A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is light green. They are positioned diagonally, with the blue one partially covering the green one.

# Exploring the Efficacy of LLMs for Expanding the CVE Dataset to Assist in Machine Learning

By: Mart Spil

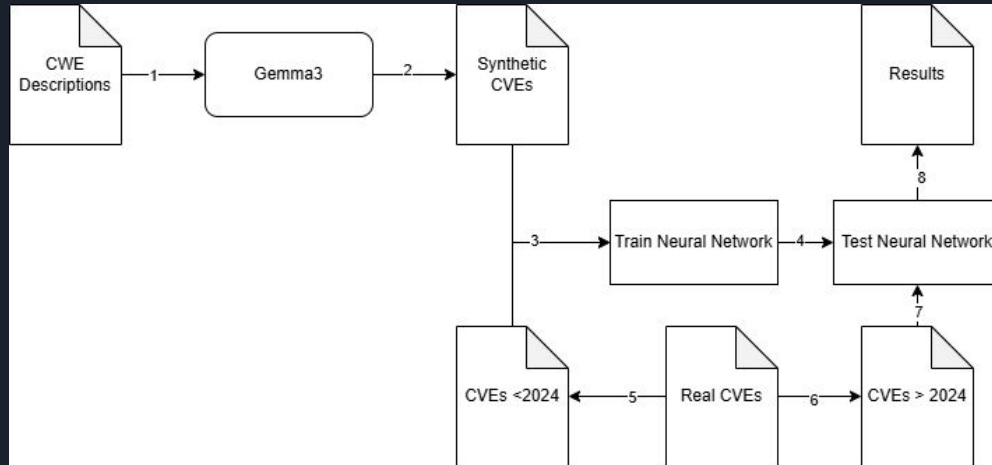


# Context

- What are CVEs
- What are CWEs
- The current problem
- How can LLMs help

# Approach

- Limit to good quality CVEs
- Generate new CVEs using a LLM with RAG
- Test Scenarios





# Results

Test Description	Precision	Recall	F1-score
Base (Capped)	$0.2041 \pm 0.0069$	$0.2754 \pm 0.0093$	$0.2065 \pm 0.0058$
Duplicated (Capped)	$0.2037 \pm 0.0048$	$0.2871 \pm 0.0082$	$0.2049 \pm 0.0054$
LLM Generated (Capped)	<b><math>0.2076 \pm 0.0055</math></b>	<b><math>0.2961 \pm 0.0055</math></b>	<b><math>0.2101 \pm 0.0038</math></b>

Table 1. Macro-averaged results for capped datasets (mean  $\pm$  standard deviation).

Test Description	Precision	Recall	F1-score
Base (Uncapped)	<b>0.3425</b>	0.2761	0.2825
LLM Generated (Uncapped)	0.3226	<b>0.2958</b>	<b>0.2876</b>



# Conclusion

- Quality of generated CVEs
- Future opportunities



Questions?