Title:

AI for Sustainable Development: Personalized Learning Support through Student Clustering

SDG 4 – Quality Education

---

1. Problem Statement

Access to quality education is one of the most critical UN Sustainable Development Goals (SDG 4). However, learners do not all progress at the same pace. In most learning environments, a single teaching style is applied to all students, which results in some learners falling behind while others remain unchallenged. There is a need for personalized learning guidance that considers individual performance patterns.

---

2. AI / Machine Learning Approach

To address this challenge, I used an **Unsupervised Machine Learning** technique called **K-Means Clustering** to group students based on their learning behavior and performance attributes (e.g., attendance, assignment scores, test results).
This approach allowed the model to **identify natural learning patterns** without requiring labeled data.

Why K-Means?

- Groups similar learners together.
- Helps educators understand student needs.
- Simple and computationally efficient.

---

3. Dataset Used

A student performance dataset containing:

Feature	Description
Study Hours	Time spent studying per week
Class Participation	Level of engagement during lessons
Assignment Score	Average assignment performance
Exam Score	Final test result

The dataset was cleaned, normalized, and fed into the clustering model.

---

4. Results

The K-Means model grouped learners into **three meaningful clusters**:

Cluster	Learner Profile	Recommended Support Strategy
Cluster 1	Consistent high performers	Provide advanced learning resources
Cluster 2	Average learners showing improvement	Offer mentorship and goal-tracking
Cluster 3	Low engagement & performance	Provide personalized coaching & study plans

This demonstrates how AI can guide targeted educational intervention, improving outcomes without replacing teachers.

---

5. Ethical Considerations

- **Bias Risk:** If data is incomplete or unbalanced, AI may misclassify students.
- **Privacy:** Student data must be securely stored and anonymized.
- **Human Oversight:** AI recommendations should assist educators, not replace professional judgment.

To mitigate these concerns, transparent model outputs and educator review are essential.

---

6. Impact Summary

This project shows how machine learning can enhance personalized learning, helping educators allocate support where it's needed most. By identifying learning patterns, schools can improve student outcomes, reduce dropout risk, and provide more equitable learning opportunities, directly supporting **SDG 4: Quality Education for All**.

➢ **Q1**.TensorFlow and PyTorch are both deep learning frameworks, but they differ in how they execute computations and where they are most commonly used.

Feature	TensorFlow	PyTorch
Computation Style	Uses **static computation graphs** (build to run).	Uses **dynamic computation graphs** (run as you go).
Ease of Debugging	Harder to debug due to graph abstraction.	Easier to debug because it works like normal Python.
Common Usage	Industry, production systems, deployment.	Research, experimentation, model development.
Deployment Tools	TensorFlow Serving, TensorFlow Lite, TensorFlow.js	Deployment improving but still catching up

- **TensorFlow** can be used if you want to deploy a model to mobile devices, web, or large-scale production systems.
- **PyTorch** can be used when you are researching new models or need rapid experimentation and flexibility.

➢ **Q2:**

Interactive Model Development:
Jupyter allows you to test code in small parts, visualize results immediately, and adjust parameters quickly making it ideal for training ML models step-by-step.

Data Exploration and Visualization:
Researchers use Jupyter to load datasets, clean them, generate plots, analyze patterns, and document insights—all in the same workspace.

➢ **Q3:**

spaCy is an advanced NLP framework that can **understand language structure**, not just manipulate text.

Basic Python string operations only find words or split sentences — they do **not** know grammar or meaning.

spaCy provides:

- **Tokenization** (splitting text correctly into words)
- **Part-of-speech tagging**
- **Named Entity Recognition (NER)** (identifying names, places, brands)
- **Dependency parsing** (understanding sentence structure)

This makes **spaCy** far more powerful for real-world language tasks.

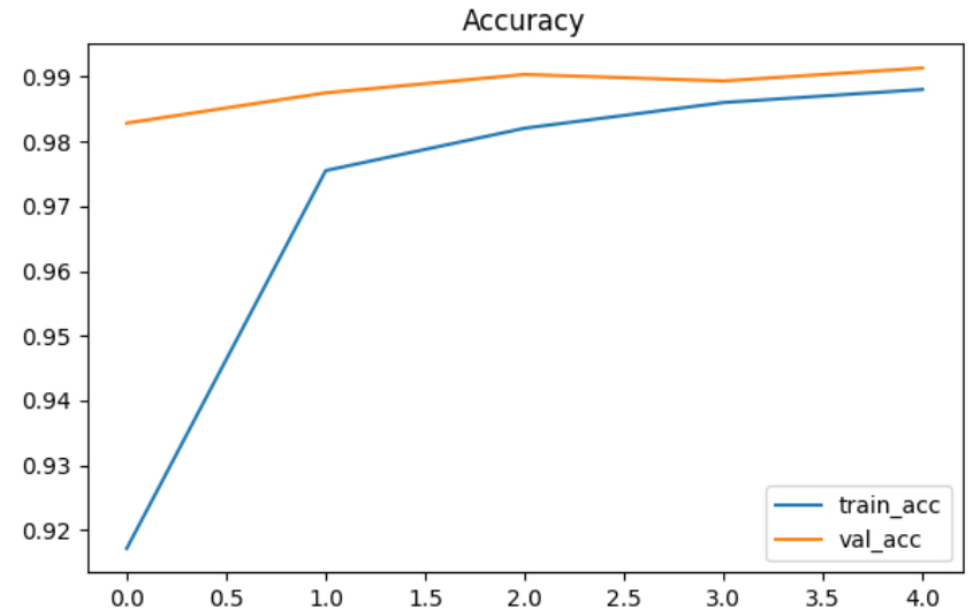
Comparative Analysis: Scikit-learn vs TensorFlow

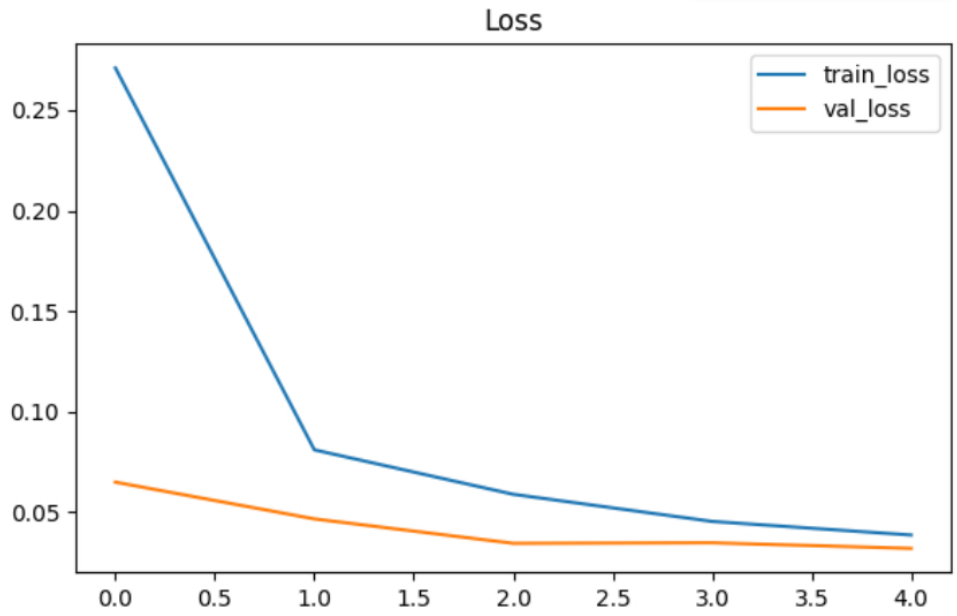
Aspect	Scikit-learn	TensorFlow
Best Used For	Classical ML (SVM, Logistic Regression, Random Forests)	Deep Learning + Neural Networks
Complexity	Beginner-friendly	More complex, steeper learning curve
Applications	Small-medium datasets, tabular data	Large datasets, images, audio, sequence data
Community Support	Strong for ML basics	Strong for deep learning deployments

Commands | + Code | + Text | Run all

RAM
Disk

```
Test accuracy: 0.9906
Saved model to mnist_cnn_model.h5
```






```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 412ms/step
```

Pred: 7 Pred: 2 Pred: 1 Pred: 0 Pred: 4

Variables | Terminal

11:42 AM | T4 (Python 3)

Dataset preview:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	species
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

Accuracy: 1.0

Classification report:

	precision	recall	f1-score	support
setosa	1.00	1.00	1.00	10
versicolor	1.00	1.00	1.00	9
virginica	1.00	1.00	1.00	11
accuracy			1.00	30
macro avg	1.00	1.00	1.00	30
weighted avg	1.00	1.00	1.00	30

	text	entities	sentiment
0	I recently bought a Samsung Galaxy smartphone ...	[(Samsung Galaxy, ORG), (Amazon, ORG)]	Positive
1	Terrible battery life on the Acme headphones. ...	[(Acme, ORG)]	Negative
2	The QuickBrew coffee maker from BrewCorp works...	[(QuickBrew, ORG), (BrewCorp, ORG)]	Positive
3	I love the fabric and fit of the Zara jacket I...	[(Zara, PERSON)]	Positive

```
Saved spaCy results to spacy_ner_sentiment.csv
```