50,000,000

50,000,000

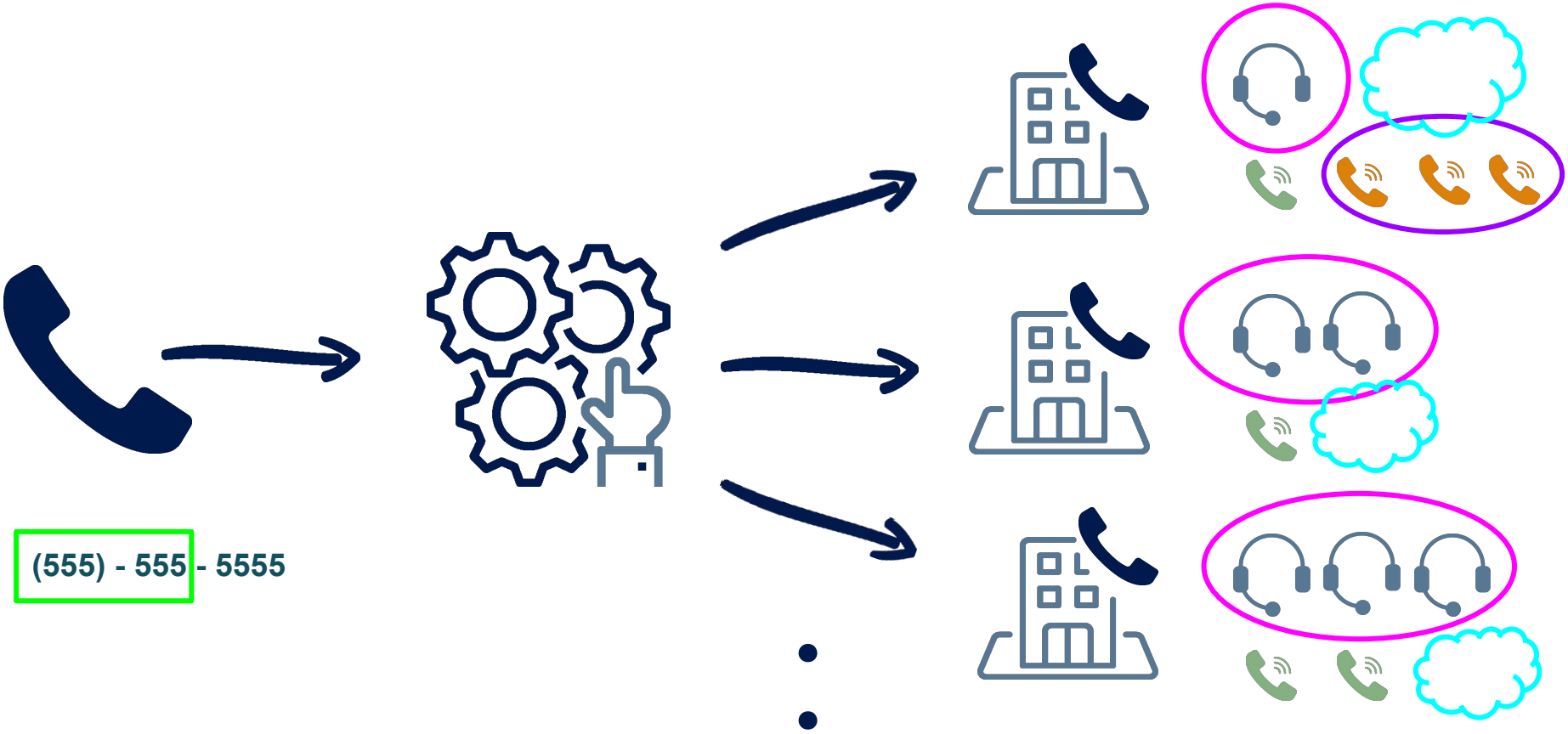
988
SUICIDE
& CRISIS
LIFELINE

2,000,000
calls

300,000
calls are abandoned

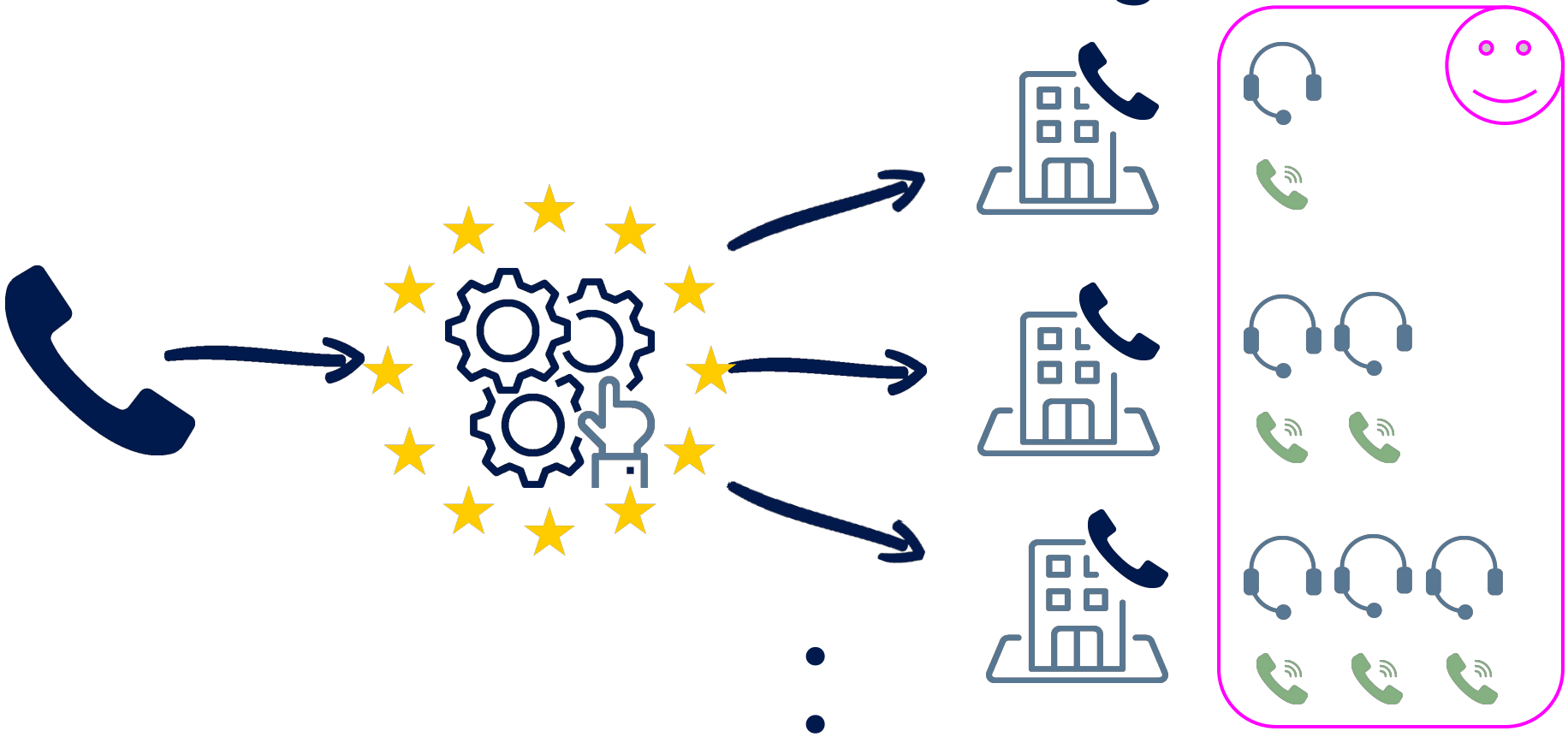
V!brant
Emotional Health

988
SUICIDE
& CRISIS
LIFELINE


(555) - 555 - 5555

# What we have: historical routing data

# What we want: a better routing table



# Data Science For Social Good

Summer Fellowship

Carnegie Mellon University

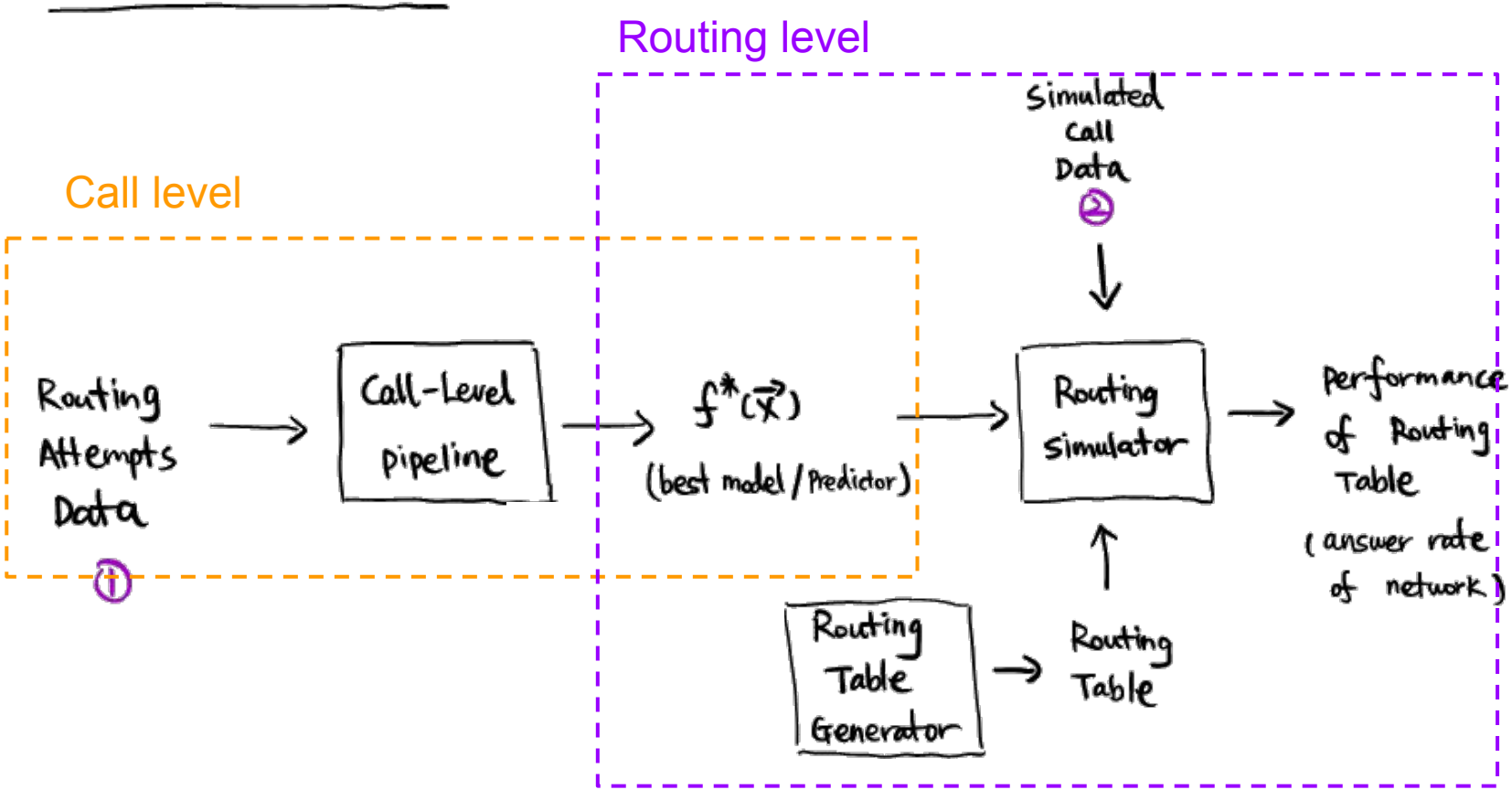
**Tejumade**
Afonja

**Paula**
Subías
Beltrán

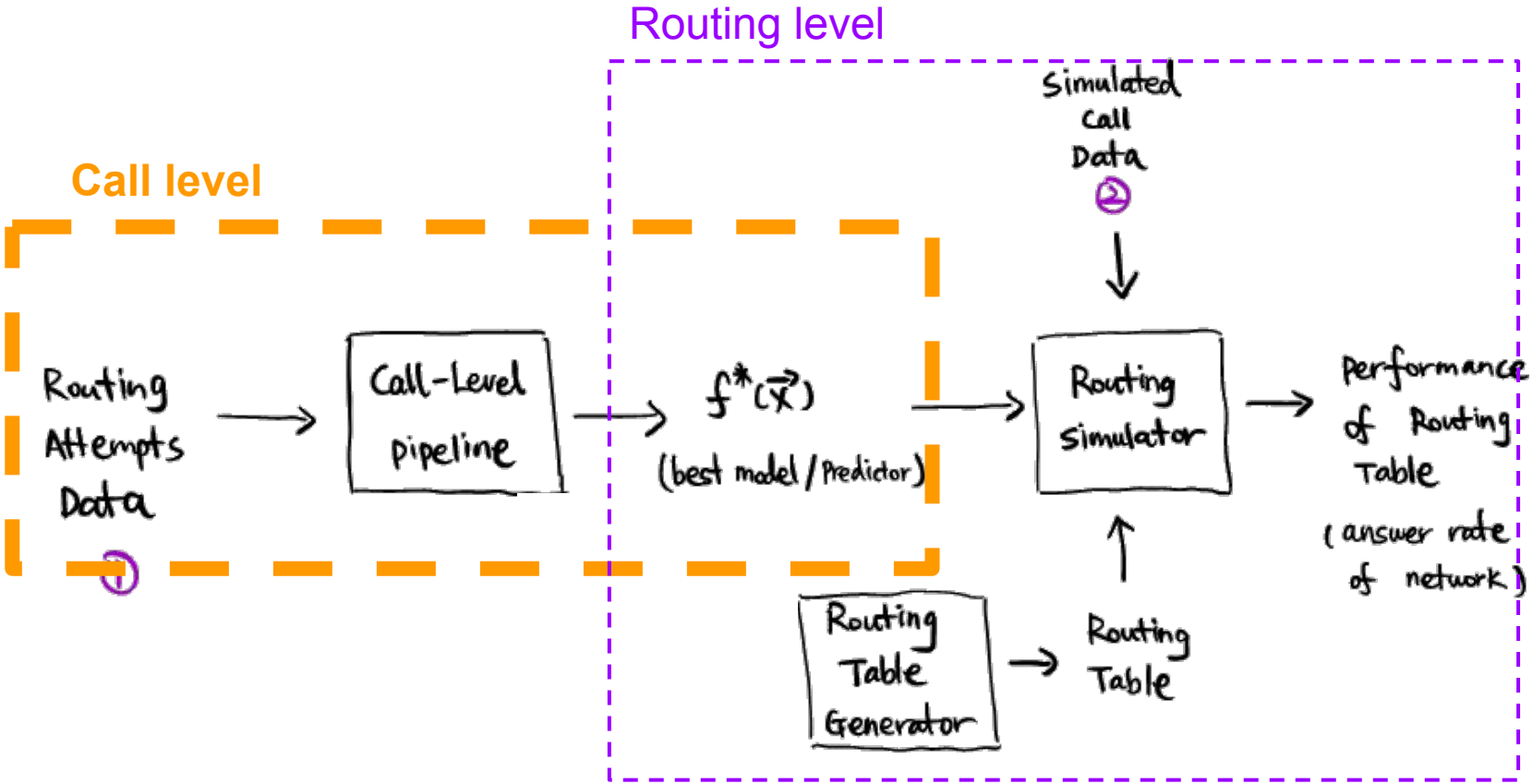
**Charles**
Cui

**Irene**
Tang

# Overall Flow

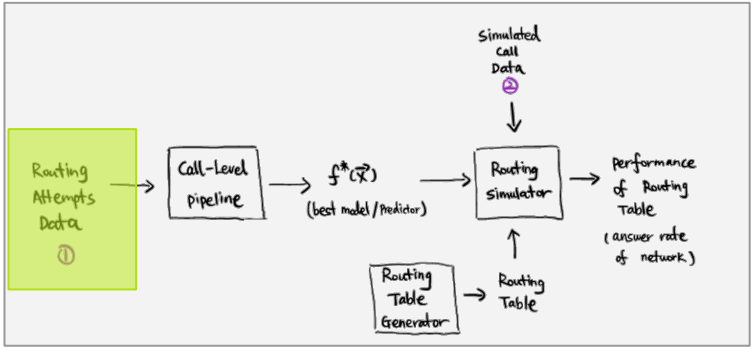


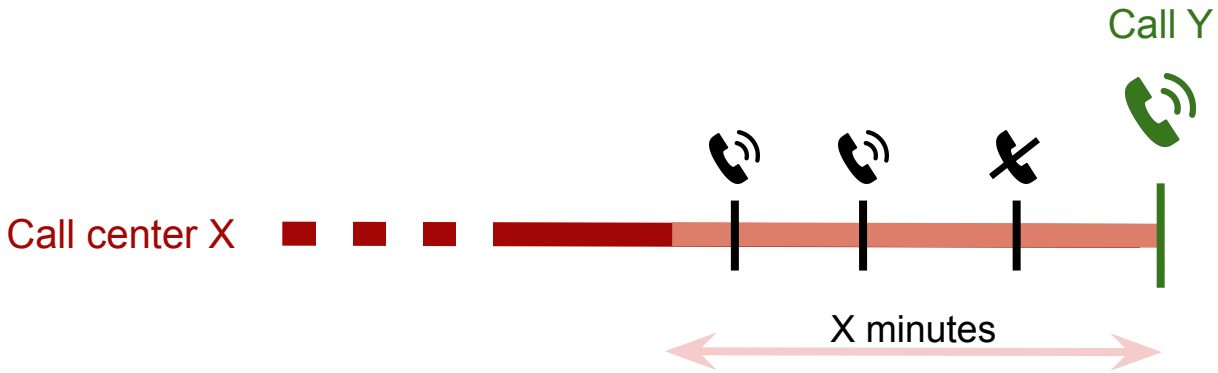
# Overall Flow

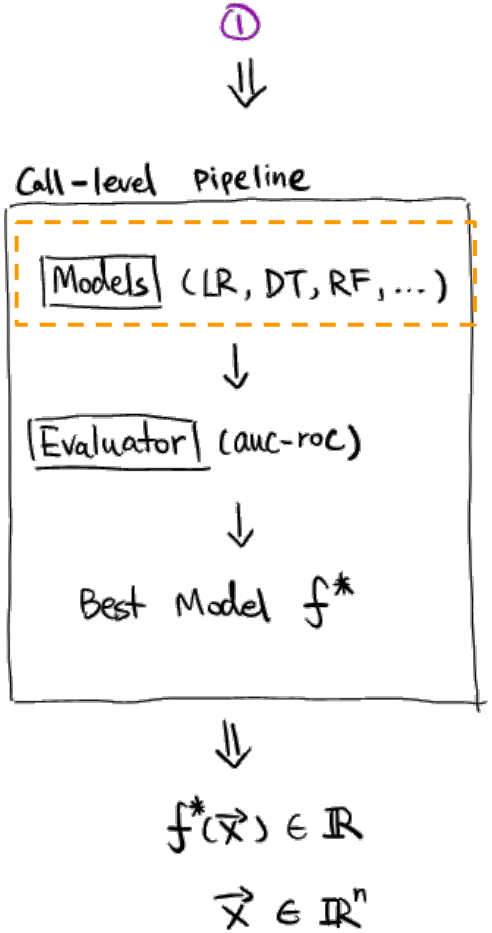


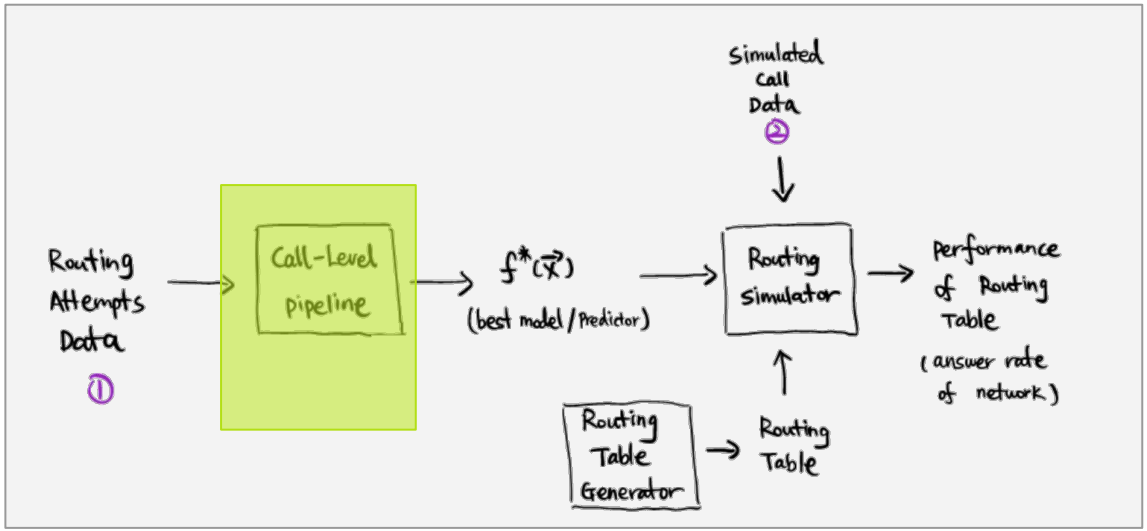
## Features

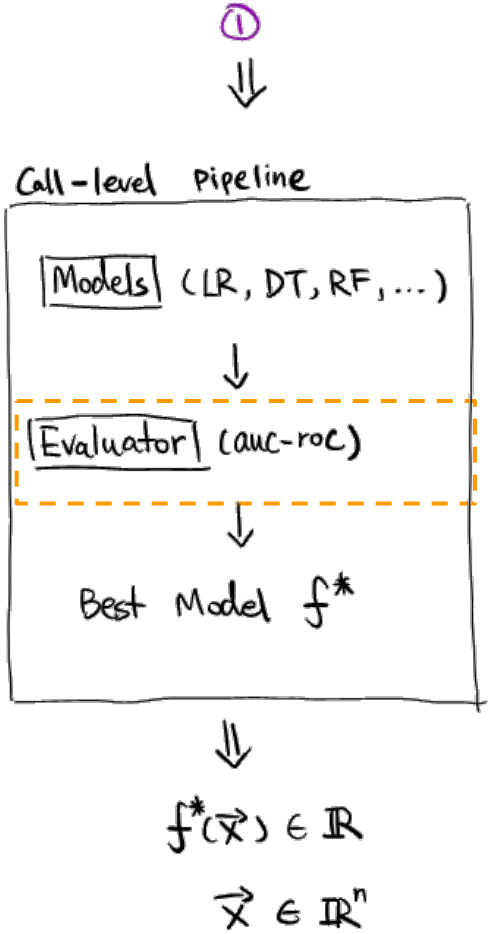
- number of past calls at center
- number of past calls flowed out center
- number of past calls abandoned at center
- …

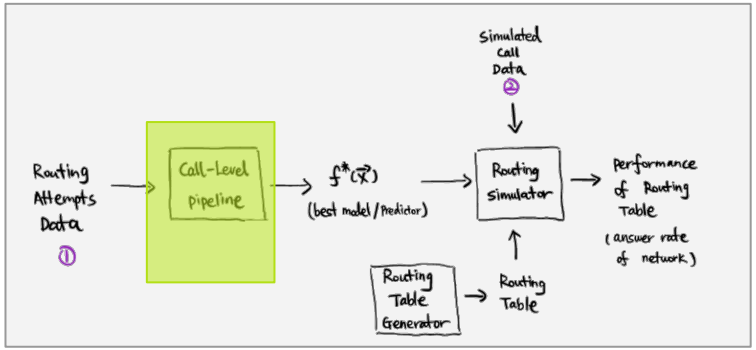


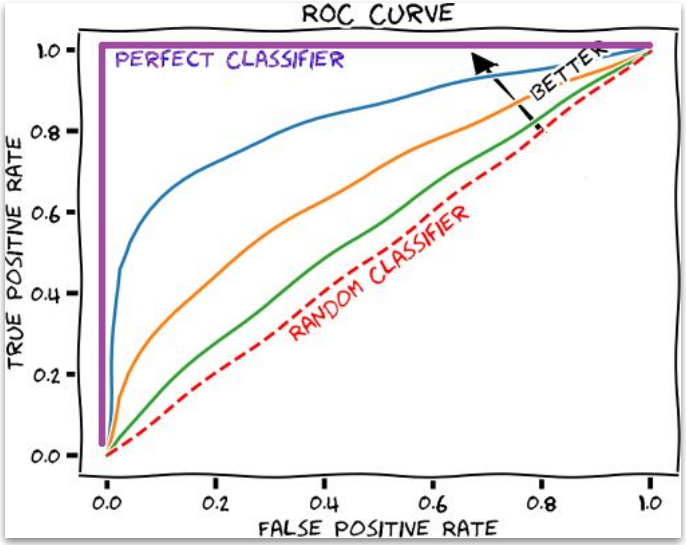



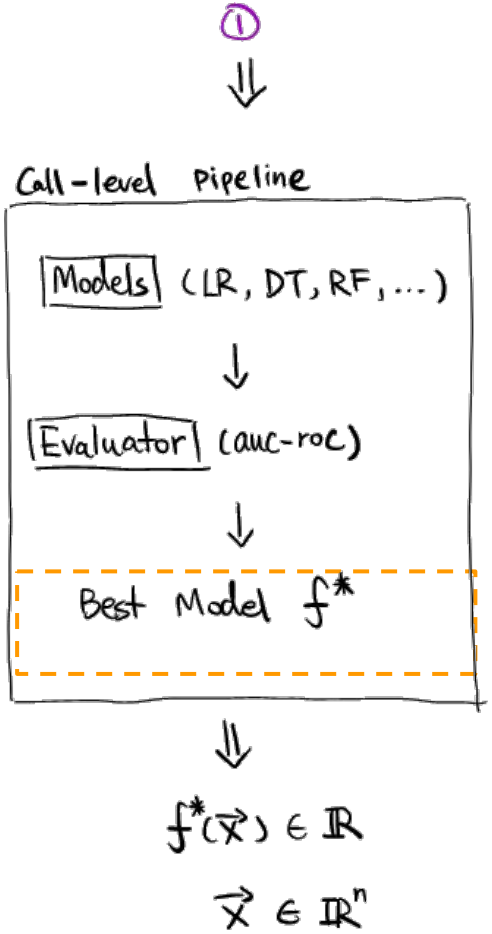

①
Call-level Pipeline
Models (LR, DT, RF, ...)
Evaluator (auc-roc)
Best Model $f^*$
$f^*(\vec{x}) \in \mathbb{R}$
$\vec{x} \in \mathbb{R}^n$

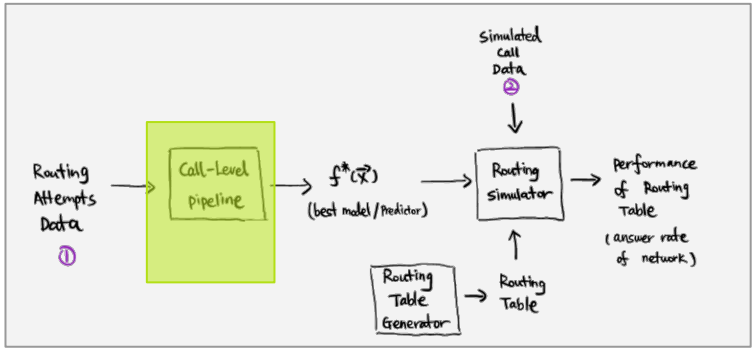

Simulated Call Data ②
Routing Attempts Data ①
Call-Level pipeline
$f^*(\vec{x})$
(best model / Predictor)
Routing Simulator
Performance of Routing Table
(answer rate of network)
Routing Table Generator
Routing Table

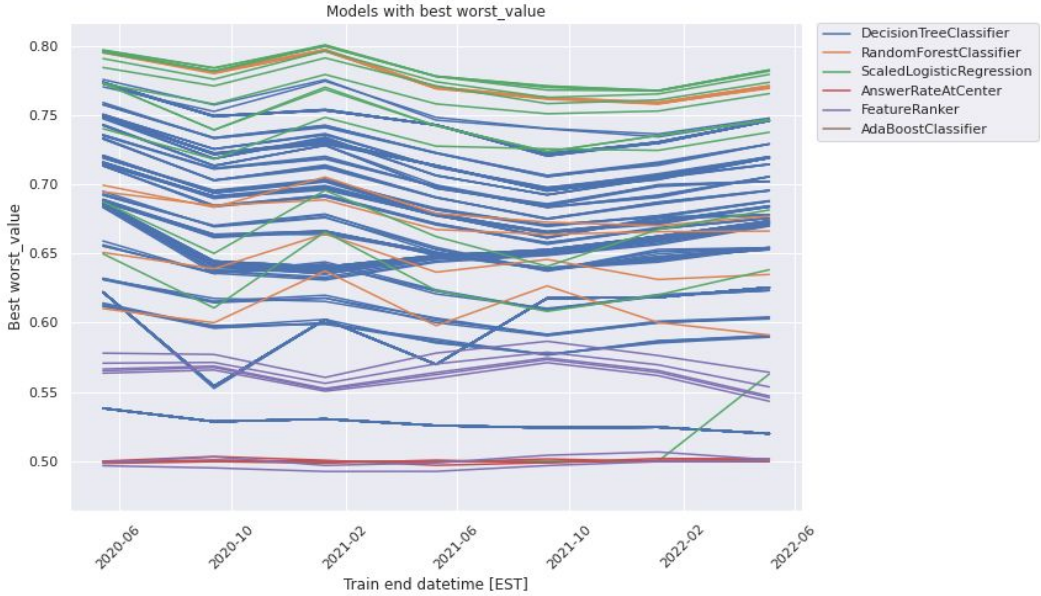

①
Call-level Pipeline
Models (LR, DT, RF, ...)
Evaluator (auc-roc)
Best Model $f^*$
$f^*(\vec{x}) \in \mathbb{R}$
$\vec{x} \in \mathbb{R}^n$


Simulated Call Data
②
Routing Attempts Data
①
Call-Level Pipeline
$f^*(\vec{x})$
(best model / predictor)
Routing Simulator
Performance of Routing Table
(answer rate of network)
Routing Table Generator
Routing Table

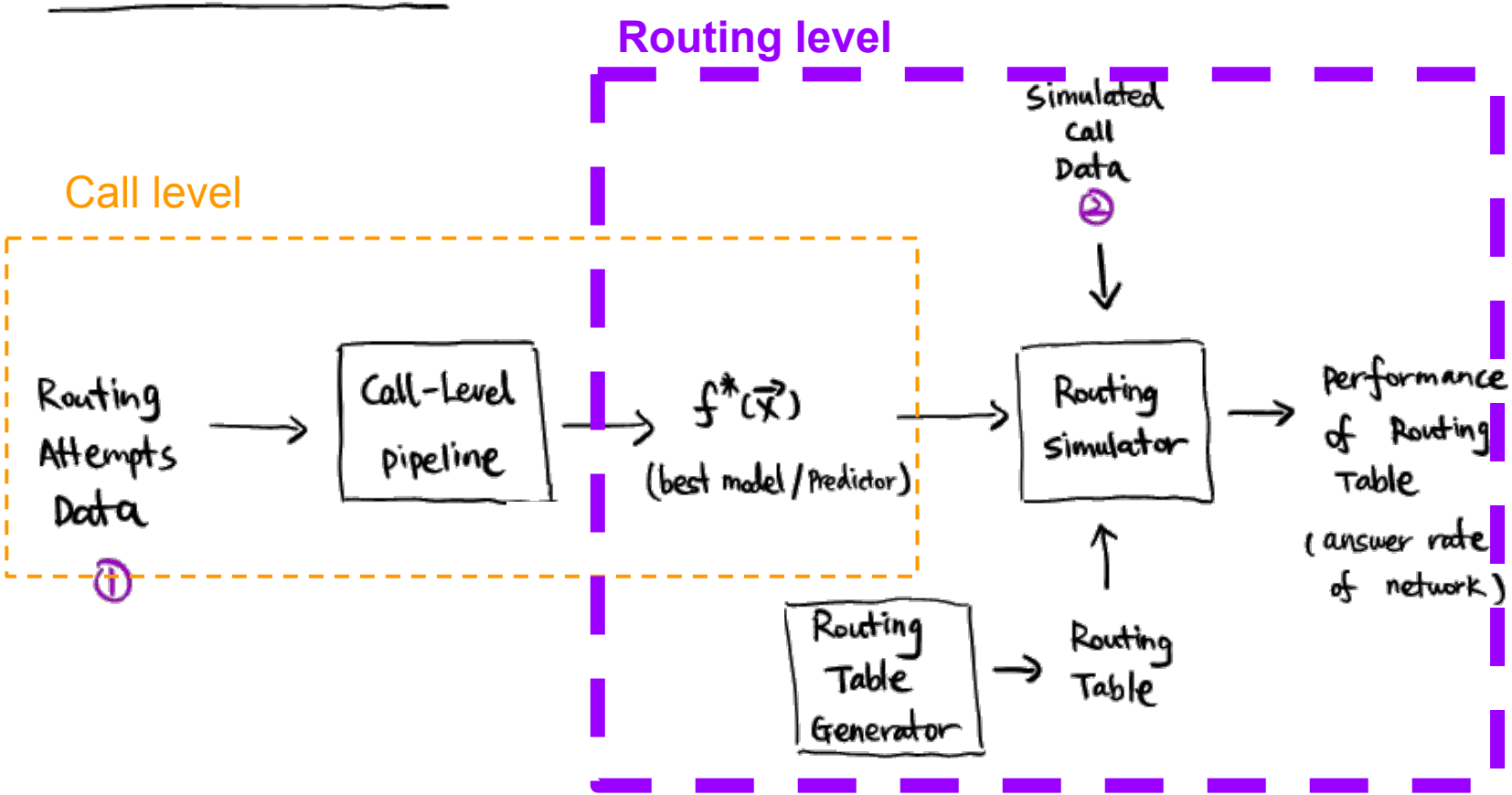

ROC CURVE
PERFECT CLASSIFIER
BETTER
RANDOM CLASSIFIER
TRUE POSITIVE RATE
FALSE POSITIVE RATE
0.0
0.2
0.4
0.6
0.8
1.0


Call-level Pipeline
Models (LR, DT, RF, ...)
Evaluator (auc-roc)
Best Model $f^*$
$f^*(\vec{x}) \in \mathbb{R}$
$\vec{x} \in \mathbb{R}^n$

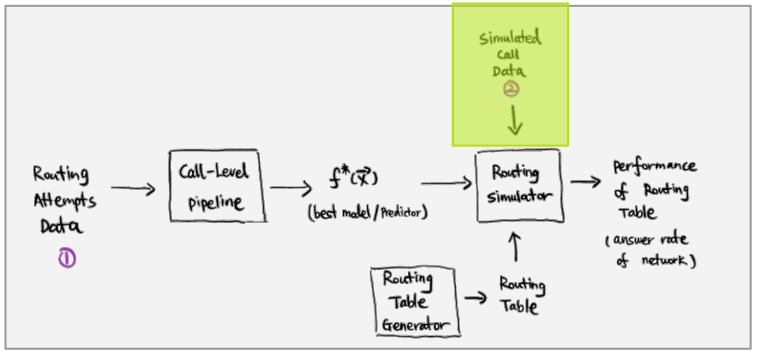

Routing Attempts Data
Call-Level Pipeline
$f^*(\vec{x})$
(best model / predictor)
Simulated Call Data
Routing Simulator
Performance of Routing Table
(answer rate of network)
Routing Table Generator
Routing Table

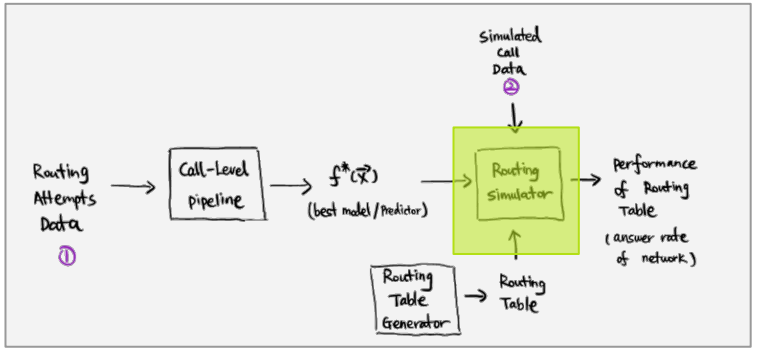

Models with best worst_value
Best worst_value
0.80
0.75
0.70
0.65
0.60
0.55
0.50
2020-06
2020-10
2021-02
2021-06
2021-10
2022-02
2022-06
Train end datetime [EST]
DecisionTreeClassifier
RandomForestClassifier
ScaledLogisticRegression
AnswerRateAtCenter
FeatureRanker
AdaBoostClassifier

# Overall Flow



| call_key | caller_npanxx | initiated_datetime_est |
|---|---|---|
| e3225-20220502083313992-426 | 607,582 | 2022-05-02 08:33:13.000 |
| e3225-20220502121556977-316 | 864,279 | 2022-05-02 12:15:56.000 |
| e3225-20220502130758752-550 | 813,972 | 2022-05-02 13:07:58.000 |
| e3225-20220502131509392-582 | 518,434 | 2022-05-02 13:15:09.000 |
| e3225-20220502153336417-160 | 929,252 | 2022-05-02 15:33:36.000 |
| e3225-20220502170350972-520 | 434,939 | 2022-05-02 17:03:50.000 |
| e3225-20220502174802577-704 | 910,594 | 2022-05-02 17:48:02.000 |
| e3225-20220502194921528-026 | 201,870 | 2022-05-02 19:49:21.000 |
| e3225-20220502200641343-068 | 330,309 | 2022-05-02 20:06:41.000 |
| e3225-20220502224047123-298 | 479,629 | 2022-05-02 22:40:47.000 |
| e3225-20220503073615623-508 | 813,215 | 2022-05-03 07:36:15.000 |
| e3225-20220503074952522-524 | 312,690 | 2022-05-03 07:49:52.000 |
| e3225-20220503085613228-686 | 916,835 | 2022-05-03 08:56:13.000 |
| e3225-20220503115611864-282 | 646,626 | 2022-05-03 11:56:11.000 |
| e3225-20220503130724863-562 | 319,217 | 2022-05-03 13:07:24.000 |






## Routing Simulator

② ⇒

$f^*$ ⇒

Routing table ⇒

for next call i in ②:

- → Get the center $S_j$ it gets routed to using routing table
- → Compute the features for (call i, $S_j$)
- → Compute the probability $P_{ij}$ that i gets picked up by $S_j$ using $f^*$ and features
- → Flip a $P_{ij}$-biased coin
  - → If heads: call i gets answered by $S_j$
  - → If tails: call i is NOT picked up by $S_j$
    - → Add 3 min to call i and insert back to ②

# Results

An improved routing table that has a **1%** higher predicted overall answer rate than the 988's current routing table.

estimated
+20,000
callers helped

**Charles** Cui

**Irene** Tang

**Tejumade** Afonja

**Paula** Subías Beltrán

# Thank you!