Dr. Anwar Said

Office: A417 ISIS building 1025, 16th Avenue, South, Nashville, TN, 37212, USA
@ anwar.said@vanderbilt.edu; Homepage Google Scholar GitHub

About

I am currently a postdoctoral research scholar at Vanderbilt University, TN. My research belongs to the area of graph machine learning (GML), an emerging field having far-reaching applications in a diverse range of industries, from drug discovery and optimization to recommendation and forecasting. My research mainly focuses on the design of novel GML approaches to enhance their performance and apply them to solve complex real-world problems. Working closely with a team of experts, I have developed several GML approaches that have been applied to problems such as circuit design completion, link prediction in Ethereum data, and fraud detection in social networks. In addition to my work in GML, I also work in data science, graph theory, and network science. Recently, I have expanded my research to include GML applications in the field of electronic design automation as well as neuroscience. Through my research, I strive to make meaningful contributions to the field of GML and improve our understanding of the complex systems that shape our world.

Education

Postdoctoral Research Scholar Mar. 2021-Present

- Institute for Software Integrated Systems,
  Department of Computer Science,
  Vanderbilt University, TN, USA.
- Advisor: Prof. Xenofon Koutsoukos
- Area: Graph machine learning and its applications

Ph.D. Computer Science Nov. 2021

- Information Technology University, Lahore, Pakistan
- Advisors: Prof. Saeed-Ul Hassan, Prof. Mudassir Shabbir
- *Dissertation:* Novel Graph Representations for Machine Learning Applications

M.Phil. Computer Science Aug. 2016

- Quaid-i-Azam University, Islamabad, Pakistan
- *Dissertation:* Discovering Community Structure of Complex Networks
- Major Area of Study: Social Network Analysis, Graph Theory
- Minor Area of Study: Artificial Intelligence

M.Sc. Computer Science with Distinction Dec. 2013

- University of Swat

Industry Experience

- **C++ Programmer** Jul. 2014 - Sep. 2017
  Shaheen Foundation, Islamabad

  As a C++ programmer, my role primarily revolved around the development and implementation of advanced QT applications for a diverse range of clients. I worked

in close collaboration with a team of professionals, where my responsibilities involved not only the programming aspects but also requirement engineering, model prototyping, testing, and seamless deployment of high-quality solutions. My proficiency in C++ programming and extensive experience with QT allowed me to deliver bespoke solutions that met the unique and complex requirements of the clients.

- **Graph Data Scientist** May 2021 - Mar. 2022
  BUILD & CODE Germany
  As a Graph Data Scientist, my responsibilities include developing advanced machine learning models for the completion of a huge knowledge graph, which contained over 10 billion data points. It was a challenging but rewarding experience to work on a project of such a massive scale, and it allowed me to gain valuable experience in handling and processing big data. I worked on a range of advanced machine learning techniques, including Graph Neural Networks (GNNs), classical ML methods, Neo4J and Cypher.

Research Experience

Vanderbilt University
Postdoctoral Research Scholar Mar. 2022 - Present
Research Interests: Graph Machine Learning, Graph Descriptors, Circuit Design Completion, Data Science, Graph Theory

Information Technology University, Lahore Sep. 2017 - Oct. 2022
PhD Fellow
Projects: Graph Machine Learning, Social Network Analysis, Altmetrics, Ethereum Network Analysis, Call Data Record Analysis, Federated Learning

Master Student: Quaid-i-Azam University, Islamabad Aug. 2016 - Sep. 2017
Research areas: Community detection in social networks, Genetic Algorithms, Machine Learning

Teaching Experience

Information Technology University, Lahore
Teaching Assistant for: Aug. 2018 - June 2020

- Advanced topics in social Network Analysis (Spring 2018)
- Information retrieval (Fall 2018)
- Database (Spring 2019)
- Data Science (Fall 2019)
- Data Structures (Spring 2020)

Instructor at Hira and Khyber Technical College, Kabal, Swat 2009-2011

- Physics (Undergrad/college level)
- Surveying (Undergrad/college level)
- Mathematics (Undergrad/college level)
- Chemistry (college level)

Awards and Honors

- Recipient of the Vanderbilt University, *Postdoctoral Research Scholar Fellowship*
- Recipient of the *travel grant* for attending the "Youth in High Dimensions: Machine Learning, High-dimensional Statistics, and Inference for the New Generation" Conference, 2020, ICTP, Trieste, Italy.
- Recipient of the Information Technology University, *Ph.D. Fellowship*

- Recipient of the *Gold Medal* from University of Swat for achieving highest academic ranking

Tools Developed

- I have developed NeuroGraph, a tool for preprocessing fMRI data and graph-based neuroimaging benchmarks. For more details, visit NeuroGraph's website, documentation and paper
- Released an initial version (WiP) of our web-based application for benchmarking graph neural networks (2023). (Link to the app)
- Released a python package with reproducability check by CodeOcean for our graph augmentation framework appeared in publication entitled "On Augmenting Topological Representations for Attributed Graphs", Applied Soft Computing, 2023. (Link to the package)
- Released a python library for our graph descriptor appeared in publication entitled "DGSD: Distributed graph representation via graph statistical properties", Future Generation Computer Systems (2021). (Link to the package)
- Released Python package for CC-GA, our community detection algorithm appeared in publication entitled "CC-GA: A clustering coefficient based genetic algorithm for community detection in social networks (2018)". Every two weeks, this repository has an average of **100** number of clones. (Link to the package)

Publications

Under review

1. **Said, A.,** Shabbir, M., Abbas, W., Derr, T., Koutsoukos, X., Empowering Graph Neural Networks using Subgraph Embeddings, ICMLA2023
2. Ali, A., **Said, A.**, Safder, I., Hassan, S., Aljohani, N.R., Shabbir, M., MSDGSD: A Scalable Graph Descriptor for Processing Large Graphs, IEEE Transactions on Computational Social Systems
3. **Said, A.**, Ahmad, U, Abbas, W., Shabbir, M., & Koutsoukos, X. Network Controllability Perspectives on Graph Representation, IEEE Transaction on Knowledge and Data Engineering (TKDE).

Published

1. **Said, A.**, Derr, T., Shabbir, M., Abbas, W., Koutsoukos, X., Graph Unlearning: A review. ArXiv preprint arXiv:5075426
2. **Said, A.**, Bayrak, R. G., Derr, T., Shabbir, M., Moyer, D., Chang, C., & Koutsoukos, X. (2023). NeuroGraph: Benchmarks for Graph Machine Learning in Brain Connectomics. arXiv preprint arXiv:2306.06202.
3. Athar, S., Abbasi, R. A., Saeed, Z., **Said, A.**, Razzak, I., & Salim, F. D. (2023). ASBiNE: Dynamic Bipartite Network Embedding for incorporating structural and attribute information. World Wide Web, 1-19.
4. Ahmed, A., **Said, A.**, Shabbir, M., Koutsoukos, X., Sequential Graph Neural Networks for Source Code Vulnerability Identification, ACM Symposium on the Science of Security 2023.
5. **Said, A.**, Shabbir, M., Hassan, S. U., Hassan, Z. R. & Ahmed, A. On Augmenting Topological Graph Representations for Attributed Graphs, Applied Soft Computing journal, Elsevier (2023).

6. **Said, A.**, Shabbir, M., Broll, B, Volgyesi, P, Abbas, W, & Koutsoukos, X., "Circuit Design Completion Using Graph Neural Networks", journal of Neural Computing and applications (2023).
7. Sandborn, M., Olea, C., **Said, A.**, Shabbir, M., Volgyesi, P., Koutsoukos, X., White, J., What a drag! Streamlining the UAV design process with design grammars and drag surrogates, the 2022 International Conference on Computational Science & Computational Intelligence (CSCI'22)
8. Attaullah, Abbasi, R A., Khattak, A. and **Said, A.** Identifying Misinformation Spreaders: A Graph-Based Semi-Supervised Learning Approach, Multimedia Evaluation Benchmark Workshop 2022
9. Mian, A., Shah, S, Ullah, S., **Said, A.**, Heimerl, K., & Crowcroft, J. A Value-Added IoT Service For Cellular Networks using Federated Learning, computer networks (2022).
10. **Said, A.**, Janjua, M. U., Hassan, S. U., Muzammal, Z., Saleem, T., Thaipisutikul, T., ... & Nawaz, R. (2021). Detailed analysis of Ethereum network on transaction behavior, community structure and link prediction. PeerJ Computer Science, 7, e815.
11. **Said, A**., Hassan, S. U., Tuarob, S., Nawaz, R., & Shabbir, M. (2021). DGSD: Distributed graph representation via graph statistical properties. Future Generation Computer Systems, 119, 166-175.
12. **Said, A.**, Hassan, S. U., Abbas, W., & Shabbir, M. (2021). NetKI: A kirchhoff index based statistical graph embedding in nearly linear time. Neurocomputing, 433, 108-118.
13. Hassan, S. U., Shabbir, M., Iqbal, S., **Said, A.**, Kamiran, F., Nawaz, R., & Saif, U. (2019). Leveraging Deep Learning and SNA approaches for Smart City Policing in the Developing World. International Journal of Information Management, 102045.
14. **Said, A.**, Bowman, T. D., Abbasi, R. A., Aljohani, N. R., Hassan, S. U., & Nawaz, R. (2019). Mining network-level properties of Twitter altmetrics data. Scientometrics, 120(1), 217-235.
15. **Said, A.**, Shah, S., Farooq, H., Mian, A., Imran, A., & Crowcroft, J. (2018). Proactive Caching at the Edge Leveraging Influential User Detection in Cellular D2D Networks. Future Internet, 10(10), 93.
16. Imran, M., Akhtar, A., **Said, A.**, Safder, I., Hassan, S. U., & Aljohani, N. R. (2018, September). Exploiting Social Networks of Twitter in Altmetrics Big Data. In 23rd STI 2018 conference, September 12-14, 2018, Leiden, The Netherlands. Centre for Science and Technology Studies (CWTS).
17. **Said, A.**, Abbasi, R. A., Maqbool, O., Daud, A., & Aljohani, N. R. (2018). CC-GA: A clustering coefficient based genetic algorithm for detecting communities in social networks. Applied Soft Computing, 63, 59-70.

Volunteering as PC Member/Co-organizer

- (PC Member) - GTA^3: The 7th Workshop on Graph Techniques for Adversarial Activity Analytics, IEEE Big Data Conference 2023, Sorrento, Italy
- (PC Member)- The 17th Annual Vanderbilt Postdoctoral Association Symposium 2023, TN, USA
- (Co-organizer) KDD23 Data Science for Social Good Workshop at Long Beach, USA

Referee Services Conferences

- International Conference on Machine Learning (ICML) 2020 - Present
- Neural Information Processing Systems (NeurIPS) 2021 - Present
- Knowledge Discovery and Data Mining (SIGKDD) 2023 - Present
- International Conference on Web and Social Media (ICWSM) 2023-Present

Referee Services Journals

- Information Sciences Elsevier 2022 - Present
- Big Data Research Elsevier 2022 - Present
- Network Science 2022 - Present
- Swarm and Evolutionary Computation (Elsevier) 2017 - Present
- IEEE Access 2019 - Present
- IEEE Transactions on Computational Social Systems 2020-Present
- ACM Transaction on Social Computing 2020-Present
- Journal of Ambient Intelligence and Humanized Computing (AIHC) 2021

Mentoring (not an advisor)

Quaid-i-Azam University, Islamabad

- Attaullah, M.Phil. Computer Science 2022
  - Co-authored "Identifying Misinformation Spreaders: A Graph-Based Semi-Supervised Learning Approach" Multimideia Eval 2022
- Sajjad Athar M.Phil. Computer Science 2021
  - Co-authored "ASBiNE: Dynamic Bipartite Network Embedding for Incorporating Structural and Attribute Information" ES2021

Information Technology University, Lahore

- Muhammad Ali, Master in Computer Science 2022
  - Co-authored "MSDGSD: A Scalable Graph Descriptor for Processing Large Graphs" ITCSS 2022
- Ammar Ahmed, Bachelor of Computer Science 2020
  - Co-authored "Sequential Graph Neural Networks for Source Code Vulnerability Identification" HotSoS2023
- Muhammad Huzaifa, Bachelor of Computer Science 2019

Invited talks/guest lectures

- Gave a guest lecture at Tennessee State University on "Machine Learning at the Forefront of Molecular Analysis" (march 2023)
- Graph Machine Learning in Circuit Design Completion 2022
  Vanderbilt Postdoctoral Association Symposium
- Graph Neural Networks for Knowledge Graph Completion 2022
  Build & Code
- On the Analysis of Cryptocurrency Networks 2021
  Blockchain Lab, ITU
- Community Detection in Social Networks 2020
  Intelligent Machine Lab, ITU
- Network Analysis for Social Good 2020
  STI conference 2020

Certification (courses)

- Responsible Conduct of Research course (RCR) (Vanderbilt) 2022
- RCR training (Vanderbilt) 2022
- Social Network Analysis (Vanderbilt)
- Deep Learning (Coursera)
- Machine Learning (Coursera)
- Probability Theory (Coursera)
- Linear Algebra (Online MIT)
- Graph Machine Learning (Online Stanford)
- Social Network Analysis (Online Stanford)
- 1-year survey diploma (KPK TTB)
- MS Office (SDC Peshawar)
- AutoCAD (SDC Peshawar)

Volunteering

- Volunteer services for Islamic Center of Nashville during Ramadan (2023)
- Member of VPA Symposium planning committee (2023)
- Volunteer at ICML 2021
- Volunteer at NeurIPS 2022
- Intro to CS @ Govt. Middle School Dhero (middle school students) 2020
- Intro to AI @ Govt. High School Kabal (high school students) 2019
- Intro to OOP @ Govt. JPG Swat (undergraduate students) 2019
- Volunteer flood relief in Swat 2011
- Volunteer with "save the children" NGO, Swat 2010
- Volunteer earthquake relief in Abbottabad 2005

Programming Tools

- Python, Networkx, TensorFlow, Scikit-learn, PyG, PyTorch, Deep Graph Library, Multiprocessing, SQL, Neo4J, Cypher, C++, AWS, Jira, iGraph, Message Passing Interface (MPI), Socket Programming, Tableau, Docker, MLOps, Streamlit

Research Impact

Number of citations: 263
Impact factor: 69