Zarni Htet
155 Claremont Ave Rm 745, New York, NY 10027 | 646.704.6716 | zh938@nyu.edu

February 1, 2017

National Capital Region
Biocomplexity Institute of Virginia Tech (0379)
900 N. Glebe Road
Arlington, VA 22203

Dear Ms. Lyman/Hiring Official**,**

My name is Zarni Htet and I am writing to apply for the Data Science for Public Good Fellowship at Virginia Tech's Social & Decision Analytics Laboratory. I am a first-year Master's student at NYU Applied Statistics for Social Science Research Program. Previously, I was an IT Consultant for 2 years and studied Economics and Computer Science for my Bachelor's. I am confident that my past experiences and longstanding passion for social good make me a perfect candidate for this fellowship.

My past work experiences have prepared me to thrive in an interdisciplinary and culturally diverse environment. At Prescio Consulting, I led a team of programmers, web designer and data collectors across different time zones to build a Bitcoin merchant directory while meeting the demands of multiple stakeholders within the company. I was highly motivated with the work due to the potential of the Bitcoin ecosystem to make a positive social impact towards remittance in developing countries and lowering transaction costs significantly. It was the same motivation during my earlier work at Grinnell where we were contributing towards the project to make programming more accessible for different people. This drive led me to pick NYU Applied Statistics for Social Science program over the more corporate-oriented Columbia Data Science program so that I can work on projects that have meaningful social impacts.

I am passionate about using data to solve societal problems. I worked on two projects from Driven Data, a data platform for social good. The first was a model to predict water pump failures in Tanzania. I used Python and pre-packaged machine learning algorithms such as logistic regression, decision tree learning and regularization techniques to improve prediction accuracy scores. The second project predicted blood donations for mobile blood collection units in Taiwan. This was a shorter project, but it required me to write a cross-validation function and a cost helper function for logistic regression. The goal of both projects was optimal resource allocation by public services, and I will be taking it a step further with my current project that will last this semester and possibly extend beyond it. I am taking on the project to explore the New York subway system using GTFS data from the MTA, weather data from Weather Underground, and additional socioeconomic data. The project is more challenging than my previous projects since it requires me to curate my own datasets and frame my own research questions.

The data for public good fellowship is the perfect next step for my data science career in public good. The Lab provides opportunities to tackle topics in direct consultation with local policy makers and government agencies. It promotes the perfect environment for growth through collaborating with interdisciplinary team of peers, mentors, and visiting scholars. I am really looking forward to the next research topic to engage and make a difference in the local communities. I am confident that my coursework in Missing Data with use of multiple imputation, Statistical Inference, and Multilevel Models along with my previous professional experiences working with cross-functional team have prepared me well for this fellowship program.

Sincerely,
Zarni Htet

Zarni Htet

155 Claremont Avenue Rm 745, New York, NY, 10027 ~ zh938@nyu.edu ~ 646-704-6716

EDUCATION

New York University – *MS candidate in Applied Statistics*, **3.94/4.00** May 2018

- Fall Coursework: Probability Theory, Large Databases in Research, Practicum in R
- Spring Coursework: Statistical Inference, Missing Data, Multi-Level models, Classification & Clustering
- Ongoing Projects: Blood Donation Prediction: http://bit.ly/2kTrA50 , NYC Subway Delays Exploration

Columbia University – *Graduate Foundations Program, Audits* May 2016

- Coursework: Econometrics, Applied Python for Engineers, Linear Algebra, Stats with Calculus
- R psets: http://bit.ly/1TZ6qMD, Python: KNN Classifier, Percolation with Numpy: http://bit.ly/1r3pEoG

Cornell College – *BA in Economics and Computer Science,* ***3.82/4.00*** May 2013

- Don Cell Award (Senior Economics Award), Omicron Delta Epsilon,Calculus Math Award
- Coursework: Intermediate Microeconomics, Databases, Linear Algebra, Differential Equations

EXPERIENCE

General Assembly – *Data Science Bootcamp Student* Aug 2015 – Oct 2015

Programming & Data Bootcamp, Washington, D.C.

Final project code & write-up: http://bit.ly/1n0O00P

- Achieved 77% Classification accuracy with AUC score of 85% in predicting water pump functionality for pumps in Tanzania with data (~50,000 x 39) from Ministry of Water
- Elaborated concepts & tools used such as regularization, decision tree, pipeline, cross-validation and feature engineering
- Certified and completed coursework in KNN, Logistic Regression, Regularization, Decision Trees, Random Forest and K-Means

Prescio Consulting – *IT Consultant* Aug 2013 – Aug 2015

Financial & IT Consultancy, Casa Grande, AZ.

- Led a team of developers to build an interactive Bitcoin merchant directory: Leptabit from wireframing, feature development, 30 table database schema design, high-level optimization to testing and deployment
- Managed the later stages of alpha stage Bitcoin wallet in feature development and testing
- Initiated application deployments to AWS EC2 web stack from cost analysis to pilot runs
- Contributed in the design of 33 table database schema on Client's Land rights web application
- Led a data scraping team to fill our marketing CRM (~ 500 x 15 - 2 months) database & Bitcoin merchants database (~ 10,000 x 35, rows x columns)

Zebra Zapps, Allen Interaction – *Software Engineering Intern* Aug 2012 – Oct 2012

Software Development, Mendota Heights, MN.

- Fixed front-end side bugs (html, css, jquery) of the website on Pure MVC framework using firebug tool
- Automated back-end processes via shell scripting
- Developed multi-threaded Flickr image downloader

Grinnell College – *Research Student* Jun 2012 – Aug 2012

Grinnell Computer Science Lab, Grinnell, IA.

Paper : http://bit.ly/1PvjW4t

- Contributed in developing the Client API for Dr.Racket using DBus to communicate with scriptable interfaces such as GIMP to manipulate images with Scheme for intro to computing students
- Presented at SIGCSE 2013 Conference as part of a larger poster presentation (March 2013)

The Cornellian – *Business Manager* Apr 2011 – May 2012

College Newspaper, Mount Vernon, IA.

- Administered the budget for the second biggest student organization on campus
- Raised external funds through securing local advertisements and coordinating the web team
- Initiated the software switch of the Editors' team onto Google docs for a more streamlined editing process

TECHNICAL SKILLS

R, Stata, Python, SQL, MySQL, RubyOnRails, Java,C, Linux (Ubuntu, Redhat, Debian), Google & Microsoft Productivity Suite.

Github: https://github.com/ZarniHtet13?tab=repositories