SIC AI Course Report 2024

ORIC CUST Islamabad

Tel: 0304 1116062

Website: https://knowledge.tech/

Tel: 051-111555666

Website: https://cust.edu.pk/oric/

Knowledge Streams

TABLE OF CONTENTS

INTRODUCTION

The Samsung Innovation Campus (SIC) Artificial Intelligence (AI) course, organized by the Office of Research, Innovation, and Commercialization (ORIC) department at Capital University of Science and Technology (CUST), Islamabad, was held on campus in collaboration with Knowledge Streams, Lahore.

Dr. S. Sohail H. Naqvi, Chief Executive Officer of Knowledge Streams, brought his extensive expertise and visionary leadership to this initiative. His unwavering commitment to empowering the youth with relevant skills and knowledge was evident throughout the course. Dr. Naqvi's dedication to creating impactful learning opportunities reflects his vision of bridging academic excellence with real-world applications, making this program a remarkable milestone for AI education.

Dr. S. Sohail H. Naqvi

Komal Masood

Komal Masood, AVP of Learning & Development at Knowledge Streams, played a pivotal role in the success of the course. With over a decade of experience in managing large scale projects across legal, financial, and education domains, Komal’s expertise in training and skills development ensured the program's effective execution. Her dedication to fostering an engaging and enriching learning environment greatly enhanced the participants' experience. Komal’s efforts in coordinating and supporting the course were instrumental in making it a resounding success, leaving a lasting impact on everyone involved.

The collaboration between CUST and Knowledge Streams, under the leadership of such seasoned professionals, ensured a high-quality learning experience. This course aimed to equip participants with cutting edge knowledge and practical skills in AI, fostering innovation and preparing students for future challenges in the technology sector.

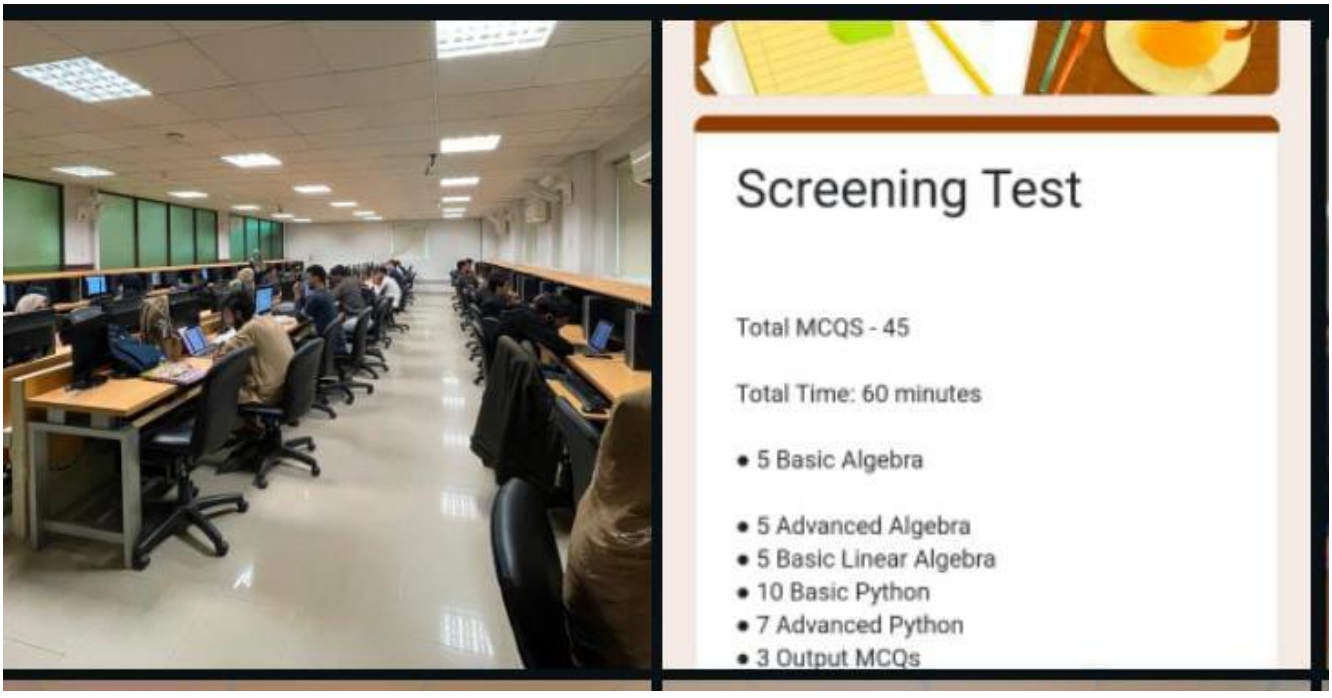
SELECTION CRITERIA

Many students initially applied for the SIC AI course, meeting the following criteria:

1. Age limit between 20 and 30 years (inclusive).
2. A programming background in OOPs and Python or C++.
3. Knowledge of linear algebra and probability theory

More than 400 students applied, out of which the initially shortlisted candidates, based on the above-mentioned criteria, took an entry test conducted on September 27th. After evaluating the test results, the selected students began the course on October 4th.

The course commenced with a total of 42 students. By the conclusion of the program on December 16th, 2024, 34 students successfully completed the course.



COURSE DURATION

The course duration spanned from September 25th to December 16th, 2024, with classes commencing on October 4th, 2024. Sessions were held Monday to Thursday from 9:00 AM to 12:00 PM for lectures, followed by a 30-minute break from 12:00 PM to 12:30 PM, and hands-on lab activities from 12:30 PM to 3:00 PM. On Fridays, lectures were conducted from 9:00 AM to 11:00 AM, followed by lab activities from 11:00 AM to 1:00 PM.

The lecture structure was divided into 70% theoretical instruction and 30% hands-on practice, while the lab work was entirely practical. To earn a certificate of completion, participants were required to maintain at least 90% attendance and achieve a minimum score of 50% in assessments.

COURSE CONTENT

The SIC AI course mainly aimed to equip students with a robust foundation in artificial intelligence and essential data science skills through an engaging blend of theoretical knowledge and practical exercises. Participants explored into the mathematical principles underpinning data science and machine learning, including linear algebra, probability, and statistics. The course emphasized hands-on experience in data preprocessing using Python libraries like NumPy and Pandas to enable efficient execution of machine learning models and data visualization.

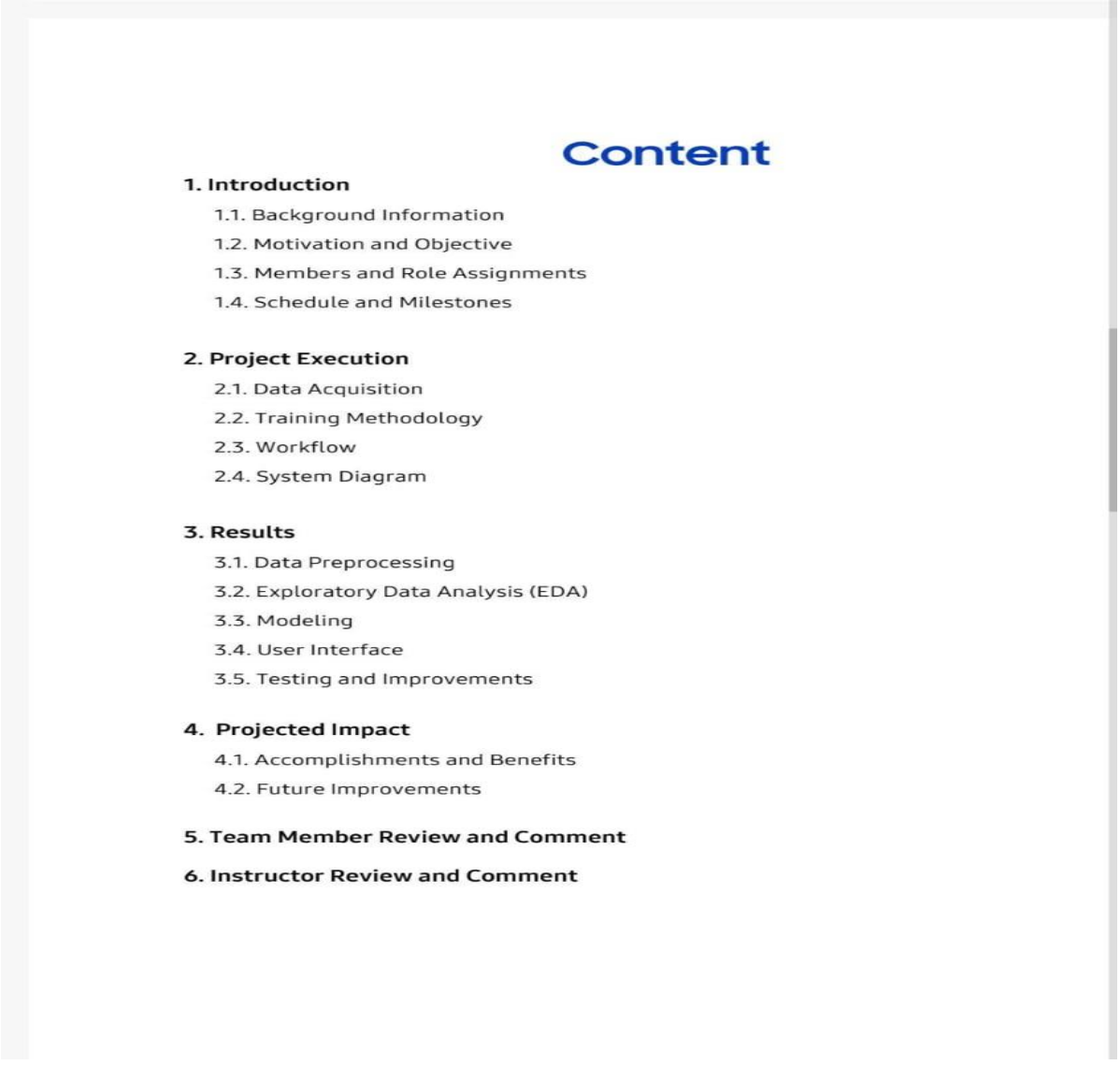
Additionally, students explored supervised and unsupervised learning techniques, gaining the ability to select and apply appropriate algorithms. The curriculum also introduced natural language processing (NLP), teaching students how to process textual data for meaningful insights applicable to real-world business scenarios. The course concluded with deep learning, providing proficiency in building and training neural networks using TensorFlow and Keras, along with exposure to advanced deep learning techniques. Throughout the course students understood datasets clearly how they can collect data from various resources how to work on different kinds of available datasets.

Participants worked thoroughly on Housing Price dataset, Digit MNIST & Fashion MNIST datasets, Species dataset, Titanic dataset and various others found by students themselves from build-in tensor-flow datasets. The course was divided into 60% weightage of class and lab work collectively and 40% for capstone project.

Empowering minds with cutting-edge AI knowledge – unlocking innovation, one skill at a time!

CAPSTONE PROJECT

The capstone project served as a critical evaluation of students' ability to apply their AI knowledge and skills to a comprehensive real-world problem. The projects were assessed based on three key criteria: the quality of the final product, the effectiveness and clarity of the presentation, and the level of collaboration and teamwork exhibited by participants. This initiative encouraged creativity, critical thinking, and the practical application of learned concepts, culminating in innovative and well-executed solutions that demonstrated a strong grasp of the course objectives.



Content

PROJECT SHOWCASE

The following projects highlight the practical applications of AI and data science concepts covered during the course. Each project represents a unique blend of foundational knowledge, technical skills, and creative problem-solving, allowing students to address real-world challenges effectively. These projects emphasize hands-on learning, teamwork, and the ability to turn theoretical knowledge into impactful solutions.

1) **Toxic Comment Classification Using Bi-LSTM and Bi-GRU**

This project addresses the critical issue of toxic content on online platforms by classifying harmful comments into six categories: toxic, severe_toxic, obscene, threat, insult, and identity_hate. Leveraging the Kaggle Toxic Comment Classification dataset, the project employs two advanced recurrent neural network (RNN) architectures—Bidirectional Long Short-Term Memory (Bi-LSTM) and Bidirectional Gated Recurrent Unit (Bi-GRU).

Driven by the increasing prevalence of online toxicity and its adverse impact on social interactions, the project aims to create safer digital environments. The models were trained and tested on the dataset, and their performance was evaluated to determine the more effective architecture for classifying toxic comments. The comparison highlights the efficiency and potential of RNN-based models in tackling harmful content moderation challenges.

2) **Automated Image Classification for Early Diagnosis Diseases**

This project addresses the growing need for rapid and accurate diagnostic methods in response to the rising prevalence of infectious diseases. Traditional diagnostic techniques are often time-consuming and labor-intensive, underscoring the importance of automated image classification systems.

Focusing on diseases such as chickenpox, measles, and monkeypox, the project leverages deep learning techniques to enable early detection and effective disease management. Using convolutional neural networks (CNNs), a proven approach in medical image analysis, the project aims to reduce diagnostic errors, enhance accuracy, and facilitate timely intervention. By automating the classification of medical images, this work highlights the transformative potential of AI in healthcare diagnostics.

3) **AI-Powered Brain Tumor Detection and Classification System**

This project focuses on automating the detection and classification of brain tumors using advanced deep learning techniques to enhance diagnostic accuracy and reduce the time

and effort required by manual methods. Leveraging MRI scans, a non-invasive imaging technique, the system aims to classify brain tumors into four categories: **Glioma Tumor, Meningioma Tumor, Pituitary Tumor, and No Tumor**.

By utilizing transfer learning with the Xception model and a custom-built convolutional neural network (CNN), the project achieves high-accuracy tumor classification. Gemini AI (Gemini-1.5-Flash) is integrated to generate detailed explanations for predictions, providing radiologists with insights into the model's reasoning. Additionally, the system features a user-friendly interface built with Streamlit, allowing real-time tumor classification and saliency map visualization to highlight critical areas in MRI scans influencing the predictions.

The project underscores the critical role of AI in healthcare, offering a robust tool to assist radiologists in early and accurate brain tumor diagnosis, ultimately improving treatment outcomes and patient survival rates.

4) **Urdu Speech Emotion Recognition Using Deep Learning**

This project pioneers emotion recognition for Urdu, addressing the underrepresentation of this language in speech emotion recognition (SER) systems. By focusing on classifying four fundamental emotions—**Angry, Happy, Neutral, and Sad**—from Urdu speech audio, it bridges a critical gap in emotion-aware AI technologies.

Leveraging features like **Mel-Frequency Cepstral Coefficients (MFCCs)** and **Mel Spectrograms**, the project implements and evaluates deep learning architectures, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and hybrid CNN-LSTM models.

Motivated by the increasing demand for emotion-aware applications in areas such as mental health monitoring and customer service, this project contributes to inclusivity in AI while advancing human-computer interaction for Urdu-speaking communities.

5) **Question-Answering System for the Legal Domain Using Advanced NLP Techniques**

This project focuses on developing a question-answering (QA) system tailored to the legal domain, enhancing accessibility and efficiency for legal professionals, businesses, and clients. The system provides quick and accurate responses to queries, retrieves relevant statutes and case laws, explains complex legal principles, and streamlines document analysis.

Leveraging **advanced NLP models like Llama and Retrieval Augmented Generation (RAG)**, the project ensures context-aware answers and supports multilingual and jurisdictional needs. The workflow includes **dataset identification and selection**, **exploratory data analysis**

(EDA), **model implementation and fine-tuning**, **prompt engineering**, and final integration of RAG for real-world applications.

The QA system's functionality extends to legal research, compliance, and client guidance, offering insights tailored to specific jurisdictions or industries. By enabling tasks like drafting legal documents, answering FAQs, and facilitating cross-border legal matters, this project exemplifies how AI can revolutionize access to legal information.

6) **Detection of Persuasive Techniques in Textual Memes Using NLP**

This project aims to address the growing concern of manipulative content in online textual memes by developing a model capable of identifying up to 20 distinct propaganda techniques. With the blend of humor and persuasion inherent in memes, detecting manipulative techniques poses a unique challenge, which recent advancements in natural language processing (NLP) aim to solve.

The project utilizes the **SemEval-2021 dataset**, specifically the **PTC and EM datasets**, with preprocessing to ensure proper encoding. The approach involves evaluating the performance of a **baseline model** and then fine-tuning **pre-trained models** such as **DistilBERT**, **Electra**, **DistilRoBERTa**, **Deberta**, and **Albert**. The model's effectiveness is measured by the **F1 score**, with the goal of surpassing baseline performance. This work contributes to the detection and understanding of persuasive techniques in memes, helping mitigate their potential for manipulation.

7) **Predicting Crop Yields Using Machine Learning for Sustainable Agriculture**

This project focuses on enhancing agricultural productivity and sustainability by leveraging machine learning (ML) for accurate crop yield prediction. Traditionally, yield forecasts were based on historical data and manual observations, which were prone to errors. With advancements in big data analytics and ML, it is now possible to create scalable and accurate models to predict crop yields, improving resource utilization and ensuring food security.

The project addresses critical challenges in agriculture, including climate change, unpredictable weather patterns, soil degradation, and outdated farming practices. By analyzing data from diverse sources such as farming practices, weather conditions, and soil quality, ML algorithms help optimize resource management—such as water, fertilizer, and pesticide use—and enhance yield predictions. The ultimate objective is to contribute to more informed decision-making in farming, promote sustainable practices, and improve productivity to meet the growing global food demand.

8) **Deep Learning for Image Classification Using Convolutional Neural Networks (CNNs)**

This project explores the use of deep learning techniques, specifically Convolutional Neural Networks (CNNs), for image classification tasks. Image classification is a core aspect of artificial intelligence, where the goal is to categorize images into predefined classes. By leveraging modern machine learning methods, the project aims to achieve high accuracy in classifying images, bypassing manual feature engineering and utilizing deep learning models to automatically learn from raw data.

The growing demand for automated solutions in industries like healthcare, retail, and autonomous driving motivates this research. In healthcare, CNNs can assist in analyzing medical images for disease detection, while in retail, they can optimize inventory management through product image classification. The project focuses on developing a robust neural network model capable of high performance across various challenging datasets. Additionally, it aims to demonstrate the model's scalability and adaptability, making it suitable for diverse applications such as defect detection, wildlife monitoring, and document digitization.

9) **Crypto Currency Price Prediction Using Long Short-Term Memory (LSTM) Networks**

This project focuses on crypto price prediction, a critical task in financial markets where accurate forecasting can help investors make informed decisions and manage risks. Stock prices are sequential in nature, making them suitable for analysis using machine learning, particularly deep learning techniques like Long Short-Term Memory (LSTM) networks.

The motivation behind this project is to explore the potential of LSTM networks in time series forecasting for predicting crypto currency prices. The main objective is to develop a model that can accurately predict the next day's closing stock price by analyzing historical price trends and technical indicators.

10) **Pneumonia Detection Using Deep Learning**

Pneumonia remains a significant global health issue, particularly for children and the elderly. Early and accurate diagnosis is critical to effective treatment and reducing mortality. Traditional methods like visual inspection of chest radiographs by radiologists can be slow and prone to human error. Recent advancements in deep learning have shown potential in automating pneumonia detection, improving diagnostic accuracy.

This project is to improve pneumonia detection using a deep learning architecture that enhances diagnostic precision. The objective is to develop a robust model that utilizes transfer learning from well-established deep learning models for binary classification. Four

models–InceptionV3, VGG19, VGG16, and ResNet50–are leveraged to create an effective pneumonia detection system.

CLOSING CERMONY HIGHLIGHTS

The closing ceremony of the three-month-long SIC AI Bootcamp, organized by Samsung Innovation Campus in collaboration with Knowledge Streams, was a momentous occasion that celebrated the hard work and dedication of all participants. Held at CUST, this event marked the successful completion of a comprehensive program designed to equip students with the latest knowledge and skills in artificial intelligence and data science. The ceremony commenced with a beautiful recitation of the Holy Quran, creating a serene and respectful atmosphere for the proceedings.

As the ceremony progressed, Dr. Muhammad Mansoor Ahmed, the Vice Chancellor of CUST, addressed the attendees. In his speech, Dr. Mansoor expressed his immense pride in the students' accomplishments, highlighting the significance of AI and data science in shaping the future. He commended the participants for their commitment to learning and praised the collaborative efforts between CUST, Samsung Innovation Campus, and Knowledge Streams in making this bootcamp a success. His words were an encouragement to the students, urging them to continue their pursuit of knowledge and innovation in the ever-evolving field of technology.

Dr. Sohail Naqvi from Knowledge Streams also shared his insights with the audience. His address focused on the importance of continuous learning and the ever-growing role of artificial intelligence in diverse industries. Dr. Naqvi's message resonated with the students, emphasizing that the skills they had acquired during the bootcamp would open up numerous opportunities for them in both academia and industry.

The ceremony culminated with the distribution of certificates and shields to the students, recognizing their hard work, perseverance, and successful completion of the bootcamp. The Vice Chancellor, Dr. Mansoor Ahmed, personally handed over the certificates, offering his congratulations and best wishes for their future endeavors. This momentous event not only celebrated the students' achievements but also marked the beginning of their journey as skilled professionals in the fields of AI and data science, poised to make significant contributions to the world. The ceremony concluded with an air of pride, accomplishment, and excitement for the future.

Celebrating the journey of innovation, learning, and growth, where every ending is a new beginning in the world of Artificial Intelligence!