**Project Name: Benchmarking Commercial AI-SYS3888**

**Project Period: 2022/8 - 2022/12**

**Sponsor(s) or Client:  Professor Simon Poon**

**Organisation: USYD**

**Project Supervisor: Yaoqi Huang**

| **Background:**  *The context for this project* | The ongoing development of AI products for medical diagnosis reveals encouraging results and these products have begun being integrated into the clinical environment. These products can be used to improve workflow, support radiologists and improve patient care. There are multiple certified AI products which diagnose tuberculosis, the performance of each differs in sensitivity, accuracy, effectiveness and other aspects.  The client is seeking more information about the product qXR, particularly its diagnostic test accuracy compared to its competitors and user experience. |
| --- | --- |
| **Aim**  *Purpose of the project.* | The purpose of this project is to evaluate the product qXR using diagnostic test accuracy and user experience. The diagnostic test accuracy of qXR will be benchmarked against its competitors through a systematic review and meta-analysis. User experience studies will be synthesised to identify factors affecting user acceptance. |
| **Objectives (SMART):**  *Goals of the project.* | * Follow the PRISMA protocol correctly for the systematic review to identify all appropriate papers(Focus on certified AI product “qXR” first). * Benchmark qXR against all other competing AI for TB diagnosis. * Perform meta-analysis to compare and benchmark at least 5 different commercial AIs and qXR used for medical imaging in TB diagnoses. * Technical team first needs to look for any evidence as well as any comparison study related to qXR. * UX team first needs to look for any evidence related to qXR and expand to other similar AI products if it’s enough. * Perform meta-analysis to identify factors affecting user acceptance of AIs used for medical imaging in TB diagnoses. * Conclude, with justification, which product or products are best. |
| **Success Criteria:**  *What does success look like for the sponsor and how can it be measured.* | * The systematic review identified all relevant papers. * Perform a well-rounded meta analysis to compare the diagnostic test accuracy of multiple products. * Perform a meta-analysis to review factors affecting user acceptance of AI used for medical diagnosis (specific to TB if possible). * Recommend the best AI product based on the evidence and explain why and include any areas where it could improve. * Successfully emphasise the difference between models of AI. Contrast user experience distinction including perceived usefulness, perceived ease of use and so on. * Well written report and effective presentation. |
| **Deliverables**  *List the outputs that will be produced as part of the project including the final product or service.* | Internal   1. Systematic review.  * Research papers collected using PRISMA.  1. Initial meta-analysis  * Compare DTA of two products * Identify factors affecting user experience  1. Final meta-analysis  * Benchmark of multiple products  1. Discussion and conclusion  * Recommend the best product * Discuss why the product was chosen * Identify any areas the product could improve in   External   1. Progress report and presentation.  * Systematic review * Initial meta-analysis  1. Final report and presentation.  * Systematic review * Final meta-analysis * Discussion and conclusion |
| **Scope**  *The work that needs to be accomplished in order to deliver/complete the project.* | The project is divided into two sub-parts.  The first part evaluates the diagnostic test accuracy of different AI products which diagnose tuberculosis in comparison to qXR and the second part evaluates factors affecting user acceptance of qXR products for diagnosing tuberculosis. A systematic review using PRISMA will be conducted to collect papers for the DTA and user experience meta-analyses. The best product/s will be recommended with justification and identification of any areas of improvement. |
| **Out Scope**  *Work you are not required to deliver as part of your project* | * Perform our own data collection. * Communicate with an expert in this industry. * Deep understanding of how the AI algorithms work. * Program and implement AI models. |
| **Milestones**  *Key checkpoints with the client e.g. client deployment, approval deadlines* | During our weekly client meetings we will update the client about our progress.  Major milestones   1. Week 3: finalise scope and have it approved by client. 2. Week 6: progress report and presentation. 3. Week 12: final report and presentation. |
| **Human Resources:**  *Are there other specialist staff or subject matter experts that will participate?* | It is the team’s understanding that an expert in the industry is not required. The client is an expert in this industry, so the team may ask the client for clarification about the project content.  The team may consult a university librarian to assist with research and how to perform systematic reviews. |
| **Other Resources:**  *Are there other resources to be utilised in the project? Data? Equipment?* | Publicly available research will be compiled using meta-analysis to determine the validity of each AI. |
| **Reporting/ Meeting Frequency:**  *With what regularity will the team meet with the client* *and report to the client.* | Weekly on Wednesdays at 9:30am. The project manager will email an update to the client prior to the meeting. |