Hossam M. Ashtawy

Curriculum Vitae

☎:517.489.2714 | ✉: ashtawy@msu.edu
🏠: www.ashtawy.com | ✉: 278 Cassia Ct, Canton, MI 48187

EDUCATION

Ph.D. Electrical and Computer Engineering

2017 · Michigan State University, East Lansing, MI
GPA: 3.94
Conducted multi-disciplinary research concerned with machine learning, data mining, cloud computing and their application to computer-aided drug discovery. Worked on innovative computational tools that can help pharmaceutical companies discover and design effective drugs faster and more cost-effectively.

M.S. Electrical and Computer Engineering

2011 · Michigan State University, East Lansing, MI
GPA: 4.00
Won a full scholarship from the University of Zawia to pursue M.S. in the U.S. This M.S. scholarship was awarded to the top ranked graduates.

B.S. Computer Engineering

2006 · University of Zawia, Zawia, Libya
Graduated second out of class of more than twenty students.
Was offered a teaching assistant position and M.S. scholarship to the U.S. after graduation.

WORK EXPERIENCE

R&D Quantitative Analyst

2016-present · Ford Motor Company, Dearborn, MI

- Developed time-series models to forecast light truck and auto sales in the U.S. The models we considered include RW, ARIMA/X, ET, NARX, RNN, etc.
- Built supervised machine-learning and time-series models to forecast recession and business cycles using historical data of macroeconomic indicators.
- Developed a data collection and processing pipeline for vehicle configuration level dynamic demand and optimization model.
- Estimated the configuration demand model using a large scale discrete choice model based on multi-level nested logit trees.
- Wrote an efficient optimization library based on TensorFlow to substantially shorten the model estimation time.
- I estimated the model parameters using maximum likelihood and Bayesian estimation methods (such as MCMC and Variational Inference). We analyzed the similarities and differences between the point estimates from maximum likelihood and posterior distributions from Bayesian inference under various data processing strategies and modeling assumptions.
- Supplemented the model estimation library with a utility that efficiently finds the optimal mix and price for hundreds to thousands of configurations using the estimated demand model.

RESEARCH EXPERIENCE

Data-driven Modeling of Biochemical Interactions and Drug Discovery

2010-2016 · Michigan State University, East Lansing, MI

- Built automated data retrieval and processing pipeline from diverse biochemical archives.
- Compared and analyzed conventional scoring functions in their ability to predict binding poses, binding affinities, bioactivities and relative rank-ordering of large numbers of the collected drug-like molecules.
- Developed a comprehensive library of molecular descriptors for a data-driven multi-perspective interaction modeling of protein-ligand complexes.
- Built highly accurate task-specific scoring functions based on large ensembles of Deep Neural Networks and other machine learning algorithms such as *k*NN, MARS, RF, BRT, and SVM.
- Developed multi-task deep neural networks for simultaneous prediction of binding poses, affinities, and bioactivities for millions of complexes.

- Integrated these models into popular molecular modeling software packages for robust virtual-screening applications.

Teaching Assistant

2010-2016 · Michigan State University, East Lansing, MI

- Taught a variety of computer engineering courses that include ECE411, ECE303, ECE345, ECE430.
- Designed and supervised lab experiments and projects, wrote exams and quizzes.

Teaching Assistant

2006-2008 · University of Zawia, Zawia, Libya

- Taught various computer engineering and science related courses to freshmen students for more than two years.
- Participated in designing several computer programming related courses.
- Supervised and evaluated senior design projects.
- Wrote and graded midterm and final exams.

KEY SKILLS

- Programming languages: C/C++, Java, R, Perl, Python, MATLAB, HTML, SQL, and VHDL.
- Optimization techniques: gradient-based methods, dynamic programming, and evolutionary algorithms.
- Data analytics tools: Spark, Hadoop, Hive, Pig, and NoSQL.
- Deep Learning libraries: developed my own, TensorFlow, Keras, MXNet, Theano, and the R package darch.
- Computational platforms: PCs, HPCC (Grid and/or Cluster) systems, and cloud infrastructures such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform.
- Molecular modeling software: Autodock, GOLD, MOE, OpenEye, Discovery Studio, Dock, SYBYL, PyMOL, Chimeara, and several others.
- Operating systems: Linux, Windows, and macOS.

PROFESSIONAL SOCIETY AFFILIATIONS

- Institute of Electrical and Electronics Engineers (IEEE), member.
- IEEE Computer Society, member.
- American Chemical Society (ACS), member.

AWARDS

- M.S. Scholarship for distinguished students from the University of Zawia, Libya in 2008.
- Ph.D. Dissertation Completion Fellowship ($5000) from Michigan State University, East Lansing, MI in 2015.
- Summer Research Fellowship ($2000) from Michigan State University, East Lansing, MI in 2015.
- Amazon AWS Research Grant ($4200) to support my research in data-driven drug discovery from Amazon.com, Inc. in 2016.

SAMPLE PROJECTS AND ON-LINE ACCOUNTS

- Descriptor Data Bank: descripordb.com
- Task-specific scoring functions: github.com/ashtawy/task-specific_scoring_functions
- Large Scale Discrete Choice Model: github.com/ashtawy/dcmflow
- Personal website: ashtawy.com
- Linked-in website: linkedin.com/in/ashtawy

WORK AUTHORIZATION

- A permanent U.S. resident: authorized to work in the US on a permanent basis.
- Willing to re-allocate within the U.S.
- Willing to travel domestically and internationally.

PUBLICATIONS

[1] H. M. Ashtawy and N. R. Mahapatra, "Task-specific scoring functions for predicting ligand binding poses and affinity and for screening enrichment," *Journal of chemical information and modeling*, vol. 58, no. 1, pp. 119–133, 2018.

[2] H. M. Ashtawy and N. R. Mahapatra, "Descriptor data bank (ddb): A cloud platform for multiperspective modeling of protein–ligand interactions," *Journal of chemical information and modeling*, vol. 58, no. 1, pp. 134–147, 2018.

[3] H. M. Ashtawy and N. R. Mahapatra, "Boosted neural networks scoring functions for accurate ligand docking and ranking," *Journal of bioinformatics and computational biology*, vol. 16, no. 2, p. 1850004, 2018.

[4] H. M. Ashtawy and N. R. Mahapatra, "BgN-Score and BsN-Score: Bagging and boosting based ensemble neural networks scoring functions for accurate binding affinity prediction of protein-ligand complexes," *BMC bioinformatics*, vol. 16, no. Suppl 4, p. S8, 2015.

[5] H. M. Ashtawy and N. R. Mahapatra, "Machine-learning scoring functions for identifying native poses of ligands docked to known and novel proteins," *BMC Bioinformatics*, vol. 16, no. Suppl 6, p. S3, 2015.

[6] H. M. Ashtawy and N. R. Mahapatra, "A comparative assessment of predictive accuracies of conventional and machine learning scoring functions for protein-ligand binding affinity prediction," *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, vol. PP, no. 99, pp. 1–1, 2014.

[7] H. M. Ashtawy and N. R. Mahapatra, "Ensemble neural networks scoring functions for accurate binding affinity prediction of protein-ligand complexes," in *Pattern Recognition in Bioinformatics: 9th IAPR International Conference, PRIB 2014, Stockholm, Sweden, August 21-23, 2014. Proceedings*, vol. 8626, p. 129, Springer, 2014.

[8] H. M. Ashtawy and N. R. Mahapatra, "Molecular docking for drug discovery: Machine-learning approaches for native pose prediction of protein-ligand complexes," in *Computational Intelligence Methods for Bioinformatics and Biostatistics*, pp. 15–32, Springer, 2014.

[9] H. M. Ashtawy and N. R. Mahapatra, "Does accurate scoring of ligands against protein targets mean accurate ranking?," in *Proc. 9th International Symposium on Bioinformatics Research and Applications (ISBRA 2013)*, pp. 298–310, Springer, 2013.

[10] H. M. Ashtawy and N. R. Mahapatra, "Enn-score: An ensemble neural networks scoring function for accurate binding affinity prediction of protein-ligand complexes," in *Proc. 9th International Symposium on Bioinformatics Research and Applications (ISBRA 2013)*, pp. 54–61, 2013.

[11] H. M. Ashtawy and N. R. Mahapatra, "A comparative assessment of ranking accuracies of conventional and machine-learning-based scoring functions for protein-ligand binding affinity prediction," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, vol. 9, no. 5, pp. 1301–1313, 2012.

[12] H. M. Ashtawy and N. R. Mahapatra, "A comparative assessment of ranking powers of conventional and machine-learning-based scoring functions on diverse and homogeneous test sets," in *Proc. 10th Asia-Pacific Bioinformatics Conference (APBC 2012)*, pp. 241–254, 2012.

[13] H. M. Ashtawy and N. R. Mahapatra, "A comparative assessment of conventional and machine-learning-based scoring functions in predicting binding affinities of protein-ligand complexes," in *Bioinformatics and Biomedicine (BIBM), 2011 IEEE International Conference on*, pp. 627–630, IEEE, 2011.