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: 600+

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 comprehensive certification program encompasses a wide range of essential skills in **Data Science**, focusing on **Data Visualization**, **Data Management**, **Machine Learning**, and **Data Analysis**. Participants gain hands-on experience with various tools and techniques, including **Python**, **SQL**, and **CRISP-DM methodology**. The coursework covers topics like **Data Pipelines**, **Feature Engineering**, **Data Augmentation**, **Big Data**, and **Model Deployment**. Participants also engage in practical projects, such as the **Applied Data Science Capstone**, where they perform **data collection**, **wrangling**, and exploratory analysis using real-world datasets, including predicting Falcon 9 rocket landings. This certification is ideal for developing a robust foundation in data science and preparing for a successful career in the field.

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 comprehensive certification equips participants with job-ready skills in **Data Analytics**, emphasizing practical experience in **data cleaning**, **data visualization**, and **dashboards**. The program covers essential tools such as **Python**, **Excel**, and **SQL**, with advanced training in **Python libraries** (e.g., **Pandas**, **NumPy**, and **scikit-learn**), **Jupyter Notebooks**, **Google Looker** and **Cognos Analytics**. Participants develop proficiency in **exploratory data analysis**, **predictive modeling**, **generative AI**, and **machine learning**, and complete hands-on projects, including building interactive dashboards and analyzing real-world datasets. The program also offers **interview preparation** and career support to ensure a smooth transition into the field of data analytics.

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 concepts and techniques in **Deep Learning**, including **Neural Networks**, **Back-propagation**, **Hyperparameters**, **Regularization**, **Optimization**, and frameworks like **TensorFlow**. Participants learn to implement various architectures, including **Convolutional Neural Networks (CNNs)** and **Recurrent Neural Networks (RNNs)**. Advanced topics such as **GRU**, **LSTM**, **Attention Models**, and **Transformers** for natural language processing (NLP) are also explored. The coursework emphasizes practical implementation and optimization strategies to achieve high performance in deep learning tasks.

1. Neural Networks and Deep Learning - Completed - Jul. 2021 -
2. Improving Deep Neural Networks - Completed - Aug. 2021 -
3. Structuring Machine Learning Projects - Completed - Oct. 2021 -
4. Convolutional Neural Networks - Completed - 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 provides a comprehensive overview of essential concepts in **Blockchain** technology. Participants delve into the foundations of **Cryptography**, exploring techniques that ensure secure transactions and data integrity. They gain insights into **Consensus Protocols** such as Proof of Work (PoW) and Proof of Stake (PoS), which are crucial for maintaining the integrity of decentralized networks. The program emphasizes the development and deployment of **Smart Contracts** using **Solidity**, focusing on best practices for creating secure and efficient contracts. Additionally, participants learn to build and manage **Decentralized Applications (Dapps)** leveraging frameworks like **Truffle Suite** and platforms such as **Hyperledger Fabric**. The curriculum also covers important topics like **Blockchain Security**, ensuring a solid understanding of vulnerabilities and protective measures, while providing a thorough overview of 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 essential concepts in **Web3**, including the foundational technologies that support decentralized applications, the roles of **smart contracts**, **digital assets**, and **governance tokens** within the ecosystem. Participants explore the implications of **DAOs** (Decentralized Autonomous Organizations) and identify key **blockchain design principles** along with the challenges associated with implementing blockchain technology in real-world scenarios.

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

This course provides a comprehensive overview of **business networks** utilizing blockchain technology, emphasizing **Hyperledger Composer** and **Hyperledger Fabric**. Participants learn about essential concepts, key use cases, and the process of transferring assets within a blockchain network. The course also covers **access control** mechanisms, **network consensus** methods, and the roles and responsibilities of individuals involved in building and maintaining a blockchain business network.

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 offers a comprehensive introduction to **Python syntax**, focusing on programming fundamentals and automation tasks relevant to IT roles. Participants learn about essential concepts such as **Python automation**, **code reuse**, and **refactoring**. The curriculum covers error handling techniques and includes a structured **problem-solving framework** to tackle complex programming challenges. Hands-on exercises enable learners to apply their skills in writing efficient Python scripts and manipulating data effectively.

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

These certifications covers essential skills in **Python Basics**, including **Data Types**, **Data Visualization**, and libraries such as **Pandas**, **NumPy**, **Seaborn**, and **Matplotlib**. It also includes fundamental concepts of **Exploratory Data Analysis (EDA)**, **SQL**, **Statistical Thinking**, and **Statistical Analysis**. The curriculum emphasizes **Relational Databases**, **SQL JOINs**, **SQL Aggregation**, as well as tools like **Git/GitHub** and command line operations including **CLI piping**.

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

PROJECTS & PORTFOLIO -

SpaceX Falcon 9 ML Project - Kaggle Notebook - , Dash App - - This project focuses on **SpaceX Falcon 9 launches**, covering 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 peoject - Simple CNN models - Happyface and Digit hand Signs - - using **TensorFlow** and **Keras** for image classification tasks. Demonstrates practical use of both the **Sequential** and **Functional** APIs for model development.

DL peoject - ResNet - Digit hand Signs - Kaggle Notebook - - development, training, and testing of the **ResNet** Model using **TensorFlow** and **Keras**.

Tesla and GameStop Stock/Revenue Data and Dashboard - Kaggle Notebook - - **exploratory data analysis (EDA)** of **Tesla and GameStop Stock/Revenue Data**, including **data fetching** via 'yfinance', analysis of key metrics and trends, and a summary of **market behavior** and **financial performance**.

Socioeconomic Indicators in Chicago (2008-2012) - Kaggle Notebook - - EDA of **socioeconomic indicators in Chicago (2008-2012)** through **pairplots**, **heatmaps**, **correlation matrix**, and **descriptive statistics**.

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** and **profit by dealer**. It also covers **recalls per model**, **customer sentiment**, and trends in **monthly sales** and **profit**.

Products and Sales Analysis Report for Customer Loyalty Program - Looker Report - - Detailed insights into **product sales** and **customer loyalty** metrics, 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.

Whitelist DApp - , GitHub - , , Sepolia Ether Scan -
This DApp allows users to whitelist up to **10 addresses** for the presale of **NFTs**. It is built using **React.js**, **Next.js**, and **ether.js**, featuring **Web3Modal** integration for seamless connection to users' wallets.

NFT Collection DApp - , GitHub - , , Sepolia Ether Scan -
This DApp mints up to **20 NFTs**, allowing only whitelisted addresses from the above DApp to mint during the presale period. Once the presale ends, it opens up for public minting. Built using **React.js**, **Next.js**, and **ether.js** with **Web3Modal** for wallet connections.

Basic DApp - , GitHub - , Sepolia Ether Scan -
A basic DApp that sets a person's mood in a smart contract. It utilizes **ethers.js** for interacting with the Ethereum blockchain, allowing users to store and retrieve mood data securely.

ERC20 Based Cryptocurrency - GitHub - , Sepolia Ether Scan -
This project involves creating a fungible token adhering to the **ERC-20 standard** as a custom cryptocurrency. Developed using **Remix IDE** and **MetaMask** for deployment and testing.

Basic NFT Contract - GitHub - , Sepolia Ether Scan -
This project focuses on building a basic **NFT (Non-Fungible Token)** contract on the Ethereum network using **Hardhat** and **OpenZeppelin Contracts**, demonstrating the creation and management of NFTs compliant with 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)