

## SUPPLEMENTARY MATERIALS

## Anonymous authors

Paper under double-blind review

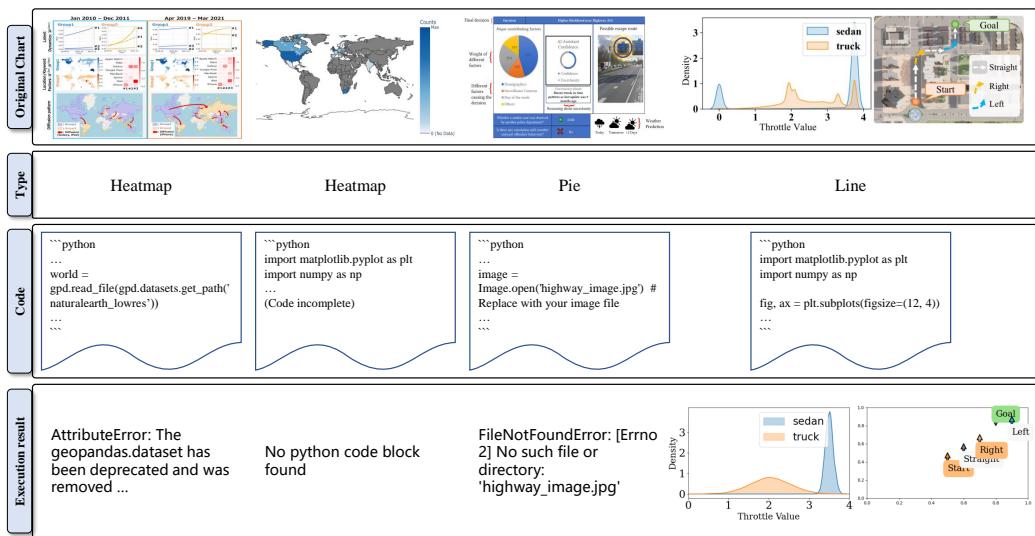
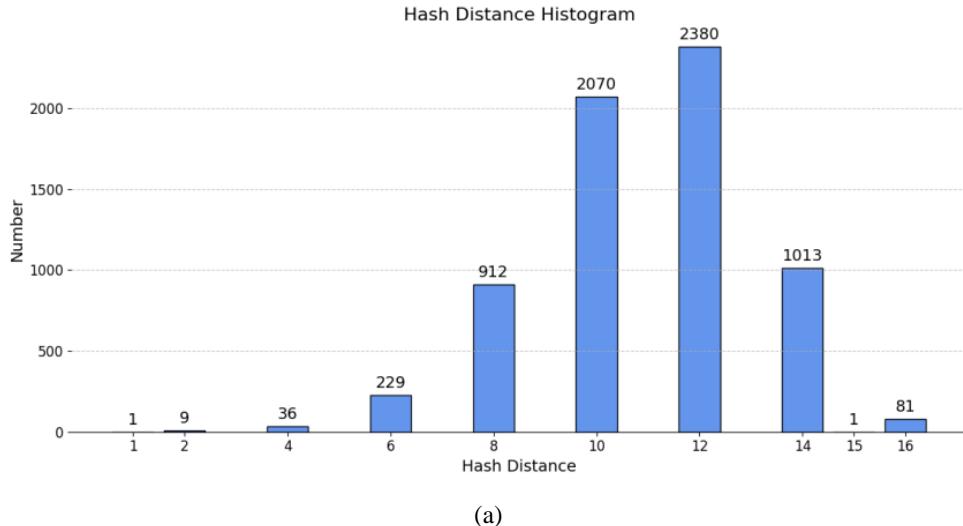


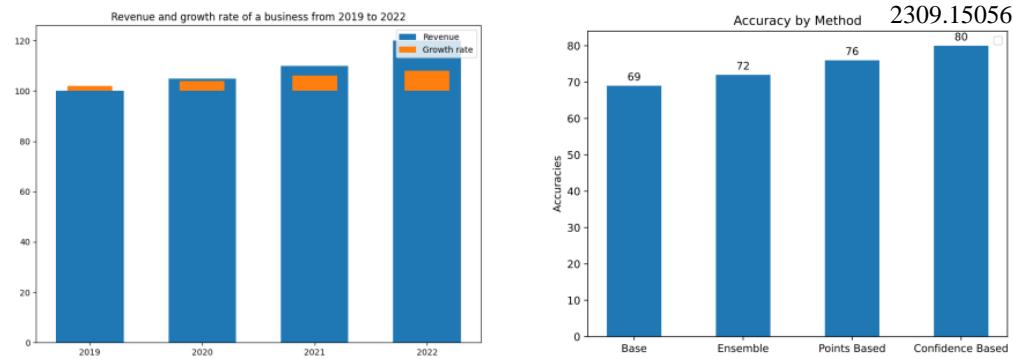
Figure A: Examples of misclassified chart images. Each column shows: the original image (Higashiguchi et al., 2025; Olson et al., 2025; Mehrotra et al., 2025; Shin et al., 2024), assigned chart type, extracted code, and code execution result. Most misclassified cases either fail to execute due to missing files or code errors, ensuring they are filtered out at the code validation stage.

054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074



(a)

075  
076  
077  
078  
079  
080  
081  
082  
083  
084  
085  
086  
087  
088  
089  
090  
091  
092  
093  
094  
095  
096  
097  
098  
099  
100

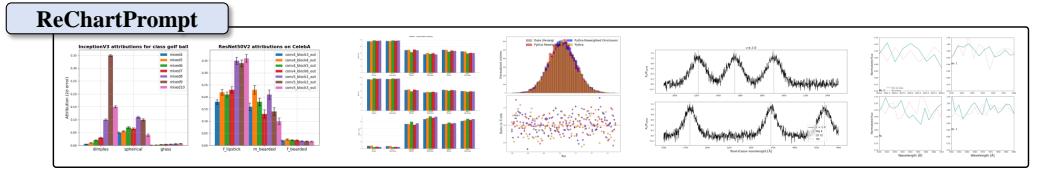


(b)

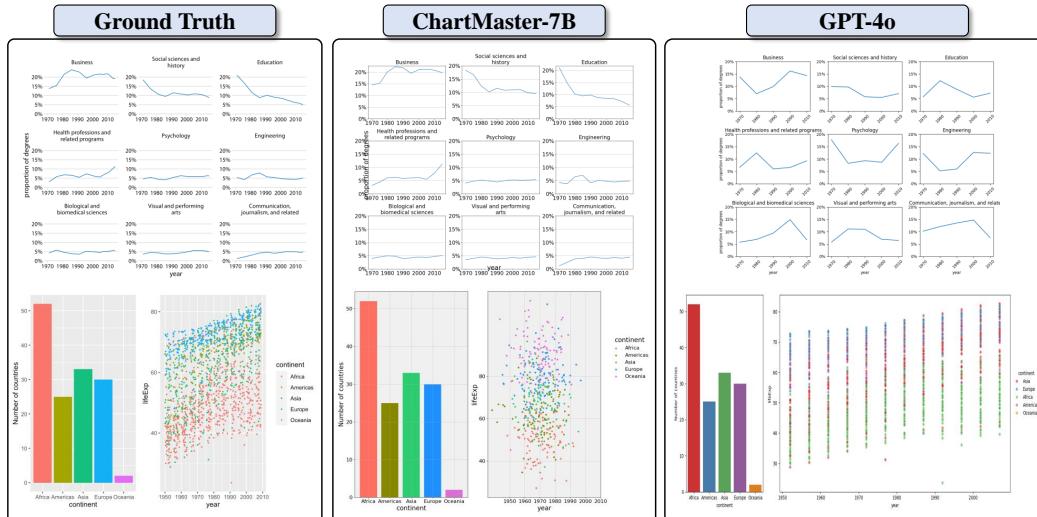
101  
102  
103  
104  
105  
106  
107

Figure B: (a) Histogram of minimum hash distances between each test image and its most similar chart (Silver et al., 2022; Hahlin et al., 2025) in the 288,992 charts set, indicating no near-duplicate images. (b) Examples of test images and their closest charts with hash distance 1–2, showing that even the most similar samples are visually distinct.

108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121



(a) complex charts from our dataset



(b) stress test on unseen complex chart

Figure C: (a) Examples of complex charts in the ReChartPrompt-240K dataset, illustrating the diversity and difficulty beyond simple chart types. (b) Stress test results on manually collected complex charts. ChartMaster-7B demonstrates strong generalization and visual consistency, successfully reconstructing multi-panel and faceted charts, and outperforming GPT-4o in structure fidelity.

145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161

162 REFERENCES  
163

164 A Hahlin, O Kochukhov, P Chaturvedi, E Guenther, A Hatzes, Ulrike Heiter, A Lavail, Evangelos  
165 Nagel, N Piskunov, K Pouilly, et al. Comparative study of small-scale magnetic fields on  $\xi$  boo a  
166 using optical and near-infrared spectroscopy. *Astronomy & Astrophysics*, 696:A4, 2025.

167 Shingo Higashiguchi, Yasuko Matsubara, Koki Kawabata, Taichi Murayama, and Yasushi Saku-  
168 rai. D-tracker: Modeling interest diffusion in social activity tensor data streams. *arXiv preprint*  
169 *arXiv:2505.00242*, 2025.

170 Siddharth Mehrotra, Ujwal Gadiraju, Eva Bittner, Folkert van Delden, Catholijn M. Jonker, and  
171 Myrthe L. Tielman. “even explanations will not help in trusting [this] fundamentally biased  
172 system”: A predictive policing case-study. In *Proceedings of the 33rd ACM Conference on User*  
173 *Modeling, Adaptation and Personalization*, pp. 51–62, 2025.

174

175 Lauren Olson, Ricarda Anna-Lena Fischer, Florian Kunnenman, and Emitzá Guzmán. Who speaks  
176 for ethics? how demographics shape ethical advocacy in software development. In *Proceedings of*  
177 *the 2025 ACM Conference on Fairness, Accountability, and Transparency*, pp. 2847–2862, 2025.

178 Sangwoo Shin, Minjong Yoo, Jeongwoo Lee, and Honguk Woo. Semtra: A semantic skill translator  
179 for cross-domain zero-shot policy adaptation. In *Proceedings of the AAAI Conference on Artificial*  
180 *Intelligence*, volume 38, pp. 15000–15008, 2024.

181

182 Daniel Silver, Tirthak Patel, and Devesh Tiwari. Quilt: Effective multi-class classification on quan-  
183 tum computers using an ensemble of diverse quantum classifiers. In *Proceedings of the AAAI*  
184 *Conference on Artificial Intelligence*, volume 36, pp. 8324–8332, 2022.

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215