**
- We learned basic Bayesian methods - prior and posterior densities, choice of a prior, the asymptotic limit of the posterior distribution, bayesian tests, confidence regions, and hierarchical methods. We were also introduced to some Bayesian computing methods, mostly in a simulation framework.

Summer 2022 **Sentiment Analysis**, An Independent Study.
(Available at:Sentilysis)

Talks and Presentations

Summer 2023 **(Invited talk) Application of Machine Learning in ecological modelling**.
At NIBMG, Kolkata

Summer 2022 **(Invited talk) Introduction to Machine Learning and its applications**.
At Open day, ISI Kolkata

Technical Skills

Operating System: **Windows**, Linux.

Programming Language: **R**, Python, MATLAB(Basic), Scilab(Basic), C etc..

Software: LaTeX, Word, Excel etc..

Class Projects

Fall 2023 **Variation of Poverty Index over Household Size through Estimated Engel Curve for Chhattisgarh**.
- **Instructor: Dr. Nachiketa Chattopadhyay, SOSU, ISIK**
- Available here

Fall 2023 **Analysis of the Car Price Data**.
- **Instructor: Dr. Debashis Paul, ASU, ISIK, & Department of Statistics, University of California, Davis**
- Available here

Fall 2023 **Estimation (along with SE) of number of people having medium of language in school same as their native language for State of Chhattisgarh from 71st Education Survey Data of Whole India (Ongoing)**.
- **Instructor: Dr. Kajal Dihidar, SOSU, ISIK**

Spring 2023 **Finding Representative Basket for the State of Chhattisgarh from 68th Round Consumer Survey Data of Whole India**.
- **Instructor: Dr. Nachiketa Chattopadhyay, SOSU, ISIK**
- Avaiable here

Summer 2023 **Assessment of Human Values**, The research employs both an individual differences approach and a general comparative study to investigate value preferences within the age group 18-28 and compare the findings to existing literature..
- **Instructor: Dr. Debdulal Dutta Roy, PRU, ISIK**
- Avaiable here

Summer 2023 **Item Analysis of Syllogistic Reasoning Aptitude Test**, This study focuses on evaluating the syllogistic reasoning aptitude of 63 graduate or masters or class 11-12 students, specifically concentrating on male students aged 18 to 25 years. The primary objective is to analyze the psychometric properties of syllogistic reasoning items through item analysis, aiming to identify areas for improvement in measuring this cognitive ability..
- **Instructor: Dr. Debdulal Dutta Roy, PRU, ISIK**
- Available here

Summer 2023 **A Regression-based approach of Peak Detection in humming audio data**, Given an audio file, which is humming, I have to identify the time stamps of note change. We developed a new methodology based on linear regression and that worked well.
- **Instructor: Dr. Arnab Chakraborty, ASU, ISIK**
- Available here

Spring 2023 **Predicting Position(GPS), Weather Prediction, Prediction of Number of Covid Cases, Stock Price Prediction using Linear Kalman Filter** , We learned the method of Linear Kalman Filter, applied that on two simulated and two real datasets, and presented as a class presentation..
- **Instructor: Dr. Arnab Chakraborty, ASU, ISIK**
- Available here

Winter 2022 **Forward and Backward Digit Span**, Collected 29 samples of forward and backward digit span from 29 persons using R and analyzed the relationships and the summaries.
- **Instructor: Dr. Debdulal Dutta Roy, PRU, ISIK**
- Available here

Fall 2022 **Predicting Breast Cancer using Logistic, Probit, t-distribution(df = 43), double exponential regression(classification) and model selection**, Analyzed and Classified data of breast cancer using Logistic, Probit and similar distributional regression techniques and the **Matthews Correlation Coefficient(MCC)** and analyzed significance levels along with use of some modified resampling techniques.
- **Instructor:Dr. Shyamal Krishna Dey, ASU, ISIK**
- Available here

Spring 2022 **Face Recognition using PCA and LDA**, We used Principal Component Analysis and Linear Discriminant Analysis to make clusters of similar images from a group of 50 photos each of 200×180 dimension.
- **Instructor:Dr. Arnab Chakraborty, ASU, ISIK**
- Available here

Learning

- **First Semester:** Analysis of one variable in different topological sets, Discrete Random variables, PGF, SSRW1, SARW1, Branching process, Simple Linear Regression, Vector and Matrices 1, C Programming.
- **Second Semester:** Analysis of one variable, Sequence of Functions, Power Series, An Introduction to Fourier Analysis, Continuous Random variables, MGF, CHF, All definitions of Expectation, convergence a.s. and in probability, bivariate random variables, MLE, Simulation, Chi-Square Test, Multiple Linear Regression, Vector Matrices 2 with PCA, LDA, etc, Numerical Analysis, Basic Clustering Methods.
- **Third Semester:** Analysis [Multivariate Calculus] with differential forms, Multivariate Random variables, Large Sample models, L^p and in distribution convergence, order of different notions of convergences, Homogeneous Poisson process, Elements of Algebraic Structures, Point and Interval Estimation, Hypothesis Testing, LAD, Probit, Logit Regression, Microeconomics.
- **Fourth Semester:** Stochastic process on countable space, Non-Homogeneous Poisson process and an introduction to Poisson process in multidimensional spaces, An Introduction to Copula, Kalman Filter, Shrinkage Methods, Multivariate Statistics, Resampling Methods, Chi-Square proof, Time Series analysis, Fast Fourier Transform, Discrete Fourier Transform, Audio handling in R, Discrete Mathematics, Differential Equation, Psychology
- **Fifth Semester:** Linear Statistical Models with Random effect and Mixed effect Models. Parametric Inference (with a bit of Bayesian), Design and Analysis of Algorithms, Economic and Official Statistics, Sample Survey.

Other Information

Languages	**Bengali(Mother Tongue), English(Fluent)**, Hindi(Conversational).
Teaching	**Trains**, a number of motivated high school students for mathematical olympiads, ISI, CMI, JEE, IIT(JAM, Statistics) etc..
Hobbies	**Singing, Listening to Music, Playing Keyboard, Drawing**.

References of Ongoing Work

Further details of any field or overviews of ongoing research are available upon request.

I hereby declare that all the information provided above are true to the best of my knowledge.