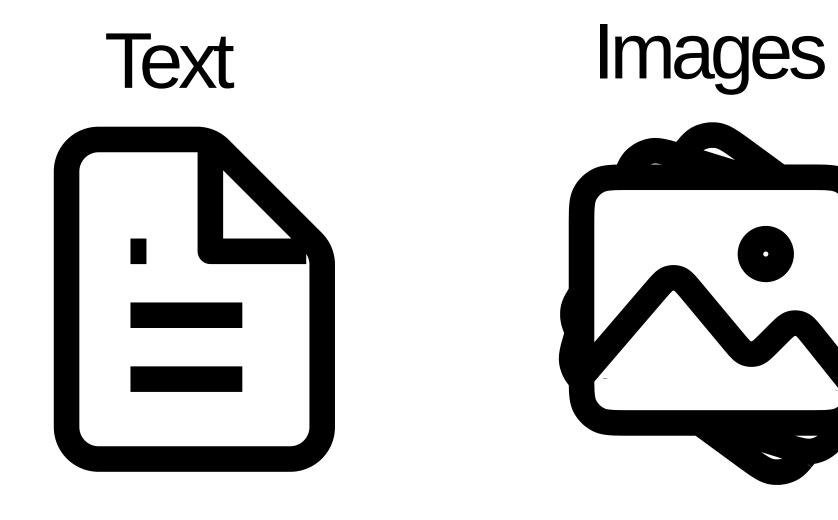


EMMa

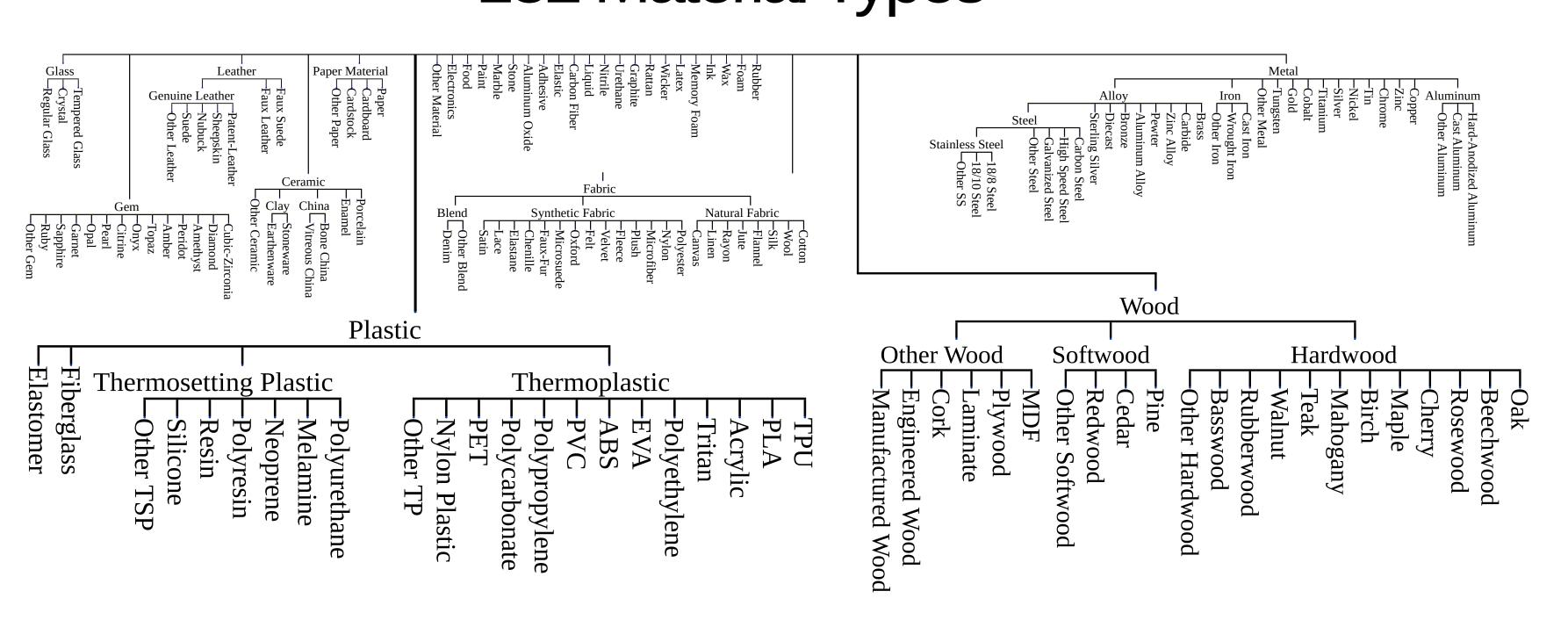
An Extensible, Multimodal dataset with Materials

Inputs



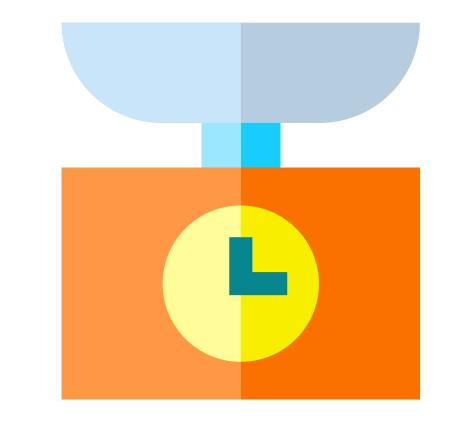


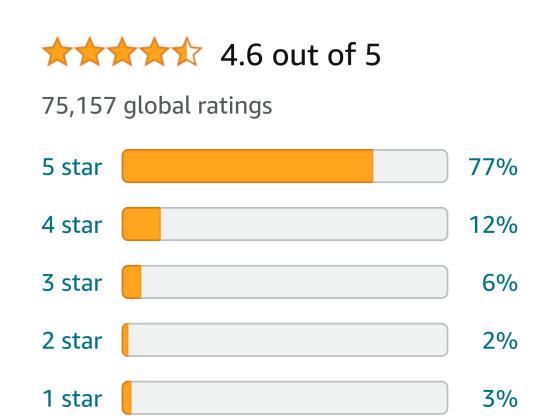
182 Material Types



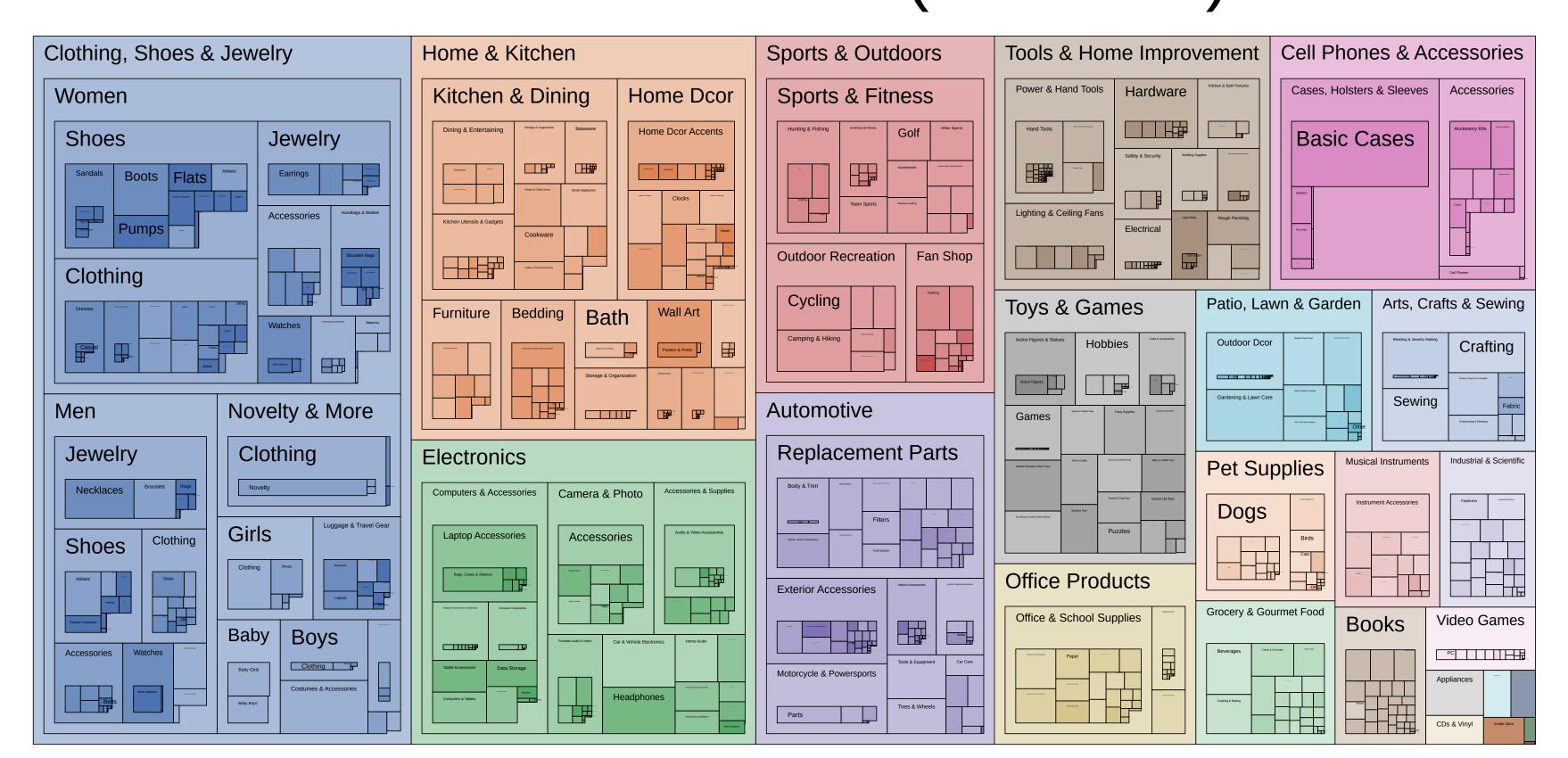
Price, Weight, and Ratings







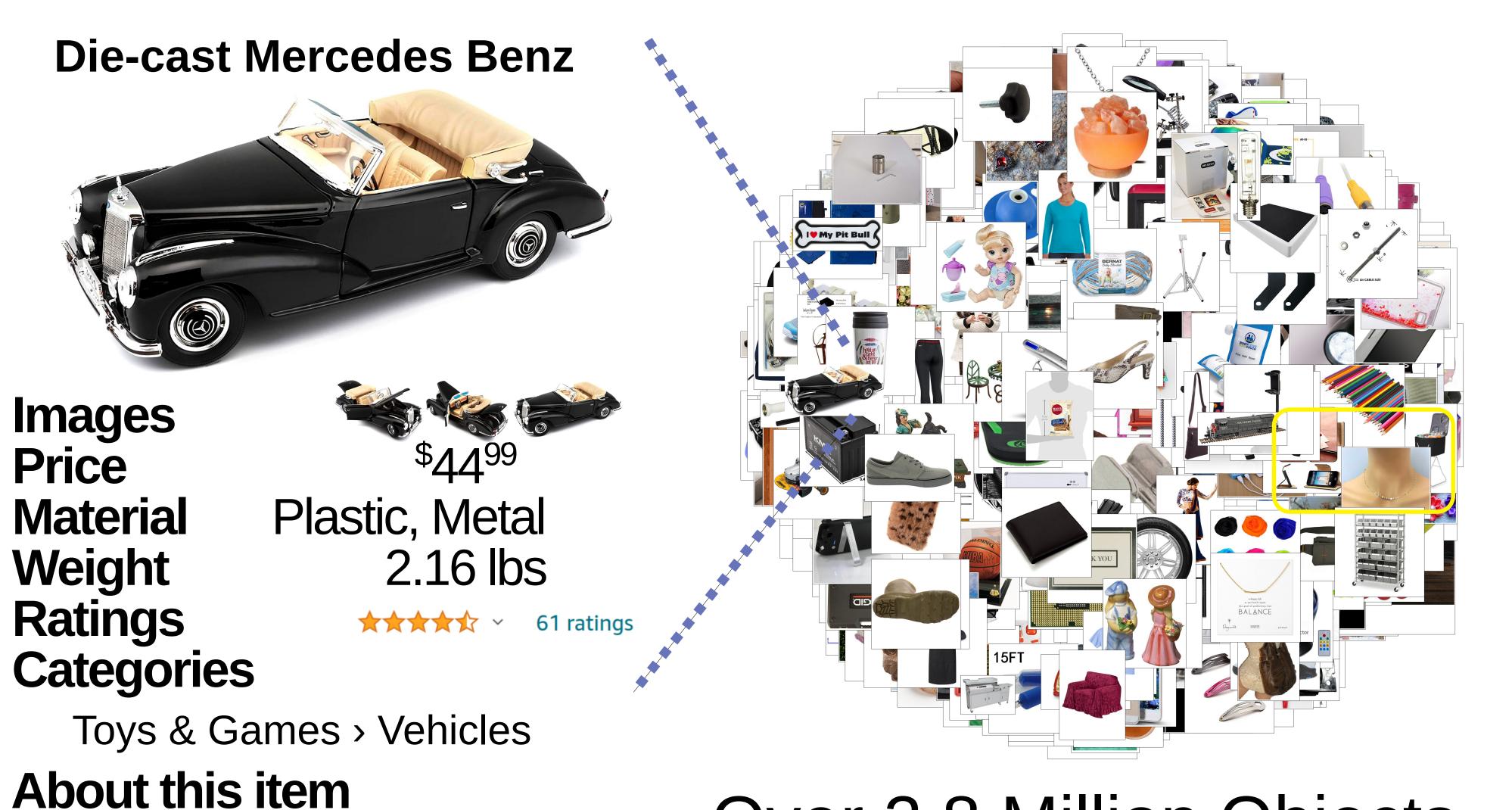
Taxonomic Classification (11k nodes)





An Extensible Multimodal Multi-task Object Dataset with Materials

Dawn Chen² Ruohan Gao¹ Jiajun Wu¹ Silvio Savarese^{1,3} Trevor Standley¹



Over 2.8 Million Objects 7.4 Million Images

Extensibility

Sharp	No
Transparent	Yes
Deformable	No
Articulating	Yes
Electronic	No
Your property	

Label ~1k Objects



Propagate to all 2.8 Million





EMMa is a large multi-task dataset for NLP and computer vision. It contains 2.8 million objects. Every object has one or more images, discriptive text, and labels for price, weight, materials, taxinomic class, and customer ratings. We train NLP models, computer vision models, and hybrid models that take images, text, and all other attributes. We use hybrid models to accurately fill in missing data. With all attributes filled in, we train a powerful object embedding. We leverage this embedding and active learning to allow users to quickly and accurately add custom features.

Filled in missing attributes with Expectation Maximization

	<u> </u>				
	Price (MnRE) ↑	Mass (MnRE) †	Materials (F1) ↑	Category (Acc) ↑	Ratings (KL-d) ↓
SINGLE IMAGE GEN 1	0.649	0.640	60.9	82.6	0.155
SINGLE IMAGE GEN 2	0.653	0.649	72.7	84.5	0.153
TEXT GEN 1	0.710	0.693	79.8	91.6	0.148
TEXT GEN 2	0.704	0.695	85.3	91.9	0.147
EVERYTHING GEN 1	0.736	0.773	80.0	92.1	0.138
EVERYTHING GEN 2	0.743	0.788	86.0	92.2	0.136

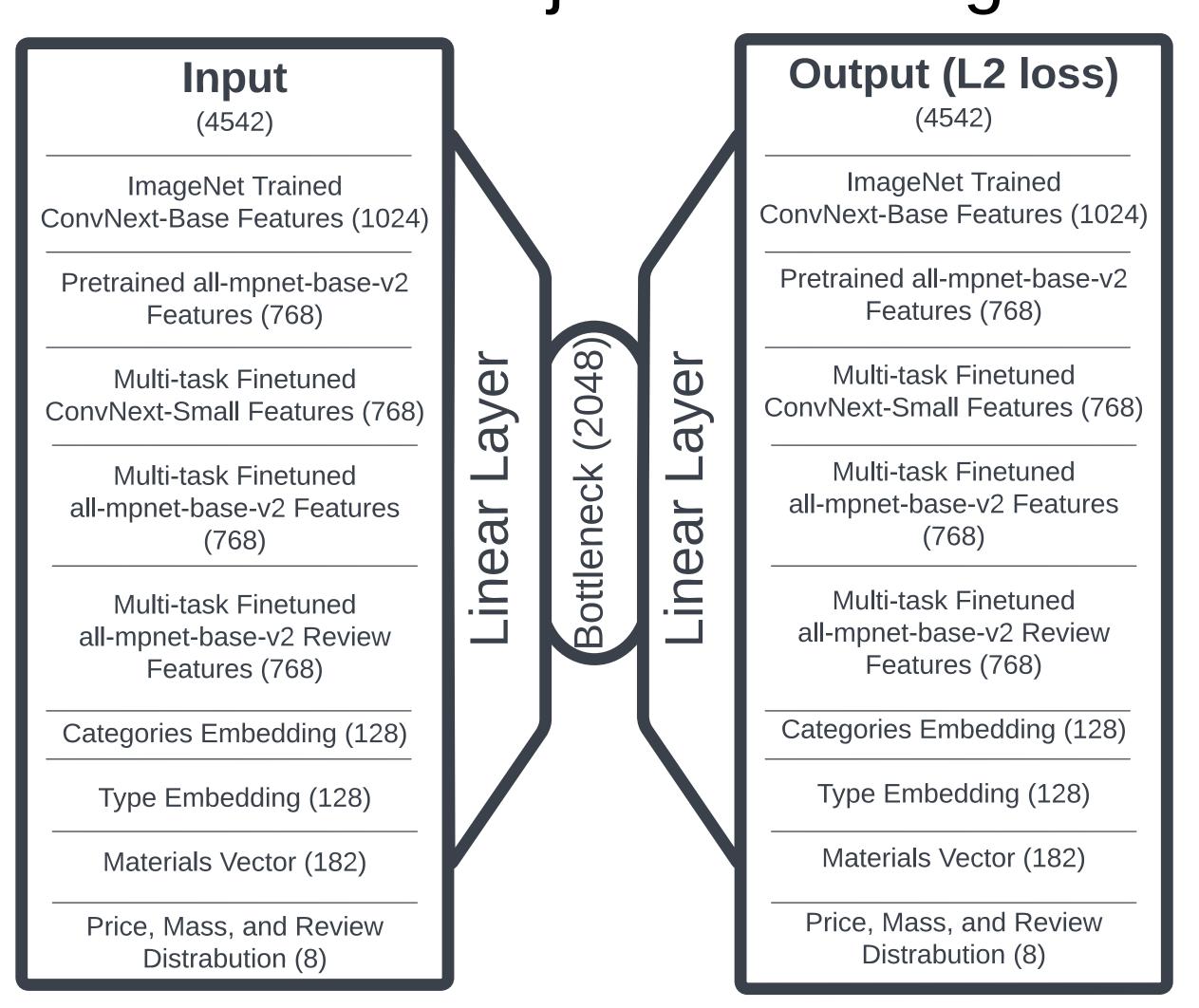
Powerful Object Embedding

Comes mounted on plastic stand

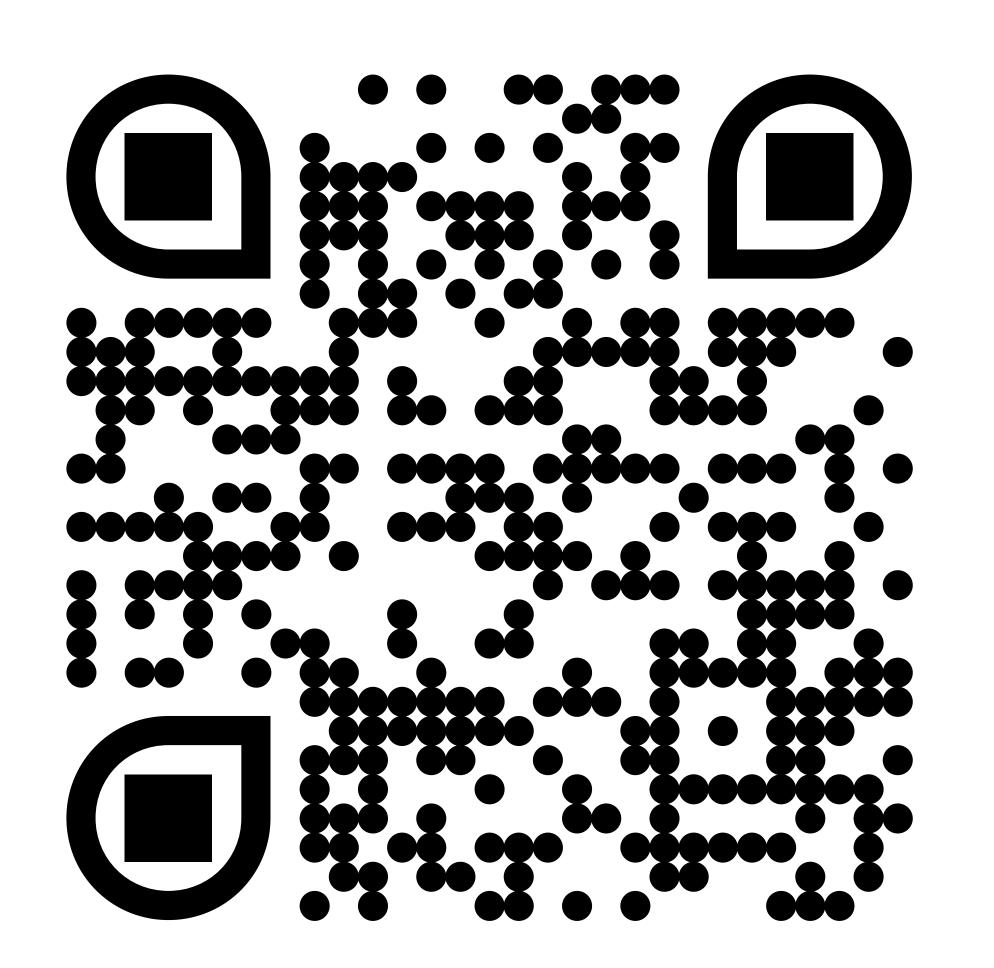
Opening doors and engine compartment

Detailed chassis with separate exhaust system

Full function steering with Four wheel spring suspension

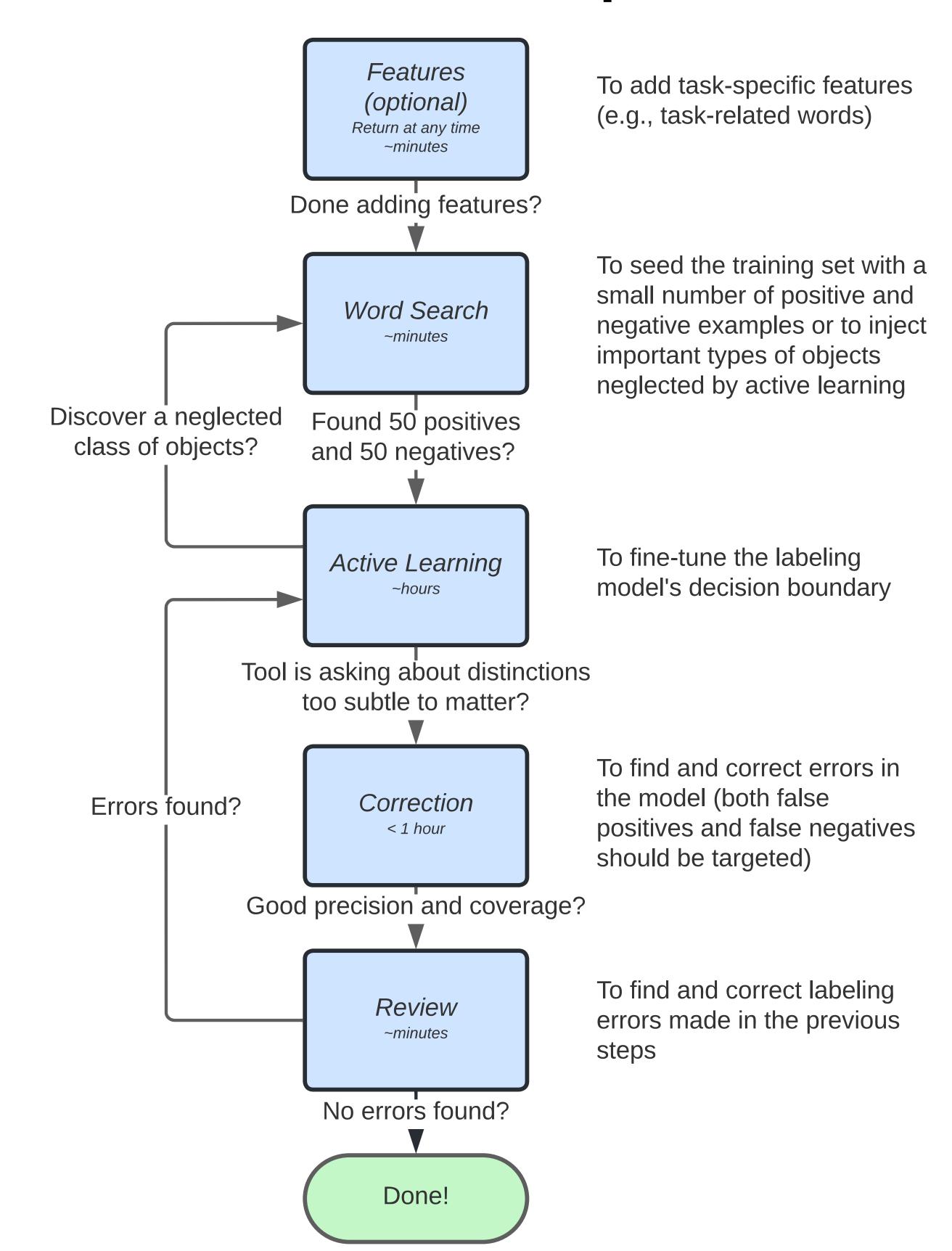


Website & Interactive Dataset Explorer

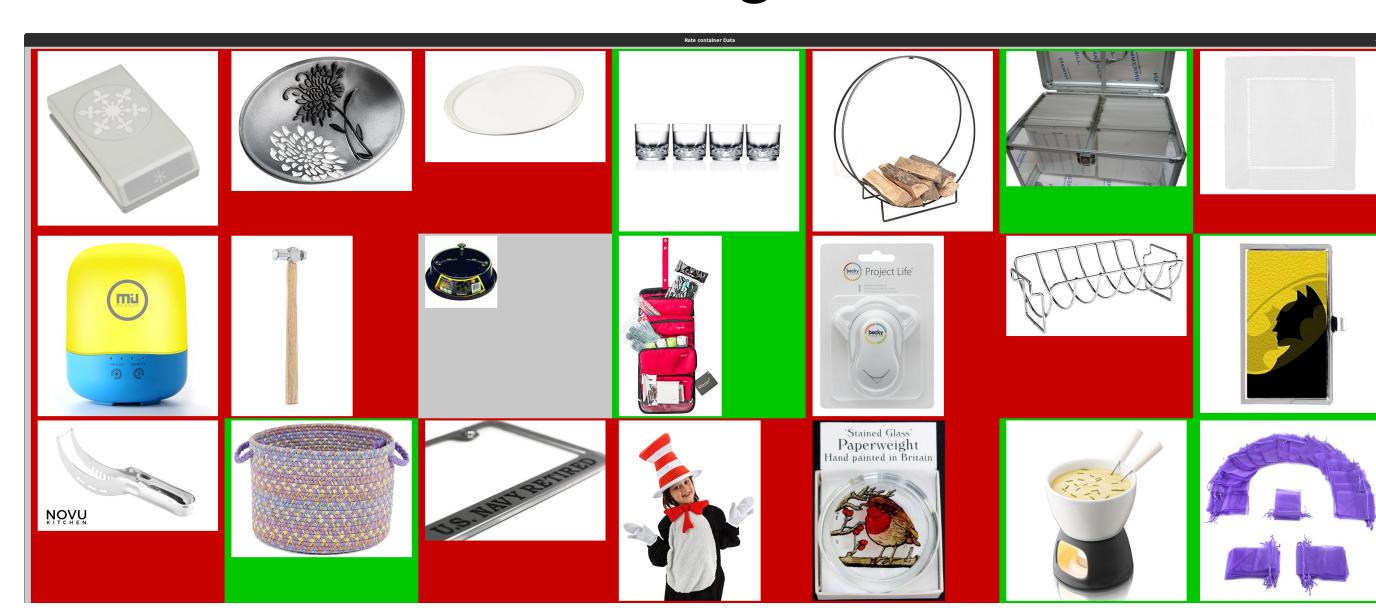




Add Your Own Properties



Active Learning Platform



Smart Labeling Results

	# Train	# Train +	# Test	# Test +	Precision	Recall	F1	Acc
SHARP	1331	633	1504	25	92.0	92.0	92.0	99.7
TRANSPARENT	2217	877	706	46	93.5	93.5	93.5	99.1
DEFORMABLE	636	350	253	150	96.7	97.3	97.0	96.4
ARTICULATING	1399	817	403	148	93.3	94.6	94.0	95.5
ELECTRONIC	525	248	301	63	96.8	95.2	96.0	98.3

