Abdullah Arshad

abdullaharshad2409@gmail.com | /in/Abdullah Arshad | https://abdullah2409.github.io/

Education

Lahore University of Management Sciences (LUMS) Sep. 2021 – Jul. 2025
Bachelor of Science in Computer Science *Lahore, Pakistan*

- **Relevant Coursework:** Artificial Intelligence, Computer Networks, Distributed Systems, LLM Systems, Machine Learning, Generative AI, Operating Systems, Software Engineering

Publications

Mental Health Assessment Using Patient Records in Services Dataset
Paper Under Review at British Journal of Psychiatry (Cambridge/RCPsych)
Abdullah Arshad, Shiza Ihtisham, Basmaa Ali, Clifton Chow

Research Experience

Mental Health Assessment Using Patient Records in Zanjabee Dataset Sep. 2024 – Present
Advisors: Basmaa Ali, Clifton Chow, Shiza Ihtsham

- Scaled depression severity modelling to the **Zanjabee** EMR dataset: 10 years of records from a suburban Boston practice covering **4,750** patients and **16,450** encounters.
- Replaced brittle rule-based extraction with an LLM-assisted pipeline using **GPT-5** and **Gemini 2.5 Pro** to surface context-aware symptom descriptions and map them to **PHQ-9** severity categories.
- Improved severity classification on imbalanced data by collaborating with physicians to validate features and training **Random Forest**, **Gradient Boosting**, and **SVM** classifiers with SMOTE and class weighting.

Mental Health Assessment Using Patient Records in Services Dataset Jan. 2025 – Jan. 2026
Advisors: Basmaa Ali, Clifton Chow, Shiza Ihtsham

- Quantified depression prevalence/treatment and somatic vs. cognitive symptom expression in a tertiary-care medicine cohort in urban Pakistan (**N=356**), benchmarking patterns against Global North reports.
- Built Python NLP pipelines to preprocess unstructured clinical notes into an encounter-level **binary symptom vector** feature matrix and generate interpretable visualizations (word clouds, normalized heatmaps).
- Predicted PHQ-9 scores for unscreened patients (**n=59**) by training a **Random Forest** on observed PHQ-9 encounters (**n=297**) within the extracted feature matrix.

Multimodal Prediction of Anxiety/Depression Using LLMs Sep. 2024 – Jul. 2025
Advisors: Basmaa Ali, Agha Ali Raza

- Developed a multimodal Urdu anxiety/depression prediction system on **450** structured clinical interviews collected at Chughtai Lab, Pakistan's largest diagnostic center.
- Fine-tuned **Qwen2-Audio** in PyTorch to capture richer prosodic and spectral cues than MFCC-based baselines.
- Augmented acoustic features with **Google Gemini Speech** transcripts (**8.3%** WER) and spaCy-derived lexical markers to strengthen model robustness.
- Applied **SMOTE** to address class imbalance, yielding a **15%** improvement in F1 over baseline models for early mental health screening.

Work Experience

SensAI – AI Physician Assistant Platform Sep. 2025 – Present
Lead Software Engineer at DariaTech *Lahore, Pakistan*

- Led design and implementation of SensAI, an AI-powered medical scribing platform that converts multilingual doctor–patient conversations into structured notes, problem lists, ICD-10 codes, and order sets.
- Refactored the monolithic transcription flow into 11 parallel Gemini bin extractors with real-time WebSocket streaming, cutting total processing time from **60–75s** to **30–35s** and time-to-first-result from atleast **60s** to **3–5s**.
- Improved reliability by adding exponential-backoff retries and an IndexedDB-based offline audio queue, raising API success rate from **92%** to **99.5%** and eliminating offline data loss.
- Reduced Gemini token usage per request by **40%** via context caching and domain-specific few-shot prompts, while improving extraction accuracy by **15–20%** on complex clinical fields (medications, diagnoses, ICD codes).
- Optimized backend performance with atomic database updates, smart polling with WebSocket fallback, and conditional bin extraction, cutting database round-trips from **15** to **6** queries per visit and lowering redundant polling traffic by **75%**.

Teaching Experience

Teaching Assistant for CS-100: Introduction to Programming Jan. 2023 - Jul. 2024
LUMS

- Conducted tutorials, held office hours, and graded assignments for over 200 students each semester.
- Guided students through fundamental programming concepts and problem-solving techniques in C++.

Head of IT, SPADES Feb. 2023 – Jan. 2024
LUMS

- Led the front-end development team and supervised the design and implementation of the main SPADES website.
- Assigned projects and monitored team progress to ensure timely and high-quality deliverables.
- Conducted workshops on front-end development, focusing on HTML, CSS, and JavaScript.

Private Tutor Jan. 2022 – Present
Self-Employed

- Teaching programming languages (Java, Python, C++) to A-level and university-level students.
- Preparing students for academic success by providing tailored lessons and hands-on coding exercises.
- Mentoring students on best practices in software development and problem-solving techniques.

Development Projects

CS100 Virtual Teaching Assistant | *C++, Generative AI, Google Gemini*

- Developed a virtual teaching assistant tailored for CS-100 students at LUMS to provide coding guidance, evaluate solutions, and offer reinforcement exercises.
- Implemented features such as sanity and sufficiency checks, problem-solving guidance, and quizzes, focusing on C++ and aligning with the course curriculum.

Content Moderation and Toxicity Classification | *BERT, Transformers, Python, Jigsaw Toxic Comment Dataset*

- Built a machine learning pipeline to classify toxic content (e.g., threats, insults, identity hate) using the Jigsaw Toxic Comment Dataset, leveraging models like Logistic Regression, RNNs, and Transformer-based architectures (BERT, DeBERTa).
- Designed preprocessing workflows and implemented Transformers for feature extraction and model training, achieving high accuracy in toxicity detection.

Distributed Key-Value Store | *Golang*

- Implemented a distributed, fault-tolerant key-value store supporting leader elections and log replication with persistence.

KIYA - Succession Planning System | *MERN Stack*

- Developed a web application for automated employee succession planning based on promotion criteria.

Command-line Shell | *C*

- Created a shell in C capable of handling conjugated commands and pipelining.

Awards

- Recognized as a High Achiever for securing 13 A*/A grades in O Levels.
- Awarded a 100% merit-based scholarship for the entire duration of A Levels.

Technical Skills

Programming Languages: Java, Python, C, C++, SQL (PostgreSQL), JavaScript, HTML/CSS, R
Frameworks and Tools: React, Node.js, Flask, FastAPI, WordPress, Material-UI, JUnit
Developer Tools: Git, Docker, TravisCI, Google Cloud Platform (GCP), VS Code, PyCharm, IntelliJ, Eclipse
Libraries and Packages: pandas, NumPy, Matplotlib, SciKit-Learn