Haluk Dogan

1106 Harbor Landing, Roswell, GA, 30076
+1 404-421-8408
hdogan@vivaldi.net
https://haluk.github.io
https://github.com/haluk

Research Interest

My general areas of interest include machine learning, big data, artificial intelligence, and Bayesian model optimization. I have experience in building various models using decision trees, random forests, SVM, CRF, Naive Bayes, Bayesian Network, and deep learning architectures including but not limited to CNN, Bi-LSTM with Attention, Seq2Seq models, and VAE/CVAE for big data problems. Data grows rapidly and the need for efficient data processing and modeling increases. I am planning to direct my research toward building efficient systems that can deal with large volumes of data.

Education

PhD	Computer Science	University of Nebraska-Lincoln, NE, USA	2018 – 2021
MS	Computer Engineering	Bogazici University, Istanbul, Turkey	2011 – 2013
BS	Computer Science	Istanbul Bilgi University, Istanbul, Turkey	2006 – 2010

Experience

➤ **Machine Learning Scientist**
Software Developer **Prorize, LLC, GA, USA** **May 2021 – Current**

- Improved existing forecast by developing a new ML based method by 10% (MAPE)
- Developed ML methods for time-to-event data analysis
- Developed a central notification system
- Built a data layer API
- Developed front-end components
- Technologies
  - Programming Languages: Python, C#, Kotlin, Java
  - Development Stack: ASP.NET Core, React, Flask, Spring Boot
  - Deep Learning Frameworks: PyTorch
  - Version Control: Git

➤ **Machine Learning Scientist** **University of Nebraska-Lincoln, NE, USA** **Aug 2013 – May 2021**

- Built machine learning models with a focus on graphical models and Bayesian statistics
- Built discriminative and generative deep learning architectures
- Applied design patterns (Builder, Decorator, Strategy) and event driven programming to maximize development productivity
- Research resulted in 6 journal/conference publications
- Technologies
  - Programming Languages: Python, R, Java, Bash, Anglican, Pyro
  - Deep Learning Frameworks: Tensorflow, PyTorch
  - Machine Learning and Numeric Libraries: Scikit-learn, Pandas, Numpy, Scipy, pyAgrum, Orange
  - Data Visualization: Matplotlib, ggplot2, Seaborn
  - Version Control: Git
  - Virtualization: Docker, Google Cloud Platform
  - Development and Runtime Environment: Linux, Emacs, Open Science Grid

➤ **Co-founder/Python Developer** **Roomkita, Istanbul, Turkey** **Aug 2013 – Aug 2019**

- Backend development for a travel agency website using Model-View-Controller (MVC) design pattern
- Developed machine learning models to improve search results that prioritize user preferences based on user clicks
- The company was featured in tnooz, a global provider of news related to travel technology
- Technologies
  - Programming Languages: Python
  - Database: PostgreSQL
  - Version Control: Git, Subversion
  - Web Server: NGINX
  - Development and Runtime Environment: Linux, Emacs

➤ **Software Developer** **Istanbul Bilgi University, Istanbul, Turkey** **Nov 2010 – Aug 2013**

- Lead recitation/lab hours for “Introduction to Programming (Racket/Lisp)”, “Probability and Statistics (Python)”, “Bioinformatics (Python)” courses
- Developed software for research activities in the department
- Participated in the development and maintenance of college website

➤ **Java Software Developer** **i2i Systems, Istanbul, Turkey** **Mar 2010 – June 2010**

- Converted billing rules defined by analysts in plain text to LL grammars
- Developed a program that parses plain text using defined grammars and update billing database
- Software was incorporated into routine operations of the billing department to facilitate billing
- Technologies
  - Programming Languages: Java
  - Libraries: Spring Framework, Hibernate ORM
  - Database: Oracle
  - Version Control: Subversion
  - Build System: Maven
  - Development and Runtime Environment: Linux, Eclipse, Cron

➤ **Data Scientist** **GNA, Istanbul, Turkey** **Aug 2009 – Mar 2010**

- Performed Extract, Load, Transfer operations
- Built data warehouse to prepare weekly business reports
- Added custom features to an open source business intelligence tool
  - Programming Languages: Java
  - Business Intelligence Tool: Pentaho
  - Database: Oracle
  - Build System: Maven
  - Development and Runtime Environment: Linux, Oracle Software Developer

➤ **Java Developer** **Aradiom, Istanbul, Turkey** **Mar 2008 – Jun 2008**

- Developed backend/frontend of a regex editor to create cron jobs
  - Programming Languages: Java
  - Libraries: JBoss Seam Framework
  - Version Control: Subversion
  - Development and Runtime Environment: Linux, Eclipse, Cron

Services

- Workshop Co-Organizer
  - The International Workshop on Expository Representation Learning of Biomedical Data, IEEE BIBM 2019, San Diego, CA, USA (`http://sbbi-panda.unl.edu/bibm2019/`)
  - Interactive Workshop on Support Vector Machine (SVM) for Classification and Regression Problems, UNMC 2018, Omaha, NE, USA (`http://sbbi-panda.unl.edu/svm-workshop/`)
- Reviewer
  - Cancer Medicine reviewer 2020 – Current
  - BMC Bioinformatics reviewer 2019 – Current
  - NeurIPS sub-reviwer 2020 – Current
  - IJCAI sub-reviewer 2018 – Current

Publications

15. Madadjim, R, H Dogan, and J Cui (2022). Computational learning of small RNA regulation in pancreatic cancer progression. In: *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, pp.162–167.
14. Zhou, F, P Ebea, E Mutai, H Wang, S Sukreet, S Navazesh, H Dogan, W Li, J Cui, P Ji, DMO Ramirez, and J Zempleni (2022). Small Extracellular Vesicles in Milk Cross the Blood-Brain Barrier in Murine Cerebral Cortex Endothelial Cells and Promote Dendritic Complexity in the Hippocampus and Brain Function in C57BL/6J Mice. *Frontiers in Nutrition* **9**.
13. Dogan, H, Z Hakguder, R Madadjim, S Scott, M Pierobon, and J Cui (Aug. 2021). Elucidation of dynamic microRNA regulations in cancer progression using integrative machine learning. *Briefings in Bioinformatics*.
12. Dogan, H, J Shu, Z Hakguder, Z Xu, and J Cui (Oct. 2020). Elucidation of molecular links between obesity and cancer through microRNA regulation. *BMC Medical Genomics* **13**(1).
11. Cui, J, J Shu, T Gao, and H Dogan (July 2019). Unraveling exosome-enabled cancer signaling: An integrated genomic approach. In: *Molecular and Cellular Biology / Genetics*. American Association for Cancer Research.
10. Dogan, H, Z Hakguder, S Scott, and J Cui (Nov. 2019). Elucidation of MicroRNA-Gene Regulation in Human Cancer with Integrative Network Models. In: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE.
9. Li, H, H Dogan, and J Cui (Nov. 2019). A New Approach to Batch Effect Removal Based on Distribution Matching in Latent Space. In: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE.
8. Quint, E, D Xu, H Dogan, Z Hakguder, S Scott, and M Dwyer (2019). Formal language constraints for markov decision processes. *arXiv preprint arXiv:1910.01074*.
7. Xu, D, E Quint, Z Hakguder, H Dogan, S Scott, and M Dwyer (2018). Constraining Action Sequences with Formal Languages for Deep Reinforcement Learning.
6. Tomov, ML, ZT Olmsted, H Dogan, E Gongorurler, M Tsompana, HH Otu, M Buck, EA Chang, J Cibelli, and JL Paluh (Dec. 2016). Distinct and Shared Determinants of Cardiomyocyte Contractility in Multi-Lineage Competent Ethnically Diverse Human iPSCs. *Scientific Reports* **6**(1).
5. Wang, F et al. (Feb. 2016). Detecting Microbial Dysbiosis Associated with Pediatric Crohn Disease Despite the High Variability of the Gut Microbiota. *Cell Reports* **14**(4), 945–955.
4. Dogan, H, H Can, and HH Otu (Jan. 2014). Whole Genome Sequence of a Turkish Individual. *PLoS ONE* **9**(1), e85233.
3. Nalbantoglu, U, A Cakar, H Dogan, N Abaci, D Ustek, K Sayood, and H Can (Aug. 2014). Metagenomic analysis of the microbial community in kefir grains. *Food Microbiology* **41**, 42–51.

2. Dogan, H and HH Otu (Aug. 2013). “Objective Functions”. In: *Methods in Molecular Biology*. Humana Press, pp.45–58.

1. Isci, S, H Dogan, C Ozturk, and HH Otu (Nov. 2013). Bayesian network prior: network analysis of biological data using external knowledge. *Bioinformatics* **30**(6), 860–867.