Shoumik Roychoudhury

CONTACT INFORMATION

3592 Cresson Street
Apt D4
Philadelphia, PA 19129

Phone: +1(856) 905-4890
*E-mail:*shoumik.rc@gmail.com
shoumik.rc@temple.edu

EDUCATION

Temple University, Philadelphia, PA, USA

PhD, Computer and Information Science, May 2020

- Research expertise: Machine Learning, Data mining, Deep Learning, Temporal pattern discovery, Time-series analysis, Time-series classification, Sequence modeling, Health informatics.
- Dissertation Title: *Leveraging Temporal Subsequences for Time-series Classification.*

MS, Electrical and Computer Engineering, December 2011

- Research Area: Computer vision, Moving object tracking, Thermal video analysis
- Thesis Topic: *Tracking Human in Thermal Vision using Multi-feature Histogram.*

TECHNICAL SKILLS

Programming Languages : C/C++, Java, Python, Matlab, SQL.
Frameworks used: TensorFlow, Keras, PyTorch, MySQL, PostgreSQL, Hive.

RESEARCH EXPERIENCE

Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA
Research Intern — May 2018 - August 2018

- Investigated fast pattern matching methods to identify and extract unique temporal patterns characterizing home electrical appliances from signals collected through Home Energy Management System (HEMS) for modeling smart home behaviors.

Temple University, Philadelphia, PA, USA
PhD Research Assistant — January 2013 - May 2020

- **US Dept. of the Navy, Office of Naval Research, Auxiliary System Sensor Fusion (subcontract to Technical Documentation Inc.)**
  - Proposed interpretable cost-sensitive framework for early classification of cardiac arrhythmia alarms from bedside monitors in ICU implemented in Matlab.
  - Statistically significant improvement in terms of classification accuracy over state-of-the-art methods achieving 34% false alarm suppression with 100% true alarm detection rates.

- **National Science Foundation funded BIGDATA project**
  - Proposed an algorithm implemented in Java which significantly improved the time-series classification accuracy by extracting novel temporal subsequence order information from multivariate time-series data.
  - Improved identification of Poll-score trends in 2016 US Presidential election from temporal information extracted from large scale twitter data of 12 million tweets via an ensemble based multivariate time-series classification model implemented in Java.

- Major improvement in classification accuracy for across 18 highly imbalanced time-series datasets via a novel cost-sensitive learning framework implemented in Java.

- **IQVIA funded research project**
  - Created and analyzed longitudinal patient visits from a multi-domain EHR repository of 40 million patients by leveraging the OMOP CDM architecture using HiveQL.
  - As a member of a 4-person team developed disease-agnostic multi-domain stacked deep sequence model using PyTorch which significantly improved disease detection predictive performance in terms of AUPRC by more than 10% on average compared to the individual domain models as well as joint domain model.

- **Defense Advanced Research Projects Agency (DARPA) funded project**
  - Proposed a novel data driven approach to discover proxies for target diagnosis from large scale hospital discharge records databases achieving 94% prediction accuracy.

Peer-reviewed Publications

- **Roychoudhury, S**., Zhou, F., Obradovic, Z. "Leveraging Subsequence-orders for Univariate and Multivariate Time-series Classification," *Proc. 19th SIAM Intl Conf. Data Mining (SDM)*, Calgary, Canada, May 2019.

- **Roychoudhury, S**., Ghalwash, M., Obradovic, Z. "Cost-sensitive Time-series classification," *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, Skopje, Macedonia, September 2017.

- **Roychoudhury, S**., Ghalwash, M., Obradovic, Z. "False Alarm Suppression in Early Prediction of Cardiac Arrhythmia," *Proc. 15th IEEE International Conference on Bioinformatics and Bioengineering*, Belgrade, Serbia, November 2015.

- Kezunovic, M., Obradovic, Z., Dokic, T., **Roychoudhury, S**. "Systematic Framework for Integration of Weather Data into Prediction Models for the Electric Grid Outage and Asset Management Applications," *Proc. 51st IEEE Hawaii International Conference on System Science (HICSS)*, Big Island, Hawaii, January 2018.

- Mirowski, T., **Roychoudhury, S**., Zhou, F., Obradovic, Z. "Predicting Poll Trends using Twitter and Multivariate Time-series Classification," *Proc. 8th Int'l Conf. Social Informatics (SocInfo)*, Seattle, WA, November 2016.

- Ramljak, D., Davey, A., Uversky, A., **Roychoudhury, S**., Obradovic, Z. "Casting a Wider Net: Data Driven Discovery of Proxies for Target Diagnoses," *AMIA 2015 Annual symposium*, San Francisco, November 2015.

- Ramljak, D., Davey, A., Uversky, A., **Roychoudhury, S**., Obradovic, Z. "Hospital Corners and Wrapping Patients in Markov Blankets," *4th Workshop on Data Mining for Medicine and Healthcare at SIAM SDM*, May 2015.

Publications (under submission)

- **Roychoudhury, S**., Zhou, F., Obradovic, Z. "Learning Shapelets and Temporal Dependencies from Randomly Initialized Subsequences," *in review.*

- **Roychoudhury, S**., Cao, X.H., Ljubic, B., Pavlovski, M., Glass, L., Nair, R., Obradovic, Z. "Multi-domain Stacking Deep Sequence Model for Disease Diagnosis," *in preparation.*

- Ljubic, B., **Roychoudhury, S**., Cao, X.H., Pavlovski, M., Nair, R., Glass, L., Obradovic, Z. "Influence of Cohort Selection on Deep Learning for Alzheimer's Disease Prediction," *submitted to Elsevier Journal of Computer Methods and Programs in Biomedicine.*
- Cao, X.H., Ljubic, B., Pavlovski, M., **Roychoudhury, S.**,Glass, L., Obradovic, Z. "Learning Input and Output Kernels for Time-to-Event Prediction on High-Dimensional Gene Expression Data," *submitted to IEEE Journal of of Biomedical and Health Informatics.*
- Ljubic, B., Alshehri, J., **Roychoudhury, S**., Bajik, V., Pavlovski, Neste, C., V., Obradovic, Z. " Genetics and Comorbity Network of Colorectal Cancer," *in review.*

Services

- PC member: ECML-PKDD 2020 (Research Track)
- Invited Reviewer: ECML-PKDD 2020 (Research Track), Start Talking Science, AMIA, ICTAI, Mary Ann Liebert Big Data
- Ad-Hoc reviewer: KDD, IEEE Big Data, ICDM

Teaching Experience

Temple University, Philadelphia, PA, USA
Teaching Assistant

• Electrical Engineering Science Lab 1	September 2009 - December 2009
• Classical Control System Lab	January 2010 - May 2010
• Signals	September 2010 - December 2010
• Digital Circuit Design Lab	January 2011 - May 2011
• Math for a Digital World	September 2012 - May 2013
• Mathematical Concepts in Computing II	January 2015 - December 2015
• Data Structure	September 2019 - May 2020

Awards and Scholarships

SIAM International Conference on Data Mining (SDM14) Student Travel Award.

References

Dr. Zoran Obradovic (PhD Advisor)
Professor and Director, Center for Data Analytics and Biomedical Informatics
Department of Computer and Information Science
Temple University
Philadelphia, PA
Email: zoran.obradovic@temple.edu

Dr. Slobodan Vucetic
Professor, Department of Computer and Information Sciences
Temple University
Philadelphia, PA
Email: vucetic@temple.edu

Dr. Mohamed Ghalwash
Research Scientist, AI for Healthcare
IBM Research Center
Yorktown Heights, NY
Email: Mohamed.Ghalwash@ibm.com