Jun-Hwan Choi

https://choi-junhwan.github.io/homepage
choi.junhwan@gmail.com

Objectives

Data Scientist/ML Engineer with a background in computational Astrophysics.

Data Science Employment & Experiences

Data Scientist at Walmart Nov. 2021 - Present
- Developing next best action recommendation for Sam's Club memeber personalization using a response model and a renewal uplift model.
- Working on Sam's club membership renewal prediction model explainability.
- Developing a general purpose causal analysis framework. Applying the framework for marketing, CRM, and strategy use cases in Walmart.
- Developing and deploying W+ membership benefit recommendation system using a large-scale nearest neighbor search. W+ member churn analysis (model prediction and model explanation).
- Walmart online purchase basket analysis based on association rules with PySpark pipeline for PoC for the Contextual intelligence project.

Data Scientist at SparkCognition Nov. 2016 - Nov. 2021
- 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). It particularly focuses on anomaly detection and forecast.
- 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 *Darwin*.

Data Incubator Mar. - May 2016
- Highly competitive data science bootcamp: Python, SQL, Machine Learning (including NLP and Time Series), Data Visualization, MapReduce/Apache Spark, Capstone Project

Patents & Publication

1 "Diverse clustering of a data set (US20230112096A1)"
J. Choi, T. McDONNELL, Y. Lan, K. D. Moore, & C.-Y. Ho

2 "Execution of a genetic algorithm having variable epoch size with selective execution of a training algorithm (US11853893B2) "
S. Andoni, K. D. Moore, E. M. Bonab, & **J. Choi**

3 "Ensembling of neural network models (US11610131B2)"
S. Andoni, K. D. Moore, E. M. Bonab, **J. Choi**, & T. S. McDonnell

4 "Execution of a genetic algorithm with variable evolutionary weights of topological parameters for neural network generation and training (US11106978B2)"
S. Andoni, K. D. Moore, E. M. Bonab, **J. Choi**, & E. O. Korman

5 "Automated model building search space reduction (US10657447B1)"
T. S. McDonnell, S. Andoni, **J. Choi**, J. Goode, Y. Lan, K. D. Moore, & G. Sellers

6 "Divide and conquer: neuroevolution for multiclass classification"
T. McDonnell, ..., 2018, **J. Choi**, et al, Proceedings of the Genetic and Evolutionary Computation Conference, 474

Computer Skills

Languages & Software: Python (including Pytorch/Tensorflow), C/C++, Fortran, SQL, Matlab/Octave, R,
High Performance Computing Experience: Develop and implement numerical simulations in national super computing facilities (TACC, NCSA, and OLCF). Cloud computing in GCP/AWS.

Academic Projects

Published more than 20 peer review academic journals including 10 leading author.
Post Doctoral Scholar in University of Texas Aug. 2013 - Nov. 2016
Reionization and Galaxy Formation in the Local Universe: Performing data analysis for the large cosmological simulation for the early Universe.
Post Doctoral Scholar in University of Kentucky Jul. 2010 - Jul. 2013
The Early Massive Black Holes: Implementing massively parallelized N-body/Hydrodynamics simulations and developing the data analysis (regression and power spectrum) to investigate a new channel of black hole formation in the early Universe.
Post Doctoral Scholar in UNLV Sep. 2007 - Jun. 2010
Galaxy formation in early Universe Developing massively parallelized cosmological N-body/Hydro simulation to study evolution of early galaxies.

Other Experiences

Co-instructor, who organizes, mentors, and lectures, for *Freshman Research Initiative* course in Department of Astronomy, University of Texas Austin, Jan. - Dec. 2014
Organizer for Astronomy Journal Club, UNLV, 2008 - 2009
Refereed Papers: MNRAS, ApJ (2013 -)

Computer Skills

Languages & Software: Python, C/C++, Fortran, SQL, Matlab/Octave, R, Scala
Python Data Science Tools: Scipy, Pandas, Matplotlib, StatsModels, Scikit-Learn, Keras/Tensorflow, Pytorch, EconML
High Performance Computing Experience: Develop and implement numerical simulations in national super computing facilities such as TACC, NCSA, and OLCF.

EDUCATION

Ph.D. in Astronomy, University of Massachusetts at Amherst (MA, USA), Aug. 2007
M.S. in Astronomy, Yonsei University (Seoul, Korea), Feb. 1999
B.S. in Astronomy (minor in Physics), Yonsei University (Seoul, Korea), Feb. 1997