Halid Ziya Yerebakan

Contact   halidziya@gmail.com

Interests   Machine Learning, Natural Language Processing, Computer Vision
Bayesian Non-Parametrics, Clustering, Topic Models, Representation Learning, Language Models, Transfer Learning, Parallel Computing, Metric Learning, Speech Recognition.

Education   **Purdue University**, West Lafayette, IN

Ph.D. Computer Science, GPA : 3.97, Spring 2017
Hierarchical Non-parametric Bayesian Mixture Models and Applications

Fatih University, Istanbul, Turkey

B.S., Electronics , GPA : 3.92, June 2010

B.S., Computer Engineering , GPA : 3.91 , July 2010

Experience   **Research Professional**   4/2019 - Present , **SIEMENS Healthineers**, Malvern, PA
- Developed machine learning algorithms for image and text data.
- Released algorithms for NLP engine in AI-Pathway Companion product.

Senior Scientist   6/2017 - 4/2019 , **SIEMENS Healthineers**, Malvern, PA
Research Assistant   2011 - 2017, **Computer and Information Science**, IUPUI
- **I2GMM** : The Infinite Mixture of Infinite Gaussian Mixtures, a non-parametric generative Bayesian clustering model that allows an arbitrary number of clusters with arbitrary shapes.
- **hLDA** : Developing new inference algorithm for non-parametric tree structured topic model hierarchical latent Dirichlet allocation.

Teaching Assistant   2011 - 2017 , **Computer and Information Science**, IUPUI
- **Courses :** Operating Systems, Discrete Mathematics, Data Mining, Machine Learning

Machine Learning Intern   Summer 2015 and 2016, **SIEMENS Healthcare**, Malvern, PA
- Developed machine learning methods on radiology reports
- Collected public text datasets for radiology report analysis
- Developed a new bayesian nonparametric model for hierarchical word clustering
- Developed deep learning algorithms in Python/Theano on GPU for scalable inference

Statistics Intern   Summer 2014 , **Dow Agrosciences (via Kelly Services)**, IN
- Improved a statistical simulation by making it robust, faster, and more visual
- Awarded best intern poster presentation
- Worked in a high performance cluster environment using R

Machine Learning Intern   Summer 2012 , **Bashpole Software Inc**, North Webster, IN
- Improved performance of record matching using various machine learning and NLP methods
- Developed a Javascript Chrome extension for information extraction, released in Web Store
- Conducted research on different classifier and distance metric alternatives for record matching

Software Engineer Intern,   Summer 2009, **C Tech Information Technologies**, Gebze, TURKEY
- TetikG : RTSP Stream transfer using BeagleBoard embedded system and its GPIO driver

Software Engineer Intern,   Summer 2008, **Surat Technology**, Istanbul, TURKEY
- Developed software for storage management systems

Publications

1. **Halid Yerebakan**, Anna Jerebko, Parmeet Bhatia "Processing a medical image to provide a visual indicator". EP Patent Application EP3869453A1 (2021)

2. **Halid Yerebakan**, Yoshihisa Shinagawa. "Encoding textual information for text analysis." U.S. Patent Application No. 16/846,756, US20200334410A1 (2020).

3. Yoshihisa Shinagawa, **Halid Ziya Yerebakan** "Supervised Features For Text Classification", EP Patent Application EP3657354A1 (2020).

4. Colin B Hansen, Yiyuan Zhao, **Halid Yerebakan**, Luca Bogoni, Anna Jerebko "False positive reduction of vasculature for pulmonary nodule detection", Proceedings Volume 11314, Medical Imaging 2020: Computer-Aided Diagnosis; 113142B (2020)

5. **Halid Ziya Yerebakan**, Yoshihisa Shinagawa, Parmeet Bhatia, Yiqiang Zhan "Visualization framework based on document representation learning", US Patent App. 15/865,539.

6. **Halid Ziya Yerebakan**, Yoshihisa Shinagawa, Parmeet Bhatia, Yiqiang Zhan "Document Representation Learning For Patient History Visualization", Coling 2018 Demo Paper.

7. **Halid Ziya Yerebakan**, Murat Dundar "Partially Collapsed Parallel Gibbs Sampler for Dirichlet Process Mixture Models", Pattern Recognition Letters, 2017

8. **Halid Ziya Yerebakan**, Yiqiang Zhan, Fitsum Reda, Yoshihisa Shinagawa "Hierarchical Latent Word Clustering", NIPS Non-parametric Bayesian Workshop, NIPS'15.

9. **Halid Z. Yerebakan**, Bartek Rajwa, Murat Dundar, "The Infinite Mixture of Infinite Gaussian Mixtures", (I2GMM) NIPS'14

10. Murat Dundar, **Halid Ziya Yerebakan** , Bartek Rajwa, "Batch Discovery of Recurring Rare Classes toward Identifying Anomalous Samples" , SIGKDD 2014.

11. Murat Dundar, Ferit Akova, **Halid Z. Yerebakan**, Bartek Rajwa, "A Non-parametric Bayesian Model for Joint Cell Clustering and Cluster Matching: Identification of Anomalous Sample Phenotypes with Random Effects," BMC Bioinformatics 15 (1), 314, 2014

Awards

Ranks:
- Ranked 169 in nationwide University Placement Exam (OSS) among about 2M participants
- Ranked 658 in nationwide Graduate Examination among 225,930 participants
- Ranked 1500 in High School Placement Exam in TURKEY among about 500,000 students
- Ranked 2 upon graduation, Electronics Engineering in Fatih University

Scholarships:
- Research Assistant, IUPUI (NSF Funded)
- Graduate scholarship, Tubitak (Turkey's NSF) award, 2010, TURKEY
- Full scholarship, Fatih University, and financial support funding
- %80 Scholarship, private school Yildirim Han College (High School)
- Turkish Government Scholarship (DPY) , Middle School

Contests:
- Winner AWS Challenge HackOhio Hackathon, Ohio State University, 2016
- Top10 HackOhio Hackathon, Ohio State University, 2016
- Ranked 1st in MayIchallenge Artificial Intelligence contest, METU, 2011
- Ranked 2nd, Bilkent University Robotic Contest 2010, Robotic arm project, TURKEY.
- Second stage, Tubitak National Computer Olympiads and bootcamp(top 50 in the nation)
- Middle School Computer Contest 2. Degree in City Mersin

Technical Skills

- Python, Numpy, Tensorflow, TFLite, Keras, PyTorch, C, C++, Matlab, Java, multi-core programming, R, Mallet, QT, JavaScript, Jquery, J2ME, OpenGL, .NET, Weka, SDL, PHP, Docker, HTML5, XML, CSS, SQL, cluster environment (HPC), Linux, Latex, LyX, Overleaf, Julia, Git, Vmware, MySQL, Oracle, GNU make, Arduino, BeagleBoard, Rasberry pi, Pic, X86, Verilog HDL, AutoCAD, Photoshop, Wireshark, Nginx, C#, Pascal, AWS, Azure, Django, REST, Apache:(UIMA, Lucene, CTakes, Http Server), Android, Kotlin, Dart

URLs

- https://github.com/halidziya
- https://sourceforge.net/u/halidziya/profile/
- https://halidziya.blogspot.com
- https://scholar.google.com/citations?user=jgI0b8cAAAAJhl=en
- https://www.linkedin.com/in/halid-ziya-yerebakan-a7124226/
- http://www.enderunix.org/docs/devbook/
- https://lessontime.ai

Service

Academic:
- SIAM SDM 2014 Reviewer
- NIPS 2016 Reviewer

Extracurricular:
- Open source book translation : FreeBSD Developers Handbook, Academy member
- Volunteer, Tutoring Classes