**Dataset**

|  | Accepted | Rejected | Total |
| --- | --- | --- | --- |
| CVPR |  |  |  |
| NeurIPS(2023-2024) | 600 | 150 | 750 |
| ICLR(2023-2024) | 600 | 150 | 750 |

NeurIPS

Accepted

Train : NeurIPS 2023년 Oral 67, Spotlight 133, Poster 100, NeurIPS 2024 Oral 61, Poster 119 -> Total 480

Test : NeurIPS 2024 Spotlight 120

-> Total 120

Rejected

Train : NeurIPS 2023 Rejected 120

Test : NeurIPS 2024 Rejected 30

ICLR

Accepted

Train : ICLR 2024 Oral 85 Spotlight 114 Poster 101, ICLR 2025 Oral 90 Poster 90

-> Total 480

Test : ICLR 2025 Spotlight 120

-> Total 120

Rejected

Train : ICLR 2025 Rejected 120

Test : ICLR 2024 Rejected 30

비율이 4:1인 이유는 3.7:1이라는 기존 논문의 방식에 따라 최대한 조정하였다.

NeurIPS 데이터셋을 보면 대체적으로 Train 셋이 Test셋에 비해 1년 빨리 투고된 논문인데에 반해 ICLR에서 Rejected 데이터셋에서 Train 셋이 2025년이고 Test셋이 2024년 것인데 기존과 다른 이유는 ICLR 2024 Rejected 논문을 120개 크롤링하려 했으나 갯수가 86개이고, 그렇다고 2024년 2025년 논문을 같이 하기에는 번거롭다고 여겨 기존 계획(Train Dataset: 2024, Test Dataset: 2025)에서 변경했다.