

# ANUDEEP KONERU

(240) 438-8675 | anudeepk@umd.edu | LinkedIn | Portfolio Website | Github

## SUMMARY

Data Scientist with a 4.0 GPA in Information Systems at UMD, combining mathematics, machine learning, data analytics, and graph-based AI. Skilled in building scalable pipelines, predictive models, and knowledge graphs using Spark, Neo4j, and HuggingFace. Strong in translating complex data into insights with TA experience and real-world impact in infrastructure and healthcare.

## EDUCATION

### University of Maryland, Robert H. Smith School of Business

Master of Science in Information Systems | *Terrapin Scholar* | GPA 4.0

College Park, Maryland, USA

01/2024 – 05/2025

### Mahindra University

Bachelor of Technology in Electrical and Electronics Engineering | GPA 3.0

Hyderabad, Telangana, India

08/2019 – 06/2023

## SKILLS

<b>Programming &amp; Scripting:</b>	Python, R, SQL, JavaScript, Bash
<b>Machine Learning &amp; AI:</b>	Scikit-learn, PyTorch, TensorFlow, Keras, XGBoost, Transformers, HuggingFace, OpenCV
<b>Data Engineering &amp; Pipelines:</b>	Apache Spark, Kafka, Hadoop, Airflow, ETL Workflows, Streaming Data, Data Cleaning
<b>Graph Analytics &amp; Retrieval:</b>	Neo4j, Cypher, GraphFrames, NetworkX, FAISS, ChromaDB, LangChain, RAG, PyG
<b>Cloud &amp; Databases:</b>	AWS (S3, EC2, Lambda), GCP, BigQuery, Snowflake, Databricks, PostgreSQL, MySQL, MongoDB
<b>Visualization &amp; Tools:</b>	Tableau, Power BI, Seaborn, Matplotlib, Plotly, Dash, Git, Docker, Jupyter, Postman, Agile

## EXPERIENCE

### Teaching Assistant – Market Microstructure

04/2025 – 06/2025

University of Maryland

College Park, Maryland, USA

- Mentored 50+ graduate students on applying ML algorithms to simulate market behavior, model order flow, and analyze liquidity across multi-million-row financial datasets in weekly lab sessions.
- Delivered hands-on tutorials covering regression, classification, and time-series forecasting; improved student performance by 25% through hyperparameter tuning, model evaluation, and real-world finance case studies.

### Teaching Assistant – Data Mining & Predictive Analytics

01/2025 – 04/2025

University of Maryland

College Park, Maryland, USA

- Designed and evaluated 100+ student assignments in Python (Pandas, NumPy, Scikit-Learn) and R, improving model evaluation accuracy by 25% through real-world applications and metrics such as RMSE and precision.
- Built SQL and Jupyter pipelines to evaluate 100+ student submissions; integrated GraphFrames for clustering, led 20+ Neo4j/Cypher labs, and improved project quality by 30% through hands-on training in GCP, AWS SageMaker, and Colab workflows.

### Data Scientist – Infrastructure Analytics

05/2023 – 01/2024

GKC Pvt. Ltd

Hyderabad, Telangana, India

- Designed automated ETL workflows using Apache Airflow and PostgreSQL; improved data pipeline efficiency by 35%, while improving site monitoring through Power BI dashboards and graph analytics with GraphFrames and NetworkX.
- Built forecasting models with Python, SQL, Scikit-Learn; cut budget variance 20% across 10+ projects via insights and Agile teamwork.

### Data Engineering Intern

10/2021 – 03/2022

Cognizance Institute of Mental Health and Neurosciences Pvt. Ltd

Hyderabad, Telangana, India

- Designed ETL workflows in Python and AWS Glue to transform and validate 5,000+ clinical records; supported KPI tracking and anomaly detection through interactive Tableau dashboards consumed by leadership teams.
- Automated workflows using Lambda and S3 data lakes, reducing reporting delays by 40% and enabling real-time quality control.

## PROJECTS

### CourseConnect AI: Academic Chatbot using Neo4j Graphs [GitHub]

02/2025 – 05/2025

- Built an AI platform transforming course catalogs into knowledge graphs using BERT-based NER, triplet extraction, and RAG via LLaMA 3.
- Developed a Streamlit app with PyVis, NetworkX, and an OpenRouter chatbot for real-time academic planning via Web services.

### Friend Recommendation System using Neo4j and PySpark [GitHub]

03/2025 – 04/2025

- Modeled 200-user Neo4j graph with 500+ links; built features and negative samples for link prediction using PySpark and logistic regression.
- Deployed Streamlit dashboard using Pandas, Spark MLlib, and Neo4j Bolt driver for dynamic friend recommendations.

## CERTIFICATIONS

**Neo4j Certified Professional** — Issued April 2025, Credential ID: a734ab9b-8118-4e0c-ab64-46ac4c44b3f90

**Neo4j Graph Data Science Certification** — Issued April 2025, Credential ID: 5d5dc1db-fd1b-4c48-9f58-687bf87f0756

## ACHIEVEMENTS & HONORS

**Phi Kappa Phi Honor Society** – Inducted among the top 7.5% of graduate students at the University of Maryland for academic excellence.

**Co-founder, Stonk Society NFT Project** – Sold 3,333 NFTs in 24 hours (2022), generating 11+ FTM each and growing a collector community.