Kanchan Chowdhury

Webpage: kanchanchy.github.io
Linkedin: linkedin.com/in/kanchan-chowdhury-5729699a

Email : kchowdh1@asu.edu
Mobile : +1-xxx-xxx-xxxx

RESEARCH INTERESTS

Machine Learning Systems, Database Systems, and Geospatial Data Analytics

EDUCATION

- **Arizona State University** Tempe, Arizona
  PhD & MS in Computer Science *Aug. 2018 – July 2024*
  Advisor: Prof. Mohamed Sarwat & Prof. Jia Zou, CGPA: 4.00

- **Chittagong University of Engineering and Technology** Chittagong, Bangladesh
  Bachelor of Science in Computer Science and Engineering *Mar. 2010 – Nov. 2014*
  Advisor: Prof. Mohammed Moshiul Hoque, CGPA: 3.76

EXPERIENCE

- **Arizona State University** Tempe, Arizona
  Postdoctoral Research Scholar *Aug. 2024 - Present*
  Mentor: Prof. Jia Zou
  - **Responsibilities**:
    1) Conducting research on optimizing end-to-end pipelines consisting of database queries and machine learning models for inference workloads, reducing runtime and memory usage of in-database machine learning systems.
    2) Helping the mentor in writing funding proposals, reviewing research papers, and co-mentoring graduate students.

- **Arizona State University** Tempe, Arizona
  Research Assistant *Aug. 2018 - July 2024*
  - **Research Projects**:
    1) Co-optimization of machine learning and join queries based on model decomposition and join push-down.
    2) Designing and implementing a deep-learning and data processing system for raster imagery and vector datasets.
    3) Re-partitioning geospatial datasets to reduce spatial model training time and memory usage.

- **Wherobots Inc.** Scottsdale, Arizona
  Research and Development Intern *Jan. 2023 - Aug. 2023*
  - **Responsibilities**: Designing and developing spatial machine learning and deep learning tools, scalable map-matching, and geospatial data analytical algorithms. Integrating the developed tools into the Wherobots cloud platform.

- **Gagagugu PTE LTD** Dhaka, Bangladesh
  Software Engineer *Jan. 2017 - Jun. 2018*
  - **Responsibilities**: Developing Android Apps with social networking features such as calling, messaging, and posting.

- **Le Chef Plc** Dhaka, Bangladesh
  Android Application Developer *Jan. 2015 - Dec. 2016*
  - **Responsibilities**: Developing Android Apps featuring online order and reservation services for restaurants in the UK.

TEACHING

- **Instructor** Fall 2022
  Spatial Data Science and Engineering (CSE 594) *Arizona State University*
  - **Responsibilities**:
    1) Preparing lecture slides and conducting lectures.
    2) Preparing assignments, projects, exam questions, and grading rubrics.
    3) Office hours to help students understand lectures and projects.

- **Teaching Assistant**   Fall 2021, Spring 2021, Fall 2020, and Spring 2020
  Distributed Database Systems (CSE 512)   *Arizona State University*

- **Teaching Assistant**   Fall 2019
  Data Processing at Scale (CSE 511)   *Arizona State University*

- **Teaching Assistant**   Spring 2019
  Object-Oriented Programming & Data Structure (CSE 205)   *Arizona State University*

- **Teaching Assistant**   Spring 2019
  Principles of Programming with C++ (CSE 100)   *Arizona State University*

- **Teaching Assistant**   Fall 2018
  Principles of Programming with Java & Python (CSE 100)   *Arizona State University*

PUBLICATIONS

- **Kanchan Chowdhury**, Lulu Xie, Jia Zou; ExBoost: Out-of-Box Co-Optimization of Machine Learning and Join Queries. *To Appear in International Conference on Database Systems for Advanced Applications 2025*

- Hong Guan, Lixi Zhou, **Kanchan Chowdhury**, Lulu Xie, Jia Zou; Privacy and Accuracy-Aware AI/ML Model Deduplication. *To Appear in ACM SIGMOD International Conference on Management of Data 2025*

- Hong Guan, Lixi Zhou, **Kanchan Chowdhury**, Jia Zou; Declarative Privacy-Preserving Inference Queries. *To Appear in International Conference on Database Systems for Advanced Applications 2025*

- **Kanchan Chowdhury**, Mohamed Sarwat; Deep Learning with Spatiotemporal Data: A Deep Dive into GeotorchAI. *40th International Conference on Data Engineering (ICDE)*, 2024

- Lixi Zhou, Qi Lin, **Kanchan Chowdhury**, Saif Masood, Jia Zou; Serving Deep Learning Models from Relational Databases. *27th International Conference on Extending Database Technology (EDBT)*, 2024

- **Kanchan Chowdhury**, Mohamed Sarwat; A Demonstration of GeoTorchAI: A Spatiotemporal Deep Learning Framework. *ACM SIGMOD International Conference on Management of Data*, 2023

- **Kanchan Chowdhury**, Vamsi Meduri, Mohamed Sarwat; A Machine Learning-Aware Data Re-partitioning Framework for Spatial Datasets. *38th International Conference on Data Engineering (ICDE)*, 2022

- **Kanchan Chowdhury**, Mohamed Sarwat; GeoTorch: A Spatiotemporal Deep Learning Framework. *30th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '22)*, 2022

- Vamsi Meduri, **Kanchan Chowdhury**, Mohamed Sarwat; Evaluation of Machine Learning Algorithms in Predicting the Next SQL Query From the Future. *ACM Transactions on Database Systems (TODS)*, 2021

- Jia Yu, **Kanchan Chowdhury**, Mohamed Sarwat; Tabula in Action: A Sampling Middleware for Interactive Geospatial Visualization dashboards. *46th International Conference on Very Large Databases (VLDB)*, 2020.

- Vamsi Meduri, **Kanchan Chowdhury**, Mohamed Sarwat; Recurrent Neural Networks for Dynamic User Intent Prediction in Human-Database Interaction. *22nd International Conference on EDBT, 2019*

- **Kanchan Chowdhury**, Lamia Alam, Shyla Sarmin, Safayet Arefin, Mohammed Moshiul Hoque; A Fuzzy Features Based Online Handwritten Bangla Word Recognition Framework. *18th ICCIT*, 2015

MENTORING & TRAINING

- Co-mentoring Ph.D. and MS students as part of my Postdoc responsibilities at Arizona State University.

- Co-mentored an MS student in the Data Systems Lab at Arizona State University.

- Trained more than 150 undergraduate students in Bangladesh on Android Development across two universities.

Additional Services

- **Research Grant Proposal Writer**: Worked with my Postdoc mentor in writing proposals for research grants submitted to NSF under categories such as *Division of Information and Intelligent Systems (IIS)*.

- **Paper Reviewer**: Served as a reviewer for the journal IEEE TKDE, Journal of Advances in Information Technology, and the conferences ICCAD 2023, DAC 2025.

- **External Reviewer**: Reviewed papers as an external reviewer for the following conferences and journals - SIGMOD 2020, 2022, PVLDB 2019-2025, ICDE 2020, SIGSPATIAL 2021, VLDB Journal, and TSAS Journal.

- **Grant Reviewer**: Served as a travel grant reviewer for the Graduate and Professional Student Association (GPSA) at Arizona State University from May 2022 to August 2023.

- **Conference Volunteer**: Volunteered to organize two conferences - SIGSPATIAL 2022 and SIGMOD 2023.

- **Presentations & Talks**: Five conference presentations - SIGMOD 2023, FOSS4GNA 2023, ICDE 2022, SIGSPATIAL 2022, and ICCIT 2015.

- **Open Source Contribution**: Contributed to Apache Sedona, an open-source geospatial cluster computing framework with 2k+ GitHub Stars, by adding support for two new spatial data types.

- **Membership**: I hold a membership in the professional organizations such as ACM, IEEE, and SIGMOD

Participation and Awards

- Recipient of ACM SIGMOD 2023 student travel award to attend the conference and present a paper.

- Recipient of ACM SIGSPATIAL 2022 travel award to attend the conference and present a paper.

- Received ASU Graduate and Professional Student Association (GPSA) travel grant twice - 2022 and 2023

- Recipient of CIDSE Doctoral Fellowship at Arizona State University for the academic year 2018-2019.

- 2nd Runner-up at National Hackathon organized by ICT Division of Bangladesh in 2014. The challenge of the hackathon was to design a project-based solution to solve a national problem of the country.

- 2nd Runner-up at National Mobile Application Code Hub organized by BUET, Bangladesh in 2014.

- Recipient of Honors award from my undergraduate university for maintaining academic excellence.

- Recepient of merit scholarship in all four years of my undergraduate education.

- Recepient of the government merit scholarship in Secondary School Certificate examination, Higher Secondary School Certificate examination, 8th-grade public examination, and 5th-grade public examination.

Technical Skills

- **Programming**: Python, Java, C, C++, Scala, SQL, and HTML

- **Databases**: PostgreSQL, SparkSQL, and MySQL

- **Machine Learning**: PyTorch, Scikit-learn, Keras, TensorFlow, ML & DL Models, and ML Statistics

- **Data Analytics**: Apache Spark, Apache Sedona, PySpark, GeoPandas, Pandas, Matplotlib, and Plotly

- **Teaching**: Course Curriculum and Course Material Development, Managing Courses on Canvas

- **Others**: Distributed Computing, SDE Design Principles, CI/CD Pipeline, Docker, Jira, and Rest API

Important CourseWorks

- **AI**: Statistical Machine Learning, Fundamentals of Statistical Learning, Data Intensive Systems for Machine Learning

- **DBMS**: Distributed Database Systems, Big Data Analysis with Scala and Spark, & Database Management Systems

- **Core**: Discrete Mathematics, Software Engineering, Operating Systems, Data Communications, Computer Architecture, Networking, Digital Logic Design, Data Structure, Algorithms, & OOP

Projects

- **GeoTorchAI**: A deep learning and data preprocessing framework for raster imagery and spatiotemporal vector datasets, with **400+ GitHub Stars**. It enables spatiotemporal machine learning practitioners to easily and efficiently implement spatiotemporal deep learning models, besides supporting scalable data preprocessing.

- **NLIDB-Bench**: A benchmark for evaluating state-of-the-art approaches of SQL query generation from natural language queries. It proposes a set of evaluation metrics and conducts experiments with four datasets.

- **Named Entity Recognition**: This work tunes various steps of state-of-the-art methods for named entity recognition and performs evaluation with two popular datasets: CoNLL-2003 and OntoNotes-5.0.

- **Hotspot Analysis on Apache Sedona**: This work performs spatial queries and range joins between two spatial datasets and calculates the Getis-Ord statistic of NYC Taxi Trip dataset to perform hot-cell analysis.

- **Data Fragmentation**: It performs data partitioning operations on top of an open-source relational database management system (PostgreSQL). It loads data into a database using Range and round-robin partitioning.

- **Climate Change Forecasting**: A data science project to perform data cleaning, feature engineering, and data preprocessing operations on raw temperature data and predict temperature trends with the LSTM model.

- **Fake News Detection**: A data science project to perform data cleaning, feature engineering, and data preprocessing operations on news datasets and classifying fake and real news with a Bidirectional LSTM model.

- **Data Visualization in Python**: It performs statistical data analysis using useful visualization patterns with three Python libraries: Matplotlib, Seaborn, and Plotly Express.

- **Sentiment Analysis with BERT**: A PyTorch project to classify emotions using the Twitter smile dataset. It contains ML Engineering steps, such as data preprocessing, tokenization, encoding, training, and evaluation.

Referees

- **Dr. Jia Zou (jia.zou@asu.edu)**
  Assistant Professor of Computer Science, Arizona State University

- **Dr. Mohamed Sarwat (mo@wherobots.com)**
  Former Associate Professor of Computer Science, Arizona State University
  Founder and CEO, Wherobots Inc

- **Dr. Venkata Vamsikrishna Meduri (vamsi.meduri@ibm.com)**
  Research Scientist, IBM Almaden Research Center