



Evidence-hierarchy audit of public DGA benchmarks for trustworthy AI-assisted transformer assessment

Yebin Zou¹, Richang Xian^{1,*}, Yiming Qin¹, Zihao Chu², Tao Yu³ and Chuancai Jing³

¹ College of Electrical and Electronic Engineering, Shandong University of Technology, Zibo, China

² STEM College and Royal Melbourne Institute of Technology University, Melbourne, Australia

³ Shandong Huineng Electric Co., Ltd., Zibo, China

* Correspondence author; E-mail: xianrc@sdut.edu.cn.

Supplementary tables

Table S1. Explicit mechanism-supported feature sets M(c) used for the Kaggle five-gas MCS calculation.

Class	Raw-gas features	Ratio features	Severity features	Rationale
Normal	H ₂ , CH ₄ , C ₂ H ₆ , C ₂ H ₄ , C ₂ H ₂	None	TDCG, THC	Low total gas and low hydrocarbon evidence support normal-condition agreement in the public-label benchmark.
PD	H ₂	CH ₄ /H ₂	None	Hydrogen-dominant low-energy gas evidence supports PD-like separation.
D1	H ₂ , C ₂ H ₂	C ₂ H ₂ /C ₂ H ₄ , C ₂ H ₂ /CH ₄	None	Acetylene and acetylene-rich ratios support discharge discrimination.
D2	H ₂ , C ₂ H ₂ , C ₂ H ₄	C ₂ H ₂ /C ₂ H ₄ , C ₂ H ₂ /CH ₄	None	Acetylene and ethylene-related evidence support high-energy discharge separation.
T1	CH ₄ , C ₂ H ₆	CH ₄ /H ₂ , C ₂ H ₆ /CH ₄	THC	Methane and ethane evidence supports low-temperature thermal behavior.
T2	CH ₄ , C ₂ H ₄	C ₂ H ₄ /C ₂ H ₆ , C ₂ H ₆ /CH ₄	THC	Ethylene growth and hydrocarbon ratios support medium-temperature thermal behavior.
T3	C ₂ H ₄ , C ₂ H ₂	C ₂ H ₄ /C ₂ H ₆ , C ₂ H ₂ /C ₂ H ₄	TDCG, THC	High ethylene and total-gas severity support high-temperature thermal evidence; CO and CO ₂ were unavailable in the five-gas benchmark.

Table S2. Representative confidence-MCS selective-reporting results on the Kaggle public-label split.

Policy	tp	tm	Coverage	Review	n	Accuracy	Macro-F1	Error
All samples	0.00	NA	100.0%	0.0%	1038	0.9114	0.8967	8.9%
Confidence only	0.80	NA	78.0%	22.0%	810	0.9728	0.9663	2.7%
Confidence only	0.90	NA	63.9%	36.1%	663	0.9849	0.9823	1.5%
MCS only	NA	0.50	88.2%	11.8%	916	0.9159	0.9052	8.4%
MCS only	NA	0.70	62.0%	38.0%	644	0.9208	0.8903	7.9%
Confidence and MCS	0.80	0.50	71.5%	28.5%	742	0.9717	0.9655	2.8%
Confidence and MCS	0.90	0.50	59.2%	40.8%	615	0.9837	0.9811	1.6%



Copyright©2026 by the authors. Published by ELSP. This work is licensed under Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.