



Seemab Yamin

📍 Home : Attock, Pakistan

✉ Email: seemab.y1@gmail.com 🌐 Website: <http://seemab-yamin.com/>

👤 Github: <https://github.com/seemab-yamin/>

Gender: Male Date of birth: 09/07/2000 Place of birth: Attock, Pakistan Nationality: Pakistani

ABOUT ME

With a **Bachelor of Science in Computer Science** from **COMSATS University Islamabad**, my career as a **Web Scraping and Automation Engineer** and subsequently a **Data Engineer** has sparked my interest in the real-world applications of **Data and AI**. Through these positions, I have honed my expertise in **Python** for developing secure and scalable data acquisition and processing pipelines.

Currently, as a freelance software developer, I have expertise in developing AI-based custom solutions, which has sparked my interest in research and development on Large Language Models (**LLMs**), Natural Language Processing (**NLP**), and Computer Vision (**CV**). I am looking forward to shifting to a research setup to delve into these areas at a more fundamental level.

EDUCATION AND TRAINING

[01/09/2018 – 15/09/2022]

Bachelor of Science in Computer Science

COMSATS University of Islamabad <https://www.comsats.edu.pk/>

City: Attock | Country: Pakistan | Field(s) of study: Information and Communication Technologies: • *Software and applications development and analysis* | Level in EQF: EQF level 6

WORK EXPERIENCE

Luxevision Consulting LLC

City: Remote | Country: United States

[01/11/2024 – 01/05/2025]

Data Engineer

- Managed and automated data processing ETL pipelines using Apache Airflow (Python) within Agile team environments.
- Monitored pipeline performance to ensure high levels of data integrity and system reliability.
- Integrated Amazon S3 with the Open Metadata platform and orchestrated ETL jobs using Matillion.

Codeaza Technologies

City: Rawalpindi | Country: Pakistan

[01/09/2022 – 06/2024]

Web Scraping and Automation Engineer

- Maintained and deployed robust core Python ETL pipelines utilizing AWS cloud services.
- Executed data remodeling, built and managed data warehouses, and ensured high-quality web scraping and efficient crawling processes.
- Developed and maintained automated bots, supported software deployments, and performed rigorous data validation and cleaning.
- Integrated diverse databases and led agile teams to deliver scalable, efficient data solutions.
- Engineered an LLM-based smart data solution by implementing the ChatGPT API.

6th Sense Technologies

City: Remote | Country: Pakistan

[01/01/2022 – 01/01/2023]

Junior Python Developer

- Designed and built APIs to support scalable applications, ensuring robust and extensible architecture.
- Maintained and optimized web scrapers for consistent and high-volume data extraction.
- Deployed and managed server infrastructure to maintain high availability and performance.
- Developed reliable backend solutions for web applications using the Django framework.

SKILLS

Data Science

Data Science | Large Language Models | Natural Language Processing | Machine Learning | Python Programming

Data Engineering

Apache Airflow | Pandas | Data Warehousing | Scrapy | Data Modeling | ETL | Data Pipeline | Agile Scrum

Backend

MySQL | Flask | FastAPI | SQL | Django

Frontend

HTML | CSS | JavaScript | Bootstrap

PROJECTS

[01/09/2020 – 01/09/2022]

Smart Topic Extractor

Core Function: Automatically identifies and extracts the most frequent and significant words and phrases from text to summarize content and detect main subjects.

Purpose: Efficiently processes large volumes of text data that are impractical to analyze manually due to time, effort, and error constraints.

LANGUAGE SKILLS

Mother tongue(s): Urdu , Panjabi; Punjabi

Other language(s):

English

LISTENING C2 READING C2 WRITING C1

SPOKEN PRODUCTION C1 SPOKEN INTERACTION C1

Chinese

LISTENING A1 READING A1 WRITING A1

SPOKEN PRODUCTION A1 SPOKEN INTERACTION A2

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

PUBLICATIONS

[2025]

Book Title: Smart Infrastructure Management

- Connects a robust theoretical foundation with real-world application efforts spanning various critical assets, including tracks, bridges, and roads.
- Leverages the latest developments in technology and infrastructure management best practices to address current challenges.
- Offers valuable insights into future trends, fostering further research endeavors.
- Acknowledges the pressing need to correlate economics, resilience, and sustainability facets into project decision-making

Authors: Shi Qiu, Qasim Zaheer, Jin Wang, Chengbo Ai, Seemab Yamin | **Publisher:** Elsevier

Link: <https://doi.org/10.1016/C2023-0-52401-2>