Kwangmin Kim

Data Scientist/ Data Analyst

Email: kmink3225@gmail.com
Website: kmink3225.netlify.app
Linkedin: https://www.linkedin.com/in/kwangmin-kim-a5241b200/

ABOUT ME

Over 7 years of experience in data modeling, data pre-processing, data analytics, and statistical/ML modeling. As a product owner, I communicated with other departments and directly managed and completed projects. I was selected for the Platform Planning TF Team under the Strategic Planning Office and gained a high understanding of the platform. I also collaborated with the in-house patent center to invent technologies helpful for corporate assetization.

I am interested in gaining a mathematical understanding of algorithms and modeling. I communicate with the collaborating departments using detailed objective facts obtained through data analysis. I aim to work in a planned, systematic way.

SKILLS

Data Scinece	Database	etc
• R, Python	• SQlite	• Ubuntu, Powershell, Git/Github, Conda
• SAS	• Oracle-SQL	• Quarto, R markdown, Jupyter,

EXPERIENCE

2020.12 - Present

Seegene, Diagnosis IT General Research Institute, **Data Science Team**
Data Scientist / Data Analyst
- Algorithm DHF documentation and FDA verification & validation reporting.
- Diagnosis algorithms management and development.
- Device QC (Quality Control) algorithms management and development.
- Data analytic and statistical analysis.
- IP (Intellectual Property) planning and filed 9 patent inventions.

2018.12 - 2020.04

Columbia University Irving Medical Center,
Taub Institute for Research on Alzheimer's Disease and the Aging Brain
Research Statistician
- Clinical data analytics pipeline construction.
- Clinical data analytics: statistics, machine learning(ML), data mining, pathway analysis.
- ML and statistics consulting to medical doctors, epidemiologists, and neurologists.

EDUCATION

2017.08 - 2019.05 • Columbia University in the City of New York, Biostatistics(MS), Department Head Award
2015.08 - 2017.05 • Baruch College, The City University of New York, Mathematics(BA)
2006.03 - 2012.02 • Kangwon National University, Biochemistry(BS), Summa Cum Laude, Full Scholarship

PROJECTS

2022.07 - present

Algorithm DHF & FDA Verification & Validation Documentation
Seegene, Diagnosis IT General Research Institute
- Diagnosis algorithm DHF planning and documentation.
- FDA verification & validation report documentation.
- Designed algorithm testing and constructed statistical analysis pipelines for algorithm verification as the product manager.
- Dynamic documentation using Quarto, R, and python.

2021.12 - 2022.07

Platform Planning TF
Seegene, Strategy Planning Office & In-house Patent Center
- Platform strategy planning and intellectual property planning.
- 5/16 inventions have been filed(the remaining 11 are ongoing).
- Database system, statistics, and ML knowledge consultation for executives, planners, and attorneys.

2021.01 - 2021.09

Data-Driven Diagnosis Algorithm Development
Seegene, Strategy Planning Office & In-house Patent Center
- Planned and developed a data-driven signal processing algorithm.
- Reflected optical characteristics of diagnosis device.
- Reflected the unique technology of SG reagents.

2019.05 - 2020.04

Diagnosis Device Quality Control (QC) Platform Construction
Seegene, Diagnosis IT General Research Institute
- led all processes throughout the project as the product owner.
- Developed noise measurement algorithms.
- Collaborated with other departements to automate the two-stage QC process, visualize it, and reduce turnaround time by 11x.
- Classified device failures, human errors, and production line errors.

2018.12 - 2019.05

Long Life Family Study (LLFS) Project
Columbia University Irving Medical Center, Taub Institute
- Used statistics and ML to identify metabolic profiles significantly associated with dementia
- Analytics pipeline consturction: missing value analysis, statistical analysis, ML classification, and pathway analysis
- Using data mining, a strong confounder was discovered that had not been found by the research institute for 8 months

Analytic Project on Alzheimer's Disease and the Aging Brain
Columbia University in the City of New York, Biostatistics
- Selected as one of the top 3 out of about 100 graduate students in the annual research competition for master's graduate students
- Comparative study of optimal ML methods for Alzheimer's Disease and the Aging Brain and metabolomics

2015.01 - 2015.06

Using Tea Leaves, Heavy Metal Removal Algorithm Development
The City University of New York, Mathematics
- Developed a mechanistic model that reflects the adsorption process of heavy metals into tea leaves using differential equation and non-linear least square algorithm

AWARDS

Invention Compensation

R&D Division Excellence Award
Invention Compensation

Department Head Award
$1,000 Stipend
Job Offer

$1,000 Stipend