

Shoujun Gu, Ph.D

shoujun.gu@gmail.com | 210-315-2598 | Rockville, Maryland

<https://www.linkedin.com/in/shoujun-gu-bb624643/> | <https://github.com/shoujungu>

Expertise

- Data Mining and Analysis Tools
- Machine Learning Algorithms
- Biological and Genomic Data Analysis
- Data Munging and Visualization Tools
- Predictive Modeling

Related Projects

- Team member of The Cancer Genome Atlas (TCGA) data mining and analysis project.
The manuscript of this project entitled 'A Pan-Cancer Atlas of Genomic, Epigenomic and Transcriptomic Alterations in the TGF- β Pathway', is under reviewing¹
- Build model to predict impact factor range based on the abstract of a research manuscript.
Details available at: https://github.com/shoujungu/IF_Predict
- Data mining and analysis on all healthcare publications (Year: 2017) in PubMed database.
Details available at: https://github.com/shoujungu/2017_Pubmed_Analysis

Education

The George Washington University, Washington, DC

Oct 2017 – Apr 2018

Data Analytics Boot Camp

An intensive 24-week long boot camp dedicated to Data Mining and Analytics Skills on a variety of real-world problems.

UT Health Science Center at San Antonio

Aug 2005 – May 2012

Ph.D. Department of Biochemistry

Achievements:

- Oral presentation in American Society for Bone and Mineral Research 2008 Annual Meeting
- First Author publication on *The Journal of Biological Chemistry*

Working Experience

The George Washington University, Washington, DC

Oct 2015 -

Postdoctoral Researcher

Data mining on various biological databases and next-generation sequencing data analysis.

Achievements:

- Multiple publications on various peer-review journals

Of Note

Data and Analytics Tools: Python, MySQL, Excel/VBA, Tableau, Linux Bash

Publications: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3375531/>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4839765/>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4960540/>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5724803/>