Report

Part One: Regression Model for Metal Lifespan Estimation

Objective:
The goal of Part One is to develop a regression model that estimates the lifespan of a metal based on various parameters provided in the dataset.

Data Description:
- The dataset includes features such as 'partType', 'microstructure', 'coolingRate', 'quenchTime', 'forgeTime', 'smallDefects', 'largeDefects', 'sliverDefects', 'seedLocation', and 'castType'.
- The target variable is 'Lifespan', representing the lifespan of the metal parts.

Approach:
1. Loaded the dataset from the provided CSV file.
2. Performed data exploration to understand the characteristics of the dataset.
3. Identified and handled categorical features using one-hot encoding.
4. Split the dataset into training and testing sets.
5. Implemented a linear regression model for lifespan prediction.
6. Evaluated the model using metrics such as Mean Squared Error (MSE) and R-squared.

Results:
- The linear regression model provides a foundation for estimating metal lifespan based on the given features.
- Evaluation metrics (MSE, R-squared) indicate the model's performance on the test set.

Part Two: Image Classification for Defect Detection

Objective:
In Part Two, the objective is to create a machine learning solution for classifying metal parts as defective or non-defective based on provided scans.

Data Description:
- The dataset includes images of metal parts along with labels indicating whether a part has defects ('Yes' or 'No') and the type of defect.

Approach:
1. Loaded the image dataset along with corresponding labels.
2. Preprocessed the images, including resizing and normalization.
3. Implemented a Convolutional Neural Network (CNN) for image classification.
4. Utilized data augmentation to improve model generalization.
5. Trained the CNN on the labeled image dataset.

6. Evaluated the model using accuracy and other relevant metrics.

Results:

- The CNN model demonstrates effectiveness in classifying metal parts as defective or non-defective.
- Evaluation metrics provide insights into the model's accuracy and performance on unseen data.

This aims to provide the manufacturing company with machine learning solutions to estimate metal lifespan and automate the classification of defective parts.