

Jun-Hwan Choi

choi.junhwan@gmail.com
(413) 219-6355

Objectives	Data Scientist with a background in computational Astrophysics.
Data Science Employment & Experiences	<p>Data Scientist at SparkCognition Nov. 2016 - Present</p> <ul style="list-style-type: none">- Building predictive machine learning model with time series data to predict future events using Random Forest, Decision Tree Boosting, and Artificial Neural Network Methods (including CNN and RNN). Unsupervised anomaly detection for future events using clustering algorithms, one-class SVM, t-sne, and variational autoencoder.- Developing a AutoML/Neural Evolution Machine Learning Platform <i>Darwin</i>. <p>Data Incubator Mar. - May 2016</p> <ul style="list-style-type: none">- highly competitive data science bootcamp- mini-projects: SQL, Machine Learning (including NLP and Time Series), Visualization, MapReduce, Apache Spark- capstone project: Consumer's Complaints Analysis (http://jhc-complaints.herokuapp.com)
Academic Projects	<p>Published more than 23 peer review academic journals including 10 leading author.</p> <p>Post Doctoral Scholar in University of Texas Aug. 2013 - Nov. 2016 <i>Reionization and Galaxy Formation in the Local Universe</i>: Perform data reduce and analysis (including visualization) for the large cosmological simulation for the early Universe and generates 100×15 TB outputs in Titan Supercomputer at OLCF.</p> <p>Post Doctoral Scholar in University of Kentucky Jul. 2010 - Jul. 2013 <i>The Early Massive Black Holes</i>: Improvement and implementation of massively parallelized N-body/Hydrodynamics simulations to investigate a new channel of black hole formation in the early Universe. Developed python scripts that characterize the simulated gas density probability distribution with linear and polynomial regression, and the power spectrum of the gas perturbation.</p> <p>Post Doctoral Scholar in UNLV Sep. 2007 - Jun. 2010 <i>Galaxy formation in early Universe</i>: Using massively parallelized cosmological N-body/Hydro simulation to study evolution of early galaxies. Developed and implemented key subroutines for the galaxy formation physics and on-the-fly density based clustering algorithm to extract the galaxy properties from simulations.</p>
Other Experiences	<p>Co-instructor, who organizes, mentors, and lectures, for <i>Freshman Research Initiative</i> course in Department of Astronomy, University of Texas Austin, Jan. - Dec. 2014</p> <p>Organizer for Astronomy Journal Club, UNLV, 2008 - 2009</p> <p>Refereed Papers: MNRAS, ApJ (2013 -)</p>
Computer Skills	<p>Languages & Software: Python, C/C++, Fortran, SQL, Matlab/Octave, R, Scala</p> <p>Python Data Science Tools: Scipy, Pandas, Matplotlib, StatsModels, Scikit-Learn, Keras(Tensorflow/Theano), Pytorch</p> <p>High Performance Computing Experience: Develop and implement numerical simulations in national super computing facilities such as TACC, NCSA, and OLCF.</p>
EDUCATION	<p>Ph.D. in Astronomy, University of Massachusetts at Amherst (MA, USA), Aug. 2007</p> <p>M.S. in Astronomy, Yonsei University (Seoul, Korea), Feb. 1999</p> <p>B.S. in Astronomy (minor in Physics), Yonsei University (Seoul, Korea), Feb. 1997</p>