

TUSHAR AHUJA

A software engineer with a passion for creating delightful products.

ta1302@nyu.edu | www.tusharahuja.in | www.linkedin.com/in/tushar-ahuja | www.github.com/holmes0078

EDUCATION

New York University Tandon School of Engineering

Master of Science in Computer Science, 3.5/4.0

May 2018

RELATED COURSEWORK: Artificial Intelligence, Machine Learning, Programming for Big Data, Web Search Engines

University of Pune, India

Aug 2016

Bachelor of Engineering in Computer Engineering, 3.8/4.0

RELATED COURSEWORK: Software Design Methodologies and Testing, Data Structures and Problem Solving, Software Engineering, Operating System Design, Discrete Mathematics

EXPERIENCE

Technology Summer Analyst – Barclays, New York

Jun–Aug 2017

Risk and Analytics - Barclays International

- Created a Proof of Concept for a tool using Apache Spark that automates Mapping and Running a Quantitative Calculation reducing turnaround time for the client from 3 days to a less than 3 hours
- Implemented the tool as a full stack Web App with Scala Play as Backend and Angular JS on Apache Zeppelin as Front-End
- Designed a Web Dashboard which provided the tool access in a multi-tenant environment leading to a reduction in the number of meetings to almost zero for the team dealing with the process

Teaching Assistant – New York University, Brooklyn

Sep–Dec 2016

Applied Data Science – NYU Urban Science

- Assisted the faculty in the ongoing development and design of Data Science lab assignments in Python in a manner supporting a research-led and scholarly approach to student learning
- Graded assignments, and taught extra classes on Linear Regression
- Held weekly office hours to provide tutoring, counseling, or assistance to students in need

ACADEMIC PROJECTS

Small Search Engine, Web Search Engines (<https://goo.gl/AdXL4V>)

Sep–Dec 2017

- Implemented a multi-threaded Web Crawler to crawl 8 million web pages at a rate of 500 URLs per second
- Built an inverted index from 8 million web pages, compressed the index using variable byte encoding leading to 30 percent smaller index size on disk
- Designed a DAAT query processor returning top 10 pages for 3-word queries under half a second

Detecting Twitter Bots, Machine Learning (<https://goo.gl/4MUvQv>)

Jan–Mar 2017

- Implemented ensemble Machine Learning Models to predict whether a Twitter user is a bot or not
- Compared the generalization error produced by various models and chose Random Forest Trees for classification
- Tuned the model parameters and obtained an accuracy of 95% in predicting bots securing the 3rd rank in class Kaggle competition

Citi Bike Descriptive Analysis, Big Data (<https://goo.gl/XjffR2>)

Sep–Dec 2016

- Performed Descriptive Data Analysis of Citi Bike Data sizing 40 Gb using Apache Spark and Python
- Created an R-tree index on geo-location data reducing analysis time from 30 minutes to under 3 minutes
- Built a story map to visualize the Analysis using ArcGIS (<http://arcg.is/2qzEJzA>)

SKILLS

Programming (Over 5000 Lines):

C/C++, Java, Python

Programming (Over 1000 Lines):

JavaScript, PHP, Scala

Web Development:

HTML, CSS, JavaScript

Tools:

Intellij, Visual Studio, Vim, Amazon AWS, Microsoft Azure, Docker

Familiar:

React, Polymer, Node.js, MySQL, Git