UMER MAJEED

Data Scientist
AI & Web3 Engineer

umermjd11.github.io

Islamabad, Pakistan

+92 311 1577 484

umermjd11@gmail.com

/in/umermjd11

Kaggle umermjd11

CV umermjd11.github.io/cv

github.com/umermjd11

github.com/umermajeedkhu

scholar.google.com/citations?user=LrsLEJgAAAAJ

Citations: 650+

SUMMARY

An experienced Data Scientist and innovative Web3 Developer with a Ph.D. candidacy in Computer Science & Engineering. Proficient in utilizing Python, R, SQL, and JavaScript for data analysis, machine learning, and blockchain projects. Skilled in data visualization tools like Plotly and Dash, with a solid foundation in Pandas and NumPy for data manipulation. Expertise in Solidity and Node.js for blockchain technologies and DApps, with a proven track record in implementing smart contracts, DAOs, ERC-20, and ERC-721 projects. Demonstrated success in developing predictive models, conducting in-depth exploratory data analysis, and advancing federated learning and blockchain applications. Published researcher with a focus on leveraging data science, AI, and blockchain for innovative solutions. Eager to apply expertise in statistical analysis, machine learning, and Web3 development to drive impactful insights and contribute valuable insights to dynamic and collaborative teams.

SKILLS

PLs & Frameworks:	Python, R, SQL, C++, Julia, Dash, TensorFlow, PyTorch, Keras, Solidity, JavaScript, Node.js, TypeScript, React.js, Next.js, Microsoft Excel, IBM Cognos Analytics, Google Looker Studio.
Libraries & Technologies:	NumPy, pandas, Matplotlib, Plotly, Seaborn, scikit-learn, NLTK, ggplot2, hardhat, brownie, Web3.js, ethers.js, MetaMask, Infura, Alchemy, Ethernal, Chai, Ganache, surya, openzeppelin-solidity, Truffle.
Familar IDEs:	JupyterLab/ Jupyter Notebook, PyCharm, Remix, RStudio, VS Code, Google Colab
Familar OS:	Ubuntu, Windows

SELECTED PUBLICATIONS

International Journals

Umer Majeed et al., "DAO-FL: Enabling Decentralized Input and Output Verification in Federated Learning with Decentralized Autonomous Organizations," TechRxiv. Preprint, Dec 2023. - Developed **DAO Membership Tokens (DAOMTs)** for governance, implementing **mintable and soul-bound tokens** to facilitate decentralized decision-making. Engineered a decentralized framework for **input and output verification** in federated learning, leveraging **DAOs and ERC-721 tokens** to enhance **security and transparency**.

Umer Majeed et al., "FL-Incentivizer: FL-NFT and FL-Tokens for Federated Learning Model Trading and Training," IEEE Access, Jan 2023. - Incentivized learners to submit local models to the federated learning server by implementing a **reward system** using **ERC-20 tokens** for participants. Developed a mechanism to commercialize the federated learning global model by tokenizing it as **ERC-721 based dynamic NFT**.

Umer Majeed et al., "ST-BFL: A Structured Transparency empowered cross-silo Federated Learning on the Blockchain framework," IEEE Access, Nov. 2021. - Developed a **blockchain-based framework** enhancing data privacy in federated learning through **structured transparency** and **homomorphic encryption**. Implemented **smart contracts** and output verification mechanisms to ensure accountability and integrity in collaborative machine learning processes.

Umer Majeed et al., "Blockchain for IoT-based Smart Cities: Recent Advances, Requirements, and Future," Journal of Network and Computer Applications, Vol. 181, pp.1-22, May 2021. - Conducted a comprehensive literature review to formulate blockchain genesis and enhancements in **blockchain technology** in terms of **consensus algorithms** and platforms. Identified applications and challenges for **blockchain-enabled smart cities**.

International Conferences

Umer Majeed et al., "Cross-Silo Model-Based Secure Federated Transfer Learning for Flow-Based Traffic Classification," ICOIN 2021. - Developed a **federated transfer learning** scheme for traffic classification on **time-related statistical features** using DL and **TensorFlow Federated** on multi-organizational datasets, enhancing accuracy and efficiency through knowledge transfer in a cross-silo setting. Ensured data privacy in federated learning by implementing a **secure aggregation** protocol.

Umer Majeed et al., "Cross-Silo Horizontal Federated Learning for Flow-based Time-related-Features Oriented Traffic Classification," APNOMS 2020. - Developed a **horizontal federated learning** model for **traffic classification** on **TensorFlow Federated**, utilizing flow-based **time-related statistical features** to enhance **data privacy** and **security**. Demonstrated the effectiveness of deep learning techniques in traffic classification in **cross-silo** settings.

Umer Majeed et al., "FLchain: Federated Learning via MEC-enabled Blockchain Network," APNOMS 2019. - Introduced **FLchain**, a novel architecture that combines **blockchain** with **Federated Learning** to enhance **data security** and **privacy** through the use of **channel-specific ledgers** and a global **model state trie**. Demonstrated that **FLchain** outperforms traditional Federated Learning methods by providing robust **provenance** and maintaining an **immutable, auditable learning model**.

Domestic Conferences

Umer Majeed et al., "Vanilla Split Learning for Transportation Mode Detection using Diverse Smartphone Sensors," KCC 2021. - Implemented a **split learning** framework for **transportation mode detection** leveraging **smartphone sensors** to enhance **data privacy** and reduce **client-side computation**. Showed that the **split neural network** achieves comparable performance to traditional deep learning models while being more robust against **inference attacks**.

Umer Majeed et al., "Blockchain-assisted Ensemble Federated Learning for Automatic Modulation Classification in Wireless Networks," KCC 2020. - Proposed an **ensemble federated learning** scheme for **automatic modulation**

classification (AMC) using **deep learning** techniques on **decentralized data**. Leveraged a **blockchain network** to enhance model training and demonstrated improved performance of the ensemble model over base federated models in **wireless communication** systems.

SELECTED CERTIFICATIONS AND MOOCS -

IBM Data Science Professional Certificate - Coursera - Audit Completed with Labs -

This certification covers essential **Data Science** skills, including **Data Visualization**, **Data Management**, **Machine Learning**, and **Data Analysis**. It emphasizes hands-on experience with **Python**, **SQL**, and **CRISP-DM**, exploring **Data Pipelines**, **Feature Engineering**, **Big Data**, and **Model Deployment**. Practical projects involve **data collection**, **wrangling**, and **exploratory analysis**, building a solid foundation for a career in data science.

1. What is Data Science? - April 2024
2. Tools for Data Science - April 2024
3. Data Science Methodology - April 2024
4. Python for Data Science, AI & Development - April 2024
5. Python Project for Data Science - April 2024
6. Databases and SQL for Data Science with Python - May 2024
7. Data Analysis with Python - May 2024
8. Data Visualization with Python - June 2024
9. Machine Learning with Python - June 2024
10. Applied Data Science Capstone - August 2024
11. Generative AI: Elevate Your Data Science Career - July 2024
12. Career Guide and Interview Preparation - August 2024

IBM Data Analyst Professional Certificate - Coursera - Audit Completed with Labs -

This certification provides job-ready **Data Analytics** skills, focusing on **data cleaning**, **visualization**, and **dashboards**. It covers tools like **Python**, **Excel**, **SQL**, and libraries such as **Pandas**, **NumPy**, and **scikit-learn**, along with **Jupyter Notebooks**, **Google Looker**, and **Cognos Analytics**. Skills in **EDA**, **predictive modeling**, **generative AI**, and **machine learning** are applied in projects involving dashboard creation and real-world data analysis.

1. Introduction to Data Analytics - Sep. 2024
2. Excel Basics for Data Analysis - Sep. 2024
3. Data Visualization & Dashboards - Excel & Cognos - Sep. 2024
4. Generative AI: Enhance your Data Analytics Career - Sep. 2024
5. Career Guide & Interview Preparation - Oct. 2024

Deep Learning Specialization - Coursera -

This specialization covers key **Deep Learning** concepts like **Neural Networks**, **Back-propagation**, **Regularization**, and **Optimization**, using frameworks such as **TensorFlow**. It includes architectures like **CNNs** and **RNNs**, advanced topics like **GRU**, **LSTM**, **Attention Models**, and **Transformers** for NLP, with a focus on practical implementation and **optimization strategies**.

1. Neural Networks and Deep Learning - Jul. 2021 -
2. Improving Deep Neural Networks - Aug. 2021 -
3. Structuring Machine Learning Projects - Oct. 2021 -
4. Convolutional Neural Networks - Oct. 2021 -
5. Sequence Models - In Progress

AI For Everyone - Andrew Ng - Coursera - - Completed - Dec. 2019 -

This course provides an overview of **AI terminology**, **strategy**, and **workflows** for machine learning and data science. It addresses **ethical considerations** and **societal impacts** of AI, including **bias** and its effects on various sectors.

Ethereum Developer Degree - learnweb3.io - In Progress -

1. ***Freshman Graduate*** - - *Fundamentals of **blockchain**, **Ethereum**, and **Solidity** for building **dApps** and understanding **decentralized systems**.*
2. ***Sophomore Graduate*** - - *Deep understanding of **gas**, **mining**, **PoW**, **PoS**, and **EVM**. Learn to build full **dApps** with **custom contracts**, **NFTs**, **DAOs**, **ICOs**, and **DEXs** using **React** and **Next.js**.*
3. ***Junior*** - In Progress - *Exploring **Layer 2** solutions, **ENS** integration, local **smart contract testing**, **IPFS**, **Ceramic**, **Chainlink VRF**, and **The Graph's Indexer**.*
4. ***Senior*** - In Progress - *Mastering advanced Web3 topics including **Merkle Trees**, **Flash Loans**, **Smart Contract Security**, **MEV**, and **Gas Optimization**.*

Blockchain Specialization - University of Buffalo - Coursera - -

This specialization covers essential **Blockchain** concepts, including **Cryptography** for secure transactions, **Consensus Protocols** like **PoW** and **PoS**, and the development of **Smart Contracts** using **Solidity**. It also focuses on building **Dapps** with frameworks like **Truffle Suite** and **Hyperledger Fabric**, and emphasizes **Blockchain Security** and the broader **Blockchain Ecosystem**.

1. Blockchain Basics - Completed - Dec. 2018 -
2. Smart Contracts - Completed - July 2019 -
3. Decentralized Applications - Completed - Jan. 2020 -
4. Blockchain Platforms - Completed - Feb. 2020 -

Web3 and Blockchain Fundamentals - INSEAD - Coursera - Audit Completed - Feb. 2024 -

This course covers key **Web3** concepts, including **smart contracts**, **digital assets**, and **governance tokens**. It explores **DAOs**, **blockchain design principles**, and the challenges of implementing blockchain in real-world scenarios.

IBM Blockchain Foundation for Developers - IBM - Coursera - Completed - Aug. 2018 - -

This course offers an overview of **business networks** using **Hyperledger Composer** and **Hyperledger Fabric**. It covers key concepts, use cases, asset transfer, **access control**, **network consensus**, and the roles in building and maintaining blockchain business networks.

Cryptography, Private & Secure AI/Data Science Courses - OpenMined -

1. ***Our Privacy Opportunity*** - Completed - Mar. 2021 - Explore structured transparency, **privacy techniques**, and the **privacy-transparency trade-off**.
2. ***Foundations of Private Computation*** - Ongoing - Progress 80% - Implement **federated learning**, **secure multi-party computation**, **homomorphic encryption**, and **differential privacy**.
3. ***Introduction to Remote Data Science*** - Completed - Feb. 2022 - Use **remote execution tools**, deploy **Domain Nodes**, and apply **privacy-preserving techniques** for distributed data science.

Crash Course on Python - Google - Coursera - Completed - March 2020 - -

This course introduces **Python syntax**, emphasizing programming fundamentals and automation for IT roles. It covers key concepts like **automation**, **code reuse**, and **refactoring**, along with error handling and a **problem-solving framework**. Hands-on exercises focus on writing efficient scripts and data manipulation.

DataCamp - - https://www.datacamp.com/portfolio/umermajeed 2017-2018

These certifications cover **Python Basics**, including **Data Types**, **Data Visualization**, and libraries like **Pandas**, **NumPy**, **Seaborn**, and **Matplotlib**. They also include **EDA**, **Statistical Thinking**, **Statistical Analysis**, **Relational Databases**, **SQL**, **SQL JOINs**, **SQL Aggregation**, and tools like **Git/GitHub** and **CLI piping**.

1. Introduction to Python - - 2017
2. Intermediate Python - - 2017
3. Intermediate SQL - - 2017
4. Introduction to Shell - - 2018
5. Functions in Python - - 2017
6. Python Toolbox - - 2017
7. Statistical Thinking (1) - - 2017
8. Statistical Thinking (2) - - 2017
9. Version Control - Git - - 2018
10. Data Types in Python - - 2017
11. Data Visualization - - 2017
12. Data Visualization - Seaborn - - 2018

PROJECTS & PORTFOLIO -

ML Project - - SpaceX Falcon 9 launches - Kaggle Notebook - , Dash App - ** - This project covers key aspects of **machine learning such as **data collection** (via API and web scraping), **data wrangling**, **exploratory data analysis (EDA)**, and the creation of **visualizations** and **interactive dashboards** using **Plotly Dash** and **Folium**. The project also applies **predictive analysis** through classification techniques to forecast launch success rates.

**DL projects - ** - using TensorFlow, keras, PIL, transformers

1. Simple CNNs - Happyface & Digit hand Signs - - Github
2. ResNet - Digit hand Signs - - Kaggle NB -
3. Transfer Learning - MobileNet - - Kaggle NB -
4. Object Detection using yolov2 - - Github NB
5. Image segmentation using Unet - - Kaggle NB -
6. Face recognition using facenet - - Github NB
7. DL Art - Neural Style Transfer - - Kaggle NB -
8. RNN from Scratch - Dinosaur Island - - Kaggle NB -
9. Text generation - LSTM based RNN - - Kaggle NB -
10. Music Generation - LSTM based RNN - - Kaggle NB -
11. Word Embeddings - Similarity & Debiasing - - Github NB -
12. Emojifier: Expressiveness with Emoji - - Github NB -
13. Neural Machine Translation with Attention - - Github NB -
14. Trigger word detection - from voice - - Kaggle NB -
15. Transformer from Scratch - - Github NB -
16. Explore Positional Encodings - Transformer - - Github NB -

**Exploratory Data Analysis (EDA) Projects - ** - using matplotlib, plotly, pandas

1. **Tesla and GameStop Stock/Revenue Data** - Kaggle Notebook - : involves data fetching via `yfinance`, analysis of key metrics, trends, and a summary of **market behavior** and **financial performance**.
2. **Socioeconomic Indicators in Chicago (2008-2012)** - Kaggle Notebook - : using **pairplots**, **heatmaps**, **correlation matrix**, and **descriptive statistics**.

**Dashboard & Visualization Projects - ** - using Google Looker

1. **Sales and Service Analysis Report for SwiftAuto Traders** - Looker Report - : A comprehensive dashboard analyzing **car sales and service performance**, featuring KPIs like **total profit**, **quantity sold**, and visualizations of **sales by model**, **profit by dealer**, **recalls per model**, **customer sentiment**, and trends in **monthly sales** and **profit**.
2. **Products and Sales Analysis Report for Customer Loyalty Program** - Looker Report - : Detailed insights into **product sales** and **customer loyalty**, with data on **total revenue**, **quantity sold**, and visualizations including **line charts**, **bar charts**, **treemaps**, **gender slicers**, and **revenue by geography** through maps and word clouds.

**Web3 Projects - ** - using `React.js`, `Next.js`, `ether.js`, `Web3Modal`, `Hardhat`, `OpenZeppelin`, `Remix IDE`

1. **Whitelist DApp** - , GitHub - , , Sepolia Etherscan - : Allows users to whitelist up to **10 addresses** for NFT presale. Built with **React.js**, **Next.js**, and **ether.js** with **Web3Modal** for wallet integration.
2. **NFT Collection DApp** - , GitHub - , , Sepolia Etherscan - : Mints up to **20 NFTs** for whitelisted addresses during presale, later opens for public minting. Developed with **React.js**, **Next.js**, and **ether.js**.
3. **Basic DApp** - , GitHub - , Sepolia Etherscan - : Sets a user's mood in a smart contract. Uses **ethers.js** for Ethereum blockchain interaction.
4. **ERC20 Based Cryptocurrency** - GitHub - , Sepolia Etherscan - : Implements a fungible token on Ethereum following the **ERC-20 standard** with **Remix IDE** and **MetaMask** for testing.
5. **Basic NFT Contract** - GitHub - , Sepolia Etherscan - : Demonstrates NFT creation on Ethereum with **Hardhat** and **OpenZeppelin Contracts**, following the **ERC721 standard**.

EDUCATION

2017 - Present — **Master & Ph.D. (Combined) in Computer Science & Engineering** — **CGPA 4.11/4.3**
Department of Computer Science & Engineering, Kyung Hee University, Yongin, South Korea

2011 - 2015 — **BS Electrical (Telecommunication) Engineering** — **CGPA 3.83/4.00**
National University of Sciences & Technology (NUST), Islamabad, Pakistan

EXPERIENCE

2015 – 2016 — **PHP developer** — **Artologics, Islamabad, Pakistan**

- Developed robust back-end applications using Core PHP and CodeIgniter framework.
- Implemented jQuery and JavaScript to facilitate smooth communication between the user interface and server-side components via AJAX requests, enhancing the interactivity of web application.
- Employed SQL queries to interface with MySQL databases, ensuring data integrity and reliability while developing robust solutions for efficient data management.

`PHP / SQL / CodeIgniter / jQuery / AJAX / JavaScript / APIs`

BADGES

Founder's Badge - LearnWeb3 Badges
www.opensea.io/assets/matic/0x60f028C82f9f3bF71e0C13fE9e8E7f916b345C00/262556
The founder's badge was airdropped to students who were early adopters of LearnWeb3.

LANGUAGES

English - Proficient (written and verbal), **Urdu** - Native, **Korean** -Beginner (TOPIK Level 2)