Colby Ford

www.colbyford.com
colby.ford@uncc.edu
colbytylerford@gmail.com
(828) 773-8009

EDUCATION

UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE

Doctor of Philosophy
Computing & Information Systems - Bioinformatics & Computational Biology
Wayland H. Cato Doctoral Fellow

Master of Science
Data Science & Business Analytics
First Graduate of the Program

Bachelor of Arts
Applied Mathematics
Minor: Psychology
Coffey Foundation Scholar
C.L. Robbins Scholar

CALDWELL COMMUNITY COLLEGE & TECHNICAL INSTITUTE

Associate in Science
Associate in Arts

LINKS

Github:// colbyford
LinkedIn:// colbyford
Research Gate:// colby_ford
Google Scholar:// Colby T. Ford

SKILLS

PROGRAMMING

Languages:
R • Python • SQL • Scala
Perl • SAS • Visual Basic • Java
Parallel Computing:
Spark • Hadoop • MPI • SNOW
Markup/Design:
HTML5 • CSS3 • LaTeX

OCCUPATIONAL EXPERIENCE

UNC CHARLOTTE | Bioinformatics Researcher & Teaching Faculty

November 2018 - Present | Charlotte, NC

- Researched in a broad range of fields including infectious diseases, epistasis, and human phylogenetics.
- Teaching *Cloud Computing for Data Analysis* (DSBA 6190) for the Data Science Initiative Master's program. Received a funding grant from Microsoft to modernize and rebuild the course content on the Azure cloud platform.

BLUEGRANITE | Senior Data Scientist, AI Solution Architect

January 2017 - Present | Charlotte, NC

- Develop data and AI solutions using the Microsoft Azure platform such as Microsoft ML Server, Microsoft Azure Machine Learning, HDInsight (Hadoop), and Azure Databricks (Apache Spark).
- Manage client engagements including requirements gathering, project planning, and budgeting.
- Host training workshops, give conference presentations and demonstrations on Microsoft advanced analytics technologies.

LASH GROUP | Senior Data Science Lead

February 2016 - February 2017 | Fort Mill, SC

- Develop machine learning experiments and data analysis workflows to aid in client analytics.
- Discover, understand, and present insights into patient drug adherence.
- Consult in company data architecture including master data management and governance for future business strategy to grow value for client accounts.

MARINER | Data Scientist

October 2014 – February 2016 | Charlotte, NC

- Consult in the development of machine learning experiments, including parametric and non-parametric models, statistical predictions, and data mining.
- Build Azure cloud-based solutions for data collection, processing, and storage using Microsoft Azure Technologies such as Data Science Virtual Machines, Machine Learning Studio, Azure SQL Database, and more.
- Design and create interactive visualizations for both dashboarding and reporting using Microsoft PowerBI, Tableau Desktop & Tableau Server.

NORTH CAROLINA NEW SCHOOLS | Technology Facilitator & Mathematics Instructional Assistant

April 2014 – October 2014 | Hudson, NC

- Responsible for entire technology inventory: ordering, maintenance, management, etc. and maintained school website.
- Liaison between college & high school technology departments.
- Taught NCVPS mathematics courses, held additional teaching sessions in math and science and SAT & ACT preparation.

COURSEWORK

DOCTORAL

Bioinformatics Programming - Perl
Molecular Sequence Analysis
Bioinformatics Programming - Python
Energy and Interaction
Research Rotation I
Research Rotation II
Drugs: Molecular & Cell Mechanics
Computational Structural Biology
Statistics for Bioinformatics

MASTERS

Database Systems
Visual Analytics
Big Data Analytics
Business Intelligence
Database Design & Management
Consumer Marketing Analytics
Innovation Analytics
Advanced Business Analytics
Machine Learning
Data Warehousing

UNDERGRADUATE

Mathematical Dataset Modeling
Mathematical Thought
Applied Regression
Financial Mathematics
Regression Analysis
Industrial & Organizational Psychology
Deductive Logic
Deterministic Modeling
Cognitive Science - Perception
Business Programming - Visual Basic

RESEARCH

PUBLICATIONS

- **Ford C.**, Smith, K., Zenarosa, G., Williams, J., Janies D., (Coming Soon) (*Frontiers of Microbiology*) Persistence of Genes that Confer Antimicrobial Resistance in the History of *Escherichia coli* Genomes
- Smith, K., Zenarosa, G., **Ford C.**, Williams, J., Janies D., (April 2019) (*Presentation - National Council on Undergraduate Research - 2019 Conference*) Phlogenetic Analysis of the Genetic Variation of Multi-Drug Resistant *Escherichia coli*
- **Ford C.**, Wen J., Janies D., Shi X., (2018) (*Coming Soon*) parEBEN: A Parallelized Strategy for Improving Epistasis Analysis Based on Empirical Bayesian Elastic Net Models
- Wen J., **Ford, C.**, Janies D., Shi X., (2018) (*ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*) New strategies toward scaling up epistasis analysis on large-scale genomic datasets
- **Ford, C. T.** (2018). An integrated phylogeographic assessment of the Bantu migration. (Order No. 10748780, The University of North Carolina at Charlotte). *ProQuest Dissertations and Theses*, 120.
- **Ford, C.** & Lathrop, A. (2017). (*Presentation - Analytics for Social Good, University of Cincinnati.*) Predictive modeling of vegetation density using R and a cloud data platform.
- Janies, D.A., **Ford, C.,** Damodaran, L., Witter, Z., (2016) (*Online Journal of Public Health Informatics*) Spread of Middle East Respiratory Coronavirus: Genetic versus Epidemiological Data
- **Ford, C**., Xue, M., Whiteley, P.M., Wheeler, W., Janies, D.A., Shi, X. (2016) *(Society for Anthropological Sciences)* Visualizing Linguistic Disparity of Uto-Aztecan Languages and Bantu Languages
- **Ford, C.**, Nodzak, C.M., Uppal, A., Zhao, L., (2016) (*Critical Assessment of Genome Interpretation*) Prediction of the effect of naturally occurring missense mutations on cellular n-acetyl-glucosaminidase enzymatic activity.

OTHER WORK

- **Ford, C.** (2015). Demand forecasting using machine learning to reduce working capital. *Mariner White Paper.*
- **Ford, C.** (2015). The allure of machine learning, now within reach in Microsoft Azure. *Mariner White Paper.*
- **Ford, C.**, & Snyder, W. (2015). Revenue protection using machine learning for utilities management. *Mariner White Paper.*
- **Ford, C.**, (2016). Assessment of retail out-of-stock conditions using statistical inference. *Mariner White Paper.*
- Blog Posts Written for BlueGranite

BOOKS

- **Ford, C.** (2010). *Caesura, late-intermediate piano technique book.* Hudson, NC: Author.
  ISBN: 978-0-557-36832-7

TEACHING

Cloud Computing
Advanced Functions & Modeling
Discrete Mathematics
Precalculus
Machine Learning
SAT/ACT Prep

TRAINING

Microsoft ML Server
Azure Databricks
Microsoft Power BI
Tableau

RESEARCH

JANIES LAB | Disease Transmission Research

August 2015 – Present | Charlotte, NC

Working with **Dr. Dan Janies** to understand the disease transmission patterns of pathogens. Some examples include antimicrobial resistance of *E. coli*, pathogenicity changes in *P. vivax*, and the Middle Eastern Respiratory Syndrome coronavirus.

DARPA | Bantu Migration - Genomics and Language Research

November 2015 – May 2018 | Charlotte, NC

Working with researchers from DARPA, The American Museum of Natural History (**Drs. Ward Wheeler** and **Peter Whitley**), and The University of North Carolina at Charlotte (**Drs. Dan Janies** and **Xinghua (Mindy) Shi**) to understand the genetic, geographic, and linguistic relationship between languages in both Bantoid and Uto-Aztecan groups.

SHI LAB | Parallel Computing for Epistasis Detection

May 2018 – Present | Charlotte, NC

Working with **Drs. Xinghua (Mindy) Shi** and **Jia Wen** in the Shi Lab in epistasis research. We developed the parEBEN package as a parallel approach of the empirical Bayesian elastic net for detecting gene-gene interaction.

PROFESSIONAL MEMBERSHIPS

2013	The Society for Industrial and Applied Mathematics
2014	The American Statistical Association
2015	UNCC Data Science Initiative Advisory Board
2016	Northeastern University LEVEL Board Member and Lecturer
2016	The Society for Anthropological Sciences
2018	American Association for the Advancement of Science

CONFERENCE ATTENDANCES

Nov. 2014	Microsoft Roadmap Event (*Speaker*)	Charlotte, NC
Sep. 2015	Microsoft Cortana Analytics Conference	Seattle, WA
Sep. 2016	Advanced Pharma Analytics	Newark, NJ
Feb. 2017	Analytics for Social Good - U. Cincinnati (*Speaker*)	Cincinatti, OH
Mar. 2017	Society for Applied Anthropology (*Speaker*)	Santa Fe, NM
Nov. 2017	Nissan Analytics Expo (*Speaker*)	Nashville, TN
Jan. 2018	rstudio::conf 2018	San Diego, CA
Jun. 2018	Spark+AI Summit 2018	San Francisco, CA
Sep. 2018	Big Data Ignite (*Speaker*)	Grand Rapids, MI
Sep. 2018	Microsoft Ignite	Orlando, FL
Jan. 2019	rstudio::conf 2019	Austin, TX

CERTIFICATIONS

May 2018	Databricks Certified Developer - Apache Spark 2.x for Python

FUNDING

Jan. 2019	Azure Funding for Cloud Computing Course	$30,000