HEPro AI/ML Internship

Activity 3 – Machine Learning: Student Segmentation & Risk Detection

Cluster Interpretation Report

Submitted By: Subhajit Das
Project: Dedicated Mentoring System for Students (HEPro AI+)
Technique Used: K-Means Clustering

---

1. Introduction

The purpose of this activity is to identify hidden behavioral and performance patterns among students using machine learning techniques. By applying unsupervised learning to readiness scores, students are segmented into meaningful groups that enable targeted mentoring strategies.

---

2. Methodology

Student data was first preprocessed and normalized using StandardScaler to ensure balanced feature contribution. K-Means clustering was then applied on the following dimensions:

- Academic Performance Score (APS)

- Wellness & Wellbeing Score (WWS)

- Productivity & Time Management Score (PTMS)

- Career Readiness Score (CRS)

The model grouped students into four clusters representing distinct mentoring profiles.

---

3. Cluster Interpretations

🔴 Cluster 1 – At-Risk Students

Key Indicators:

- Lower Student Readiness Index (SRI)

- Reduced academic and career readiness

Interpretation:
Students in this cluster require immediate intervention through structured mentoring, academic support, and career guidance.

🟡 Cluster 3 – Productivity Risk

Key Indicators:

- Low PTMS scores

- Average academic performance

Interpretation:
These students demonstrate potential but struggle with time management and engagement. Focused productivity mentoring can significantly improve outcomes.

🔵 **Cluster 0 – Stable Students**

Key Indicators:

- Balanced scores across readiness dimensions

- Moderate SRI values

Interpretation:
Stable students benefit from regular mentoring and performance monitoring rather than urgent intervention.

🟢 **Cluster 2 – High Performers**

Key Indicators:

- Higher SRI values

- Strong wellness and career readiness

Interpretation:
These students are ideal candidates for advanced mentoring, leadership opportunities, and career acceleration programs.

---

4. ML vs Rule-Based Comparison

Machine learning segmentation closely aligns with rule-based SRI categorization. This confirms that clustering enhances insight generation while maintaining explainability within the mentoring system.

---

5. Conclusion

The clustering process successfully reveals meaningful student groups and provides a strong foundation for personalized mentoring interventions. Integrating ML insights with rule-based scoring enables scalable and interpretable student support.