528 Riverside Drive,
Apt. 3B, New York,
NY 10027

AMEYA KARNAD

(646)492-0972
ameya.karnad@columbia.edu
http://ameyakarnad.github.io

EDUCATION

Columbia University New York, NY
Master of Science, Data Science, GPA: 3.71/4 Aug 2018 - Dec 2019 (Expected)
Courses: Algorithms, Exploratory Data Analysis & Visualization, Personalization, Probability and Statistics, Machine Learning, Statistical Inference & Modeling, Applied ML, Data Science & Public Policy
This Fall: Analysis of Social Networks and Crowds, Applied Deep learning (Audit), Capstone Project (Bloomberg)

Visvesvaraya Technological University Hubli, IN
B. V. Bhoomaraddi College of Engineering and Technology Jul 2012 – Jun 2016
Bachelor of Engineering, Information Science, GPA: 9.42/10 (Silver Medalist)

SKILLS

Languages and technologies: Python (pandas, numpy, scikit-learn, nltk, tensorflow, networkx, surprise), R (ggplot2, dplyr, data.table, Shiny), JavaScript (SVG, D3), SQL, MongoDB, Docker, Kubernetes

PROFESSIONAL EXPERIENCE

Data Science Researcher – *EdLab, Columbia University* (Columbia Data Science Scholar's Program) Jun 2019 - Present
Technologies: *Apache PredictionIO, MongoDB, EZProxy, docker*
- Development and research on Search and recommendation Systems, Social network analysis and Topic modelling
- Currently working on Auto-tagging of documents and designing a package for research metadata retrieval

Teaching Assistant – Applied Data Science, *Columbia University* *New York, US*
Awards: *"Excellence in Course Assistantship" for the course. Link* Jan 2019 – May 2019
- Guided students in the use of Data preprocessing, Feature engineering, machine learning, and Communicating results in various fields like Education, Health Care, Marketing, Advertising and Social Media using R

Software Engineer - Micro Focus (spun-off from Hewlett Packard Enterprise) Bangalore, IN
Awards: *2nd place in a site-wide Hackathon* Technologies: *Kubernetes, Docker, python* Sept 2016 – Jun 2018
- Worked with Agile teams to design, test and automate REST APIs for providing analytics insights into software security compliance and risk data for a datacenter automation Software.

SELECT DATA SCIENCE PROJECTS

Editorial Classifier, Bloomberg – Columbia University Ongoing
Languages*: Python* Applications: *NLP, Classification*
- Building classifiers to identify Editorial content in news and uncovering trends over time across news sources for Bloomberg.
- Anticipated developments include working with Bert Algorithm, Pytorch, Deep learning and Topics over time

Educational Recommendation System, Edlab
Languages*: Python (nltk, predictionIO), SQL, MongoDB* Applications: *Recommendation systems, Topic modelling*
- Developed a multi-platform database pipeline for physical and digital user behavior, and research content using AWS.
- Built and evaluated a hybrid recommendation system on multiple digital learning platforms

Analysis of debates in Ugandan parliament, UPeßßßßßßnn
Languages*: Python (nltk), R (ggplot2)* Applications: Data Preprocessing, *NLP*
- Analyzed debates in the Ugandan parliament to identify speeches made by MPs on the discovery of oil in Uganda in 2007
- Project part of a paper on the effects of Uganda's Oil Discovery in 2007 on voting behavior in 2011 by Prof. Guy Grossman

Does Climate Change and Natural Resource Storage Cause Conflict? *Columbia University*
Languages*: Python (Scikit-learn), R (Shiny, data.table)* Applications: *Regression, SVM, Decision Trees, Dashboard*
- Found correlations between factors such as Rainfall, Temperature, food prices and Bio-mass and Conflicts in Senegal. Link
- Built a Dashboard to be used by policymakers to take effective decisions to avoid natural resources shortages and conflicts.

Mini Projects in Data Science and Public Policy, *Columbia University*
Languages*: Python (Scikit-learn, nltk), R (ggplot2, dplyr)* Applications: *NLP, Visualizations, SVM, Decision Trees*
- "Russian Twitter trolls and 2016 U. S. Election" – analyzing FiveThirtyEight's 3 million tweets dataset Link
- "Evaluating Northpointe Inc's Compass system recidivism predictor" Link
- "Predicting Economic factors related to Poverty using Night Lights dataset" Link

How America Flies? - Creating Insightful Visualizations on Airline Performance Data, *Columbia University*

Languages*: R (ggplot2, dplyr), JavaScript (D3, SVG)* Applications: *Data visualizations, Exploratory data analysis*

- Conducted Exploratory data analysis and found insights on Airline On-time performance data Project Link
- Designed a Visualization tool in JavaScript and D3 to find flight delays between top airports Tool Link

Beer Personalization and Recommendation System, *Columbia University*

Languages*: Python (surprise, nltk, scikit-learn)* Models: *Matrix factorization(SVD, NMF), Content based, KNN*

- Developed a comparative analysis of the models mentioned above with regards to accuracy and time.
- Tackled recommendation concerns such as cold- start, variety and serendipity Link

RESEARCH EXPERIENCE

- **Karnad, A.**; Yadappanavar, S.; Hiremath P. G. S., Evaluation and validation of problem solving and thinking skills based on student academic performance. *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology 2017*, Bangalore, IN, Link
- **Karnad, A.**; Chen, Y.; Chae, H.; Natriello, G., The Application of Social Network Analysis to support collaborative e-learning, *American Educational Research Association 2020, San Francisco, CA*. Under Review

HOBBIES

Debating, reading about history and current affairs, creating educational YouTube Videos