NURBEK TASTAN

+7(747) 187-4690, +971 54 346-9180 ⋄ Abu-Dhabi, UAE
email: nurbek.tastan@mbzuai.ac.ae ⋄ linkedin: nurbek-tastan ⋄ github: tnurbek

PROFILE

Doctoral researcher with 4+ years of experience in machine learning with an understanding of the modern machine learning pipeline and deep mathematical/probabilistic thinking. Expert in modeling complex machine learning, computer vision, deep learning, and optimization algorithms.

EDUCATION

PhD in Machine Learning, Mohamed bin Zayed University of Artificial Intelligence — Expected 2027
Abu-Dhabi, United Arab Emirates
Relevant Coursework: Foundations and Advanced Topics in Machine Learning, Advanced Probabilistic and Statistical Inference, Advanced Topics in Continuous Optimization, Federated Learning, Safe and Robust Computer Vision, Advanced Machine Learning

MSc in Machine Learning, Mohamed bin Zayed University of Artificial Intelligence — 2021 - 2023
Abu-Dhabi, United Arab Emirates
Relevant Coursework: Mathematical Foundations for AI, ML, CV, Advanced ML, Trustworthy AI, Probabilistic and Statistical Inference, Causal Learning, Reinforcement Learning, Optimization

BSc in Systems of Information Security, International Information Technology University — 2017 - 2021
Almaty, Kazakhstan
Relevant Coursework: Programming (Python, C++, Java), ML, Cybersecurity, a lot of Maths
GPA **3.94**/4.00 (Top 1st of 2021 graduating class (out of 900 students))

EXPERIENCE

Graduate & Research Assistant — Jan 2022 - Present
MBZUAI — *Abu-Dhabi, United Arab Emirates*

- Creating mathematical problems and lab materials, guiding students on their research projects
- Teaching concepts of Probability theory, RVs, inference, and evaluating the goodness of estimators
- Teaching ML Basics, Super/Unsupervised learning, Neural Networks, CNNs, RNNs, RL, Graphical models
- Big Data and Parallel Computing, ML in Spark, Data Mining, Link analysis (PageRank) and LSH

Data Scientist — May 2020 - Sep 2021
InCyberService (Healthcare) — *Almaty, Kazakhstan*

- Created a model that can classify highly imbalanced-class data (chest x-ray) accurately (towards Fair AI)
- Achieved 8.5% growth in the recognition rate (14% improvement of f1 score)
- Created explainable models to get reasonable outcomes (saliency map: LRP, Grad-CAM, LIME)
- Developed object detection models to extract meaningful information from paper-based documents
- Created a new pipeline: controlling data processes using Airflow, Apache Spark (PySpark), Hadoop

Python Developer and Teacher — Jan 2018 - Mar 2019
Bolashak School, IITU — *Almaty, Kazakhstan*

- Developed a school management system to control internal processes using Python/Django
- Trained students Python and C++, including OOP, functional programming, and web parsing

SKILLS

Languages	Python, C++, Scala
Technologies	PyTorch, Tensorflow, ML, DL and CV frameworks/libraries, PySpark
Experienced areas	Federated Learning, ML, DL, CV, Privacy-Preserving ML, Explainable AI, Safety and Robustness of AI, Fairness, Anomaly Detection, Causal Learning
Other skills	Big Data, Apache Spark, Airflow, Hadoop, Spark Streaming, Apache Kafka, SQL DBMSs, Postgres, Cassandra, Clickhouse, Docker, Git

RESEARCH PROJECTS

MSc. Thesis: Secure Decentralized Learning Created a confidential and private decentralized cross-silo federated learning algorithm. I utilized concepts of Federated Learning, Knowledge Distillation, and Fully Homomorphic Encryption (FHE) on CV tasks using PyTorch.

Development of healthcare system using blockchain technology Created a healthcare framework that is secure and ensures users' privacy. The whole system consists of a web application and a telegram bot that is built using Python/Django/Tensorflow. And it has an additional feature of diagnosing heart problems using patient data. The system is fully decentralized, meaning that the data chunks are saved as blocks and controlled using hash values.

PROJECTS

Collaborative Learning with Robustness to Poisoning Attacks Implemented several poisoning attacks (label flip, random, noise, backdoor) and came up with a defense mechanism against them. Used anomaly detection on the data being shared between participants and eliminated malicious parties. Implemented using PyTorch.

Registration Plate Number Detection Real-time video processing by extracting frames and detecting vehicle registration plates. Then, fed it into OCR to extract the plate number. Used my own data and Google Dataset. Built an application that opens the barrier and controls cars entering the building. Used Yolo and Tesseract OCR.

Causality Problem: Cyber-Security Attacks on Graph Data Built an application that helps the cyber security community in analyzing attacks in an efficient way. The input to this problem is security vulnerabilities (e.g., buffer overflow, gateway attacks, etc.) represented as DAGs. Used two well-known methods: greedy equivalence search (GES) (score-based: BIC and BDeu) and Peter Clark (PC) (constraint-based) algorithm.

Density Estimation in Continuous Bayesian Network Reasoning under uncertainty. Built models to learn parametric probability density functions (pdf) of a continuous Bayesian network. Used linear Gaussian models with an assumption that all pdfs are Gaussian. It outperformed the current methods by showing smaller TVD values.

Weather Forecasting Implemented LSTM model that predicts combined conditions of the atmosphere in Almaty. Used Google's Earth Engine to get the historical data.

PUBLICATIONS

Redefining Contributions: Shapley-Driven Federated Learning. IJCAI, 2024.

Collaborative Learning of Anomalies with Privacy (CLAP) for Unsupervised Video Anomaly Detection: A New Baseline. CVPR, 2024.

A Coarse-to-Fine Pseudo-Labeling (C2FPL) Framework for Unsupervised Video Anomaly Detection. WACV, 2024.

CaPriDe Learning: Confidential and Private Decentralized Learning based on Encryption-friendly Distillation Loss. CVPR, 2023. N. Tastan, K. Nandakumar.

Valid and Invalid Bitcoin Transactions. ACM ICEMIS, 2020. N. Tastan, S. Amanzholova.

Burglary Detection Framework for House Crime Control. IEEE ICCSA, 2019. N. Tastan, A. Razaque.

HONORS & AWARDS

"Hack The Space" Hackathon, Dubai, United Arab Emirates, "The Best Math Team"	2022
Republican scientific competition among students, Kazakhstan, II place (Silver)	2021
Scientific competition among students in IITU, Kazakhstan, II place	2021
Republican scientific competition among students, Kazakhstan, II place (Silver)	2020
Scientific competition among students in IITU, Kazakhstan, III place	2020
Presidential Scholarship, Ministry of Education, Kazakhstan	2020
Republican mathematics competition among bachelor students, III place (Bronze)	2019
ICPC regional, II place	2019
Mathematics competition, IITU, I place	2019
Mathematics competition, IITU, II place	2018
Award "Altyn Belgi", Ministry of Education, Kazakhstan	2017
National Presidential Olympiad, Ministry of Education, Kazakhstan (I - regional, II - final)	2016
Republican Applied Mathematics scientific competition, III place	2016
Regional Applied Mathematics scientific competition, I place	2015
Mathematics Olympiad, I place	2013