# Intel AI for Manufacturing Certificate Course

## Project Report

## 1. Project Overview

### a. Project Title

Cost Forecasting for Source Components (Local VS Import).

### b. Project Description

This project provides a predictive recommendation system to assist manufacturers in deciding whether to source PCB (Printed Circuit Board) components locally or through imports. Using synthetic data, machine learning, and explainable AI, it delivers business insights through a user-friendly Streamlit dashboard.

### c. Timeline

* 1-3 Day: Problem understanding and dataset design.
* 4-7 Day: Data generation and preprocessing.
* 8-12 Day: Model development, tuning, and evaluation.
* 13-15 Day: Dashboard creation and testing.
* 16-18 Day: Documentation and final review.

### d. Benefits

* Informed decision-making for component sourcing.
* Improved cost-efficiency and lead time prediction.
* Business-friendly UI for non-technical users.
* Enhanced transparency via model explainability (SHAP).

### e. Team Members

* + Namra Patel (220160107050)
  + Nilesh Parmar (220160107063)
  + Darshit Panchal (220160107054)
  + Divya Panchal (220160107055)

### f. Risks

* + Dependence on synthetic data; may need retraining on real-world datasets.
  + Generalization risk if sourcing dynamics vary across industries.
  + Model bias due to imbalanced classes in sourcing decisions.

## 2. Objectives

### a. Primary Objective

To build a sourcing recommendation system for PCB components that predicts whether to source locally or import based on cost, lead time, and taxes.

### b. Secondary Objectives

* + Generate realistic synthetic datasets.
  + Implement SHAP-based explainability.
  + Provide business insights through visualization.
  + Design an interactive dashboard for real-time usage.

### c. Measurable Goals

* + Model accuracy ≥ 85%.
  + Interactive dashboard with real-time prediction and SHAP plot.
  + Visual KPIs including sourcing trends, model performance, and distribution.

## 3. Methodology

### a. Approach

Agile development methodology with weekly deliverables and continuous testing.

### b. Phases

1. Requirement Analysis
2. Synthetic Data Generation
3. Model Training & Evaluation
4. Dashboard Development
5. Testing and Deployment

### c. Deliverables

* + synthetic\_sourcing\_data.csv
  + sourcing\_recommendation\_model.pkl
  + SHAP visualizations
  + Streamlit dashboard (app.py)
  + Project documentation

### d. Testing and Quality Assurance

* + Manual and automated testing on edge cases.
  + Evaluation using accuracy, precision, recall, F1-score.
  + Validation through SHAP explainability.

### e. Risk Management

* + Use of class balancing techniques (e.g., SMOTE).
  + Feature scaling and hyperparameter tuning.
  + Designed dashboard to flag uncertain predictions (low probability).

## 4. Technologies Used

### a. Programming Languages

* Python

### b. Development Frameworks

* + Streamlit: For building the interactive dashboard interface.

### c. Database Management Systems

* + Not applicable (synthetic CSV data is used).

### d. Development Tools

* + VS Code (IDE).
  + Git & GitHub (Version Control).

### e. Testing Tools

* Manual testing using different component input scenarios.
* SHAP for model interpretability and validation.

### f. Cloud Services

* Planning to implement it in any AWS service in near Future.

### g. Security

* Local environment only. In production, secure data handling with HTTPS, access control, and user authentication will be necessary.

### h. APIs and Web Services

* No external APIs used. All processing is done using local scripts and trained models.

## 5. Results

### a. Key Metrics

* Model Accuracy: ~92%.
* Precision: ~91%.
* Recall: ~90%.
* F1 Score: ~90%.
* Confusion Matrix visualization included in dashboard.

### b. ROI

* Helps manufacturers make data-driven sourcing decisions, reducing costs and lead time.
* Speeds up procurement strategy with intelligent recommendations.

## 6. Conclusion

### a. Recap the Project

* Developed an AI-powered recommendation system that simulates sourcing data, trains a model, and provides explainable sourcing decisions (Local vs. Import) via a dashboard.

### b. Key Takeaways

* SHAP explainability greatly enhances model transparency for non-technical stakeholders.
* Business insights from synthetic data can guide procurement strategies.
* Streamlit offers a rapid way to create business-facing AI tools.

### c. Future Plans

* Integrate real sourcing datasets from manufacturing partners.
* Add user authentication and data upload features.

### d. Successes and Challenges

* Successes: Achieved high model accuracy, built user-friendly interface, generated useful synthetic data.
* Challenges: Synthetic data lacks real-world noise, feature engineering for imbalanced class distribution.

## 7. Project Specifics

### a. GitHub URL

* [Github Link](https://github.com/Pranjul9947/PCB-Component-Sourcing-Recommendation-System)

### b. Screenshot Drive URL

* [Project resources](https://drive.google.com/drive/folders/1kHvFQhXVkyqrakrMr2QXYsRsGX6JHwdx?usp=sharing)

### c. Dataset URL

* [Drive Link](https://drive.google.com/file/d/1W3e-PoBW7nWf4JKGGmhmmeen5Zo5xMz4/view?usp=drivesdk)