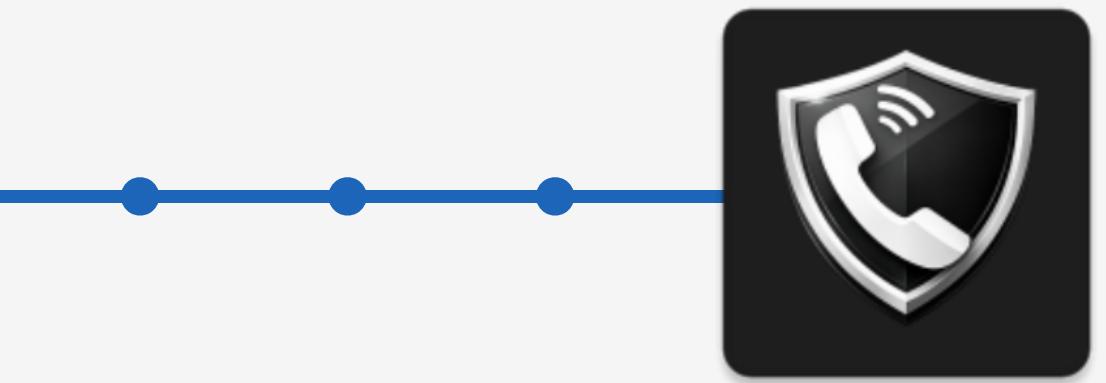


다계층 AI 기반
실시간 보이스피싱 탐지



TEAM

Dropout(0.25)

목차 안내

TABLE OF CONTENTS



01 프로젝트 개요

- 보이스피싱이란?
- 피해 현황 분석

02 서비스 소개

- 전략 체계
- 기존 유사 서비스와 차별점

03 개발

- 분석
- 아키텍처
- 모델링 및 성능 지표
- 파이프라인
- 서버 & 데이터베이스

04 UI / UX

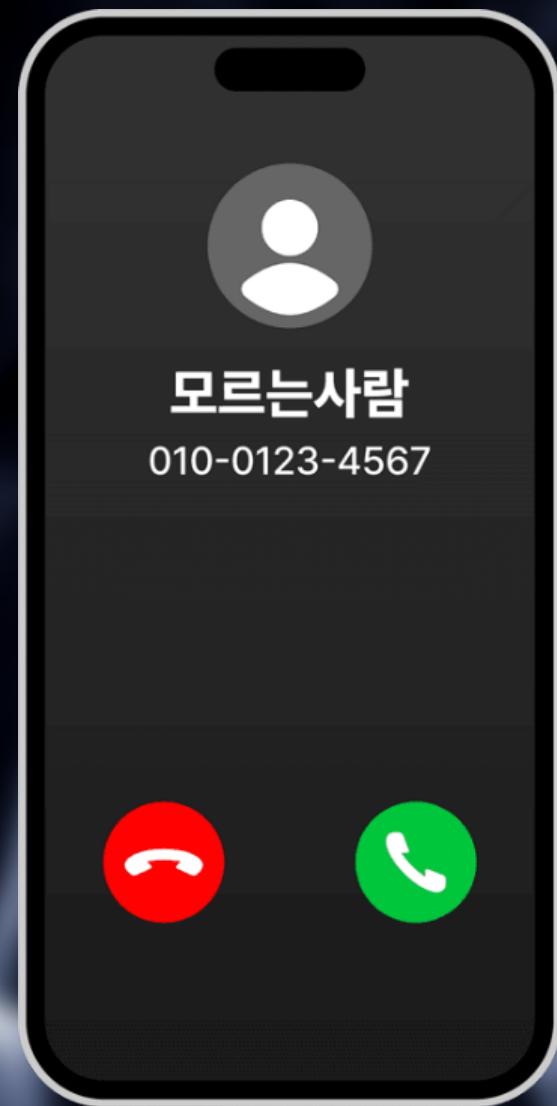
- 실제 서비스 화면

05 기대효과

- 보완 필요점
- 기대효과

06 마무리

- 시연
- 질의응답

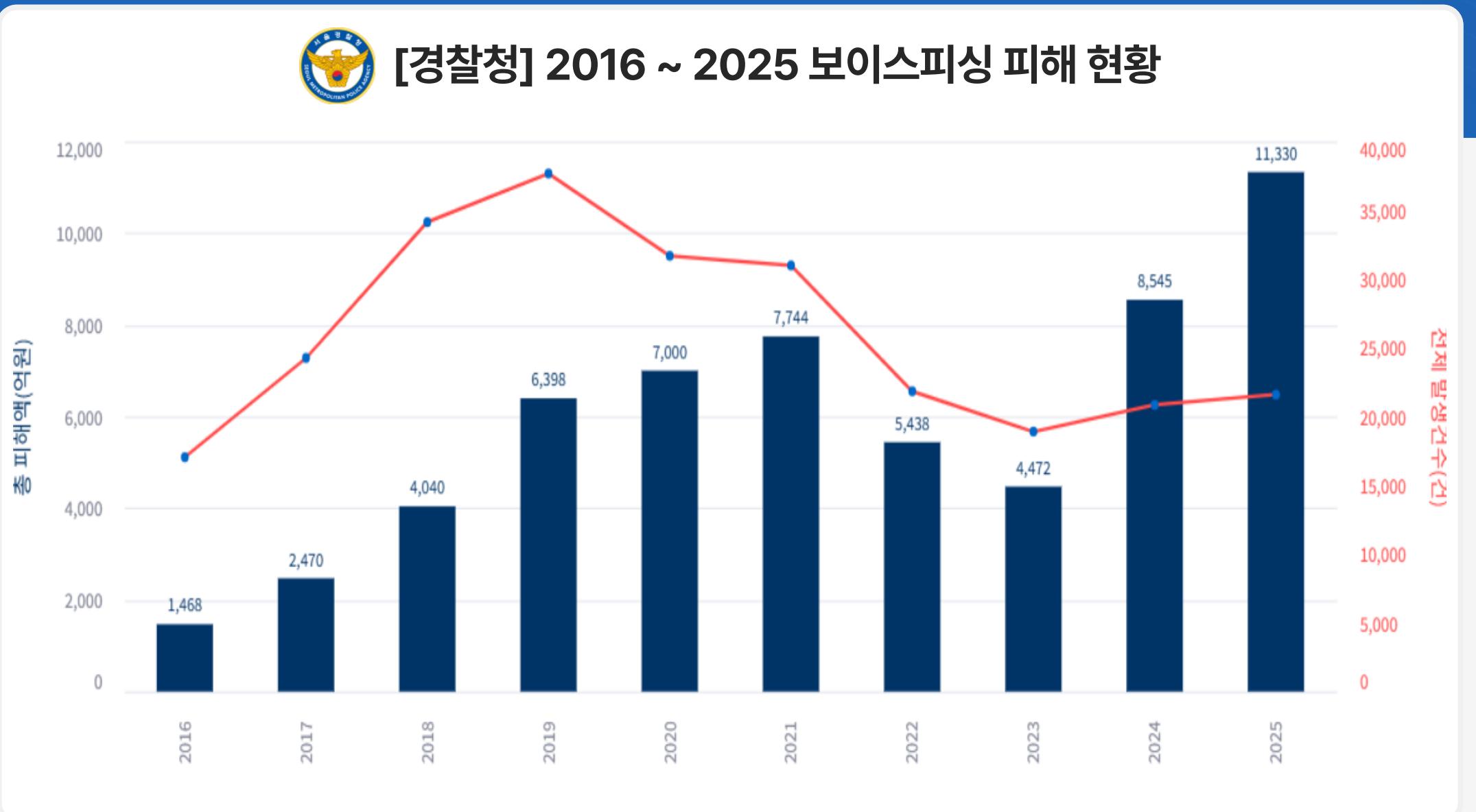


프로젝트 개요 보이스피싱이란?

보이스피싱(Voice Phishing) = 음성(Voice) + 개인정보(Private Data) + 낚시(Fishing)
 전화, 문자, 메신저 등 전기통신수단을 이용해 피해자를 속여 재산상의 이득을 취하는 사기 범죄



[경찰청] 2016 ~ 2025 보이스피싱 피해 현황



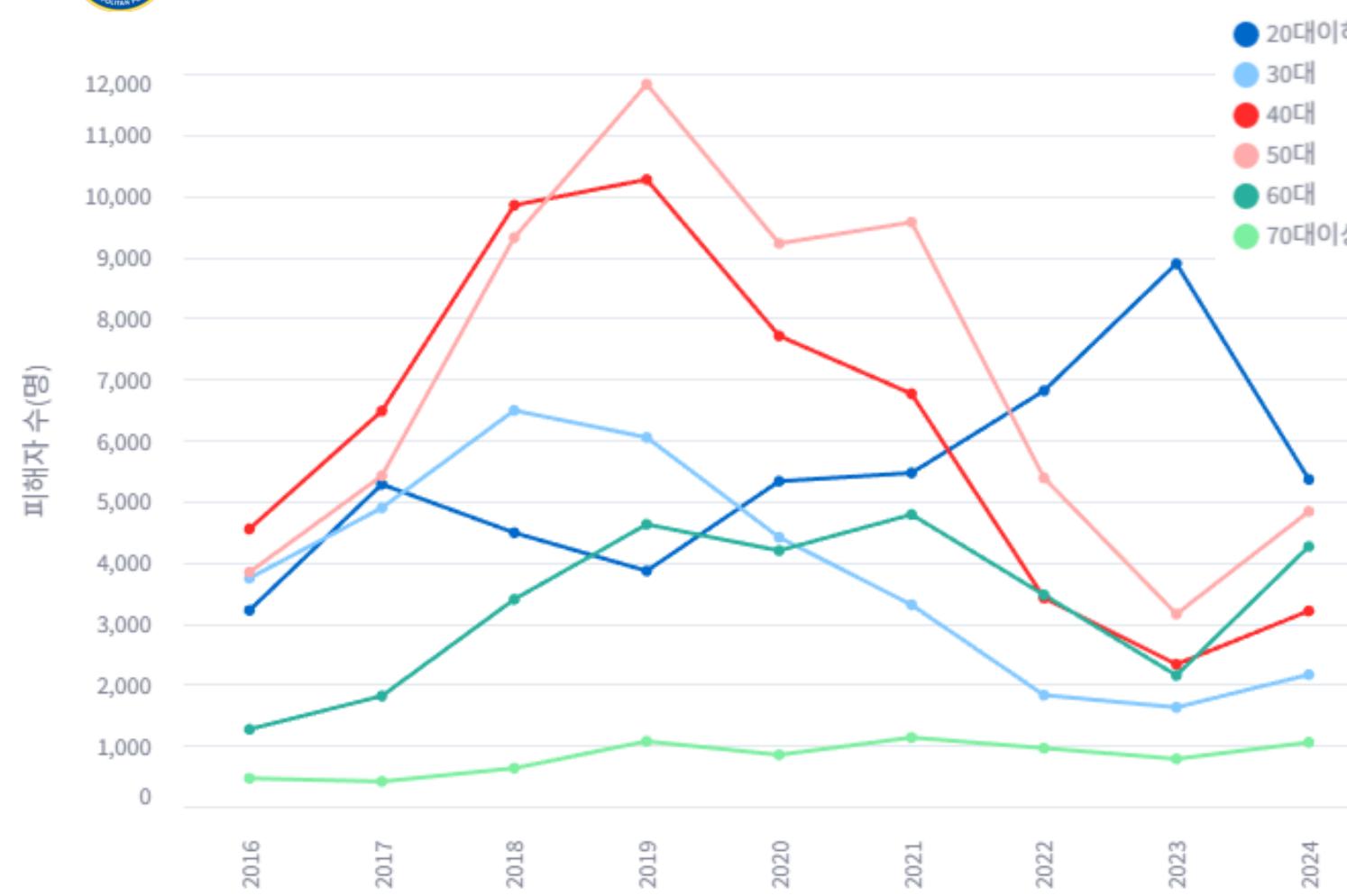
건당 피해액 추이



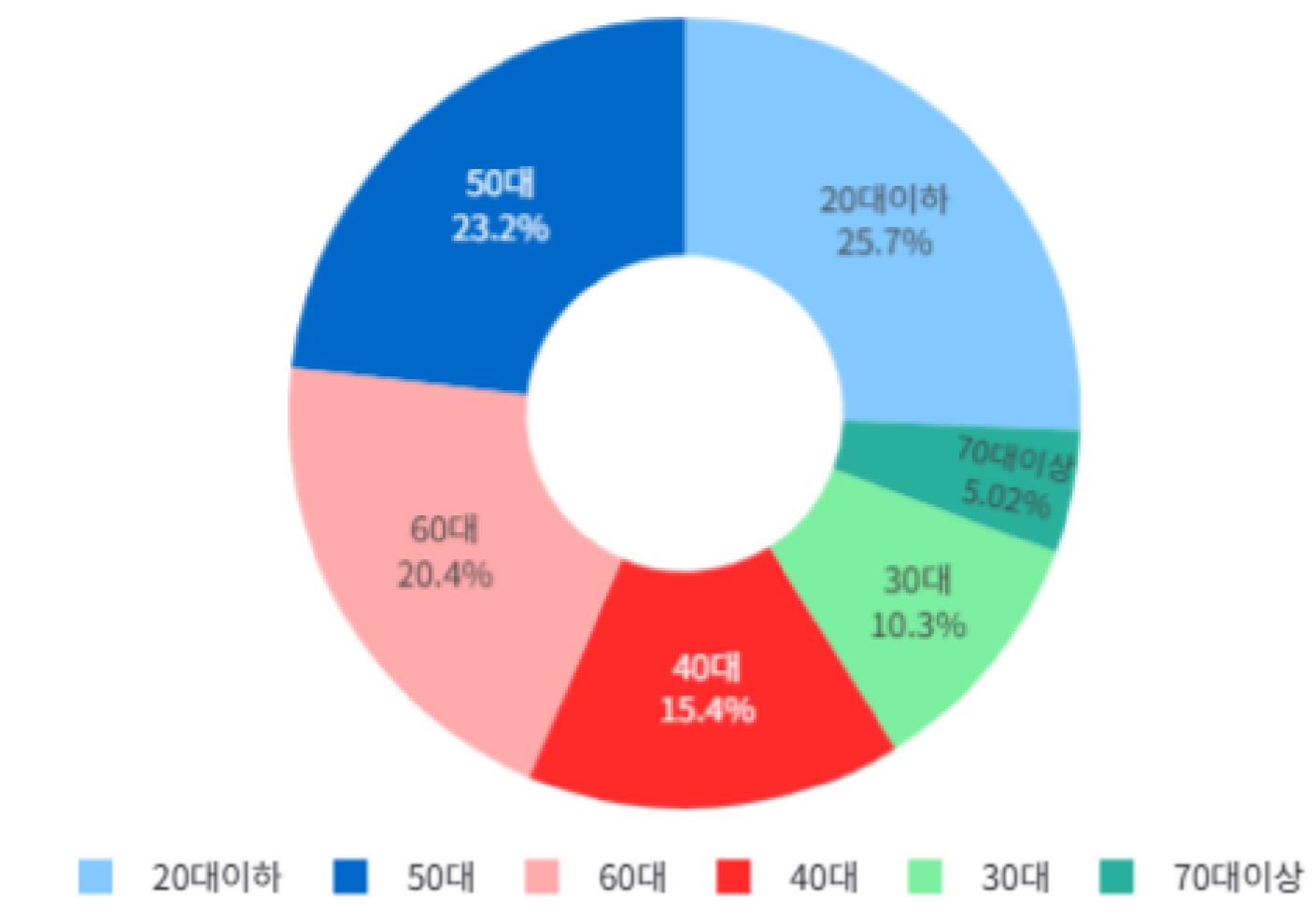
"누구나 범죄 대상이 될 수 있다."



[경찰청] 2016 ~ 2024 연령대 피해 추이



[경찰청] 2024 보이스피싱 피해



"보이스피싱 왜 당하나요?"

최근 경찰청에서 실시한 보이스피싱 인식 조사 결과
대략 안다(64.1%), 구체적으로 안다(25.9%)로 전체 응답자 중 **약 90%가 보이스피싱 범죄를 알고 있다는 답변**



수법도 많이 알려져있는데 부주의하다

알고 있는데 왜 당할까?

들으면 알아서 난 안당할 것 같다



AI를 활용한 최신 범죄 기법

AI 기반 음성 합성 기술로 가족 또는 지인의 목소리와 말투를 모방하여 금전을 갈취하는 신종 보이스피싱



고도로 발달된 AI

3초 분량의 목소리 샘플만 있어도 특정인의 말투, 문장 등을 구현할 수 있음

억양, 호흡, 감정 표현까지 가능하여 실제 음성 구분이 어려움

Why this matters



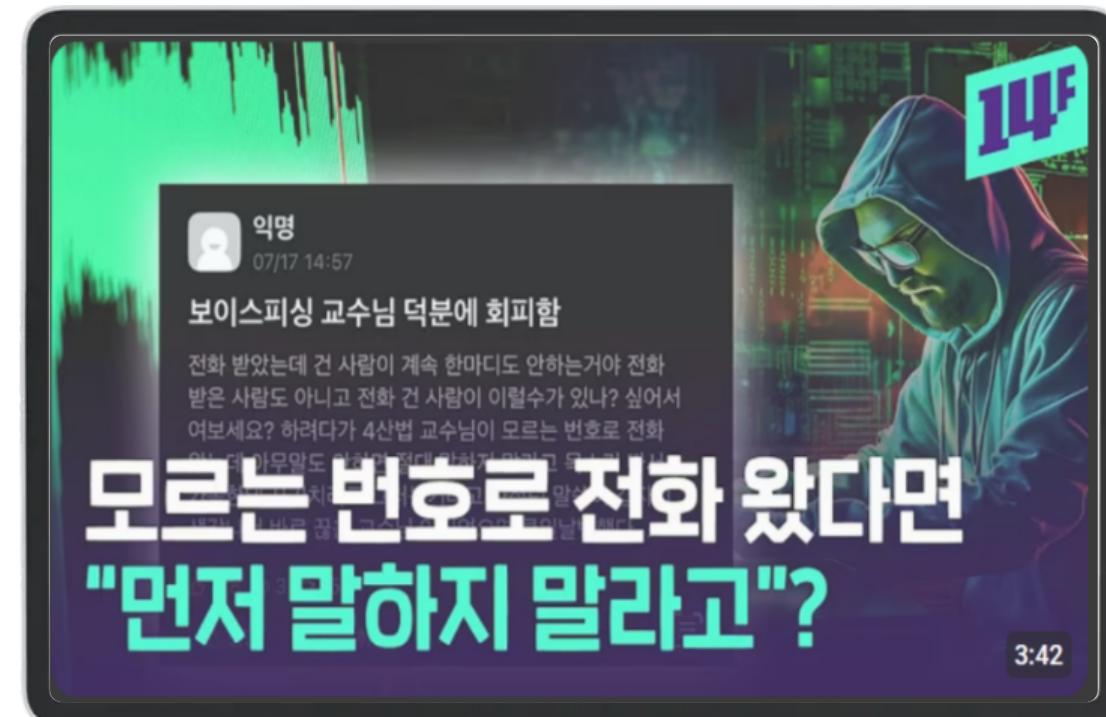
of your voice is all scammers need to create a deepfake.¹



of people say they are more concerned about deepfakes now than they were a year ago.²

AI와 같이 발전하는 범죄 기법

짧은 단어라도 말하는 순간 목소리가 녹음돼 AI로 음성이 복제되고 이를 활용해 가족과 지인에게 보이스피싱 시도



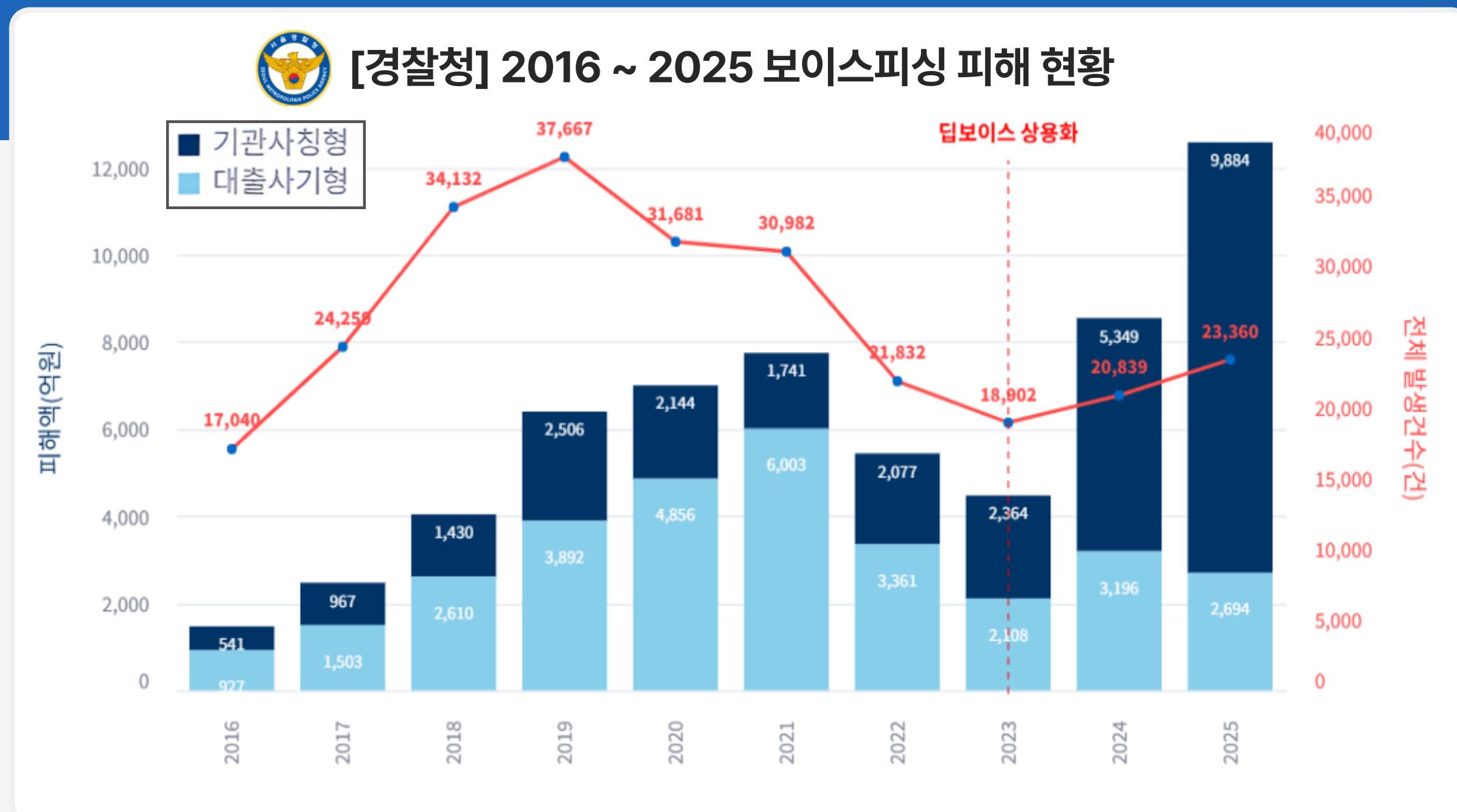
딥보이스피싱 주요 피해 사례

2019년 3월, AI를 사용해 독일 모회사 최고 경영자(CEO)의 목소리를 복제한 뒤 영국 자회사의 CEO에게 전화하여 헝가리의 한 공급업체에 긴급하게 자금을 송금하라고 지시하여 **약 430억 원의 피해** 발생



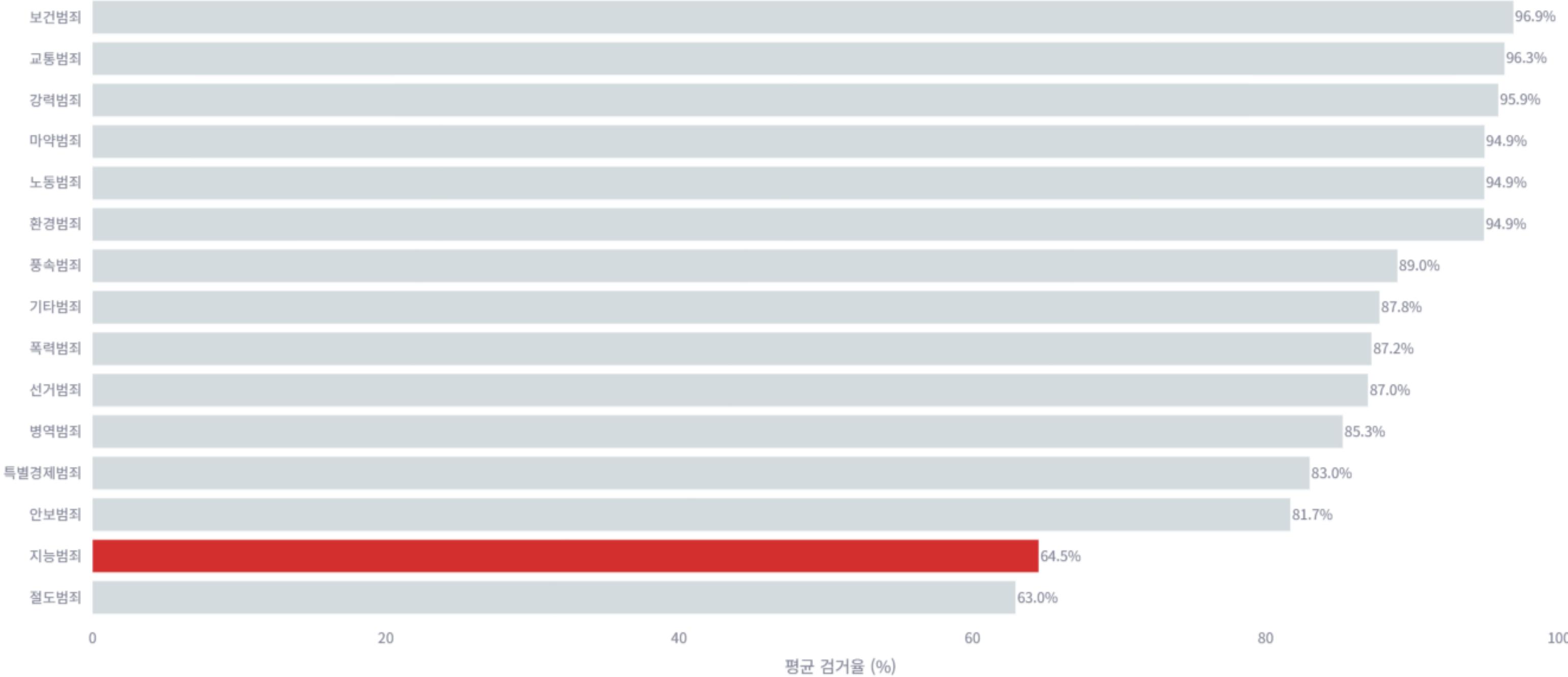
2023년 피해자 딸의 목소리를 AI로 복제하여 납치를 당했다며 돈을 요구하였고 60대 여성은 딸의 목소리를 100% 믿고 2천만원을 송금

2023년은 딥보이스 기술이 급격히 발전하여 구분하기 어려운 수준에 도달함에 따라,
이를 활용한 상용 서비스뿐만 아니라 이를 탐지하는 안티딥보이스 기술이 본격적으로 상용화된 시점





범죄 유형별 검거율



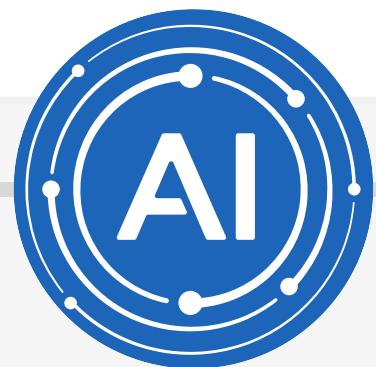
서비스 소개

솔루션 제공

AI를 활용한 딥보이스 탐지

정교해진 범죄로 인한
기존 탐지 기술의 한계

딥보이스 탐지 모델 결합으로
보이스피싱 탐지 범위 확대



늦은 피해 인식으로 놓치는 골든 타임

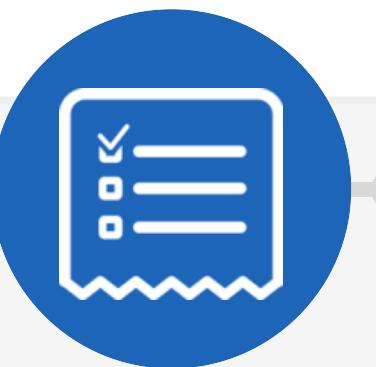
통화 내용 기반 보이스피싱
경고 및 정보 제공으로
빠른 판단 지원



보이스피싱 시나리오 활용

보이스피싱 대응방안
접근의 어려움

통화 내용 결과에 따른
가이드 제공으로 즉각 대응



의심 통화에 대한 사후관리 필요

의심 통화 신고
신고 번호 제공으로
반복 피해 방지



기존 안티피싱 서비스와 차별점

	탐지 중심	핵심 기술	신종 수법	개인정보	가중치 방식	판정
개발 앱	 대화 맥락	AE + koBERT 	방어	강력한 비식별화 	의미 유사도 	실시간 알림
익시오	 키워드	온디바이스 AI	취약	기기 내 처리	단순 포함 여부	실시간 알림
시티즌코난	 악성 APK	앱 시그니처 스캔	취약	권한 대량 요구		발견 즉시
에이닷	 키워드	서버형 AI 분석	취약	서버 전송 분석	단순 포함 여부	실시간 알림
후후	 발신자 번호	블랙리스트 DB	취약	번호/스팸 정보 위주		스팸/안심

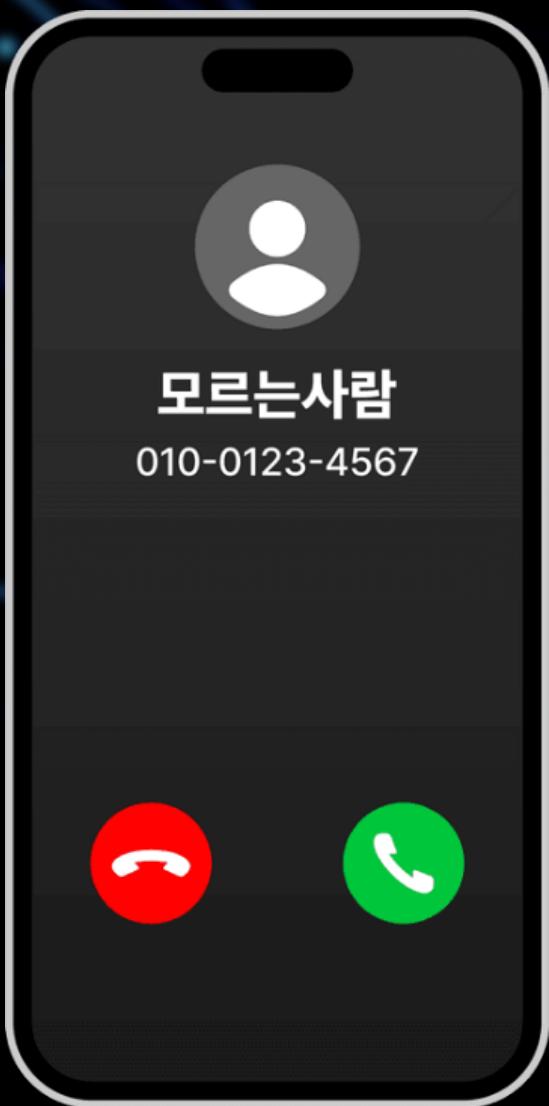
서비스 소개

1Q

2Q

3Q

4Q



신고이력 조회

딥보이스 탐지

통화 내용 실시간 분석

보이스피싱 탐지

보이스피싱 유형별 가이드 제공

통화 요약 리포트 제공

CONTENT

개발

A screenshot of a Java Integrated Development Environment (IDE) showing code for a Spring Boot application. The code includes annotations like `@Entity`, `@Table`, and `@Column`, indicating the mapping of database tables and columns. There are also `getters` and `setters` for various fields. The code is written in Java and is part of a larger class, likely a domain model.

```
1 package com.example.demo;
2
3 import org.springframework.data.annotation.Id;
4 import org.springframework.data.jpa.domain.EntityBase;
5 import org.springframework.data.jpa.domain.SimpleEntity;
6
7 import javax.persistence.*;
8
9 @Entity
10 @Table(name = "User")
11 @Id
12 @GeneratedValue(strategy = GenerationType.IDENTITY)
13 @Column(name = "id")
14 @Column(name = "name")
15 @Column(name = "email")
16 @Column(name = "password")
17 @Column(name = "role")
18
19 public class User extends SimpleEntity {
20
21     private String name;
22     private String email;
23     private String password;
24     private String role;
25
26     public User() {
27     }
28
29     public User(String name, String email, String password, String role) {
30         this.name = name;
31         this.email = email;
32         this.password = password;
33         this.role = role;
34     }
35
36     public String getName() {
37         return name;
38     }
39
40     public void setName(String name) {
41         this.name = name;
42     }
43
44     public String getEmail() {
45         return email;
46     }
47
48     public void setEmail(String email) {
49         this.email = email;
50     }
51
52     public String getPassword() {
53         return password;
54     }
55
56     public void setPassword(String password) {
57         this.password = password;
58     }
59
60     public String getRole() {
61         return role;
62     }
63
64     public void setRole(String role) {
65         this.role = role;
66     }
67
68     @Override
69     public String toString() {
70         return "User{" +
71                 "name='" + name + '\'' +
72                 ", email='" + email + '\'' +
73                 ", password='" + password + '\'' +
74                 ", role='" + role + '\'' +
75                 '}';
76     }
77
78     public static void main(String[] args) {
79         SpringApplication.run(DemoApplication.class, args);
80     }
81 }
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
259
260
261
262
263
264
265
266
267
268
269
269
270
271
272
273
274
275
276
277
278
279
279
280
281
282
283
284
285
286
287
288
289
289
290
291
292
293
294
295
296
297
298
299
299
300
301
302
303
304
305
306
307
308
309
309
310
311
312
313
314
315
316
317
318
319
319
320
321
322
323
324
325
326
327
328
329
329
330
331
332
333
334
335
336
337
338
339
339
340
341
342
343
344
345
346
347
348
349
349
350
351
352
353
354
355
356
357
358
359
359
360
361
362
363
364
365
366
367
368
369
369
370
371
372
373
374
375
376
377
378
379
379
380
381
382
383
384
385
386
387
388
389
389
390
391
392
393
394
395
396
397
398
399
399
400
401
402
403
404
405
406
407
408
409
409
410
411
412
413
414
415
416
417
418
419
419
420
421
422
423
424
425
426
427
428
429
429
430
431
432
433
434
435
436
437
438
439
439
440
441
442
443
444
445
446
447
448
449
449
450
451
452
453
454
455
456
457
458
459
459
460
461
462
463
464
465
466
467
468
469
469
470
471
472
473
474
475
476
477
478
479
479
480
481
482
483
484
485
486
487
488
489
489
490
491
492
493
494
495
496
497
498
499
499
500
501
502
503
504
505
506
507
508
509
509
510
511
512
513
514
515
516
517
518
519
519
520
521
522
523
524
525
526
527
528
529
529
530
531
532
533
534
535
536
537
538
539
539
540
541
542
543
544
545
546
547
547
548
549
549
550
551
552
553
554
555
556
557
558
559
559
560
561
562
563
564
565
566
567
568
569
569
570
571
572
573
574
575
576
577
578
579
579
580
581
582
583
584
585
586
587
588
589
589
590
591
592
593
594
595
596
597
598
599
599
600
601
602
603
604
605
606
607
608
609
609
610
611
612
613
614
615
616
617
618
619
619
620
621
622
623
624
625
626
627
628
629
629
630
631
632
633
634
635
636
637
638
639
639
640
641
642
643
644
645
646
647
648
649
649
650
651
652
653
654
655
656
657
658
659
659
660
661
662
663
664
665
666
667
668
669
669
670
671
672
673
674
675
676
677
678
679
679
680
681
682
683
684
685
686
687
688
689
689
690
691
692
693
694
695
696
697
697
698
699
699
700
701
702
703
704
705
706
707
708
709
709
710
711
712
713
714
715
716
717
718
719
719
720
721
722
723
724
725
726
727
728
729
729
730
731
732
733
734
735
736
737
738
739
739
740
741
742
743
744
745
746
747
748
749
749
750
751
752
753
754
755
756
757
758
759
759
760
761
762
763
764
765
766
767
768
769
769
770
771
772
773
774
775
776
777
778
779
779
780
781
782
783
784
785
786
787
788
789
789
790
791
792
793
794
795
796
797
798
799
799
800
801
802
803
804
805
806
807
808
809
809
810
811
812
813
814
815
816
817
818
819
819
820
821
822
823
824
825
826
827
828
829
829
830
831
832
833
834
835
836
837
838
839
839
840
841
842
843
844
845
846
847
848
849
849
850
851
852
853
854
855
856
857
858
859
859
860
861
862
863
864
865
866
867
868
869
869
870
871
872
873
874
875
876
877
878
879
879
880
881
882
883
884
885
886
887
888
889
889
890
891
892
893
894
895
896
897
897
898
899
899
900
901
902
903
904
905
906
907
908
909
909
910
911
912
913
914
915
916
917
918
919
919
920
921
922
923
924
925
926
927
928
929
929
930
931
932
933
934
935
936
937
938
939
939
940
941
942
943
944
945
946
947
948
949
949
950
951
952
953
954
955
956
957
958
959
959
960
961
962
963
964
965
966
967
968
969
969
970
971
972
973
974
975
976
977
978
979
979
980
981
982
983
984
985
986
987
988
989
989
990
991
992
993
994
995
996
997
998
999
999
1000
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1089
1090
1091
1092
1093
1094
1095
1096
1097
1097
1098
1099
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1189
1190
1191
1192
1193
1194
1195
1196
1196
1197
1198
1199
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1289
1290
1291
1292
1293
1294
1295
1296
1297
1297
1298
1299
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1389
1390
1391
1392
1393
1394
1395
1396
1397
1397
1398
1399
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1489
1490
1491
1492
1493
1494
1495
1496
1497
1497
1498
1499
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1549
1550
```

모든 과정은 통화가 진행되는 실시간으로 이루어져야 하며, 경고 알림도 즉시 전송되어야 피해를 막을 수 있다.



윤리 및 법적 문제

- 개인 정보 보호
- 통신 비밀 보호
- 구글 보안 정책



오탐 문제

- 결정 임계값을 신중하게 설정하여 정상 통화를 피싱으로 오인하는 상황이 없어야 한다.



사후 관리

- 신종 수법을 방어할 수 있어야 하고, 반복적인 피해를 방지해야 한다.

기술 스택

API



앱



협업

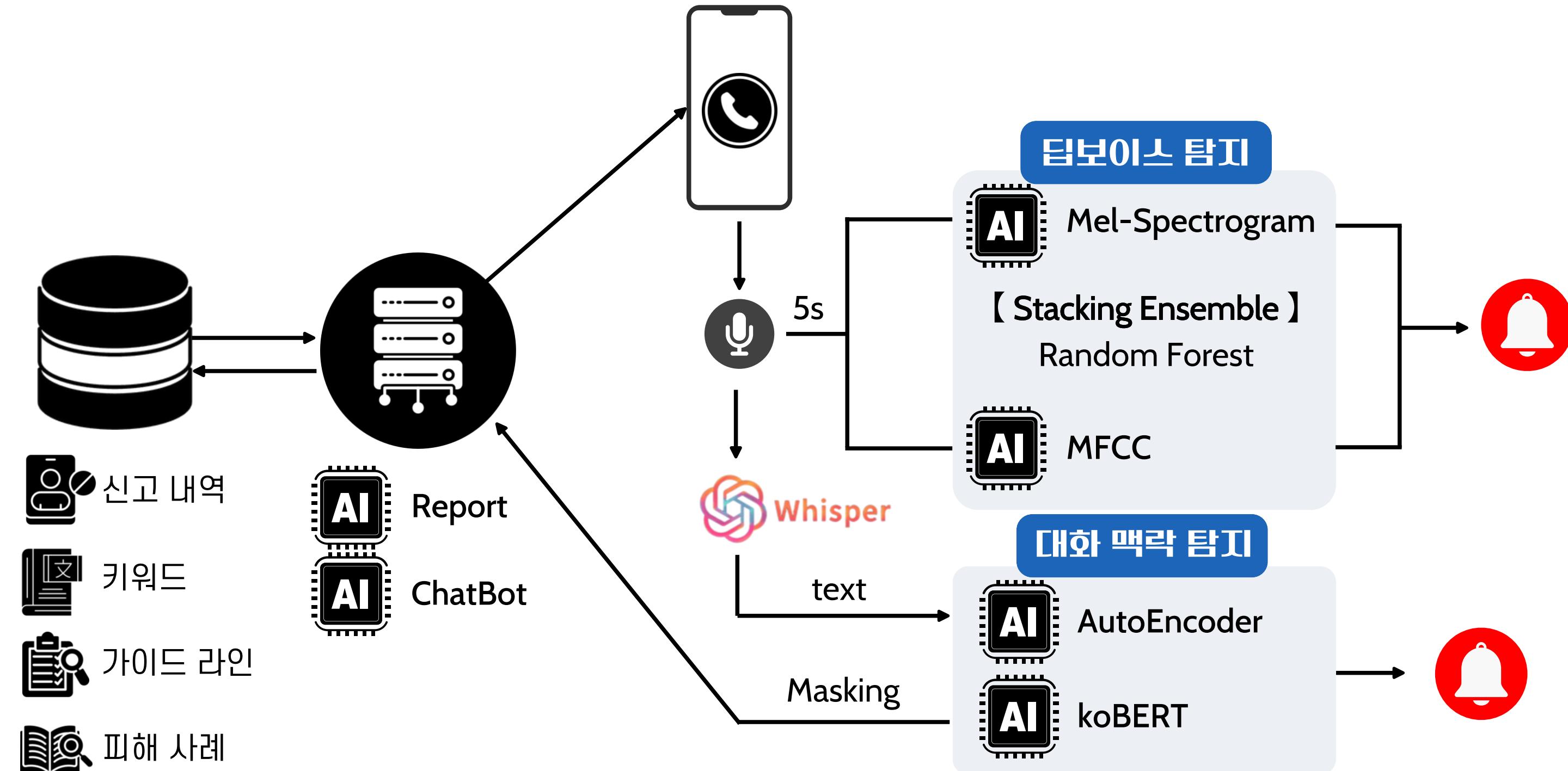


모델링



데이터





팀원 소개 및 역할 분담

TEAM

Dropout(0.25)

AI 엔지니어



팀장 허성욱

프로젝트 총괄 및 일정 관리
딥보이스 탐지 모델 개발
대화 맥락 탐지 모델 개발
전체 파이프라인 설계

풀스택



팀원 이경민

애플리케이션 개발
데이터베이스 설계 및 구축
FastAPI 서버 구축
유지 보수 파이프라인 구축

데이터 분석



팀원 김정안

피싱 데이터 수집 및 분석
딥보이스 데이터 분석
데이터 시각화 및 브리핑
분석 보고서 대시보드 작성

AI 엔지니어



팀원 김나영

딥보이스 탐지 모델 개발
요약 리포트 모델 개발
챗봇 모델 개발
문서화 작업

사전 DB 구축

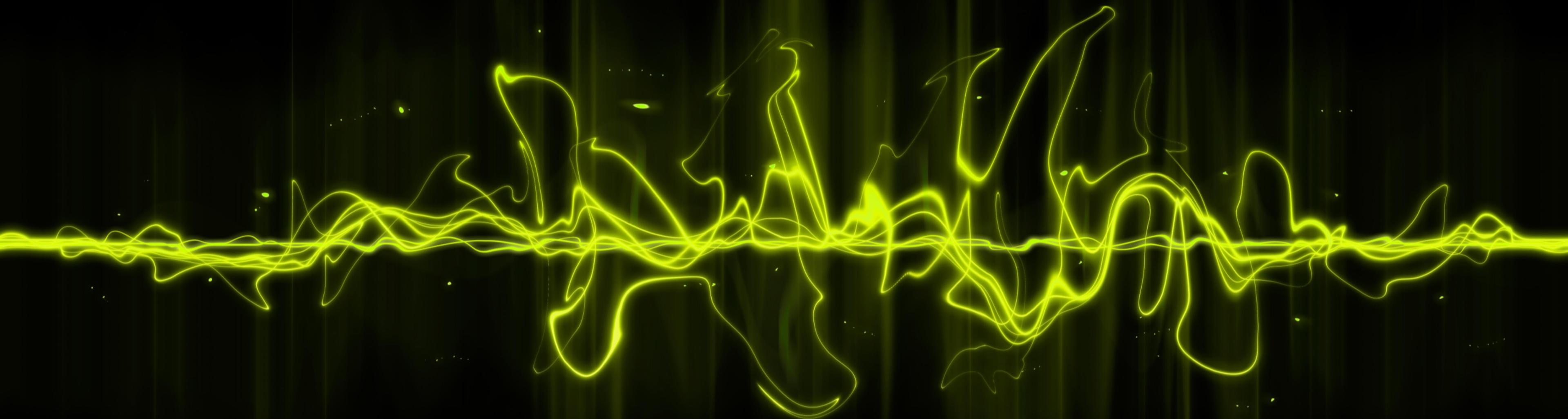
- 전기통신금융사기 통합대응단
크롤링
 - 공공기관 트래픽 공격 이슈
 - 블랙리스트 DB 임시 번호 기입

실시간 오디오 접근

- 단말기 권한 문제로 통화 오디오
직접 접근 제한
 - 스피커폰을 통한 접근 우회
 - 반자동 서비스 우회
 - 루팅을 통한 직접 접근

감성 분류

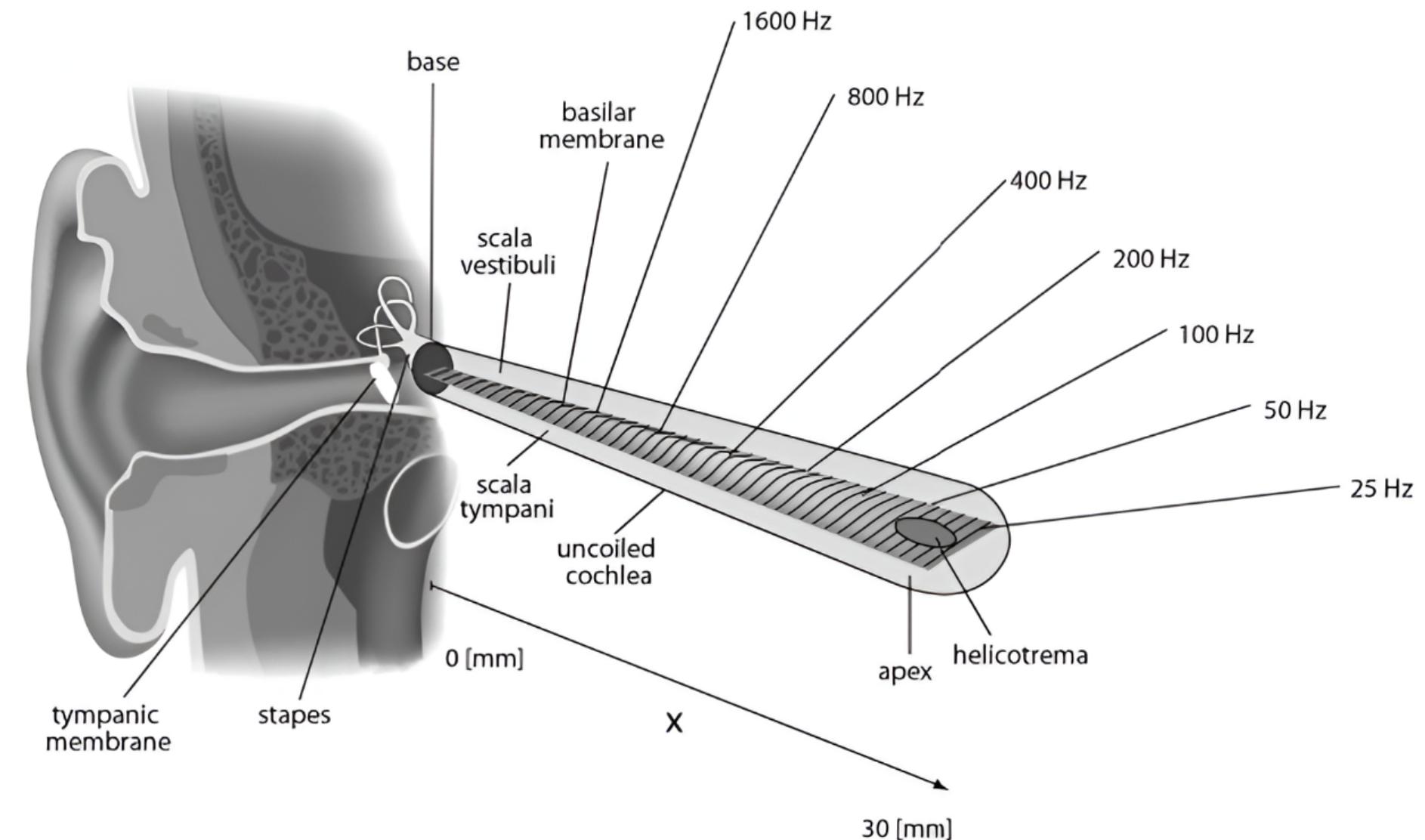
- 발화자 구분 및 감성 분류
 - 초기 학습 실패
 - 데이터 품질 모호
 - 분석 소요 시간 모호
 - 기존 아키텍처에서 제외



"기계는 음성을 어떻게 다루는가?"

MFCC / MEL-SPECTROGRAM

[논문] Automatic Phoneme Recognition using Mel-Frequency Cepstral Coefficient and Dynamic Time Warping



Mel

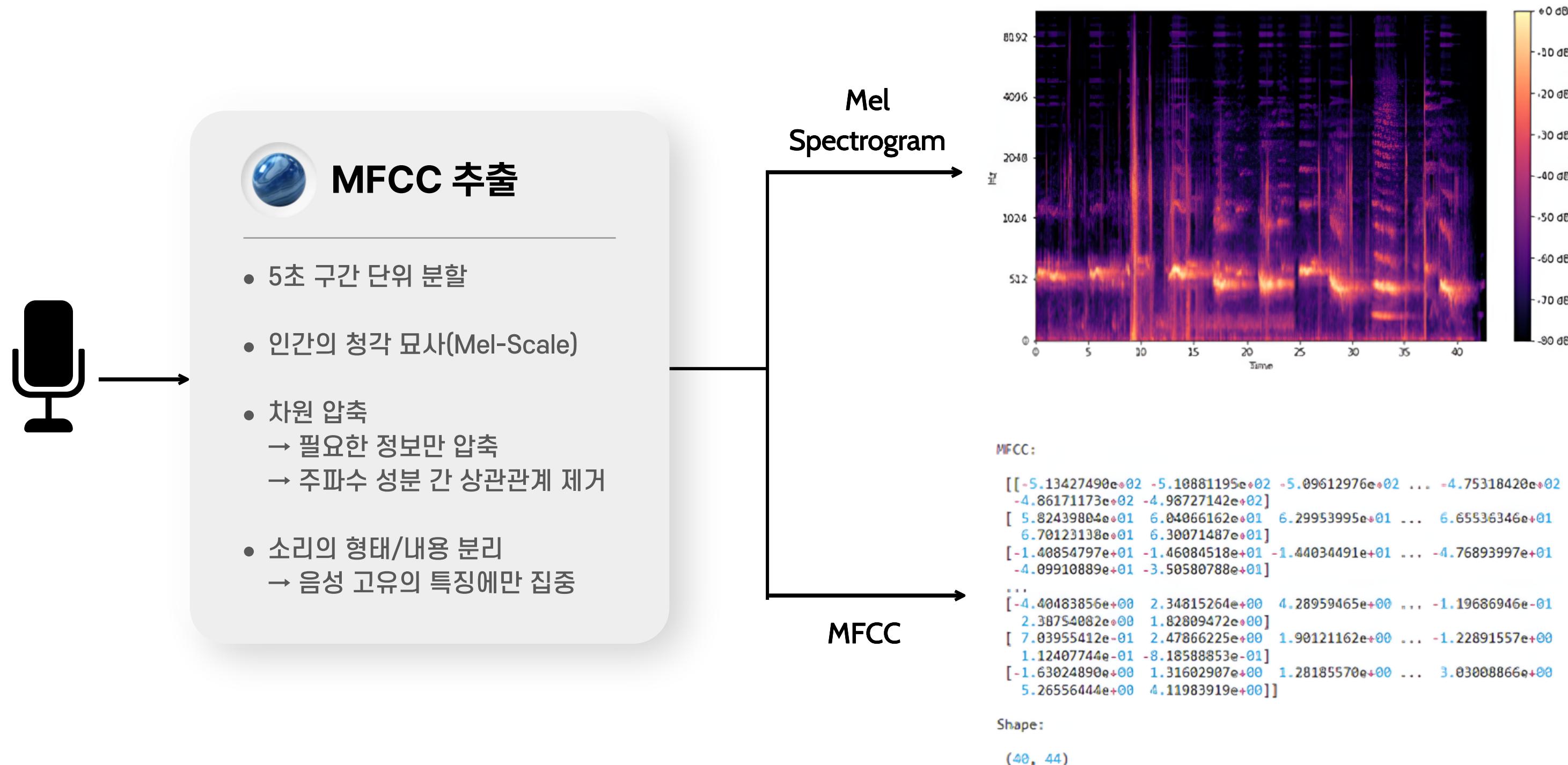
인간의 귀로 측정한 주파수 단위

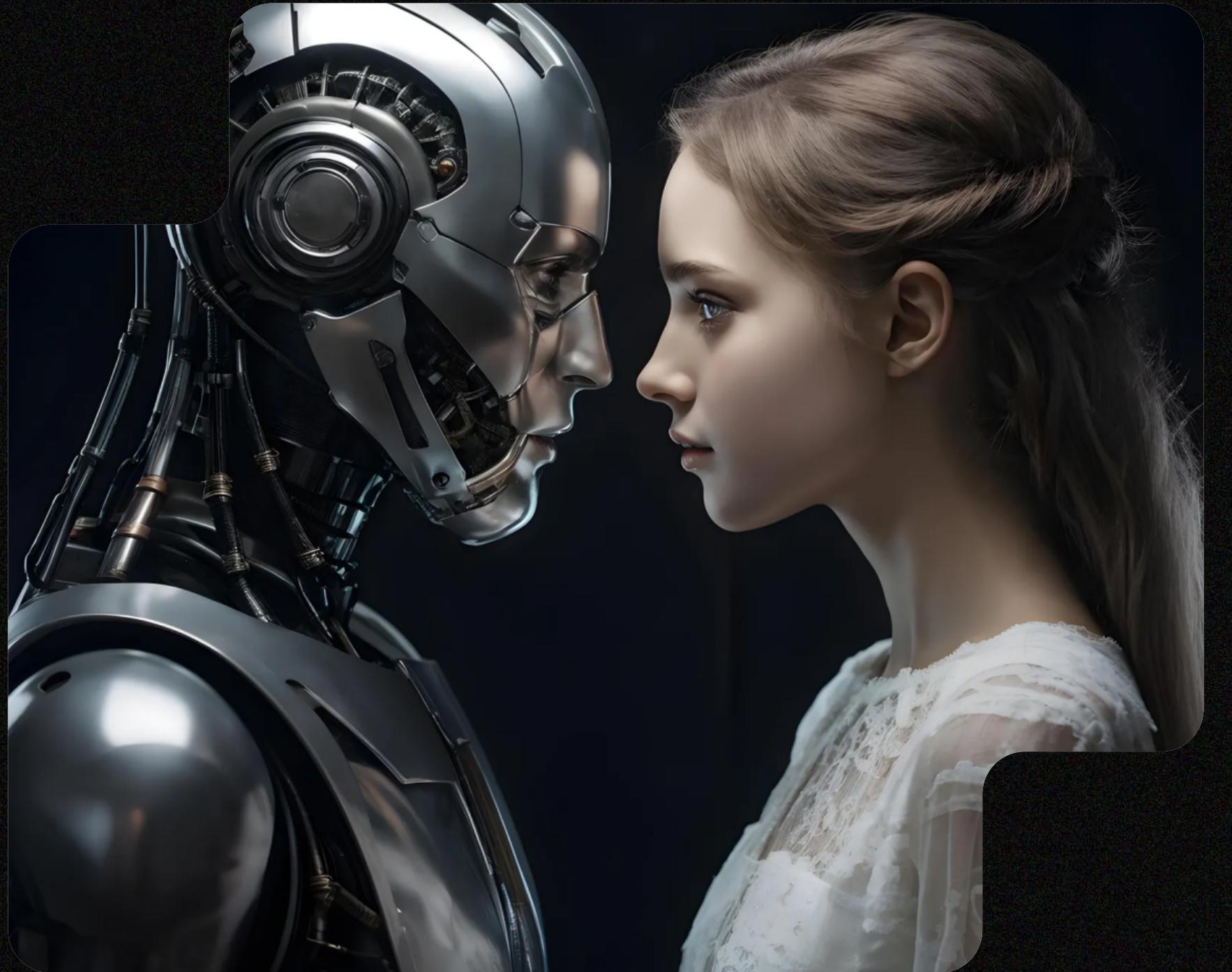
Mel-Scale

인간의 달팽이관 특성을 고려

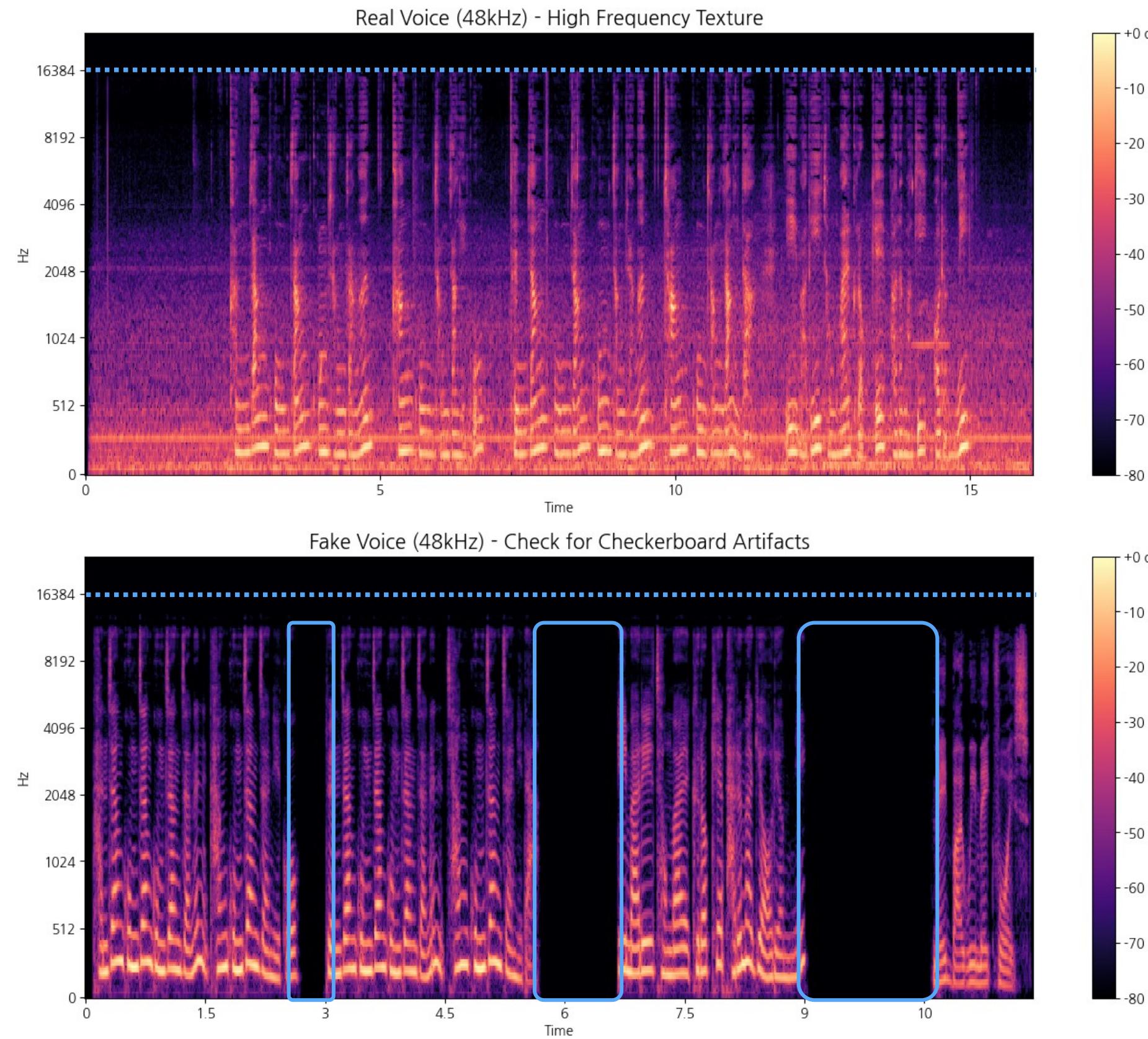
MFCC / MEL-SPECTROGRAM

[논문] Voice Recognition Using MFCC Algorithm



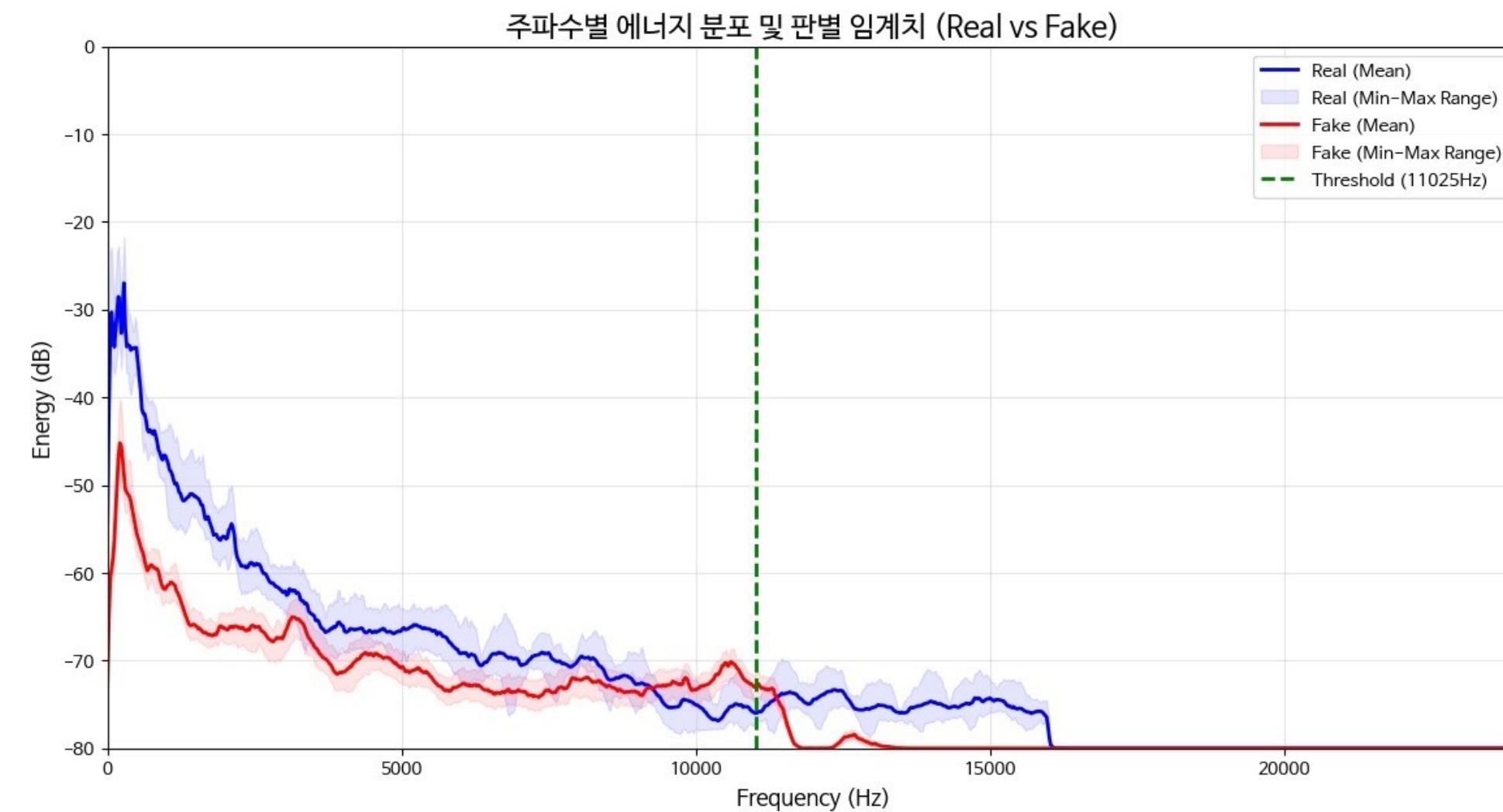


변조 음성 VS 실제 음성



변조 흔적

- 실제 음성 대비 고주파 구간이 검은색으로 완전히 비어있음
- 변조 음성은 문장 사이의 무음 구간이 완벽하게 검정색(0)



변조 흔적

- 11~12kHz 사이에서 에너지가 급격하게 떨어짐
- 10~11kHz 사이에서 에너지가 솟구치는 구간 발생
- 전반적인 에너지 레벨이 실제 음성에 비해 낮게 형성
- 에너지 분포 구간 폭이 좁음



CNN

- 핵심** 고정 크기 필터링
- 분석** 특정 시점 주파수 특징
- 장점** 지역 특성 포착 유리
- 이슈** 특정 환경에 과적합



Transformer

- 핵심** Self-Attention
- 분석** 전체적인 맥락
- 장점** 정교한 합성음 탐지
- 이슈** 성능 지표 모호



Res2Net

- 핵심** 계층적 멀티스케일
- 분석** 다양한 시간/주파수
- 장점** 데이터 효율 높음
특징 추출 극대화
높은 파라미터 효율
과적합에 강한 구조



Res2Net50

- 다중 척도 추출
- 50계층의 깊은 네트워크



SE Block

- 전체 정보 압축
- 변조 채널에만 집중



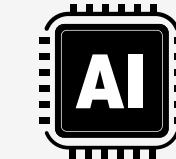
Res2Net50-SE

- 미세한 합성 흔적 포착
- 도메인 강건성
- 학습 효율 및 일반화 성능



Dataset

- 랜덤 가우시안 노이즈
- 랜덤 볼륨 증폭



학습 결과

Train Acc 99.8% | Train Loss 0.005 | Val Acc 99.03% | Val Loss 0.0371

테스트

정답 / 테스트 데이터 : (59,349 / 60,014)

Acc

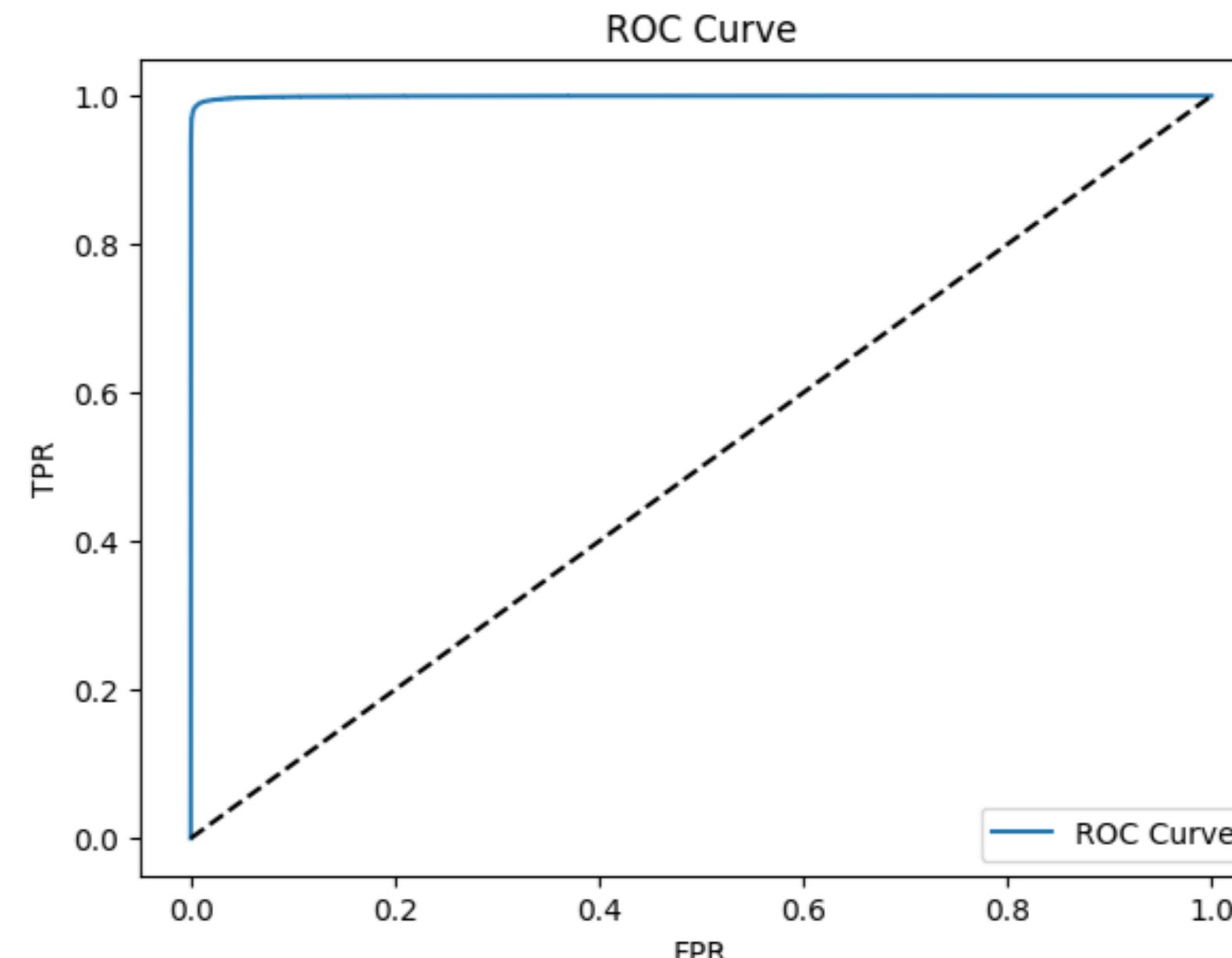
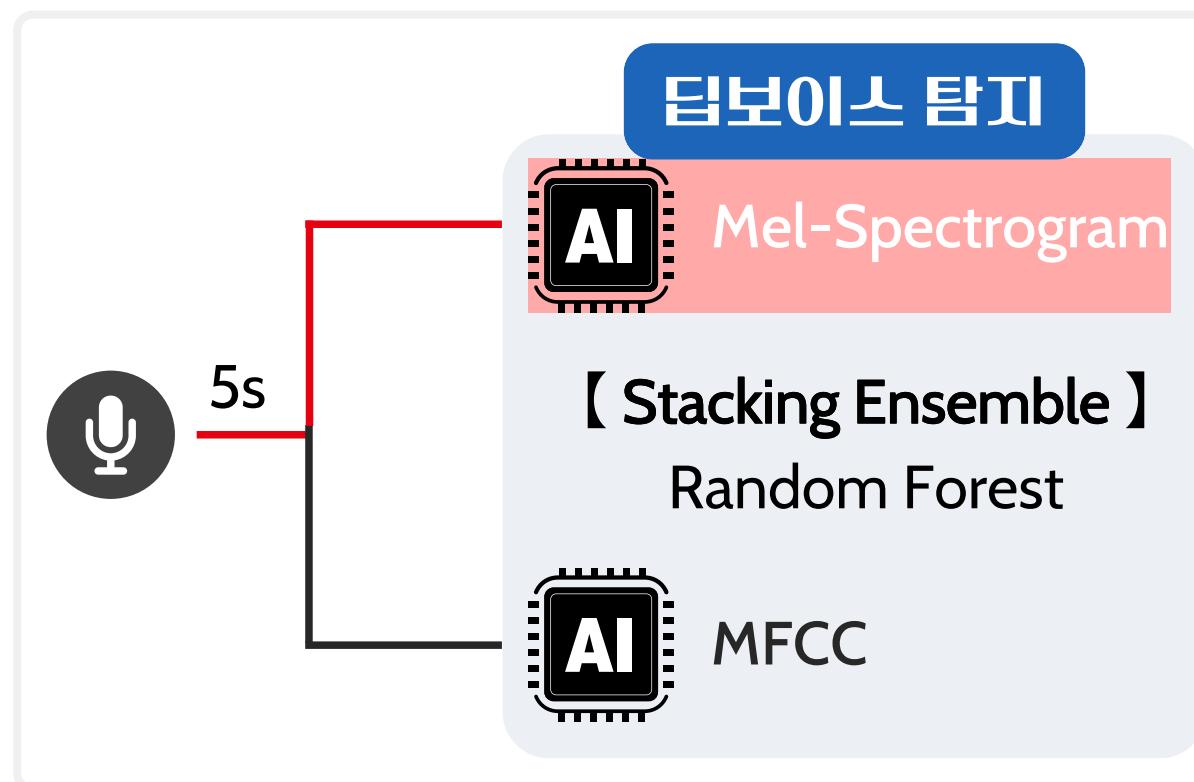
98.89%

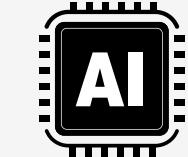
EER

0.05%

임계치

0.7297





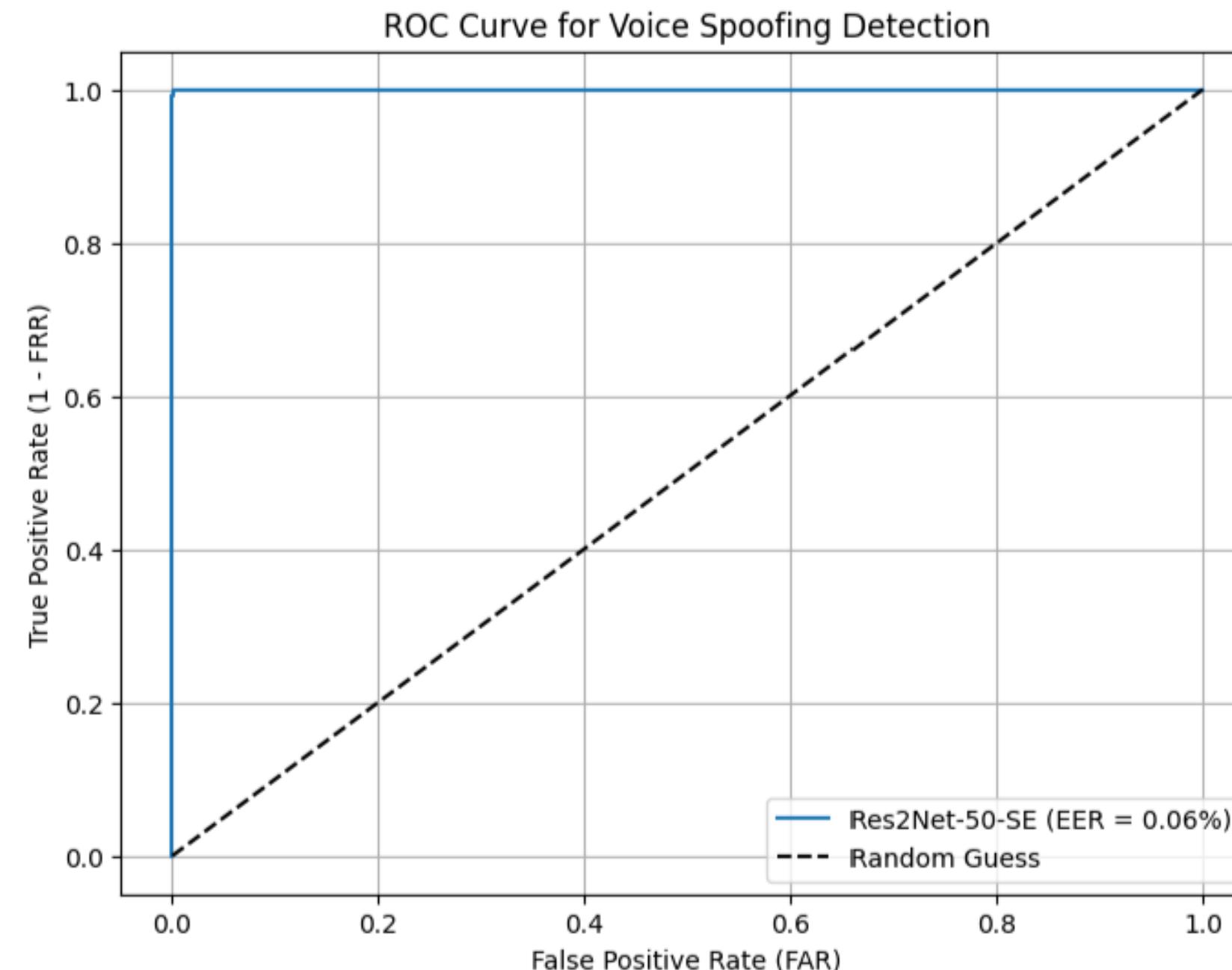
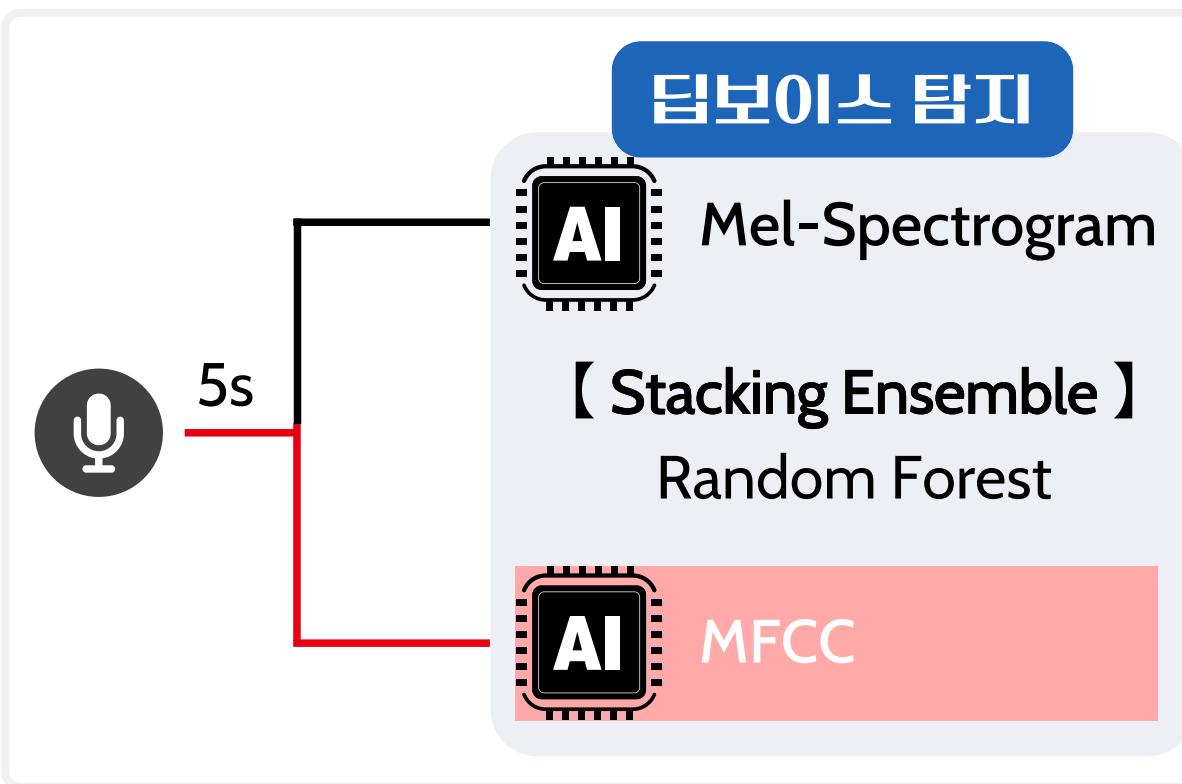
학습 결과 Train Acc 99.94% | Train Loss 0.0019 | Val Acc 99.88% | Val Loss 0.0035

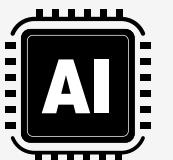
테스트 정답 / 테스트 데이터 : (59,954 / 60,000)

Acc 99.92%

EER 0.0633%

임계치 0.2893

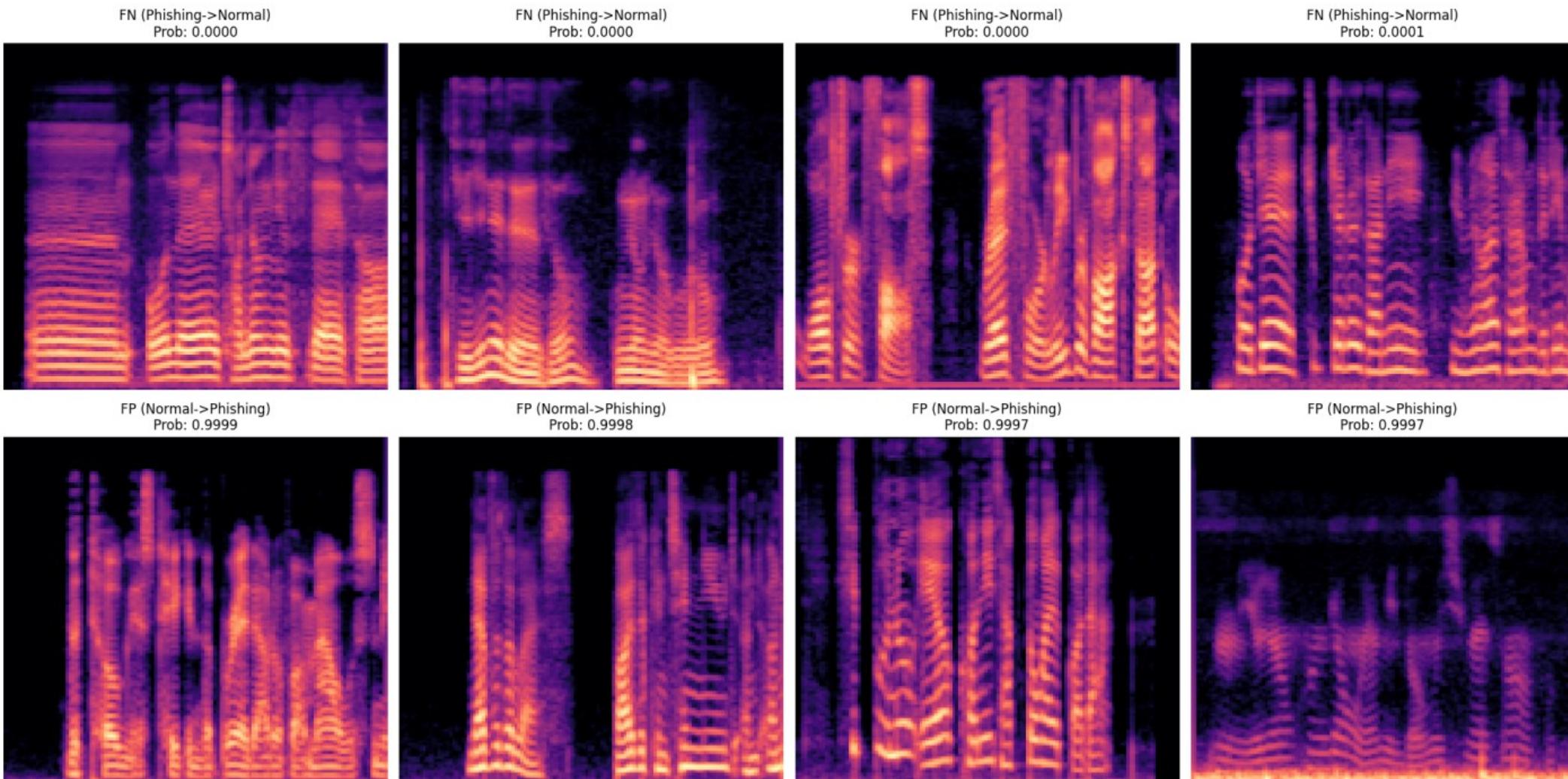




오답 유형

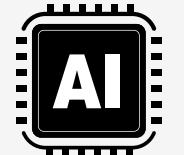
False Positive : 433, False Negative : 232 - (665 / 60,014)

Model Error Analysis: False Negatives vs False Positives



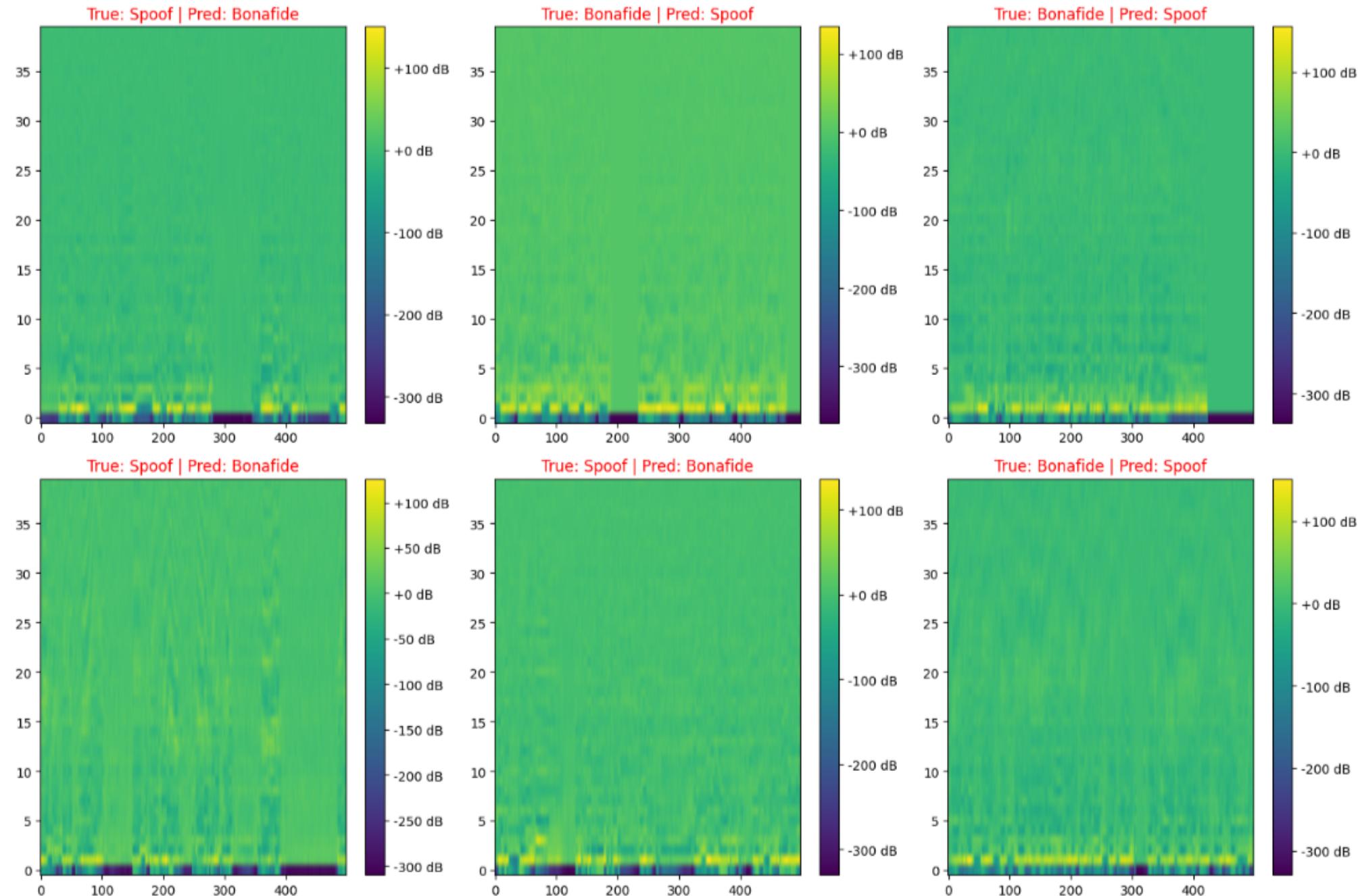
오답

- 한국어 정상 329개
- 영어 정상 104개
- 한국어 변조 183개
- 영어 변조 49개



오답 유형

False Positive : 13, False Negative : 33 - (46 / 60,000)



오답

- 한국어 정상 10개
- 영어 정상 3개
- 한국어 변조 28개
- 영어 변조 5개

MFCC

파일: /data/통화테스트

결과: 정상(본인음성)

상세 확률: [정상: 100.00% / 변조: 0.00%]

판단 기준: 임계값 0.2893 적용

파일: /data/딥보이스/딥보이스_제작.mp3

결과: ⚠️ 딥보이스(변조음)

상세 확률: [정상: 0.00% / 변조: 100.00%]

판단 기준: 임계값 0.2893 적용

파일: /data/딥보이스/딥보이스_여성.mp3

결과: ⚠️ 딥보이스(변조음)

상세 확률: [정상: 0.01% / 변조: 99.99%]

판단 기준: 임계값 0.2893 적용

Mel-Spectrogram

파일명: 딥보이스_남성.mp3

↳ [예측] 확률: 99.97% -> ⚠️ 보이스피싱

↳ [정답] ⚠️ 보이스피싱

↳ [결과] ⭕ 일치

파일명: 딥보이스_여성.mp3

↳ [예측] 확률: 100.00% -> ⚠️ 보이스피싱

↳ [정답] ⚠️ 보이스피싱

↳ [결과] ⭕ 일치

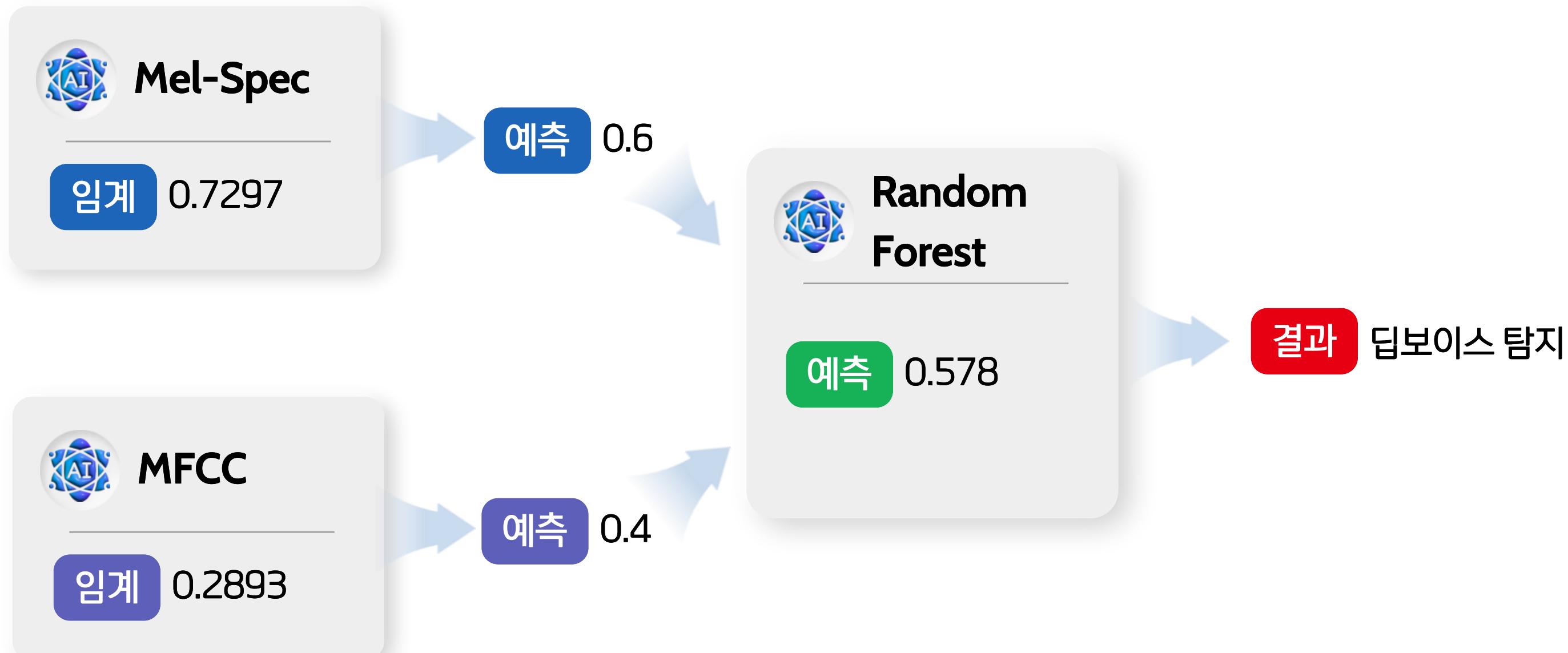
파일명: 안녕하세요 저는 여자 목소리.mp3

↳ [예측] 확률: 100.00% -> ⚠️ 보이스피싱

↳ [정답] ⚠️ 보이스피싱

↳ [결과] ⭕ 일치

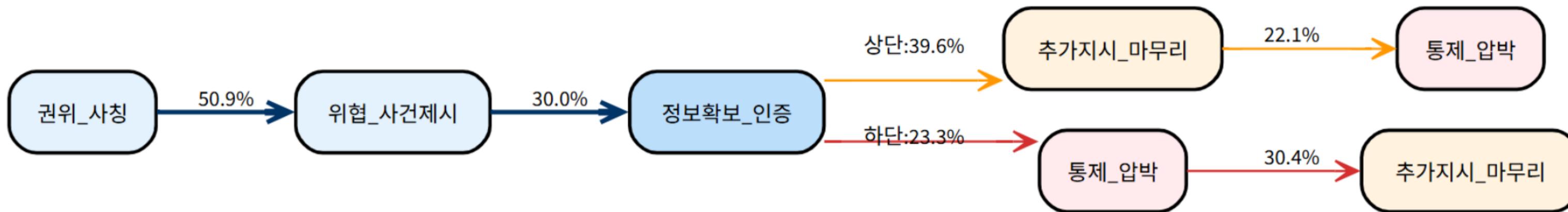
Stacking Ensemble





대화 맥락 탐지

기관사칭형



1. 권위 사칭

- 기관을 사칭해 신뢰 형성
- 서울중앙지검
 - 첨단범죄수사팀
 - 금융감독원
 - 과장/수사관
 - 사건번호
 - 공문 발송
 - 녹취 시작
 - 사건 번호

2. 위협·사건

- 범죄 연루 사실 통보
- 대포통장 개설
 - 명의 도용
 - 중고나라 사기
 - 피해자 입증
 - 자금 세탁
 - 계좌 동결
 - 고소장 접수
 - 출석 요구서

3. 정보확보

- 개인/금융정보 탈취
- 자산 내역 확인
 - IP 주소 추적
 - OTP 번호
 - 신분증 촬영
 - 팀뷰어(원격)
 - 본인 인증
 - 계좌 비밀번호
 - 공인인증서

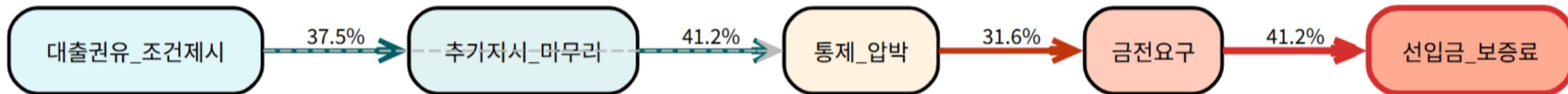
4. 통제·압박

- 심리적 고립 및 협박
- 조용한 곳 이동
 - 제3자 발설 금지
 - 전화 끊지 마세요
 - 주변 소음 차단
 - 수사 방해
 - 공무집행방해
 - 구속 영장 청구

5. 마무리

- 금전 탈취 및 증거인멸
- 국가안전계좌 이체
 - 현금 인출 후 전달
 - 상품권 핀번호
 - 악성 앱 설치
 - 대출 실행
 - 카카오톡 탈퇴
 - 통화 기록 삭제

대출사기형



회유:26.3%

1. 대출권유

저금리 대출 유혹

- 정부지원 자금
- 저금리 대환 대출
- 햄살론/버팀목
- 신용등급 상향
- 1금융권
- 마이너스 통장
- 신청 대상자
- 무보증/무담보

2. 추가자서

개인정보 및 앱 설치

- 금융기관 앱 설치
- 신분증 사본 전송
- 기존 대출 상환
- 카톡 친구추가
- 신청서 작성
- 입출금 내역
- 재직 증명서

3. 통제·압박

가짜 심사 및 교란

- 심사 진행 중
- 신용 평점 부족
- 법무사 통화
- 금융법 위반
- 전산 처리
- 모니터링 감지
- 부결 사유
- 중복 신청

4. 금전요구

각종 비용 청구

- 보증 보험료
- 예치금 납부
- 공탁금 설정
- 인지세/수수료
- 상환 처리 비용
- 신용 보증금
- 계좌 해지 비용

5. 선입금

최종 금전 갈취

- 가상계좌 발급
- 편법 상환 처리
- 선입금 입금
- 담당자 계좌
- 무통장 입금
- 즉시 이체
- 현금 인출 전달

1 강도 높은 비식별화

- 보안/개인정보 보호
- 강도 높은 마스킹 기법 적용



2 TF-IDF 벡터라이저

- 등장 빈도 가중치 벡터



3 파인 투닝

- 문맥 기반의 파인 투닝



koBERT

- ✓ 한국어 맞춤형 사전 학습 모델
- ✓ 한국어 특유의 말투/문맥 이해도 높음
- ✓ 문맥의 양방향 이해
- ✓ 문장 내 단어들 간의 관계를 다각도로 분석



비용/속도 이슈

- 1억개 이상의 파라미터
- 피싱 대화 전체를 무거운 모델이 전부 처리하는 데 서버 비용이 많이 들고 속도가 느림



과적합

- 맥락보다 특정 키워드에 집중하여 정상 대화를 오탐



신종 수법 취약

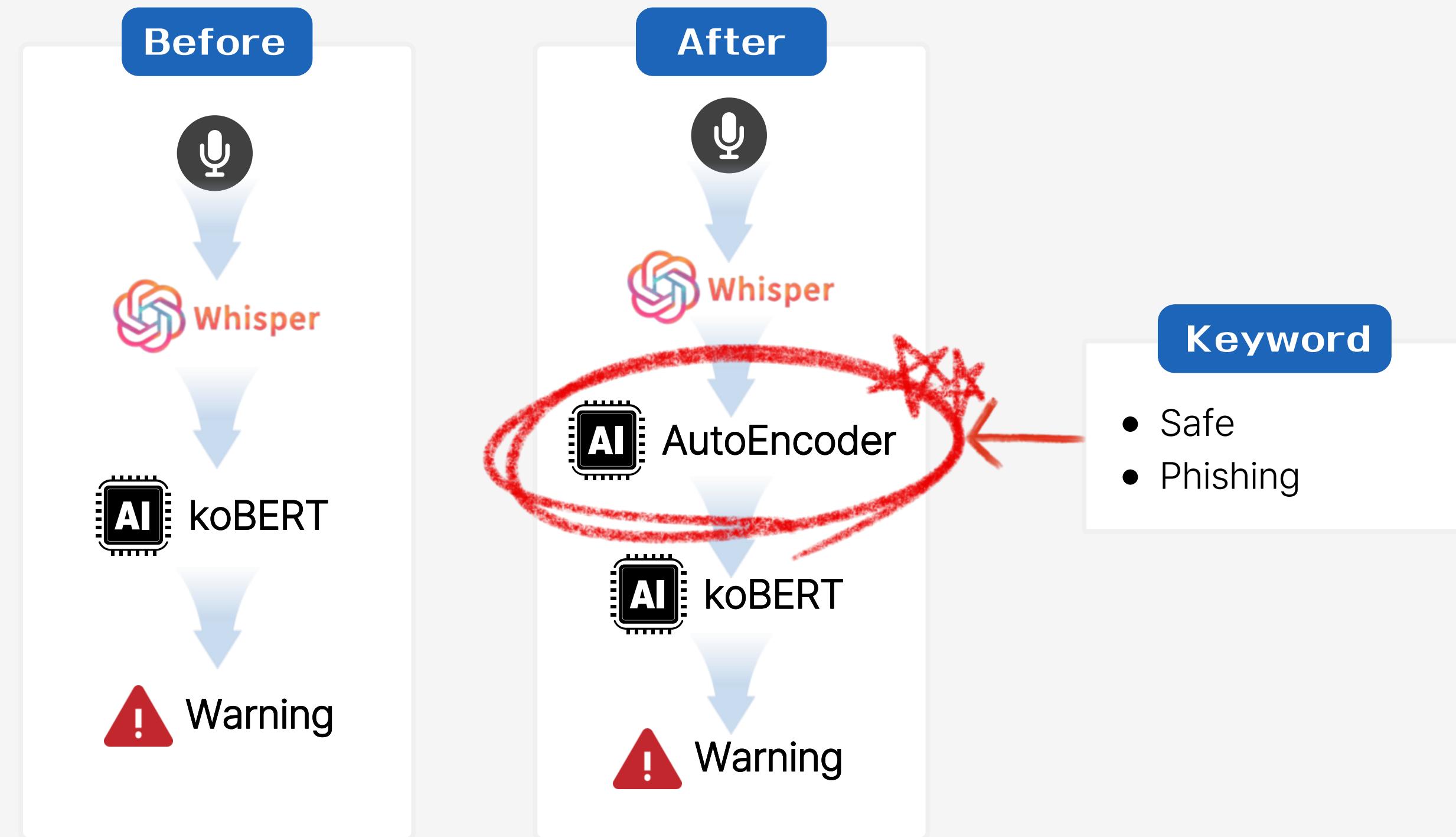
- 학습되지 않은 신종 수법이 등장할 경우, 어떤 결과를 예측할 지 모름

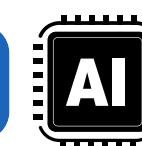


유지 보수 어려움

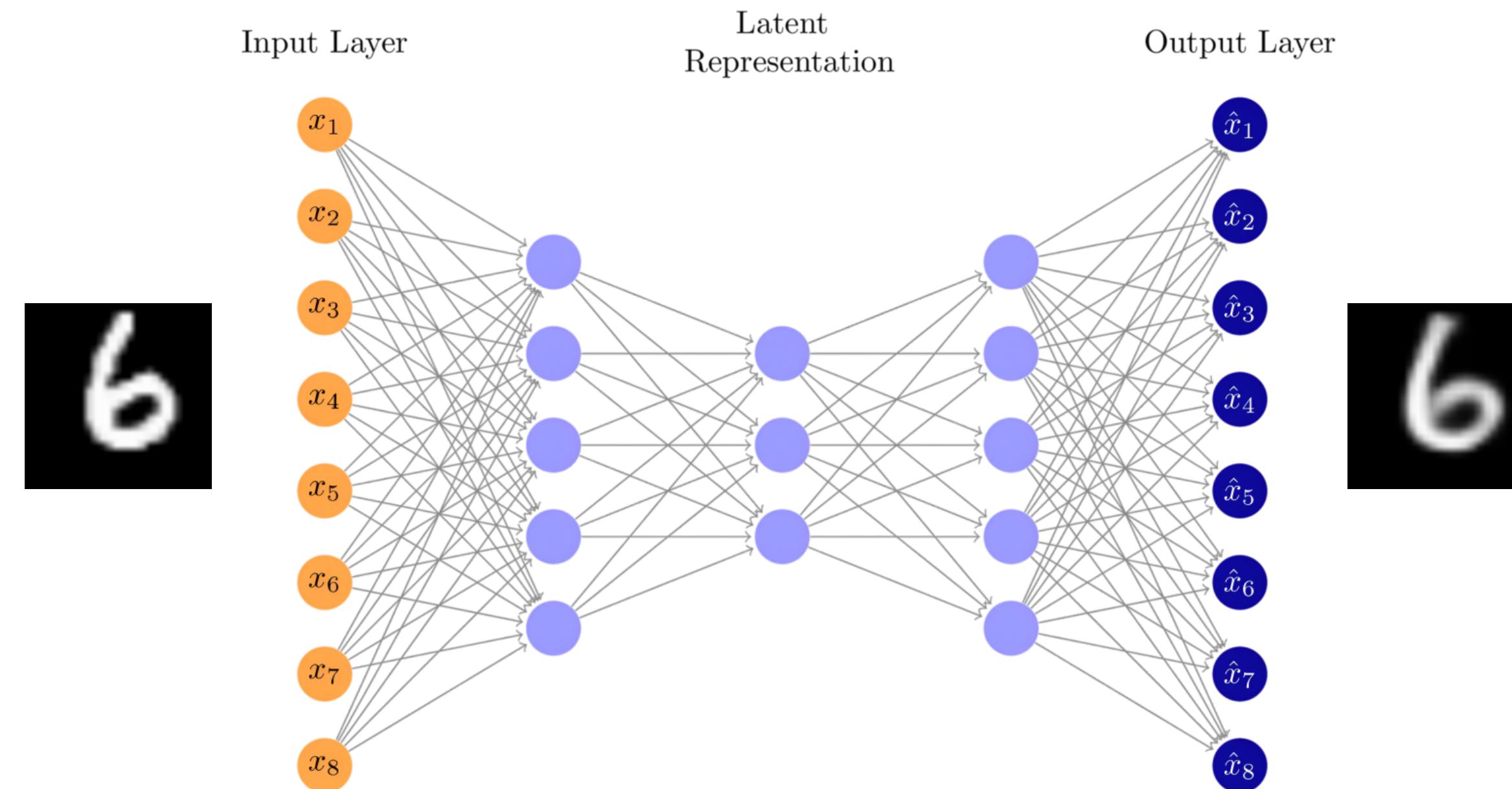
- 신종 수법이 등장할 때마다 재학습시키는 점이 번거로움

보이스피싱 탐지 알고리즘 개선





AutoEncoder (Anomaly Detection)



구조 변경 후 개선점



연산 효율 증가

- 통화의 대다수를 차지하는 일상적인 대화는 가벼운 AE가 처리하므로 연산량/서버비용 대폭 감소
- 사용자 경험 개선



신종 수법 방어

- 신종 수법이 등장하더라도 AE가 1차적으로 차단하기 때문에 대응 가능



오탐 감소

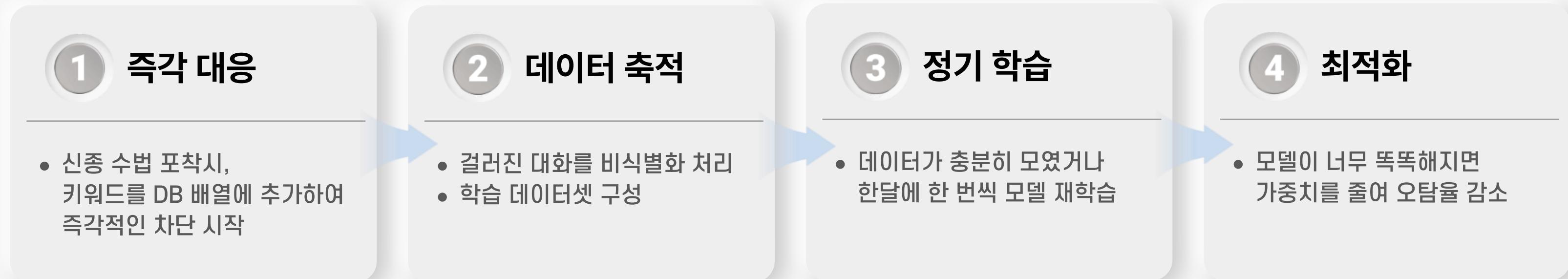
- 두 모델이 다른 관점으로 분석
- 단순 키워드 등장뿐만 아니라, 문맥적 이상함까지 탐지해야 피싱으로 판정



유지 보수 용이

- 신종 수법이 등장해도 단순히 키워드 업데이트만으로도 즉각 차단 가능
- 가벼운 AE 모델 재학습

운영 프로세스



테스트

AE + koBERT Pipeline

===== ● 정상 시나리오 (정탐 테스트) =====		
[<input checked="" type="checkbox"/> SAFE]	오늘 미세먼지 심한데 산책은 무리겠지? 그냥 집에서 영화나 보자.	L 분석: 일상 대화(Safe-list) (L:7599.88 / P:0.9%)
[● WARNING]	고객님, 신한은행입니다. 요청하신 통장 사본 이메일로 발송해 드렸습니다.	L 분석: 의심 정황 포착 (L:10593.59 / P:0.96%)
[<input checked="" type="checkbox"/> SAFE]	엄마 나 오늘 늦어. 김치찌개 남은 거 데워 먹을게 맛있게 먹어!	L 분석: 일상 대화(Safe-list) (L:6427.62 / P:1.92%)
[<input checked="" type="checkbox"/> SAFE]	주말에 축구 경기 보러 가기로 한 거 티켓 예매 완료했어.	L 분석: 일상 대화(Safe-list) (L:6191.59 / P:1.89%)
[● WARNING]	본인 확인을 위해 생년월일 6자리만 말씀해 주시겠습니까?	L 분석: 의심 정황 포착 (L:12097.4 / P:1.04%)
===== ● 피싱 시나리오 (수법별 복합 공격) =====		
[⚠ CRITICAL]	정부지원금 혜택 대상입니다. 기존 대출 상환하셔야 추가 대출 가능합니다.	L 분석: 핵심 위험 키워드 탐지 (강력 차단) (L:6835.85 / P:0.79%)
[⚠ CRITICAL]	금융감독원입니다. 본인 명의 계좌가 범죄 연루되어 국고 환수 예정입니다.	L 분석: 핵심 위험 키워드 탐지 (강력 차단) (L:5761.46 / P:1.07%)
[⚠ CRITICAL] [해외결제]	950,000원 승인 완료. 본인 아니면 상담원 연결 후 원격 제어 받으세요.	L 분석: 핵심 위험 키워드 탐지 (강력 차단) (L:10151.43 / P:1.73%)
[⚠ CRITICAL]	나 핸드폰 고장나서 수리비 급해. 편의점에서 기프트 카드 사서 핀번호 보내줘.	L 분석: 핵심 위험 키워드 탐지 (강력 차단) (L:9216.72 / P:1.17%)
[⚠ CRITICAL]	수사 협조 안 하시면 구속 수사 진행됩니다. 국가 안전 계좌로 예치하세요.	L 분석: 핵심 위험 키워드 탐지 (강력 차단) (L:6470.27 / P:0.97%)

피싱 탐지의 3단계		
	SAFE	안전한 대화
	Warning	주의
	Critical	피싱 탐지

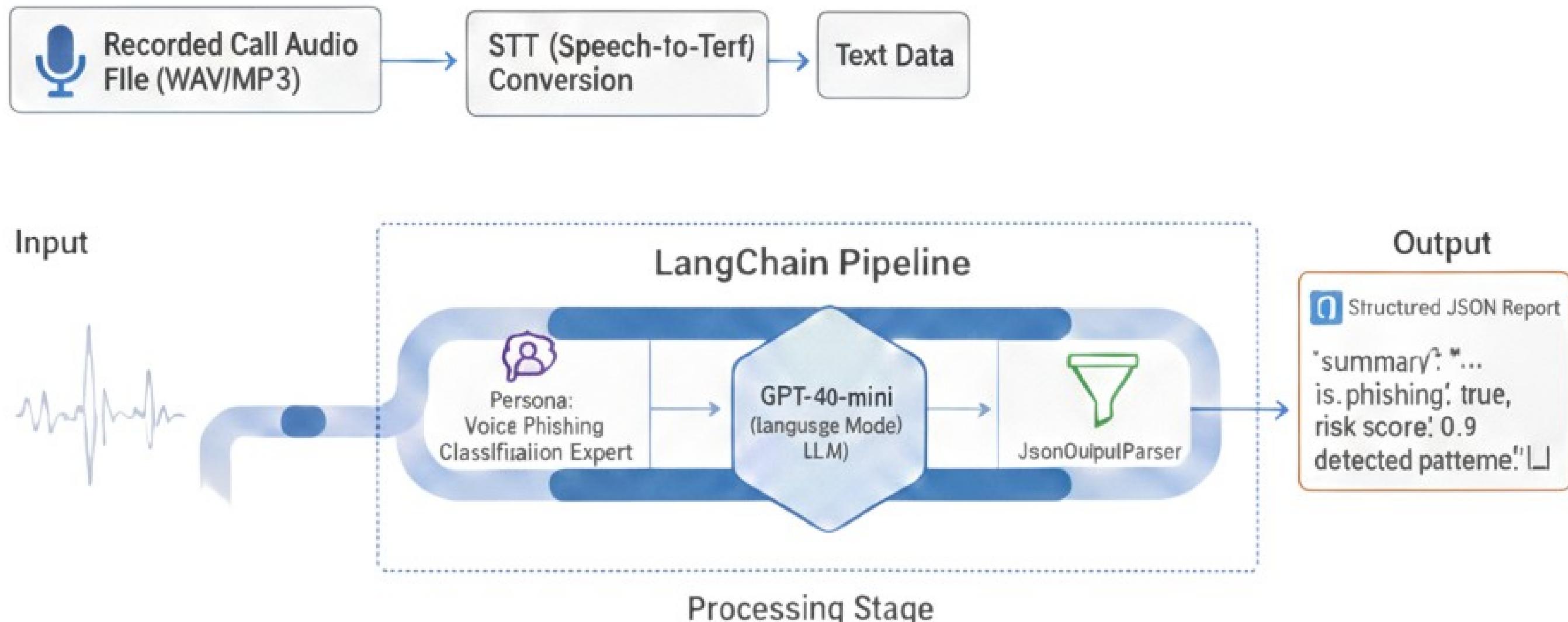
CONTENT



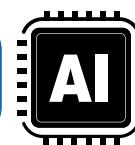
통화 종료 이후

리포트

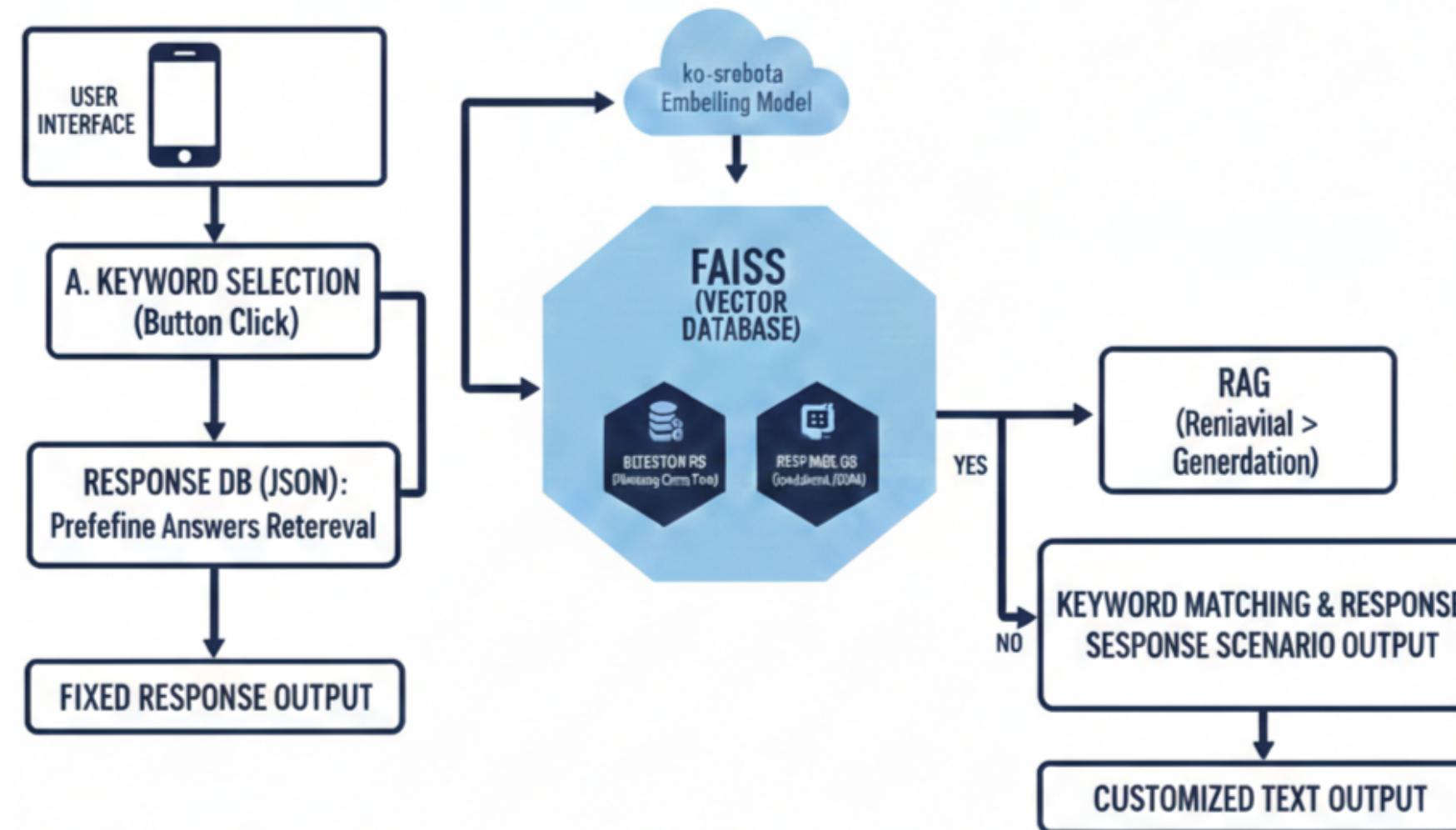
LangChain + OpenAI GPT-4o-mini



챗봇



LangChain + jhgan/ko-sroberta-multitask

**RAG + VectorDB**

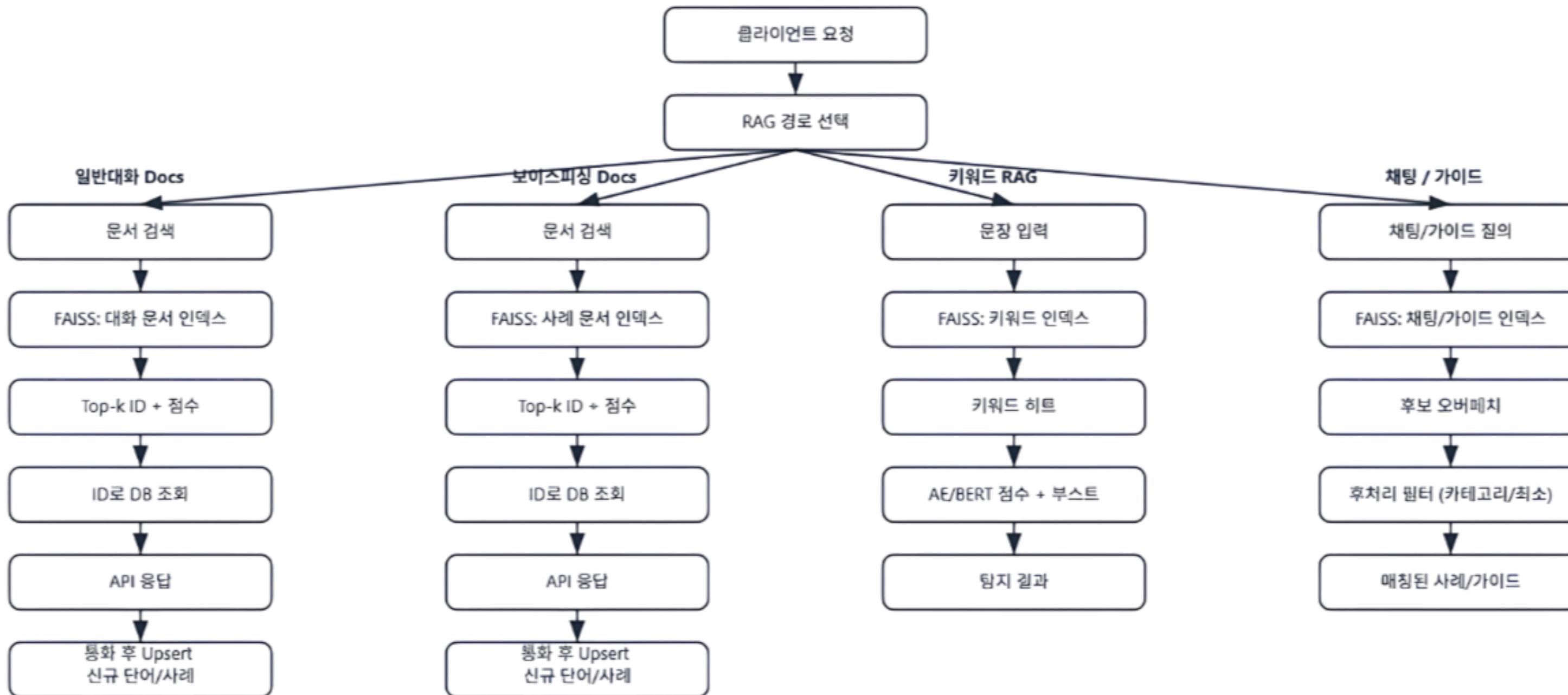
- 리포트 기반 키워드 버튼 생성
- 사용자 질문에 대한 내용을 벡터DB에서 탐색 후 유사한 통화 내용을 찾아 답변 제공

DB

VectorDB

RAG 흐름 (FAISS)

FAISS가 후보를 빠르게 찾고 DB/스코어링 로직이 결과를 정제합니다

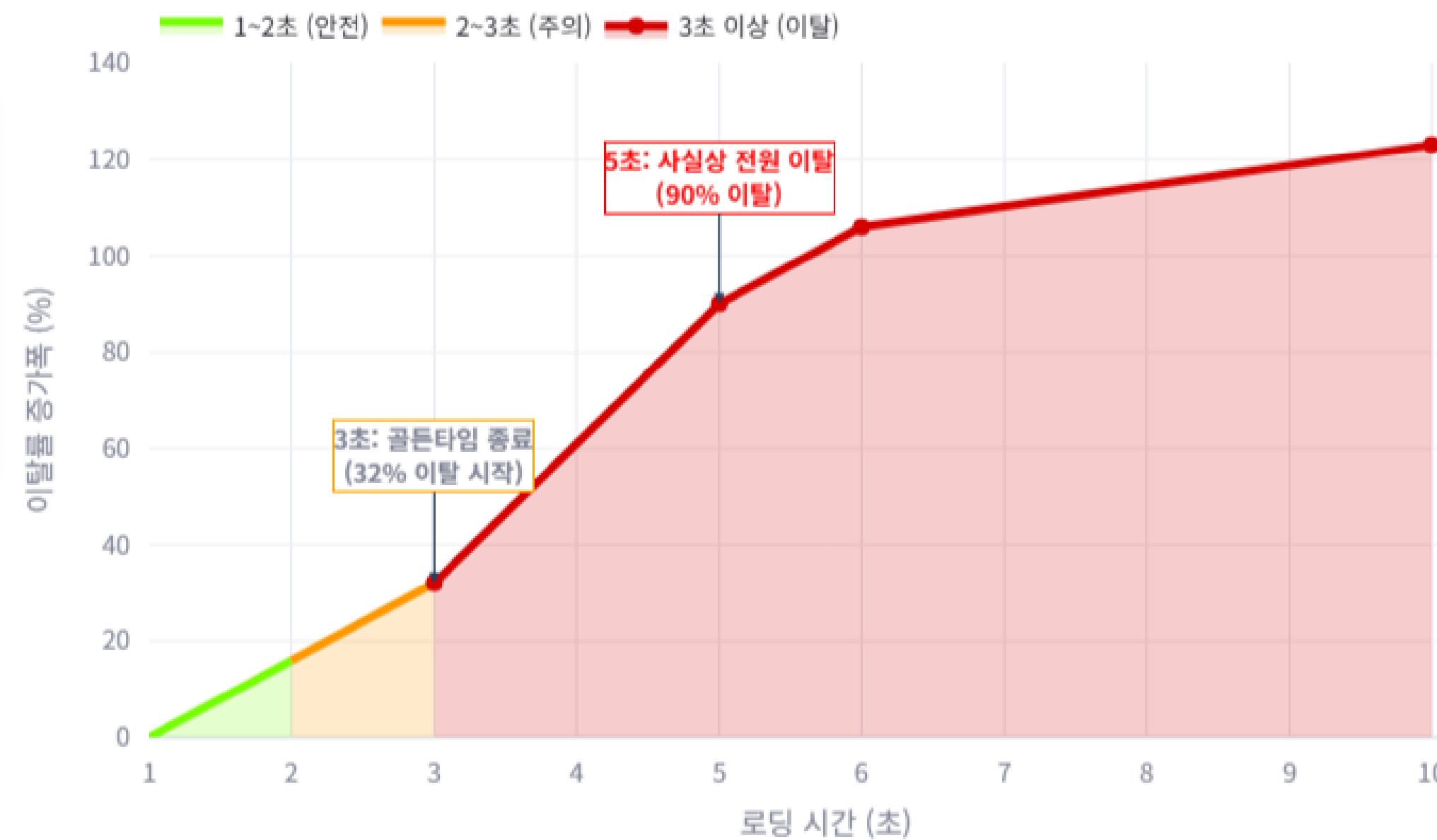


서버



Python

로딩 시간에 따른 이탈율 증가 추이 (NVISAGE)



고려사항

- API는 빨라야 한다.
- 사용자는 어떤 기술 스택을 사용하는지 신경쓰지 않는다.
- 응답속도가 1초라도 느려지면 사용자 이탈 가능성 발생 有

Python



Python

django**FastAPI**

Performance speed

Normal

Faster than Django

The fastest out there

Async support

YES
with restricted latencyNO
needs AsyncioYES
native async support

Packages

Plenty
for robust web appsLess than Django
for minimalistic appsThe least of all
for building web apps faster

Popularity

The most popular

The second popular

Relatively new

Learning

Hard to learn

Easier to learn

The easiest to learn

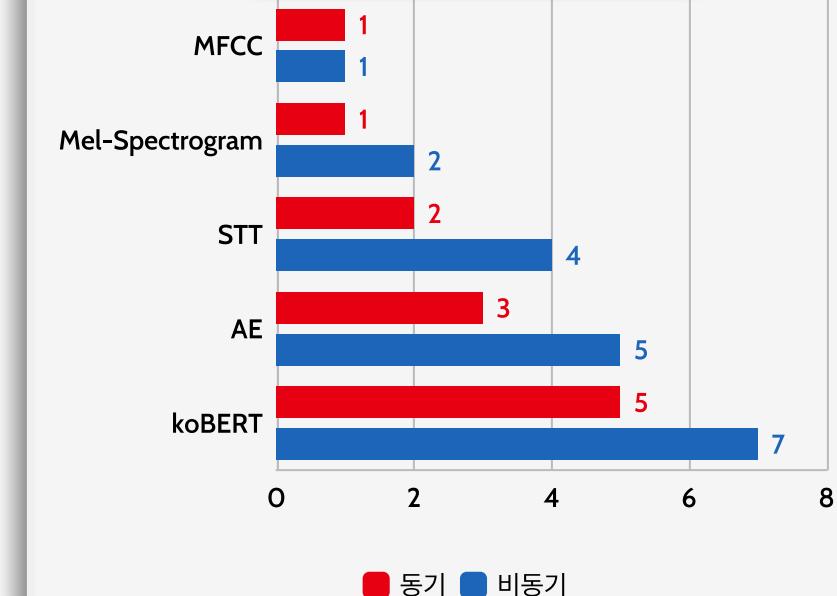
모바일



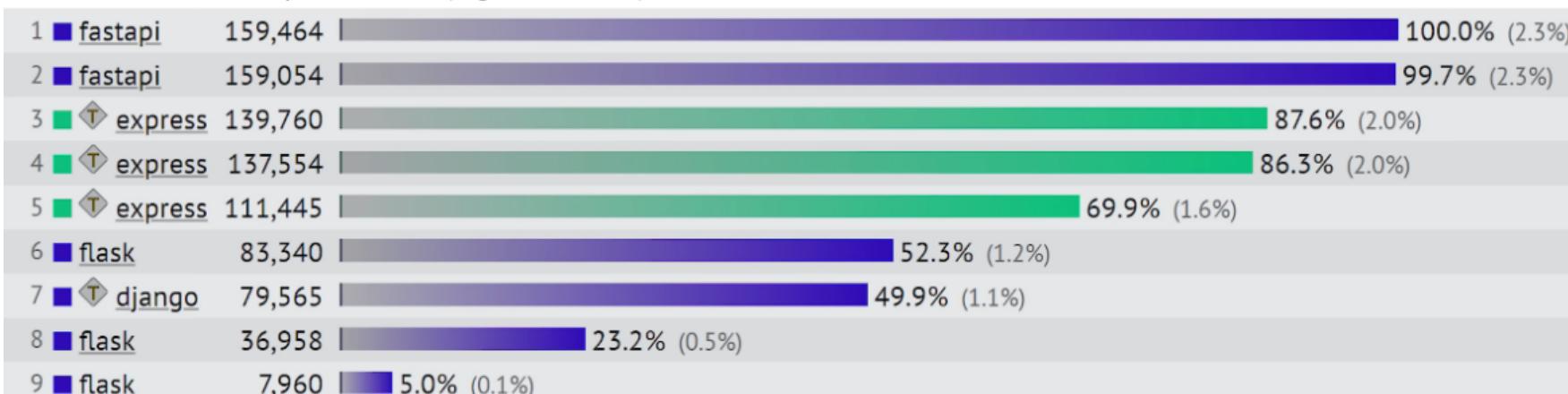
Python

Model	실행 속도(s)
MFCC	1.0
Mel Spectrogram	1.0
STT	2.0
AE	1.0
koBERT	2.0

동기/비동기 차이



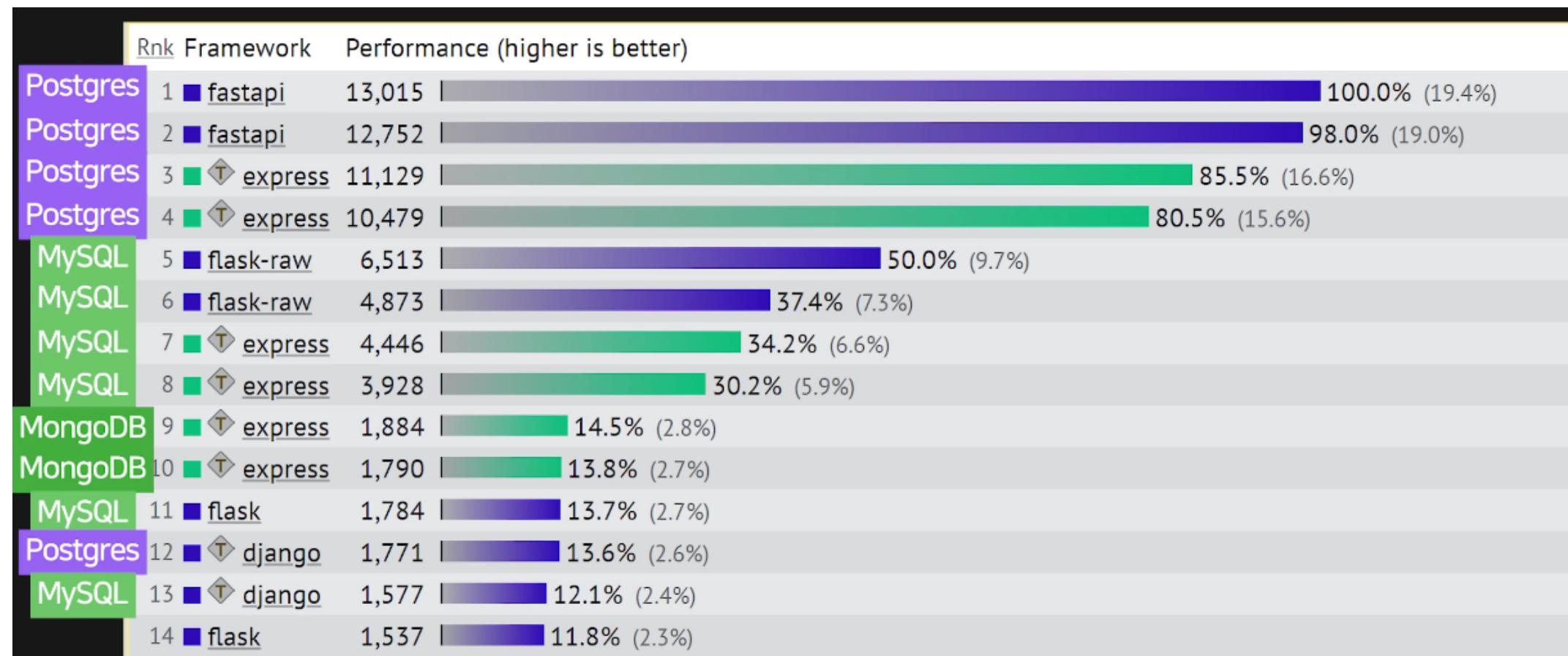
Rnk Framework Best performance (higher is better)



서버 → 유저 문자 전송량/sec

DB

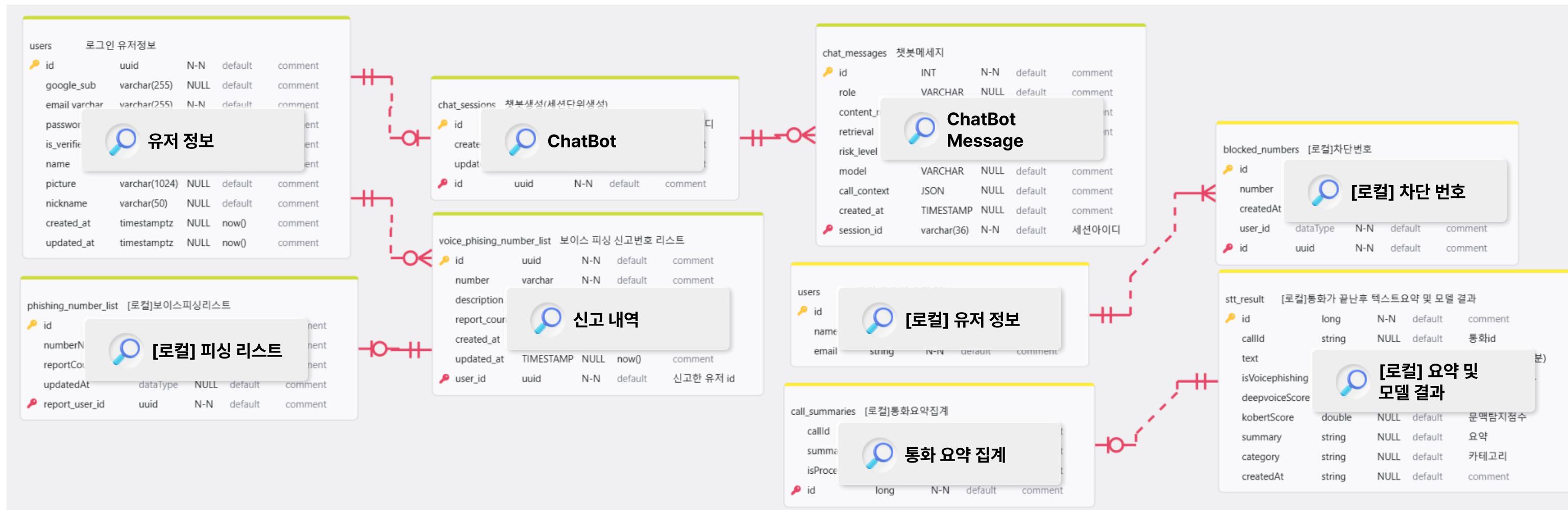
PostgreSQL



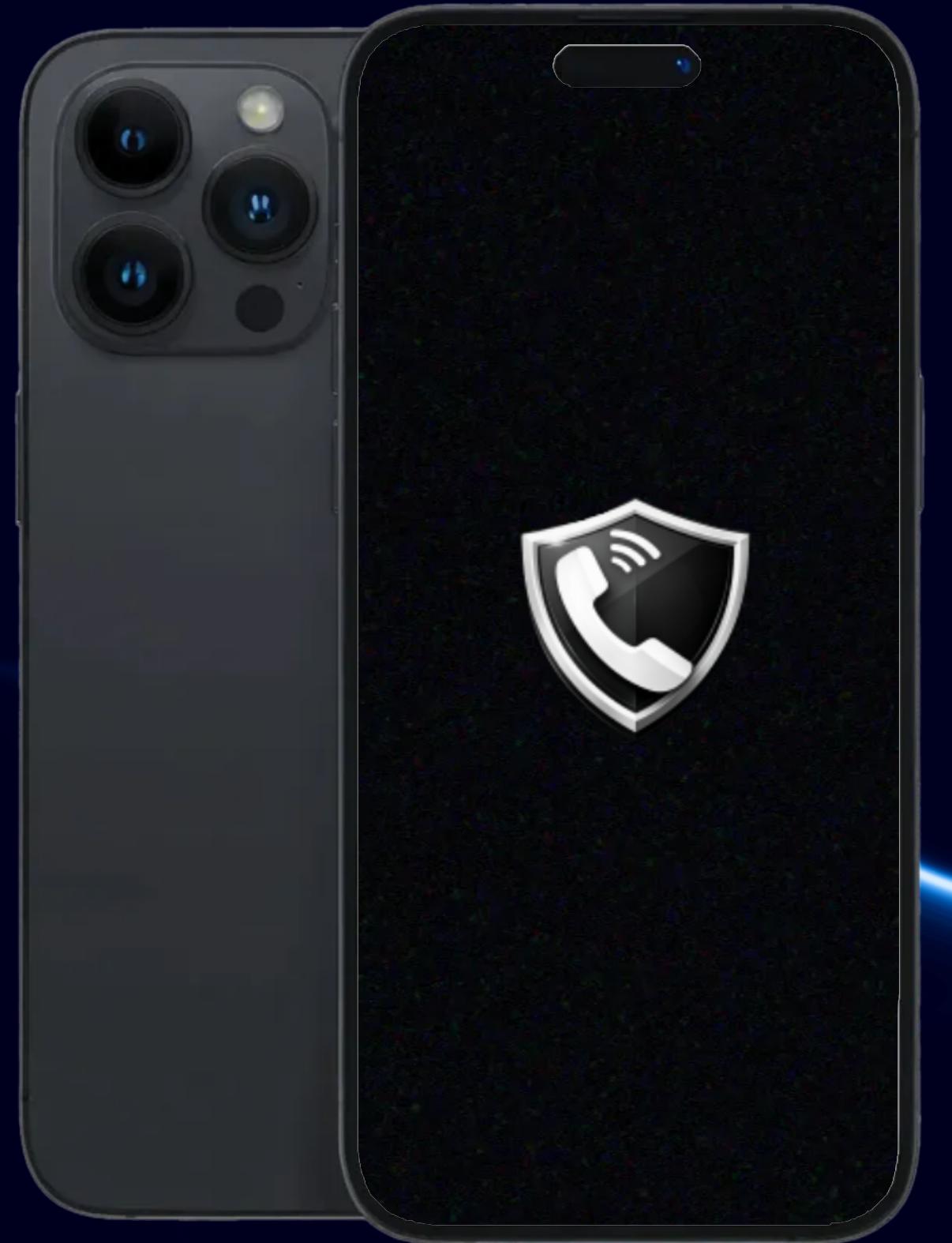
DB 데이터 출력량/sec

DB

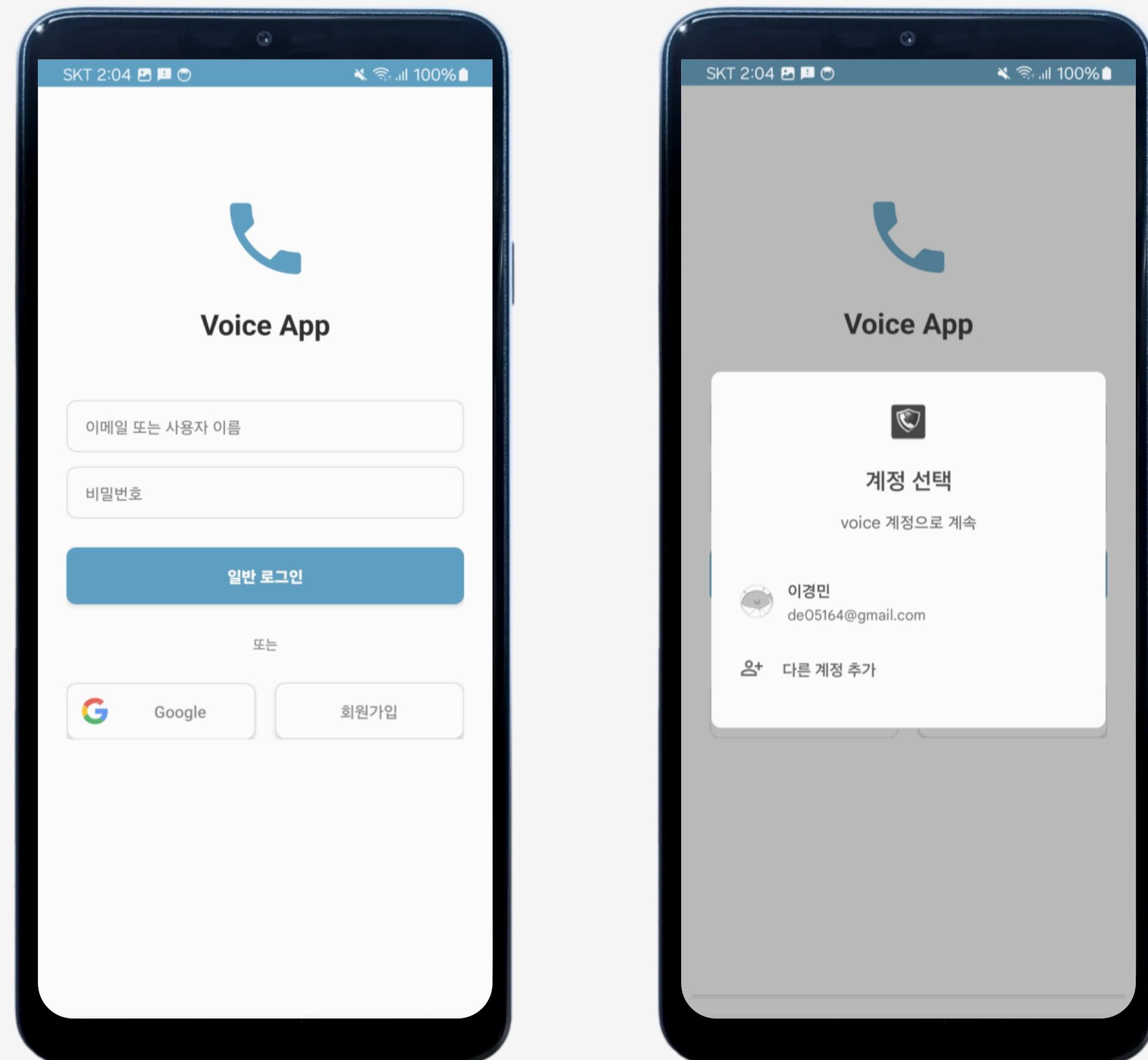
ERD Diagram

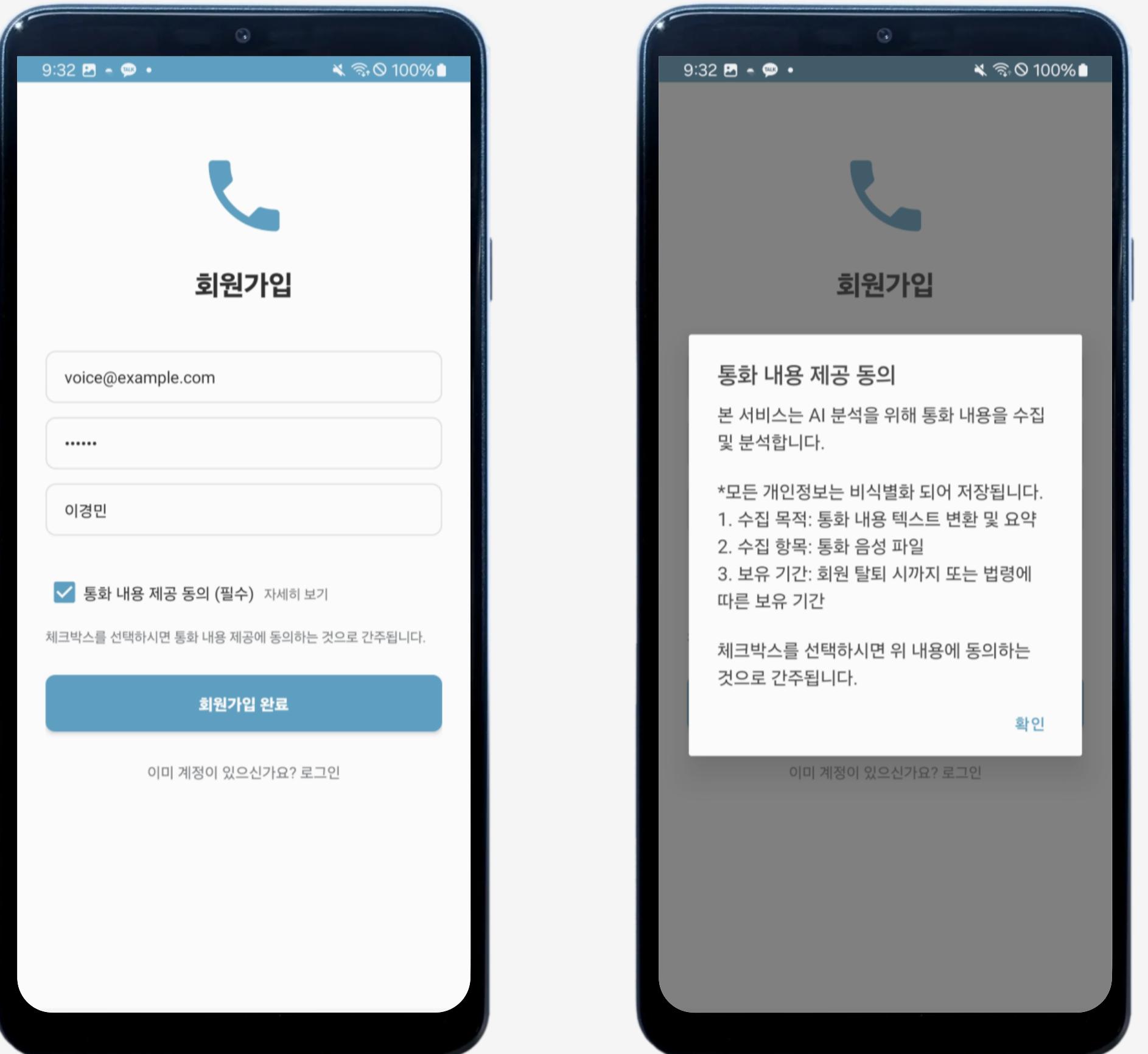


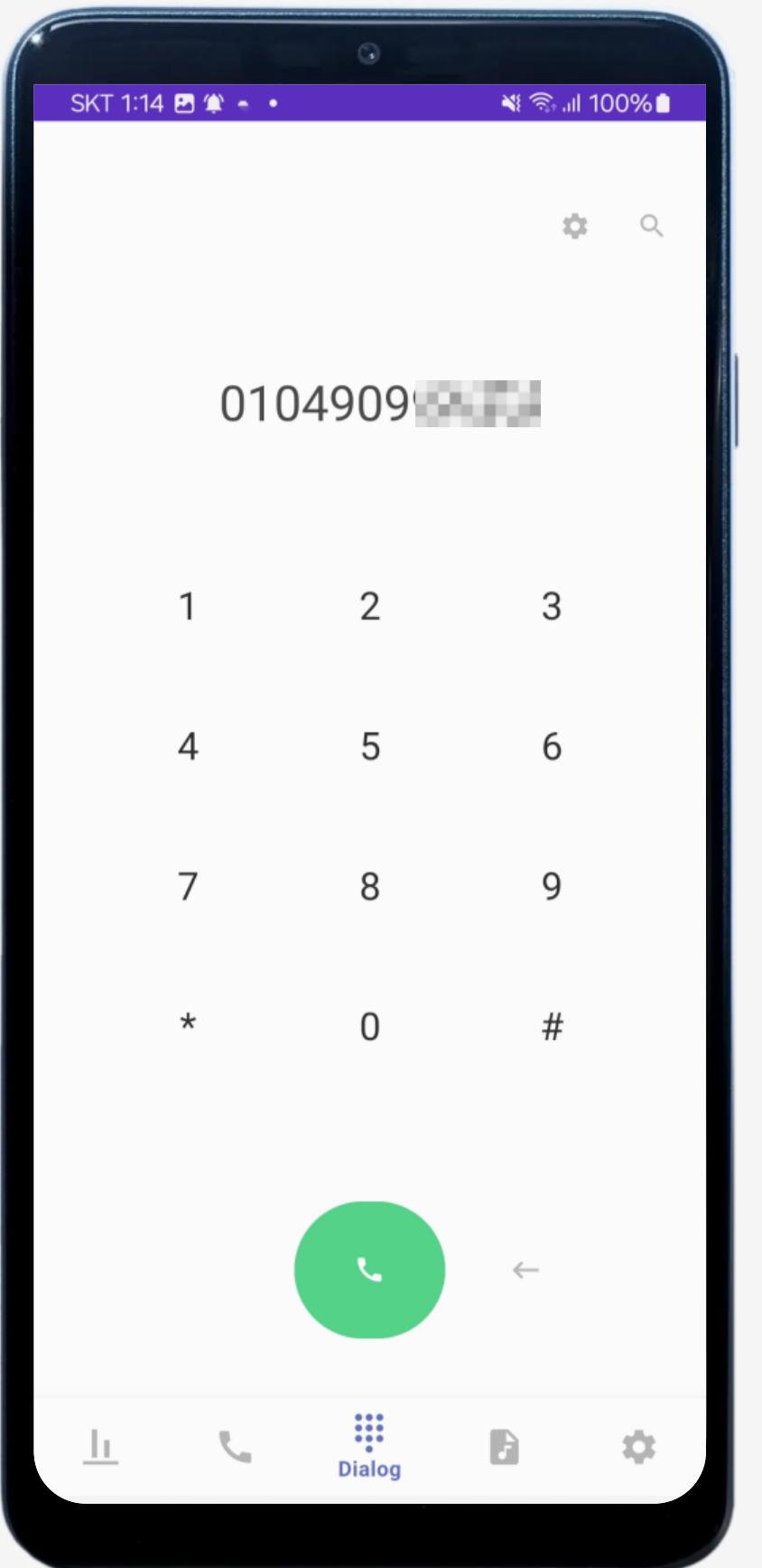
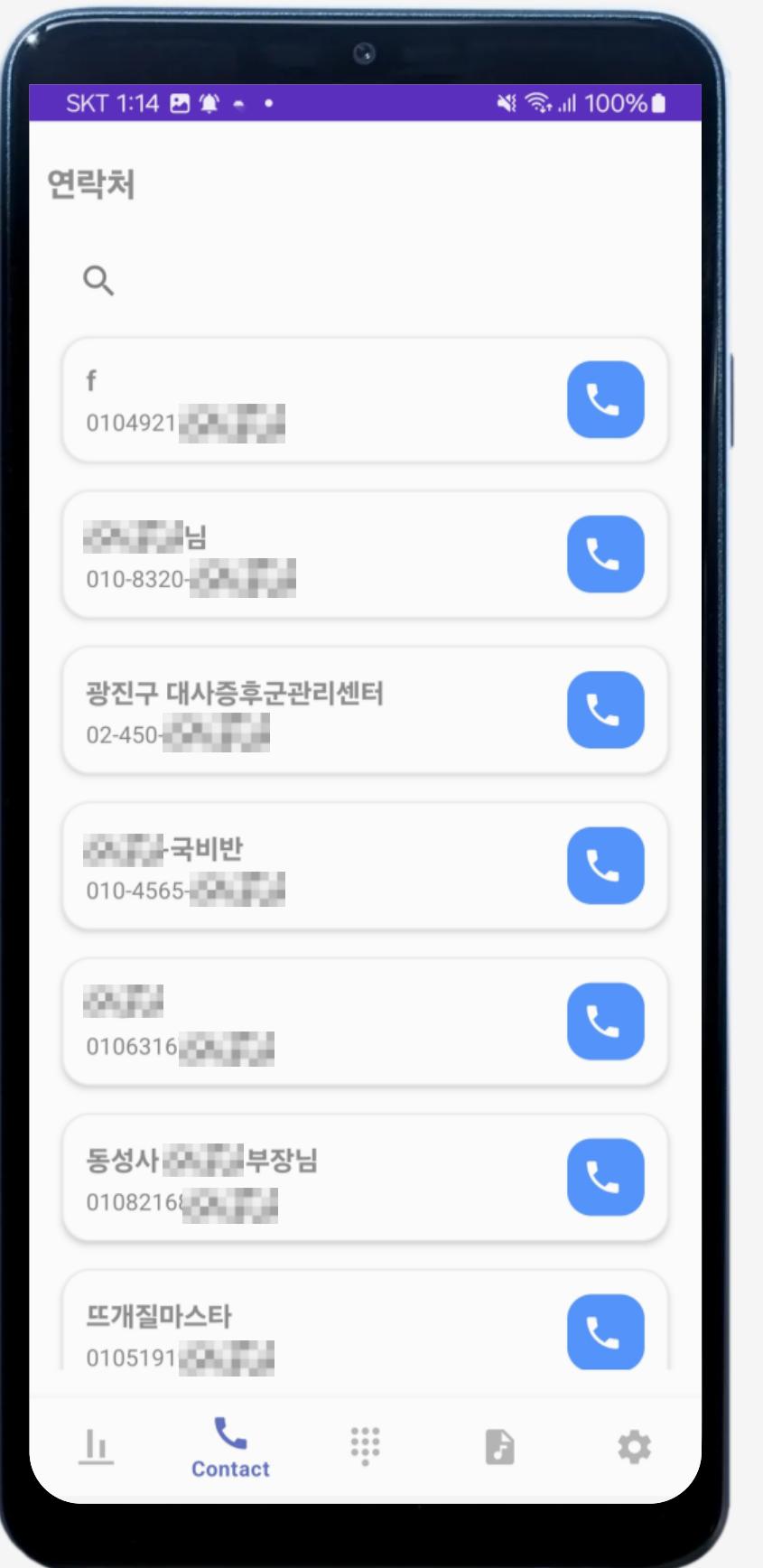
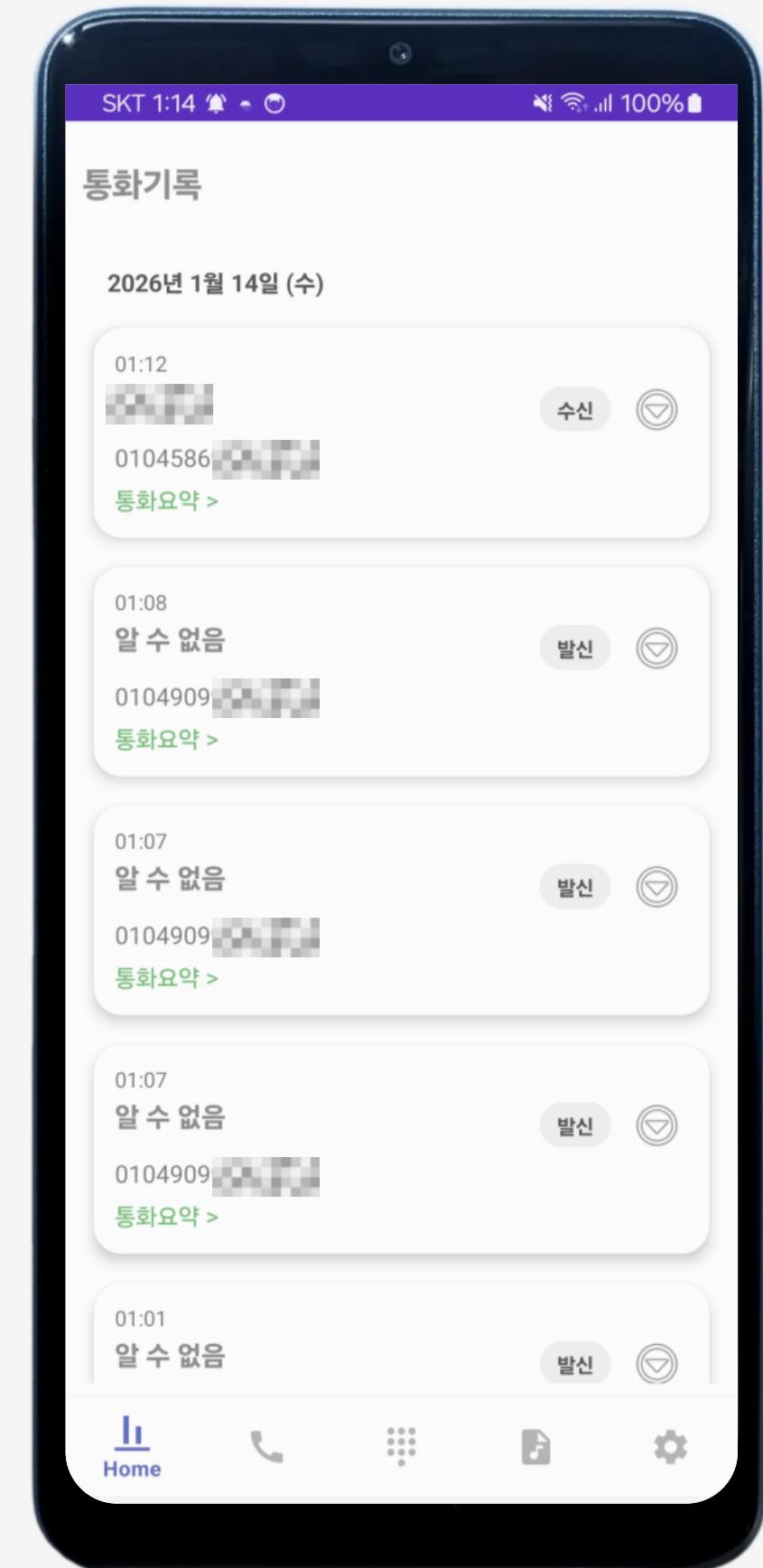
CONTENT

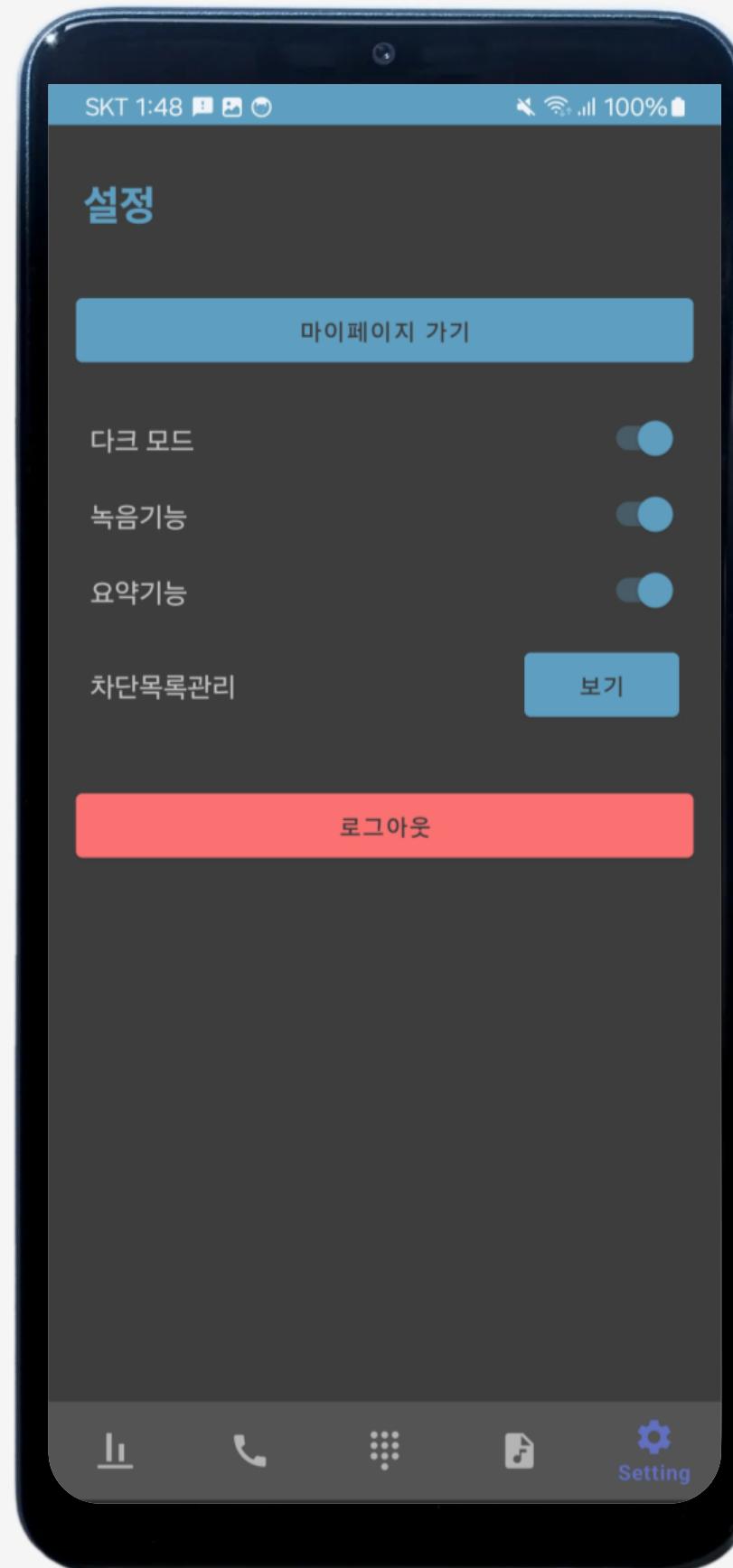
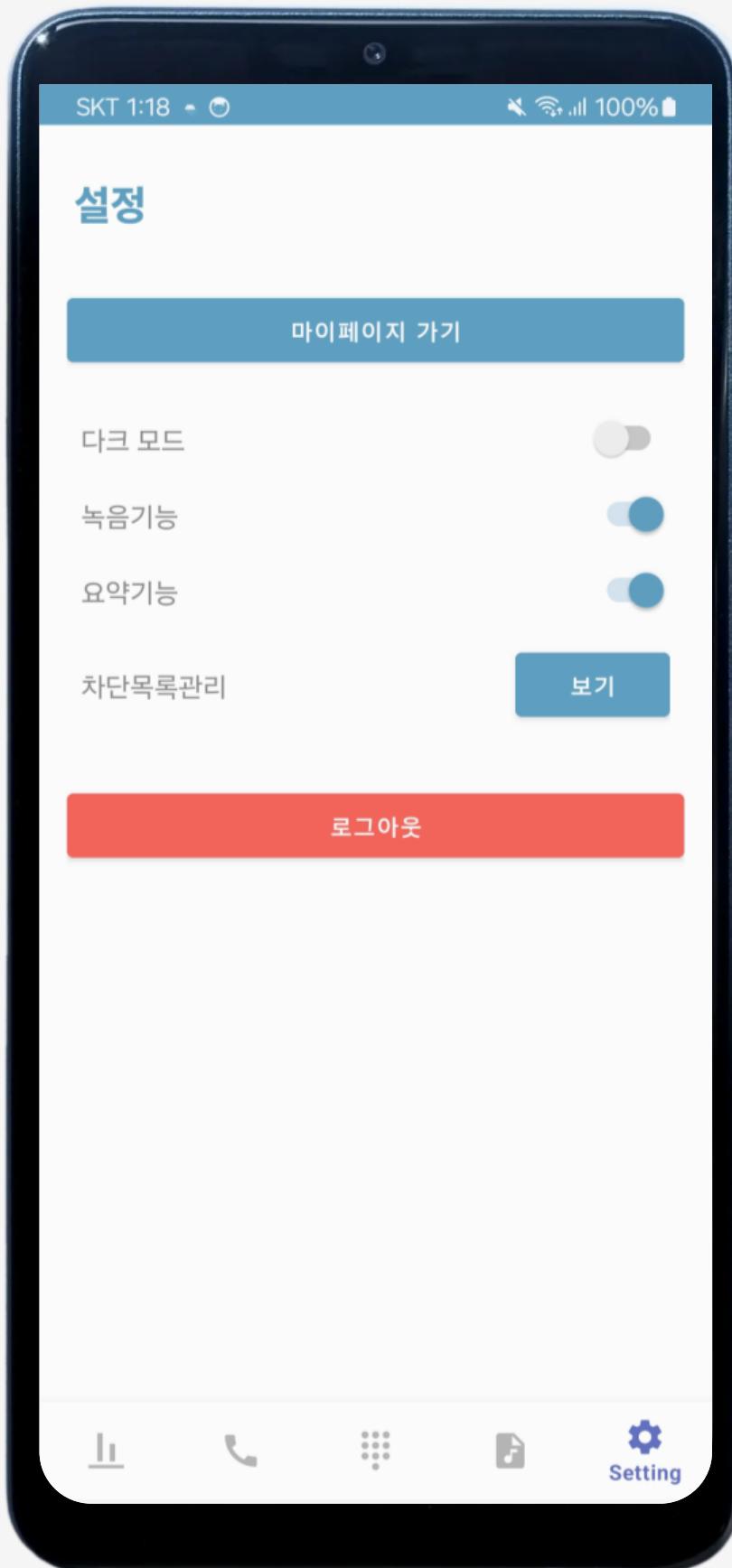
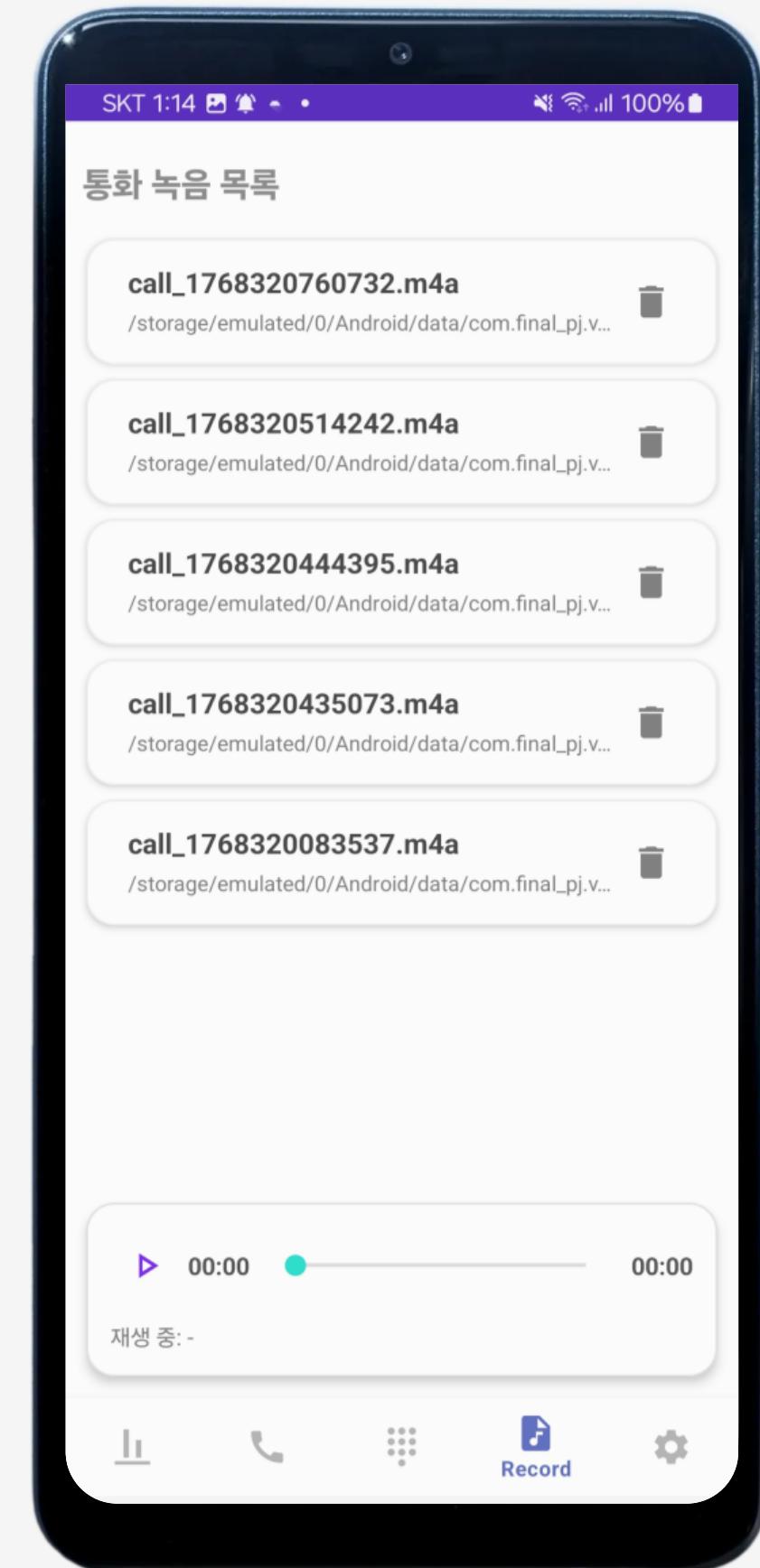


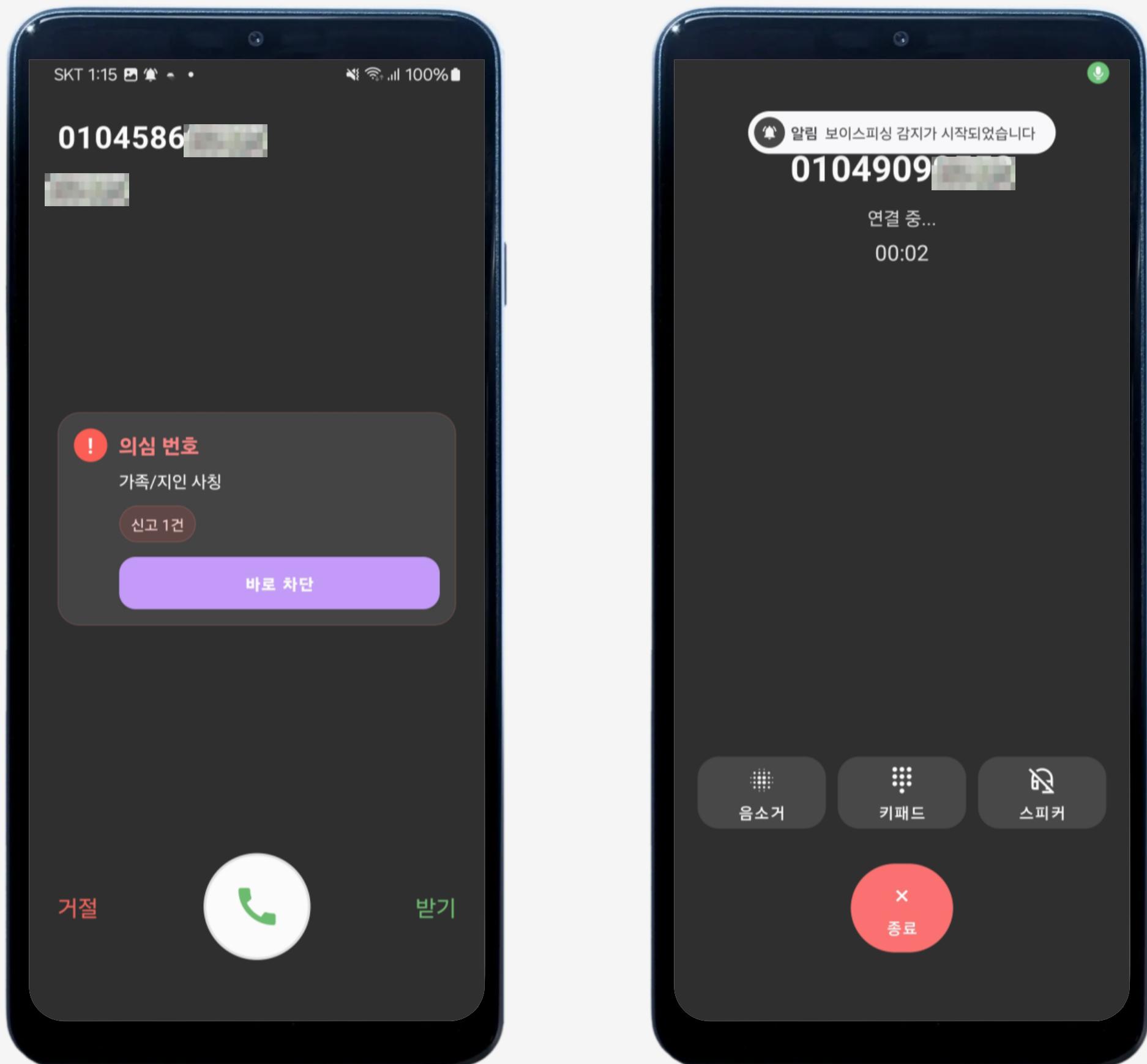
UI/UX



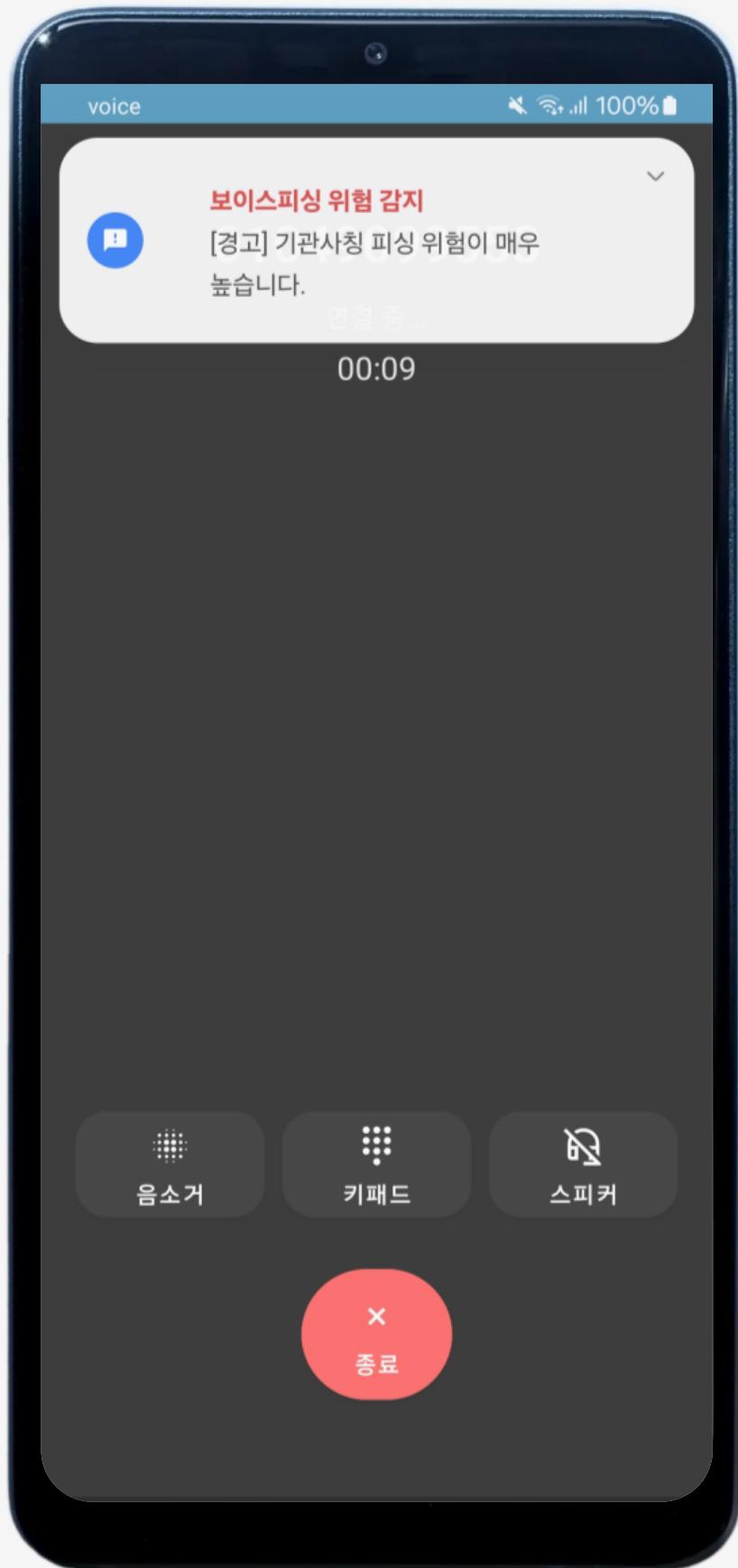
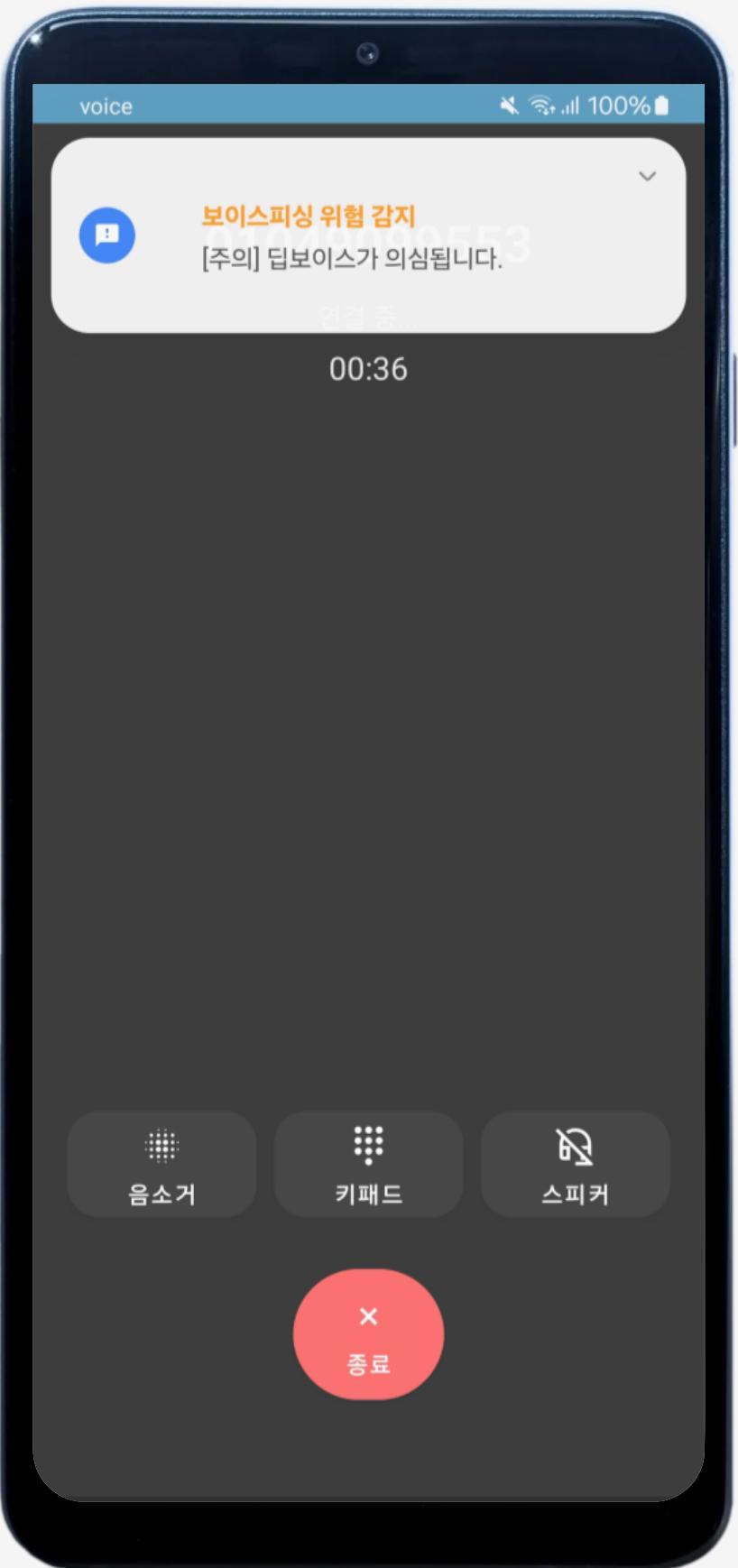
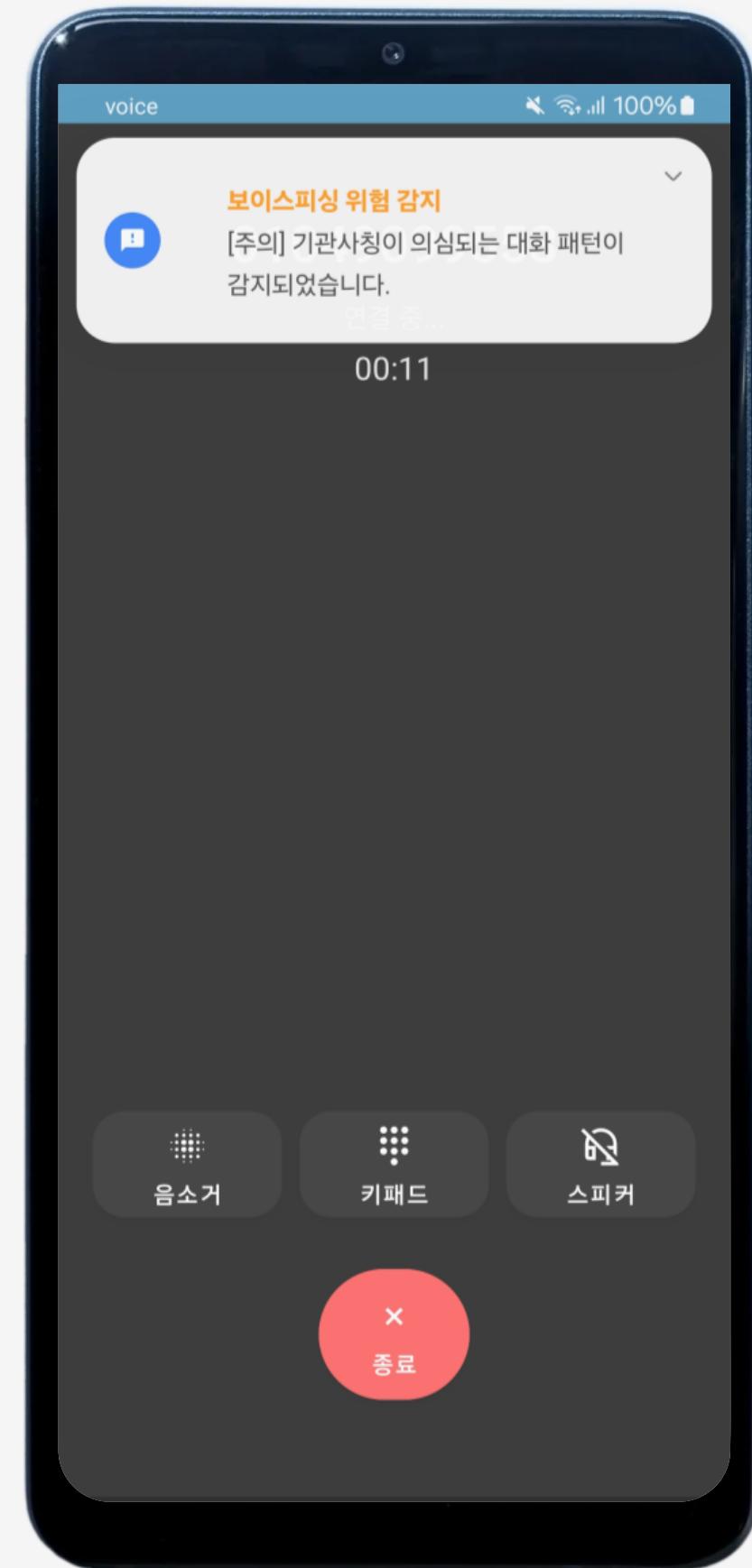


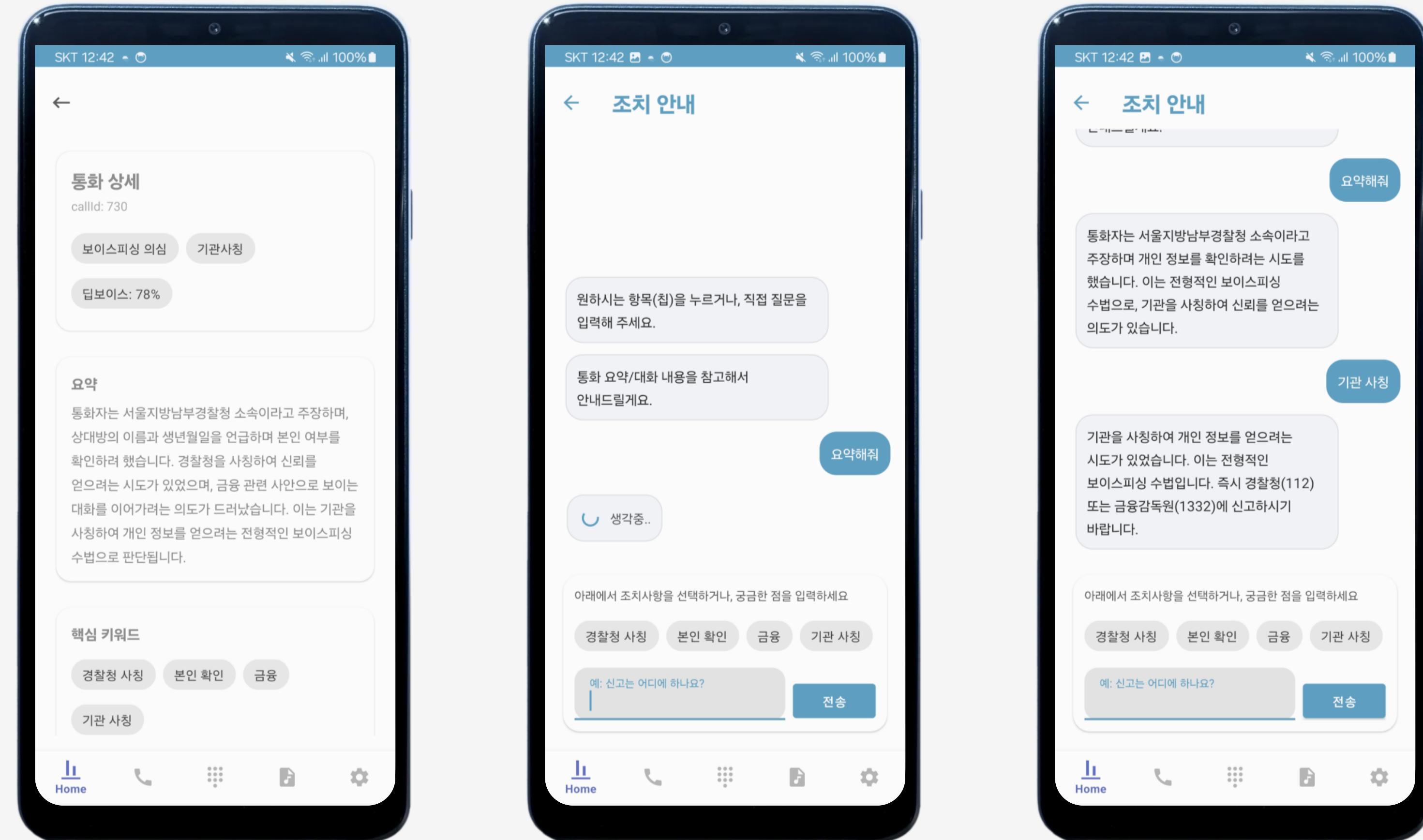


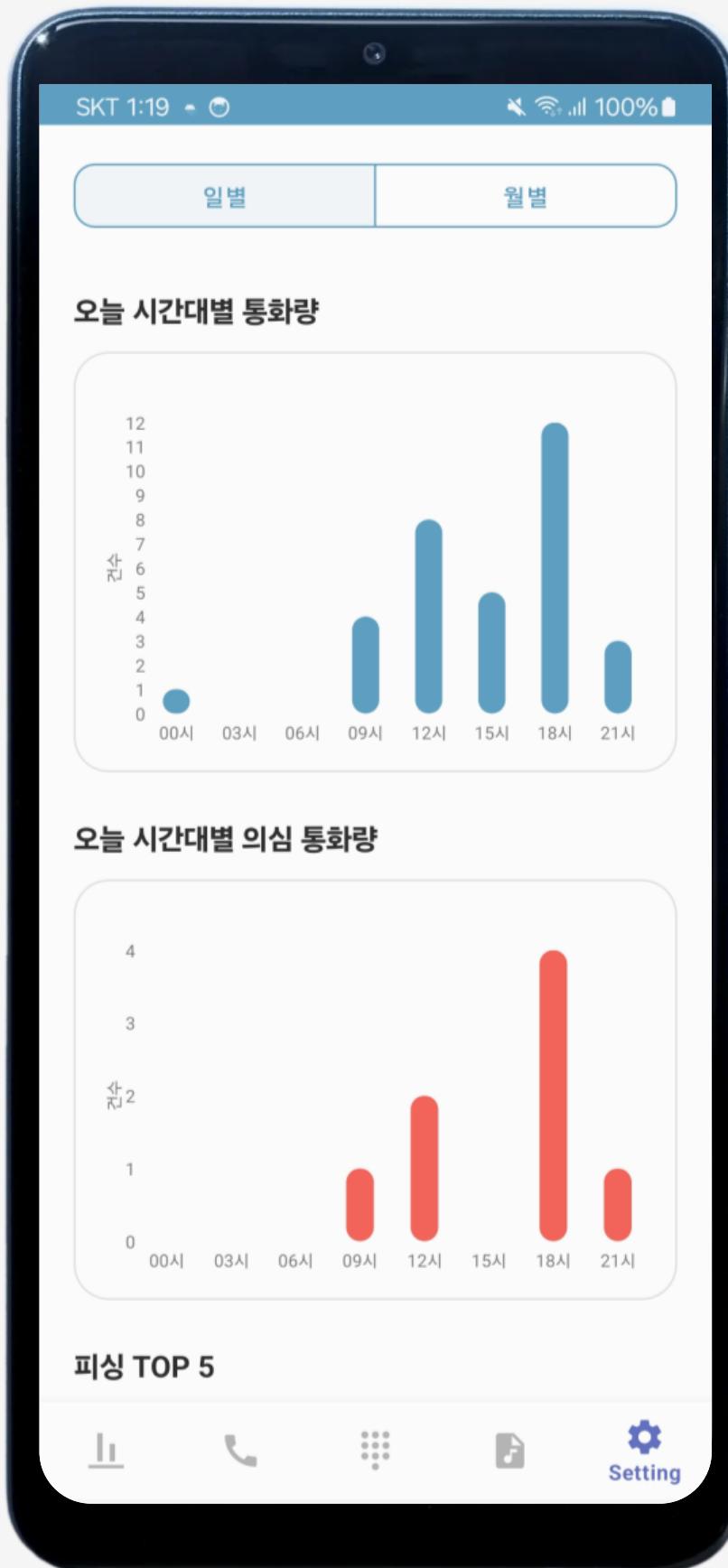


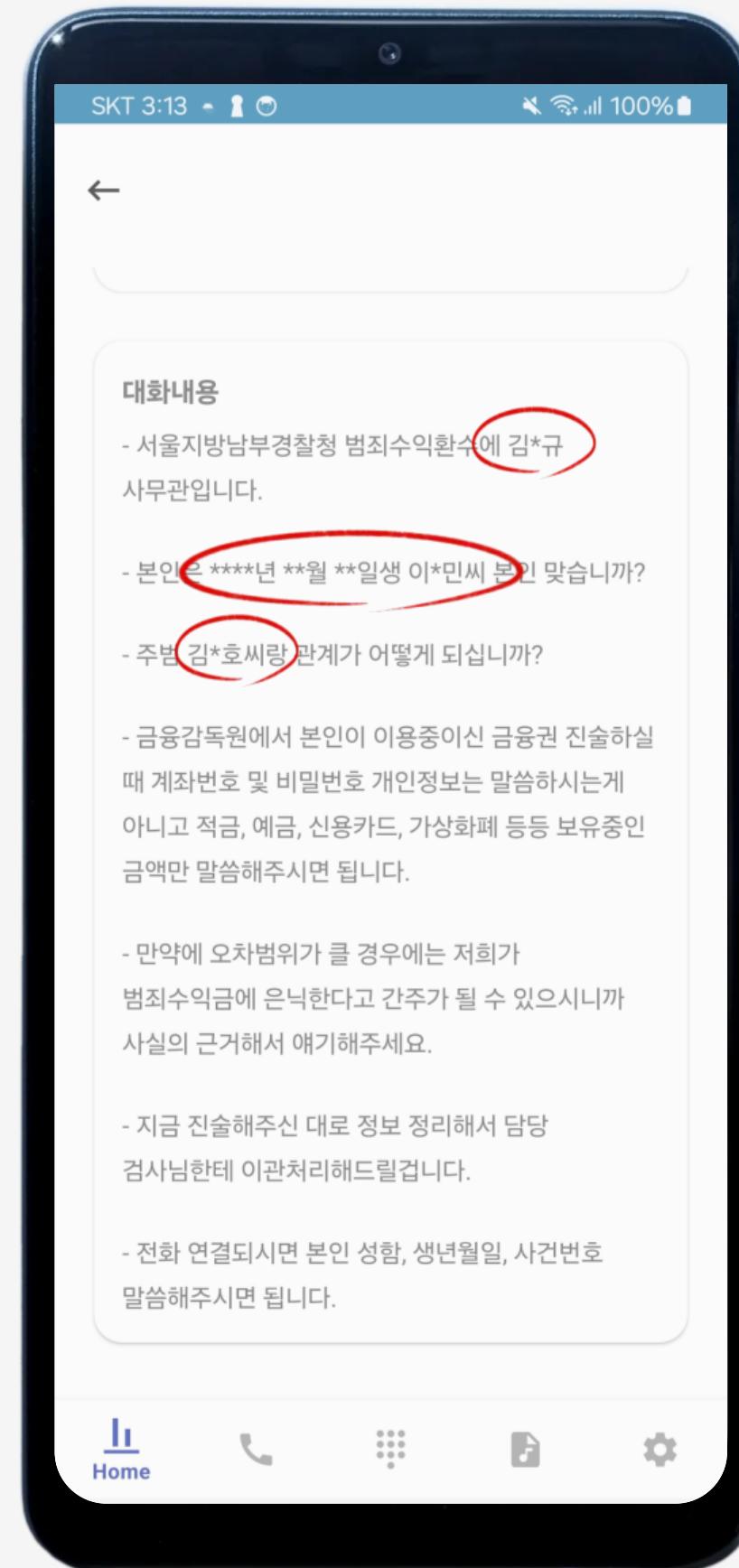
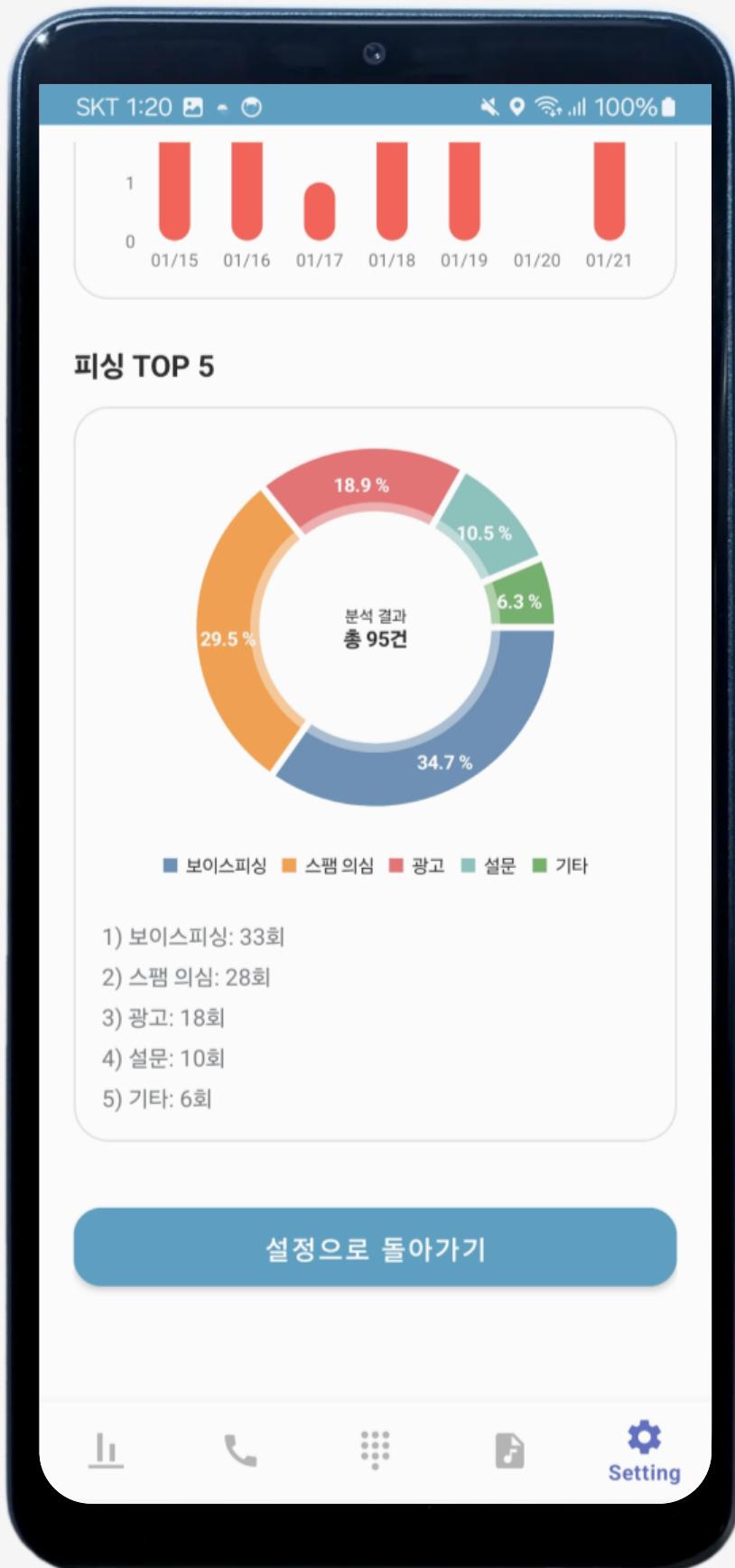
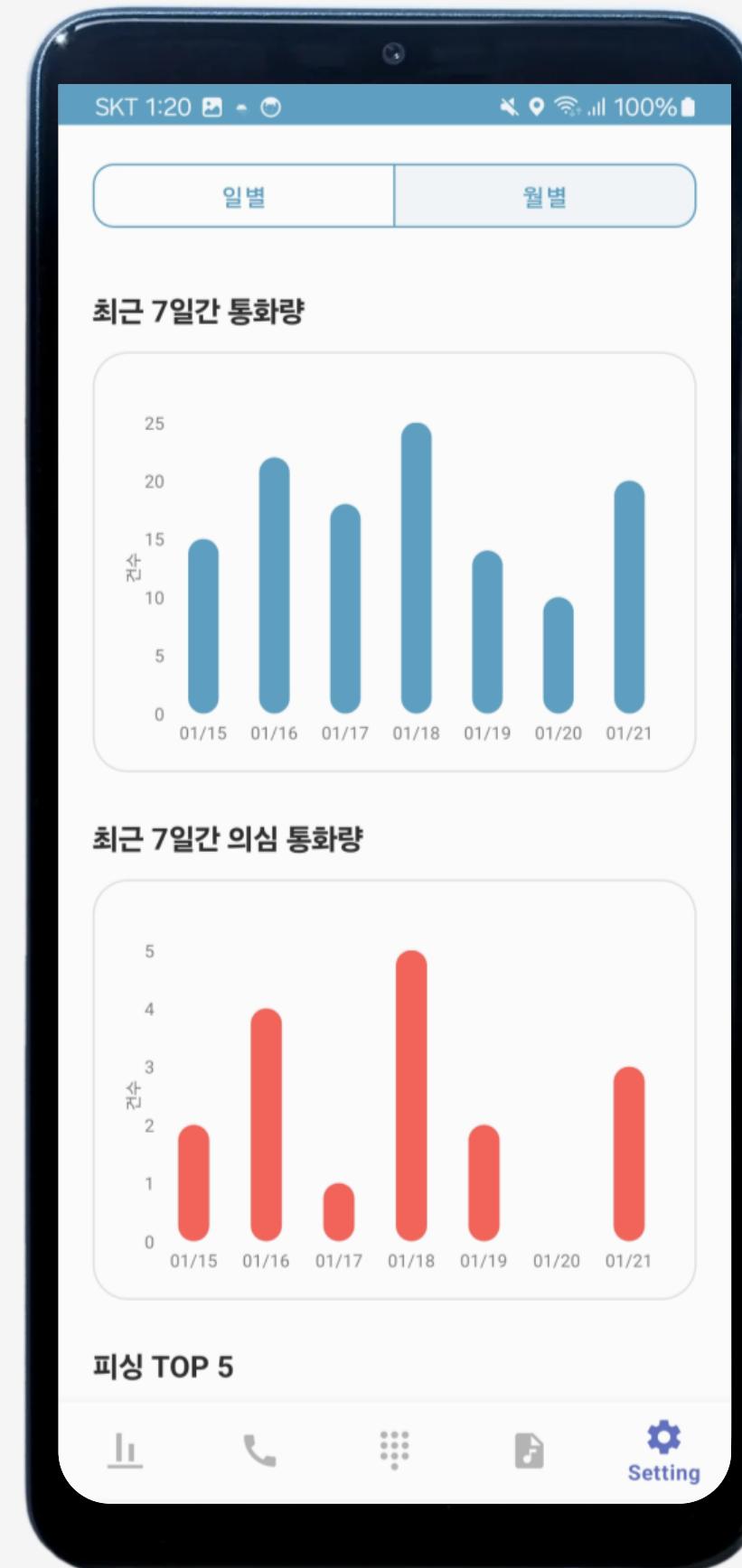


UI/UX









보완 필요점



데이터

- 더 넓은 범주의 데이터 확보
- 다양한 증강 기법 비교
- 전처리 기법 비교



더 깊이있는 분석

- 데이터 기반의 더 깊이 있는 분석을 통해 새로운 인사이트를 도출하고 이를 활용한 실용적인 결과 산출 기대



AI 감성 분석

- 피싱범의 압박적인 말투, 공격적, 권위적인 태도를 이용하여 보다 더 정밀한 피싱 탐지 효과 기대

우리 기술의 가치 및 기대 효과

✓ 정성적 기대효과

- ✓ 보안 인식 증대와 금융 사기 예방 의식 강화
- ✓ 사회적 안정망 강화

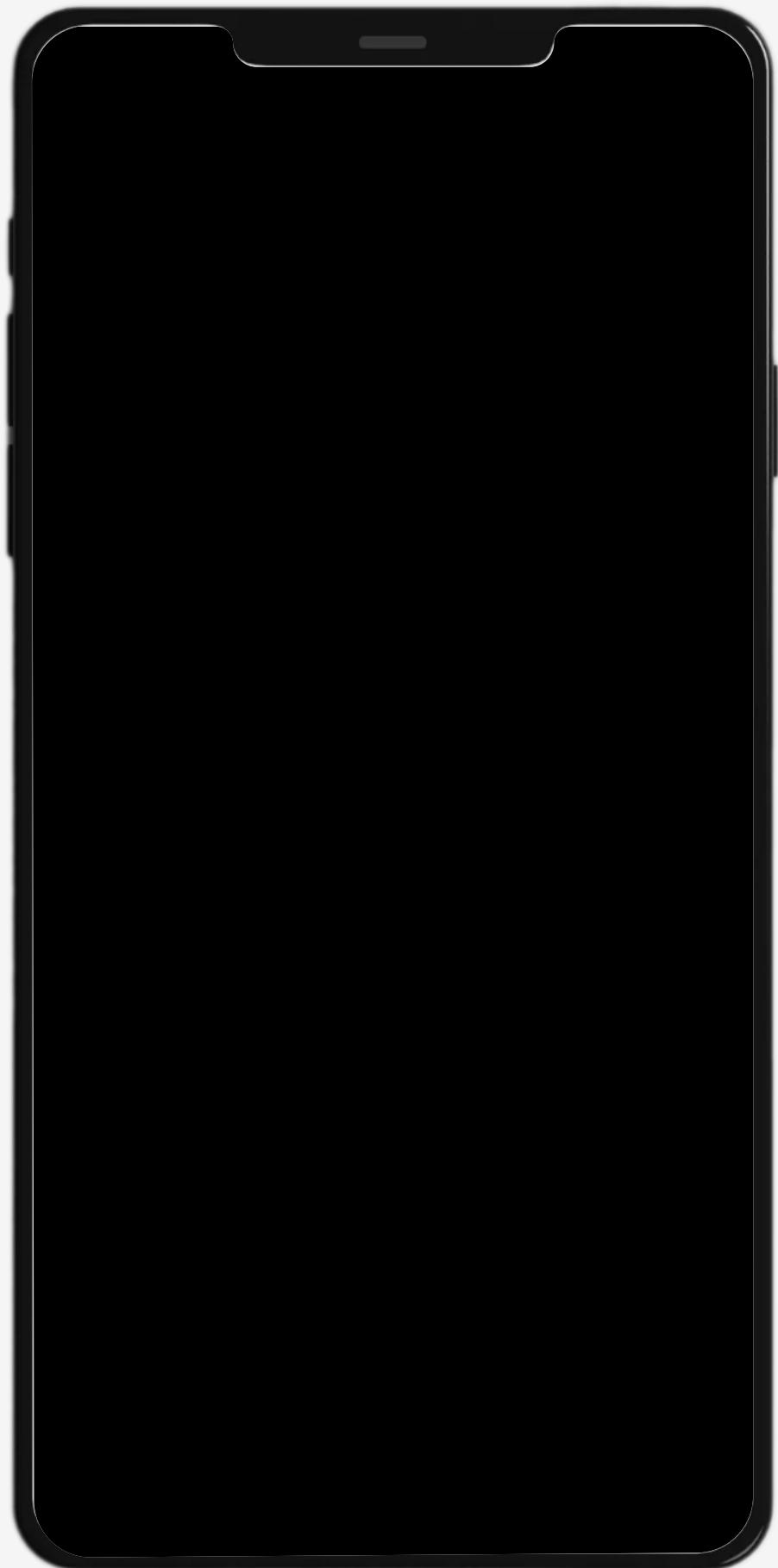
✓ 정량적 기대효과

- ✓ 높은 보이스피싱 탐지 성공률을 통한 피해 감소
- ✓ 범죄 조사 등의 사회적 비용 절감

- 직관적 UI 및 실시간 경고로 사기 피해 사전 예방
- AI보안 솔루션 도입으로 금융 보안 기술 혁신 가속화
- 탐지 데이터로 범죄 예방 협력 확대

- 보이스피싱 1건 탐지시 8,071만원의 사회적 비용 절감
(직접 피해액 5,381만원 × 사회적 비용 계수 1.5배)
- 경찰의 검거는 피해금을 돌려받기 어렵지만,
AI의 탐지는 피해액 전액을 보존
- 단순한 "범죄 예방"을 넘어 "금융 자산 보호 솔루션" ✓

나연



Q&A

Github



Dashboard



가용 데이터

[AI-Hub] 한국인 대화, 다화자 음성, 자유 대화,
한국어 대화 요약, 민원 콜센터 질의응답,
상담 음성, 저음질 전화망 음성
[Kaggle] ASVspoof2021, 2024
[RVC] 한국어 대화 음성 변조
[금융감독원] 실제 보이스피싱 대화
외 SNS, 유튜브 등 보이스피싱 시나리오 크롤링



https://github.com/dokpe01/voice_phishing_detection



<https://appdashboardgit-brkccx8rxbadykpepe68en.streamlit.app/>