

Seojin Bang

PhD candidate in Computational Biology
School of Computer Science, Carnegie Mellon University
✉ seojinb@cs.cmu.edu | 🌐 SeojinBang | 📄 seojinb

RESEARCH INTEREST

My research interest is developing machine learning and deep learning approaches for disease diagnosis and characterization. In particular, I am interested in developing interpretable machine learning approaches for providing a rationale for biomedical decision system and extract human-understandable explanation from the model to support biomedical decision making. I am also interested in developing multimodal/multiview machine learning approaches for combining different types of biomedical data.

EDUCATION

Present	Carnegie Mellon University PhD candidate in Computational Biology	Advisor: Wei Wu
2015	University of Connecticut MS in Statistics	Advisor: Haim Bar
2013	Seoul National University , Republic of Korea MS in Statistics	Advisor: Taesung Park
2010	Sungkyunkwan University , Republic of Korea BS in Mathematical Education · BE in Statistics	

PUBLICATIONS/PREPRINTS

- [1] **Explaining a black-box using deep variational information bottleneck approach.**
preprint arXiv:1902.06918, 2019.
Bang, Seojin and Xie, Pengtao and Lee, Heewook and Wu, Wei and Xing, Eric.
- [2] **Multiple kernel k -means clustering using min-max optimization with l_2 regularization.**
preprint arXiv:1803.02458, 2018.
Bang, Seojin and Yu, Yaoliang and Wu, Wei.
- [3] **Phased-lstm based predictive model for ehra with complex missing patterns.**
preprint, 2019.
Bang, Seojin and Yang, Yang and Wang, Yuchuan.
- [4] **Ontoan: An ontology-based neural network model for patient need detection.**
preprint, 2019.
Jang, Hyeju and **Bang, Seojin**.
- [5] **Multiview cluster analysis identifies variable corticosteroid response phenotypes in severe asthma.**
American Journal of Respiratory and Critical Care Medicine, 2019.
Wu*, Wei and **Bang* (co-first)**, **Seojin** and Bleecker, Eugene and Castro, Mario and Denlinger, Loren and Erzurum, Serpil and Fahy, John and Fitzpatrick, Anne and Gaston, Ben and Hastie, Annette and Israel, Elliot and Jarjour, Nizar and Kerr, Sheena and Levy, Bruce Meyers, Deborah and Moore, Wendy and Peters, Michael and Phipatanakul, Wanda and Sorkness, Ronald and Wenzel, Sally.

- [6] **A mixture model to detect edges in sparse co-expression graphs.**
Biostatistics, under review, preprint arXiv:1804.01185, 2018.
 Bar, Haim and **Bang, Seojin**.
- [7] **Phosphoproteomic analysis of the amygdala response to adolescent glucocorticoid exposure reveals g-protein coupled receptor kinase 2 as a target for reducing motivation for alcohol.**
Proteomes, 6(4), 2018.
 Bertholomey, Megan L. and Stone, Kathryn and Lam, TuKiet T. and **Bang, Seojin** and Wu, Wei and Nairn, Angus C. and Taylor, Jane R. and Torregrossa, Mary M.
- [8] **Naïve bayes ensemble: A new approach to classifying unlabeled multi-class asthma subjects.**
 In *Bioinformatics and Biomedicine (BIBM), 2016 IEEE International Conference on*, pages 460–465. IEEE, 2016.
Bang, Seojin and Wu, Wei.
- [9] **Ethnic variability in the allelic distribution of pharmacogenes between korean and other populations.**
Pharmacogenetics and genomics, 22(12):829–836, 2012.
 Kim, In-Wha and Im Kim, Kyung and Chang, Hyeu-jin and Yeon, Bora and **Bang, Seojin** and Park, Taesung and Kwon, Ji-sun and Kim, Sangsoo and Oh, Jung Mi.
- [10] **Joint selection of snps for improving prediction in genome-wide association studies.**
 In *Bioinformatics and Biomedicine Workshops (BIBMW), 2012 IEEE International Conference on*, pages 852–858. IEEE, 2012.
Bang, Seojin and Kim, Yong-Gang and Park, Taesung.

RESEARCH EXPERIENCE

Interpretable Machine Learning

2018 – 2018

Research Intern advised by Dr. Pengtao Xie

Petuum

- Developed a system-agnostic interpretable machine learning approach to explain a decision made by a black box system.
- Build a pytorch-based software for the interpretable learning approach.
- Preprint: [1] (in progress)

Severe Asthma Research Program

2015 – 2019

Research Assistant advised by Dr. Wei Wu

Carnegie Mellon University

- Developed a multiple kernel k-means clustering approach to identify asthma subtypes by combining different types of clinical data.
- Build an R-package: **MKKC** for the multiple kernel k-means clustering approach.
- Publication: [2, 5, 8]

Dimensional Reduction Analysis of Ultra High-dimensional Bioinformatics Data

2012 – 2013

Research Scientist advised by Dr. Taesung Park

Seoul National University

- Developed a joint variable selection method in high-dimensional data using the elastic-net regularization technique.
- Publication: [10]

Complex Biomarker Analysis for Pancreatic Cancer Diagnosis Modeling

2012 – 2013

Research Scientist advised by Dr. Taesung Park

Seoul National University

- Identified complex biomarkers of miRNA, mRNA, and protein for pancreatic cancer diagnosis using a statistical approach.
- Investigated different subtypes of intraductal papillary mucinous neoplasm (IPMN) using longitudinal clinical data using a time dependent survival model.
- Developed a prognostic and prediction model with miRNA, mRNA, and protein markers using a statistical approach.

The Pharmacometric Study (PK/PD Modeling & Simulation) of Immune modulating Agents Utilizing Pharmacogenomics

2012

Research Assistant advised by Dr. Taesung Park

Seoul National University

- Examined differences in allele frequencies of pharmacogenes among populations using the size-modified index.
- Publication: [9]

WORKING EXPERIENCE

Jan 2016 – Present	Research Assistant for Prof. Wei Wu <i>Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA</i> Publication/Preprint: [5, 2, 7, 8]
May 2018 – Dec 2018	Research Intern in Artificial Intelligence and Machine Learning Solution Team <i>Petuum, Pittsburgh, PA</i> Preprint: [1]
Aug 2013 – Aug 2015	Research Assistant for Prof. Haim Bar <i>Department of Statistics, University of Connecticut, Storrs, CT</i> Publication: [6]
Feb 2013 – Aug 2013	Research Scientist in Bioinformatics and Biostatistics <i>BIBS at Seoul National University, Korea</i>
Aug 2011 – Feb 2013	Research Assistant for Prof. Taesung Park <i>Department of Statistics, Seoul National University, Korea</i> Publication: [9, 10]

SOFTWARES

Python	VIBI : python implementation of VIBI https://github.com/SeojinBang/VIBI
R-package	MKKC : multiple kernel k -means clustering on a multi-view data https://github.com/SeojinBang/MKKC
R-package	edgefinder : a method for recovering a gene networks structure from co-expression data. https://arxiv.org/abs/1804.01185
LaTeX template	TidyCV : simple and tidy LaTeX template for your curriculum vitae https://github.com/SeojinBang/TidyCV

TECHNICAL STRENGTHS

Computer Languages	Python, R, C++, MATLAB, Bash, HTML
Library & Tools	Pytorch, Tensorflow, Keras, Git, LaTeX, Excel

HONORS AND AWARDS

2018	The Center for Machine Learning and Health Fellowships in Digital Health full tuition and stipend for 12 months and research-related expenses (total \$75,200)	
2013	The Korean Statistical Society Paper Awards (3st Place)	
2012	The Korean Statistical Society Poster Awards (1st Place)	
2006 – 2010	National Science and Engineering Undergraduate Scholarship full tuition for 8 semesters an additional \$500 grant for a high GPA an additional \$500 grant for a high GPA	2006 – 2010 2009 2008

PROFESSIONAL SERVICE

2019	Reviewer , <i>IEEE Access</i>
2019	Reviewer , <i>IEEE Transactions on Neural Networks and Learning Systems</i>
2018	Admission Committee , <i>Computational Biology Department, Carnegie Mellon University</i>
2018	Reviewer , <i>The 9th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB)</i>
2012	Program Committee , <i>International Symposium on Statistical Genetics</i> , Rep of Korea
2012	Program Committee , <i>Microarray Analysis Workshop: Statistical Analysis using R language</i> , Korea
2011	Program Committee , <i>The 2011 Spring Conference of the Korean Statistical Society</i> , Korea

TEACHING EXPERIENCE

Carnegie Mellon University	Teaching Assistant
2018 Quantitative Cell and Molecular Biology Lab	
2017 Computational Methods for Proteogenomics and Metabolomics	
University of Connecticut	Teaching Assistant
2014 Mathematical Statistics	
2014 Introduction to Mathematical Statistics	
2013 Elementary Concepts of Statistics	
2013 Introduction to Statistics I and II	
2013 Statistical Methods	
Seoul National University, Korea	Teaching Assistant
2012 Statistics Laboratory	
2012 Regression and Analysis and Laboratory	
2011 Statistics	
Bongyoung Girls' Middle School, Korea	Student Teacher
2009 Middle School Mathematics	
Sungkyunkwan University, Korea	Teacher
2006 Alternative Elementary/Middle School Mathematics	