Abbas Rizvi, Ph.D.

Data Scientist

EDUCATION

Ph.D.	Pharmaceutical Sciences	The Ohio State University Columbus, OH, USA	2015-2019
M.S.	Cancer Biology	University at Buffalo Buffalo, NY, USA	2013-2015
M.Sc.	Integrated Systems Biology	Université du Luxembourg Luxembourg, Luxembourg	2013-2015
B.S.	Biology	SUNY Fredonia Fredonia, NY, USA	2008-2012

ADDITIONAL EDUCATION

Pending RStudio Certified Tidyverse Instructor *RStudio PBC Education, Online*	Sep 2020
Certificate of Advanced Study in Applied Biostatistics *University at Buffalo, Buffalo, NY*	Aug 2015-Dec 2016
Computational Genomics Winter Institute *University of California Los Angeles, Los Angeles, CA*	Feb 26-Mar 3, 2018
Advanced Gene Mapping *The Rockefeller University, New York, NY*	Jan 23-27, 2017
Analysis of Next Generation Sequencing for Complex Traits *Max Delbrük Center for Molecular Medicine, Berlin, Germany*	July 4-8, 2016
Bioinformatics, Exchange Semester (CanSys Program) *Vrije Universiteit Amsterdam, Amsterdam, Netherlands*	Aug 2014-Dec 2014
Data Science Specialization Certificate (completed 6 of 9 courses) *Johns Hopkins University (Coursera)*	Feb 2014-Present

WORK EXPERIENCE

AI Engineer/Data Scientist, Covail (Columbus Collaboratory) June 2018-Present
Columbus, OH, USA

- Built, designed, and deployed dozens of AI/machine learning models
- Conducted several natural language processing (NLP) projects, including using TF-IDF, topic modeling, Word2Vec, embeddings, transformer models (BERT, seq2seq, RNN-LSTM)
- Built custom seq2seq model for large organization and is in production on their environment, greatly improving their business process function
- Built NLP model that mined a large government R&D lab SOP violations and created a model that included the most prevalent words, yielding a model that was highly accurate at predicting error/failure models

- Built an NLP model at a Social Worker organization that including TF-IDF and Word2Vec embeddings. Built a gradient boosting model that was highly accurate at predicting if a document was narrative or non narrative
- Built an NLP model using regular expressions to pull rates from contract and furthermore predict if text was predictive of a transaction error occurring.
- Automated dozens of projects in several different industries including aviation, insurance, biotech, internal audit, international libraries, healthcare, supply chain, and more.
- Scoped projects and communicated with stake holders and project sponsors throughout the entire build process.

Graduate Research Associate, Department of Pharmaceutical Sciences June 2016-May 2019
The Ohio State University, Columbus, OH, USA
Advisor: Lara Sucheston-Campbell, PhD
Research: genome-wide association studies, survival analysis, data visualization

Pre-Doctoral Trainee, Department of Pharmacology and Therapeutics Aug 2015-June 2016
Roswell Park Cancer Institute, Buffalo, NY, USA
Research: -omics in prostate cancer (RNA-seq, ChIP-seq), pharmacogenomics in breast cancer, R package development

Data Scientist Intern Aug 2015-Nov 2015
TROVE Predictive Data Science LLC, Buffalo, NY, USA
Project: Recapitulate power grid network after large scale outages using machine learning (igraph, markov models) and statistical techniques (linear models) using R and SQL

Graduate Research Assistant, Department of Pharmacology and Therapeutics Aug 2013-Sep 2015
Roswell Park Cancer Institute, Buffalo, NY, USA
Adviser: Barbara Foster, PhD
Research: Transcriptomics (RNA-seq) in advanced prostate cancer

Undergraduate Research Assistant, Department of Biology Aug 2011-May 2012
SUNY Fredonia, Fredonia, NY, USA
Adviser: Scott Medler, PhD
Research: Effects of high fat diet on muscle fiber isoforms in mice

COMPUTING EXPERTISE

R, R/Bioconductor, R ShinyApps, R Markdown, UNIX/Linux, Docker, Python, LaTeX, Git, PLINK, IMPUTE2, MATLAB, Microsoft Office Suite

PUBLICATIONS

Karaesmen E, Hahn T, Dile AJ, **Rizvi AA**, Wang J, Wang T, Haagenson MD, Preus L, Zhu Q, Liu Q, Yan L, Liu S, Haim CA, Stram D, Pooler L, Sheng X, Van Den Berg D, Brock G, Webb A, McCarthy PL, Pasquini MC, Spellman SR, Lee SJ, Paczesny S, Sucheston-Campbell LE. "Multiple functional variants in the IL1RL1 region are pretransplant markers for risk of GVHD and infection deaths. (2019) Blood Advances 3 (16) 2512-2524

Tang H, Hahn T, Karaesmen E, **Rizvi AA**, Wang J, Paczesny S, Wang T, Preus L, Zhu Q, Wang Y, Haiman CA, Stram D, Loreal P, Sheng X, Van Den Berg D, Brock G, Webb A, Pasquini MC, McCarthy PL, Spellman SR, Sucheston-Campbell LE. (2019). "Validation of genetic associations with acute GVHD and nonrelapse mortality in DISCOVeRY-BMT" Blood Advances 3 (15) 2337-2341.

Rizvi AA, Karaesmen E, Morgan M, Preus L, Wang J, Sovic M, Hahn T, Sucheston-Campbell LE. "gwasurvivr: an R package for genome wide survival analysis". (2019) Bioinformatics 35 (11), 1968-1970

Long MD, Singh PK, Russel JR, Llimos G, Rosario S, **Rizvi AA**, van den Berg PR, Kirk J, Sucheston-Campbell LE, Smiraglia DJ, Campbell MJ. (2019). *"The miR-96 and RARγ signaling axis governs androgen signaling and prostate cancer progression"*. Oncogene 38 (3) 421-444

Sucheston-Campbell LE, Clay-Gilmour AI, Barlow WE, Budd GT, Stram DO, Haiman CA, Sheng X, Yan L, Zirpoli G, Yao S, Jiang C, Owzar K, Hershman D, Albain KS, Hayes DF, Moore HC, Hobday TJ, Stewart JA, **Rizvi AA**, Isaacs C, Salim M, Gralow JR, Hortobagyi GN, Livingston RB, Kroetz DL, Ambrosone CB (2017) "Genome-Wide Meta-Analyses Identifies Novel Taxane-Induced Peripheral Neuropathy Associated Loci". *Pharmacogenetics and Genomics*. pgen-2017

Karaesmen E*, **Rizvi AA***, Preus L, McCarthy PL, Pasquini MC, Onel K, Zhu X, Spellman S, Haiman CA, Stram DO, Pooler L, Sheng X, Zhu Q, Yan L, Liu Q, Hu Q, Webb A, Brock G, Clay-Gilmour AI, Battaglia S, Tritchler D, Liu S, Hahn T, Sucheston-Campbell LE (2017). Replication and validation of genetic polymorphisms associated with survival after allogeneic blood or marrow transplant. *Blood*, blood-2017-05-784637. https://doi.org/10.1182/blood-2017-05-784637. *Denotes co-first authorship

Clay-Gilmour AI, Hahn T, Preus LM, Onel K, Skol A, Hungate E, Zhu Q, Haiman CA, Stram, DO, Pooler L, Sheng X, Yan L, Liu Q, Hu Q, Liu S, Battaglia S, Zhu X, Block AW, Sait SN, Karaesmen E, **Rizvi AA**, Weisdorf DJ, Ambrosone CB, Tritchler D, Ellinghaus E, Ellinghaus D, Stanulla M, Clavel J, Orsi L, Spellman S, Pasquini, MC, McCarthy PL, Sucheston-Campbell LE (2017). Genetic association with B-cell acute lymphoblastic leukemia in allogeneic transplant patients differs by age and sex. *Blood Advances,* 1(20), 1717-1728. https://doi.org/10.1182/bloodadvances.2017006023.

DeNies MS, Johnson J, Maliphol AB, Bruno M, **Rizvi AA**, Rustici K, Medler S, "Diet-induced obesity alters skeletal muscle fiber types of male but not female mice". (2014) *Physiological Reports*. PMID: 24744883

PEER REVIEW

2017-2018	Reviewer for Biology of Blood and Marrow Transplantation

Talks

Artificial Intelligent Expert Panel, Columbus State, April 2020 (presentation)
"AI at Covail"

Columbus State, November 2019 (presentation)
"How to build a career in data science"

useR2019! (Lightning Talk), July 2019
"photon: electron meets shiny"

BioC2019 (Lightning Talk), June 2019
"photon: electron meets shiny"

BioC2018 (Contributed Talk), July 2018
"gwasurvivr: an R package to perform survival association testing on imputed genetic data."

BMT Tandem Meetings 2018 (Oral Presentation), Feb 2018
"Genetic Associations with Day +100 Transplant Related Mortality (TRM) after HLA-Matched Unrelated Donor (MUD) Blood and Marrow Transplantation (DISCOVeRY-BMT Study)"

The Ohio State University Comprehensive Cancer Center Annual Scientific Meeting, Mar 2017
(Poster Presentation)
"*Replication or validation of previously reported associations between genetic polymorphisms and survival outcomes after URD HSCT*"

American Society of Hematology 58th Annual Meeting, Dec 2016 (Oral Presentation)
"Replication of Candidate SNP Survival Analyses and Gene-Based Tests of Association with Survival Outcomes after an Unrelated Donor Blood or Marrow Transplant: Results from the DISCOVeRY-BMT Study"

Inaugural The Ohio State University Pharmaceutics Retreat, Oct 2016 (Oral Presentation)
"Visualizing genomic results for team science: an example from the DISCOVeRY-BMT study"

Roswell Park Cancer Institute 2015 Retreat, July 2015 (Poster Presentation)
"Reprogramming the AR and LSD1 transcriptome in Castration Resistant Prostate Cancer"

12th Annual SUNY Fredonia Creative Activity and Research, Apr 2012 (Poster Presentation)
"Muscle Fiber Types in Obese Mice: Does a High Fat Diet Cause Changes in Skeletal Muscle Phenotype?"

FELLOWSHIPS/SCHOLARSHIPS/HONORS

Elected to Phi Kappa Phi Honor Society – Chapter 155	Sept 2017
Scholarship Recipient, BioC 2017: Where Software and Biology Connect, Boston, MA	July 2017
Advanced Gene Mapping - Travel Stipend ($1000)	June 2017
Cancer and Systems Biology (CanSys) Program ($12,000 stipend)	Feb 2014-Dec 2014

EU-US Atlantis Global Scholars Program between the University at Buffalo,
Université du Luxembourg, and VU University Amsterdam

ORGANIZER

Columbus Predictive Modelers (Co-organizer)
satRdays Columbus 2020 (Co-organizer)

CONFERENCES ATTENDED

BioC 2016: Where Software and Biology Connect, Stanford, CA – June 24-26, 2016
American Society of Hematology 58th Annual Meeting, San Diego, CA – Dec 2-5, 2016
BioC 2017: Where Software and Biology Connect, Boston, MA – July 27-28, 2017
BMT Tandem 2018, Salt Lake City, UT — Feb 21-25, 2018
BioC 2018: Where Software and Biology Connect, Toronto, ON, Canada – July 25-27 2018
BioC 2019: Where Software and Biology Connect, New York, NY – June 24-27 2019
useR2019! Toulouse, France – July 9-12, 2019
rstudio::conf2020 San Francisco, California – Jan 27-30, 2020