# Gaukhar Nurbek

gaukharanurbek@gmail.com | LinkedIn | GitHub | Portfolio

## Education

**University of Texas Rio Grande Valley**, Edinburg, TX
PhD in Computer Science | Expected May 2028 | GPA: 4.0/4.0

**University of Texas Rio Grande Valley**, Edinburg, TX
MS in Computer Science | May 2024 | GPA: 4.0/4.0

**Kazakh-British Technical University**, Almaty, Kazakhstan
BEng in Information Systems | May 2018 | GPA: 3.5/4.0

## Technical Skills

**Languages:** Python, C++, Go, TypeScript, JavaScript, MATLAB, SQL, Bash
**ML/AI:** PyTorch, TensorFlow, PyTorch Geometric, LangChain, scikit-learn, OpenAI API, Hugging Face
**Tools:** Git, Docker, Jupyter, CUDA, MuJoCo, Weights & Biases, Linux, HPC Systems
**Areas:** Deep Learning, Time Series Analysis, Foundation Models, Self-Supervised Learning, Computer Vision, Signal Processing, Reinforcement Learning, Graph Neural Networks, Multimodal Learning, NLP

## Publications

**An Efficient Self-Supervised Learning Framework for Swarm Robot Trajectory Analysis**
IEEE ICDM 2025 Workshop (Accepted)

**TRACE: Grounding Time Series in Context for Multimodal Embedding and Retrieval**
NeurIPS 2025 (Accepted)

**Search for Core-Collapse Supernovae Signals Using a Multiclass CNN**
Physical Review D 110, 064055, 2024

## Professional Experience

**Software Engineering Intern** | Uber | Sunnyvale, CA | May 2023 – August 2023

- Developed production feature for internal developer productivity tool serving 15,000+ weekly active users using Go backend and TypeScript frontend with RESTful APIs
- Wrote comprehensive unit tests achieving 85% code coverage; participated in code reviews and cross-functional collaboration in agile environment with distributed microservices

**Data Scientist** | Kazdream Technologies | Astana, Kazakhstan | May 2019 – August 2019

- Trained multilingual speech-to-text deep learning model on 200,000+ audio samples using wav2letter, Python, and CUDA; containerized with Docker for deployment

**Data Analyst** | Center for Sustainable Capital Development | Astana, Kazakhstan | August 2018 – February 2019

- Built time series regression model with 90% accuracy deployed as Flask REST API; implemented web crawler processing 1M+ rows using Python, Selenium, Beautiful Soup, and SQL

## Research Experience

**Graduate Research Assistant** | University of Texas Rio Grande Valley | September 2024 – Present

- Co-authored NeurIPS 2025 and IEEE ICDM 2025 papers on self-supervised learning and multimodal time series embedding
- Developed 14 benchmark datasets for out-of-distribution anomaly detection; evaluated 6 foundation models (Chronos, MOMENT, TimeMOE, Moirai, TimerXL, TimeFM) for cross-domain generalization
- Built TypeScript leaderboard interface for time series forecasting benchmark; optimized neural network inference achieving 4x speedup using distributed computing
- Collaborated with LIGO Scientific Collaboration applying ML for gravitational wave and supernova signal detection

**Graduate Research Assistant** | University of Texas Rio Grande Valley | August 2023 – May 2024

- Designed reinforcement learning framework for 3D locomotion control using Temporal Graph Neural Networks with PPO; implemented graph-structured representations in PyTorch Geometric and MuJoCo

**Research Assistant** | University of Texas Rio Grande Valley | August 2019 – July 2022

- Improved gravitational wave signal detection algorithm by 5% using signal processing techniques in MATLAB, C++, and Python on NSF-funded project

- Enhanced classification pipeline using CNNs achieving 80% noise reduction; developed automated preprocessing pipelines on HPC clusters

## Selected Projects

**ML Interview Prep Chatbot** | June 2024

Built RAG-based interview preparation assistant using Python, Django, LangChain, and OpenAI API with vector embeddings for semantic search over ML interview questions

## Teaching & Leadership

**Graduate Teaching Assistant** | UTRGV | Fall 2022 – Spring 2024
Mentored 300+ students in Deep Learning, Reinforcement Learning, and Object-Oriented Programming

**Memberships:** Women in Data, Data Mining Research Lab (UTRGV)