

<http://dx.doi.org/10.17703/JCCT.2024.10.1.489>

JCCT 2024-1-59

## 사회적 변수와 개개인의 감정지수를 함께 고려한 딥러닝 기반 행복 지수 모델 설계

### Deep Learning-based Happiness Index Model Considering Social Variables and Individual Emotional Index

오수민\*, 박민서\*\*

Sumin Oh\*, Minseo Park\*\*

**요약** 행복 지수는 집단적인 행복 정도를 직관적으로 파악하는데 효과적인 측정 시스템이다. 가치관의 변화에 따라 행복 지수에 행동의 가치를 추가한 연구들이 제안되고 있으나, 개인이 느끼는 감정을 활용하여 관계성을 분석한 연구는 부족한 실정이다. 따라서 본 연구는 행동의 가치를 나타내는 사회적 변수와 개개인의 감정지수를 함께 고려해 행복 지수를 예측하는 딥러닝 모델을 설계한다. 첫째, 2005년 1월 ~ 2020년 12월의 사회적, 감정적 변수를 수집한다. 둘째, 데이터 전처리 및 유의변수 탐색을 수행한다. 셋째, 딥러닝 기반의 회귀 모델로 학습하고, 5-Fold 교차 검증(Cross Validation)으로 학습 모델을 평가한다. 본 연구의 제안 모델은 테스트 데이터에서 90.65%의 높은 예측 정확도를 보인다. 향후 이 연구는 국가별 데이터로 확대 적용하여 행복 지수 주요 요인 분석 등의 연구에 활용될 수 있을 것으로 기대된다.

**주요어** : 행복 지수, 감정지수, 사회적 변수, 딥러닝, 회귀 모델

**Abstract** Happiness index is a measurement system for understanding collective happiness. As values change, studies have been proposed to add the value of behavior to the happiness index. However, there is a lack of studies analyze the relationship using individual emotions. Using a deep learning model, we predicted happiness index using social variables and individual emotional index. First, we collected social and emotional variables from January 2005 to December 2020. Second, we preprocessed the data and identified significant variables. Finally, we trained deep learning-based regression model. Our proposed model was evaluated using 5-fold cross validation. The proposed model showed 90.86% accuracy on test sets. Our model will be expected to analyze the significant factors of country-specific happiness index.

**Key words** : Happiness index, Emotional index, Social variables, Deep learning, Regression model

\*준회원, 서울여자대학교 데이터사이언스학과 학부생 (제1저자)Received: October 15, 2023 / Revised: November 2, 2023

\*\*정회원, 서울여자대학교 데이터사이언스학과 조교수 (교신저)Accepted: November 30, 2023

자)

\*\*Corresponding Author: mpark@swu.ac.kr

접수일: 2023년 10월 15일, 수정완료일: 2023년 11월 2일

Dept. of Data Science, Seoul Women's Univ, Korea

게재확정일: 2023년 11월 30일

## I. 서론

1970년대 부탄에서 만든 행복 개념인 국민 총 행복 지수(Gross National Happiness)는 한 국가의 집단적인 행복과 안녕을 측정하는 시스템이다[1]. UN 산하 지속가능 발전 해법 네트워크(Sustainable Development Solutions Network, SDSN)가 공개한 '2023 세계행복보고서'에 따르면, 대한민국의 행복 지수는 5.951로 OECD(Organisation for Economic Co-operation and Development) 회원 38개국 중 35위를 차지했다[2]. 반면, 국제통화기금(International Monetary Fund, IMF)에서 공개한 국내총생산(Gross domestic product, GDP)은 2023년 기준 약 1.7조 달러로 OECD 회원국 중 13위를 차지하였다[3]. 이처럼 한국은 경제적으로 높은 성장을 보임에도 낮은 행복 지수를 가진다[4].

높은 경제 성장에 따른 가치관의 변화로 사람들은 물질적인 가치 외에도 불평등, 공정 등과 같은 사회적 가치를 중시하게 되었다[5]. 이에 따라 행복의 결정요인 분석을 위해 질적인 측면을 고려할 필요가 있다.

행복을 결정하는 질적 요인으로 사회적 행동 가치가 주목받고 있다. 이에 사회적 행동의 가치를 활용한 연구들이 활발하게 이뤄지고 있다[6-9]. 그러나, 기존 연구는 행복을 느끼는 대상인 개개인의 긍정/부정적인 감정을 배제한 사회적 변수에 한정하여 행복을 결정하는 요인을 탐색한다는 한계가 존재한다. 감정 인식과 감정 표현의 수준에 따라 인간관계와 행복에 차이가 존재하기 때문에 사람의 행복 정도를 측정하려면 사람이 느끼는 감정까지도 포함되어야 한다[10].

따라서 본 연구에서는 사회적 변수와 개개인의 감정을 함께 고려한 딥러닝 기반의 행복 지수 예측 모델을 제안하고자 한다. 행복 지수와 관련된 사회적 변수 및 긍정/부정적인 감정 데이터를 수집한 후, 딥러닝 기반의 회귀 모델을 개발하고, K-fold 교차 검증을 통해 정확도를 향상시킨다.

본 논문은 다음과 같이 구성된다. 제 2장에서는 행복 지수를 결정하는 요인을 분석한 선행 연구에 대해 살펴본다. 제 3장에서는 본 연구에서 제안하는 딥러닝 기반의 회귀 모델을 제안한다. 제 4장에서 연구 결과를 설명하고, 제 5장에서 결론을 언급한다.

## II. 관련 연구

본 장에서는 행복 지수를 결정하는 요인을 분석한 연구들을 살펴본다.

김영숙 외[6]는 간호대 학생 220명을 대상으로 수집한 2015년 3월 16일 ~ 2015년 4월 10일의 간호대 학생의 사회학적 특성, 사회적 지지, 자아 탄력성, 스트레스, 전공 만족도 등 5개 변수를 활용하여 행복 지수를 예측하고 각 요인이 행복 지수에 미치는 영향을 분석하였다. 계층적 회귀 분석 결과 사회적 지지, 자아 탄력성, 대인 관계, 스트레스, 전공 만족도 순으로 행복 지수에 영향을 미치는 것을 확인하였다. 사회적 지지가 많을수록, 자아 탄력성과 전공 만족도가 높을수록 행복지수가 높아지고, 스트레스가 많을수록 행복 지수가 낮아지는 것을 증명하였다.

Zyl 외[7]는 남아프리카공화국 대학생 334명을 대상으로 수집한 성별, 나이, 인종, 학습 시간(Full-time/Part-time student), 자기효능감, 삶에 대한 만족도 등 5개 변수를 활용하여 행복 지수를 예측하고 각 요인이 행복 지수에 미치는 영향을 분석하는 회귀 모델을 제안하였다. 자기효능감과 삶에 대한 만족도, 행복 지수 사이에 양의 상관관계가 있는 것을 확인하였고, 자기효능감이 행복에 유의한 영향을 끼친다는 것을 증명하였다.

송은아 외[8]는 보건복지부 소속 전체 5개 국립정신병원에 근무 중인 전체 간호사 249명을 대상으로 수집한 2016년 5월 16일 ~ 2016년 6월 30일의 감정노동의 빈도, 간호업무환경 지수, 자아효능감 등 3개 변수를 활용하여 행복 지수를 예측하고, 각 요인이 행복 지수에 미치는 영향을 분석하였다. 위계적 회귀 분석 결과 자아효능감, 간호업무 환경 지수, 감정노동의 빈도순으로 행복 지수에 영향을 끼치는 것을 확인하였다. 간호업무 환경 지수, 자아효능감이 높을수록 행복 지수가 높아지고, 감정노동의 빈도가 높을수록 행복 지수가 낮아지는 것을 증명하였다.

김진미[9]는 조현병, 기분장애(우울), 불안장애, 외상 후 스트레스 장애의 정신질환을 가지고 있는 환자 223명을 대상으로 수집한 2019년 6월 10일 ~ 2019년 10월 4일의 성별, 연령, 직업, 경제적 상황, 정신질환 진단명, 사회관계망 측정, 사회적 지지 척도, 삶의 만족도 등의 4개 변수를 활용하여 행복 지수를 예측하고, 각 변인

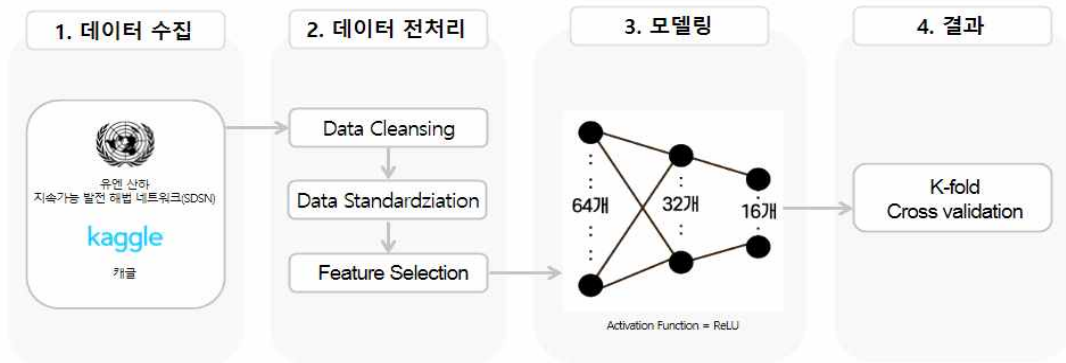


그림 1. 사회적 변수와 개 개인의 감정지수를 활용한 행복 지수 예측을 위한 딥러닝 모델링  
 Figure 1. Flow diagram of deep learning modeling for predicting happiness index using social variables and emotional index

의 영향을 파악하였다. 회귀 분석 결과 삶의 만족도와 행복 지수 사이에는 양의 상관관계가 있음을 확인하였다. 정신질환을 가지고 있는 환자의 개인적 특성을 고려하는 경우 행복 지수를 더 잘 예측할 수 있음을 증명하였다.

### III. 딥러닝 기반 행복지수 모델 설계

본 연구는 딥러닝을 활용하여 행복 지수 모델을 만들고자 한다. 본 연구의 프로세스는 그림 1과 같다.

#### 1. 데이터 수집

지속가능 발전 해법 네트워크(Sustainable Development Solutions Network, SDSN)에서 발표한 ‘세계행복보고서’를 기준으로 2005년 ~ 2020년의 각 국가의 행복 지수, 1인당 실질 GDP(Gross domestic product), 사회적 지원, 삶의 선택을 할 수 있는 자유도, 기대수명, 긍정적인 감정의 정도, 부패 지각도, 부정적인 감정의 정도, 타인에 대한 배품 정도(관대함) 등 11개 변수를 수집하였다. 해당 데이터는 총 1,949개의 수치 데이터로 구성되어 있다.

#### 2. 데이터 전처리

전체 데이터의 수집 시점을 기준으로 일부 데이터가 수집되지 않은 240개의 데이터의 경우, 결측치 처리를 위해 제거하였다.

수집한 데이터를 그대로 모델에 활용하는 경우 너무 큰 값이나 범위를 가지는 변수로 인해 편향(Bias)된 모

델이 도출될 수 있으며, 모델의 성능을 하락시킬 수 있다. 이에 서로 다른 분포를 가지는 변수의 범위를 조정하여 그 차이를 감소시키기 위해 독립변수들의 값의 범위를 평균 0, 표준편차 1의 정규 분포로 변환하는 데이터 표준화(Data Standardization)를 수행하였다.

표준화 한 독립변수를 분산팽창지수(Variance Inflation Factor, VIF)를 활용하여 다중공선성 기준 유의성 평가를 수행하였다. 일반적으로 VIF가 10 이상인 경우, 다중공선성이 존재하는 것으로 간주된다. 11개 변수 중 유의한 변수는 사회적 지원, 삶의 선택을 할 수 있는 자도, 기대수명, 긍정적인 감정의 정도, 부패 지각도, 부정적인 감정의 정도, 타인에 대한 배품 정도(관대함) 등 7개 변수이다. 표 1은 본 연구에서 사용한 7개의 유의한 독립변수와 해당 변수의 VIF이다.

표 1. 독립변수의 분산팽창지수

Table 1. Variance Inflation Factors of Independent variables

독립변수	Variance Inflation Factor
사회적 지원	2.2
삶의 선택을 할 수 있는 자유도	2.0
기대수명	1.9
긍정적인 감정의 정도	1.9
부패 지각도	1.5
부정적인 감정의 정도	1.4
타인에 대한 배품 정도(관대함)	1.3

#### 3. 모델링

데이터 전처리 후 도출된 7개 변수로 구성된 1,708개의 데이터를 행복 지수 예측을 위한 딥러닝 모델에 적용하였다. 실험을 위하여 전체 2005년 ~ 2020년의 데이터

세트에서 훈련 데이터(Training Sets)와 테스트 데이터(Test Sets)를 무작위로 각각 80%, 20%의 비율로 나누어 구성하였다.

실험에 사용된 모델은 입력층과 3개의 은닉층, 출력층으로 구성된다. 입력층에 입력된 데이터는 64개, 32개, 16개의 노드로 구성된 은닉층을 거친 후, 출력층을 통해 예측 값을 출력한다. 각 은닉층의 활성화 함수로 ReLU(Rectified Linear Unit)를 사용한다. 제안 네트워크에 사용된 학습률(Learning Rate)은 0.001, 배치(Batch) 크기는 16을 사용하였다. 최적화 함수는 Adam Optimizer를, 손실함수는 MSE(Mean Squared Error) Loss를 사용하였으며, 총 100번의 반복 학습을 수행하였다. 분석 모델의 정확도를 검증하기 위하여 5-Fold 교차 검증(Cross Validation)을 수행한다.

#### IV. 실험 및 결과 분석

제안 모델의 설명력을 평가하고 검증하기 위해 모델의 설명력을 나타내는 결정 계수(R-squared)와 모델의 예측력을 나타내는 평균 제곱 오차(Mean Squared Error, MSE), 정확도(Accuracy)를 측정하였다. 표 2는 사회적 변수와 개개인의 감정지수를 기반으로 행복 지수를 예측하는 제안 모델의 정량적 평가 결과이다.

표 2. 사회적 변수와 개개인의 감정지수를 활용한 행복 지수 예측 모델의 정량적 평가

Table 2. Quantitative evaluation of happiness index prediction model using social variables and emotional index

데이터 세트	Evaluation		
	R-squared	MSE	Accuracy
Training Sets	0.8841	0.3879	94.03%
5-fold cross validation	0.7890	0.5220	88.82%
Test Sets	0.8217	0.4746	90.65%

결정 계수 값이 1에 가까울수록 설명력이 좋은 모델이며, 평균 제곱 오차 값이 작을수록 실제값과 예측값 간의 차이가 작은 모델이므로 예측 정확도가 높다고 해석할 수 있다. 본 연구의 실험 결과는 표 2에서 볼 수 있듯이, 5-Fold 교차 검증(Cross Validation)에서 결정 계수 0.7890, 평균 제곱 오차 0.5220, 정확도 88.82%로 우수한 성능을 보였으며, 테스트 데이터 셋의 실험 결과 역시 결정 계수 0.8217, 평균 제곱 오차 0.4746, 정확

도 90.65%으로 모두에서 우수한 성능을 보였다.

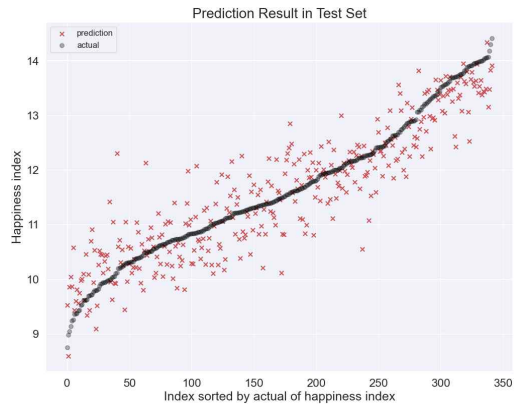


그림 2. 행복 지수 예측 결과 그래프, 빨간색 X는 예측값, 회색 O는 실제값을 의미함

Figure 2. Happiness index prediction result graph, red X means predicted value, gray O means actual value

그림 2는 테스트 데이터를 기반으로 제안한 모델의 성능을 보여주는 그래프로 테스트 데이터의 예측값과 실제값의 차이를 나타낸다. x-축은 랜덤하게 추출한 테스트 데이터를 실제값 기준으로 정렬하여 새롭게 부여한 Index이고, y-축은 해당 Index에 해당하는 행복 지수를 의미한다. 테스트 데이터의 실제값을 회색 O로, 예측값을 빨간색 X로 표현하였다. 이 그래프를 통해 실제 값과 예측값이 유사한 추세로 보여지는 것을 알 수 있다.

#### V. 결론

본 논문에서는 딥러닝을 활용하여 행복 지수를 예측하는 모델을 제안하였다. 행복 지수와 이를 측정하는데 사용되는 사회적 변수 및 긍정/부정적인 감정의 정도를 나타내는 데이터를 수집하였다. 일부 데이터가 수집되지 않아 모델의 예측력에 문제가 될 수 있는 결측치를 제거하고, 데이터 편향(Bias) 문제를 방지하기 위해 데이터 표준화(Data Standardziation)을 수행하였다. 유의한 독립변수를 탐색하기 위해 분산팽창지수(Variance Inflation Factor, VIF)를 기준으로 유의성 평가를 진행하여 변수를 다시 한번 선처리하였다. 해당 변수를 딥러닝에 적용하고, 5-Fold 교차 검증(Cross Validation)을 수행하였다. 실험 결과, 행복 지수 예측을 위한 제안

모델이 우수한 성능을 가짐을 확인하였다.

향후 연구에서는 각 국가별 행복에 영향을 끼치는 요인을 추가적으로 수집하여 국가마다 주요하게 여기는 행복의 조건 및 국가별 특성을 파악하고자 한다.

## References

- [1] Bates, Winton, "Gross national happiness," *Asian Pacific Economic Literature* Vol. 23, No. 2, pp. 1-16, 2009. DOI:10.1111/j.1467-8411.2009.0235.
- [2] Rowan, Andrew N, "World Happiness Report 2023," *WellBeing News* Vol. 5, No. 3, 2023.
- [3] IMF: International Monetary Fund, Available online: <https://www.imf.org>
- [4] Jihyun Lee, "Study on Happiness and Corruption Characteristics: Analysis of OECD Countries and Implications for Korea," *Korean Society of Public Administration Winter Conference 2022*, pp. 1503-1530, 2022. DOI:10.38162/JOME.52.3.2
- [5] Inglehart, Ronald, and Christian Welzel, "Changing mass priorities: The link between modernization and democracy," *Perspectives on politics* Vol. 8, No. 2, pp. 551-567, 2010. DOI:10.1017/S1537592710001258
- [6] Young Suk Kim and Mi Young Han, "Factors influencing happiness index of nursing students," *Journal of Korean Academy of Nursing Administration* Vol. 21, No. 5, pp. 501-510, 2015. DOI:10.11111/jkana.2015.21.5.501
- [7] van Zyl, Yolande, and Manilall Dhurup. "Self-efficacy and its relationship with satisfaction with life and happiness among university students," *Journal of Psychology in Africa* Vol. 28, No. 5, pp. 389-393, 2018. DOI:10.1080/14330237.2018.1528760
- [8] Eun A Song, et al, "Influence of emotional labor, nursing work environment, and self-efficacy on happiness index among national psychiatric hospital nurses in Korea," *Journal of Korean Academy of Psychiatric and Mental Health Nursing* Vol. 26 No. 1, pp. 46-54, 2017. DOI:10.12934/jkpmhn.2017.26.1.46
- [9] Jin mi Kim, "A Study on Social Network Characteristics, Social Support, Functional Recovery, and Life Satisfaction of People with Mental Illness," *Journal of the Convergence on Culture Technology (JCCT)*, Vol. 9, No. 6, pp.

- 85-96, 2023. DOI:10.17703/JCCT.2023.9.6.85
- [10] Ji Su Lee, Seo Yeon Choi, Seung Kyu Nam, "Effects of Relationship Experience and Emotional Expression on Depression, Anxiety, and Happiness," *Journal of Human Studies*, Vol. 45, pp. 71-103, 2021. DOI:10.21738/JHS.2021.12.45.71