Weishi SHI

ws7586@rit.edu

EDUCATION

AUGUST 2016 – NOW	PhD student of Computing and Information Sciences
	Rochester Institute of Technology, Rochester, NY.
	Research topic: Active learning in knowledge-rich domains.
JUNE 2016	Master of Science
	Filed of study: Information Sciences And Technologies
	Rochester Institute of Technology, Rochester, NY
	Master Project: Programming knowledge tag recommendation
2005 – 2009	Bachelor of Engineering
	Filed of study: Software Engineering
	Xi'an Jiao Tong University,China

RESEARCH INTEREST

Machine Learning, Data mining, Active Learning, Bayesian Methods.

PUBLICATIONS

Google scholar: https://scholar.google.com/citations?user=nAPZIPsAAAAJhl=en

- Shi, Weishi, et.al. "A Gaussian Process-Bayesian Bernoulli Mixture Model for Multi-Label Active Learning". Neural Information Processing Systems (NIPS) 2021

- Shi, Weishi, & Yu, Qi. "Active Learning with Maximum Margin Sparse Gaussian Processes". International Conference on Artificial Intelligence and Statistics (AISTATS) 2021

- Shi, Weishi, et.al. "Multifaceted Uncertainty Estimation for Label-Efficient Deep Learning". Neural Information Processing Systems (NIPS) 2020

- Moayad Alshangiti, Shi, Weishi, et.al. "A Bayesian learning model for design-phase service mashup popularity prediction" Expert Systems with Applications (ESWA) 2020

- Shi, Weishi, et.al. "Presenting and Evaluating the Impact of Experiential Learning in Computing Accessibility Education" International Conference on Software Engineering (ICSE) 2020

- Shi, Weishi, & Yu, Qi. "Integrating Generative and Discriminative Sparse Kernel Machines for Multi-class Active Learning". Neural Information Processing Systems (NIPS) 2019

- Shi, Weishi, & Yu, Qi. "Fast Direct Search in an Optimally Compressed Target Space for Efficient Multi-Label Active Learning". International Conference on Machine Learning (ICML) 2019.

- Lima, E., Shi, Weishi., Liu, Xumin., & Yu, Qi. "Integrating Multi-level Tag Recommendation with External Knowledge Bases for Automatic Question Answering". ACM Transactions on Internet Technology (TOIT) 2019.

- Shi, Weishi, & Yu, Qi. "An Efficient Many-Class Active Learning Framework for Knowledge-Rich Domains." IEEE International Conference on Data Mining (ICDM) 2018.

- Obot, N., O'Malley, L., Nwogu, I., Yu, Q., Shi, Weishi., & Guo, X. "From Novice to Expert Narratives of Dermatological Disease". IEEE International Conference on Pervasive Computing and Communications Workshops 2018.

- Shi, Weishi., Liu, X., & Yu, Q. "Correlation-Aware Multi-Label Active Learning for Web Service Tag Recommendation. In 2017 IEEE International Conference on Web Services" (ICWS) 2017.

- Liu, X., Shi, Weishi., Kale, A., Ding, C., & Yu, Q. "Statistical Learning of Domain-Specific Quality-of-Service Features from User Reviews." ACM Transactions on Internet Technology (TOIT) 2017.

RESEARCH PROJECTS

A Gaussian Process-Bayesian Bernoulli Mixture Model for Multi-Label Active Learning
We propose a novel integrated Gaussian Process-Bayesian Bernoulli Mixture model (GP-B^{2}M) and a principled

sampling function for multi-label classification active learning. The proposed method could accurately quantifies a data sample's overall contribution to a correlated label space and choose the most informative samples for cost-effective annotation. **Multifaceted Uncertainty Estimation for Label-Efficient Deep Learning**
We propose a multi-source uncertainty prediction approach that enables deep learning (DL) models to be actively trained with 30% labelled data instances as required by passive learning.

An Efficient Many-Class Active Learning Framework for Knowledge-Rich Domains
We propose an active learning framework to reduce the experts' data annotation costs using real-world text data extracted from dermatology physician's narratives.

Correlation-Aware Multi-Label Active Learning for Web Service Tag Recommendation
This work focus on multi-label active learning for web service tag recommendation. We reduce the expert's labeling efforts by learning and leveraging the correlations among tags.

Statistical Learning of Domain-Specific Quality-of-Service Features from User Reviews
A statistical learning approach is developed to extract domain-specific QoS features from user-provided web service reviews. The work automatically groups relevant terms and returns the results to web service users.

Multi-level Tag Recommendation with External Knowledge for Auto-Question Answering
We propose a multi-level post-tag co-clustering system to cluster unstructured data form a Q&A website (i.e., Stack Overflow). The proposed system recommends semantic-rich tags to a given post .

Fast Direct Search in Optimally Compressed Target Space for Efficient Multi-Label Active Learning
A novel framework that integrates compressed sensing and Bayesian principal component analysis is proposed for active learning. The proposed work reduce the human experts' data annotation effort for building multi-label machine learning models.We use Bayesian OPT to achieve 3-5X faster model selection speed.

Integrating Generative and Discriminative Sparse Kernel Machines for Multi-class Active Learning
We proposes a novel machine learning model, Kernel Machine Committee (KMC), that integrates generative and discriminative kernel machines for fast and accurate multi-class active learning. The proposed work reduce the human experts' data annotation effort for building multi-class machine learning models.

Presenting & Evaluating the Impact of Experiential Learning in Computing Accessibility Education
In this work we design and practice an educational component for computer science students to improve their in-class learning performance on computing accessibility.

Skills

PROGRAMMING LANGUAGE:	Python, R, JAVA.
ANALYTICAL TOOLS & IDEs:	Scikitlearn, numpy, scipy, pandas, cvxopt, pytorch, Matlab, SQL
DATABASES:	Mongo DB, SQLServer, MySQL

Working Experience

AUGUST 2016 –NOW	ROCHESTER INSTITUTE OF TECHNOLOGY *Research Assistant*
JUN 2017 –AUG 2017	ROCHESTER INSTITUTE OF TECHNOLOGY GCCIS:ISTE.600 *Lecturer*
JUN 2019 –AUG 2019	ROCHESTER INSTITUTE OF TECHNOLOGY GCCIS:ISTE.612 *Lecturer*
SPRING 2015 –MAY 2019	ROCHESTER INSTITUTE OF TECHNOLOGY GCCIS: ISTE.612, ISTE.600 *Teaching Assistant* Grading and tutoring.
JUL 2012 –OCT 2013	K-BOXING MEN'S WEAR (SHANGHAI, CHINA) CO., LTD. *Supervisor* Manage SAP system for human resource department.
MAY 2010 –NOV 2011	TNT LOGISTICS (SHANGHAI, CHINA) *Technical Support Supervisor* Develop functions of human resource information system; Launch System to 56 companies; Train system users.