F1-score vs. Elapsed Time [gpt5]