

Position: Don't Use the CLT in LLM Evals With Fewer Than a Few Hundred Datapoints

Sam Bowyer¹ Laurence Aitchison^{1,*} Desi R. Ivanova^{2,*}
University of Bristol¹ University of Oxford² *Equal Contribution



A block title

Some block contents, followed by a diagram, followed by a dummy paragraph.



Figure 1. A figure caption.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi ultricies eget libero ac ullamcorper. Integer et euismod ante. Aenean vestibulum lobortis augue, ut lobortis turpis rhoncus sed. Proin feugiat nibh a lacinia dignissim. Proin scelerisque, risus eget tempor fermentum, ex turpis condimentum urna, quis malesuada sapien arcu eu purus.

A block containing a list

Nam vulputate nunc felis, non condimentum lacus porta ultrices. Nullam sed sagittis metus. Etiam consectetur gravida urna quis suscipit.

- **Mauris tempor** risus nulla, sed ornare
- **Libero tincidunt** a duis congue vitae
- **Dui ac pretium** morbi justo neque, ullamcorper

Eget augue porta, bibendum venenatis tortor.

A highlighted block

This block catches your eye, so **important stuff** should probably go here.

Curabitur eu libero vehicula, cursus est fringilla, luctus est. Morbi consectetur mauris quam, at finibus elit auctor ac. Aliquam erat volutpat. Aenean at nisl ut ex ullamcorper eleifend et eu augue. Aenean quis velit tristique odio convallis ultrices a ac odio.

Fusce dapibus tellus vel tellus semper finibus. In consequat, nibh sed mattis luctus, augue diam fermentum lectus.

In euismod erat metus non ex. Vestibulum luctus augue in mi condimentum, at sollicitudin lorem viverra.

Suspendisse vulputate mauris vel placerat consectetur. Mauris semper, purus ac hendrerit molestie, elit mi dignissim odio, in suscipit felis sapien vel ex.

Aenean tincidunt risus eros, at gravida lorem sagittis vel. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae.

A block containing an enumerated list

Vivamus congue volutpat elit non semper. Praesent molestie nec erat ac interdum. In quis suscipit erat. **Phasellus mauris felis, molestie ac pharetra quis**, tempus nec ante. Donec finibus ante vel purus mollis fermentum. Sed felis mi, pharetra eget nibh a, feugiat eleifend dolor. Nam mollis condimentum purus quis sodales. Nullam eu felis eu nulla eleifend bibendum nec eu lorem. Vivamus felis velit, volutpat ut facilisis ac, commodo in metus.

1. **Morbi mauris purus**, egestas at vehicula et, convallis accumsan orci. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.
2. **Cras vehicula blandit urna ut maximus**. Aliquam blandit nec massa ac sollicitudin. Curabitur cursus, metus nec imperdiet bibendum, velit lectus faucibus dolor, quis gravida metus mauris gravida turpis.
3. **Vestibulum et massa diam**. Phasellus fermentum augue non nulla accumsan, non rhoncus lectus condimentum.

Fusce aliquam magna velit

Et rutrum ex euismod vel. Pellentesque ultricies, velit in fermentum vestibulum, lectus nisi pretium nibh, sit amet aliquam lectus augue vel velit. Suspendisse rhoncus massa porttitor augue feugiat molestie. Sed molestie ut orci nec malesuada. Sed ultricies feugiat est fringilla posuere.

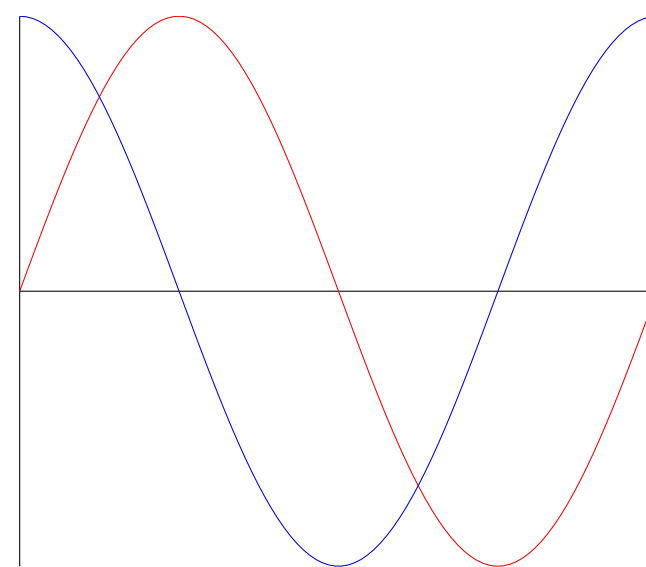


Figure 2. Another figure caption.

Nam cursus consequat egestas

Nulla eget sem quam. Ut aliquam volutpat nisi vestibulum convallis. Nunc a lectus et eros facilisis hendrerit eu non urna. Interdum et malesuada fames ac ante *ipsum primis* in faucibus. Etiam sit amet velit eget sem euismod tristique. Praesent enim erat, porta vel mattis sed, pharetra sed ipsum. Morbi commodo condimentum massa, *tempus venenatis* massa hendrerit quis. Maecenas sed porta est. Praesent mollis interdum lectus, sit amet sollicitudin risus tincidunt non.

Etiam sit amet tempus lorem, aliquet condimentum velit. Donec et nibh consequat, sagittis ex eget, dictum orci. Etiam quis semper ante. Ut eu mauris purus. Proin nec consectetur ligula. Mauris pretium molestie ullamcorper. Integer nisi neque, aliquet et odio non, sagittis porta justo.

- **Sed consequat** id ante vel efficitur. Praesent congue massa sed est scelerisque, elementum mollis augue iaculis.
 - In sed est finibus, vulputate nunc gravida, pulvinar lorem. In maximus nunc dolor, sed auctor eros porttitor quis.
 - Fusce ornare dignissim nisi. Nam sit amet risus vel lacus tempor tincidunt eu a arcu.
 - Donec rhoncus vestibulum erat, quis aliquam leo gravida egestas.
- **Sed luctus, elit sit amet** dictum maximus, diam dolor faucibus purus, sed lobortis justo erat id turpis.
- **Pellentesque facilisis dolor in leo** bibendum congue. Maecenas congue finibus justo, vitae eleifend urna facilisis at.

A highlighted block containing some math

A different kind of highlighted block.

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

Interdum et malesuada fames {1, 4, 9, ...} ac ante ipsum primis in faucibus. Cras eleifend dolor eu nulla suscipit suscipit. Sed lobortis non felis id vulputate.

A heading inside a block

Praesent consectetur mi $x^2 + y^2$ metus, nec vestibulum justo viverra nec. Proin eget nulla pretium, egestas magna aliquam, mollis neque. Vivamus dictum **uTv** sagittis odio, vel porta erat congue sed. Maecenas ut dolor quis arcu auctor porttitor.

Another heading inside a block

Sed augue erat, scelerisque a purus ultricies, placerat porttitor neque. Donec $P(y \mid x)$ fermentum consectetur $\nabla_x P(y \mid x)$ sapien sagittis egestas. Duis eget leo euismod nunc viverra imperdiet nec id justo.

Nullam vel erat at velit convallis laoreet

Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Phasellus libero enim, gravida sed erat sit amet, scelerisque congue diam. Fusce dapibus dui ut augue pulvinar iaculis.

First column	Second column	Third column	Fourth
Foo	13.37	384,394	α
Bar	2.17	1,392	β
Baz	3.14	83,742	δ
Qux	7.59	974	γ

Table 1. A table caption.

Donec quis posuere ligula. Nunc feugiat elit a mi malesuada consequat. Sed imperdiet augue ac nibh aliquet tristique. Aenean eu tortor vulputate, eleifend lorem in, dictum urna. Proin auctor ante in augue tincidunt tempor. Proin pellentesque vulputate odio, ac gravida nulla posuere efficitur. Aenean at velit vel dolor blandit molestie. Mauris laoreet commodo quam, non luctus nibh ullamcorper in. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos.

Nulla varius finibus volutpat. Mauris molestie lorem tincidunt, iaculis libero at, gravida ante. Phasellus at felis eu neque suscipit suscipit. Integer ullamcorper, dui nec pretium ornare, urna dolor consequat libero, in feugiat elit lorem euismod lacus. Pellentesque sit amet dolor mollis, auctor urna non, tempus sem.

References

[1] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.