Dialogue Manager: Leaderboard

Data	Calls	Duration (hours)
Fisher (American English)	117	19.72
Cult (Hinglish)	100	27.93
Total	217	47.65

Class	Distribution
Turn End	6244
Interruption	1261
Backchannel	5650
Pause	11856

Models:

1. **DM_V1** - 110525_baseline_training1_RANDOM_BATCHES_EQUAL_TASKS_98000 (Latest)
   a. We introduced synthetic interruption and better labelling of Pauses.
2. **DM_V0** - DM_v2_run2_uniform_batching_20032025_100K_predicitons_fisher_cult_v2_1488 (Last version of DM)
3. **VAD**: Runs in Production

Results in Tabular Format

1. F1 Scores

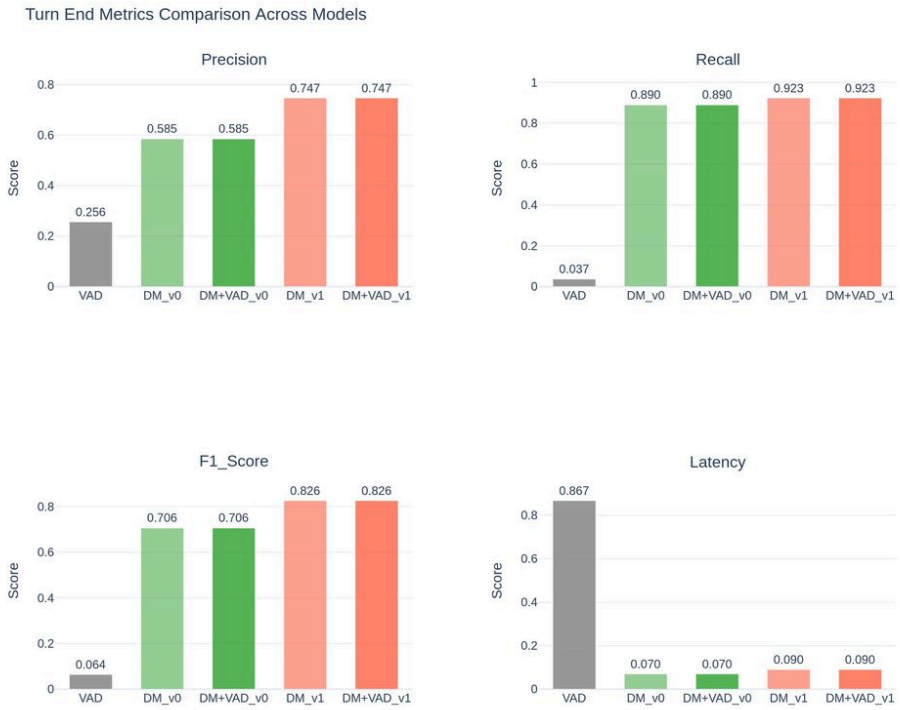
Model	Interruptions F1	Turn End F1	Pauses F1	Backchannel F1	Synthetic Interruptions Accuracy	Break Accuracy
DM_v1	0.37	0.82	0.89	0.89	0.98	0.99
DM+VAD_v1	0.35	0.82	0.89	0.88	0.95	0.99
DM+VAD_v0	0.36	0.70	0.74	0.47	0.96	0.95
DM_v0	0.37	0.70	0.74	0.40	0.98	0.85
VAD	0.34	0.06	0.77	0.63	0.81	1.00

Model	Interruptions Precision	Turn End Precision	Pauses Precision	Backchannel Precision
DM_v1	0.51	0.75	0.95	0.86
DM+VAD_v1	0.41	0.75	0.95	0.85
DM_v0	0.23	0.58	0.93	0.98
DM+VAD_v0	0.22	0.58	0.92	0.92
VAD	0.23	0.26	0.65	0.87

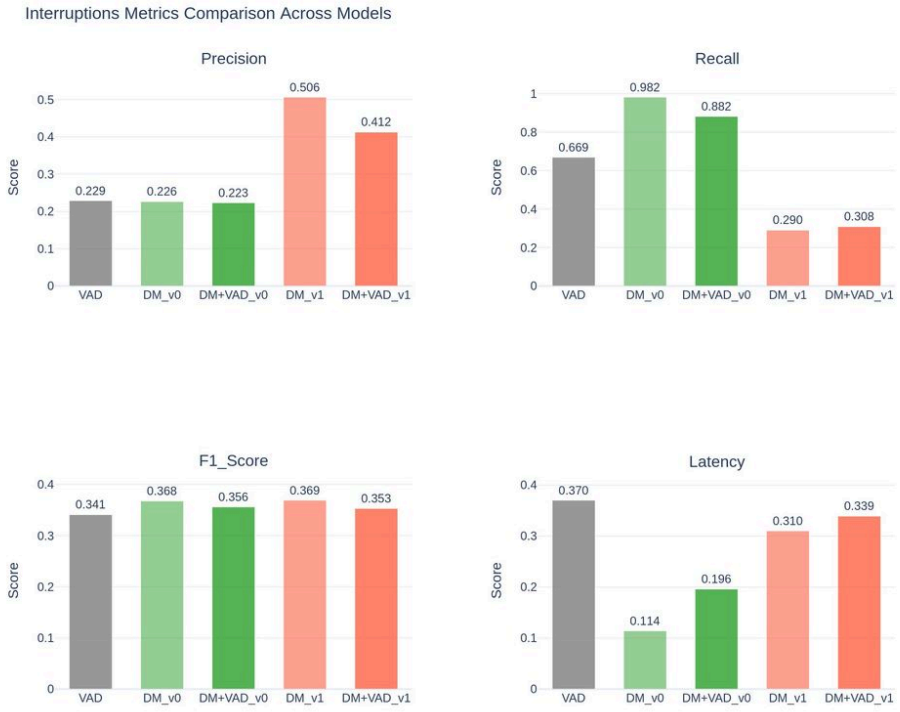
Model	Interruptions Recall	Turn End Recall	Pauses Recall	Backchannel Recall
DM_v1	0.29	0.92	0.83	0.94
DM+VAD_v1	0.31	0.92	0.83	0.90
DM_v0	0.98	0.89	0.62	0.25
DM+VAD_v0	0.88	0.89	0.63	0.31
VAD	0.67	0.04	0.94	0.50

Model	Interruptions Latency	Turn End Latency	Pauses Latency	Backchannel Latency
DM_v0	0.11	0.07	0.18	0.32
DM+VAD_v0	0.20	0.07	0.19	0.31
DM+VAD_v1	0.34	0.09	0.18	0.40
DM_v1	0.31	0.09	0.31	0.41
VAD	0.37	0.87	0.28	0.35

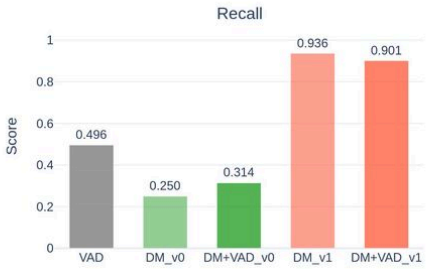
Results in Graphical Format

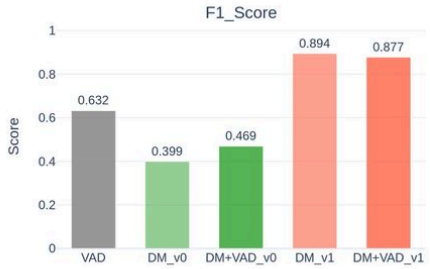


6.

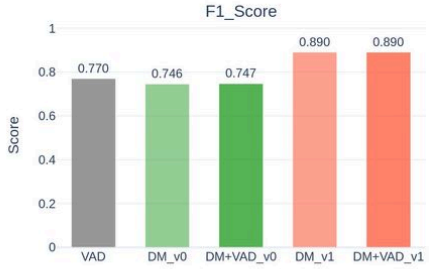


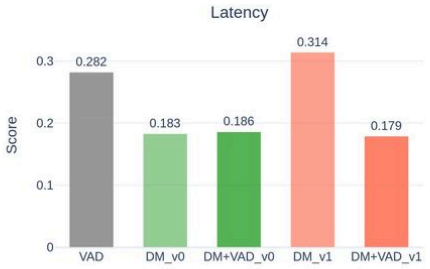
Backchannel Metrics Comparison Across Models





Pauses Metrics Comparison Across Models





Understanding the Data: