

# A Study on the Korean Interest Rate Spread Prediction Model Using the US Interest Rate Spread : SVR-Ensemble (RNN, LSTM, GRU) Model based

Sun-Ho Jeong · Young-Hoo Kim · Myung-Jin Song · Yun-Jae Chung · Sung-Seok Ko<sup>†</sup>

Department of Industrial Engineering, Konkuk University

## 미국 금리 스프레드를 이용한 한국 금리 스프레드 예측 모델에 관한 연구 : SVR-앙상블(RNN, LSTM, GRU) 모델 기반

정순호 · 김영후 · 송명진 · 정윤재 · 고성석<sup>†</sup>

건국대학교 산업공학과

Interest rate spreads indicate the conditions of the economy and serve as an indicator of the recession. The purpose of this study is to predict Korea's interest rate spreads using US data with long-term continuity. To this end, 27 US economic data were used, and the entire data was reduced to 5 dimensions through principal component analysis to build a dataset necessary for prediction. In the prediction model of this study, three RNN models (BasicRNN, LSTM, and GRU) predict the US interest rate spread and use the predicted results in the SVR ensemble model to predict the Korean interest rate spread. The SVR ensemble model predicted Korea's interest rate spread as RMSE 0.0658, which showed more accurate predictive power than the general ensemble model predicted as RMSE 0.0905, and showed excellent performance in terms of tendency to respond to fluctuations. In addition, improved prediction performance was confirmed through period division according to policy changes. This study presented a new way to predict interest rates and yielded better results. We predict that if you use refined data that represents the global economic situation through follow-up studies, you will be able to show higher interest rate predictions and predict economic conditions in Korea as well as other countries.

**Keywords :** Interest Rate-Spread, SVR-Ensemble, RNN, PCA

### 1. 서 론

미국은 1992년 통화 위주의 경제 정책에서 금리 위주의 경제 정책으로 전환하였다. 그 후 금리 정책은 미국 경제를 이해하는 핵심 키워드가 되었다. 일반적으로 경기는 불황과 호황을 반복하며 순환하고 있으며, 경기에 따라 각국의 중앙은행에서는 유기적인 금리 정책을 펼치고 있다. 장단

기의 국채금리 또한 금리 정책과 해당 국가와 세계 경기 상황에 따라 변동을 하고 있다. 그러나 일반적으로 장기와 단기 금리의 변동은 비슷한 방향성을 가지고 있으나, 그 변동 폭은 다르게 나타나고 있다. 따라서 장기 국채금리와 단기 국채금리의 차이가 발생하고 있으며, 이러한 차이를 금리 스프레드라 정의하고 있다. 금리 스프레드는 미래의 경기 예측에 유의미한 결과를 가져다준다고 Harvey[9]에 의해 실증적인 분석이 이루어졌으며, Kim and Lee[16]의 연구에서도 신용 스프레드는 경기 예측력에 유의미한 결과가 있음을 보여주고 있다.

Received 3 July 2020; Finally Revised 2 September 2020;

Accepted 10 September 2020

<sup>†</sup> Corresponding Author : ssko@konkuk.ac.kr

장기 금리가 단기 금리보다는 높은 경우가 일반적인 현상이나, 단기 국채의 금리는 상승하고 장기 국채의 금리는 하락할 때 장, 단기 금리 간의 역전이 발생하여, 금리 스프레드가 음의 값을 갖는다. 이는 미래 경기에 대한 비관론이 확산하고, 기대 인플레이션은 하락하는 구간에서 주로 발생하고 있다. 미국의 경우 장단기 금리 역전 이후 대체로 12~24개월 안에 경기침체가 뒤따랐다. 80년 이후 장단기 금리가 역전되고 나면 평균적으로 15개월 내외에서 경기침체가 발생하고 있음을 경제 데이터를 통해 확인할 수 있다. 따라서 금리 스프레드의 역전을 예측할 수 있다면 불황에 미리 대비할 수 있고 불황에 도달하지 않도록 적절한 정책을 사용할 수 있을 것이다.

따라서 본 연구에서는 딥러닝 기법을 이용한 한국 금리 스프레드 예측을 위한 모델에 관한 연구를 진행하였다. 하지만 국내의 경제지표들은 딥러닝 기법을 적용하기 위한 자료의 양이 부족한 측면이 존재한다. 그러나 세계화로 인해 전 세계 국가들의 경제는 상당한 정도로 동조화되고 있으며, 경제도 유기적으로 연결되었다. 한 나라의 경제 불황은 그 나라에만 국한된 것이 아니고, 그 나라의 경제와 무역 규모에 따라 다른 나라에 많은 영향을 미친다. 따라서 국내 금융 상황은 미국의 금융 상황과 상당 부분 동조화되어, 충분한 분석 자료 확보가 쉬운 미국의 금리 스프레드를 이용하여, 간접적으로 한국의 금리 스프레드를 예측하는 모델을 연구하였다.

본 논문은 제 2장에서 금리 스프레드에 관한 선행연구와 시계열 데이터 예측에 관한 연구를 조사, 분석하였으며, 제 3장에서는 본 연구에서 사용하는 요인들과 차원 축소를 위해 주성분 분석을 하였고, 시계열 예측에 활용되는 RNN(Recurrent Neural Network)가 모델, 분류와 회귀분석에 사용되는 SVR(Support Vector Regression) 모델을 제시하였다. 제 4장에서는 3장에서 제시한 분석 모델에 관한 결과를 논하였다. 마지막으로 결론에서는 본 연구의 시사점과 향후 연구 과제를 제시하였다.

## 2. 선행연구

Stock and Watson[21]에 의해 금리 스프레드를 이용한 경제 상황 예측이 처음 시도된 이후, 금리 스프레드와 관련된 선행 연구가 활발히 진행되어오고 있다. Estrella and Hardouvelis[6]은 단순 회귀모형을 통해 금리 스프레드가 미국의 경제성장률과 불황 예측 지표로서의 가능성을 시사하였으며, Gabraith and Tkacz[8]은 금리 스프레드와 경제성장률 간의 비선형 관계를 실증분석하였다. Far and Ariff[7]의 연구에 따르면 금리 스프레드에 영향을 미치는 변수로는 GDP(Gross Domestic Product) 성장, 화폐 공급, 산업

생산 및 무역수지가 있으며, Ahmad et al.[1]는 금리 스프레드는 소비자 물가 지수와 이자율의 수익률 변동에 가장 큰 영향을 미친다고 주장했다. 또한 Min[18]은 GDP 대비 부채 비율, 부채 서비스 비율, 순 외국 채무, 국제 준비금 대비 GDP 비율, 인플레이션율, 유가, 환율의 유동성, 용해도 및 거시경제 변수 측면에서 금리 스프레드에 크게 영향을 미치는 것을 확인하였다.

국내에서 또한 한국의 금리 스프레드 관련 연구들이 진행되었다. Kim et al.[15]은 다변량 시계열 모형(VAR : Vector Autoregressive Regression 모형)을 사용하여 한국 환율, 생산자물가지수, 거시경제 변수들 사이 관련성에 대한 사전 지식 없이 통계적 절차에 따라 금리 스프레드를 예측하였다. Ji and Park[12]는 금리 스프레드와 경기변동에 대한 불황 확률을 알아보기 위해서 Probit 모형을 이용하여 불황 확률을 추정하였다. Park et al.[19]은 Nelson-Siegel 모형을 이용하여 우리나라 국고채 수익률 곡선 추정하였으며, 중앙은행 기준금리, 소비자물가지수(CPI), 산업생산지수, 거시경제 변수를 고려하여, 더 나은 예측력을 보여주었다.

다양한 시계열 금융 데이터 분석을 위한 분석 기법은 ARIMA(Auto Regressive Integrated Moving Average) 기반 분석의 기법을 통해 발전하였으나, 최근에는 딥러닝 기반의 알고리즘이 주를 이루고 있다. Dagum[5]은 경제 시계열 데이터의 예측 및 계정 조정을 위하여 자기 회귀 누적 이동평균(ARIMA) 모형을 사용하였다. 또한, Wang and Leu[22]은 주식 시장의 추이를 인공지능망 기반의 ARIMA 모형을 통해 예측하였다. Shin et al.[20]는 RNN과 LSTM(Long Short Term Memory Models)을 활용하여 15%의 추가 예측률 향상을 보였으며, Joo and Choi[13]는 양방향 LSTM 모형을 통해 예측에서 발생하는 오차를 감소시켰다. 최근 연구에서는 다수의 딥러닝 모델을 조합하여 성능을 높여려는 연구가 이어지고 있다. Kim et al.[14]는 MLP (Multi Later Perception), CNN(Convolutional Neural Network), RNN의 기반 앙상블 모형을 통해 주가를 예측하였다. 또한 Althelaya et al.[2]은 LSTM과 GRU(Gated Recurrent Unit)의 앙상블 모형을 통해 주식 시장을 예측하고자 하였다.

## 3. 연구모델 설계

본 연구모델은 한국 금리 스프레드를 예측하기 위해 미국의 금리 스프레드를 이용한다. 연구모델은 두 가지 과정을 거치게 되는데, 첫 번째는 딥러닝 모델을 이용하여 미국의 금리 스프레드를 예측하는 것이다. 이 과정에서 미국 금리 스프레드에 영향을 주는 요인들을 PCA(Principal

Component Analysis) 분석을 통해 차원을 축소한 후 딥러닝 모델의 독립변수로 사용하여 미국 금리 스프레드를 예측하는 BasicRNN, GRU, LSTM 세 가지 딥러닝 모델을 구현하였다. 다음 단계로 예측된 미국 금리 스프레드를 이용하여 SVR 앙상블 모델을 통해 한국 금리 스프레드를 예측하는 모델을 설계하였다.

### 3.1 한국 미국 금리 스프레드 상관분석

한국 단기 국채 1년 물은 2000년 이후로 발행되지 않았고, 3년 물의 단기 국채는 1999년부터 발행되었다. 따라서 단기 국채로는 3년 물의 국채를, 그리고 장기 국채로는 10년 물의 국채를 사용하여 한국 장단기 금리 스프레드를 정의하였다. 그러나 한국 금리 스프레드를 예측하는데 데이터의 수가 부족하다. 이를 보완하기 위해 데이터의 양이 상대적으로 많은 미국 금리 스프레드를 이용하여 한국 금리 스프레드를 예측하였다.

이를 위해서 한국과 미국 금리 스프레드 상관분석을 실시하였다. 미국의 장단기 금리 스프레드의 장기 국채는 10년 물로 하였고 단기 국채는 3개월물로 정의되고 있다. 2001년 10월부터 2019년 10월까지 매월 말일 기준 미국과 한국의 금리 스프레드에 대한 상관분석을 실시한 결과는 <Figure 1>에 나와 있다. <Figure 1>에서 x축은 한국 금리 스프레드를 의미하고, y축은 미국 금리 스프레드를 나타낸다.

2009년 이전에는 한국과 미국의 금리 스프레드 차이가 상관분석 결과 상관계수가 0.5121이었다. 하지만 2009년 3분기 이후 금리 스프레드의 상관분석 결과는 상관계수가 0.8185로 계산되었다. 상관분석의 결과는 2009년 3분기 이후 한국의 금리 스프레드는 미국의 스프레드에 상당히 동조화되었음을 보여주고 있다. 이것은 2008년 서브프라임 모기지 사태 이후 미국의 금융 정책에 대한 영향이 이전보다 훨씬 증가하였음을 보여주고 있다. 그리고 이 분석을 통해 미국의 금리 스프레드를 이용한 국내 금리 스프레드 예측이 유효한 전략임을 보여주고 있다.

### 3.2 데이터 구성

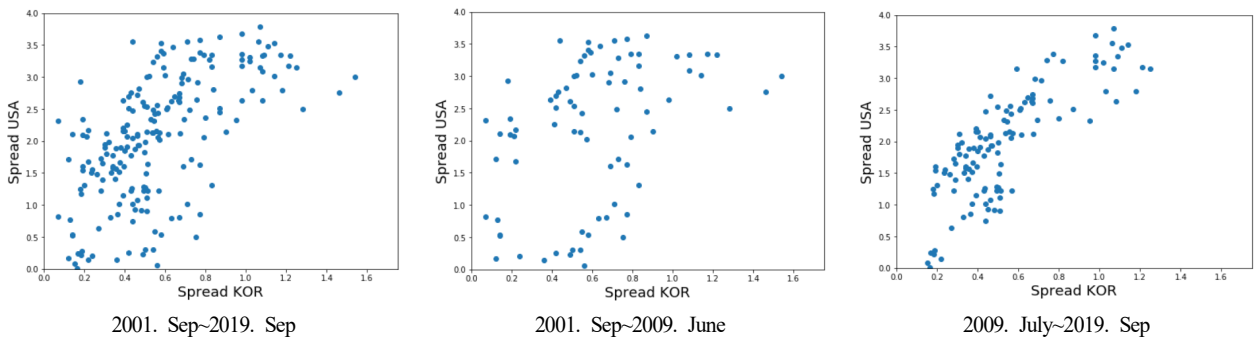
본 연구에서는 금리 스프레드를 예측하기 위해 미국의 정책이 통화정책에서 금리정책으로 바뀐 시기인 1992년 1월 1일부터 미국 국고채 3개월 물과 10년 물의 스프레드를 월 단위로 분석하였다. 금리 스프레드에 영향을 미치는 요인으로 선행연구에서 학자들이 제시한 요인들과 미국 콘퍼런스 보드(Conference board)에서 매달 발표하는 경기 동향 지표에 영향을 주는 요인들을 사용하였다. 이 과정에서 차원 축소와 변수 추출 기법으로 널리 쓰이고 있는 주성분 분석(PCA)을 활용하여 차원이 축소된 주성분을 통해 금리 스프레드 예측 모델을 구현하였다.

#### 3.2.1 요인 구성 및 특징

데이터를 구성하는 요인은 총 27가지로 앞에서 조사한 금리 스프레드에 영향을 주는 요인 중 중복되는 요인을 제외한 13가지의 요인을 사용했다. Far and Arriff[7]의 연구에서는 무역수지, GDP 성장, 소비자 생산 지수, 구매자 관리지수 Min[18]의 부채 비율, 상업 및 산업 대출, 유가, 환율, 금리 Cheng et al.[3]의 다우 존슨 지수 Kim et al. [15]의 생산자 물가 지수 Ahmad et al.[1]의 소비자 물가 지수 Park et al.[19]의 산업생산지수의 데이터가 본 연구에 사용되었다.

나머지 14가지 요인은 경기지표 중 선행, 동행, 후행지수를 사용하였다. 경기 동향 지표는 미국의 각종 경제지표에 대해서 조사, 분석 발표하는 미국 콘퍼런스보드의 자료를 이용하였다. 이 지표들은 경기 선행지수, 경기 동행지수, 경기 후행지수로 이루어져 있다. 경기 선행지수는 경제 변화를 선행하는 경향을 보이는 지표들로 구성되어 있으며 호황과 침체를 예측하는 높은 정확도를 보인다. 본 연구에 사용된 경기 선행 지표는 산업 생산품, 개인 소득 공제 이체 지분, 실제 제조 및 무역 판매 지수가 있다.

경기 동행지수는 현재 경기 상황 판단에 이용되며, 비농업 고용자 수, 화폐 공급량, 제조업체 신규주문, 신규



<Figure 1> Correlation Analysis between US Interest Rate Spread and Korea Interest Rate Spread

주택 건설, 소비심리, 실업 보험이 사용되었다. 경기 후 행지수는 현재 경기의 사후 확인에 이용되며, 평균 실업 기간, 생산 단위 노동비용, 은행 평균 대출 금리, 서비스 소비자 가격 지수, 무역 재고 판매 비율이 사용되었다. 위 요인들은 세인트루이스 연방 준비은행에서 제공하는 데이터베이스, FRED(Federal Reserve Economic Data)에서 추출했으며, 1992년 1월 1일부터 2019년 10월 1일까지 월간 데이터의 증가를 사용했다.

### 3.2.2 주성분 분석

주성분 분석은 상관관계가 높지 않고, 독립성이 약한 다수의 변수 조합을 통해, 그 변수들의 정보를 가장 많이 나타낼 수 있는 새로운 변수를 생성하는 것으로, 차원을 축소하거나 단순화하는 다변량 자료 분석 기법이다[17]. 본 연구를 위해 수집한 27개 요인 간의 관계가 분명치 않아 금리 스프레드를 예측하는 모형의 데이터로 모두 사용하는 것은 비효율적인 측면이 존재한다. 또한, 딥러닝에서 하나의 요인이 증가하면, 하나의 차원이 증가하게 되는데, 이는 데이터 간의 거리를 증가시키고 모델을 복잡하게 하여 과적합 위험이 커지는 경향이 존재한다. 따라서 주성분 분석을 통해 축약된 주성분을 금리 스프레드 예측을 위한 지표로 사용하였다. 그리고 실증분석 결과는 다음 장에서 제시하였다.

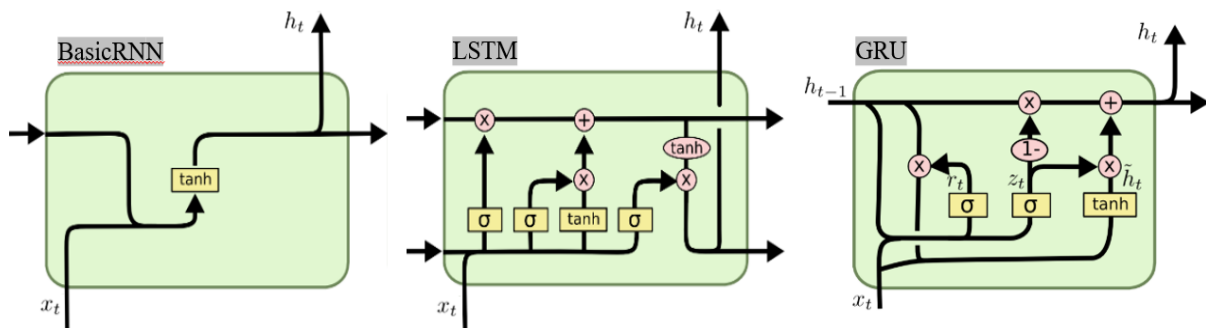
### 3.3 RNN 모델 : 미국 금리 스프레드 예측

본 연구모델에서 사용되는 데이터는 시간의 순서, 즉 시계열의 속성을 가지고 있다. 따라서 미국의 금리 스프레드 예측을 위해서는 시간의 흐름에 따라 변화하는 시계열 데이터 처리에 좋은 성능을 보이는 RNN 모델을 활용하고자 한다. RNN은 은닉 노드가 순환구조 형태를 보이는 인공신경망의 한 종류이며, 과거 학습 결과가 현재의 학습에 활용되는 딥러닝 네트워크이다.

BasicRNN은 RNN의 기본적인 모델로서 입력층, 은닉층, 출력층이 존재하며, 은닉층의 노드에서 활성화 함수

를 통해 나온 결과값을 출력층 방향으로 보내고 다시 은닉층 노드의 다음 계산 입력으로 보내는 특징이 있다. 하지만 장기적인 시간 의존성을 갖는 데이터를 학습해야 하는 경우, 기울기 소실 문제가 발생할 수 있다. LSTM은 BasicRNN에서 발생할 수 있는 기울기 소실 문제를 해결하기 위하여 제안된 모델로서, Hochreiter and Schmidhuber[10]은 장기 의존성을 학습하는 구조의 인공신경망 LSTM을 제안하였다. LSTM 모델의 은닉층은 과거의 정보를 고려한 망각 게이트(Forget Gate), 입력할 크기를 정해주는 입력 게이트(Input Gate), 출력할 양을 결정하는 출력 게이트(Output Gate)로 구성되어 있다. GRU는 Cho et al.[4]이 의존성 문제에 대한 해결책을 유지하면서, 은닉 상태를 업데이트하는 계산을 줄여 LSTM을 단순화한 인공신경망이다. GRU 모델의 은닉층은 과거의 정보를 초기화하는 리셋 게이트(Reset Gate)와 과거와 현재 데이터의 최신화 비율을 결정하는 업데이트 게이트(Update Gate) 두 가지 게이트만이 존재하며, LSTM보다 간략하지만 유사한 성능을 보인다. 각 모델의 은닉층 노드의 내부 구조는 <Figure 2>를 통해 확인할 수 있다. 다수의 경제지표 데이터를 입력받아 금리 스프레드 예측값을 출력하는 다대일(many-to-one) 모델을 설계한다.

BasicRNN, LSTM, GRU의 경우 파라미터의 조정에 따라 학습의 결과가 바뀌기 때문에 구조는 고정하되, 그리드서치(grid search) 방법을 통해 정해진 기준 내에서 적합 파라미터를 찾고자 하였다. 미국 금리 차 예측 시, 정규화 작업을 거치는 것이 효과적이기 때문에 MinMaxScaler를 통해 데이터의 범위가 0과 1 사이의 값을 갖도록 변환시켜주었으며, 학습 횟수(Iterations)의 수는 500으로, 학습률(Learning Rate)을 0.01로 설정하였다. 활용한 금리 스프레드 예측을 위한 변수들의 기간을 고려하여 시퀀스(Sequence)는 1년, 12달 데이터를 기준으로 학습하여 12로 설정하였다. 각 순환신경망 모형의 은닉층은 10개의 층을 사용하였으며, 활성화 함수로는 tanh, 손실 함수로는 RSME(Root Mean Square Error), 최적화 방법은 Adam Optimizer를 활용하였다.



<Figure 2> RNN Models Internal Structure

### 3.4 SVR 앙상블 모델 : 한국 금리 스프레드 예측

앞서 언급한 것과 같이 세 가지 딥러닝 모델을 통한 미국 금리 스프레드 예측값을 독립변수로 활용하여, SVR 앙상블 모델을 설계하였다. 딥러닝 모델 간의 장단점이 존재하므로 단점을 감소시키고, 장점을 증가시키기 위해 앙상블 모델을 사용하고자 하였다. 일반적인 앙상블 모형은 ‘1/요인 수’를 곱하는 형태로 세 가지 모델의 영향력의 비중이 같아지도록 하였다. 그러나 본 연구모델은 각각의 요인들이 한국 금리 스프레드에 끼치는 영향력을 개별적으로 고려할 수 있는 기계 학습 기법의 SVR을 통해 각 요인의 정확한 계수들을 구하여 앙상블 모델을 사용하였다.

SVR은 분류, 회귀, 특이점 판별에 쓰이는 지도 기계 학습 방법의 하나다. 학습데이터를 비선형 매핑을 통해 고차원으로 변환한다. 새로운 차원에서 초평면을 최적으로 분리하는 선형분리를 찾으며, 의사결정 영역을 찾는다. SVR의 일반적인 함수는 다음과 같다.

$$f(w, b) = w^T \Phi(x) + b$$

여기서  $w, b$ 는 회귀식의 모수로,  $w$ 는 가중치 벡터를 의미하며,  $b$ 는 편향 상수이다. 그리고  $\Phi(x)$ 는 입력 변수의  $x$ 를 고차원으로 변환하는 함수를 의미한다. 그리고 본 논문에서 사용한 SVR 모델은 RBF kernel을 사용한다. RBF kernel은 재무 관련 시계열 데이터를 예측할 때 더욱 광범위하게 사용되며, 다른 유형의 kernel보다 우수하다고 할 수 있다[11]. RBF kernel은 다음과 같다.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

$\gamma$ 는 kernel 방정식의 분산을 의미한다.

## 4. 실증분석

### 4.1 주성분 분석 결과

<Table 1>은 수집한 27개의 요인에 대한 주성분 분석

결과이다. SPSS(Statistical Package for the Social Sciences)를 사용하여 분석하였고, 요인 회전은 Varimax방식을 채택하여 고유값이 1 이상인 주성분만 제시하였다. 주성분 분석을 통해 생성된 주성분 1은 전체 27개의 인자 특성 중 약 57%를 나타내고 있고, 주성분 2는 약 14%를 나타내고 있다. 주성분 1과 주성분 2는 전체 인자의 71.30%를 담당하고 있으며 주성분 1~5는 전체 데이터의 약 91.516%를 나타내고 있다.

<Table 1> Results of Principal Component Analysis

Component	Total	Variance(%)	Cumulative(%)
1	15.341	56.820	56.820
2	3.911	14.485	71.305
3	2.296	8.502	79.807
4	1.980	7.333	87.140
5	1.182	4.377	91.516

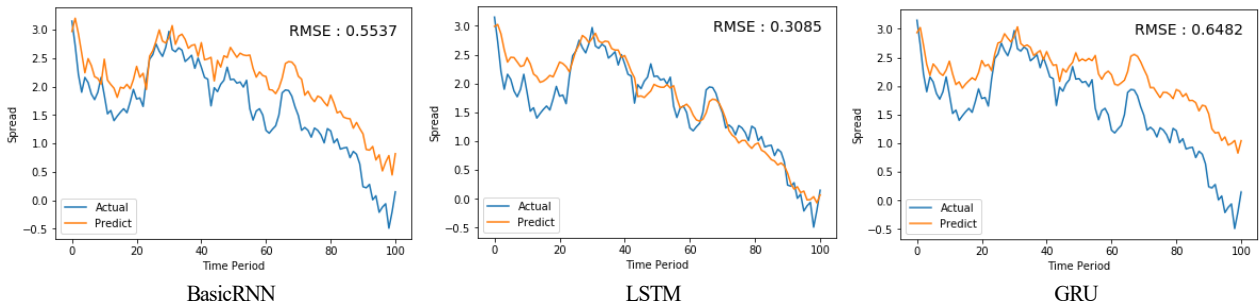
<Table 2>는 주성분 분석을 통해 27개의 요인이 5개의 주성분으로 나누어진 것을 제시하였다. 5개의 주성분을 구성하는 요인은 공통점을 가지고 있어 그것을 바탕으로 제목을 명명하였다. 주성분 1은 총 15개의 요인과 높은 연관성이 있으며, Dow Jones 지수, GDP 성장률, 소비자 물가 지수 등 전반적인 경제 상황을 보여주어 경제지표로 명명하였다. 주성분 2는 경기 상황에 민감한 신규 주택 건설, 엔·달러 환율, 평균 실업 기간 등과 관계가 있으므로 경기지표로 명명하였다. 주성분 3은 무역 재고 판매 비율, 무역수지에 영향을 받으므로 무역지표라고 명명하였고, 주성분 4는 실업보험, 구매관리지수, 생산노동자의 한 주 노동시간에 연관성이 있어 제품 생산에 영향을 미치므로 생산지표로 명명하였다. 주성분 5는 금리지표로 미국 Prime 금리와 미국 기준금리와 관계가 있다.

### 4.2 미국 금리 스프레드 예측 결과

실제 데이터와 예측 데이터 간 오차의 평균 제곱합 RMSE를 활용하여 개별 모델들을 평가하는 지표로 사용했다.

<Table 2> Group of Principal Component Analysis Factors

Main Index	# of Factors	Factors
Economic Index 1	15	Dow Jones Average, Commercial And Industrial Loan, M2, GDP Growth, Real Manufacturing And Trade Industries Sales, Consumer Price Index(Service), Producer Price Index, Personal Income Less Transfer Payments, Industrial Production, Import Price Index, Manufacturer New Orders(Consumer), Consumer Price Index, New Order For Non Defence, Federal Debt, Labor Cost Per Unit Of Output
Economica Index 2	5	New Private Housing, Yen Dollar Rate, Average Weeks Unemployed, Consumer Expectation, Crude Oil
Trade Indicator	2	Ratio Of Trade Inventories Sales, Trade Balance
Production Index	3	Purchase Management Index, Insurance Unemployment, Average Weekly Hours Of Production Workers
Interest Rate Index	2	Average Prime Rate Charged By Banks, EFF Rate



<Figure 3> Prediction of US Spread Using RNN

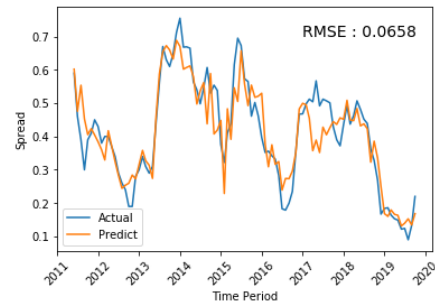
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

주성분 분석에서 도출된 5가지 주성분을 독립변수로 사용하여 앞서 제시한 3가지 모델(BasicRNN, LSTM, GRU)을 구현하였다.

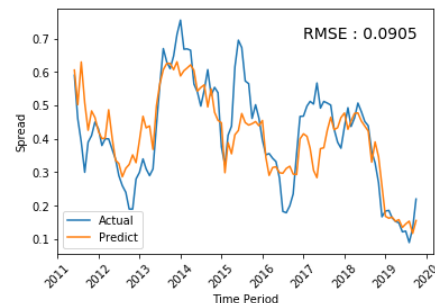
<Figure 3>에서는 이 3가지 모델에 대한 그래프이며, 2011년 6월부터 2019년 10월까지 총 101개월의 미국 금리 스프레드를 나타내었다. 주황색 그래프는 예측 데이터이고, 하늘색 그래프는 실제 데이터이다. 동일한 데이터를 독립변수로 사용하였지만, 딥러닝 모델에 따라 그래프의 추세와 RMSE 값이 상이했다. RMSE 값이 가장 낮은 모델은 LSTM을 사용한 모델로 RMSE 값이 0.3085이며, 세 가지 모델 중 미국 금리 스프레드를 가장 잘 예측하였다고 볼 수 있는 모델이다. 하지만 그래프의 전반적인 등락을 세밀하게 예측하는 것은 한계가 있다. BasicRNN과 GRU의 RMSE는 각각 0.5537, 0.6482로 LSTM보다 RMSE 값이 높고, 실제 금리 스프레드 값과는 차이가 있다. 그러나 두 모델 모두 실제 금리 스프레드의 등락 측면에서는 LSTM 모델보다 정확하게 예측한 측면이 존재한다. Basic RNN은 2011년~2012년 구간을, GRU는 Time Period 2013년 2분기부터 2014년 2분기 구간을 더 잘 예측한 것으로 보인다.

### 4.3 한국 금리 스프레드 예측 결과

앞서 제시한 BasicRNN, LSTM, GRU 모델의 결과를 통해 획득한 미국 금리 스프레드 예측 데이터를 조합하여 SVR 앙상블 모델을 통해 한국의 금리 스프레드 예측 모델을 설계하였다. 파라미터는 그리드 서치를 통해  $\gamma$ (gamma)는 0.333, C(Regulation Parameter)는 30을 사용하였으며, 결과는 <Figure 4>와 같다. 실제 한국 금리 스프레드의 추세를 따라가고 있으며, RMSE 값은 0.0658로 실제 데이터와 근접하게 예측하였다. 하지만 일부 구간 실제 데이터보다 과하게 예측한 구간이 있었으며, 비록 스프레드의 등락은 예측하였지만, 실제 값과 차이가 나는 구간도 존재한다.



<Figure 4> SVR Ensemble



<Figure 5> Normal Ensemble

선행연구에 따르면 다수의 연구에서 예측 모델의 성능 향상을 위하여 균등하게 가중치를 부여한 앙상블 모델을 사용하였다. 이를 일반 앙상블 모델(Normal Ensemble)이라 명명한다. 본 연구의 한국 금리 스프레드의 예측을 위해 사용한 SVR 앙상블 모델의 우수성을 입증하기 위해 일반 앙상블 모델을 대조군으로 설정하여 비교 분석을 하였다. 일반 앙상블 모델을 통해 한국 금리 스프레드를 예측한 결과, <Figure 5>와 같이 실제 한국 금리 스프레드를 예측하였다. RMSE 값은 0.0905로 SVR 앙상블 모델에 비해 높은 값을 나타내었으며, 실제 데이터의 추세를 반영하지 못하는 구간이 존재하였다. SVR 앙상블 모델이 일반 앙상블 모델보다 RMSE와 경향성 측면에서 뛰어난 예측력을 보이며 등락에 기민하게 예측한다. 하지만 두 모델 모두 2017년~2018년 구간에서 예측의 정확도가 다른 기간에 비해 낮음을 확인하였다.

#### 4.4 2017년 전후 비교

2017년 1월부터 4월까지의 금리 스프레드 예측은 실제 금리 스프레드의 등락과 정반대인 것을 알 수 있으며, 이에 대한 보정을 위해 예측 모형을 수정하였다. 실제 금리 스프레드 값과 차이가 가장 많이 났던 2017년을 기점으로 기존 SVR 모델을 나누어 기간별로 다르게 학습한 모델을 설계하였고 이것을 수정된 SVR(Modified SVR)로 명명하였다. <Figure 6(a)>, <Figure 6(c)>는 2011년부터 2019년까지 전체기간을 학습하여, 금리 스프레드를 예측한 결과를 2017년 기준으로 나누어서 나타냈다. <Figure 6(b)>는 2011년부터 2017년까지의 데이터를 학습하여 실제 금리 스프레드와 비교했으며, <Figure 6(d)>는 2017년 이후부터 2019년까지의 데이터를 학습하여 실제 금리 스프레드와 비교한 결과를 나타냈다. 수정된 SVR과 이전 SVR의 모델 비교는 <Table 3>에 정리되어 있다.

<Table 3> Comparison Result of SVR Model : RMSE

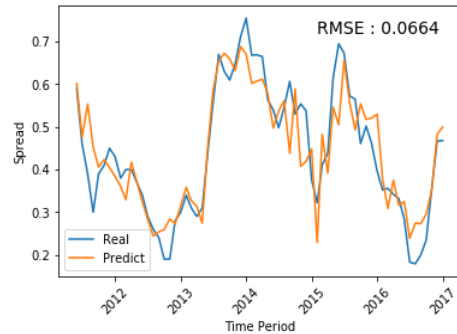
	SVR	Modified SVR
Before 2017	0.0664	0.0602
After 2017	0.0645	0.0392

<Figure 6(a)>에서 제시한 기존 SVR 모델의 금리 예측 스프레드 모델보다 <Figure 6(b)>에서 예측한 금리 스프레드 모델에서 2016년 3분기의 실제 금리 스프레드를 우수하게 예측하였으며, RMSE값도 0.0664에서 0.0602로 감소하였다. <Figure 6(c)>에서는 실제 금리 스프레드와 가장 크게 차이가 났던 2017년 상반기 부분을 우수하게 예측했으며, <Figure 6(d)>와 비교했을 때 RMSE값이 0.0645에서 0.0392로 감소하였다.

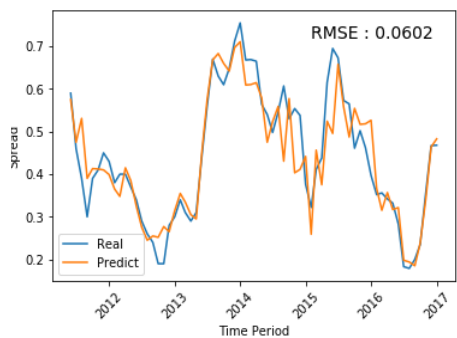
2009년 서브프라임 모기지 사태 이후 미국의 금리시장과 한국의 금리시장이 동조화되었다. 이에 본 연구는 미국 데이터를 이용하여 2009년 이후의 한국 금리 스프레드를 예측하는 모델을 제안하였다. <Figure 4>와 <Figure 5>에서 볼 수 있듯이 한국 금리 스프레드를 예측하는 앙상블 모델과 SVR 모델 모두 2017년을 예측하지 못한 것을 볼 수 있는데, 이는 45대 미국 대통령 트럼프 당선 시기와 맞물리는 것을 알 수 있다. 이 이후의 미국의 경제 정책 방향이 변화가 큰 영향을 미친 것으로 파악할 수 있다.

#### 5. 결론

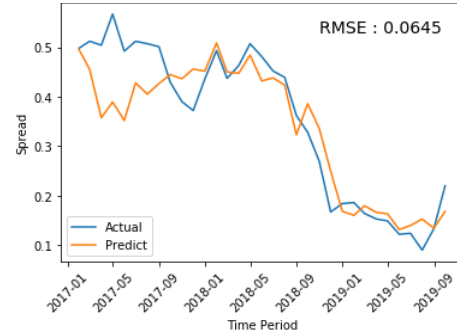
본 연구는 미국 금리 스프레드에 영향을 주는 27가지 요인을 주성분 분석을 통해 새로운 5가지 차원의 독립변수로 차원을 축소하여, 이를 이용한 RNN 모델을 통해



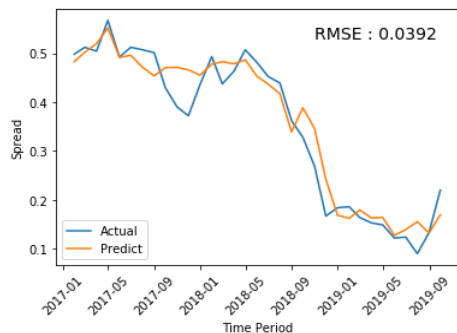
<Figure 6.a> SVR : Before 2017



<Figure 6.b> Modified SVR : Before 2017



<Figure 6(c)> SVR : After 2017



<Figure 6(d)> Modified SVR : After 2017

미국 금리 스프레드를 예측하였다. 그리고 미국 금리 스프레드 예측값들로 SVR 앙상블 모델을 통해 한국 금리 스프레드를 예측하여, 상당히 높은 예측력을 보여주었다.

본 연구에서 구현 모델을 통해 다음과 같은 시사점을 도출할 수 있었다. 첫째, 장기간 수집된 미국의 데이터를 통하여 한국의 금리 스프레드에 대한 매우 높은 예측도를 보여주고 있다. 이는 세계적으로 진행된 경제 동조화의 결과로 보인다. 둘째, 미국의 정책이 급변함에 따라 한국의 금리 스프레드를 예측할 수 없는 구간이 일부 존재한다. 정책의 급격한 변화는 예측에 변동을 발생시키므로, 예측 시 발생할 수 있는 각 국가의 상황을 고려한 모델이 필요하다. 셋째, RNN 모델(BasicRNN, LSTM, GRU) 3가지를 조합한 SVR 앙상블 모델을 통해 향상된 성능의 예측 모델을 제시하였으며, 이는 단순 앙상블보다 높은 성능을 보여주었다.

정확한 예측이 필요한 금리 스프레드를 딥러닝을 이용하기보다 정확성을 높이기 위해 새로운 방법을 추가하였다는 데 연구의 의의가 있다.

이 모델 연구를 통해 서브프라임 모기지 사태 이후 실행한 미국의 정책에 의해 전 세계의 여러 나라가 동조화되어있음을 보여주었고, 미국 금리 스프레드의 요인들을 활용하여 다른 나라의 금리 스프레드를 예측에도 큰 도움이 된다는 것을 보여주었다. 또한, 국가별 예측뿐만 아니라 국제 경제 상황을 나타낸 글로벌 데이터를 사용하여 금리 스프레드를 예측할 수 있을 것이다. 세계화로 인해 전 세계의 경제는 연동되어 있으므로 한나라의 데이터가 아닌 글로벌 데이터를 이용하면 다양한 모델을 구현할 수 있을 것이다.

## Acknowledgement

This paper was written as part of Konkuk University's research support program for its faculty on sabbatical leave in 2019.

## References

- [1] Ahmad, N., Muhammad, J., and Masron, T.A., Factors influencing yield spreads of the Malaysian bonds, *Asian Academy of Management Journal*, 2009, Vol. 14, No. 2, pp. 95-114.
- [2] Althelaya, K.A., El-Alfy, E.S.M., and Mohammed, S., Stock market forecast using multivariate analysis with bidirectional and stacked (LSTM, GRU), In *2018 21<sup>st</sup> Saudi Computer Society National Computer Conference (NCC)*, IEEE, 2018, pp. 1-7.
- [3] Cheng, W., Wagner, W., and Lin, C.H., Forecasting the 30-year U.S. treasury bond with a system of neural networks, *NeuroVe \$t Journal*, 1996, Vol. 1, No. 2.
- [4] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y., Learning phrase representations using RNN encoder-decoder for statistical machine translation, *arXiv preprint arXiv:1406.1078*, 2014.
- [5] Dagum, E.B., Modelling, forecasting and seasonally adjusting economic time series with the X-11 ARIMA method, *Journal of the Royal Statistical Society. Series D(The Statistician)*, 1978, Vol. 27 No. 3/4, pp 203-216.
- [6] Estrella, A. and Hardouvelis, G.A., The term structure as a predictor of real economic activity, *The Journal of Finance*, 1991, Vol. 46, No. 2, pp. 555-576.
- [7] Fah, C.F. and Ariff, M., Factors correlated with treasury bond spreads in an emerging capital market, *International Journal of Humanities and Social Science*, 2008, Vol. 1, No. 5, pp. 154-164.
- [8] Galbraith, J.W. and Tkacz, G., Testing for asymmetry in the link between the yield spread and output in the G-7 countries, *Journal of International Money and Finance*, 2000, Vol. 19, No. 5, pp. 657-672.
- [9] Harvey, C.R., Term structure forecasts economic growth, *Financial Analysts Journal*, 1993, Vol. 49, No. 3, pp. 6-8.
- [10] Hochreiter, S. and Schmidhuber, J., Long short-term memory, *Neural Computation*, 1997, Vol. 9, No. 8, pp. 735-1780.
- [11] Ince, H. and Trafalis, T.B., Short term forecasting with support vector machines and application to stock price prediction, *International Journal of General Systems*, 2016, Vol. 37, No. 6, pp. 677-687.
- [12] Ji, H.J. and Park, S.K., Evaluate the predictive power of the interest rate spread, *The Korean Journal of Financial Management*, 2002, Vol. 19, No. 2, pp. 233-251.
- [13] Joo, I.T. and Choi, S.H., Stock prediction model based on bidirectional LSTM recurrent neural network, *The Journal of Korea Institute of Information, Electronics, and Communication Technology*, 2018, Vol. 11, No. 2, pp. 204-208.
- [14] Kim, H.J., Jung, J.H., Ko, E.N., Cho, M.J., and Lee, K.H., Stock price prediction using deep learning ensemble, *Database Research*, 2018, Vol. 34, No. 2, pp. 111-120.
- [15] Kim, J.H., Jin, D.L., Lee, J.S., Kim, S.J., and Son, Y.S., Prediction of interest rate spread by vector autoregression model, *Journal of the Korean Data and Information Science Society*, 2012, Vol. 23, No. 6, pp. 1093-1102.



- [16] Kim, J.Y. and Lee, H.S., The predictive power of credit spread for economic activity, *POSRI Business and Economic Review*, 2011, Vol. 11, No. 3, pp. 36-63.
- [17] Lee, C.Y., Song, G., and Kim, J., Analyses of power consumption of the heat pump dryer in the automobile drying process by using the principal component analysis and multiple regression, *Journal of Society of Korea Industrial and Systems Engineering*, 2015, Vol. 38, No. 1, pp. 143-151.
- [18] Min, H.G., *Determinants of Emerging Market Bond Spread : Do Economic Fundamentals Matter?*, The World Bank. 1999.
- [19] Park, W.J., Im, B.I., and Jeon, Y.I., Estimation and forecast of korean government bond yields using a macro-finance model of term structures, *Journal of Korean Economic Studies*, 2013, Vol. 31, No. 2, pp. 187-210.
- [20] Shin, D.H., Choi, K.H., and Kim, C.B., Deep learning model for improving stock price prediction rate using RNN and LSTM, *The Journal of Korean Institute of Information Technology*, 2012, Vol. 15, No. 10, pp. 9-16.
- [21] Stock, J.H. and Watson, M.W., New indexes of coincident and leading economic indicators, *NBER Macroeconomics Annual*, 1989, Vol. 4, pp. 351-394.
- [22] Wang, J.H. and Leu, J.Y., Stock market trend prediction using ARIMA-based neural networks, In *Proceedings of International Conference on Neural Networks (ICNN'96)*, 1996, Vol. 4, pp. 2160-2165.

**ORCID**

- Sun-Ho Jeong | <http://orcid.org/0000-0001-9522-3570>
- Young-Hoo Kim | <http://orcid.org/0000-0003-3071-7160>
- Myung-Jin Song | <http://orcid.org/0000-0002-4703-7690>
- Yun-Jae Chung | <http://orcid.org/0000-0003-1140-9779>
- Sung-Seok Ko | <http://orcid.org/0000-0001-8659-007X>