

# Abhishek Ajay

[abhishek.ajay@columbia.edu](mailto:abhishek.ajay@columbia.edu) | 347-841-4516 | [LinkedIn](#) | [Github](#)

718 West 171st, New York, NY, 10032

## EDUCATION

### Columbia University, Mailman School of Public Health

New York, NY

*Master of Science (MS) in Biostatistics*

*Expected May'20*

**Relevant Coursework:** Topics in Statistical Learning and Data Mining, Analysis of Longitudinal Data, Statistical Machine Learning, Statistical Inference, Introduction to Applied Statistics, Advanced Statistical Methods, Statistical Genetics Modeling

### Indian Institute of Technology, Guwahati

Assam, India

*Bachelor of Technology in Biotechnology*

*July'14 - May'18*

**Thesis:** Analysis of transcriptomics data of a natural *Daphnia magna* population under varying predation pressure.

## SKILLS

- **Tools:** R (glmnet, caret, nnet, pscl, dplyr, tidyverse, ggplot, broom, shiny), Python<sup>1</sup>, SQL<sup>1</sup>, MS Excel (pivot tables), MS Word, MS PowerPoint
- **OS:** Windows, MacOS

## PROJECTS (COLUMBIA UNIVERSITY)

### Communicating Narrative Concerns Entered by Registered Nurses (CONCERN) in Psychiatry Ward

*Oct'19 - Present*

- Performing analytics of existing nursing documentation structure within six years of electronic health record (EHR) data obtained from the psychiatry wards at New York-Presbyterian Allen and Milstein Hospital.
- Building a back-end model for a clinical decision support system to predict patient deterioration (subjection of patients to Seclusion or Mechanical Restraint.) in the next 24 hours.

### Santander Customer Transaction Prediction (Kaggle Competition)

*Nov'19 - Dec'19*

- Analyzed high-dimensional anonymized data with 200 predictors using Logistic Regression, Elastic Net, Random Forest and XGBoost.
- Performed analysis on down-sampled data using the *caret* package.
- XGBoost gave the highest test AUC (0.888) using the down-sampled data.

### Forecasting Stock Price Direction using Clustering Algorithms

*May'19*

- Computed prediction models using Regularized Logistic Regression, K-Nearest Neighbor, Naive Bayes, Support Vector Machines (Radial and Linear Kernel), Random Forest (RF) and Tree based classification, etc. to predict whether the day's closing price will be higher or lower than the day before.
- Analyzed 5 years of historical daily price attributes of Amazon.com, Inc. taken from Yahoo Finance and found the RF Classification model to be having the highest Cross Validation accuracy.

## EXPERIENCE

### Columbia University, School of Nursing

New York, NY

*Program Assistant*

*March'19 - Present*

- Collaborate with a two person team in managing the school's global initiative, sending out nearly 100 students to 17 different international sites.
- Transformed the whole application process and document processing by migrating from paper-based forms to electronic forms (DocuSign).
- Carry out information tracking on MS Excel, report making on MS PowerPoint for the school's leadership and survey design on Qualtrics.

### Columbia University, Mailman School of Public Health

New York, NY

*Teaching Assistant for Probability*

*Sept'19 - Dec'19*

- Collaborated with a six-person team in managing weekly office hours, assignments, quizzes and exams for the graduate-level Probability course, P8104 with 120 registered students.

### Katholieke Universiteit Leuven

Leuven, Belgium

*Visiting Scholar (Intern), Department of Biology*

*May'17 - July'17 & Dec'17 - Jan'18*

- Analyzed high throughput transcriptomics (RNA-seq) data of a natural *Daphnia magna* population with around 30000 genes to understand the variation in gene expression level with varying predation pressure.
- Performed Concordance Analysis, Principal Component Analysis (PCA) for sample quality control.
- Produced visualizations using ggplot2 in R to understand the variation in gene expression over time using reaction norms.

## ADDITIONAL EXPERIENCES

**Student Mentor-** Student Welfare Board, IIT-Guwahati

*Sept'16 - April'17*

**Chief Organizer-** TechExpo, Techniche (Annual Technical Festival), IIT-Guwahati

*Oct'15 - Sept'16*

**Head-** Industrial Conclave and Exhibitions, Techniche, IIT-Guwahati

*Oct'15 - Sept'16*

**Member-** Montage, The Photography Club of IIT-Guwahati

*July'14 - May'18*

<sup>1</sup>On-going/Basic Proficiency