## REPORT

Evaluating letter grades as a classification tool

When we think about letter grades like ABC and so on, they are basically a way to classify students into success levels. It’s like turning a student’s performance into a label.

But real life isn't always that clean, a student who barely gets 85 might receive an “a” just like someone who scored a perfect 100. That doesn't always feel fair especially when the effort or understanding is very different so while the system is useful, it’s not always precise. It groups together students who might not be on the same level.

Interview summary faculty and students

Faculty interviews

1. Faculty 2 interviews one professor emphasizes that precise grading EG b + a minus etcetera motivates students and better reflects effort in borderline cases.
2. Another professor mentioned that having too many grade differences can cause unnecessary stress and doesn't really make a big difference in student outcomes.

Student 5 interviews

1. Three students preferred more precise letter grades as they felt it fairly rewarded small differences in efforts like A= 95 and more.
2. Two students favored broader labels for example ABC to reduce anxiety and simplify GBA understanding.
3. All agreed that in the job market product classifications are more meaningful than plus minus distinctions.

KNN Classification implementation

1. manual labeling

All students were manually classified into outstanding successful or failing based on their total grade. None of the students got outstanding because they all got less than 85.

1. sampling method

Out of 141 students 21 students were selected as training data

And 120 students were selected as test data

For training students successful and failed students were selected to offer a wide range of training data.

The model used only two features projects and final exam.

These two components together contributed 65% of the total grade had minimal missing data they were chosen to ensure stable classification without noisy quiz input.

I used Euclidean distance

\text{Distance} = \sqrt{(x\_{\text{proj}} - y\_{\text{proj}})^2 + (x\_{\text{final}} - y\_{\text{final}})^2}

KNN classification

The KNN algorithm was used with k = 3 neighbors. The model predicted test students categories based on the most common label among the three nearest training examples.

Confusion matrix and evaluation

I first did 20 students for testing and used the KNN formula. It was fairly accurate with manual classification, but since I was doing 111 students, it was hard. I did the remaining students using colab. Initially the intention was to calculate true positive, true negative, false positive and false negative using collab and also finding the KNN for the data but collab treated all comparison like KNN results, manual category as simply true or false even when trying to label them as true positive or false negative using conditionals collapse continued returning only true/false.

As a result I've found that doing these steps manually in excel was more effective. It allowed me to;

* visually check each case manually
* assign the correct confusion label i.e. (true positive for correctly predicted “failing”)
* be confident in my final confusion matrix analysis but with a lot of classification data it was hard to do it manually