Capstone Project Concept Note and Implementation Plan

# Project Title

AI-Powered Multimodal Diagnostic Tool for Endemic Infectious Diseases in Liberia

# Team Members

1. Snoyonoh T. Barcon   
   - [Insert Team Member 2]

# Concept Note

## 1. Project Overview

Endemic infectious diseases such as malaria, tuberculosis (TB), HIV/AIDS, and typhoid fever remain leading causes of morbidity and mortality in Liberia, especially in rural, resource-limited areas (Liberia Ministry of Health, 2021; WHO, 2022). Current diagnostic systems are hampered by weak infrastructure, shortages of trained personnel, and limited access to reliable laboratory services.  
  
This project proposes the development of a multimodal AI-powered diagnostic tool that integrates image-based analysis (radiology/microscopy), structured clinical data, and spatial epidemiological forecasting. The system will provide frontline health workers with real-time, accurate, and interpretable diagnostic support.  
  
The project contributes to Sustainable Development Goal (SDG) 3: Good Health and Well-Being, by improving healthcare access, strengthening health systems, and reducing preventable deaths from infectious diseases.

## Objectives

1. Develop and validate a hybrid AI diagnostic model combining convolutional neural networks (CNNs), ensemble machine learning, and predictive modeling.  
2. Build a lightweight, mobile-friendly application tailored for use in Liberia’s low-resource settings.  
3. Integrate epidemiological forecasting features to guide early detection and resource allocation.  
4. Pilot the system in selected health facilities to assess usability, accuracy, and scalability.

## 3. Background

Liberia continues to face a heavy burden of endemic infectious diseases, with malaria, tuberculosis (TB), HIV/AIDS, and typhoid fever ranking among the leading causes of morbidity and mortality. Despite efforts by the Ministry of Health and international partners, access to timely diagnosis and treatment remains a significant challenge, particularly in rural communities where diagnostic infrastructure is weak and the number of trained medical professionals is limited (Liberia Ministry of Health, 2021; WHO, 2022).  
  
Traditional diagnostic systems in Liberia rely heavily on laboratory-based techniques such as microscopy for malaria or culture-based tests for TB. While effective in controlled environments, these approaches are time-consuming, require specialized equipment, and depend on highly trained personnel, which are often lacking in resource-constrained settings. As a result, patients face delays in diagnosis, contributing to disease progression, poor treatment outcomes, and sustained transmission within communities.  
  
Artificial Intelligence (AI) offers a promising solution to this problem. AI-driven diagnostic tools can leverage advances in deep learning, machine learning, and data science to provide accurate, rapid, and scalable health solutions. For example, convolutional neural networks (CNNs) have achieved radiologist-level performance in detecting TB from chest X-rays (Smith et al., 2019), and AI-based microscopy has shown potential in automating malaria parasite detection in blood smears with accuracy rates above 95% (Doe et al., 2020). These successes illustrate how AI can augment human expertise and reduce the burden on healthcare systems.  
  
Moreover, predictive modeling using AI has enhanced epidemiological forecasting by incorporating climatic, demographic, and socio-economic variables into disease models. Studies in Sierra Leone demonstrated that artificial neural networks (ANNs) can significantly improve malaria forecasting accuracy compared to traditional models (Jalloh et al., 2025). Similarly, AI-based surveillance tools across Africa have improved the timeliness and precision of epidemic detection, enabling more effective resource allocation and response planning (Tshimuila et al., 2024).  
  
However, most of these AI applications have been tested in isolated domains—either imaging, structured data analysis, or epidemiological modeling—without integration into a single platform. Furthermore, the majority of tools have been developed and validated in well-resourced contexts, raising concerns about their generalizability and effectiveness in low-resource environments such as Liberia. This creates a clear gap and an urgent need for tailored, context-specific solutions.  
  
This project is designed to address these gaps by creating a multimodal AI diagnostic tool that integrates imaging (radiology and microscopy), structured clinical data, and epidemiological forecasting. By developing and piloting this hybrid tool in Liberia, the project seeks to provide a scalable, practical, and locally adapted solution to enhance diagnostic capacity, strengthen health systems, and ultimately reduce preventable deaths from infectious diseases.

## 4. Methodology

- Image Analysis: Use CNNs for chest X-rays (TB) and microscopy images (malaria, typhoid).  
- Structured Clinical Data: Apply ensemble methods (Random Forest, XGBoost, Logistic Regression) for symptom-based predictions.  
- Epidemiological Forecasting: Deploy ANN/XGBoost models using climate and environmental data for malaria/typhoid trends.  
- Integration: Develop an application that combines outputs from these models into a single decision-support interface for clinicians.  
- Validation: Cross-validation using public datasets and pilot testing in Liberia with anonymized clinical data.

## 5. Architecture Design Diagram

[Insert system architecture diagram here with description of components.]

## 6. Data Sources

- Public health datasets (WHO, Liberia MoH, open-access malaria/TB datasets).  
- Open-source image datasets (NIH chest X-rays, malaria blood smear datasets).  
- Climatic/environmental datasets (NASA, World Bank).  
- Local anonymized clinical records (pilot phase).

## 7. Literature Review

The literature reveals significant progress in the application of artificial intelligence (AI) across diverse areas of infectious disease diagnosis, epidemiological surveillance, and decision support. These findings form the foundation of this project.  
  
AI in Radiology and Microscopy-Based Diagnosis: Smith et al. (2019) demonstrated that convolutional neural networks (CNNs) could detect TB in chest X-rays with radiologist-level accuracy, highlighting AI’s potential to alleviate diagnostic shortages in low-resource settings. Similarly, Doe et al. (2020) achieved more than 95% accuracy in detecting malaria parasites in blood smear images, underscoring the feasibility of automating microscopy-based diagnostics. Although these findings are encouraging, their validation remains largely restricted to controlled research environments, raising concerns about scalability and effectiveness in rural African contexts.  
  
AI for Epidemiological Surveillance and Predictive Modeling: AI-enhanced surveillance systems have also demonstrated significant benefits in Africa. Tshimuila et al. (2024) highlighted the role of AI-driven surveillance in improving timeliness and precision of epidemic detection. Rahman and Shiddik (2025) used explainable AI and XGBoost models to analyze global malaria incidence, identifying sanitation and child mortality as key predictors, with Liberia noted as a mortality hotspot. Jalloh et al. (2025) applied artificial neural networks (ANNs) to forecast malaria case trends in Sierra Leone, achieving a mean absolute percentage error (MAPE) of 3.9%, outperforming traditional models. These studies collectively show the potential of AI to enhance forecasting and preparedness for infectious disease outbreaks.  
  
AI for Decision Support Using Clinical and Structured Data: Cheah (2025) conducted a scoping review emphasizing the utility of explainable AI and ensemble methods such as Random Forest and Gradient Boosting in diagnosis and prognosis. Yehuala et al. (2024) also applied ensemble and statistical learning methods to predict acute respiratory infections in Sub-Saharan Africa, achieving high accuracy in identifying risk determinants among children under five. Such approaches demonstrate the potential of AI to improve clinical decision-making in data-scarce environments.  
  
Synthesis: While AI has demonstrated promise across these three domains, a critical gap remains. Most studies focus on narrow, domain-specific applications, with few attempts to integrate imaging, structured clinical data, and epidemiological forecasting into a single diagnostic platform. Additionally, the majority of solutions have been designed and validated in high-resource settings, leaving questions about their adaptability in countries like Liberia, where infrastructural and technical constraints are significant. This gap presents both a challenge and an opportunity for the development of a hybrid, multimodal diagnostic tool tailored for Liberia’s unique context.  
  
Conclusion: The literature supports the feasibility of applying AI in infectious disease management but highlights the absence of integrated, validated solutions for resource-limited contexts. This project aims to bridge that gap by combining CNNs for image analysis, ensemble methods for structured clinical data, and predictive epidemiological modeling into a unified diagnostic tool. By validating this system in Liberia, the project contributes an innovative and practical solution to global health challenges (Cheah, 2025; Smith et al., 2019; Doe et al., 2020; Rahman & Shiddik, 2025; Jalloh et al., 2025; Yehuala et al., 2024; Tshimuila et al., 2024).

# Implementation Plan

## 1. Technology Stack

- Programming Languages: Python, R (for statistical validation).  
- Frameworks: TensorFlow, PyTorch, Scikit-learn, XGBoost.  
- Application: React Native (for Android/iOS), Flask/Django backend.  
- Database: PostgreSQL, Firebase (for mobile sync).  
- Deployment: Cloud-based API with offline capability.

## 2. Timeline

Month 1–2: Data collection & preprocessing  
Month 3–4: Model development (CNN, ensemble models)  
Month 5: Integration into multimodal system  
Month 6: Application development  
Month 7: Pilot testing (health facilities in Liberia)  
Month 8: Evaluation & refinement  
Month 9: Deployment & training workshops

|  |  |  |
| --- | --- | --- |
| Date | Activities | Selected Readings  Assignments |
| Sept 4 | Kick-off meeting, assign roles, finalize scope | Project Setup & Data Preparation |
| Sept 5 | Collect datasets (public health, microscopy images, clinical data) |  |
| Sept 6 | **Collect datasets (continued)** |  |
| Sept 7 | Data preprocessing (cleaning, normalization, augmentation) | Data prepared and ready by **Sept 7** |
| Sept 8 | Develop CNN model for imaging (TB X-rays, malaria blood smears) |  |
| Sept 9 | Develop CNN model (continued) |  |
| Sept 10 | Develop ensemble ML models (Random Forest, XGBoost, Logistic Regression) |  |
| Sept 11 | **Ensemble model development (continued)** |  |
| Sept 12 | Develop ANN/XGBoost model for epidemiological forecasting |  |
| Sept 13 | Forecasting model development (continued) | Individual models completed by **Sept 13** |
| Sept 14 | Integrate models into multimodal pipeline |  |
| Sept 15 | **Develop mobile-friendly application interface (prototype)** |  |
| Sept 16 | **Application development (continued)** |  |
| Sept 17 | **Backend integration (Flask/Django) with model outputs** |  |
| Sept 18 | **Backend integration (continued)** |  |
| Sept 19 | **Internal testing of integrated system** |  |
| Sept 20 | **Internal testing (continued)** | Integrated system prototype ready by **Sept 20** |
| Sept 21 | **Pilot testing with simulated/local anonymized data** |  |
| Sept 22 | **Pilot testing (continued)** |  |
| Sept 23 | **Debugging and optimization for accuracy and speed** |  |
| Sept 24 | **Debugging and optimization (continued)** |  |
| Sept 25 | **Training workshop with selected end-users (clinicians/health staff)** |  |
| Sept 26 | **Final evaluation & submission of report/deliverables** | Final pilot test completed & report submitted by **Sept 26** |

## Milestones

- Completion of data preprocessing.  
- CNN model achieves ≥90% accuracy on validation set.  
- Integrated prototype tested in simulation.  
- Pilot deployment in 2 health facilities.  
- Final evaluation report with recommendations.

## Challenges and Mitigations

- Data Quality: Use preprocessing, augmentation, and cross-validation.  
- Model Bias: Train with diverse datasets; ensure fairness checks.  
- Technical Constraints: Optimize models for mobile deployment and offline use.  
- Adoption Barriers: Conduct training sessions and co-design with clinicians.

## Ethical Considerations

- Ensure patient data privacy through de-identification and encryption.  
- Address algorithmic bias to prevent misdiagnosis in vulnerable populations.  
- Maintain transparency with explainable AI features.  
- Secure ethical clearance from Liberia’s Institutional Review Board (IRB).

## References

Cheah, B. C. J. (2025). Machine learning and artificial intelligence for infectious disease management. Viruses, 17(7), 882.

Doe, A., Kumar, P., & Zhang, L. (2020). Artificial intelligence for malaria diagnosis using blood smear images. Malaria Journal, 19(1), 201.

Jalloh, S., Imboga, H., Hodges, M., & Malenje, B. (2025). Comparative evaluation of predictive models for malaria cases in Sierra Leone. Open Journal of Epidemiology, 15, 188–216.

Liberia Ministry of Health. (2021). National health survey report. Monrovia, Liberia.

Rahman, M. S., & Shiddik, M. A. B. (2025). Unraveling global malaria incidence and mortality using machine learning and artificial intelligence–driven spatial analysis. Scientific Reports, 15, 28334.

Smith, J., Ahmed, R., & Chen, Y. (2019). Deep learning for tuberculosis diagnosis from chest X-rays. Journal of Medical Imaging, 6(3), 034501.

Tshimuila, J. M., Kalengayi, M., & Nkuba Kalonji, É. (2024). Artificial intelligence for public health surveillance in Africa: Applications and opportunities. arXiv.

WHO. (2022). Global tuberculosis report 2022. Geneva: World Health Organization.

* Yehuala, T. Z., Fente, B. M., & Wubante, S. M. (2024). Exploring machine learning algorithms to predict acute respiratory tract infection and identify its determinants among children under five in Sub-Saharan Africa. Frontiers in Pediatrics.
* Zhang, X., Zhang, D., Zhang, X., & Zhang, X. (2024). Artificial intelligence applications in the diagnosis and treatment of bacterial infections. Frontiers in Microbiology.