

1300 Fayette St.  
Apt. 52, Conshohocken,  
PA, 19428

# AMEYA KARNAD

(646)492-0972  
ak4251@columbia.edu  
<http://ameyakarnad.github.io>

## EDUCATION

### Columbia University

Master of Science, Data Science, GPA: 3.77/4

Courses: Algorithms, Exploratory Data Analysis & Visualization, Personalization, Probability and Statistics, Machine Learning, Statistical Inference & Modeling, Applied ML, Data Science & Public Policy, Analysis of Social Networks and Crowds, Applied Deep learning (Audit), Capstone Project (Bloomberg - Editorial Classifier)

New York, NY

Aug 2018 - Dec 2019

### Visvesvaraya Technological University

B. V. Bhoomaraddi College of Engineering and Technology

Bachelor of Engineering, Information Science, GPA: 9.42/10 (Silver Medalist)

Hubli, IND

Jul 2012 - Jun 2016

## SKILLS

**Languages and Technologies:** Python (pandas, numpy, scikit-learn, nltk, tensorflow, networkx, genism, spacy), R (ggplot2, dplyr, data.table, R-shiny), JavaScript (SVG, D3), SQL, MongoDB, Docker, Kubernetes, AWS

## PROFESSIONAL EXPERIENCE

### Senior Software Engineer - Oracle

**Technologies:** Python, JAVA

- Working on the Oracle Transportation Management (OTM) Software.
- Building Machine learning, and Operational Research Algorithms and Designing Dashboards for OTM.

Conshohocken, PA

Mar 2020 - Present

### Data Science Research Engineer - EdLab, Columbia University (Data Science Scholar's Program)

**Technologies:** Apache PredictionIO, MongoDB, EZProxy, Docker

- Development and research on Search and Recommendation systems, Social Network Analysis, and Topic Modeling.
- Worked on building Data pipelines for End-to-End Data Science Applications.

New York, NY

Jun 2019 - Dec 2019

### Teaching Assistant - Applied Data Science, Columbia University

**Awards:** "Excellence in Course Assistantship" for the course. [Link](#)

- Guided students in the use of Data preprocessing, Feature engineering, and Machine learning using R.

New York, NY

Jan 2019 - May 2019

### Software Engineer - Micro Focus (spun-off from Hewlett Packard Enterprise)

**Awards:** 2<sup>nd</sup> place in a site-wide Hackathon

**Technologies:** Kubernetes, Docker, Python, SQL

- Worked with agile teams to design, test and automate REST APIs for providing analytics insights into software security compliance and risk data for a data-center automation product.

Bangalore, IND

Sept 2016 - Jun 2018

## SELECT DATA SCIENCE PROJECTS ([GITHUB LINK](#))

### Editorial Classifier, Bloomberg - Columbia University

**Languages:** Python(Pytorch, TensorFlow, spacy), R (ggplot)

**Applications:** NLP, Classification, Deep Learning

- Worked with unstructured news sources to build classifiers to identify editorial content for Bloomberg Terminal.
- Developed traditional classification algorithms with text-based features like n-grams, Parts of speech, Named entities etc.
- Trained Natural Language Processing based Bidirectional LSTM, BERT, and XLNET Models for classification [Link](#).

### Educational Recommendation System, Edlab

**Languages:** Python (nltk, predictionIO), SQL, MongoDB, AWS

**Applications:** Recommendation systems

- Developed a multi-platform database pipeline for physical and digital user behavior, and research content using AWS.
- Built a hybrid recommendation system on multiple digital learning platforms and evaluated it using AB testing.

### 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

- Developed Machine Learning models to find correlations between Natural, and Economic factors and Conflicts in Senegal [Link](#).
- Built a Dashboard to be used by policymakers to take effective decisions to avoid natural resource shortages and conflicts.

### 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](#).

## OTHER ACHIEVEMENTS

- Finalist at Annual Columbia Data Science Society (CDSS) Hackathon held in September 2019
- Conducted workshops on Social Network Analysis and Data Science in Education