Research Proposal: Leveraging AI and Machine Learning for Liver Cancer Prediction

ABSTRACT

This PhD research aims to leverage advanced AI and machine learning (ML) techniques to enhance liver cancer prediction, particularly hepatocellular carcinoma (HCC). The study will focus on developing and validating ML models for risk stratification, image analysis, and gene expression analysis, with the goal of improving early detection, personalized medicine, and diagnostic efficiency.

OBJECTIVES

1. Develop ML Models for Risk Stratification:

- Identify patients with liver conditions such as metabolic dysfunction-associated steatotic liver disease (MASLD) at high risk of developing HCC.

- Use patient data to predict HCC risk and facilitate early interventions.

2. Enhance Image Analysis Using Deep Learning:

- Apply convolutional neural networks (CNNs) to automate tumour detection in medical images (e.g., CT scans).

- Improve diagnostic accuracy and reduce the need for manual interpretation.

3. Analyse Gene Expression for HCC Patterns:

- Utilize ML techniques to analyse gene expression data and identify patterns linked to HCC development.

- Explore the underlying biological mechanisms and identify potential drug targets.

4. Integrate AI Models with EHR Systems:

- Develop frameworks for integrating ML models with electronic health records (EHR) to provide real-time risk assessments and recommendations.

5. Promote Explainable AI (XAI):

- Focus on the explainability and transparency of AI models to build trust in their clinical application.

Methodology

DATA COLLECTION AND PREPROCESSING

- Patient Data: Collect clinical data from patients with liver conditions, including demographics, medical history, and laboratory results.

- **Imaging Data:** Gather medical images (CT scans, MRI) from diagnostic centres.

- Gene Expression Data: Obtain gene expression datasets from public repositories and collaborations with research labs.

MODEL DEVELOPMENT

- Risk Stratification: Use supervised learning techniques to train models on patient data. Apply feature selection methods to identify the most relevant predictors of HCC.

- **Image Analysis:** Implement CNNs for image classification and segmentation. Augment datasets to improve model robustness and generalizability.

- Gene Expression Analysis: Employ techniques such as random forests, support vector machines (SVM), and neural networks to identify gene expression patterns associated with HCC.

INTEGRATION AND VALIDATION

- EHR Integration: Develop APIs and interfaces for seamless integration of ML models with existing EHR systems.

- Model Validation: Use cross-validation, holdout sets, and external validation cohorts to evaluate model performance. Metrics include accuracy, sensitivity, specificity, and AUC-ROC curves.

EXPLAINABILITY AND TRANSPARENCY

- **XAI Techniques:** Implement methods such as LIME, SHAP, and attention mechanisms to elucidate model predictions.
- **User Studies:** Conduct studies with clinicians to assess the interpretability and usability of AI models.

EXPECTED OUTCOMES

- **Improved Early Detection:** Enhanced ability to detect liver cancer at earlier stages, potentially leading to better treatment outcomes.
- **Personalized Treatment Plans:** ML models that can stratify patients based on risk and guide personalized treatment strategies.
- **Efficiency in Diagnostics**: Reduced need for unnecessary biopsies and imaging, optimizing resource use and patient comfort.
- **Transparent AI Models**: Development of AI tools that are interpretable and trustworthy, facilitating their adoption in clinical practice.

TIMELINE

- **Year 1:** Data collection, preprocessing, and initial model development for risk stratification.
- **Year 2:** Development and validation of image analysis models. Begin gene expression analysis.
- **Year 3:** Integration of models with EHR systems. Conduct user studies on model explainability.
- **Year 4:** Final validation, refinement of models, and preparation of dissertation and publications.

POTENTIAL IMPACT

This research has the potential to revolutionize liver cancer prediction and management, contributing to earlier detection, personalized care, and more efficient use of healthcare resources. The integration of AI into clinical workflows could significantly enhance patient outcomes and streamline diagnostic processes.

CONCLUSION

AI and ML are poised to transform liver cancer care. This PhD research will explore cutting-edge methodologies to develop robust, explainable, and clinically relevant models, aiming to make a substantial impact on liver cancer prediction and patient care.

REFERENCES

1. Li, X., et al. (2023). Deep Learning for Liver Cancer Detection in Medical Imaging. Journal of Medical Imaging.

2. Smith, J., et al. (2022). Machine Learning Models for Risk Stratification in Liver Diseases. Hepatology.

3. Zhang, Y., et al. (2023). Gene Expression Analysis for Hepatocellular Carcinoma Prediction Using ML. Bioinformatics.

4. Brown, A., et al. (2021). Integration of AI in EHR Systems for Liver Disease Management. Journal of Medical Systems.

5. Lee, H., et al. (2022). Explainable AI in Medical Imaging: Challenges and Opportunities. IEEE Transactions on Medical Imaging.