

Sen (Forrest) Yang

700 E Middlefield Rd, Mountain View, CA | +1 (848)-565-5991 | sy358@rutgers.edu

[LinkedIn](#) | [Google Scholar](#)

Summary

Experienced Engineer and Researcher in Machine Learning. Currently working at LinkedIn AI as a Machine Learning Engineer. Open to Machine Learning Engineer and Research Scientist roles.

TECHNICAL SKILLS

Programming: Java, Python, Spark (Scala), Tensorflow, R, Matlab
Specialty: Solving large-scale problems in Data Mining, NLP, Recommender System, Process Mining
Data Engineering: Data pipeline, Model Versioning, A/B test, Service Maintenance (online, near-line, offline)

EDUCATION

Rutgers University, the State University of New Jersey Piscataway, NJ
• *Ph.D.*, in Electrical & Computer Engineering, GPA 3.9/4.0 Sept 2013 – Oct. 2018
• Advisor: [Ivan Marsic](#) and [Hui Xiong](#)

Rutgers University, the State University of New Jersey Piscataway, NJ
• *M.S.*, in Electrical & Computer Engineering, GPA 3.9/4.0 Sept 2013 – Jun. 2015

Nanjing University of Posts and Telecommunications, China Nanjing, China
• *B.A.*, in Communication Engineering, GPA 84/100 Sept 2008 – June 2012

WORKING EXPERIENCE

LinkedIn Mountain View, CA
• Machine Learning Engineer @ [Data Standardization and Knowledge Graph Team](#), LinkedIn AI Dec 2018 – Present
• Deployed multiple end2end ML models on LinkedIn data standardization, inference and embedding
• Standardized LinkedIn member data with deep-n-wide models and multilingual NLP techniques
• Improve entity (e.g., degree) vocabulary with in-house linguistic specialists and external crowdsourcing
• Sponsored Updated Revenue: + 0.43%, Gold Members: +1.4M (~\$32M in revenue)

Nokia Bell Labs Murray Hill, NJ
• Data Scientist Intern & Machine Learning Engineer Co-op @ Data Science Research Group May 2017 – Feb 2018
• Participated Project: Automated Machine Learning & May 2018 - Present
• Selected to present to [Marcus Weldon](#), president of Bell Labs and CTO of Nokia.

Rutgers University & Children's National Medical Center (CNMC) Piscataway, NJ
• [NIH project](#): Automatic Workflow Capture & Analysis for Improving Trauma Resuscitation Outcomes Aug 2014 – Oct. 2018

Huawei Technologies Co. Ltd. Shenzhen, China
• Technical Engineer on 4G network Aug 2012 – June 2013

RESEARCH AND PROJECTS (selected)

- 1. Automated Feature Extraction from Time Series with Deep Learning** *Bell Labs, 2017 - Present*
 - Research the use of Autoencoder (with LSTM, Bi-LSTM, CNN, etc.) for unsupervised feature learning for time series
 - Compare the feature learning performance of deep learning approaches with traditional statistical approaches (TSFresh)
 - Evaluate in both supervised and unsupervised ways with UCR data collections and two other datasets collected at Bell labs
- 2. Deep Learning Framework for Next Medical Treatment Activity Recommendation** *NIH Project, 2017 - 2018*
 - Developed a context-aware RNN based recommender system to provide runtime treatment recommendations
 - Invented Act2vec, a method to embed human activities or events into numerical vectors via a neural net
 - Proposed a novel data augmentation algorithm that can fabricate synthetic patient data that closely resembles authentic data
 - Evaluated our system on two medical processes and achieved a top-1 accuracy of 0.46 and a top-3 accuracy of 0.77 (from 15 classes)
- 3. Smart Trauma Resuscitation Decision Support System** *NIH Project, 2014 - Present*
 - Design a computerized decision support system to monitor trauma resuscitation workflows and alter medical team of errors
 - Develop a sensor-based (RFID, Video, and microphone) system to automatically identify medical team activities

- Analyze medical team activities, discovering treatment patterns, detecting human errors and extracting medical knowledge

4. Prototype-based Recommender System for Medical Treatment Procedures

NIH Project, 2015 - 2017

- Presented a data-driven recommender system that provides step-by-step treatment recommendations
- Developed a time-warping-based temporal sequence similarity measure (TwS-PT) and a method for calculating prototype sequence
- Tested our methods on three real-life medical processes and achieved accuracy up to an F_1 score of 0.77 (0.37 for baseline)
- Implemented as a web app ([VIT-PLA 2.0](#)) using D3.js, JSP, Java and includes interactive visual functions

SELECTED PUBLICATIONS ([Full List](#))

1. Leveraging Adversarial Training in Self-Learning for Cross-Lingual Text Classification	2020
Dong X, Zhu Y, Zhang Y, Fu Z, Xu D, Yang S , De Melo G <i>43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)</i>	
2. An Approach to Automatic Process Deviation Detection in a Time-Critical Clinical Process	2018
Yang S , Sarcevic A, Farneth RA, Chen S, Ahmed OZ, Marsic I, Burd RS <i>Journal of Biomedical Informatics, Elsevier (JBI)</i>	
3. Discovering Urban Travel Demands through Dynamic Zone Correlation in Location-based Social Networks	2018
Hu W, Yao Z, Yang S , Chen S, Jin PJ <i>European Conference on Machine Learning and Knowledge Discovery in Databases (PKDD 2018)</i>	
4. Automated Mining of Approximate Periodicity on Numeric Data: A Statistical Approach	2018
He R, Yang S , Yang J, Cao J <i>2018 ACM International Conference on Compute and Data Analysis (ICCD2018)</i>	
5. A Data-driven Process Recommender Framework	2017
Yang S , Dong X, Sun L, Zhou Y, Farneth RA, Xiong H, Burd RS, Marsic I <i>2017 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2017)</i>	
6. Medical Workflow Modeling Using Alignment-Guided State-Splitting HMM	2017
Yang S , Zhou M, Chen S, Dong X, Ahmed O, Burd RS, Marsic I <i>IEEE International Conference on Healthcare Informatics (ICHI 2017)</i>	
7. Process-oriented Iterative Multiple Alignment for Medical Process Mining	2017
Chen S, Yang S , Zhou M, Burd R, Marsic I <i>IEEE International Conference on Data Mining Workshop (ICDM Workshop 2017)</i>	
8. VIT-PLA: Visual Interactive Tool for Process Log Analysis	2016
Yang S , Dong X, Zhou M, Li X, Chen S, Webman R, Sarcevic A, Marsic I, Burd RS <i>KDD 2016 Workshop on Interactive Data Exploration and Analytics (KDD Workshop 2016)</i>	
9. Duration-Aware Alignment of Process Traces	2016
Yang, S. , Zhou, M., Webman, R., Yang, J., Sarcevic, A., Marsic, I. and Burd, R.S <i>Industrial Conference on Data Mining. Springer International Publishing, 2016</i>	

Professional Activity

Program Committee: KDD 2019, KDD 2020

Reviewer: ACM Transactions on Knowledge Discovery from Data;

ACM Transactions on Management Information System;

Data & Knowledge Engineering;

IEEE Transactions Big Data;

IEEE Intelligent Systems.

> During my Ph.D. study, I mentored 28 graduate and undergraduate students on their research. Most of them are now software developers or data analysts in top IT companies.