Najeebullah Shah

Jabl e noor scheme, spinny road, Quetta ·Pakistan ·87300 najeebullahshah@gmail.com +923219526702

Education

2018/09–2024/03 Ph.D. in Bioinformatics (Computational Biology)
Tsinghua University Beijing, China

2012/09–2015/09 M.S. in Computer Science
COMSATS Islamabad, Pakistan

2007/08–2011/08 B.S. in Telecom Engineering
NUCES FAST Islamabad, Pakistan

Professional Experience

2022/07–2023/09 Lecturer Computer Science
University of Balochistan Mastung, Pakistan

2017/03–2018/08 Lecturer Computer Science
University of Loralai Loralai, Pakistan

2015/02–2022/05 Senior Android Developer
LeCodeur Rawalpindi, Pakistan

2011/09–2012/10 Software Engineer
Visionary Integrations Lahore, Pakistan

Software & Development Resources

- **Self Sufficient Candidate Selection Method:** ML-based ab-initio algorithm to automatically discover cell lineage developmental story without prior knowledge
- **Exhaustive Search Method:** Modified version of self sufficient candidate selection method
- **clsGAN:** ML-based Customized Generative Adversarial Network model by incorporating the ideas of Conditional GAN with Label Smoothing and Wasserstein GAN with Gradient Penalty for generating synthetic data on limited and complex single cell data
- **Latent Space Interpretation Scheme:** A ML-based framework to acquire two dimensional easy-to-understand interpretation plots that establishes direct link between latent space variables of GAN and important data semantics
- **Medical Medicine App:** ML-based app that uses Natural Language Processing (NLP) library spacy to create NER for medicines in 12GB file data and extract contraindication relations between different medicines and also between medicines and medical conditionals in python. Develop android app for clients to use the medical app using Flask for backend API
- **OCR for Restaurant Receipts:** ML-based app that uses Optical Character Recognition to extract text from restaurant receipt images and then use Natural Language Processing (NLP) library spacy to extract key information. Then utilize Flask for backend api and display the extract information on website
- **SDK development:** Android app that provides third party services to clients for different actions at the launch of the android apps.
- **Fretbay App:** Android app that creates a system for items/luggage transportation with job listings from customers and bidding from transporters. It includes profiles, messaging, evaluations, etc

- **Myboxman App:** Android app with item delivery system for jobs, biddings, real time messaging, profile in mongo db, etc
- **Fretbay Now App:** Android app with urgent item transportation system with listing, biddings, messaging, profiles, etc
- **Friferie App:** Android & iOS app with hotel management system for European countries

IT Skills

Research Experience

2018/09–2024/03 Ph.D. Student , Xuegong Lab, Bioinformatic Division

Tsinghua University Beijing, China

- Worked on different projects with single-cell omics data, spatial transcriptomics and Liver Tumor CT scan dataset

2012/09–2015/09 M.Phil. Student

COMSATS Islamabad, Pakistan

- Worked on Brain Tumor MRI scan dataset

Publications & Preprints

- **N Shah**, J Li, F Li, W Chen, H Gao, S Chen, K Hua, X Zhang An Experiment on Ab Initio Discovery of Biological Knowledge from scRNA-Seq Data Using Machine Learning In *Patterns from Cell Press*, Elsevier, July 2020, pp. 100071.
- **N Shah**, F Li, X Zhang Hidden Knowledge Recovery from GAN-generated Single-cell RNA-seq Data In *bioRxiv*, December 2023
- **N Shah**, Q Meng, Z Ziheng, X Zhang Systematic analysis on the horse-shoe-like effect in PCA plots of scRNA-seq data In *Bioinformatics Advances from Oxford Press*, February 2024 ***[Under Review]***.
- **N Shah**, S Ziauddin, AR Shahid Brain tumor segmentation and classification using cascaded random decision forests. In *2017 ECTI-CON*, IEEE, June 2017, pp. 718-721.

Awards/Honors

- CSC scholarship for Ph.D. (Chinese Scholarship Council)
- National ICT R&D Fund scholarship for BS

Workshop and Posters

- Ab-initio Discovery of Biological Knowledge from scRNA-Seq Data Using Machine Learning In *NeurIPS 2021 AI for Science Workshop*, 2021 **[WORKSHOP/POSTER]**
- Hidden Knowledge Recovery from GAN-generated Single-cell RNAseq Data In *Machine Learning in Computation Biology (MLCB)*, 2022 **[POSTER]**

LANGUAGES

Pashto (Native)
English (Fluent)
Urdu (Native)
Mandarin (Elementary)

INTERESTS

Playing Chess
Playing Basketball
Running
Trekking
Cooking