

Project Initialization and Planning Phase

Date	15 October 2024
Team ID	SWTID1727274979
Project Title	Deep learning techniques for breast cancer risk prediction
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

This project proposal outlines a solution to address risk prediction on one of the most severe problem globally called as breast cancer . With a clear objective and a concise problem statement, the proposed solution details the problem statement , approach and its impact , key features, future scope , data sources and resource requirements, including hardware as well as software requirements .

Project Overview	
Objective	<p>The primary objective of this project is to develop a deep learning-based model using Convolutional Neural Networks (CNNs) to predict breast cancer risk from Histopathology images .</p> <p>The key objectives are :</p> <ul style="list-style-type: none"> • To collect , train and preprocess a diverse dataset of Breast Histopathology Images . • To design and implement a deep learning architecture using CNNs tailored for breast cancer risk prediction. • To train and validate the model on labeled datasets to predict the likelihood of breast cancer and classify whether they are benign (Non - Cancerous) or Malignant (Cancerous) . • To evaluate the model's performance using appropriate metrics i.e Accuracy . • To demonstrate the practical potential of the CNN model in assisting clinicians in early detection and risk assessment .
Scope	<p>This project basically focus on building a predictive model using deep learning techniques to assess breast cancer risk, utilizing publicly available datasets. It will cover data preprocessing, model training, validation, and performance evaluation with Accuracy .</p>
Problem Statement	

Description	Breast cancer remains one of the most common cancers globally, and early detection is crucial for effective treatment. Traditional methods have limitations in accuracy and scalability , especially for high-risk breast cancer patients. Computer-aided diagnosis systems showed potential for improving diagnostic accuracy , but early detection and prevention can significantly reduce the chances of death . It is important to detect breast cancer as early as possible .
Impact	A reliable deep learning-based prediction model can significantly improve early diagnosis and personalized treatment, reducing mortality rates and healthcare costs associated with advanced breast cancer.
Proposed Solution	
Approach	This project will utilize a Convolutional Neural Network (CNN) architecture for feature extraction from breast histopathology images and patient health records. Techniques like deep learning will be applied to improve model accuracy , and the model will be trained and tested using Python frameworks like TensorFlow and Keras and python Libraries like Pandas , Numpy , Scikit-Learn , Matplotlib , etc.
Key Features	<ul style="list-style-type: none"> • High accuracy and scalability for large image datasets . • Integration of image for improved risk prediction . • Use of deep learning techniques to enhance model efficiency with limited labeled data .

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	NVIDIA GPUs are preferred Cores : Atleast 4 and above
Memory	RAM specifications	16 GB RAM or above
Storage	Disk space for data, models, and logs	Minimum 512 GB SSD and more is preferred
Software		
Frameworks	Python frameworks	Flask Framework
Libraries	Additional libraries	Pandas , Numpy , Matplotlib ,

		Tensorflow , Keras , Seaborn , Scikit-Learn .
Development Environment	IDE, version control	IDE : Visual Studio Code (VS Code) ,Google Colab , Git Virtual Environment : venv
Data		
Data	Source, size, format	Source : Breast Histopathology Images Dataset from Kaggle . Size : 3 GB Format : 278K Files of breast histopathology images divided into 2 categories (0 and 1) .