Dong Hu

(206)890-1181 · hud3@rpi.edu · https://jurohd.github.io

RESEARCH INTERESTS

Matrix Completion, Tensor Decomposition, Low-rank Approximation, Sketching, Label Complexity Reduction, High-dimensional Statistics, Randomized Linear Algebra.

EDUCATION

Rensselaer Polytechnic Institute Fall 2019 – Spring 2024 (expected)
Ph.D. in Computer Science, GPA: 3.91/4.0 Advisor: *Prof.* Alex Gittens

Rensselaer Polytechnic Institute Spring 2016 – Spring 2019
B.S. in Mathematics and Computer Science, GPA: 3.86/4.0

RESEARCH EXPERIENCE

Rensselaer Polytechnic Institute *Graduate Research Assistant*

- Reduced label complexity for nonlinear machine learning Summer 2023 – Present
  Adapted a polynomial reject-sampling algorithm to nonlinear scenarios, specifically multi-layer neural networks, by linearizing the system. Empirically showed the heuristic maintained tight approximations with reduced label complexity, analogous to the scenario in ℓ_2-regression case. Numerically demonstrated its superior performance in comparison to uniformly dropping the same number of rows.
- Noisy-CUR algorithm for matrix completion Fall 2019 – Summer 2020
  Proposed and investigated in a regression-based matrix completion algorithm (*NoisyCUR Algorithm*) for low budget matrix completion setting and experimentally verify the performance of our algorithm on both synthetic and real data, compared our algorithm with state-of-the-art Matrix completion algorithms and showed its superior performance in the low budged settings. Paper accepted by ECML 2020.

IBM, Yorktown Heights *Artificial Intelligence Research Collaboration(AIRC) scholar*

- Low-rank Tucker decomposition via sketching Fall 2021 – present
  Proposed a proximal regularized sketched alternating least squares (Tucker-ALS) algorithm for the low-rank Tucker decomposition of large tensors and proved that a sublinear rate of convergence of proximally regularized sketched CPD algorithms also holds for the proposed algorithm. Showed that the iterative nature of the Tucker-ALS approach can be algorithmically exploited to choose more performant sketching rates at different iterations. A journal paper submission to TMLR in 2023 is in preparation.
- Sparse graph based sketching Summer 2020 – Spring 2021
  Investigated sparse sketching matrices obtained from bipartite graphs, and explored two popular classes of them: expander graphs and magical graphs. Proved that for a subspace with arbitrary dimension, the minimum right vertices (the sketch size) that satisfies these two graphs for yielding a $(1 \pm \varepsilon)$ ℓ_2-subspace embedding. Empirically showed their low distortion property and low reconstruction error to the state-of-the-art sketching techniques when dealing with low rank matrix approximation problems. Publicized my implementation at *Python Matrix-Sparse-Graph-Sketching Toolkit*. Paper accepted by ICASSP 2021.

PUBLICATIONS

D. Hu, S. Ubaru, A. Gittens, K. Clarkson, L. Horesh, and V. Kalantzis. "Sparse graph based sketching for fast numerical linear algebra." in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* 2021.

D. Hu, A. Gittens, and M. Magdon-Ismail, "NoisyCUR: An algorithm for two-cost budgeted matrix completion," in *Machine Learning and Knowledge Discovery in Databases - European Conference(ECML-PKDD),* 2020