Team Information

1. Diki Fernandi 20124023
2. Aulia Defitri Wulandari 20224001
3. Kamilah Nada Maisa 20224010
4. Sudarminto Senlau 20224007

Project Overview

Source of your data:

https://pdftolink.online/share/tmp-mbhhzbcu-21b2040a-f4c7-4033-9a43-6419a0fdf6dc

Method : Utilized SEM (Scanning Electron Microscopy) images of nanofibers as input data. Image preprocessing was performed using ImageJ to enhance visual quality. Feature extraction was carried out using the Inception V3 deep learning model, followed by dimensionality reduction with Principal Component Analysis (PCA). Three machine learning models—Linear Regression, Random Forest, and XGBoost—were applied to predict the average fiber diameter. Model performance was evaluated using R^2, RMSE, MAE, MAPE, and processing time.

Machine learning category : supervised learning

Estimated timeline for completing the project :

No.	Method	Project Month				
		1	2	3	4	5
1	Literature Review					
2	SEM image collection					
3	Image adjustment using ImageJ software					
4	Features extraction using inception V3					
5	Using PCA to reduce features					
6	Applying Various Machine Learning Models : Linear Regression, Random Forest, and Gradient Boosting					
7	Performance Evaluation (RMSE, MAE. MAPE, Time, and R^2)					
8	Task : Medium story, YouTube video, Gist code, OSF presentation slides					

Content

1) Problem definition

Nanofiber fabrication methods such as *electrospinning* often produce wide variations in fiber diameter. This inconsistency makes it difficult to control product quality and optimize performance. Conventional measurement methods require significant time and effort — for example, manually measuring 100 points per sample — which is inefficient. Accurate diameter prediction is crucial, as nanofibers are used in filtration, energy storage, biomedical engineering, and sensors.

2) Proposed solution

Our team proposes a machine learning-based model to predict the average diameter of nanofibers using visual information extracted from Scanning Electron Microscopy (SEM) images. By applying feature extraction via *InceptionV3* and dimensionality reduction using PCA, we tested models such as Linear Regression, Random Forest, and Gradient Boosting to predict the fiber diameter. The performance was evaluated using metrics including RMSE, MAE, MAPE, R^2 score, and computational time.

3) Discussion

Early Insights:

- **Linear Regression** outperformed other models with an R^2 of 0.98 and the fastest computation time (1 second).
- A strong linear correlation was observed between the extracted features and the target diameter (average correlation ≈ 0.72).

Challenges:

- More complex models such as Random Forest and XGBoost underperformed compared to Linear Regression, both in accuracy and speed.
- The dataset appeared to follow a highly linear trend, limiting the effectiveness of non-linear models

Decisions Made:

- Linear Regression was selected as the most efficient and effective model.
- Future work should focus on curating or generating datasets with non-linear patterns to better assess the value of more advanced models.

4) Understanding

Our team has demonstrated a solid understanding of:

- **The topic:** The importance of nanofiber diameter in functional applications and the need for automation.
- **The data:** SEM images converted into numerical features via *InceptionV3*, followed by PCA for dimensionality reduction.
- **The method:** Appropriately selected machine learning models, standard performance evaluation metrics, and mindful computational considerations.

Task Assignment

Medium story: Sudarminto Senlau

YouTube video: Diki Fernandi

Gist code: Aulia Defitri Wulandari

OSF presentation slides: Kamilah Nada Maisa