HKBK COLLEGE OF ENGINEERING
(Approved by AICTE & Affiliated to VTU)
22/1, Opposite, Manyata Tech Park Rd, Nagawara,
Bengaluru-560045
Email: info@hkbk.edu.in URL: www.hkbk.edu.in

Department of Artificial Intelligence & Machine Learning
And
Electronical Communication Engineering

Presents

Nexora Hackathon

Problem Statements

Domain: Artificial Intelligence and Machine Learning

Problem Statement 1: Disaster Response Resource Allocator

Objective:
Extract structured crisis data from chaotic social media inputs and intelligently allocate limited emergency response resources.

Description:
During natural disasters, social media platforms often become critical channels for distress signals and real-time updates. While AI models can analyze individual posts—such as identifying that someone is trapped or requires urgent medical assistance—they typically lack the capability to independently coordinate response logistics.
In this challenge, participants will leverage AI to classify the type of emergency and extract key entities such as location from unstructured social media data. Under a "no-code" constraint, participants must then design and build a dynamic dashboard that operationalizes these insights.
The system should map identified incidents to a simulated database of available emergency resources (e.g., fire trucks, ambulances, rescue boats). Participants must implement decision logic that prioritizes incidents based on severity and geographic proximity, ensuring efficient and context-aware allocation of resources.

Resources:
- Hugging Face Zero-Shot Classification models (e.g., BART-large-MNLI)
- Named Entity Recognition (NER) models (e.g., RoBERTa-based NER)

Problem Statement 2: Multilingual Civic Services Copilot

Objective:
Build an intelligent, multilingual conversational copilot that assists citizens in navigating complex government services—enabling them to search schemes, check eligibility, and understand application requirements using natural language (text or voice).

Description:
AI systems today are highly capable of transcribing regional speech, translating across languages, and generating conversational responses. However, large language models (LLMs) are susceptible to hallucination—producing incorrect or fabricated policy details—which is unacceptable in high-stakes civic contexts.
In this challenge, participants must use no-code tools to design a robust and secure chatbot or web application interface. The core architecture should follow a **Retrieval-Augmented Generation (RAG)** paradigm:

- Build a structured mock database of government schemes
- Translate and process user queries across multiple Indian languages
- Retrieve relevant, verified information from the database
- Generate grounded responses strictly based on retrieved data

Participants must also implement validation layers—such as rule-based checks or secondary classification steps—to ensure:

- Personally Identifiable Information (PII) is not stored or logged
- The system does not fabricate eligibility criteria, benefits, or entitlements

The solution should support both voice and text-based interaction, ensuring accessibility across diverse user groups.

Resources:

- Hugging Face Inference API
- LLMs (e.g., Mistral-7B-Instruct-v0.2)
- Automatic Speech Recognition (ASR) for Indian languages
- Machine Translation models

Domain: Cybersecurity

Problem Statement 1: Multilingual Phishing & Threat Intelligence Hub

Objective:
Build an automated threat intelligence system that analyzes employee communications (emails/SMS), detects phishing attempts, and triggers appropriate response actions through a simulated security workflow.

Description:
Organizations frequently face phishing attacks delivered through emails and SMS. These attacks may include malicious links, fake invoices, or impersonation attempts, often adapted to local languages and contexts. While AI models can effectively detect phishing patterns and extract suspicious elements, they cannot independently execute security actions such as blocking domains or initiating incident response workflows.
In this challenge, participants must design a no-code workflow that simulates a **Security Orchestration, Automation, and Response (SOAR)** system. The solution should:

- Ingest simulated employee communications (emails or SMS)
- Use AI models to classify messages as phishing or safe
- Extract critical entities such as URLs, email addresses, and financial details
- Assign a risk score to each message

Participants must also implement rule-based logic to simulate real-world security operations.

For example:

- If a message exceeds a defined risk threshold, the system should
    - Update a live Security Operations Center (SOC) dashboard
    - Trigger high-priority alerts (e.g., via Slack, Discord, or Jira)
    - Simulate response actions such as domain blocking or ticket creation

The focus is on combining AI-based detection with structured decision logic to create a reliable and explainable threat response system.

Resources:

- Hugging Face Inference API
    - Text Classification (phishing/spam detection)
    - Token Classification (entity extraction)
    - Document Question Answering (invoice and attachment analysis)

Problem Statement 2: Multilingual Financial Fraud and Scam Detection for India's UPI and SMS Ecosystem

Objective:

Design a multilingual system to detect and prevent financial fraud in SMS, chat, and UPI-related communications, while providing clear risk alerts to users and actionable insights to institutions.

Description:

Users in India frequently receive fraudulent messages through SMS, WhatsApp, and in-app chats. These scams often involve fake UPI requests, KYC updates, loan offers, job opportunities, or investment schemes. Many messages are written in regional languages or mixed scripts (e.g., Hinglish), making them harder to identify as fraudulent—especially for new digital users and senior citizens.

In this challenge, participants are required to build a simple and effective system that can identify such fraudulent messages in real time. The system should:

- Classify messages as **fraudulent** or **legitimate**
- Extract important details such as links, UPI IDs, bank names, and amounts
- Detect suspicious patterns (e.g., fake links, impersonation of banks, unusual requests)
- Handle multiple Indian languages and mixed-language inputs

The solution should also be able to recognize emerging scam patterns and adapt over time. Additionally, participants should provide a user-facing feature (such as a chatbot or API) where users can check suspicious messages and receive a clear risk score along with a short explanation in their preferred language.

Resources:

- Hugging Face text-classification models
- Multilingual and translation models
- Named Entity Recognition (NER) models
- Embedding models for similarity and clustering

Domain: Smart education & edutech

Problem Statement 3: AI-Driven Multilingual Adaptive Learning Companion

Objective:
Develop a multilingual, adaptive learning system that explains STEM concepts in local languages, answers student queries, and tracks individual learning progress for educators.

Description:
Students often struggle to understand complex STEM concepts due to language barriers and lack of personalized support. While AI models can translate content, transcribe speech, and generate simplified explanations, they do not inherently provide structured learning workflows or track student progress.
In this challenge, participants must design a no-code solution that combines AI capabilities with an educational framework. The system should include:

- A student-facing application (mobile or tablet-based)
- Support for both text and voice inputs
- AI-driven translation and explanation of concepts in regional languages
- Real-time responses to student queries

In addition, the workflow must capture and store student interactions and assessment data. This data should power a teacher-facing dashboard that:

- Visualizes individual and class-wide learning progress
- Identifies knowledge gaps
- Allows educators to review and override AI-generated responses when necessary

The focus is on building an interactive, data-driven learning environment that supports both students and teachers.

Resources:

- Hugging Face Inference API
    - Machine Translation models (e.g., NLLB-200)
    - Automatic Speech Recognition (ASR) for Indian languages
    - Question Answering / Text Generation models

Problem Statement 2: Inclusive Multimodal Learning Hub for Students with Disabilities in Higher Education

Objective:
Develop an AI-powered system that converts academic content into accessible, multimodal formats to support students with visual, hearing, and learning disabilities.

Description:
Students with disabilities in higher education often face challenges accessing lectures, PDFs, diagrams, and assessments. Accessibility is frequently handled in an inconsistent or manual manner, placing additional burden on both students and faculty.
In this challenge, participants must design a no-code solution that automatically transforms educational content into accessible formats. The system should:

- Convert scanned documents, images, and diagrams into readable text
- Generate simplified versions of complex academic material
- Provide audio outputs and captions for lecture content
- Support structured outputs that can be used for sign language interpretation

Additionally, the solution should include a faculty-facing interface that:

- Analyzes uploaded content for accessibility gaps (e.g., missing alt text, poor structure)
- Suggests or applies automated improvements
- Enables easy conversion of materials into multiple accessible formats

The goal is to create a scalable and consistent approach to inclusive education.

Resources:

- Hugging Face vision–text (OCR) models
- Text summarization and simplification models
- Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) models
- NLP and vision models for accessibility analysis

Domain: Healthcare & Medtech

Problem Statement 1: Custom Multimodal LLM for Affordable X-ray Analysis

Objective:
Develop a custom multimodal AI system using Hugging Face frameworks to enable low-cost, offline X-ray analysis for rural healthcare settings, supporting early detection and triage of critical conditions.

Description:
Rural healthcare systems in India face a severe shortage of radiologists, resulting in delayed diagnosis of conditions such as tuberculosis (TB), pneumonia, and fractures. Existing AI solutions are often cloud-dependent and expensive, making them unsuitable for Primary Health Centers (PHCs) and Community Health Centers (CHCs).
In this challenge, participants must design a deployable multimodal system that combines medical image analysis with basic patient data to assist frontline healthcare workers. The system should:

- Process chest X-ray images along with patient metadata (e.g., age, symptoms, oxygen levels)
- Detect key abnormalities such as lung opacities, consolidations, and structural irregularities
- Classify conditions including TB (high priority), pneumonia (urgent), cardiomegaly, and fractures
- Generate structured diagnostic outputs with confidence scores and referral recommendations
- Provide reports in regional languages for accessibility

The solution must be optimized for low-resource environments and capable of running offline on standard devices.

Core Requirements:

- Multimodal processing of image and clinical data
- Risk-based triage and prioritization of cases
- Explainable outputs (e.g., highlighted regions of concern)
- Detection of poor-quality X-rays (e.g., blur, low exposure)
- Lightweight deployment suitable for mobile or edge devices

Resources:

- Hugging Face models (vision, text, and generation)
- Public chest X-ray datasets (e.g., CheXpert, MIMIC-CXR, TB datasets)
- Multilingual models for report generation
- Tools for model optimization and quantization

Problem Statement 2: Prescription Digitizer and Conflict Checker

Objective:
Develop a system that extracts text from handwritten or printed medical prescriptions and identifies potential drug conflicts by cross-referencing a medication database.

Description:
Medical prescriptions are often handwritten or poorly formatted, making them difficult to read and prone to interpretation errors. While AI can perform Optical Character Recognition (OCR) to extract text from such documents, the extracted output is typically unstructured and requires further processing.
In this challenge, participants must design a no-code solution that converts prescription images into structured data and analyzes it for safety risks. The system should:

- Extract text from prescription images using OCR
- Identify and parse drug names from unstructured text
- Match extracted drugs against a database of known drug interactions
- Detect and flag potentially harmful combinations

Participants are responsible for designing the logic to structure the extracted data and validate it against a mock or predefined database of drug conflicts.

Resources:

- Hugging Face Inference API
- Dataset available on Kaggle
    - OCR / Document Question Answering models (e.g.,LayoutLM-based models)

Domain: Embedded Systems & IOT

Problem Statement 1: Edge-AI Road Safety and Emissions Monitoring for Indian Smart Cities

Objective:
Develop an edge-based AI system that analyzes traffic conditions in real time to detect safety violations, monitor congestion, and assess emissions-related risks in Indian urban environments.

Description:
Indian cities face significant challenges related to road safety, traffic congestion, and vehicular emissions. These challenges are intensified by heterogeneous traffic conditions, including mixed vehicle types, high density, and frequent rule violations. Many existing surveillance systems rely on models trained on non-Indian datasets, resulting in poor performance in local conditions.
In this challenge, participants must design a system that uses India-specific computer vision models and edge computing to provide accurate, real-time traffic insights**.**

The system should:

- Detect and classify vehicles in complex traffic environments
- Identify traffic violations (e.g., red-light violations, lane indiscipline, helmet non-compliance)
- Measure traffic density and congestion patterns
- Integrate additional sensor data (e.g., air quality, noise levels)
- Generate a risk index for monitored locations

The solution should be optimized for deployment on edge devices and support integration with existing city infrastructure.

Resources:

- Hugging Face object detection models trained on Indian traffic datasets (e.g., UVH-26)
- Edge deployment platforms (e.g., Jetson devices, smart cameras)
- IoT sensor data (air quality, noise monitoring)
- Time-series or forecasting models for traffic analysis

Problem Statement 2: Forest Fire Early Warning System

Objective:
Develop an intelligent system to detect forest fires at an early stage and provide timely alerts to authorities and nearby communities.

Description:
Forest regions in the Western Ghats, including areas near Theni, are highly vulnerable to fires during dry seasons. These incidents threaten biodiversity, wildlife, and local populations. Existing detection methods are often slow, resulting in delayed response and increased damage.
In this challenge, participants must design a system that combines sensor data, satellite inputs, and AI-based analysis to enable early detection and rapid response.

The system should:

- Detect early signs of fire using IoT sensors (e.g., smoke, temperature)
- Analyze satellite imagery or aerial data for fire indicators
- Use predictive models to assess fire risk and spread
- Generate real-time alerts for forest officials and nearby communities

The focus is on improving detection speed, accuracy, and communication efficiency**.**

Resources:

- IoT sensors (smoke, temperature, humidity)
- Thermal imaging (e.g., drones or cameras)
- Satellite data sources
- AI/ML models for prediction and anomaly detection
- Communication channels (SMS, mobile applications)

Domain: Open Innovation

Description:
This domain invites participants to propose and develop their own problem statements across any sector or discipline. It is designed to encourage creativity, interdisciplinary thinking, and exploration of novel ideas beyond predefined themes.
Participants are free to identify real-world challenges and build innovative solutions using

relevant technologies, frameworks, or methodologies. Submissions should clearly define the problem, outline a feasible solution approach, and demonstrate potential impact.

Focus Areas:

- Open to all domains (e.g., healthcare, education, sustainability, fintech, governance, etc.)
- Encourages originality and unconventional problem-solving
- Supports experimental and high-impact ideas

Expected Outcome:

- A well-defined problem statement
- A functional prototype or solution concept
- Clear articulation of impact, feasibility, and scalability

Note:
If a submitted problem statement is not approved by the evaluator, participants are required to revise or replace it based on the feedback provided and resubmit for approval.

For any enquiries reach out for Organizers at nexora@hkbk.edu.in