

# Sejoon Oh

Computational Biology Department  
Carnegie Mellon University  
5000 Forbes Ave GHC 7412, Pittsburgh, PA 15213  
Email: ohhenrie@cmu.edu

## RESEARCH INTERESTS

Data Mining, Machine Learning, Parallel and High-Performance Computing, Computational Biology

## EDUCATION

**Carnegie Mellon University**, Pittsburgh, PA

- Ph.D. Student in Computational Biology

Aug. 2018 – Present

**Seoul National University**, Seoul, Korea

- Bachelor of Science (B.S.) in Computer Science and Engineering
  - Overall GPA: 3.68 / 4.0, Major GPA: 3.67 / 4.0

Mar. 2012 – Aug. 2018

## PUBLICATIONS

### JOURNAL PAPERS

- [J4] Kijung Shin, **Sejoon Oh**, Jisu Kim, Bryan Hooi, and Christos Faloutsos, “Fast, Accurate and Provable Triangle Counting in Fully Dynamic Graph Streams”, **ACM Transactions on Knowledge Discovery from Data (TKDD)**, 2019 (**Under Revision**).
- [J3] **Sejoon Oh**, Namyong Park, Jun-Gi Jang, Lee Sael, and U Kang, “High-Performance Tucker Factorization on Heterogeneous Platforms”, **IEEE Transactions on Parallel and Distributed Systems (TPDS)**, 2019.
- [J2] Namyong Park, **Sejoon Oh**, and U Kang, “Fast and Scalable Method for Distributed Boolean Tensor Factorization”, **VLDB Journal**, 2019.
- [J1] **Sejoon Oh\***, Jungwoo Lee\*, and Lee Sael, “GIFT: Guided and Interpretable Factorization for Tensors with an Application to Large-Scale Multi-platform Cancer Analysis”, **Bioinformatics**, 2018 (\* these authors contributed equally to this work).

### CONFERENCE PAPERS

- [C1] **Sejoon Oh**, Namyong Park, Lee Sael, and U Kang, “Scalable Tucker Factorization for Sparse Tensors - Algorithms and Discoveries”, *IEEE International Conference on Data Engineering (ICDE 2018)*, Paris, France, Apr. 2018.
  - **Gold Prize Winner (1st in CS) from Samsung Humantech Paper Award**
  - **Best Undergraduate Thesis Award from Seoul National University**
- [C2] Namyong Park, **Sejoon Oh** and U Kang, “Fast and Scalable Distributed Boolean Tensor Factorization”, *IEEE International Conference on Data Engineering (ICDE 2017)*, San Diego, California, USA, Apr. 2017.

## RESEARCH EXPERIENCE

**Graduate Research Assistant**, Carnegie Mellon University

- Research area: machine learning for scRNA-seq and age prediction

Aug. 2018 – Present

**Undergraduate Research Intern**, Seoul National University

- Data Mining Lab. (Advised by Prof. U Kang)

July 2016 – May 2018

- Research area: tensor analysis, recommender system, and high-performance computing

## RESEARCH PROJECTS

- **ML-based Cell Type Prediction Using Single Cell RNA-seq Data**

Aug. 2018 – Dec. 2018

- Implemented and analyzed dimensionality reduction methods (PCA and NMF)
- Term project for CMU 10701 ‘Graduate Machine Learning’

- **Developing Big Data Engine Based on High-Performance Computing**

Jan. 2017 – May 2018

- Core developer of sparse matrix and tensor operations
- Funded by Korea Ministry of Science and ICT

- **Anomaly Detection Techniques on I/O Trace Time Series**

Mar. 2017 – June 2017

- Core developer of the project, cooperated with SK Telecom company

- **Accelerator Programming Winter School**

Feb. 2017

- Implemented convolutional neural network (CNN) on heterogeneous platforms

- **Deep Writing Algorithm Using Word-Level LSTM**

Sept. 2016 – Dec. 2016

- Term project for a class “Introduction to Machine Learning”

	<ul style="list-style-type: none"> <li>▪ <b>Personalized Recommender System via Coupled Matrix Factorizations</b> Aug. 2016 – Dec. 2016 <ul style="list-style-type: none"> <li>• Core developer of the project, cooperated with Hyundai card company</li> </ul> </li> </ul>
<b>AWARDS &amp; SCHOLARSHIPS</b>	<ul style="list-style-type: none"> <li>▪ <b>Best Thesis Award (among all CSE undergraduate students)</b> Aug. 2018 Awarded by Seoul National University, Korea</li> <li>▪ <b>Humantech Paper Award (Gold Prize, 1st in Computer Science)</b> Feb. 2018 Awarded by Samsung, Korea</li> <li>▪ <b>National Scholarship for Science and Engineering</b> Dec. 2017 Awarded by Ministry of Science and ICT, Korea</li> <li>▪ <b>Final Top-10 Winner</b> Feb. 2017 Awarded at Accelerator Programming Winter School (APWS), Korea</li> <li>▪ <b>The 5th Place Winner</b> Aug. 2016 Awarded at Samsung Collegiate Programming Cup (SCPC), Korea</li> <li>▪ <b>Merit-based Scholarship</b> Aug. 2012 Awarded by Seoul National University, Korea</li> <li>▪ <b>Superior Academic Performance Scholarship</b> Mar. 2012 Awarded by Seoul National University, Korea</li> <li>▪ <b>Silver Medalist of Asia-Pacific Informatics Olympiad</b> May 2011 Awarded at the 5th Asia-Pacific Informatics Olympiad (APIO), Iran</li> <li>▪ <b>Gold and Silver Medalist</b> July 2008 – July 2011 Awarded at Korea Olympiad in Informatics (KOI), Korea</li> <li>▪ <b>Candidate for International Olympiad in Informatics (IOI)</b> Aug. 2008 – Aug. 2010 Trained at IOI Summer and Winter School, Korea</li> </ul>
<b>PROFESSIONAL SERVICES</b>	<p><u>Journal Reviewer</u></p> <ul style="list-style-type: none"> <li>▪ European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2018; Guest Reviewer) 2018</li> </ul>
<b>PATENTS</b>	<p><u>KOREA</u></p> <p><b>Sejoon Oh</b>, Namyong Park, U Kang, “Apparatus for Supporting Multi-dimensional Data Analysis through Parallel Processing and Method for the Same”, Korean patent number: 10-2017-0158951 (filed on Nov. 2017).</p>
<b>WORK EXPERIENCE</b>	<p><b>Korean Augmentation To the United States Army (KATUSA)</b>, Seoul, Korea</p> <ul style="list-style-type: none"> <li>▪ Interpreter &amp; Administrative Assistant Oct. 2014 – July 2016 <ul style="list-style-type: none"> <li>• Mandatory military service, served as a sergeant</li> <li>• Received Army Commendation Medal by the U.S. Army brigade commander, acknowledging outstanding leadership</li> </ul> </li> </ul>
<b>RELEVANT COURSEWORK</b>	<ul style="list-style-type: none"> <li>▪ Graduate Artificial Intelligence (CMU - 15780) Spring 2019</li> <li>▪ Graduate Machine Learning (CMU - 10701) Fall 2018</li> <li>▪ Artificial Intelligence Spring 2018</li> <li>▪ Introduction to Linear Algebra Fall 2017</li> <li>▪ Advanced Topics in Algorithms (Graduate Coursework) Spring 2017</li> <li>▪ Introduction to Data Mining Spring 2017</li> <li>▪ Engineering Research Practice 1 Spring 2017</li> <li>▪ Introduction to Machine Learning Fall 2016</li> <li>▪ Database Spring 2014</li> <li>▪ Software Applications Spring 2014</li> </ul>
<b>TECHNICAL SKILLS</b>	<ul style="list-style-type: none"> <li>▪ C, Python, and OpenCL (Advanced)</li> <li>▪ Java, C++, and MATLAB (Experienced)</li> <li>▪ Scala, R, and CUDA (Intermediate)</li> </ul>
<b>LANGUAGES</b>	<ul style="list-style-type: none"> <li>▪ Korean: ILR Level 5 – Native proficiency</li> <li>▪ English: ILR Level 4 – Full professional proficiency <ul style="list-style-type: none"> <li>• TOEFL score: 108 (Reading: 28, Listening 29, Speaking 24, Writing 27) Aug. 2017</li> </ul> </li> </ul>