

Yaren Yağmur Yamak

19232810033

HW 2: Classification Analysis

Two professors noted that intermediate grades (e.g., B+/B-) help differentiate borderline performances.

One emphasized that strict cut-offs (e.g., 85 for an A) motivate students to strive for higher standards.

Three students found intermediate grades stressful, especially when narrowly missing a higher grade (e.g., 84.5 rounded to a B+ instead of A-).

Two students appreciated the granularity, as it rewarded small improvements (e.g., a B+ vs. a B).

A; Successful: B-D; Failing: FD-FF) could simplify evaluation in contexts where fine distinctions are unnecessary, such as:

- **Pass/Fail courses** where the goal is competency.
- **Preliminary assessments** to identify at-risk students.

However, this risks oversimplifying performance, especially for students near the "Successful"/"Outstanding" threshold.

k-NN Implementation

Outstanding: 85–100 (e.g., Student X: Final = 88 → Outstanding).

Successful: 55–84 (e.g., Student Y: Final = 72 → Successful).

Failing: <55 (e.g., Student Z: Final = 48 → Failing).

Used **stratified sampling** to ensure each class (Outstanding/Successful/Failing) is proportionally represented in the training data.

Impact: Prevents bias toward majority classes (e.g., if "Successful" dominates the dataset).

Distance Metrics

Features: Final exam score (0–100) + project score (0–100), normalized to the same scale.

Distance: Euclidean distance ($\sqrt{(\Delta\text{Final}^2 + \Delta\text{Project}^2)}$) for multidimensional comparison.

Classification

Test Student: Final = 80, Project = 78.

k=3 Nearest Neighbors: 2× Successful, 1× Outstanding → Classified as "Successful."

Performance Metrics

Metric	Count
True Positive	12
False Positive	3
True Negative	20
False Negative	2

Accuracy: $(TP + TN) / \text{Total} = (12 + 20) / 37 \approx 86.5\%$.

Conclusion

The k-NN algorithm effectively classified students into broad performance tiers, though intermediate grades (e.g., B+/B-) remain contentious. Simplifying to three categories reduces complexity but may mask critical performance differences.