

# 시계열 예측 모델을 이용한 반도체 수요 예측

조인환, 신오순

승실대학교 IT융합학과

ciw0802@naver.com, osshin@ssu.ac.kr

## Demand Forecasting Using Time Series Forecasting model for Semiconductor Industry

Cho In Hwan, Shin Oh soon

Dept. of IT Convergence, Soongsil Univ.

### 요 약

본 연구의 목적은 반도체에 대한 수요를 예측함으로써, 대규모의 R&D 투자가 요구되는 반도체 산업 기업체들의 의사 결정에 기반이 되기 위한 객관적인 수요 예측 모형을 구축하는 것이다. Time Series Forecasting 과제에 활용되고 있는 Transformer, NLinear, SCINet, FiLM 모델에 반도체 데이터를 적용하여 성능을 비교하고, 이 중 가장 성능이 뛰어난 모델에 레이어 추가, Dropout 등 최적화 알고리즘을 추가하여 반도체 데이터에 대한 성능 향상을 목표로 한다.

### I. 서 론

2021년 반도체 가격이 연평균 30% 상승하는 슈퍼 사이클이 온 후, 반도체 산업 전반은 활황을 예상했다. 하지만 2022년 하반기 세계적인 경기 침체로 반도체 수요가 급격히 위축되면서 활황에 대한 예상은 깨지게 되었다. 본 연구의 목적은 반도체에 대한 수요를 예측함으로써, 대규모의 R&D 투자가 요구되는 반도체 산업 기업체들의 의사 결정에 기반이 되기 위한 객관적인 수요 예측 모형을 구축하는 것이다.

Time Series Forecasting 과제에 활용되고 있는 Transformer, NLinear, SCINet, FiLM 모델에 반도체 데이터를 적용하여 성능을 비교하고, 이 중 가장 성능이 뛰어난 모델을 레이어 추가, Dropout 등 최적화 알고리즘을 추가하여 반도체 데이터에 대한 성능 향상을 목표로 한다.

### II. 기존 연구 및 활용 데이터

#### 1. 기존 연구

김중선(2022.6)<sup>[1]</sup>은 한국 반도체산업의 수출 경쟁력을 제고하기 위한 방안으로, 세계 시장 점유율, 무역특화지수 등 관세 통관 통계 데이터를 사용하였다. 반도체 품목군별 시장점유율을 분석하기 위해 반도체 수출입실적의 81.7%를 차지하는 한국, 미국, 중국, 대만, 홍콩, 일본, 독일, 싱가포르의 데이터를 분석하였다.

박상현(2010.8)<sup>[2]</sup>은 반도체 가격에 신경망을 이용한 모형을 구축하여 예측을 하였다. 반도체 가격 예측 모델에

대해 반도체 가격을 목표 변수로 두고, 누적 반도체 생산량, 환율, 금리, DOW 지수, KOSPI 지수를 입력 변수로 선택하여 단기 예측에서 신경망 알고리즘이 높은 성능을 보이는 것을 확인했다.

기태우(2022.8)<sup>[3]</sup>은 ANN, ARIMA, 이동평균법, 지수평활법 모델을 사용하여 반도체 수요 예측 연구를 진행하였다. 연구에 활용한 데이터는 월간 판매 데이터이고, 생산 수량에 대한 의사결정 기간을 고려해 3개월 단기 수요 예측을 진행하였다. 연구 결과 ANN모델, 지수평활법, ARIMA 순으로 반도체 수요 예측에 대한 높은 정확도를 확인하였다.

기존 연구들을 토대로 반도체 수요 예측에 활용할 데이터를 선택한다.

#### 2. 활용 데이터

반도체 가격 및 수요 데이터에서 수요에 상관성 있는 변수를 추출한다. 반도체 산업의 75%를 차지하는 한국, 미국, 중국, 대만, 홍콩, 싱가포르의 1인당 GDP 데이터와 반도체 생산, 반도체 가격, 반도체 수출, 시장점유율 데이터를 사용한다.

목표 변수는 반도체 수출 데이터이고, 공급 요인 변수는 반도체 생산, 반도체 가격, 시장 점유율을 사용한다.

수요 요인 변수는 1인당 GDP 데이터를 사용한다.

### III. 시계열 예측 모델

#### 1. Transformer

Transformer 모델은 RNN이 각광받을 때 제시된 모

델이다. RNN은 병렬화(parallelization)가 불가능해진 sequence의 데이터를 계산할 때 효율이 많이 떨어지는 문제가 있었다. Transformer 모델은 병렬화가 가능한 구조로 인코더와 디코더를 연결한 구조이다. 시계열 예측을 위한 공개 데이터인 Electricity Transformer Temperature 데이터셋(ETT dataset)에서 평균 절대 오차(MAE) 0.4213의 낮은 오차를 보인다.

## 2. NLinear

NLinear모델은 먼저 소개한 Transformer 기반의 모델이 Time Series Forecasting에 적절한지 의문을 제시하며 상대적으로 간단한 선형 모델로 Transformer 보다 높은 성능을 달성한 모델이다. 시계열 데이터가 가진 정보를 Trend와 Remainder로 분해하고 간단한 Layer network 구조를 통해 분석한다. ETT dataset에서 MAE 0.226을 달성하였다.

## 3. SCINet

SCINet 모델은 Convolution 레이어를 사용하는 TCN 모델을 기반으로 하여 시계열 데이터를 Trend와 분기 정보, 변칙정보로 나눠 적용하는 모델이다. 현실 세계의 시계열 데이터는 노이즈가 많아 변칙정보를 고려한 SCINet의 잠재력은 높다고 할 수 있다. ETT dataset에서 MAE 0.316을 달성하였다.

## IV. 반도체 수요 예측

### 1. 활용 데이터 선정

수요 예측에 활용하는 모델은 ETT dataset에 최적화된 모델로, 해당 데이터셋은 시간 변수, 목표 변수 외의 6개의 입력변수를 필요로 한다. 앞서 선정한 변수는 목표 변수와 시간 변수를 제외하면, 공급 요인 변수 3개, 수요 요인 변수 6개로 총 9개의 변수를 고려하고 있다. 이 중 목표 변수인 반도체 수요에 상관성이 큰 변수를 6개 선정한다. 상관성 계산은 pearson 상관계수를 이용하였고, 반도체 수출(억불)이라는 목표변수에 대하여 홍콩 GDP, 반도체생산(조원), 대만GDP, 반도체시장점유율(%), 중국GDP, 미국GDP 순으로 상관계수가 높게 나타났다. 이 6개의 변수를 공급, 수요 요인 변수로 사용하여 목표 변수를 예측한다.

변수명	상관계수
반도체수출(억불)	1.000000
홍콩	0.971466
반도체생산(조원)	0.965341
대만	0.955282
반도체시장점유율(%)	0.949949
중국	0.945195

변수명	상관계수
미국	0.922119
year	0.912264
대한민국	0.901469
싱가포르	0.883375

## 2. 예측 모델에 사용

예측 모델에 활용 데이터를 입력하여 가장 오차가 작은 예측 모델을 선정한다.

## V. 결론

연구 모델 1차 선정에 사용된 ETT 공개 데이터셋은 변압 부하에 따른 온도 예측이라는 물리적인 수치를 예측하기 위한 데이터셋이다. 이 때문에 ETT 데이터셋에 최적화된 모델이라 하더라도 반도체 수요와 같은 복합적인 변수가 영향을 미치는 데이터에는 예상 성능과 다른 성능을 보일 수 있다. 상관성이 적은 변수라고 여긴 대한민국 GDP와 싱가포르GDP를 실제로 사용했을 때 예측 성능은 어떻게 변하는지, 예측 모델에 레이어 추가 등 모델을 수정하였을 때 성능 변화는 어떤지, drop out과 같은 Deep learning 최적화 알고리즘이 성능을 향상시키는 지 등 예측 성능을 끌어올리기 위한 추가 과제에 대해 생각해볼 수 있겠다.

## 참 고 문 헌

- [1] Joong-Seon Kim, "A Study on the Export Competitiveness of Korean Semiconductor Industry", June. 2022.
- [2] S. H. Park, "A Neural Network-based semiconductor Price Prediction System", Aug. 2010.
- [3] TaeWoo Ki, "Demand Forecasting Using Artificial Neural Networks for a Material Supplier in Semiconductor Industry", Aug. 2022.