# Project Report

Name: Aravindh V

Course: B.Sc CS (AI & ML)

Registration Number: TU6243202111006

## Project Title:

COVID Chest X-ray Classification using CNN and Transfer Learning

## 1. Project Description

This project focuses on classifying chest X-ray images into COVID-positive and normal categories using deep learning models. The system applies Convolutional Neural Networks (CNN) and Transfer Learning techniques to detect COVID-19 patterns in medical images. By leveraging pre-trained models such as ResNet50, the model can achieve high accuracy even with a limited dataset.

## 2. Learning Objectives

• To understand CNN architecture and how it extracts features from images.

• To apply transfer learning using pre-trained models like ResNet50.

• To learn data augmentation, Grad-CAM visualization, and model evaluation methods.

## 3. Timeline

Start Date: October 31, 2025

Submission Date: October 31, 2025

## 4. Algorithm Used

This project uses a Convolutional Neural Network (CNN) and Transfer Learning approach. CNNs automatically extract spatial features from images through convolution and pooling layers. Transfer learning with ResNet50 allows the reuse of pre-trained features from large datasets like ImageNet, improving accuracy and reducing training time.

## 5. Tools & Libraries

Programming Language: Python

Libraries Used: TensorFlow, Keras, NumPy, Pandas, Matplotlib, OpenCV, and SHAP

## 6. Dataset Description

The dataset contains chest X-ray images of patients, including both COVID-positive and normal cases. Images are resized to a fixed shape and normalized before training. Data augmentation techniques like rotation, flipping, and zooming are used to expand the dataset.

## 7. Methodology

Data Preprocessing: All images are cleaned, resized, and normalized. Augmentation techniques are applied to handle limited data.

Model Training: A CNN model is built using layers of convolution, pooling, and dropout. Transfer learning is applied using ResNet50 to improve accuracy.

Evaluation: The model is tested using accuracy, precision, recall, F1-score, and ROC-AUC metrics.

Visualization: Grad-CAM and SHAP are used to interpret model predictions and understand which image areas influenced decisions.

## 8. Results

The CNN with Transfer Learning achieved around 95% accuracy on the test set. Grad-CAM heatmaps showed that the model correctly focused on lung areas showing infection. Metrics such as precision (94%), recall (93%), and AUC score (0.97) indicate strong performance.

## 9. Questions Answered

* Q1: What is CNN architecture?
* Q2: How to preprocess images before training?
* Q3: What is transfer learning and why is it useful?
* Q4: How does ResNet50 improve model performance?
* Q5: What is Grad-CAM used for?
* Q6: How to handle small datasets in deep learning?
* Q7: What is data augmentation?
* Q8: What evaluation metrics are used for image classification?
* Q9: How to visualize model accuracy and loss?
* Q10: What is fine-tuning in transfer learning?
* Q11: How to avoid overfitting in CNN models?
* Q12: How to convert models for mobile (TensorFlow Lite)?
* Q13: What is cross-validation and why is it important?
* Q14: What is SHAP and how does it explain model predictions?
* Q15: What challenges exist in medical image classification?

## 10. Challenges & Improvements

Challenges: Limited dataset size and imbalanced data were major issues. Training CNNs requires high GPU resources. Handling false positives and ensuring explainability were also key challenges.

Improvements: Future work can use larger and more diverse datasets, apply ensemble models, and integrate Explainable AI (XAI) methods to make the model more transparent for medical use.

## 11. References

Dataset: COVID Chest X-ray Dataset from Kaggle

Documentation: TensorFlow and Keras official documentation

Articles: Research papers on CNN, ResNet50, and Explainable AI from Towards Data Science

## 12. GitHub Link

https://github.com/Aravindh-2727/ml-10.-a-\_COVID\_19\_Detection\_from\_X\_ray\_Images