Revisions made in this submission.

Below we list **all changes introduced since the last review cycle**, grouped by the issues reviewers raised. Each bullet notes the exact additions or rewrites and where they appear in the revised, fully anonymised manuscript.

- **Finer-grained embedding analysis (R-9YM2)** — Added a **mean-centring experiment**. Figure 1 in section 3.1, before plots was cosine-similarity distributions which is moved to appendix and *later* added mean centring plot, and describe the outcome and quantifies how centring removes global bias or ADA/PaLM but not most open-source LLMs.

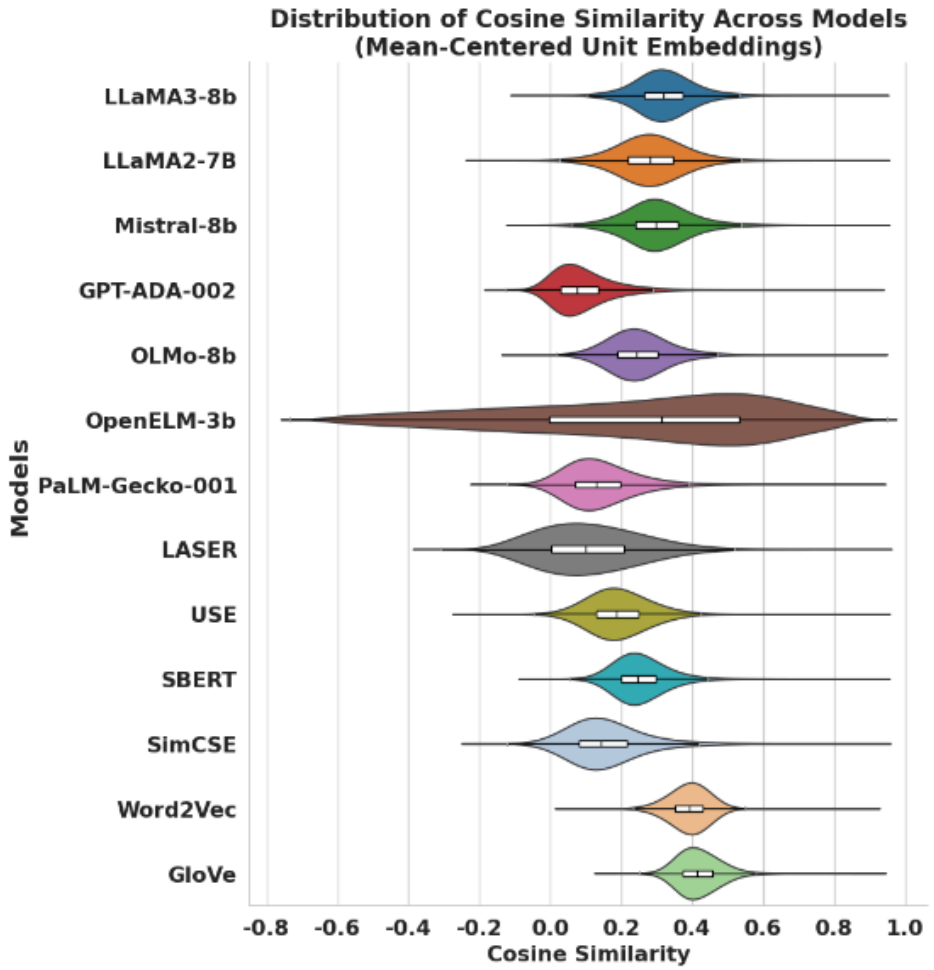

Figure 1: The Mean-Centered Embedding distribution of cosine similarities between all pairs of words.

- **New downstream-task showcase (meta-reviewer)** — introduces **Table 3 and 5 showed the "Evaluation of LLMs on downstream task."** The table reports performance on two categories; Reasoning focused and Embedding focused. We discussed the outcome and showed how we use our criteria as a model performance predictor on downstream task.

Model	Sbert	SimCSE	LlaMA2-7b	LlaMA3-8b	Mistral-7b	OLMo-7b	OpenELM-3B
ARC-c	-	-	54.2	**79.3**	**78.6**	48.5	35.58
ARC-e	-	-	84.0	**92.4**	**90.8**	65.4	59.89
BoolQ	-	-	86.1	**87.5**	**89.3**	74.4	67.4
HellaSwag	-	-	78.9	**81.8**	**83.0**	76.4	72.44
MMLU	-	-	46.2	**66.6**	**64.0**	40.5	26.76
PIQA	-	-	57.8	77.2	**80.6**	**78.4**	78.24
SIQA	-	-	77.5	81.6	**82.8**	78.5	**92.7**
WinoGrande	-	-	71.7	**76.2**	**77.9**	67.9	65.51
Clustering	42.35	29.04	45.24	**46.45**	**54.93**	32.0	18.71
Pair classification	82.37	70.33	**88.03**	87.8	**88.59**	49.32	56.71
Reranking	**58.04**	46.47	57.38	**59.68**	50.15	33.91	37.0
STS	78.9	74.33	**83.73**	83.58	**84.77**	27.04	38.31
Summarization	30.81	**31.15**	28.49	30.94	**36.32**	20.83	18.71

Table 5: Model Evaluation Results Across Various Tasks. **Blue** is top scorer and **black** is second best.

- A short "Discussions & Final Words" subsection distils practical take-aways for model selection in low-resource settings and outlines plans to release task scripts.

- **Minor corrections** — replace blue finding boxes with simple blue text.