What does A, B, C & X actually intrinsically mean?

… it's arbitrary & completely subjective

"The oil additives **indicate wrong oil reported**. All other test results appear normal. Continue to sample at consistent intervals."

=

"Oil add pack suggests **wrong oil name**. All other test results appear normal. Continue to sample at consistent intervals."

A ≠ **B**

… and a lot of the time doesn't even concern the machine's health

"The change in viscosity and additives indicate **mixed up compartments**. The zero soot and history confirm that this sample is not from an engine compartment. Suggest resample."

//

Classifying machinery condition using oil samples and binary logistic regression

How would such a model even help when put in practice?

… caveats inclusive; where's the value here?

Quite literally useless to almost all customers as they're already paying an okay price for the oil sample analysis with the **interpretation included**!

They don't stand to gain much by automating this process unless there's an initiative to move the oil sampling analytics in-house.

The **labs** themselves can definitely benefit, and I'd say that's where Phillips et al. were going with this, but…

… why not frame it as an **anomaly detection problem**? Model the 'normal' or 'healthy' (A-class) very confidently, save that interpretation work, and then bring attention to the samples which show differences!

??