Confusion Matrices - mistral-7b (Mistral)
Rows: Human verdict | Cols: Judge verdict | Φ = FaviScore