

# CURRICULUM VITAE - ASHRAF YASEEN

Updated: October 12, 2024

CONTACT INFORMATION Associate Professor of Data Science  
Center for Big Data in Health Sciences  
Department of Biostatistics and Data Science Office Phone: (713) 500-9583  
School of Public Health. University of Texas Cell Phone: (757) 609-1793  
Health Science Center at Houston. E-mail: ashraf.yaseen@uth.tmc.edu  
1200 Pressler, RAS E807 Website:  
Houston, TX 77030-3900 <https://sites.google.com/site/ashrafsyaseen/>

## A. Background

1. EDUCATION **Old Dominion University (ODU)**, Norfolk, VA  
**PhD., Computer Science**, Dec 2014  
Dissertation: *Improving Structural Features Prediction in Protein Structure Modeling using Deep Learning*. Advisor: Dr. Yaohang Li

**New York Institute of Technology (NYIT)**  
**M.S., Computer Science**, August 2003  
With Distinction  
Research project: *Multithreaded Heuristic Search* - Multithreaded programming techniques to enhance the performance of heuristic searching algorithms

**Jordan University of Science and Technology (JUST)**  
**B.S., Computer Science & Information Systems**, July 2002  
On the Honors List of the School of Information Technology  
Project: *E-Auto Insurance System* – A database application using Oracle  
Research: *Distributed Heuristic Search* – multi-process programming on distributed systems for searching algorithms

2. PROFESSIONAL EXPERIENCE **University of Texas Health Science Center (UTHealth)** – Houston, TX 2023-present  
**Center for Big Data in Health Sciences**  
**Associate Professor of Data Science (Tenured)**

**UTHealth** – Houston, TX 2018-2023  
**Center for Big Data in Health Sciences**  
**Assistant Professor of Data Science**

- Research in Data Analysis, Data Management, Machine Learning, Big Data, and High-Performance Computing
- Develop and teach courses: Introduction to Statistical and Data Science Programming, Fundamentals of Data Analytics and Predictions, Data Science Computing, Machine Learning in Practice
- Advise students (academic and research thesis advisor)
- Coordinator of the Data Science program & certificates
- Collaborate with researchers in academia and industry
- Lead a team of programmers and data managers
- Develop curricula, serve on departmental, school, and university committees, and perform scholarly and service activities

2014-2018

**Texas A&M University Kingsville (TAMUK) – Kingsville, TX**

**Assistant Professor of Computer Science**

**Director of the Computational Sciences Lab**

- Develop and teach courses: Introduction to Bioinformatics, Bioinformatics Computing, Database Systems, Cloud Computing, Operating Systems, Computer Communication Networks
- Research in Bioinformatics, Machine Learning, Security Engineering, High Performance Computing
- Advise students (graduate research project and thesis advisor)
- Develop curricula, serve on departmental, college, and university committees, and perform scholarly and service activities

**Central State University – Wilberforce, OH**

2013-2014

**Assistant Professor of Computer Science**

- Teach courses: Computer Programming in C++, Database Systems, Contemporary Operating Systems, Computer Networks, Computer Architecture
- Advise students, develop curricula, serve on departmental, college, and university committees, and perform scholarly and service activities

**Old Dominion University - Norfolk, VA**

2010-2013

**Research Assistant**

- Exploring machine learning methods in Protein modeling
- Using GPUs in Bioinformatics

**Old Dominion University - Norfolk, VA**

2007-2010

**Teaching Assistant and Lab Instructor** for Problem Solving and Programming

**Jordan University of Science & Technology - Irbid, Jordan**

2003-2007

**Lecturer of Computer Information Systems**

- Teach courses: Knowledge-based Systems, Database Applications using Oracle, Programming in C++, Programming in Visual Basic
- Senior project advisor
- Advise students, develop curricula, serve on departmental, college, and university committees, and perform scholarly and service activities

**Jordan University of Science & Technology - Irbid, Jordan**

2002-2003

**Teaching Assistant:** Introduction to Computer Science

**Al-Najjar Center - Irbid, Jordan**

**Database Developer:** Design and implement database applications using Oracle, train on using Oracle (Part-time) 2001-2002

**Creative Systems - Irbid, Jordan**

2000-2001

**Computer Programmer** using Visual Basic (Part-time)

---

### 3. AWARDS

- ❑ Professional Development Award. Center for Teaching Effectiveness, New Faculty Investment Program, TAMUK, 2016.
  - ❑ Summer Research Award. Office of Research and Sponsored Programs, TAMUK, 2016.
-

---

## B. Research

---

### RESEARCH INTERESTS

- Data Management, Data Analysis, Machine Learning, and Big Data
  - Bioinformatics
  - High Performance Computing
- 

### 1. PUBLICATIONS

† = first/co-first author

§ = corresponding author

\* = students I have advised or mentored

1. \*Zitong Zhang, §Ashraf Yaseen, Hulin Wu. Scholarly recommendation system for NIH funded grants based on biomedical word embedding models. *Natural Language Processing Journal*. August 2024. <https://doi.org/10.1016/j.nlp.2024.100095>
2. Sarah Messiah, Rhiana Abbas, Emma Bergqvist, Harold W Kohl, Michael D Swartz, Yashar Talebi, Rachit Sabharwal, Haoting Han, Melissa A Valerio-Shewmaker, Stacia M Desantis, Ashraf Yaseen, Henal A Gandhi, Ximena Flandes Amavisca, Jessica Ross, Lindsay N Padilla, Michael O Gonzalez, Leqing Wu, Mark A Silberman, David Lakey, Jennifer A Shuford, Stephen Pont and Eric Boerwinkle. Factors Associated with Elevated SARS-CoV-2 Immune Response in Children and Adolescents. *Frontiers in Pediatrics*. 14 August 2024. Volume 12 - 2024 | <https://doi.org/10.3389/fped.2024.1393321>
3. †Ashraf Yaseen, Stacia M. DeSantis, Rachit Sabharwal, Yashar Talebi, Michael D. Swartz, Shiming Zhang, Luis Leon Novelo, Cesar L Pinzon-Gomez, Sarah E. Messiah, Melissa Valerio-Shewmaker, Harold W. Kohl, Jessica Ross, David Lakey, Jennifer A. Shuford, Stephen J. Pont and Eric Boerwinkle. Baseline characteristics of SARS-CoV-2 vaccine non-responders in a large population-based sample. *PLoS One*. 2024 May 13;19(5):e0303420. doi: 10.1371/journal.pone.0303420. PMID: 38739625; PMCID: PMC11090326.
4. †Ashraf Yaseen, Claudia Robertson, Jovany Cruz Navarro, Jingxiao Chen, Brian Heckler, Stacia DeSantis, Nancy Temkin, Jason Barber, Brandon Foreman, Ramon Diaz-Arrastia, Randall Chesnut, Geoff Manley, David Wright, Mary Vassar, Adam Ferguson, Amy Markowitz, Jose-Miguel Yamal. Integrating, Harmonizing, and Curating Studies with High-Frequency and Hourly Physiological Data: Proof of Concept from Seven Traumatic Brain Injury Datasets. *Journal of Neurotrauma*. 2023 Aug 16. doi: 10.1089/neu.2023.0023. PMID: 37341031.
5. \*Zitong Zhang, \*Rachit Sabharwal, Miryoung Lee, Kehe Zhang, Paul McGaha, Michelle Crum, Cici Bauer, Susan P. Fisher-Hoch, Joseph B. McCormick, Belinda M Reininger, Samantha Thomas, Esmeralda Guajardo, Daniel Pinon, §Ashraf Yaseen. An Interactive Online Dashboard with Covid-19 Trends and Data Analysis in Northeast and South Texas. *The Texas Public Health Journal (TPHJ)*. Volume 76 Issue 2. 2023.
6. \*Zitong Zhang, §Ashraf Yaseen. A Content-Based Dataset Recommendation System for Biomedical Datasets. *2023 6th International Conference on Information and Computer Technologies (ICICT)*, Raleigh, NC, USA, 2023 pp. 198-202. doi: 10.1109/ICICT58900.2023.00040
7. \*Jie Zhu, §Ashraf Yaseen, Luis Leon-Novelo. Incorporating uncertainty quantification for actionable insights and performance improvement of academic recommenders. *Knowledge* 2023, 3, 293-306. <https://doi.org/10.3390/knowledge3030020>
8. \*Zitong Zhang, Braja Gopal Patra, §Ashraf Yaseen, \*Jie Zhu, \*Rachit Sabharwal, Kirk Roberts, Tru Cao, and Hulin Wu. Scholarly Recommendation Systems: A Literature Survey. *Knowledge and Information Systems* (2023), <https://doi.org/10.1007/s10115-023-01901-x>
9. Stacia DeSantis, †Ashraf Yaseen, Tianyao Hao, Luis León-Novelo, Yashar Talebi, Melissa Valerio-Shewmaker, Cesar Pinzon Gomez, Sarah Messiah, Harold Koh, Steven Kelder, Jessica Ross, Lindsay Padilla, Mark Silberman, Samantha Tuzo, David Lakey, Jennifer Shuford, Stephen Pont, Eric Boerwinkle, Michael Swartz. Incidence and predictors of breakthrough and severe breakthrough infections of SARSCoV-2 after primary series vaccination in adults:

- A population-based survey of 90,000 participants. *Journal of Infectious Diseases*. 2023 May 12;227(10):1164-1172. doi: 10.1093/infdis/jiad020. PMID: 36729177
10. \*Jie Zhu, Braja Patra, Hulin Wu, **Ashraf Yaseen**. A novel NIH research grant recommender using BERT. *PLoS One*. 2023 Jan 17;18(1):e0278636. doi: 10.1371/journal.pone.0278636
  11. \*Jie Zhu, **Ashraf Yaseen**. A Recommender for Research Collaborators Using Graph Neural Networks. *Frontiers in Artificial Intelligence*. 2022 Aug 1;5:881704. doi: 10.3389/frai.2022.881704. PMID: 35978654; PMCID: PMC9376356.
  12. \*Jie Zhu, Hulin Wu, **Ashraf Yaseen**. Sensitivity Analysis of a BERT-based scholarly recommendation system, *FLAIRS Conference Proceedings*, 35. 2022. <https://doi.org/10.32473/flairs.v35i.130595>.
  13. Stacia DeSantis, Luis Leon-Novelo, Michael Swartz, **Ashraf Yaseen**, Melissa Valerio, Yashar Talebi, Frances Brito, Jessica Ross, Harold Kohl III, Sarah Messiah, Steve Kelder, Leqing Wu, Shiming Zhang, Kimberly Aguillard, Michael Gonzalez, Onyinye Omega-Njemnob, David Lakey, Jennifer Shuford, Stephen Pont, Eric Boerwinkle. Methodology to estimate natural- and vaccine-induced antibodies to SARS-CoV-2 in a large geographic region. *PLOS ONE*, 2022 Sep 9. PMID: 36084125 PMCID: PMC9462720 DOI: 10.1371/journal.pone.0273694
  14. Sarah Messiah, Tianyao Hao, Stacia DeSantis, Michael Swartz, Yashar Talebi, Harold Kohl, Shiming Zhang, Melissa Valerio-Shewmaker, **Ashraf Yaseen**, Steven Kelder, Jessica Ross, Michael Gonzalez, Leqing Wu, Lindsay Padilla, Kourtney Lopez, David Lakey, Jennifer Shuford, Stephen Pont, Eric Boerwinkle. Comparison of Persistent Symptoms Following SARS-CoV-2 Infection by Antibody Status in Nonhospitalized Children and Adolescents. *The Pediatric Infectious Disease Journal*. 2022;INF.0000000000003653. doi:10.1097/INF.0000000000003653
  15. Michael Swartz, Stacia DeSantis, **Ashraf Yaseen**, Frances Brito, Melissa Valerio-Shewmaker, Sarah E Messiah, Luis G Leon-Novelo, Harold Kohl, Cesar Pinzon-Gomez, Tianyao Hao, Shiming Zhang, Yashar Talebi, Joy Yoo, Jessica Ross, Michael O Gonzalez, Leqing Wu, Steven H Kelder, Mark Silberman, Samantha Tuzo, Stephen J Pont, Jennifer Shuford, David Lakey, Eric Boerwinkle. Antibody duration after infection from SARS-CoV-2 in the Texas Coronavirus Antibody Response Survey. [published online ahead of print, 2022 May 6]. *Journal of Infectious Diseases*. 2022;jiac167. doi:10.1093/infdis/jiac167
  16. Sarah Messiah, Stacia DeSantis, Luis Leon-Novelo, Yashar Talebi, Frances Brito, Harold Kohl, Melissa Valerio-Shewmaker, Jessica Ross, Michael Swartz, **Ashraf Yaseen**, Steven Kelder, Shiming Zhang, Onyinye Omega-Njemnobi, Michael Gonzalez, Leqing Wu, Eric Boerwinkle, David Lakey, Jennifer Shuford, Stephen Pont; Durability of SARS-CoV-2 Antibodies From Natural Infection in Children and Adolescents. *Pediatrics* June 2022; 149 (6): e2021055505. 10.1542/peds.2021-055505
  17. Melissa Valerio-Shewmaker, Stacia DeSantis, Michael Swartz, **Ashraf Yaseen**, Michael Gonzalez, Harold Kohl, Steven Kelder, Sarah Messiah, Kimberly Aguillard, Camille Breaux, Leqing Wu, Jennifer Shuford, Stephen Pont, David Lakey, Eric Boerwinkle. Strategies to Estimate Prevalence of SARS-CoV-2 Antibodies in a Texas Vulnerable Population: Results From Phase I of the Texas Coronavirus Antibody Response Survey. *Frontiers in Public Health*. 2021 Dec 14;9:753487. doi: 10.3389/fpubh.2021.753487. PMID: 34970525; PMCID: PMC8712464.
  18. Sarah Messiah, Melissa Valerio-Shewmaker, Stacia DeSantis, Michael Swartz, **Ashraf Yaseen**, Frances Brito, Harold Kohl, Steven Kelder, Kimberly Aguillard, Onyinye Omega-Njemnobi, Camille Breaux, Jessica Ross, Michael Gonzalez, Shiming Zhang, Leqing Wu, David Lakey, Jennifer Shuford, Stephen Pont, Eric Boerwinkle. Estimated Prevalence of SARS-CoV-2 Antibodies in the Texas Pediatric Population, 2021. *The Lancet*, available at [SSRN](https://ssrn.com/abstract=3868061): <https://ssrn.com/abstract=3868061> or <http://dx.doi.org/10.2139/ssrn.3868061>
  19. Cong Zhu, Radhe Mohan, Steven Hsheng Lin, Goo Jun, **Ashraf Yaseen**, Xiaoqian Jiang, Han Chen, Qianxia Wang, Wenhua Cao, Brian Hobbs. Identifying Individualized Risk Profiles for

- Radiotherapy-Induced Lymphopenia Among Patients With Esophageal Cancer Using Machine Learning. *JCO Clinical Cancer Informatics*. 2021 Sep;5:1044-1053. doi: 10.1200/CCI.21.00098. PMID: 34665662; PMCID: PMC8812653.
20. \*Jie Zhu, Braja Patra, Hulin Wu, **Ashraf Yaseen**. Recommender system of scholarly papers using public datasets, *AMIA Jt Summits Transl Sci Proc*. 2021 May 17;2021:672-679. PMID: 34457183; PMCID: PMC8378599.
  21. Braja Patra, Babak Soltanalizadeh, Nan Deng, Leqing Wu, Vahed Maroufy, Wenjin Jim Zheng, Kirk Roberts, Hulin Wu, **Ashraf Yaseen**. An Informatics Research Platform to Make Public Gene Expression Time-Course Datasets Reusable for More Scientific Discoveries. *Database*, Volume 2020, 2020, PMID: 33247935 PMCID: PMC7698665 DOI: 10.1093/database/baaa074.
  22. Derek W. Brown, Stacia M. DeSantis, Thomas J. Greene, Vahed Maroufi, **Ashraf Yaseen**, Hulin Wu, George Williams, Michael D. Swartz. A Novel Approach for Propensity Score Matching and Stratification in the Presence of Multiple Treatments: Application to an EHR-Derived Study of Subarachnoid Hemorrhage. *Statistics in Medicine*. 39: 2308– 2323. 2020. doi: 10.1002/sim.8540. Epub 2020 Apr 16. PMID: 32297677; PMCID: PMC7334100.
  23. George Williams, Vahed Maroufy, Laila Rasmy, Derek Brown, Duo Yu, Hai Zhu, Yashar Talebi, Xueying Wang, Emy Thomas, Gen Zhu, **Ashraf Yaseen**, Hongyu Miao, Luis Leon Novelo, Degui Zhi, Stacia DeSantis, Hongjian Zhu, Jose-Miguel Yamal, David Aguilar, and Hulin Wu. Vasopressor Treatment and Mortality Following Non-Traumatic Subarachnoid Hemorrhage: A Nationwide EHR Analysis. *Neurosurgical Focus*. 2020. doi: 10.3171/2020.2.FOCUS191002. PMID: 32357322
  24. Vahed Maroufy, Pankil Shah, Arvand Asghari, Nan Deng, Rosemarie Le, Juan Camilo Ramírez, **Ashraf Yaseen**, W. Zheng, Michihisa Umetani, Hulin Wu. Gene expression dynamic analysis reveals co-activation of Sonic Hedgehog and epidermal growth factor followed by dynamic silencing. *Oncotarget*. 11. 10.18632/oncotarget.27547. 2020.
  25. **Ashraf Yaseen**, Hao Ji, and Yaohang Li, "A Load-Balancing Workload Distribution Scheme for Three-Body Interaction Computation on Graphics Processing Units (GPU)". *Journal of Parallel and Distributed Computing*, 87: 91–101, 2016. <https://doi.org/10.1016/j.jpdc.2015.10.003>
  26. **Ashraf Yaseen**, Mais Nijim, Brandon Williams, Lei Qian, Min Li, Jianxin Wang, and Yaohang Li "FLEXc: protein flexibility prediction using context-based statistics, predicted structural features, and sequence information". *BMC Bioinformatics*, vol. 17 Suppl 8, pp. 281, 2016. doi: 10.1186/s12859-016-1117-3. PMCID: PMC5009531 PMID: 27587065
  27. Mais Nijim and **Ashraf Yaseen**, "HuBum: Energy Efficient Hybrid Mobile Storage Systems using Solid States and Buffer Disks". *Journal of Computer Communication and Collaboration*, 2015. (DOIC: 2292-1036-2015-04-001-59)
  28. **Ashraf Yaseen** and Yaohang Li, "Context-based Features Enhance Protein Secondary Structure Prediction Accuracy". *Journal of Chemical Information and Modeling*, 54 (3), pp 992–1002, 2014. doi: 10.1021/ci400647u. Epub 2014 Mar 12. PMID: 24571803.
  29. **Ashraf Yaseen** and Yaohang Li, "Template-based C8-SCORPION: a protein 8-state secondary structure prediction method using structural information and context-based features", *BMC Bioinformatics*, 15(Suppl 8):S3, 2014. doi: 10.1186/1471-2105-15-S8-S3 | PMCID: PMC4120151 PMID: 25080939
  30. Zhiqiang Wu, Bin Wang, Chi-Hao Cheng, Dr. Deng Cao, and **Ashraf Yaseen**. "Software Defined Radio Laboratory Platform for Enhancing Undergraduate Communication and Networking Curricula," 2014 *ASEE Conference*, 2014.
  31. **Ashraf Yaseen** and Yaohang Li, "Dinosolve: A Protein Disulfide Bonding Prediction Server using Context-based Features to Enhance Prediction Accuracy", *BMC Bioinformatics*, 14(Suppl 13):S9, 2013. doi: 10.1186/1471-2105-14-S13-S9. Epub 2013 Oct 1. PMID: 24267383; PMCID: PMC3849605

32. Yaohang Li and **Ashraf Yaseen**, "Pareto-based Optimal Sampling Method and Its Applications in Protein Structural Conformation Sampling". *AAAI Workshop on Artificial Intelligence and Robotics Methods in Computational Biology*, Bellevue, 2013.
33. †**Ashraf Yaseen** and Yaohang Li "Predicting Protein Solvent Accessibility with Sequence, Evolutionary Information and Context-based Features". *Biotechnology and Bioinformatics Symposium*, (BIOT2013) Provo, 2013.
34. †**Ashraf Yaseen** and Yaohang Li "Template-based Prediction of Protein 8-states Secondary Structures". 3rd *IEEE International Conference on Computational Advances in Bio and Medical Sciences* (ICCABS2013), New Orleans 2013.
35. †**Ashraf Yaseen** and Yaohang Li "Enhancing Protein Disulfide Bonding Prediction Accuracy with Context-based Features", *Proceedings of Biotechnology and Bioinformatics Symposium*, (BIOT2012), Provo, 2012.
36. †**Ashraf Yaseen** and Yaohang Li, "Accelerating Knowledge-based Energy Evaluation in Protein Structure Modeling with Graphics Processing Units," *Journal of Parallel and Distributed Computing*, 72(2): 297-307, 2012. <https://doi.org/10.1016/j.jpdc.2011.10.005>
37. Weihang Zhu, **Ashraf Yaseen** and Yaohang Li "DEMCMC-GPU: An Efficient Multi-Objective Optimization Method with GPU Acceleration on the Fermi Architecture" *New Generation Computing*, 29(2): 163-184, 2011.
38. †**Ashraf Yaseen**, Kurt J. Maly, Steven J. Zeil and Mohammad Zubair "Performance Evaluation of Oracle Semantic Technologies with respect to User Defined Rules". *Proceeding of Database and Expert Systems Applications*, DEXA, International Workshops, Toulouse, France, August 29, 2011.

### **Books**

39. Hulin Wu, Jose-Miguel Yamal, **Ashraf Yaseen**, and Vahed Maroufy. *Statistics and Machine Learning Methods for EHR Data, From Data Extraction to Data Analytics*. United States: CRC Press, 2020.

## 2. FUNDING

- 
- |    |   |           |
|----|---|-----------|
| 1. | Texas Department of State Health Services (TDSHS) / Centers for Disease Control & Prevention (CDC)  | 2021-2024 |
|    | <b>Texas SARS-CoV-2 Variant Sequencing Study</b>  |           |
|    | Expand genomic sequencing of the virus that causes COVID-19. Sequence and study more COVID-19 samples from around the state to provide a better picture of circulating and emerging variants of the COVID-19 virus. Total: <b>\$14,124,185</b> . Co-Investigator (PI: Boerwinkle)<br>Avg %effort: 2021-2022: 15%, 2022-present: 10% |           |
| 2. | Texas Department of State Health (TDSHS)  | 2020-2025 |
|    | <b>Texas Coronavirus Antibody Response Surveillance</b>   |           |
|    | The main objective is to understand person, place, time, disparities and trends of COVID-19 to inform public health action and policy. Total: <b>\$8,409,755</b> . Co-Investigator (PI: Boerwinkle)<br>Avg %effort: 2020-2021: 25%, 2021-present: 30%   |           |
| 3. | Office of the National Coordinator for Health Information Technology - U.S. Health Resources and Services Administration (HRSA) - US Department of Health and Human Services (DHHS)   | 2021-2025 |
|    | <b>The PHIT Workforce Development Program: Creating a diverse and inclusive health information technology (IT) workforce in Texas</b>   |           |
|    | Total: <b>\$9,213,935</b> . Co-Investigator (PI: Boerwinkle)<br>Avg %effort: 10%  |           |

4. US Department of Defense (DOD) 2020-2024  
**Leveraging FITBIR Data to Improve Clinical Practice of Severe TBI**  
 Aims of this study: 1. Integrate and harmonize data from various multi-center TBI studies 2. Curate data from various multi-center TBI studies 3. To assess the association between the ways ICP is treated and long-term neurological outcomes  
 Total: **\$748,708**. Co-Investigator (PI: Yamal).  
 Avg %effort: 2020-2022: 15%, 2022-present: 10%
5. National Institute of Health (NIH) - Center for Advancing Translational Sciences 2021-2023  
**RADx-UP Phase II (COVID)**  
 This study will leverage longstanding academic-community engaged partnership to examine SARS-CoV-2 infection patterns and identify dynamic disease hotspots and testing deserts in racially diverse neighborhoods of three Texas regions (Houston/Harris County, South Texas and Northeast Texas) and evaluate the rapid adaptation and deployment of multilevel intervention strategies to SARS-CoV-2 testing in vulnerable populations.  
 Total: **\$3,204,351**. Co-Investigator (PI: Fernandez)  
 Avg %effort: 2%
6. U.S. Health Resources and Services Administration (HRSA) - US Department of Health and Human Services (DHHS) 2021-2023  
**Community-Based Workforce Development and Mobilization to Increase COVID-19 Vaccination Equity in Texas.**  
 The goal is to increase COVID-19 vaccinations through the development and mobilization of existing community-based health and outreach workforces in the state of Texas.  
 Total: **\$11,623,660**. Co-Investigator (PI: Fernandez)  
 Avg %effort: 8%
7. Texas Department of State Health Service (TDSHS) 2022-2022  
**COVID-19 Vaccine Hesitancy and Confidence (COVAHC) Survey: A Rapid Community Assessment in Texas.**  
 Total: **\$487,364**. Co-Investigator (PI: Cuccaro)  
 Avg %effort: 15%
8. National Institute of Health (NIH) - Center for Clinical and Translational Sciences 2020-2022  
**RADx-UP Phase I (COVID)**  
 This study leverages longstanding academic-community engaged partnership to examine SARS-CoV-2 infection patterns and identify dynamic disease hotspots and testing deserts in racially diverse neighborhoods of three Texas regions (Houston/Harris County, South Texas and Northeast Texas) and evaluates the rapid adaptation and deployment of multilevel intervention strategies to SARS-CoV-2 testing in vulnerable populations.  
 Total: **\$4,998,788**. Co-Investigator (PI: Fernandez)  
 Avg %effort: 3%
9. Centers for Medicare and Medicaid Services (CMS) – US Department of Health and Human Services (DHHS) 2020-2022  
**Assistance, Addressing Social Needs of High-Risk Patients through Screening and Navigation to Community Social Service Organizations (Track 2)**

The UTHealth School of Public Health team proposes to address the social needs in the ACH model, Assistance. In partnership with our UT Physicians, Memorial Hermann Hospital and Texas Children's Hospital, we seek to apply an innovative asynchronous platform for screening and navigation of patients.

Total: **\$2,559,327**. Co-Investigator (PI: Highfield)

Avg %effort: 15%

10. Harris County 2020-2021

**SARS-CoV-2 Surveillance Testing Program for Harris County**

The main objective is to enhance COVID-19 testing and understand the epidemiology and dynamics of COVID-19 in our Harris County and the city of Houston.

Total: **\$16,985,172**. Co-Investigator (PI: Boerwinkle)

Avg %effort: 20%

11. National Heart, Lung, & Blood Institute (NHLBI) - National Institute of Health (NIH) / Clinical Pathology Labs (CPL) 2020-2021

**Rapid Expansion of Existing Framework for Deploying Large-Scale COVID-19 RT-PCR Testing Platforms and Distributing Capacities.** RADx Tech NIH Grant Sub.

Total: **\$206,930**. Co-Investigator (PI: Melissa Valario)

Avg %effort: 25%

12. National Institute of Health (NIH) - NIDCR 2019-2022

**CATCH Healthy Smiles: A cluster-RCT of an elementary school oral health intervention**

This grant will allow us to plan for, and test the efficacy of an elementary school-based oral health intervention using a cluster-randomized controlled trial design across children from ethnically-diverse, low-income families in Houston, Texas.

Total: **\$4,096,889**. Co-Investigator (PI: Sharma)

Avg %effort: 7%

13. Office of Research and Sponsored Programs, Texas A&M University-Kingsville 2016-2018  
**Computational Sciences Lab**

A project to establish a High-Performance Computing Multidisciplinary Research lab. Provides computing services, facilitates multidisciplinary research through collaboration, and trains faculty members and member students

Total: **\$93,872**. (PI: Ashraf Yaseen)

14. Office of Research and Sponsored Programs, Texas A&M University-Kingsville 2016-2017  
Development of **Javelinas-Server for Predicting Protein Structural Features.**

Total: **\$3,000** (PI: Ashraf Yaseen)

15. US Department of Homeland Security (DHS) 2015-2018

**Security Engineering: Development of Curriculum and Research for Homeland Security**

Security Engineering is a multidisciplinary minor program within the College of Engineering at TAMUK offered in support of preparing engineering and science students for careers in areas related to our nation's security.

Total: **\$698,000**. Co-Investigator (PI: Selahattin Ozcelik)

Avg %effort: 15%

16. National Science Foundation (NSF)-TUES 2013-2014

**Collaborative: TUES: Software Defined Radio Laboratory Platform for Enhancing Undergraduate Communication and Networking Curricula**

Evolvable wireless laboratory design and implementation for enhancing undergraduate wireless engineering education in which the team developed and demonstrated lower cost, software defined radio (SDR) based laboratories for undergraduate courses.

Participating institutions: Wright State University (Lead), Miami University Oxford Campus, and Central State University.

Total: **\$100,000**. Role: Co-investigator.

---

## C. Invited Talks, Presentations, and Posters

---

### Invited Talks

1. Lack of antibody response in those vaccinated or with natural exposure. [Session: Examining SARS-CoV2 Response Over Time Using a Longitudinal Design: Texas CARES Survey]. American Public Health Association (APHA). November 7, 2022. Boston, MA.
2. Texas CARES Community Update, Lessons Learned and Next Steps. Healthier Texas Summit. October 21, 2022. Austin TX.
3. Epidemiology Special Session: Understanding the Human Antibody Response to Sars-Cov-2 in Diverse Populations: The Texas Coronavirus Antibody Response Survey (CARES). Data Management and Visualization. American Public Health Association (APHA). October 25, 2021, Denver, Colorado.
4. Understanding the Human Antibody Response to SARS-CoV-2 in Diverse Populations: The Texas Coronavirus Antibody Response Survey (CARES). World Health Organization (WHO) Solidarity II. August 27, 2021.
5. Texas C.A.R.E.S. Coronavirus Antibody REsponse Survey. Texas CARES Portal: An Interactive Platform with Visualizations, Maps, and Summary Statistics to Illustrate and Understand the Human Response to COVID-19. Texas Department of State Health Services (DSHS) Grand Rounds. June 16, 2021.

### Presentations at National or International Conferences

1. International Conference on Big Data and Information Analytics. Houston, Texas. December 17-19, 2018.
  - Presentation: Collaborative Platform for GEO Big Data Project: An innovative platform with Scalable Analytic Tools to Efficiently Promote Use/Reuse of Time Course Gene Expression Data for Scientific Discoveries. Session: Big Data at UTHealth: Use of the Public Genetic Database, GEO, for Big Data Research.
2. Applied Statistics Symposium. Houston, Texas. December 13-16, 2020.
  - Short course: Statistics and Machine Learning Methods for EHR Data: From Data Extraction to Data Analytics/Predictions. Full-day. Hulin Wu, Ashraf Yaseen and Vahed Mafoury.
  - Poster: Recommender system of scholarly papers using public datasets. Jie Zhu, Braja Patra, Hulin Wu and Ashraf Yaseen.
  - Presentation: Big Data to answer Big Questions: Experience with Anuerysmal SAH. Session: Statistical and Machine Learning models on EHR and Insurance Claim databases. Vahed Mafoury, Ashraf Yaseen and George Williams.
3. International Symposium on Bioinformatics Research and Applications. Norfolk, Virginia. June 7-10, 2015
4. AAAI Workshop on Artificial Intelligence and Robotics Methods in Computational Biology. Bellevue, Washington. July 14-18, 2013
5. BIOT: Biotechnology and Bioinformatics Symposium Provo, Utah. December 5-6, 2013
6. IEEE International Conference on Computational Advances in Bio and Medical Sciences. New Orleans, Louisiana. June 12-14, 2013

7. Annual Tidewater Student Research at Christopher Newport University. Newport, Virginia. November, 2012
8. BIOT: Biotechnology and Bioinformatics Symposium Provo, Utah. October 25-26, 2012

### Posters

1. \*Jie Zhu, Braja Patra, Hulin Wu and Ashraf Yaseen. Recommender system of scholarly papers using public datasets. ICSA Applied Statistics Symposium. Houston, Texas. December 13-16, 2020.
2. \*Praveenraj Uthamarajan and Ashraf Yaseen, "Analysis of Systems using Distributed Consensus Algorithms". College of Engineering-TAMUK, 2017.
3. \*Megha Lalluvadia and Ashraf Yaseen, "Applications of Text Classification". College of Engineering-TAMUK, 2017.
4. \*Varun Agrawal, Gaurav Dokania, and Ashraf Yaseen, "Predicting protein flexibility and disorder". Texas A&M University System 12th Annual Pathways Student Research Symposium, Corpus Christi, TX, 2015.
5. \*Anurag Gupta, Hridya Gopalakrishna, and Ashraf Yaseen, "Predicting protein solvent accessibility". Texas A&M University System 12th Annual Pathways Student Research Symposium, Corpus Christi, TX, 2015.
6. †Ashraf Yaseen, Mais Nijim, Brandon Williams, Lei Qian, and Yaohang Li "Predicting Protein Flexibility using Context-based Statistics, Predicted Structural Features, and Sequence Information". 11th International Symposium on Bioinformatics Research and Applications (ISBRA), Norfolk, Virginia, 2015. (Awarded #1 best poster).
7. †Ashraf Yaseen, Akeem Edwards and Yaohang Li, "Improving Intermediate Steps in ab initio Protein Molding", 14th Annual Tidewater Student Research Poster Session at Christopher Newport University. Nov, 2012.

---

## D. Teaching

---

### 1. CLASSES

- At UTHealth:
 

<u>Course</u>	<u>Semester (teaching score/5)</u>
- <b>Fundamentals of Data Analytics and Predictions</b> Course co-developer and instructor 50%	Spring 2019 (4.13), 2020 (4.32), 2021 (4.46), 2022 (4.30), 2023 (4.51)
- <b>Machine Learning in Practice</b> Course developer and lead instructor 100%	Fall 2018 (4.07), 2019 (4.62), 2020 (4.38), 2021 (4.42), 2022 (4.61), 2023 (4.4)
- <b>Data Science Computing</b> Course developer and lead instructor 50%	Spring 2020 (4.18)
- <b>Introduction to Statistical and Data Science Programming (Python &amp; R)</b> Course developer and lead instructor 50%	Fall 2020 (4.14)
- At TAMUK:
 

- <b>Introduction to Bioinformatics</b> Course developer and lead instructor 100%	Summer 2016 (4.30)
- <b>Bioinformatics Computing</b> Course developer and lead instructor 100%	Summer 2015 (4.25, 4.27), 2016 (4.36, 4.31)
- <b>Cloud Computing</b> Course developer and lead instructor 100%	Summer 2017 (NA)

- **Database Systems** Spring 2015 (4.35, 4.22), 2016 (4.35), 2017 (4.44)  
Lead instructor 100%
- **Operating Systems** Fall 2014 (4.27, 4.22), 2015 (4.42, 4.37), 2016 (4.5), 2017 (4.41)  
Lead instructor 100%
- **Computer Communication Networks** Fall 2014 (4.27)  
Lead instructor 100%

## 2. ADVISING

### ▪ **Thesis/Dissertation Supervisor at UTHealth**

Ph.D. (Data Science)	Graduation Date
1. Jie (Ginny) Zhu Virtual Research Assistant (VRA): a platform for recommending datasets, grant announcements, and collaborators for population health professionals	Summer 2022
2. Zitong Zhang (Biomedical) Data Sets and Research Grants Recommendation Systems: Methods and Statistical Evaluation	<expected> Fall 2024
3. Rachit Sabharwal Certainty/Uncertainty of Deep Learning Predictions in Population Health Applications	<expected> Spring 2025
4. Tzuruei Chao AI-based Modeling Predicting Adolescent High-risk Trajectory for HIV	<expected> Spring 2026
<hr/>	
M.S. (Biostatistics & Data Science)	
1. Joy Yoo Predicting the antibody test results for Covid-19 using Machine Learning Methods	Spring 2022
2. Rachit Sabharwal BIOREC: A biomedical recommendation system for academic conferences and journals	Spring 2022
3. Nitesh Enduru Association of genetic risk, midlife simple 7, and incident stroke: the atherosclerosis risk in communities (ARIC) study	Summer 2021
4. Christin Silos Predicting Length of Stay for Ischemic Stroke Patients using Machine Learning Methods	Spring 2021
5. Brandon O'Grady Survival analysis of colorectal cancer patients with liver metastasis	Fall 2020
6. Mengchen Ding Applications of statistical methods studying the impact of mobilization regimes on the total collection yield of hematologic stem cells	Fall 2020
7. Mira Baltaji ADHD Predictions in Adults and Associations with Bipolar Disorder using Machine Learning Methods	<expected> Fall 2024
8. Youssef Benfallah Estimation of Obesity Levels Based on Eating Habits, Lifestyle factors and Physical Condition	<expected> Fall 2024

### ▪ **Member of the Academic/Thesis/Dissertation Committees for Student at UTHealth-SPH**

Ph.D. (Biostatistics & Data Science)	
1. Xueying Wang	Summer 2022

Bias-Corrected Machine Learning Methods for Risk Predictions Using EHR Data with Censoring: Applications to Heart Failure Predictions Among Diabetes Patients	
2. Jeffrey Lin	Spring 2022
Machine Learning for the Joint Analysis of Multivariate Longitudinal and Survival data	
3. Liang Wu	Spring 2021
Developing Machine Learning Algorithms for Time-Course Healthcare Data	
4. Cong Zhu	Spring 2020
A machine learning based framework for studying the risk of radiotherapy induced lymphopenia and its association with survival among esophageal cancer patients	
<b>M.S. (Biostatistics &amp; Data Science)</b>	
1. Kimenpreen Kaur	Summer 2023
Predicting long COVID outcomes in diabetic patients	
2. Lakshmi Kanikkannan	Summer 2025
Analyzing the correlation between the average macronutrient consumption of numerous countries and their COVID-19 cases, hospitalization rates, and deaths	
<ul style="list-style-type: none"> <li>▪ <b>Thesis/Dissertation Supervisor at TAMUK</b></li> </ul>	
<b>M.S. (Computer Science)</b>	
1. Megha Lalluvadia	Spring 2017
Applications of Text Classification	
2. Jatin Waghela	Spring 2017
A study of web application development stacks and demonstration of MEAN stack	
3. Praveenraj Uthamarajan	Spring 2017
Analysis of Systems using Distributed Consensus Algorithms	
4. Gaurav Dokania	Fall 2015
Predicting Protein Disorder	
<ul style="list-style-type: none"> <li>▪ <b>Project Supervisor at TAMUK</b></li> </ul>	
<b>M.S. (Computer Science)</b>	
1. Rakshith Padmanabha	Fall 2016
Sentimental Mining of Social Media Data for the Detection of Malicious Behaviors and Activities of Individuals	
2. Rajan Pawar	Fall 2016
Performance Analysis of NoSQL Databases and Relational Databases	
3. Pushpak Gandhi	Fall 2016
Developing Web based Java Applications	
4. Dhara Shah	Fall 2016
Syntax and Semantic based Approach for Automatic Question Generation	
5. Anjali Shinde	Fall 2016
Selenium Automation Testing Tool	
6. Rushikesh Jawali	Fall 2016
Aroundme (mobile application)	
7. Suvabrata Dutta	Fall 2016
A Web Service Recommendation System	

8. Sakalkar Saurabh A Simple Recommendation System	Fall 2016
9. Aamani Mayakuntla Finger Print Compression based on Sparse Representation	Fall 2016
10. Sai Srinivas Maddipati Hospital Management System (HMS)	Fall 2016
11. Sandeep Reddy Takkolu A Soft-computing based Stock Market Recommender System	Fall 2016
12. Hridya Gopalakrishna Timesheet Manager	Spring 2016
13. Harish Pyneni Mobile Application Development	Spring 2015
14. Anil kumar Burra Mobile Application (Buddy Ride)	Spring 2015
15. Alvin Ahmed Prasla TAMUK Job Portal	Spring 2015
16. Praveen Kumar Kollipara Student Profile Management System	Spring 2015
17. Anurag Gupta Machine Learning Methods for Predicting Protein Solvent Accessibility	Spring 2015
18. Varun Agrawal Machine Learning Methods for Predicting Protein Disorder	Spring 2015
19. Mayur Prakash Kaware 'The Movie App' in android	Spring 2015

---

## E. Service

---

### 1. PROFESSIONAL GROWTH & SERVICE ACTIVITIES

#### Lead Administrative Roles

- Program Coordinator of UTHealth-SPH
  - Data Science and Advanced Data Science Certificates (2018-present)
  - Online Data Science and Advanced Data Science Certificates (2019-present)
  - Data Science MS. degree program (2020-present)
- Manager of the Data Science Software Development Team (2018-present)
- Director of TAMUK Computational Sciences Lab (2016-2018)

#### Internal Services

- Committee memberships at UT-SPH
  - Data Science Task Force (2018-present)
  - Data Science curriculum committee (2018- present)
  - Data Science Faculty Search Committee (2018-present)
  - SPH Faculty IT Advisory Committee (2020-present)
  - BaDS Department promotion committee (2018- 2020)
- Committee memberships at TAMUK
  - Biomed Research Group (2016-2018)
  - CS Undergraduate ABET Assistant Director (2016-2018)
  - Engineering College Council Committee (2016-2018)
  - EECS Graduate Curriculum Committee (2014-2018)
  - Undergraduate Program Review Committee (2014-2018)
  - TAMUK Graduate Faculty (2014-2018)
  - Javelina Scholarship Reviews Committee (2017-2018)

- University Expert List (2017-2018)

### **Recruitment**

- Talk at TAMUK, Kingsville TX. April 2018.

### **Service To the Profession**

#### **Leadership Roles in Professional Societies**

- 2018 BigDIA. 4th International Conference on Big Data and Information Analytics. Houston, Texas. December 17-19, 2018.
  - Chair of Program Book Committee and Website Committee
  - Co-Chair of Local Organizing Committee
  - Committee for Peer-Reviewed Track
  - Sessions Chair
- 2020 ICOSA Applied Statistics Symposium. Houston, Texas. December 13-16, 2020.
  - Co-Chair of Program Book and Website Committee
- International Symposium on Bioinformatics Research and Applications (Norfolk, VA 2015)
  - Publication Chair & Sessions Chair (2015)
  - Program Committee Member (2015-2017)
- Conference on Information and Computer Technology (2014). Program Committee Member

#### **Services as Reviewer**

- Paper Reviewer
  - ACM/IEEE Transactions on Computational Biology and Bioinformatics (2015-present)
  - BMC Bioinformatics (2015-present)
  - Journal of Information Science (2017-2019)
  - International Journal of Sensor Networks (IJSNET) (2017-2018)
  - International Journal of Cloud Applications & Computing (2017-2018)
  - International Symposium on Bioinformatics Research and Applications-ISBRA (2015-2017)
- Proposal Reviewer
  - TAMUK Biomedical Research Group (2017)
- Service as a Judge
  - Poster Judge at UTHealth-SPH Research Day (2019)
  - TAMUK Annual Engineering Senior Design Conference (2017)
  - TAMUK Graduate Students' Research Poster Competition (2016-2018)

#### **Membership in Professional Societies**

- Association for Computing Machinery (ACM)
- Institute of Electrical and Electronics Engineers (IEEE)
- American Public Health Association (APHA)

#### **Participation in Professional Meetings and Workshops**

- First Annual Healthcare Hackathon, Rio Grande Valley Health Information Exchange (RGV HIE). Weslaco, TX, November 19-20, 2019
- Texas Advanced Computing Center (TACC) Workshop. 10th annual TACC Summer Supercomputing Institute, The University of Texas at Austin. August 1-5, 2017 Austin, Texas
- NSF CAREER writing workshop Portland, OR. April 2-4, 2017
- NSF CAREER writing workshop and meeting with Program Directors at the NSF, Washington DC. March 20-21, 2017

- CCICADA Workshop: Command, Control, and Interoperability Center for Advanced Data Analysis, Reconnect 2016 program. Cybersecurity Institute. U.S. Military Academy at West Point, NY. June 12-18, 2016.
- Sustainable Energy Systems. Prof. P.K. Sen: “Energy, Electricity and Renewable Energy Resources: Sustainable Energy Systems”, TAMUK, May 2-3, 2016.
- Grant development workshops and webinars. TAMUK “New Faculty Investment Program” 2014-2016

## 2. COMMUNITY SERVICE

- ❖ Co-Lead and Developer of the Texas Pandemic platform (<http://www.texaspandemic.org/>) The website contains maps, graphs and analyses, to provide Texans with a better understanding of COVID-19 trends at the state level, the county level, and within collections of counties (trauma service areas, public health regions, and metropolitan areas). The platform contains visualizations of COVID19 cases, deaths, vaccinations, hospitalization, and more.