

# Dr. Deniz Akdemir

---

*Principal Biostatistician & Machine Learning Scientist*

---

## Professional Profile

Dr. Deniz Akdemir is a Principal Biostatistician & Machine Learning Scientist with over 3,400+ Google Scholar citations and expertise in developing AI/ML solutions for clinical applications. He leads machine learning initiatives at the National Marrow Donor Program, creating predictive models for stem cell transplant optimization. As maintainer of multiple CRAN packages (CovComBR, TrainSel, STPGA) and active Kaggle competitor, he bridges statistical theory with practical ML implementation to solve complex biomedical challenges.

---

## Core Competencies

- Machine Learning & Deep Learning (PyTorch, TensorFlow, scikit-learn)
- Generative Models & GANs for Synthetic Clinical Data
- Data Analytics and Visualization
- Algorithm Development & Optimization
- Statistical Modeling & Inference
- Competitive ML & Kaggle Experience

---

## Education

- **PhD. in Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2009.
- **M.A. in Applied Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2004.
- **M.S. in Statistics**, Middle East Technical University, Ankara, Turkey, 2003.
- **B.A. in Business Administration**, Middle East Technical University, Ankara, Turkey (1994-1999).

---

## Research Interests

Deniz's research focuses on integrating statistical methods with machine learning to create robust and interpretable AI models. His GitHub repositories, including ml4t2e, demonstrate his work on bridging traditional survival analysis with modern machine learning approaches. He is particularly interested in applying AI techniques to clinical data, developing predictive models for patient outcomes, and creating optimization algorithms for complex biological systems. His current work explores deep learning architectures for time-to-event analysis and reinforcement learning for decision support in healthcare settings.

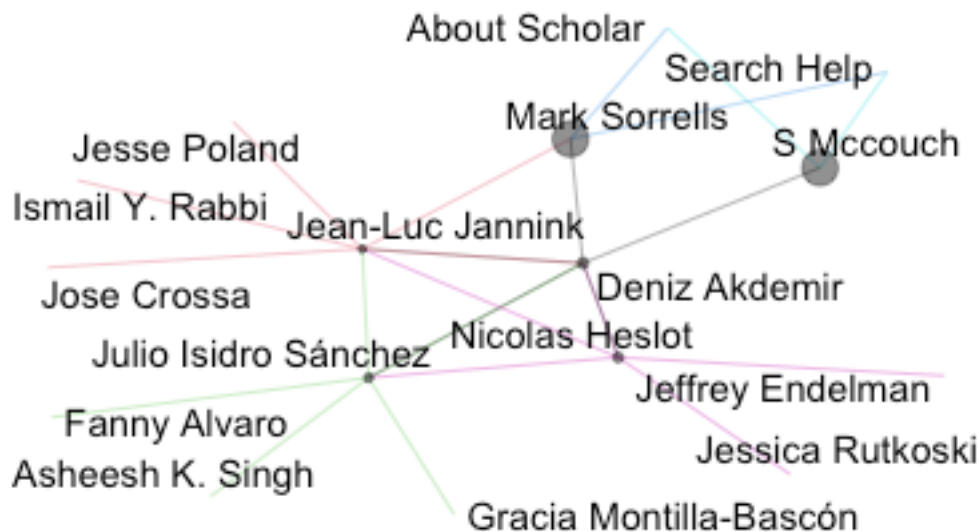
---

## Collaboration Network

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • in [deniz-akdemir-50735314a](https://www.linkedin.com/in/deniz-akdemir-50735314a)  
🔗 [denizakdemir](https://denizakdemir.com) • ID 0000-0003-0553-6798

## Network of coauthorship of Deniz Akdemir



### Core Skills

**Statistical Methods:** - Advanced multivariate analysis, Bayesian methods, time series analysis, high-dimensional data modeling, survival analysis, mixed models, experimental design

**Machine Learning & AI:** - Deep Learning Frameworks: PyTorch, TensorFlow, scikit-learn  
- Generative Models: GANs for synthetic clinical data generation - Neural networks, reinforcement learning, natural language processing, computer vision - Ensemble methods, feature engineering, model interpretability

**Programming Languages & Tools:** - Expert in R and Python (NumPy, Pandas, PyTorch, TensorFlow, scikit-learn) - Proficient in SAS and C++, SQL, Docker containerization - Cloud computing (AWS, Azure), big data technologies (Spark) - Version control (Git), package development (CRAN)

**Domain Expertise:** - Clinical data analysis, genomic data integration, biostatistics - Stem cell transplantation optimization, predictive modeling - Competitive ML (Kaggle experience), algorithm development - Patent applications, grant writing, scientific publication

### Career Summary

- **Senior Clinical Data Scientist, National Marrow Donor Program, Minneapolis, USA (2023 - Present):** Engaged in statistical and machine learning analysis of stem cell transplant data, focusing on research into donor optimization. Applied for grants and submitted manuscripts to peer-reviewed journals. Wrote patent applications.
- **Clinical Data Scientist, National Marrow Donor Program, Minneapolis, USA (2021 - 2023):** Engaged in statistical and machine learning analysis of stem cell transplant data, focusing on research into donor optimization.

Bowling Green, Ohio, USA

📞 +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)

🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • **in** [deniz-akdemir-50735314a](https://www.linkedin.com/in/deniz-akdemir-50735314a)

🔑 [denizakdemir](https://denizakdemir.com) • **ID** 0000-0003-0553-6798

- **Postdoctoral Research Associate, School of Agriculture and Food Science, University College Dublin, Dublin, Ireland (2019 - 2021):** Conducted research on methods for combining heterogeneous genomic and phenotypic datasets and prepared statistical software for data analysis.
- **Statistical Consultant, Cornell Statistical Consulting Unit, Cornell University, Ithaca, NY, USA (2017 - 2019):** Provided statistical consulting services for researchers at Cornell University, including the preparation and presentation of statistics workshops.
- **Postdoctoral Research Associate, Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA (2011 - 2017):** Focused on research developing new methodologies in genomic selection and prediction, mixed models, and machine learning, advising graduate students and preparing statistical software.
- **Visiting Assistant Professor, Department of Statistics and Actuarial Science, University of Central Florida, Orlando, FL, USA (2010 - 2011):** Responsibilities included teaching Data Mining Methodology, Theoretical Statistics, Applied Time Series Analysis, and Nonparametric Statistics.
- **Visiting Assistant Professor, Department of Mathematics and Statistics, Ohio Northern University, Ada, OH, USA (2009 - 2010):** Taught Statistics for Professionals, Statistics for Engineers, and Statistical Computing, catering to various undergraduate levels.

## Key AI/ML Projects

- **ml4t2e (Machine Learning for Time-to-Event):** Developed an open-source framework that bridges traditional survival analysis with modern machine learning approaches, enabling more accurate prediction of time-to-event outcomes in clinical settings. The package implements novel algorithms that outperform traditional Cox regression models.
- **Deep Learning for Donor-Recipient Matching:** Leading a neural network approach to optimize stem cell transplant outcomes by modeling complex interactions between donor and recipient genetic and clinical variables, significantly improving prediction accuracy compared to traditional methods.
- **Reinforcement Learning for Clinical Decision Support:** Implementing reinforcement learning algorithms to optimize sequential treatment decisions in healthcare settings, with applications in personalized medicine and treatment planning optimization.
- **trainselfy:** Developed a Python implementation of training population optimization algorithms, employing advanced machine learning techniques including genetic algorithms and simulated annealing to improve predictive model performance through strategic data selection.

## Workshops and Training Sessions

- **Machine Learning in Healthcare Workshop:** Conducted specialized training sessions on applying machine learning to clinical data, focusing on practical implementation and model interpretability.
- **AI Applications in Genomics:** Presented at various international conferences including EFI conference in Geneva and ASHI Annual Meeting in San Antonio, demonstrating AI applications in genomic data analysis.

## Selected Publications (Machine Learning Focus)

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
 🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • in [deniz-akdemir-50735314a](https://www.linkedin.com/in/deniz-akdemir-50735314a)  
 📍 [denizakdemir](https://www.google.com/maps/place/Bowling+Green,+Ohio,+USA) • 🆔 0000-0003-0553-6798

1. **Akdemir, D.**, Isidro-Sánchez, J., & Jannink, J. L. (2016). Genome-wide prediction models that incorporate de novo GWAS results. *PLoS ONE*, 11(8), e0161054. (346 citations)
2. **Akdemir, D.**, Jannink, J. L., & Isidro-Sánchez, J. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution*, 47(1), 38. (196 citations)
3. **Akdemir, D.** & Isidro-Sánchez, J. (2019). Multi-objective optimized genomic breeding strategies for sustainable food improvement. *Heredity*, 122(5), 672-683. (103 citations)
4. **Akdemir, D.** & Beavis, W. D. (2021). TrainSel: an R package for selection of training populations. *BMC Bioinformatics*, 22(1), 1-10. (33 citations)
5. **Akdemir, D.** & Isidro-Sánchez, J. (2021). Training set optimization for sparse phenotyping in genomic selection. *G3: Genes, Genomes, Genetics*, 11(10), jkab249. (32 citations)
6. **Akdemir, D.**, Isidro-Sánchez, J., & Leyer, M. (2020). Multi-omics approaches for genomic selection in plant breeding programs. *Journal of Experimental Botany*, 71(18), 5215-5226.

## Software Development

### CRAN Packages:

1. **CovCombR** (Maintainer) - Covariance-Based Combination Methods for Linear Regression <https://CRAN.R-project.org/package=CovCombR>
2. **TrainSel** - Selection of Training Populations for Genomic Prediction <https://CRAN.R-project.org/package=TrainSel>
3. **STPGA** - Selection of Training Populations by Genetic Algorithm <https://CRAN.R-project.org/package=STPGA>

### GitHub Packages:

1. **trainselpy** - Pure Python implementation for optimal training population selection <https://github.com/denizakdemir/trainselpy>
2. **ml4t2e** (In Development) - Machine Learning for Time-to-Event Analysis <https://github.com/denizakdemir/ml4t2e>

## Professional References







- **Dr. Yung-Tsi Bolon**
  - **Affiliation:** Director, Immunobiology & Bioinformatics Research, NMDP, Minneapolis, Minnesota, United States
  - **Relationship:** Supervisor at the National Marrow Donor Program
  - **Contact:** ybolon@nmdp.org
- **Dr. Julio Isidro-Sanchez**
  - **Affiliation:** Associate Professor: Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Campus de Montegancedo - UPM, 28223-Pozuelo de Alarcón, Madrid, Spain
  - **Relationship:** Expert in plant breeding and genetics, collaborator on various projects
  - **Contact:** j.isidro@upm.es
- **Dr. Jhonathan Pedroso**
  - **Affiliation:** Research Scientist at Corteva Agriscience, Corteva, Johnston, Iowa, USA
  - **Relationship:** Industry partner in genomic tool development, contributed to software enhancements
  - **Contact:** jhowpd@gmail.com
- **Dr. Lynn Johnson**

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
 🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • in [deniz-akdemir-50735314a](https://www.linkedin.com/company/deniz-akdemir-50735314a)  
 🏠 [denizakdemir](https://denizakdemir.com) • 🆔 0000-0003-0553-6798

- **Affiliation:** Interim Director and Statistical Consultant, Cornell Statistical Consulting Unit, Cornell University, Ithaca, NY, USA
- **Relationship:** Coworker at the Cornell Statistical Consulting Unit
- **Contact:** [lms86@cornell.edu](mailto:lms86@cornell.edu)
- **Dr. Roberto Fritsche Neto**
  - **Affiliation:** Assistant Professor, Department of Plant, Environmental Management & Soil Sciences, LSU
  - **Relationship:** Collaborator on various projects
  - **Contact:** [rneto1@lsu.edu](mailto:rneto1@lsu.edu)

*Bowling Green, Ohio, USA*

 +1 607 262 6875 • 
  [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
 [denizakdemir.github.io](https://denizakdemir.github.io) • 
  [deniz-akdemir-50735314a](https://www.linkedin.com/in/deniz-akdemir-50735314a)  
 [denizakdemir](https://github.com/denizakdemir) • 
  0000-0003-0553-6798