

Shaida Muhammad (Male)

✉ shaidasherpao@gmail.com | 📍 Islamabad, Pakistan | 🔗 LinkedIn.com/in/shaidamuhammad

About Me

- I am a dedicated and analytical Machine Learning Engineer with extensive experience in developing and deploying NLP and classification models. My background includes hands-on work with sentiment analysis, text classification, and large language models (LLMs), such as Meta LLAMA-2 and Phi-3.5, all of which are crucial for understanding and analyzing data.
- I possess a strong foundation in data engineering and machine learning, demonstrated through various projects involving text summarization, classification, and conversational AI. My expertise in leveraging transformer-based models allows me to extract meaningful insights from complex data sets.
- I am committed to continuous learning and staying updated with the latest research in machine learning. I thrive in collaborative environments and am eager to contribute to projects that require technical proficiency and innovation.

Work Experience

Love For Data

Rawalpindi, Pakistan

Sr. Machine Learning Engineer

January 2024 – June 2024

- Worked on Large Language Models (LLMs) for text summarization and information extraction, enhancing content comprehension.
- Implemented text classification models using HuggingFace transformers (BERT, RoBERTa) to analyze textual data, including sentiment analysis tasks.

Datacrypt

Remote, UAE

Machine Learning Engineer

October 2022 – December 2023

- Developed intelligent chatbots using the RASA Framework, improving user interaction and feedback collection.
- Designed and maintained data automation pipelines, ensuring efficient data handling and processing.
- Deployed multiple machine learning applications on AWS, Azure, and GCP, ensuring scalability and reliability.

QLU.ai

Islamabad, Pakistan

Machine Learning Engineer

April 2021 – May 2022

- Enhanced company's product with machine learning-based components, focusing on NLP for better user engagement.
- Trained and deployed NLP models for text classification and analysis, including sentiment analysis models.
- Conducted data engineering tasks on textual data to improve model performance.

Education

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

MS IN COMPUTER SCIENCE (18 YEARS)

Sep 2019 – 2023

CGPA: 3.90

Core Courses: Advanced Algorithms Analysis, Mathematical Methods for Computing, Theory of Computation, Operating Systems

Elective Courses: Machine Learning, Deep Learning, Data Mining, Natural Language Processing (NLP)

University of Peshawar

Peshawar, Pakistan

M.SC IN COMPUTER SCIENCE (16 YEARS)

Oct 2015 – Mar 2018

Marks: 80.58%

Courses: Data Structures, Algorithms, Databases, Networking, C++, E-Commerce, Digital Logic Design, Software Engineering, Compiler, Artificial Intelligence, Operating Systems

Bacha Khan University Charsadda

Charsadda, Pakistan

BACHELOR OF COMPUTER SCIENCE (14 YEARS)

Oct 2013 – August 2015

Courses: Computer Science, Mathematics-A, Physics

Skills

Programming Languages and Frameworks: Python, Java, C++, Flask, Django, FastAPI

Machine Learning Libraries: PyTorch, Scikit-learn, NLTK, SciPy, Numpy, Pandas, TensorFlow, Keras, HuggingFace

Other Skills: Git, GitHub, GitLab, SQL, Linux

Projects

Voice-Based AI Assistant for Elderly

DEVELOPED A VOICE-BASED AI ASSISTANT SYSTEM DESIGNED TO SUPPORT ELDERLY INDIVIDUALS BY ANSWERING QUERIES RELATED TO THEIR HEALTH, MEDICINES, AND OTHER RELEVANT TOPICS. THIS SYSTEM UTILIZED THE PHI-3.5 MODEL AND INTEGRATED VARIOUS OPEN-SOURCE TEXT-TO-SPEECH AND SPEECH-TO-TEXT MODELS TO CREATE A COMPREHENSIVE PIPELINE. THE ASSISTANT FUNCTIONS SIMILARLY TO SIRI, PROVIDING VOICE-ACTIVATED RESPONSES TO HELP USERS MANAGE THEIR DAILY NEEDS AND INQUIRIES.

Phi-3.5, Open Source TTS and STT Models

Document Summarization System

DEVELOPED AN ADVANCED DOCUMENT SUMMARIZATION SYSTEM USING META LLAMA-2 INTEGRATED INTO A DJANGO-BASED WEB APP. THIS PROJECT, CONDUCTED IN A SECURE ENVIRONMENT WITH LIMITED RESOURCES, INVOLVED SUMMARIZING DOCUMENTS OF VARYING LENGTHS AND FORMATS—FROM A FEW SENTENCES TO HUNDREDS OF PAGES. LEVERAGING A QUANTIZED META LLAMA-2 13B MODEL AND COLLABORATING CLOSELY WITH DJANGO AND DEVOPS TEAMS, I ENSURED EFFICIENT PROCESSING AND ACCURATE SUMMARIES TAILORED TO SPECIFIC FORMATS AND REQUIREMENTS.

Django, PyTorch, Meta LLAMA-2, Quantized Model

Domain Classification System

DEVELOPED A DOMAIN CLASSIFICATION SYSTEM FOR A COMPLEX DOCUMENT DATASET WITH NEARLY 100 DOMAINS. OVERCAME CHALLENGES OF LIMITED AND NOISY LABELED DATA BY GENERATING SYNTHETIC DATA AND FINE-TUNING ROBERTA BASE MODELS. DESIGNED A HIERARCHICAL SYSTEM WITH 16 PARENT MODELS FOR INITIAL BROAD DOMAIN CLASSIFICATION AND VARIOUS CHILD MODELS FOR SUBDOMAIN CLASSIFICATION, ACHIEVING AROUND 80% ACCURACY. THIS SOLUTION EFFICIENTLY CATEGORIZED DOCUMENTS INTO RELEVANT DOMAINS AND SUBDOMAINS DESPITE DATA LIMITATIONS.

PyTorch, roBERTa (Fine-Tuned), Synthetic Data Generation, Pandas

Sentiment Analysis System

DEVELOPED A SENTIMENT ANALYSIS MODEL TO CLASSIFY PARAGRAPHS WITHIN DOCUMENTS INTO FIVE CATEGORIES: NEGATIVE, EXTREME NEGATIVE, NEUTRAL, POSITIVE, AND EXTREME POSITIVE. THIS MODEL PROVIDED DETAILED SENTIMENT INSIGHTS FOR DOCUMENT CONTENT, ENHANCING THE ABILITY TO UNDERSTAND AND INTERPRET TEXTUAL DATA EFFECTIVELY.

PyTorch, roBERTa (Fine-Tuned), Pandas

Data Processing Pipeline System

DESIGNED AND IMPLEMENTED A DATA PROCESSING PIPELINE SYSTEM FOR THE AZADEA GROUP, INCLUDING AZADEA, DECATHLON, ADIDAS, ETC. THE SYSTEM HANDLED OVER 10GB OF XML FILES DAILY, GENERATING MULTIPLE CSV AND XML FILES. FACED WITH THE CHALLENGE OF LIMITED HARDWARE (8GB RAM ON A WINDOWS LAPTOP), I UTILIZED CREATIVE SOLUTIONS TO SEQUENTIALLY PROCESS DATA, AVOIDING MEMORY OVERLOAD. THE PROJECT INVOLVED USING PYTHON AND LIBRARIES SUCH AS NUMPY, PANDAS, AND LXML TO EFFECTIVELY MANAGE AND PROCESS LARGE DATASETS.

Python, Numpy, Pandas, lxml

Message Health Checker

DEVELOPED "MESSAGE HEALTH CHECKER," A SOPHISTICATED ROBERTA-BASED CLASSIFICATION SYSTEM FOR LINKEDIN MESSAGES. THIS TOOL CATEGORIZES MESSAGES INTO 5 MULTI-LABEL CATEGORIES WITH 5 RANKS EACH, ASSESSING TONE AND EFFECTIVENESS. IT EMPOWERS USERS TO CRAFT HIGHLY PERSONALIZED MESSAGES, BOOSTING ENGAGEMENT BY 10X AND TRACKING CAMPAIGN SUCCESS WITH DETAILED CONVERSION METRICS.

PyTorch, Transformers, RoBERTa, Numpy, Pandas

Message Assistant

DEVELOPED A MULTICLASS MULTILABEL TEXT CLASSIFICATION MODEL FOR JOB ADVERTISEMENT MESSAGES SENT BY RECRUITERS ON LINKEDIN TO JOB CANDIDATES. THE CLASSIFICATION DETERMINES THE QUALITY OF THE MESSAGE AND USES AN AI-BASED TEXT ASSISTANT. THIS PROJECT IS A COMPONENT OF QLU.AI'S PRIMARY PRODUCT.

PyTorch, Scikit-learn, Transformers, BERT, RoBERTa, Fasttext, Numpy

Job Description Splitting into Sections

Python, NLTK

CREATED A HYBRID APPROACH ALGORITHM TO SPLIT JOB DESCRIPTIONS INTO LOGICAL SECTIONS. THIS COMPONENT WORKS AS AN ASSISTANT TO IMPROVE THE PRECISION AND RECALL OF OTHER COMPONENTS OF QLU.AI'S PRIMARY PRODUCT.

Job Description Sections Classification

PyTorch, HuggingFace, Transformers

TRAINED TRANSFORMER-BASED MODELS SUCH AS BERT, ROBERTA, AND XLNET TO CLASSIFY EACH SECTION OF A JOB DESCRIPTION. THIS HELPER COMPONENT INCREASES THE ACCURACY OF OTHER COMPONENTS OF QLU.AI'S PRIMARY PRODUCT, THEREBY ENHANCING CUSTOMER TRUST IN THE PRODUCT.

Text Summarization, Paraphrasing, and Grammar Correction

HuggingFace, PyTorch

RESEARCHED PRE-TRAINED MODELS WITH GOOD ACCURACY FOR PARAPHRASING, TEXT SUMMARIZATION, AND GRAMMAR CORRECTION. THESE MODELS SERVE AS HELPER COMPONENTS FOR QLU.AI'S MESSAGE GENERATION FEATURE.

Conversational AI System (Chatbot)

RASA Framework

DEVELOPED A CONVERSATIONAL AI SYSTEM TO HANDLE COMMUNICATION BETWEEN RECRUITERS AND CANDIDATES ON LINKEDIN.

Pokémon Image Classification

PyTorch, Numpy

ACADEMIC PROJECT: APPLIED FEED-FORWARD NEURAL NETWORK (FFNN) AND CONVOLUTIONAL NEURAL NETWORK (CNN) TO CLASSIFY POKÉMON IMAGES INTO DIFFERENT CATEGORIES.

Names Generation

PyTorch, NLTK

ACADEMIC PROJECT: DEVELOPED A TEXT GENERATION MODEL BASED ON RNN TO GENERATE NAMES.

Generating Meaningful Word Embeddings

Python, Numpy, SciPy

ACADEMIC PROJECT: TRAINED SKIP-GRAM, CBoW, AND SVD MODELS ON A CUSTOM DATASET TO GENERATE MEANINGFUL WORD EMBEDDINGS.

Deployments on Cloud

AWS, Azure, GCP, EC2

DEPLOYED VARIOUS SERVICES ON CLOUD HOSTING PLATFORMS LIKE AWS, GCP, AND AZURE.

CLOUD-BASED DATA PROCESSING PIPELINES: DEVELOPED SCALABLE DATA PIPELINES USING AWS AND AZURE SERVICES. UTILIZED EC2 INSTANCES WITH EXTERNAL FTP, S3 CONNECTED TO AWS LAMBDA FOR DATA PROCESSING, AND AZURE VIRTUAL MACHINES WITH AZURE FUNCTIONS LINKED TO BLOB STORAGE FOR EFFICIENT DATA HANDLING. INTEGRATED GITHUB ACTIONS AS A CI/CD PIPELINE TO AUTOMATE DEPLOYMENT AND STREAMLINE WORKFLOWS ACROSS BOTH CLOUD ENVIRONMENTS.

Research Publications

Conference Papers

- Enhanced Human Activity Recognition using Inertial Sensor Data from Smart Wearables: A Neural Network Approach with Residual Connections — [Published]
- Mukhtasir-Khail-Net: An Ultra Efficient Convolutional Neural Network for Sports Activity Recognition with Wearable Inertial Sensors — [Published]
- Paying Attention to Human Activities: Recognizing Daily Activities with an Attention-based Neural Network and Inertial Sensors — [Published]
- Spak-Net: A Lightweight Convolutional Neural Network for Activity Recognition with Wearable Inertial Sensors — [Published]
- Machine Learning Approaches for Forecasting the Strength of Concrete Incorporating Fly Ash and Marble Slurry Waste — [Accepted]

Journal Papers

- Enhancing Recruiter Outreach: Predicting Job Ad Quality with Advanced Machine Learning and NLP Models — [Published]
- From Steps to Sentiments: Cross-Domain Transfer Learning for Activity-Based Emotion Detection in Wearable IoT Systems — [Published]

Teaching Experience

Python FREE Online Ramzan Course 2025

Remote, Worldwide

INSTRUCTOR – 4-WEEK PYTHON ONLINE COURSE

March 3, 2025 – March 28, 2025

- Delivered a 4-week online Python programming course from beginner to advanced level during the month of Ramzan.
- Covered foundational to advanced Python topics through live sessions, including real-time coding, assignments, and quizzes.
- Used Google Classroom, Google Meet, and Google Chat for content delivery and interaction with students.
- Provided all materials—slides, code examples, assignments, and quizzes—open-source on GitHub.
- GitHub Repository: github.com/ShaidaMuhammad/Python-Free-Ramzan-Course-2025

Government Shaheed Rizwan Sareer Higher Secondary School

Charsadda, Pakistan

VISITING FACULTY – COMPUTER SCIENCE

Sept 2018 – March 2019

- Taught Computer Science to F.Sc students with a focus on both theory and practical programming.
- Conducted coding practice sessions in the computer lab to reinforce programming concepts.
- Assigned and evaluated assignments and quizzes to monitor student progress.
- Managed attendance records and maintained classroom discipline and engagement.

Awards

Jan 2014	Stoori da Pakhtunkhwa (Stars of the Pukhtunkhwa) Scholarship , Secured 15th position among the Top 20 students from Government Colleges in BISE Peshawar.	BISE Peshawar
Jan 2019	Top 1% in GAT General Test , GAT is taken by all graduated students in Pakistan. The GAT Test is a requirement for applying for a Master's Degree (MS) in Pakistan. Scored 83 out of 100 marks with a 99.62 percentile.	National Testing Service

Certifications

Coursera Certifications

- Natural Language Processing Specialization on Coursera [L9XZX72XDF2A]
- Natural Language Processing with Classification and Vector Spaces [HV2BR2GHE7SU]
- Natural Language Processing with Probabilistic Models [9ZKN4QEUBSDC]
- Natural Language Processing with Sequence Models [Z84QTPRB9C6L]
- Natural Language Processing with Attention Models [7SYPZ5D8DPDN]
- Introduction to Large Language Models [LEKT5FXJHJ56]
- Generative AI with Large Language Models [RBXM49D5MZH4]
- Build Basic Generative Adversarial Networks (GANs) [Z7LPD7T6M3A4]
- Introduction to Self-Driving Cars [5AYN5M4FG8LW]
- What is Data Science? [XDT4LQRWLGNB]
- Python Essentials for MLOps [96WYAFV6J2WJ]
- AI For Everyone [46G53XPMEA6S]
- How Google does Machine Learning [G3HM7DJ4XG5Z]
- Artificial Intelligence on Microsoft Azure [2Q55W93P5JPL]
- Introduction to Deep Learning [UF89T3S2QVQG]
- Neural Networks and Deep Learning [LVB52JUQ9HU7]
- Supervised Machine Learning: Regression and Classification [Y7ECGZDYCVWY]
- Mathematics for Machine Learning: Linear Algebra [RYF7JRYNYMJS]