Stronger Together:
Where Data Meets Human Connection

Team G11

Anshu Mehta

Chen Ju (Maggie) Wang

Olivia Boe

Pin-Shiuan (Rachel) Liang

Ya Chin (Ruth) Hsu
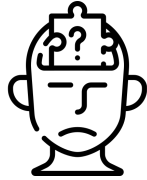

Big Brothers Big Sisters® TWIN CITIES

Goal

Foster longer, more successful matches by identifying **early signals** of relationship **health and risk**

But..

Human relationships are **complex**

The Key

Glimpse into the human brain with **comments** & **textual features**

So...   **Here is our solution...**

2

Match Health Dashboard

Our real-time monitoring dashboard enables timely, informed actions by answering:

1. **Who** requires proactive intervention?
2. **When** should we take action?

Match ID	Risk & Conflict	Engagement	Relationship Quality	Months Until Expected Closure	Predicted Closure Reason
A123	0.66	1.21	1.16	3	Success
B456	0.92	0.08	-0.49	10	Safety Concerns
C789	-0.18	2.73	0.55	8	Incompatibility

Building Block 1 - Categories Derived from Call Note Patterns

Extracted Relationship Categories

Match ID	Risk & Conflict	Engagement	Relationship Quality	Months Until Expected Closure	Predicted Closure Reason
A123	0.66	1.21	1.16	3	Success
B456	0.92	0.08	-0.49	10	Safety Concerns
C789	-0.18	2.73	0.55	8	Incompatibility

4

Key Categories for Model Training

Engagement

Tracks match support involvement

- Check-in cadence
- Calling Frequency

Risk & Conflict

Identifies relationship risks & conflicts between matches

- Unresponsiveness
- External Stressor
- Conflict Language

Relationship Quality

Measures depth & positivity of a relationship

- Engagement Consistency
- Mutual Sense of Growth
- Shared Interests

Leveraging LLM to Extract Features from Call Notes

LLM Call Note Summaries

[{"Big": "I'm doing well."}, {"Big": "I feel Little isn't enjoying time with me."}, {"Big": "We went on a walk to Sculpture Garden and got ice cream on Tuesday."}, {"Big": "I saw P/G give Little money for ice cream, but Little didn't offer to pay, so I did."}, {"Big": "We met twice, the other time was Saturday at an antique shop."}, {"Big": "Little isn't very talkative with me."}, {"Big": "Little goes on her phone in the car with headphones on, watches YouTube/movie."}, {"Big": "Little doesn't start conversation, is short with me."}, {"Big": "In the car, I said to Little, 'I get the feeling you don't really want to see me, but you are forced to join our meets.' Little didn't respond, was quiet."}, {"Big": "Yes, it would be helpful to have MEC there for a '2 month check-in.'"}, {"Big": "P/G seems super busy with their move, got a new job, etc."}, {"Big": "Little asks for food on the way home, said there wasn't dinner at home."}, {"Big": "Little had smiled about something on her show on her phone and shared it with me so I could laugh too. That made my day."}, {"Big": "I'm feeling sad as I want Little to enjoy our meets."}, {"Big": "I am not giving up, I want to continue to try. I am in this for the long haul."}]

Unresponsiveness

External Stressor
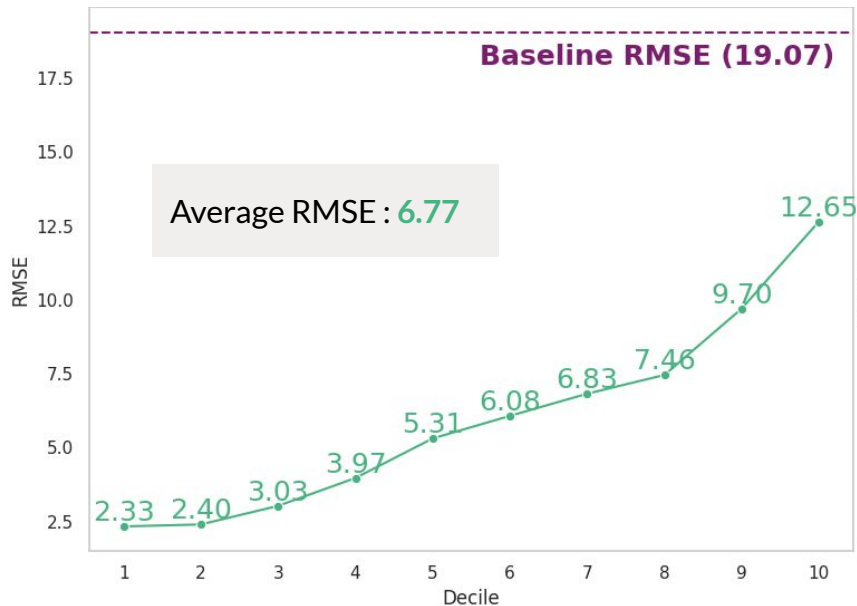
Engagement Consistency

Future Oriented

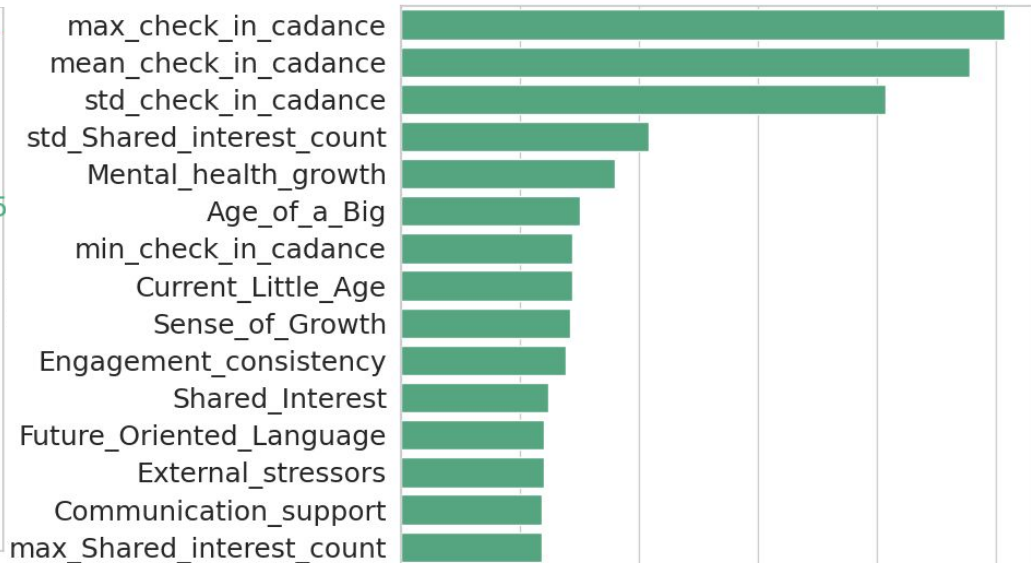# Building Block 2 - Match Length Prediction

Match Length Prediction

Match ID	Risk & Conflict	Engagement	Relationship Quality	Months Until Expected Closure	Predicted Closure Reason
A123	0.66	1.21	1.16	3	Success
B456	0.92	0.08	-0.49	10	Safety Concerns
C789	-0.18	2.73	0.55	8	Incompatibility

LightGBM Outperforms Other Models in "Match Length" Prediction

Consistently outperforms baseline

Baseline RMSE (19.07)

Average RMSE : 6.77

Decile values: 2.33, 2.40, 3.03, 3.97, 5.31, 6.08, 6.83, 7.46, 9.70, 12.65

Feature Importance (Top 15)

- max_check_in_cadance
- mean_check_in_cadance
- std_check_in_cadance
- std_Shared_interest_count
- Mental_health_growth
- Age_of_a_Big
- min_check_in_cadance
- Current_Little_Age
- Sense_of_Growth
- Engagement_consistency
- Shared_Interest
- Future_Oriented_Language
- External_stressors
- Communication_support
- max_Shared_interest_count

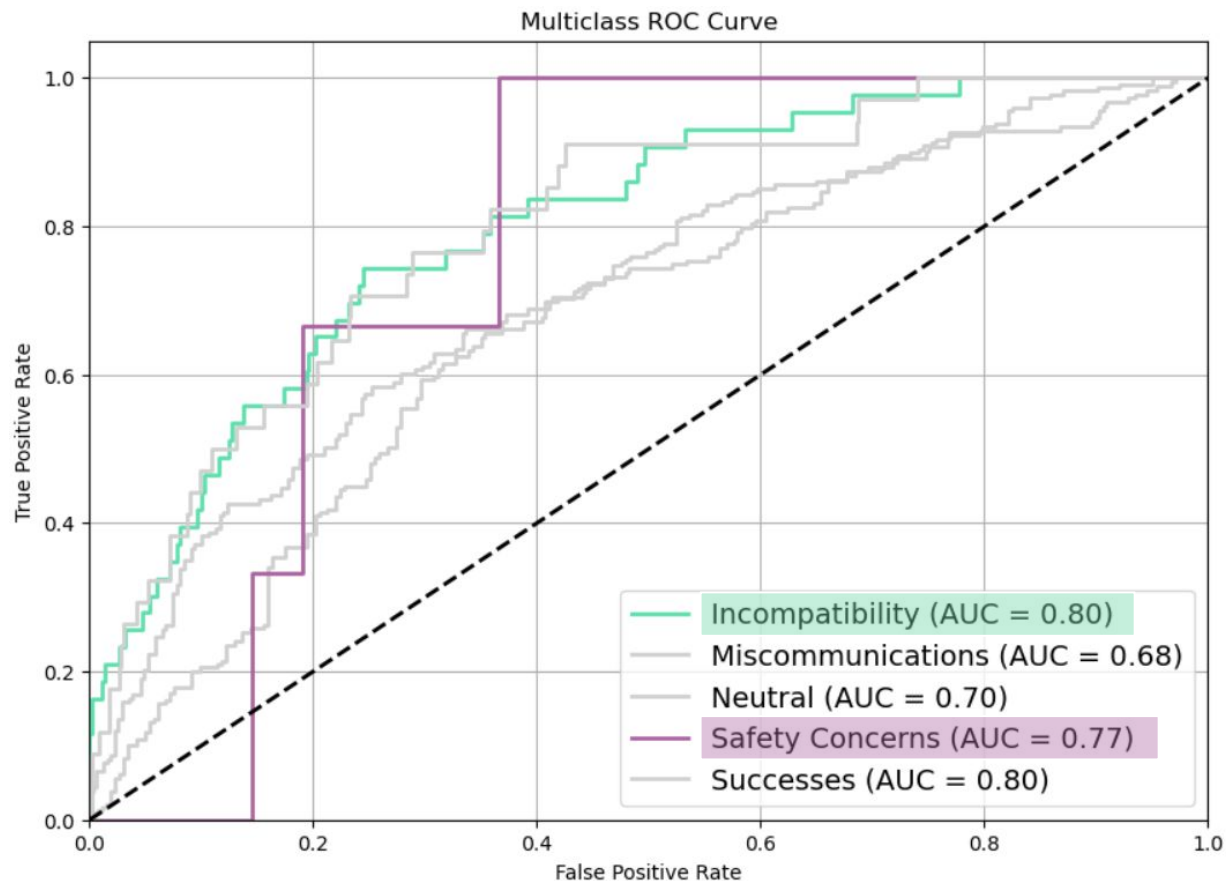*The number of data points within each decile are approximately equal

Building Block 3 - Closure Reason Prediction

Closure Reason Prediction

Match ID	Risk & Conflict	Engagement	Relationship Quality	Months Until Expected Closure	Predicted Closure Reason
A123	0.66	1.21	1.16	3	Success
B456	0.92	0.08	-0.49	10	Safety Concerns
C789	-0.18	2.73	0.55	8	Incompatibility

Strong Predictive Power on Closure Reasons

Multiclass ROC Curve

Incompatibility (AUC = 0.80)
Miscommunications (AUC = 0.68)
Neutral (AUC = 0.70)
Safety Concerns (AUC = 0.77)
Successes (AUC = 0.80)

Random Forest

Features Driving Incompatibility:

1) Age of a Big
2) Current Little Age
3) Conflict Language
4) Sentiment
5) LDA Formal Tone Topic

Features Driving Safety Concerns:

1) LDA Safety Topic
2) Age of a Big
3) Sentiment
4) Check-in frequency
5) Check-in Cadence
6) Conflict Flags

10

Understanding How Key Features Affect Match Length

- Relationship quality and engagement are associated with longer relationships
- Demographic and risk & conflict features have higher risk of early closure

	Engagement	**Risk & Conflict**	**Relationship Quality**	**Demographic**
Lower risk of early closure	Check-in Cadence Call Frequency		Engagement_consistency Future_Oriented_Language Communication_Support Relationship_quality Shared_Interest LDA Formal Tone Topic	
Higher risk of early closure		Unresponsiveness Conflict Language External Stressors LDA Friction Topic		Age of a Little Age of a Big

*Based on results from Cox Proportional Hazards Model. All features shown above are statistically significant (p-value < 0.05)

Recommendations

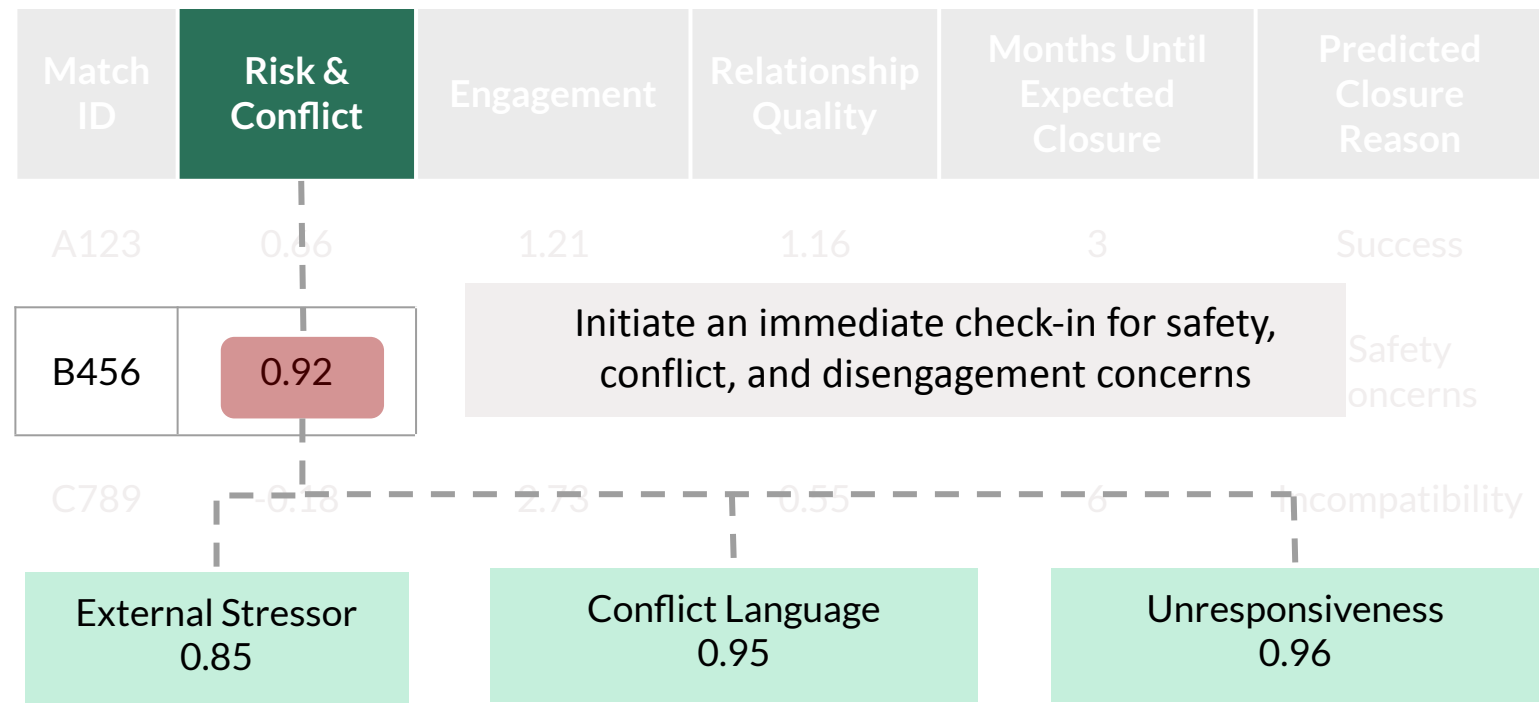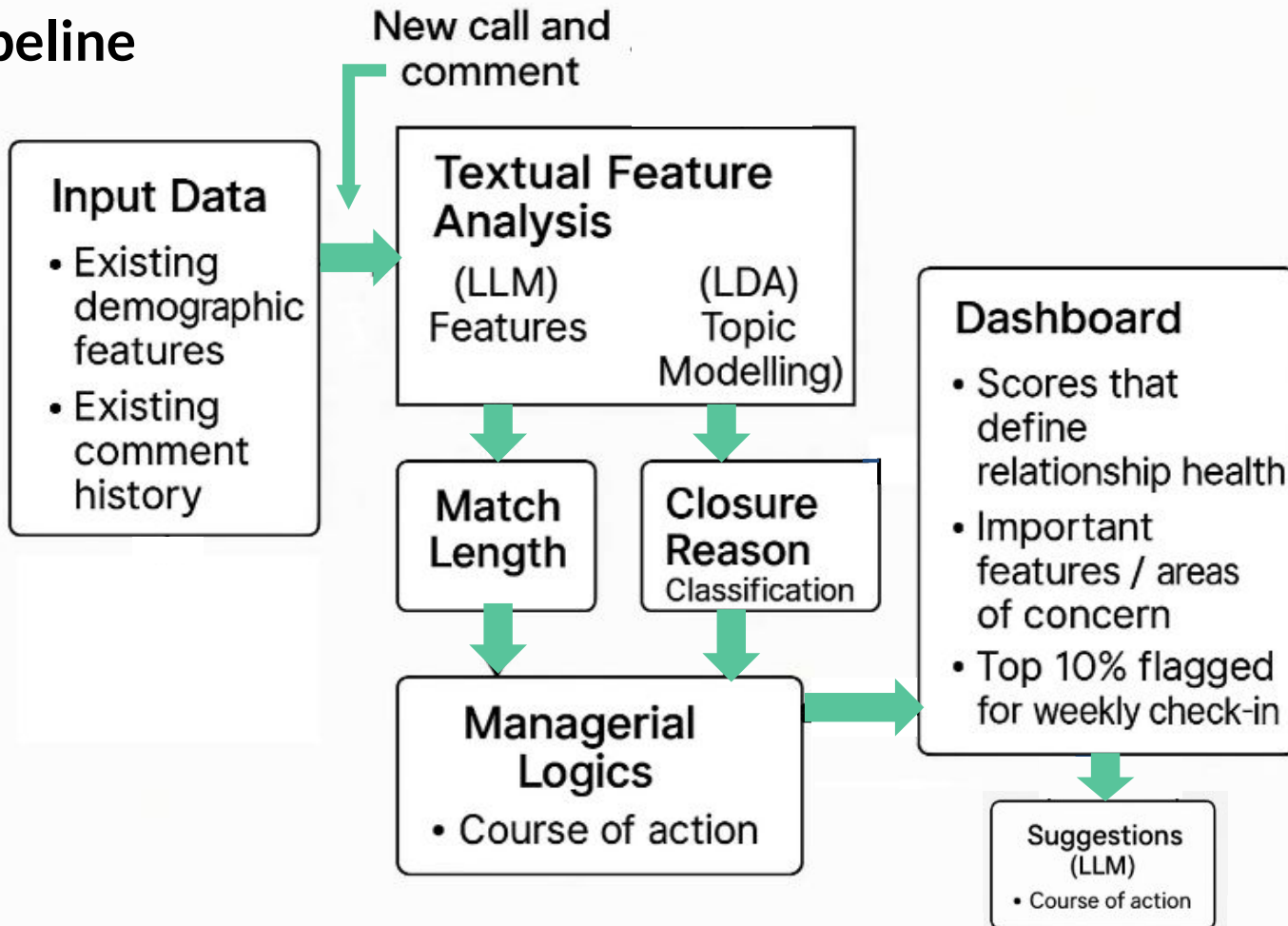| Goal	Initiatives	Actions
Timing Intervention	Closure Intervention Strategy	Implement a monitoring system: **Real-Time Dashboard** using match health score for timely interventions

Match Health Dashboard in Real-Time

An example of early intervention walk through - **WHO & WHEN** needs intervention?

Match ID	Risk & Conflict	Engagement	Relationship Quality	Months Until Expected Closure	Predicted Closure Reason
A123	0.66	1.21	1.16	3	Success
B456	0.92	0.08	-0.49	2	Safety Concerns
C789	-0.18	2.73	0.55	6	Incompatibility

Immediate Concern

Match Health Dashboard in Real-Time

An example of early intervention - **HOW** to design a data driven personalized intervention?

Match ID	Risk & Conflict	Engagement	Relationship Quality	Months Until Expected Closure	Predicted Closure Reason
A123	0.06	1.21	1.16	3	Success
B456	0.92				Safety concerns
C789	-0.18	2.73	0.55	6	Incompatibility

Initiate an immediate check-in for safety, conflict, and disengagement concerns

External Stressor
0.85

Conflict Language
0.95

Unresponsiveness
0.96

Recommendations

Goal	Initiatives	Actions
Timing Intervention	Closure Intervention Strategy	Implement a monitoring system: **Real-Time Dashboard** using match health score for timely interventions
Improve Monitoring Metrics	Data Integrity	Consolidate all data sources into single pipeline for seamless integration

Data Pipeline

Recommendations

Goal	Initiatives	Actions

Timing Intervention

Closure Intervention Strategy

Implement a monitoring system: **Real-Time Dashboard** using match health score for timely interventions

Improve Monitoring Metrics

Data Integrity

Consolidate all data sources into single pipeline for seamless integration

Support Call Framework

Create survey questions for future support calls

Risk and Conflict :
1) How well do you get along well with your Big/Little?
2) Have you noticed any changes in how your Little expresses themselves or handles situations lately?

Relationship Quality:
1) How do you enjoy your time with your big?
2) Do plans with big/little include things you enjoy?

Engagement:
1) How have you supported this relationship ?

17

Thanks for your time!

Feel free to reach out if you have any questions or thoughts

Anshu Mehta

mehta395@umn.edu

Chen Ju (Maggie) Wang

wan02518@umn.edu

Pin-Shiuan (Rachel) Liang

lian0184@umn.edu

Olivia Boe

boe00024@umn.edu
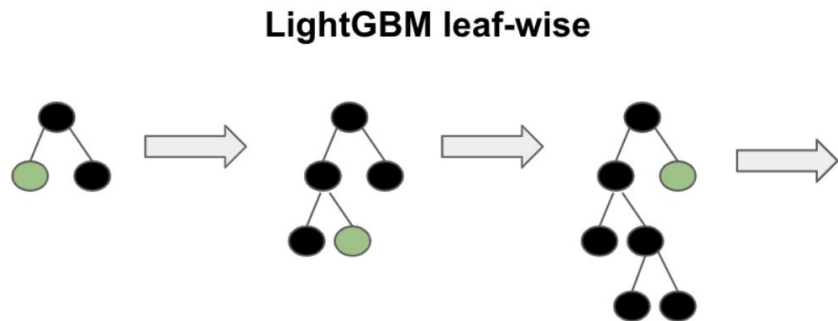
Ya Chin (Ruth) Hsu

hsu00169@umn.edu

Thank You!

Appendix

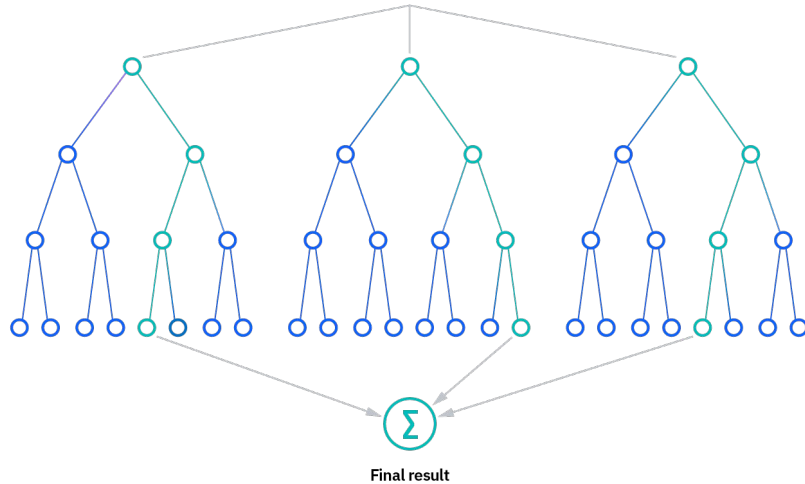# Model Strategies for Match Length Prediction - LightGBM Model

LightGBM leaf-wise

LightGBM Model
Faster training speed and better efficiency
Compatibility with Large Datasets
Capability of processing high-dimensional data
Generate feature importance automatically

Model Strategies for Closure Reason Prediction - Random Forest

Picture Reference: IBM

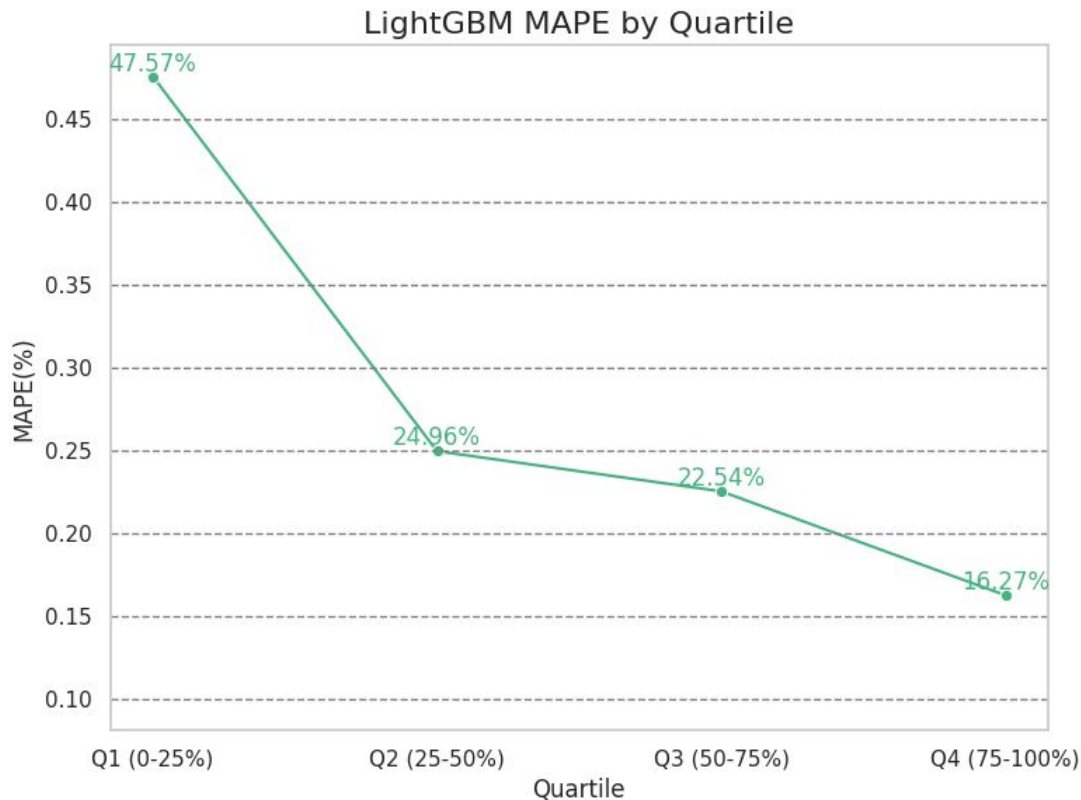| Random Forest |
|:---:|
| High Accuracy |
| Handles Missing Values & Outliers |
| Feature Importance & Interpretability |
| Works Well with Large Datasets |

Match Length Prediction - LightGBM Model MAPE

MAPE indicates that model shows best performance on longer match length.

LightGBM MAPE by Quartile

Category	Feature	Definition
Engagement	Check-in cadence	days between calls
	Check-in frequency	count of calls
Risk & Conflict	Child safety concern	drug or substance abuse by child or parent, violence, assault, or threats, rehab, hospitalization, or trauma events
	Unresponsiveness	either candidate is difficult to get a hold of, or doesn't show to scheduled events
	External stressors	family matters, school obligations, mental health, relocations, etc.
	Conflict Language	disagreement, tension, arguments, or interpersonal friction
Relationship Quality	Engagement Consistency	how often the match meets and whether they plan or look forward to future interactions
	Mental health growth	support for mental health and signs of personal/emotional development
	Mutual sense of growth	positive growth or behavior mentioned.
	Relationship quality	emotional closeness, shared activities, and a positive dynamic
	Future oriented	mention future plans, goals, long term investments
	Shared Interest	any hobbies or interests were mentioned and the speaker expresses positive emotion
	Communication support	communication consistently and empathetically