# University of Pittsburgh School of Medicine CURRICULUM VITAE

**YE, YE**

# BIOGRAPHICAL

|  |  |
| --- | --- |
| Name: Ye, Ye | Business Address: 5607 Baum Blvd, 437J, Pittsburgh, PA, 15206 |
| Email: yey5@pitt.edu | Business Phone: 412-925-2746 |

# EDUCATION and TRAINING

|  |  |  |  |
| --- | --- | --- | --- |
| **UNDERGRADUATE** |  |  |  |
| Dates Attended | Name and Location of Institution | Degree Received and Year | Major Subject |
| 09/2001-07/2006 | Peking University  Health Science Center, Beijing | Bachelor of Medicine (MD equivalent), 2006 | Preventive Medicine |
| **GRADUATE** |  |  |  |
| Dates Attended | Name and Location of Institution | Degree Received and Year | Major Subject |
| 09/2006-01/2009 | Peking University,  School of Public Health, Beijing | Master of Science, 2009 | Epidemiology and Health Statistics |
|  |  |  |  |
| 08/2009-05/2011 | Emory University,  School of Public Health, Atlanta | Master of Science in Public Health, 2011 | Public Health Informatics |
|  |  |  |  |
| 08/2011-12/2018 | University of Pittsburgh,  School of Computing and Information, Pittsburgh | Doctor of Philosophy, 2018 | Intelligent Systems |
|  |  |  |  |

**POSTGRADUATE**

|  |  |  |  |
| --- | --- | --- | --- |
| Dates Attended | Name and Location of Institution | Degree Received and Year | Major Subject |
| 03/2004-  06/2005 | Beijing Shijitan Hospital, Beijing | Non-degree | Medical Internship |
| 06/2006-  09/2006 | Haidian CDC, Beijing | Non-degree | Local Public Health Department Internship |
| 03/2006-  06/2007 | School of Mathematics, Peking University, Beijing | Non-degree | Finished Mathematics and Statistics Courses (Mathematical Analysis, Mathematical Statistics, Ordinary Differential Equations) |
| 06/2010-  08/2010 | CDC, Atlanta | Non-degree | Public Health Informatics Internship |
| 09/2014-  12/2014,  09/2020-  05/2021 | Carnegie Mellon University, Pittsburgh | Non-degree | Finished Machine Learning Courses  (Intermediate Statistics, Intermediate Deep Learning, and Convex Optimization) |
| 02/2019-  04/2022 | Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh |  | Postdoctoral Associate,  Mentors: Drs. Michael Becich (primary), Michael Wagner, Gregory Cooper |
|  |  |  |  |

# APPOINTMENTS and POSITIONS

**ACADEMIC**

|  |  |  |
| --- | --- | --- |
| Years Inclusive | Name and Location of Institution or Organization | Rank/Title |
| 2006 - 2008 | Department of Epidemiology & Health Statistics, Peking University, Beijing | Research Assistant |
| 2010 - 2011 | Department of Biostatistics & Bioinformatics, Emory University, Atlanta | Research Assistant |
| 2011 - 2017 | Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh | Graduate Student Researcher |
| 02/2019-04/2022 | Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh | Postdoc Associate |
| 05/2022 – present | Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh | Tenure-track Assistant Professor |

**NON-ACADEMIC**

NA

# CERTIFICATION and LICENSURE

**SPECIALTY CERTIFICATION:**

NA

**MEDICAL or OTHER PROFESSIONAL LICENSURE:**

NA

# MEMBERSHIP in PROFESSIONAL and SCIENTIFIC SOCIETIES

|  |  |
| --- | --- |
| Organization | Year |
| American Medical Informatics Association | 2010 – present |
| Observational Health Data Sciences and Informatics Program | 2016 - Present |
| Association for the Advancement of Artificial Intelligence | 2018 – 2019 |
| Association for Pathology Informatics | 2022 – present |

# HONORS

|  |  |
| --- | --- |
| Title of Award | Year |
| Recommended to Delta Omega Honor Society (a U.S. national honorary society in public health, founded in 1924) | 2011 |
| Finalist in Michael G. Wells Student Health Care Entrepreneurship Competition (University of Pittsburgh) | 2013 |
| CDC Best Statistical Science Theoretical Paper (3rd author) | 2014 |
| Distinguished Poster Award (2nd author) in American Medical Informatics Association Annual Symposium | 2014 |
| Best Student Paper 2nd Place (U Pittsburgh, Department of Biomedical Informatics) | 2014 |
| Distinguished Poster Award (3rd author) in American Medical Informatics Association Annual Symposium | 2015 |
| University of Pittsburgh Innovator Award | April 25, 2016 |
| University of Pittsburgh Arts and Sciences Fellowship | 2016 – 2017 |
| Best Student Paper 2nd Place (U Pittsburgh, Department of Biomedical Informatics) | 2017 |
| Andrew Mellon Fellowship (University of Pittsburgh) | 2017 – 2018 |
| Doctoral Dissertation Award Finalist (American Medical Informatics Association) | 2019 |

# PUBLICATIONS

1. ORIGINAL PEER REVIEWED ARTICLES
2. Rexit R, Tsui F, Espino J, Wesaratchakit S, **Ye Y**, Chrysanthis PK. Using a distributed search engine to identify optimal product sets for use in an outbreak detection system. In8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom) 2012 Oct 14 (pp. 560-566). IEEE.
3. Tang L, Lyles RH, **Ye Y**, Lo Y, King CC. Extended Matrix and Inverse Matrix Methods Utilizing Internal Validation Data When Both Disease and Exposure Status Are Misclassified. Epidemiol Methods. 2013 Sep 1;2(1):49-66. doi: 10.1515/em-2013-0008. PMID: 25844304; PMCID: PMC4382468.
4. Lu S, **Ye Y**, Tsui R, Su H, Rexit R, Wesaratchakit S, Liu X, Hwa R. Domain ontology-based feature reduction for high dimensional drug data and its application to 30-day heart failure readmission prediction. In9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing 2013 Oct 20 (pp. 478-484). IEEE.
5. Lu S, **Ye Y**, Tsui R, Liu X, Hwa R. Feature selection for 30-day heart failure readmission prediction using clinical drug data. NIPS Workshop on Machine Learning for Clinical Data Analysis and Healthcare, Harrahs and Harveys, Lake Tahoe, 2013.
6. **Ye Y**, Tsui FR, Wagner M, Espino JU, Li Q. Influenza detection from emergency department reports using natural language processing and Bayesian network classifiers. J Am Med Inform Assoc. 2014 Sep-Oct;21(5):815-23. doi: 10.1136/amiajnl-2013-001934. Epub 2014 Jan 9. PMID: 24406261; PMCID: PMC4147621.
7. Rexit R, Tsui FR, Espino J, Chrysanthis PK, Wesaratchakit S, **Ye Y**. An analytics appliance for identifying (near) optimal over-the-counter medicine products as health indicators for influenza surveillance. Information Systems. 2015 Mar 1;48:151-63.
8. López Pineda A, **Ye Y**, Visweswaran S, Cooper GF, Wagner MM, Tsui FR. Comparison of machine learning classifiers for influenza detection from emergency department free-text reports. J Biomed Inform. 2015 Dec;58:60-69. doi: 10.1016/j.jbi.2015.08.019. Epub 2015 Sep 16. PMID: 26385375; PMCID: PMC4684714.
9. **Ye Y**, Wagner MM, Cooper GF, Ferraro JP, Su H, Gesteland PH, Haug PJ, Millett NE, Aronis JM, Nowalk AJ, Ruiz VM, López Pineda A, Shi L, Van Bree R, Ginter T, Tsui F. A study of the transferability of influenza case detection systems between two large healthcare systems. PLoS One. 2017 Apr 5;12(4):e0174970. doi: 10.1371/journal.pone.0174970. PMID: 28380048; PMCID: PMC5381795.
10. Ferraro JP, **Ye Y**, Gesteland PH, Haug PJ, Tsui FR, Cooper GF, Van Bree R, Ginter T, Nowalk AJ, Wagner M. The effects of natural language processing on cross-institutional portability of influenza case detection for disease surveillance. Appl Clin Inform. 2017 May 31;8(2):560-580. doi: 10.4338/ACI-2016-12-RA-0211. PMID: 28561130; PMCID: PMC6241736.
11. Aronis JM, Millett NE, Wagner MM, Tsui F, **Ye Y**, Ferraro JP, Haug PJ, Gesteland PH, Cooper GF. A Bayesian system to detect and characterize overlapping outbreaks. J Biomed Inform. 2017 Sep;73:171-181. doi: 10.1016/j.jbi.2017.08.003. Epub 2017 Aug 7. PMID: 28797710; PMCID: PMC5604259.
12. Tsui F, **Ye Y**, Ruiz V, Cooper GF, Wagner MM. Automated influenza case detection for public health surveillance and clinical diagnosis using dynamic influenza prevalence method. J Public Health (Oxf). 2018 Dec 1;40(4):878-885. doi: 10.1093/pubmed/fdx141. PMID: 29059331; PMCID: PMC6676953.
13. Millett NE, Aronis JM, Wagner MM, Tsui F, **Ye Y**, Ferraro JP, Haug PJ, Gesteland PH, Cooper GF. The design and evaluation of a Bayesian system for detecting and characterizing outbreaks of influenza. Online J Public Health Inform. 2019 Sep 19;11(2):e6. doi: 10.5210/ojphi.v11i2.9952. PMID: 31632600; PMCID: PMC6788888.
14. Aronis JM, Ferraro JP, Gesteland PH, Tsui F, **Ye Y**, Wagner MM, Cooper GF. A Bayesian approach for detecting a disease that is not being modeled. PLoS One. 2020 Feb 28;15(2):e0229658. doi: 10.1371/journal.pone.0229658. PMID: 32109254; PMCID: PMC7048291.
15. Jackson BR, **Ye Y**, Crawford JM, Becich MJ, Roy S, Botkin JR, de Baca ME, Pantanowitz L. The Ethics of Artificial Intelligence in Pathology and Laboratory Medicine: Principles and Practice. Acad Pathol. 2021 Feb 16;8:2374289521990784. doi: 10.1177/2374289521990784. PMID: 33644301; PMCID: PMC7894680.
16. Bernstam EV, Shireman PK, Meric-Bernstam F, N Zozus M, Jiang X, Brimhall BB, Windham AK, Schmidt S, Visweswaran S, **Ye Y**, Goodrum H, Ling Y, Barapatre S, Becich MJ. Artificial intelligence in clinical and translational science: Successes, challenges and opportunities. Clin Transl Sci. 2022 Feb;15(2):309-321. doi: 10.1111/cts.13175. Epub 2021 Oct 30. PMID: 34706145; PMCID: PMC8841416.
17. **Ye Y**, Barapatre S, Davis MK, Elliston KO, Davatzikos C, Fedorov A, Fillion-Robin JC, Foster I, Gilbertson JR, Lasso A, Miller JV, Morgan M, Pieper S, Raumann BE, Sarachan BD, Savova G, Silverstein JC, Taylor DP, Zelnis JB, Zhang GQ, Cuticchia J, Becich MJ. Open-source Software Sustainability Models: Initial White Paper from the Informatics Technology for Cancer Research Sustainability and Industry Partnership Working Group. J Med Internet Res. 2021 Dec 2;23(12):e20028. doi: 10.2196/20028. PMID: 34860667; PMCID: PMC8686402.
18. Hartman D, Le Douget JE, **Ye Y**, Li Y, Sin-Chan P, Pronier E, Becich M. Application of deep learning models on whole slide images uncover new histological markers related to high-risk malignant pleural mesothelioma. 2022 ASCO Annual Meeting. Journal of Clinical Oncology 40, no. 16\_suppl (June 01, 2022) e13580-e13580.
19. Gao Y, Mazurek JM, Li Y, Blackley D, Weissman DN, Burton SV, Amin W, Landsittel D, Becich MJ, **Ye Y\***. Industry, occupation, and exposure history of mesothelioma patients in the US National Mesothelioma Virtual Bank, 2006–2022. Environmental Research. 2023 Aug 1;230:115085.
20. Barapatre S, Gao Y, Becich MJ, Chandran UR, Amin W\*, Li Y, **Ye Y\***. Multiple institutions’ research findings using the National Mesothelioma Virtual Bank. F1000Research. 2023 Sep 11;11:1343. (Corresponding author)
21. OTHER PEER REVIEWED PUBLICATIONS
22. Ji Y, Gao Y, Bao R, Li Q, Liu D, Sun Y, **Ye Y\***. Prediction of COVID-19 Patients' Emergency Room Revisit using Multi-Source Transfer Learning. https://arxiv.org/abs/2306.17257. 2023 Jun 29. (Accepted and presented in IEEE ICHI) (Corresponding author)
23. OTHER NON-PEER REVIEWED PUBLICATIONS
24. **Ye Y**. An evaluation plan for a pilot of CDC EMR alerting service prototype. Master’s thesis, Emory University. 2011.
25. **Ye Y**. Transfer Learning for Bayesian Case Detection Systems. Doctoral dissertation, University of Pittsburgh. 2018.
26. **Ye Y\***, Gu A. Deep Transfer Learning for Infectious Disease Case Detection Using Electronic Medical Records. arXiv preprint arXiv:2103.06710. 2021 Mar 8. (Corresponding author)
27. Bao R, Sun Y, Gao Y, Wang J, Yang Q, Chen H, Mao Z, **Ye Y\***. A Survey of Heterogeneous Transfer Learning. 2023. <https://arxiv.org/abs/2310.08459> (Corresponding author) Under Review by IEEE Transactions on Neural Networks and Learning Systems.
28. Aronis JM, **Ye Y**, Espino J, Hochheiser H, Michaels MG, Cooper GF. A Bayesian System to Track Outbreaks of Influenza-Like Illnesses Including Novel Diseases. medRxiv. 2023:2023-05. https://www.medrxiv.org/content/10.1101/2023.05.10.23289799v1
29. BOOKS, BOOK CHAPTERS AND MONOGRAPHS
30. PUBLISHED ABSTRACTS (in Scientific Journals)
31. Aronis JM, Millett NE, Wagner MM, Tsui F, **Ye Y**, Cooper GF. A method for detecting and characterizing multiple outbreaks of infectious diseases. Online Journal of Public Health Informatics. 2016 Mar 24;8(1).
32. Aronis JM, Millett NE, Wagner MM, Tsui F, **Ye Y**, Ferraro J, Haug PJ, Cooper GF. Detecting overlapping outbreaks of influenza. Online Journal of Public Health Informatics. 2017;9(1).
33. Douglas H, Le Douget JE, **Ye Y**, Li Y, Sin-Chan P, Pronier E, Becich M. Application of deep learning models on whole slide images uncover new histological markers related to high-risk malignant pleural mesothelioma. 2022 ASCO Annual Meeting I. Journal of Clinical Oncology. Published online before print June 2, 2022. DOI: 10.1200/JCO.2022.40.16\_suppl.e13580.
34. ABSTRACTS (not published in Scientific Journals)
35. Draper AJ, **Ye Y**, Ruiz VM. Using laboratory data for prediction of 30-day hospital readmission of pediatric seizure patients. 2014 AMIA, Washington, DC.
36. Ruiz VM, **Ye Y**, Draper A. The use of multiple emergency department reports per visit for improving the accuracy of influenza case detection. 2014 AMIA, Washington, DC.
37. Ruiz VM, Draper AJ, **Ye Y**. Use of diagnosis-related groups to predict all-cause pediatric hospital readmission within 30 days after discharge. 2015 AMIA, San Francisco, CA.
38. Posada J, Shi L, **Ye Y**. Use of free-text clinical reports for prediction of 30-day psychiatric readmissions. 2016 AMIA, Chicago, IL.
39. Tsui F, Ruiz V, Barda A, **Ye Y**, Butler G, Suresh S, Urbach A, Retrospective and prospective evaluations of the system for hospital adaptive readmission prediction and management (SHARP) for all-cause 30-Day pediatric readmission prediction. *2017 AMIA* paper abstract.
40. **Ye Y**, Wagner MM, Cooper GF, Tsui FC, Weiss J, Gesteland PH, Jeffrey Ferraro JP, Haug PJ. Bayesian network transfer learning to improve re-usability of computable biomedical knowledge for public health. Poster abstract accepted by Mobilizing Computable Biomedical Knowledge 2nd Annual Meeting in National Institutes of Health - Bethesda, MD, July 18-19, 2019.
41. Aronis JM, **Ye Y**, Cooper GF. A Bayesian System to Detect and Track Multiple Diseases from Electronic Medical Records. 07-09 September 2022. MIDAS annual conference.
42. **Ye Y**\*, Li Q, Gu A, Ren A, Liu D, Gao Y, Jia J, et al. An Empirical Analysis of Deep Transfer Learning for Infectious Disease Case Detection. Nov 2023. AMIA podium abstract.

# PROFESSIONAL ACTIVITIES

**TEACHING**

**Medical Student Teaching:**

NA

**Graduate Student Teaching:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Dates | Role | Title of teaching session(s) | type of learner | number of learners | number of sessions/year |
| 05/2007-07/2007 | Teaching Assistant | Multivariable Statistical Methods | MPH students | 20+ | School of Public Health, Peking University, 2007 |
| 08/2016-12/2016 | Teaching Assistant | Foundations of Clinical and Public Health Informatics | Graduate Students | About 10 | Department of Biomedical Informatics, University of Pittsburgh, 2016 |
| 06/08/2021 | Lecturer | Applying for a K Grant: Brief Introduction and Tips | Graduate Students and Postdoctoral Fellows | About 10 | Department of Biomedical Informatics, University of Pittsburgh, 2022 |
| 11/2022, 11/2023 | Lecturer | One Lecture about Public Health Informatics (Foundations of Clinical and Public Health Informatics Course) | Graduate Students | Range from 5 to 9 | Department of Biomedical Informatics, University of Pittsburgh |
| 01/2023-04/2023 | Lecturer | Biomedical Informatics Journal Club | Graduate Students | 4 | Department of Biomedical Informatics, University of Pittsburgh, 2023 Spring |
| 11/2023 | Lecturer | Two lectures about Deep Learning | Graduate Students | 9 | Department of Biomedical Informatics, University of Pittsburgh |

**Resident Teaching:**

NA

**Fellow Teaching:**

NA

**Faculty Development Teaching:**

NA

**Curriculum Development/Teaching Products/Media Products:**

NA

**Mentoring:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dates | Mentor Role | Name of Mentee | Context | Mentee Achievements |
| 08/2020-05/2022 | Research Mentor | Seemran Barapatre  (Biology Undergraduate, U Pittsburgh) | Mesothelioma Virtual Bank Research Finding Review, Open-source Software Sustainability Models Review, Artificial Intelligence in Clinical and Translational Science | Coauthored two peer-reviewed articles and one non-peer review article, graduated with a bachelor’s degree |
| 01/2022-  Present | Research Mentor | Yuhe Gao  (Information Science, Master, U Pittsburgh; now Data Scientist) | Transfer Learning,  National Mesothelioma Virtual Bank Occupational History Analysis | Coauthored two peer-reviewed articles and one non-peer review article, graduated with a master’s degree |
| 01/2022-  04/2022 | Research Mentor | Junshang Jia  (Computer Science, Undergraduate, U Pittsburgh) | Transfer Learning | Released one open-source software package and coauthored one peer-reviewed article, graduated with a bachelor’s degree |
| 01/2022-  04/2022 | Research Mentor | Disheng Liu  (Information Science, Master, U Pittsburgh) | Transfer Learning | Released one open-source software package, graduated with a master’s degree |
| 05/2022-  12/2022,  08/2023-12/2023 | Research Mentor | Yuelyu Ji  (Information Science, Master, U Pittsburgh; now PhD student) | Transfer Learning,  Large Language Models | coauthored one peer-reviewed manuscript, graduated with a master’s degree, admitted to a PhD program |
| 08/2022-  12/2022 | Research Mentor | Runxue Bao  (School of Engineering, PhD, U Pittsburgh) | Transfer Learning | coauthored one peer-reviewed manuscript, graduated with a PhD degree |
| 01/2023-12/2023 | Research Mentor | John Song  (Information Science, Master, U Pittsburgh) | Bayesian network transfer learning | Finished one open-source software package, graduate with a master’s degree in December |
| 01/2023-  present | Research Mentor | Yiming Sun  (School of Engineering, PhD program, U Pittsburgh) | Transfer Learning, Online learning | coauthored one peer-reviewed manuscript, coauthored one manuscript that is under review |

**RESEARCH**

**Current Grant Support:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grant Number (funded) | Grant Title | Role in Project    %Effort (must not  exceed 100%)    Calendar  Months (must not exceed 12 months | Years Inclusive | Source $  Amount    (include direct and indirect if  PI)    (If Co-I list subaward) |
| R00LM013383 | Transfer Learning to Improve the Re-usability of Computable Biomedical Knowledge | PI (75%)  2023 8.28 months  2024 7.89 months | 05/2022 – 04/2025 | Total Award Amount (including Indirect Costs): $747,000 |
| R01LM013509 | Automated Surveillance of Overlapping Outbreaks and New Outbreak Diseases | Co-I  2023 2.4 months  2024 2.4 months | 08/2022 – 07/2025 | Total Award Amount (including Indirect Costs): $1,008,755 |
| CDC U24 OH009077 | National Mesothelioma Virtual Bank | Researcher  2023 1.2 months  2024 1.2 months  2025 0.6  months | 09/2021 – 08/2026 | Total Award Amount (including Indirect Costs): $5,408,682 |

**Pending Grant Support** (Include status, date reviewed, and priority score)

NA

**Prior Grant Support:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grant Number (funded) | Grant Title | Role in Project    %Effort (must not exceed  100%)    Calendar Months (must not exceed 12 months | Years Inclusive | Source $  Amount    (include direct and indirect if  PI)    (If Co-I list subaward) |
| K99LM013383 | Transfer Learning to Improve the Re-usability of Computable Biomedical Knowledge | PI (100%) | May 2020-April 2022 | Year 2021  (Direct Costs: $86,650; Indirect Costs: $6,692)  Year 2020 (Direct Costs: $86,925; Indirect Costs: $5,434) |
| R01LM011370 | [Probabilistic Disease](https://www.google.com/search?es_sm=122&q=Probabilistic+Influenza+Detection&spell=1&sa=X&ei=DMr4VLX-KYzFggSewYCAAg&ved=0CBsQvwUoAA) Surveillance | Graduate Student Researcher | August 2013 – August 2017 |  |

**Non-Funded Research:**

|  |  |  |
| --- | --- | --- |
| Project Title | Role in Project | Years Inclusive |
| Mother’s and Children’s Health Status and Their Health Service Needs Study | Undergraduate Student Researcher | 07/2003-08/2003 |
| Women’s and Children’s Health Status Assessment Study | Undergraduate Student Researcher | 07/2005-08/2005 |
| Treatment of alopecia areata patients with Minoxidil Spray  (Advisor: Dr. Hongyuan Wang, Peking University) | Undergraduate Student Researcher | 02/2006-05/2006 |
| Antagonistic effect of chalybeate on lead-induced DNA injury in rat’s brain cells  (Advisor: Xuetao Wei, Peking University) | Undergraduate Student Researcher | 05/2006-06/2006 |
| Analysis of epidemiologic characteristics of Scarlet Fever in Beijing  (Advisor: Yanhui Shen, haidian CDC) | Undergraduate Student Researcher | 06/2006-09/2006 |
| ICD-9-CM Chinese version  (Advisor: Dr. Hongyuan Wang, Peking University) | Graduate Student Researcher | 09/2006-01/2009 |
| Beijing Diagnosis Related Groups/Prepayment System (DRG/PPS) program  (Advisor: Dr. Hongyuan Wang, Peking University) | Graduate Student Researcher | 09/2006-01/2009 |
| A simple computation method to estimate fraction of last year for infants  (Advisor: Dr. Hongyuan Wang, Peking University) | Graduate Student Researcher | 09/2006-01/2009 |
| Actionable Public Health Alerts for Electronic Medical Record Systems  (Advisor: Dr. Vicki Hertzberg, Emory University; Field Advisor: Nedra Y. Garrett, CDC) | Research Assistant | 06/2010-05/2011 |
| Combining Internal and External Validation Data to Correct for Misclassifications  (Advisor: Dr. Robert Lyles, Emory University) | Research Assistant | 09/2010-05/2011 |
| System for Hospital Adaptive Readmission Prediction and Management  (Advisor: Dr. Rich Tsui, U Pittsburgh) | Research Assistant | 01/2013-03/2017 |

Other research related activities

NA

**LIST of CURRENT RESEARCH INTERESTS**

**1. Transfer Learning**

As more and more machine-learned models are generated, the biggest challenge is that a model developed in one healthcare system (denoted as source) may be expected to underperform in another healthcare system (target). I developed an innovative transfer learning framework to enable sharing of both data and models between source and target. My research enables knowledge-sharing under heterogeneous scenarios and provides an approach for understanding transfer learning performance between source and target in terms of differences of features and their distributions and sample sizes. The model-transfer algorithm can be viewed as a new Bayesian network learning algorithm that has a flexible representation of prior knowledge. I am now continuing research on Bayesian and deep learning approaches to improve the re-usability of models. Re-using models is expected to benefit the public’s health by: (1) improving case detection during epidemics by enabling re-use of automatic case detectors developed in the earliest affected regions with other regions, and, more generally, (2) increasing the impact of NIH’s investment in machine learning by enabling machine-learned models to be used in more institutions and locations.

**2. Public Health Surveillance**

Traditionally, public health surveillance relies mainly on sentinel physician reporting and laboratory reporting, which usually result in reporting delays and issues due to underreporting and undertesting. I have developed a near real-time, active surveillance approach that uses machine-learned case detection systems to automatically capture cases from electronic medical records. A case detection system consists of a natural language processing parser (NLP) and a Bayesian network classifier. It uses NLP to infer the presence or absence of clinical findings from narrative notes. With these findings, a Bayesian network classifier infers each patient’s diagnosis probabilities, as well as the likelihood of patient clinical evidence, to support outbreak detection and forecast at the population-level. My case detection systems enhance the communications between clinicians and public health officials, automatically inferring patients’ diagnoses from free-text clinical notes for population monitoring and automatically providing population prevalence information to support clinical decisions on differential diagnoses. My research shows the impact of accurate natural language processing and feature selection on classification performance and demonstrates the advantages of automated predictive modeling versus experts’ simple judgement with regards to making correlations between NLP-parsed clinical findings and disease status. Regarding model parameterization, my experiments also show that using a changing prevalence of disease could increase the discriminative ability of an influenza detection model compared to using a constant prevalence. Moreover, I demonstrated high performance for influenza detection in a five-year retrospective study in both Allegheny County and Salt Lake County.

**3. Automated Readmission Risk Profiling**

Unplanned hospital readmissions are placing an unbearable financial burden on the country. We developed readmission risk assessment models for heart failure for adults and children by analyzing unstructured data (e.g., discharge reports) and structured data (e.g., laboratory orders and results, drug orders, healthcare utility information) stored in hospital electronic medical record systems. We developed a domain ontology-based feature selection algorithm that performs better than feature selection algorithms without ontology knowledge. Our work on automated readmission risk profiling received Distinguished Poster Awards at the 2014 and 2015 American Medical Informatics Annual Symposium. In addition, three main personnel, including myself, received an innovator award from the University of Pittsburgh for this work. Furthermore, our pediatric readmission prediction model has been integrated into the Cerner® Millennium electronic medical record system at the Children’s Hospital of Pittsburgh of UPMC. Both retrospective evaluation and prospective evaluation indicate that the system has the potential to identify patients with elevated risk of readmission, thereby allowing more informed and more successful discharge planning and intervention.

**4. Challenges, Opportunities of Artificial Intelligence in Biomedicine**

Growing numbers of artificial intelligence applications are being developed and applied to biomedicine. These technologies introduce risks and benefits that must be assessed and managed. In the first article, we discussed how long-standing principles of medical and scientific ethics can be applied to artificial intelligence. In the second article, we focus on the role of AI in clinical and translational research (CTR), including preclinical research, clinical research, clinical implementation, and public (or population) health. For each CTR phase, we addressed challenges, successes, failures, and opportunities for AI. We present three complementary perspectives: (1) scoping literature review, (2) survey, and (3) analysis of federally funded projects. In this third article, We reviewed the existing sustainability models for open-source software (OSS) and describes 10 OSS use cases, including 3D Slicer, Bioconductor, Cytoscape, Globus, i2b2 (Informatics for Integrating Biology and the Bedside) and tranSMART, Insight Toolkit, Linux, Observational Health Data Sciences and Informatics tools, R, and REDCap (Research Electronic Data Capture), in 10 sustainability aspects: governance, documentation, code quality, support, ecosystem collaboration, security, legal, finance, marketing, and dependency hygiene.

**INVITED SEMINARS AND LECTURESHIPS**

**SERVICE**

University and Medical School Service

NA

Diversity, Equity, and Inclusion Activities

|  |  |  |
| --- | --- | --- |
| Date | Role | Context |
| 10/2022 | Faculty interviewer | interviewed one candidate for the MSTP program. The candidate is from underrepresented and disadvantaged background. |
| 12/2022 | Self-educate | educated myself on DEI by browsing NIH webpages <https://extramural-diversity.nih.gov/diversity-matters> |

Local/Community Service Activities

|  |  |  |
| --- | --- | --- |
| Date | Role | Context |
| 10/2022 | consultant | I supported Yingci Liu’s application of an internal research grant in Rutgers. Yingci is an alumnus from DBMI and now an Assistant Professor at School of Dental Medicine, Rutgers. I met with Yingci, reviewed her proposal, provided feedbacks, and a supporting letter. |

Regional Service

NA

National Service

NA

International Service

|  |  |  |
| --- | --- | --- |
| Date | Role | Context |
| 2014, 2015, 2017 | Reviewer | American Medical Informatics Association Annual Symposium |
| 2017 - Present | Reviewer | International Journal of Medical Informatics |
| 2017 - Present | Reviewer | Journal of Public Health Management and Practice |
| 2017 - Present | Reviewer, Editorial Board | Artificial Intelligence in Medicine |
| 2017 - Present | Reviewer | Computers in Biology and Medicine |
| 2017 - Present | Reviewer | Journal of Innovation in Health Informatics |
| 2024 |  | Artificial Intelligence in Medicine |
| 2018 | Editorial Board. | Computers in Biology and Medicine |
| 2018 - Present | Reviewer | The BMJ |
| 2018 - Present | Reviewer | Online Journal of Public Health Informatics |
| 2020 - Present | Reviewer | Journal of the American Medical Informatics Association |
| 2020 – Present | Reviewer | Journal of Biomedical Informatics |
| 2021 - Present | Reviewer | Journal of Medical Internet Research |
| 2022 | Program Committee | The Eighth International Workshop on Data Management and Analytics for Medicine and Healthcare |
| 2024-2027 | Editorial Board | Journal of Biomedical Informatics |