

**Supplemental Table S2: Classification Outcomes of ChatGPT-4 for Diabetic Macular Edema Detection**

|                              | <b>Yes DME (Y)</b>                            | <b>No DME (N)</b>                             |                                 |
|------------------------------|---|---|---------------------------------|
| <b>Predicted ME (GPT)</b>    | 248 (True Positive)<br>(48.1% of total cases) | 47 (False Positive)<br>(9.1% of total cases)  | 295 (57.2% predicted as DME)    |
| <b>Predicted No ME (GPT)</b> | 46 (False Negative)<br>(8.9% of total cases)  | 175 (True Negative)<br>(33.9% of total cases) | 221 (42.8% predicted as No DME) |
|                              | 294 (57.0% of all cases)                      | 222 (43.0% of all cases)                      |                                 |

*ME: Macular Edema*