ANKIT TEWARI

Master's Student in Statistics and Operations Research
Joint Interuniversity Program by UPC and University of Barcelona

Email : ankitt.nic@gmail.com
Website: http://ankitbit.github.io
Github : github.com/ankitbit
Mobile : +34-631854450

EDUCATION

- **University of Barcelona and Polytehcnic University of Catalunya (UPC)** — Barcelona, Spain
  Master of Science in Statistics and Operations Research (Data Science Specialization) — *Aug. 2017 – Present*
- **Birla Institute of Technology** — Mesra, India
  Bachelor of Science in Mathematics and Computing; GPA: (7.4/10.0) — *Aug. 2014 – July. 2017*
- **Kendriya Vidyalaya New Cantt** — Allahabad, India
  All India Senior School Certificate Examination, CBSE; GPA: (9.3/10.0) — *Aug. 2011 – July. 2013*
- **Kendriya Vidyalaya New Cantt** — Allahabad, India
  All India Secondary School Examination, CBSE; GPA: (10.0/10.0) — *Passed in July. 2011*

EXPERIENCE

- **Faculty of Informatics, Universitat Politecnica de Catalunya (UPC)** — Barcelona, Spain
  Data Scientist — *October, 2017 - December, 2018*
  - **Social Network Analysis**: Responsible for creation of a package in R which computes a variety of quality metrics for community detection algorithms. Further respoonsible for optimization of the R codes through Rcpp framework to improve speed, scalability and efficiency
  - **Quantitative Linguistics**: Python and R implementation for the scaling of mean edge length and second degree moment in syntactic dependency trees using data from stanford university to identify the similarity and differences among 30 different globally spoken languages. This statistical analysis is considered of high importance for identifying sentence structures in languages with similar attributes.
- **Sobolev Institute of Mathematics, (SB RAS)** — Novosibirsk, Russian Federation
  Research Intern (Supervisor: Professor Vladimir Berikov, NSU and SB RAS) — *May, 2017 - August, 2017*
  - **Hyperspectral Image Analysis**: Hyperspectral images are the satellite images which needs to be processed for understanding the spatial features of a geographic terrain. Our work was focused on developing novel semi-supervised classification algorithms to accurately identify geospatial features.
- **Indian Institute of Information Technology** — Allahabad, India
  Algorithmic fairness Investigator (Instructor: Prof. Suresh Venkatsubramanian) — *November, 2016 - January 2017*
  - **FAT ML**: Attended the extensive one month Course on Fairness, Accountability and Transparency in Machine Learning under the instructor Prof. Suresh Venkatsubramanian.
  - **Fairness Aware Statistical Learning**: Implemented codes in Python for fairness aware decision tree and naive bayes algorithm. Prepared presentation on the theme of ***fairness aware statistical learning through regularization***.
- **Indian Institute of Information Technology** — Allahabad, India
  Machine Learning Intern (Superviser: Dr. Sonali Agarwal) — *November, 2016 - January 2017*
  - **Apache Spark**: This work was based on development of a distributed feature selection algorithm using spearman rank correlation technique in Apache Spark. We used the PySpark implementation.
- **National Institute of Science Education and Research (NISER)** — Bhubaneshwar, India
  Research Internship and Training Program in Mathematics (TPM-2016) — *May, 2016 - June, 2016*
  - **Training**: Participated in the national level workshop for advanced mathematical training by standing in top 1 % for selection. Attended lectures for improving fundamentals in linear algebra, real analysis, group theory and combinatorics.
  - **Research**: Covered extensive literature review of the group theoretical methods in machine learning such as Hilbert Space Learning Algorithms along with discussion on Invariance.

SALIENT PROJECTS

This list only contains the details of the few recent most important projects. For a detailed information, kindly visit the github repository www.github.com/ankitbit

- **Eigendecomposition of Facial Imagery for Gender Indentification** Barcelona, Spain
  Guide: Professor Alexander Perera, ESAII ,UPC *January, 2018 - March, 2018*
  - ***Gender Prediction***: The eigendecomposition of facial imagery is a novel technique which is easy to implement, fast, scalable and efficient for automatic gender identification. Our work was implemented in Python and we're now focusing on how to use the technique for specific applications in homeland security for passenger screening
- **Modelling Lifetime of Havells Bulbs using Bayesian Methods** Barcelona, Spain
  Guide: Professor Xavier Puig, UPC *February, 2017 - Present*
  - ***Exponential Model for observed Failure Rates + Hierarchical Model + BUGS***: This is an ongoing project which corresponds to an industrial assignment in which we are trying to model the observed failure rates using different approaches. Such a model is considered of utmost importance in reliability engineering where this kind of model helps in product planning and improvement.
- **Modelling Conflict Time Series for Counter-Insurgency Trend Prediction** Barcelona, Spain
  Guide: Professor Josep A. Sanchez, Department of Statistics and Operations Research, UPC *March,2017 - Present*
  - ***ARIMA Modelling + Intervention Analysis + Neural Networks***: This is an ongoing project that attempts to create a mathematical framework using artificial neural networks and time series for forecasting trends in insurgency which can be used for effective policy making. The initiative is highly appreciated since beginning because of it's ability to explain the military expenditure, manpower planning and operational
- **Project Qualmet: Package for quality metric in Community Detection** Barcelona, Spain
  Guide: Professor Ramon Carrer-i-Ferrera, LARCA, UPC *October, 2017 - November, 2017*
  - ***Development of R Package***: This project corresponds to the design and development of a package in R that can efficiently compute the quality metrics such as triangular partition ratio, internal density, expansion, cut ratio, conductance, flake's out degree fraction and modularity.
- **Scaling of Mean Edge Length in Syntactic Dependency Trees** Barcelona, Spain
  Guide: Professor Ramon Carrer-i-Ferrera, LARCA, UPC *November, 2017 - January, 2018*
  - ***Non-Linear Models for Syntactic Dependency Trees***: We attempted to fit an ensemble of non-linear models to data of collections of syntactic dependency trees from 30 different languages. In a syntactic dependency tree, the vertices are the words (tokens) of a sentence and links indicate syntactic dependencies.
- **Using Cluster Ensembles for Kernel Based Semi-Supervised Learning** Russian Federaion
  Guide: Professor Vladimir Berikov, Russian Academy of Sciences (SB RAS) *May, 2017 - August, 2017*
  - ***Co-association matrix as Kernel***: This project aimed at solving the semi-supervised learning problem by using cluster ensembles obtained by running a set of clustering algorithms a finite number of times and using the individual partitions obtained to create a co-association matrix which can be used for classification of the test data points.
- **Multidimensional Scaling for Identifying Low Resolution Satellite Images** Barcelona, Spain
  Guide: Professor S.K. Jain, Birla Institute of Technology *February, 2015 - May, 2015*
  - ***Geospatial Intelligence***: The idea of this project was to investigate a set of low resolution satellite images with high resolution images to identify spatial features. The proposed method simultaneously embeds the low-resolution probe images and the high-resolution gallery images in a common space such that the distance between them in the transformed space approximates the distance had both the images been of high resolution. The two mappings are learned simultaneously from high-resolution training images using an iterative majorization algorithm.

PROGRAMMING SKILLS

- **Languages**: R, Python, C++, MySQL, Postgresql,Java **Technologies**: Apache Spark, Tableau, Git, LaTeX, MATLAB

ACADEMIC ACHIEVEMENTS

- **Best Research Paper Award in IEEE's SIBIRCON-2017** — Russian Federation
  Merit certificate for standing in top 1 percent of the school consecutively for 5 years — *September, 2017*
- **Qualified for the Joint Engineering Entrance (JEE) Examination-Main** — Allahabad, India
  Qualified the national level JEE Main by standing among the top 2 % of 1.8 million candidates — *July, 2014*
- **Qualified for the National Defence Academy and Naval Academy Examination** — Allahabad, India
  Qualified by standing among top 1 % of 1 million candidates at the national level examination — *December, 2013*
- **Certificate for Overall Outstanding Academic Performance by KVS,** — Allahabad, India
  Merit certificate for standing in top 1 percent of the school consecutively for 5 years — *January, 2012*
- **Innovation in Science Pursuit for Inspired Research (INSPIRE) Award** — Allahabad, India
  Department of Science of Technology (DST), Government of India — *November, 2011*
- **Certificate of Merit by Central Board of Secondary Education (CBSE)** — Allahabad, India
  Merit certificate for standing in top 1 percent of the National Level Examination conducted CBSE — *August, 2011*
- **Letter of Recognition by Union Minister of Human Resource Development** — Allahabad, India
  Ministery of Human Resource Development, Government of India — *July, 2011*

PUBLICATIONS

- Berikov V., Karaev N., Tewari A. Semi-Supervised Classification with Cluster Ensemble — Proceedings of 2017 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). Novosibirsk Akademgorodok, Russia, 18-22 Sep 2017. P. 245-250.
- Pestunov I, Berikov V., Karaev N., Tewari A. Recognition of Hyperspectral Images using Ensemble Clustering and Semi-Supervised learning — Proceedings of the All-Russian Conference (August 29-31, 2017). Novosibirsk, 2017. 323 p. ISBN 978-5-905569-08-1

RELEVANT COURSEWORK

- Statistical Learning Theory, Bayesian Data Analysis, Time Series Analysis, Social Network Analysis, Multivariate Analysis, Advanced Statistical Inference, Linear and Generalized Linear Models, Statistical Computing, Statistical Programming with Databases, Probability and Stochastic Processes, Machine Learning etc.

LEADERSHIP

- **National Cadet Corps (NCC)** — Ranchi, India
  Officer of world's Largest Youth Organization under Ministry of Defence, Government of India — *Aug 2014 - July 2017*
  - **Cadet Engineer**: Cadet Engineer participates in regular camps for military training along with combat troops to provide them engineering support for the creation, maintenance to demolition of military infrastrucure
  - **Cadet Volunteer**: The International Day of Yoga is celebrated every year on 21st June. Our role was to ensure every possible assistance to participants and organizers ranging from logistics, security to managerial aspects
- **Association for Computing Machinery (ACM)** — Mesra, India
  Student Member and Program Volunteer — *September, 2014 - Present*
  - **Student event organization and administration**: Our role was to provide technical and administrative support during the organization of various ACM India sponsored activities ranging from coding contests to workshops and tutorials
- **Computer Society of India (CSI)** — New Delhi, India
  Regional Technology Volunteer — *July, 2015 - August, 2017*
  - **E-Governance Awareness**: The role of Regional Technology Volunteer was to spread awareness about the e-governance initiatives of the government for the public good by means of organizing talks, workshops etc.
- **Social Coder Initiative**
  Coding Volunteer — *January,2017 - Present*
  - **Code volunteering**: I have been participating in the program along the lines of United nations Online Volunteering to support the basic informatics need of NGO's, human rights organizations in developing nations