

F1-Score vs Accuracy by Model (AgenticSimLLM)

