



|           |         | Acc.  | NR $F_1$ | FR $F_1$ | TR $F_1$ | UR $F_1$ |
|-----------|---------|-------|----------|----------|----------|----------|
| Twitter15 | GPT-3.5 | 0.313 | 0.084    | 0.188    | 0.410    | 0.368    |
|           | GPT-4   | 0.419 | 0.535    | 0.376    | 0.261    | 0.431    |
|           | KPG     | 0.893 | 0.921    | 0.898    | 0.903    | 0.847    |
| Twitter16 | GPT-3.5 | 0.375 | 0.109    | 0.172    | 0.520    | 0.434    |
|           | GPT-4   | 0.394 | 0.498    | 0.369    | 0.266    | 0.397    |
|           | KPG     | 0.889 | 0.894    | 0.836    | 0.930    | 0.896    |

Table 3: The results of LLMs on Twitter15 and Twitter16.

|         |         | Acc.  | R $F_1$ | NR $F_1$ |
|---------|---------|-------|---------|----------|
| PHEME   | GPT-3.5 | 0.599 | 0.484   | 0.672    |
|         | GPT-4   | 0.615 | 0.494   | 0.689    |
|         | KPG     | 0.859 | 0.863   | 0.854    |
| Weibo22 | GPT-3.5 | 0.670 | 0.704   | 0.627    |
|         | GPT-4   | 0.745 | 0.780   | 0.696    |
|         | KPG     | 0.949 | 0.949   | 0.948    |

Table 4: The results of LLMs on PHEME and Weibo22.