

간헐적 특성을 지닌 수요의 예측: 실제 현장 데이터 활용

1. 과제 필요성

과제 선정 배경



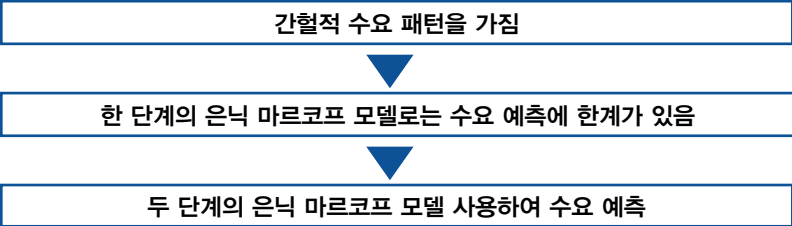
유압 브레이커 전문회사인 수산 중공업은 제품의 대부분이 높은 재고비용을 가지고 있으며 이를 보완하고자 수요예측 기법을 시도하고 있다. 하지만, 일정한 기준이 없는 수요 예측 기법의 사용으로 지속적인 손해를 보고 있다.

데이터 특징

수산중공업의 6년간 판매 데이터

훈련집합
실험집합

비고	기종	판매되지 않은 날	총 판매량
2011	20	126	1581
2012	25	133	1609
2013	26	121	4891
2014	34	150	3811
2015	33	142	3503
2016	29	111	1562



2. 본론

은닉 마르코프 모델(Hidden Markov Model, HMM)

- 은닉 마르코프 모델(Hidden Markov Model, HMM)이란?
시스템이 Hidden State와 Observation의 두 가지 요소로 이루어진 마르코프 모델

은닉 마르코프 모델(Hidden Markov Model, HMM) 선택

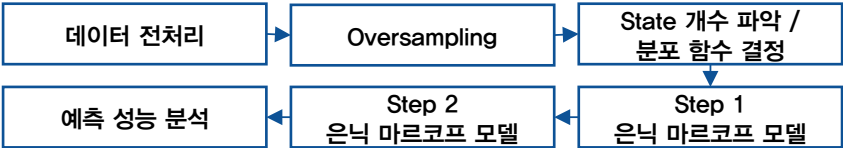
시장 수요 특징

- 시장 상태(Market State)는 예측 하기 어려움
- 수요를 통해서 시장 상태를 예측할 수 있음

HMM 특징

- 은닉 상태(Hidden State)는 눈에 보이지 않음
- 관측치(Observation)을 통해 은닉 상태를 예측할 수 있음

해결 알고리즘 - 모델링



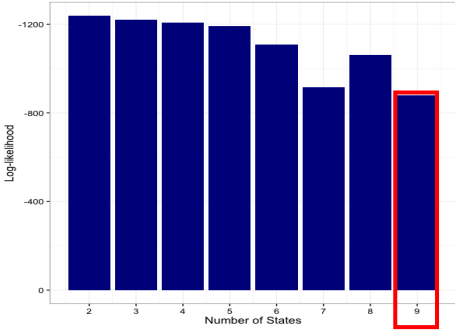
해결 알고리즘 - 데이터 전처리

S/O번호	자재내역	출고량	유용일	인도번호	출하일						
12553	BREAKER SU+55 DE/W275/L385/D100/#700~	1	20160104	1010280	2016-01-12						
12593	BREAKER SB147P SIDE/W430/L505	1	20160104	1010296	2016-01-06						
12594	BREAKER SB147P SIDE/W430/L505	1	20160104	1010297	2016-01-11						
12595	BREAKER SB40/S E/3P/W165/L200/V120??	1	20160104	1010298	2016-01-11						
12638	BREAKER SB40/S E/3P/W165/L200/V120??	1	20160105	1010300	2016-01-06						
12666	BREAKER SB30/S E/3P/W145/L120/V120??	1	20160106	1010305	2016-01-12						
12667	BREAKER SB35/S E/3P/W165/L180/V120??	1	20160106	1010305	2016-01-12						
12668	BREAKER SB40/S E/3P/W165/L200/V120??	1	20160106	1010305	2016-01-12						
12682	BREAKER SB40/S E/3P/W165/L200/V120??	1	20160106	1010305	2016-01-12						
12705	BREAKER SB50/S E/W275/L385	1	20160106	1010305	2016-01-12						
달력		SB10	SB100	SB121	SB13	SB140	SB140	SB140	SB140	SB140	SB140
2011-01-01		0	0	0	0	0	0	0	0	0	0
2011-01-02		0	0	0	0	0	0	0	0	0	0
2011-01-03		0	0	0	1	0	0	0	0	0	0
2011-01-04		0	0	0	0	0	0	0	0	0	0
2011-01-05		0	0	0	0	0	0	0	0	0	0
2011-01-06		0	0	0	0	1	0	0	0	0	0
2011-01-07		0	0	0	0	1	0	0	0	0	0
2011-01-08		0	0	0	0	0	0	0	0	0	0

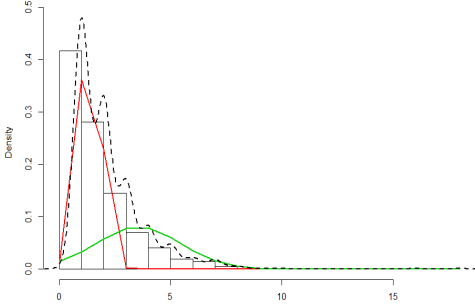


기존 데이터에서 수요 예측에 필요한 attribute (메인바디, 출고량, 출하일)을 추출 후 날짜 별 판매량 데이터 정제함

해결 알고리즘 - State 개수/ 분포 함수 결정



가장 높은 로그-우도(Log-likelihood)를 갖는 상태(state) 개수 채택



판매량 발생 빈도에 적합한 혼합 가우시안 분포(mixture Gaussian distribution) 결정

해결 알고리즘 - 은닉 마르코프 모델

Step 1 은닉 마르코프

- 훈련 데이터를 판매의 발생 유무로 분류(발생시 1, 미발생시 0)
- 분류된 데이터를 사용하여 은닉 마르코프 모델 학습 후 판매의 발생 유무만을 예측

Step 2 은닉 마르코프

- 기존의 훈련 데이터에서 0을 제외한 데이터열로 두번째 모델을 훈련
- 첫 단계에서 발생 할 것이라고 예측한 날에 대해서만 두번째 모델에서 판매 예측량 산출

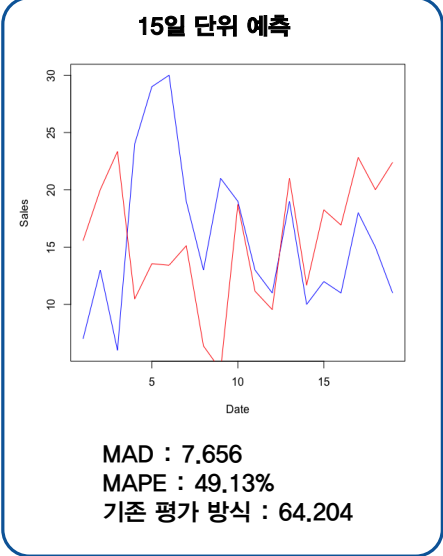
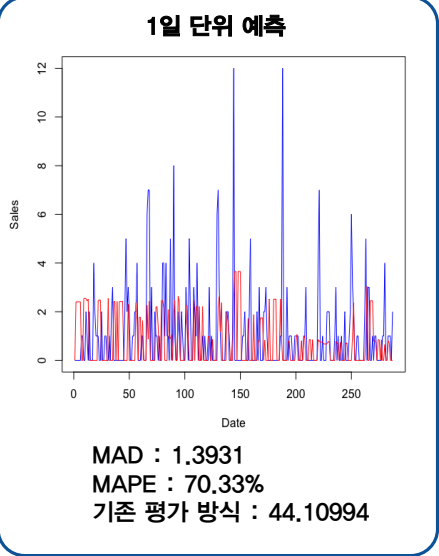
해결 알고리즘 - 예측 성능 분석

- 1일 단위의 예측 결과 분석과 15일 단위의 예측 성능 분석으로 구분
- MAD, MAPE, 기존 성능 측정 방식* 으로 모델 성능 분석

* 기존 평가 방식 = 예측판매량/실제 판매량×100 또는 실제판매량/예측 판매량×100

3. 결론

연구 결과



한계점 및 의의

- 한계점
- 과도한 오버샘플링으로 기존 데이터의 패턴 상실
 - 5년간의 판매량(훈련 집합)이 최근 1년간 판매량(실험 집합)과 비교하여 현저히 작은 경우가 많음
 - 판매량 데이터가 계절적 패턴을 띄고 있으나, 학습하기에는 패턴이 부족함

의의 기존 모델과 비교하여 두 개의 은닉 마르코프 모델 사용을 통해 간헐적 수요에 대처가 가능함