

# 머신러닝 기반의 일 별 콜 수 예측을 활용한 콜센터 인력 스케줄링 최적화

김지현\*, 박상준\*

\*JB 주식회사

e-mail : kjihyun@jbcorporation.com, psjzone@jbcorporation.com

## Optimization of call center staffing problem scheduling using machine learning-based daily call count prediction

Ji-Hyun Kim\*, Sang-Jun Park\*,

\*JB Corporation

### 요 약

콜센터에서 인력 스케줄링은 매우 중요하다. 모든 콜센터에서 인건비는 고정비 성격이 강하여 차지하는 비중이 매우 높아 콜센터의 이익을 좌지우지한다. 그렇기 때문에 콜센터의 적정 인력의 고용과 배치는 인건비 뿐만 아니라 콜 성공률 또한 직결되어 있어 콜센터 운영에서 중요한 사안이라고 할 수 있다. 대부분의 콜센터가 현재까지도 관리자의 경험에 의해 인력배치를 수립하는데, 이러한 방법은 과학적이지 않으며 인원수에 영향을 미치는 모든 변수들을 고려할 수 없다. 과거 수학적 모델을 수립하는 것이었다면, 지금은 모델을 학습시키고, 학습된 모델을 기반으로 미래의 고객과 인원수를 예측해야 한다. 본 논문에서는 수리제약식을 통해 다양한 변수들을 고려하고 비선형 정수 계획법과 딥러닝 기반의 예측 값을 이용하여 비선형 정수계획법을 통해 최적의 인력배치 스케줄링을 수립하였다.

### 1. 서론

전국 도시가스사는 콜센터의 역할이 매우 중요하다. 단순한 진출, 전입과 같은 도시가스 내적의 역할 뿐만 아니라 고객의 VOC 처리 등 도시가스사와 고객을 이어주는 소통의 창구 역할까지 하고 있기 때문이다. 이 때문에 고객의 콜을 처리하는 성공률 관리 또한 중요하다. 고객을 위해 콜 성공률 100%를 달성하려면 그만큼의 인원을 배치하면 된다. 하지만 콜센터에서 인건비가 차지하는 비중은 높기 때문에 어느 시간대에 몇 명의 근무 인원을 배치하여 콜을 처리하느냐에 대한 문제는 콜센터의 손익이 달려있을 만큼 중요한 일이다. 현재 추세는 대면 서비스에서 비대면 서비스의 비중을 높이고자 모든 산업이 노력하고 있는데, 콜센터 또한 예외는 아니다. 대부분의 도시가스사가 그렇듯 플랫폼을 운영하고 있으며, 인건비 절감과 효율적인 운영을 위해 24시간 처리가 가능한 비대면 서비스인 플랫폼으로 점차 비중을 늘려가야만 한다. 효율적인 인력 배치 스케줄링의 아이디어는 간단하다. 고객의 콜 수요에 대응해 목표 성공률을 유지하면서 오버스타핑(Over staffing)이 안되게 하면 되는 것이다. 핵심 아이디어를 사용해 모든 변수를 고려하여 콜센

터를 운영하는 위탁업체에서 쉽게 인력 배치가 만들어지지 않는다. 그렇기 때문에 데이터를 이용한 인력 스케줄링 문제를 해결해야 한다.

본 논문에서는 전통적인 최적화 방법인 비선형계획법을 이용한 스케줄링을 진행한다. 시간대에 따라 변하는 콜의 예측값 하에서 미리 지정해둔 최대 허용 포기율과 수리제약조건 식을 만족시키면서, 관리자를 포함한 상담원의 인건비를 최소화하는 인원수를 구하는 것을 목표로 한다. 하지만 전통적인 최적화 방법만을 이용하는 것이 아닌, 콜의 예측값 데이터를 머신러닝을 활용하여 과거의 콜 추세를 담은 결괏값을 사용하여 시행한다.

본 논문은 충남지역에 도시가스를 공급하고 있는 JB 주식회사의 콜센터 데이터를 이용해 연구를 진행하였으며, 데이터는 요일, 시간별로 분류하여 점심시간에는 인원 배치를 제외하였다. 또한 시스템 에러 등의 특수한 날로 보이는 이상 치들을 제거하였고, 1년간의 패턴을 학습하기 위해 공휴일이 있는 주는 제거한 데이터를 사용하였다.

2. 콜센터 인력스케줄링 최적화

2.1 정수계획모형 수립

2.1.1 제약조건 설정

본 연구에서 분석한 데이터는 JB 주식회사의 콜센터 데이터를 사용하였으며, 2015.01.01~ 2019.12.31 일까지의 총 5 개년 데이터이다. JB 주식회사의 콜센터는 인바운드 콜과 비교해 아웃바운드 콜의 비중이 매우 작기 때문에, 아웃바운드 콜 수를 제외한 인바운드 콜 수로만 분석에 사용하였다. 첫번째로, 비선형 정수계획모형을 수립하기 위해 5 년 치의 데이터의 기초 통계 결괏값을 기반으로 여러 제약조건을 설정하였다. 목적함수가 될 인건비는 상담원 한 명당 시간당 인건비를 산정하였다. 수리제약식에 들어갈 제약조건으로는 상담원의 평일 인원수는 18~25 명으로 제한하고, 토요일은 2~6 명, 일요일은 1~2 명으로 설정하였다. 목표 성공률은 JB 주식회사에서 비대면 서비스를 고려하여 90~92%로 현재 성공률보다 낮은 목표 성공률을 설정하였으며, 요일당 포기율은 월요일부터 일요일까지 각각 9.3%, 9.2%, 9.2%, 9.2%, 9.2%, 9.5%, 9.5%로 일주일 평균 포기율은 9.3%가 되도록 하였다. 상세 일주일 평균 포기율은 아래 <표 1> 과 같다.

<표 1> 요일, 시간대별 인바운드 목표 최대 허용 포기율

	월	화	수	목	금	토	일
9~10 시	25	25	20	20	20	15	15
10~11 시	25	25	20	20	20	15	15
11~12 시	25	25	20	20	20	15	15
12~13 시	100	100	100	100	100	100	100
13~14 시	25	25	20	20	20	15	15
14~15 시	15	15	15	15	15	15	15
15~16 시	15	15	15	15	15	15	15
16~17 시	15	15	15	15	15	15	15
17~18 시	15	15	15	15	15	15	15
요일 별 평균	9.3	9.2	9.2	9.2	9.2	9.5	9.5

제약조건 식은 아래와 같다.

$$\text{minimize } \sum_{i \in d, j \in t} c_n n_d + c_a \hat{y}_t^d A_t^d + n_t^d \quad (1)$$

$$A_t^d \leq B_t^d \quad \forall d, t \quad (2)$$

$$\frac{\sum_t \hat{y}_t^d A_t^d}{\sum_t \lambda_t^d} \leq B^d \quad \forall d, t \quad (3)$$

$$\frac{\sum_d \sum_t \hat{y}_t^d A_t^d}{\sum_d \sum_t \lambda_t^d} \leq B \quad (4)$$

$$A_t^d = f(I_t^d) \quad \forall d, t \quad (5)$$

$$I_t^d = \frac{\hat{y}_t^d}{n^d} \quad \forall d, t \quad (6)$$

$$L_t^d \leq n_t^d \leq U_t^d \quad \forall d, t \quad (7)$$

$$0 \leq n_t^d \leq 1 \quad \forall d, t \quad (8)$$

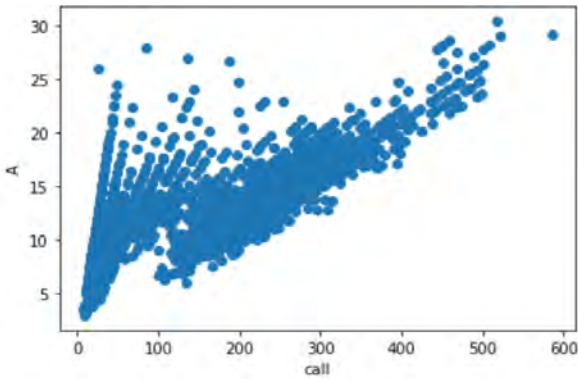
- $B_t^d = t$ 와  $d$ 에 대한 최대 허용 포기율
- $B^d = d$ 에 대한 최대 허용 포기율
- $B =$  매 주당 최대 허용 포기율
- $L_t^d(U_t^d) = t$ 와  $d$ 에 대한  $n_t^d$ 의 하한값(상한값)
- $c_n =$  상담원 한 명당 시간당 인건비
- $c_a =$  포기콜 당 초래되는 기회비용의 손실
- $n_t^d =$  관리자 한명의 시간당 인건비
- $\hat{y}_t^d =$  예측기법을 통해 구한 요일별 시간당 콜 기대값

목적함수를 표현한 식 (1) 은 상담원 한 명당 인건비와 인바운드 포기 콜 당 손실 비용, 관리자의 인건비를 더하여 최소화하는 것이다. 상담업무 이외의 다른 업무를 처리하는데 관련된 인건비는 별도로 지급한다고 가정하였다. 제약 조건식 (2)~(4) 은 목표 포기율은 미리 설정한 요일별 시간별 최대 허용 포기율에 넘지 않아야 함을 나타낸다. 즉, 일평균 포기율의 값을 몇 % 미만으로 지정하여 그 범위안에 존재해야 한다는 것이다. 식 (5)~(6) 은 포기율과 인바운드 비율을 최소자승법으로 표현한 함수인데, 2.2 에서 자세히 다루도록 한다. 식 (7) 은 근무 인원수를 평일과 주말 다르게 설정하여 지정 하한과 상한 범위안에 존재해야 한다는 식이다. JB 주식회사 콜센터의 경우, 관리자가 1명 근무하고 바쁜 시간대에 유동적으로 1시간~2시간 정도 상담업무를 한다. 그에 대한 제약식을 식(8)에 수립하였다.

본 모형에서는 의사결정변수  $n_t^d$  와  $\hat{y}_t^d$  가 필요하게 된다. 결론적으로 가장 적절한 인력  $n_t^d$  을 통하여 인바운드 콜을 처리하는데 소요되는 인건비와 한 콜당 포기 콜 손실의 합을 목적함수로 둔다.

2.2 변수 간 상관분석

2.1.1 에서 목표로 하는 포기율( $A_t^d$ )을 구하기 위해 포기율과 인바운드 콜 비율 간의 회귀분석을 통해 상관관계를 분석하였다. 시간당 걸려온 인바운드 콜 수를 그 시간에 상담했던 상담원 수로 나누어 상담원 한 명당 시간당 처리한 인바운드 콜 수(call)를 구하고, 시간당 고객이 통화 연결을 포기하는 비율을 포기율(A)이라 두고 상관분석을 실시하였다. 결과는 아래 그림과 같다.



(그림 1) 인바운드 콜 비율과 포기율 상관관계

R square 값이 0.76 으로 인바운드 콜의 비율과 포기율의 상관관계에 충분히 의미가 있으며 높은 설명력을 보여주는 것으로 해석할 수 있다. 이 같은 관계를 함수로 표현한다면 아래와 같이 표현할 수 있다. [1]

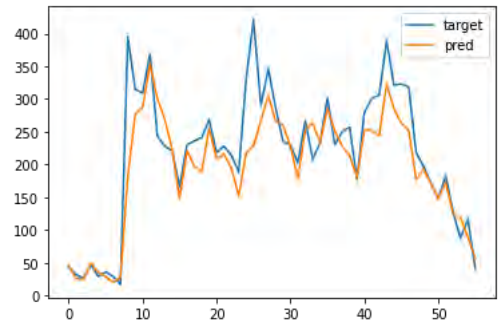
$$A = f(I) = f\left(\frac{\lambda}{n}\right)$$

즉,  $A_t^d$  포기율을 구하기 위한 함수  $f(\cdot)$ 를 최소자승법을 활용하여 회귀 식을 도출하면 아래와 같다.

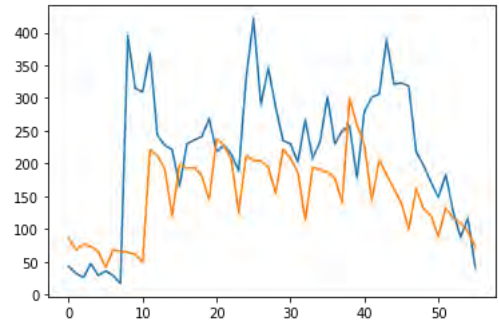
$$A = f(I) = \begin{cases} 0 & \text{if } I < 4.26 \\ 0.0258I + 8.9199 & \text{if } I \geq 4.26 \end{cases}$$

### 2.2.1 인바운드 콜 수 예측 모형

2.1.1 의 제약식에서 입력변수로 사용되는 인바운드 콜의 요일별 시간당 기댓값  $\hat{y}_t^d$  예측하기 위해 2015.01.01~2019.12.31 일까지의 데이터를 이용해 머신러닝 통계기법인 ARIMA(Auto-regressive Integrated Moving Average, 이하 ARIMA) [3] 와 딥러닝 기반의 Vanilla LSTM, Bi-directional LSTM [4] 을 실험하였다. ARIMA 의 경우 63 시점(7일\*9시간)의 장기예측인 경우, 평균으로 회귀하려는 경향이 있어 예측이 제대로 이루어지지 않고, LSTM 의 경우도 장기 예측일수록 오차값이 점점 커진다. 이와 같은 이유로 요일 당 예측으로 바꾸어 7 시점을 예측하고, 시간대의 비율에 따라 콜 예측값을 분배하였다. 앞서 언급한 세가지의 시계열 예측 알고리즘의 성능평가를 통해 최적의 예측값을 이용했다. 그 중에서 앞에서 뒤, 뒤에서 앞, 모두 고려하는 양방향(bidirectional) 네트워크인 Bi-Directional LSTM 와 ARIMA 를 비교하였다.



(그림 2) Bi-Directional LSTM Prediction Plot



(그림 3) ARIMA Prediction Plot

(그림 2) 와 (그림 3) 은 2020년 1월 둘째 주를 예측한 것이다. (데이터 전처리 시 공휴일이 있는 주는 제외하였으므로 1월 첫째 주 제외) 그림을 보면 Bi-directional LSTM 이 ARIMA 보다 실제 값의 추세를 잘 따라가는 것을 확인 할 수 있다. 또한 ARIMA 의 RMSE 값은 103 이지만 Bi-directional LSTM 의 RMSE 값은 47 로 좋은 성능을 나타내는 것을 확인할 수 있었다. 결과적으로 딥러닝을 활용한 인바운드 콜 수 예측값은 기존 데이터에 기반하여 요일과 시간대별로 평균을 내거나 머신러닝 기반 통계분석인 ARIMA 보다 실제 인바운드 콜 수에 근접한 것을 확인할 수 있었다.

아래 <표 2>는 Bi-Directional LSTM 을 활용한 시간당 요일 당 인바운드 콜 수 예측값이다.

<표 2> Bi-Directional LSTM 을 활용한 인바운드 콜 수 예측 값 ( $\hat{y}_t^d$ )

	일	월	화	수	목	금	토
9~10 시	274	242	250	280	136	38.7	378
10~11 시	256	227	240	253	113	33.1	337
11~12 시	226	205	212	231	103	32	299
12~13 시	0	0	0	0	0	0	0
13~14 시	278	256	262	284	121	45	352
14~15 시	236	210	222	242	80	32	292
15~16 시	226	209	217	227	68	28	273
16~17 시	209	195	204	215	58	24	254
17~18 시	171	159	158	167	40	21	202
평균	266	209	190	197	212	81	29

### 3. 비선형계획법 최적화 실험 결과

본 연구에서는 비선형 정수계획법의 해를 찾기 위해 쉽게 사용할 수 있는 microsoft 사의 Excel 기능인 해찾기 기능을 활용하여 지금까지 주어진 매개변수들을 이용하여 최적해를 찾았다. 총비용은 최적 스케줄링을 사용하지 않았을 때와 대비하여 16% 감소하였으며, 최적의 시간대별 상담원 수는 <표 2> 와 같다.

<표 3> 은 최적 스케줄링을 통해 도출된 요일 당 시간당 예상 포기율이며, <표 4> 는 시간대별 상담원 1 명당 얼마나 처리할 수 있는 지에 대한 콜 처리 수를 나타내었다.

<표 3> 최적화를 통한 시간대별 인원수 산정

	일	월	화	수	목	금	토
9~10 시	2	24	25	24	24	25	6
10~11 시	2	23	24	22	23	24	6
11~12 시	2	20	22	20	21	22	5
12~13 시	0	0	0	0	0	0	0
13~14 시	2	22	24	23	23	25	6
14~15 시	2	20	22	21	21	23	4
15~16 시	2	19	22	21	21	22	4
16~17 시	2	18	20	19	20	21	3
17~18 시	1	18	18	18	18	18	2

<표 4> 최적 인력 배치로 나온 예상 포기율

	일	월	화	수	목	금	토
9~10 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
10~11 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
11~12 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
12~13 시	9.0	9.0	9.0	9.0	9.0	9.0	9.1
13~14 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
14~15 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
15~16 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
16~17 시	9.2	9.2	9.2	9.2	9.5	9.5	9.3
17~18 시	9.2	9.1	9.1	9.2	9.4	9.5	9.2

<표 5> 시간대별 상담원 1 명당 콜 처리 수 (Capacity)

	일	월	화	수	목	금	토
9~10 시	11	24	23	15	11	11	11
10~11 시	11	22	23	15	11	11	11
11~12 시	11	22	23	15	11	11	11
12~13 시	0	0	0	0	0	0	0
13~14 시	11	22	23	15	11	11	11
14~15 시	11	22	23	15	11	11	11
15~16 시	11	22	23	15	11	11	11
16~17 시	11	22	23	15	11	11	11
17~18 시	10	22	21	12	10	9	9

### 4. 결론 및 향후 연구

본 연구에서는 콜센터를 운영하는 데 있어서 목표 성공률을 만족시키며 인건비를 최소화 할 수 있는 최적 인력스케줄링을 비선형계획법을 이용하여 분석하였다. 본 연구의 특징은 콜센터에서 설정한 주간, 일간, 시간대별 최대 허용 포기율 및 기타 제약 조건식들을 만족시키면서 총비용을 최소화하는 최적 인력을

산출하고 시간대별로 각 업무에 적정인원을 배치하여 통합적으로 설계했다는 점이다. 현재는 인력스케줄링을 통해 콜센터 관리자나 외주업체가 경험기반에 의존하여 적정인원을 선출하는데, 이러한 해법은 과학적이지 않으며, 모든 변수를 고려할 수 없다. 또한 본 연구는 경험 기반의 인력 스케줄링이 아닌, 데이터 기반으로 수리제약식을 수립하여 목적함수를 최소화하는 결과를 도출하였으며, 그 과정에서 머신러닝을 사용한 예측기법을 사용하였다는 것에 의의가 있다. 하지만 점심시간의 응대나, 파트타임 상담원 등 다양한 변수들이 추가될 때 결과가 달라질 수 있다는 점과 최대 허용 포기율은 미리 지정된 값이 아닌 데이터에 기반한 최대 허용 포기율을 정하는 것들은 향후 연구과제가 될 수 있다. 또한 상담원의 인건비 수준 또한 매우 다양하지만 평균을 내어 사용하여 아쉬운 부분이 있다. 마지막으로 전통적인 최적화 방법이 아닌 부분적으로 사용된 머신러닝 예측 기법을 좀 더 확장시켜 예측하는 인력스케줄링 연구도 의의가 있을 것으로 보인다.

### 참고문헌

- [1] 김성문; 나정은; 김수미. 최적화와 시뮬레이션을 이용한 콜센터의 인력배치 연구. *IE interfaces*, 2011, 24.1: 40-50.
- [2] 최규완. 정수계획법을 이용한 외식인력 스케줄링. *대한지역사회영양학회지*, 2007, 12.5: 630-638.
- [3] CONTRERAS, Javier, et al. ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems*, 2003, 18.3: 1014-1020.
- [4] ALTHELAYA, Khaled A.; EL-ALFY, El-Sayed M.; MOHAMMED, Salahadin. Evaluation of bidirectional lstm for short-and long-term stock market prediction. In: *2018 9th international conference on information and communication systems (ICICS)*. IEEE, 2018. p. 151-156.