**Yolande Tra**

MITRE Corporation

7515 Colshire Drive C: (443) 255-9764

McLean, VA 22102 ytra@mitre.org

**EDUCATION**:

**University of Missouri**, Columbia MO

Doctor of Philosophy Degree in Statistics, May 2000.

*Thesis: Bayesian Analysis for Avian Nest Survival Models*

**University of Maryland University College**, Adelphi MD

Graduate Certificate in Predictive Analytics, 2015.

**Ball State University**, Muncie IN

Master of Arts Degree in Statistics, 1994.

**Languages:** R, SAS, Python.

**Software:** SAS/SQL, SAS/BASE, SAS/MACRO, SAS/EG, SAS/EM, SAS Forecast Studio, R, Jupyter, Palantir Foundry, SPSS, AMOS, Mplus, Iveware, Lisrel, RapidMiner, Tableau, Minitab, Excel, MS Access, Web design, MS Office Suites.

### Research Interests:

Predictive Modeling, Explainable AI, Neural Network, Deep Learning, Data Mining, Machine Learning, Text Analytics, Fraud Analytics, Mixed effect Model, Longitudinal Data Analysis.

### Professional Experience:

*Senior Data Scientist* (2015 – Present) – MITRE Corporation.

### Samples of work:

* Data Scientist /Data Analyst/SME
* Enhance Foot Traffic Data with demographic Census data for potential impact on market competition and healthcare policy.
* Perform Fuzzy String Matching to support the Administration for Strategic Preparedness and Response (ASPR) agency: Artificial Intelligence, Machine Learning.
* Assist Testing and Diagnostic Working Group (TDWG): Data Analysis, Dashboard Summary and Visualization of COVID-19 Testing for the Nation.
* Inflation Reduction Act (IRA-CMMI): Rebate Analytics, Drug Negotiations.
* Attrition Survival Analysis (USDA-FSIS): Machine Learning, Predictive Analytic, Visualization.
* Demographic Frame (Census): Gap Analysis between Residential Candidate File (RCF) and Demographic Person Place Model.
* Climate Impact Frame (Census): Build residential histories (2010-2021).
* SUPPORT for Patients and Communities Act Section 2001 (CMS): research and evaluation delivered to Congress – statistical modeling.
* Person Place Model from Administrative Records (Census): comparison with 2020 Decennial Census – demographic multiple imputation.
* ATO Safety Analysis (FAA). Risk Analysis, Machine Learning, operational forecasting capability.
* AI Health Outcome Challenge Prize (CMMI). Baseline model, cheating analysis with ML, performance metrics.
* Advanced Analytics (VBA). SQL (POSTGRES), Modeling-Visualization (R).
* Part D Modernization Model Benchmark (CMMI). Developed models for predicting catastrophic spending using tree-based machine learning (ML). Benchmark Model would reduce Medicare Part D prescription drug cost and maintain healthcare quality of beneficiaries.
* Identity Theft Fraud – (IRS/ISAC): clustering of leads and related confirmed fraud, related to tax return.
* Value Based Insurance Design (VBID/CMMI)): key drivers of related and unrelated hospice cost using ML decision tree for creating capitation rates.
* National Work Queue (NWQ) claims routing Analysis (VA): clusters of claims using trajectory.
* Secondary screening (DHS/CBP): Mitigate unnecessary referrals at the land Ports of Entry for Lost and/or Stolen US-issued travel documents
* Diagnostics Utilization (MIP-BCH): EHR data, Machine Learning to predict medication Safety Events, Visualization (Tableau).
* Payment Integrity Research and Analysis capability- Education Department (PIRAC/ED): predictive model, improper payment risk for grants resulting in education grants monetary loss.
* U.S. Postal Service’s Workers Compensation program (USPS): strong indicators leading to potential fraud activity by healthcare providers
* Complex software acquisitions: reliably predict project outcomes. Insight into the impact of controllable factors/attributes of project success that hampers decision making.
* Department of Health and Human Services (HHS/OASH): association of patient-reported experience of medical team communication with disparities in hospital readmissions, Reduce Racial/Ethnic and Other Disparities in Preventable Adverse Health Care-Associated Outcomes.
* Medicare Patient Safety Monitoring System(MPSMS): factors affecting adverse drug events (ADEs)

***Statistical and Data Analytics Consultant*** (2012-2015) – Private consulting. Developed analytics solutions with data mining and machine learning techniques using Rapid Miner, R, and Tableau and SAS Enterprise Miner software. Selected achievements include

* Household annual spending prediction on Drugstore Products; score prospective households, identify dominant attribute to guide high-end advertisement.
* Text Mining, Topic modeling.
* Classification of women’s choice of contraception using neural network and random forest models: high degree of accuracy achieved with random forest.
* Determined optimum placement in a recruitment business using decision Tree algorithm to increase revenue by 10%. Interactive KPI Dashboard to display success rate by team and region helped identify geographic expansion to boost revenue.
* Build credit risk model using logistic regression to predict the probability of default, report scorecard points. Credit risk prediction using neural network: predict the “DO NOT LEND” category.
* Key drivers’ identification of cross-country variations in pharmaceutical markets; leverage key drivers to develop factors and deploy to assess market opportunities around the world (Market research analytics).

***Corporate R trainer*** (2014-2015) –NobleProg USA. Delivered online/live training to government researchers, corporate employee, and senior manager.

***Manager of Statistics and Informatics*** (2013 – 2014) -FQHC-CareFirst Patient Centered Medical Home (PCMH) Project, Total Health Care, Inc., Baltimore, MD,

Private consulting. Resolved data redundancy and missing values problems.

***Statistician*** (2010 - 2015) - University of Maryland Baltimore, School of Pharmacy.

***Assistant Professor of Statistics*** (2004-2010)- Rochester Institute of Technology

***Visiting Assistant Professor of Statistics – Lecturer*** (2000-2004) – University of Michigan

### Courses Taught:

Introduction to R; Statistics and R programming for Researchers, Microarray Data Analysis using R, Design of Experiments, SAS programming, Statistical Quality Control, Applied Statistics I and II, Nonparametric Statistics, Probability and Statistics I and II, Regression Analysis, Data Analysis I and II, Introduction to Statistics and Data Analysis, Introduction to Statistical Reasoning.

**Selected Publications from** [Google scholar](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C47&q=yolande+tra&btnG=) (pp.1-2)

1. Scannell, D., Desens, L., Day, D. S., & **Tra, Y**. (2022). Combatting Mis/Disinformation: Combining Predictive Modeling and Machine Learning with Persuasion Science to Understand COVID-19 Vaccine Online Discourse. *Medical Research Archives*, 10(5).

2. Scannell D, Desens L, Guadagno M, **Tra Y**, Acker E, Sheridan K, Rosner M, Mathieu J, Fulk M. (2021). COVID-19 Vaccine Discourse on Twitter: A Content Analysis of Persuasion Techniques, Sentiment and Mis/Disinformation. *J Health Communication*. 26(7):443-459.

3. Klein LW, **Tra Y**, Garratt KN, Powell W, Lopez-Cruz G, Chambers C, Goldstein JA; Society for Cardiovascular Angiography and Interventions. (2015). Occupational health hazards of interventional cardiologists in the current decade: Results of the 2014 SCAI membership survey*. Catheter Cardiovasc Interv*. 2015 Nov;86(5):913-24.

4. **Tra, Y**., Frisina, R. D., & D'Souza, M. (2011). A novel high-throughput analysis approach: immune response-related genes are upregulated in age-related hearing loss. *Open Access Bioinformatics*, 3, 107-122.

5. **Tra Y.V**, Evans IM. (2010). Enhancing interdisciplinary mathematics and biology education: a microarray data analysis course bridging these disciplines. *CBE Life Sci Educ*. 2010 Fall;9(3):217-26.

6. **Tra, Y**. (2007). Even You Can Learn Statistics: A Guide for Everyone Who Has Ever Been Afraid of Statistics. *The American Statistician*, 61:2, 182.