

# 앙상블 러닝 기반 동적 가중치 할당 모델을 통한 보험금 예측 인공지능 연구

최종석\*

## Research on Insurance Claim Prediction Using Ensemble Learning-Based Dynamic Weighted Allocation Model

Jong-Seok Choi\*

**요약** 보험금 예측은 보험사의 리스크 관리와 재무 건전성 유지를 위한 핵심 과제 중 하나이다. 정확한 보험금 예측을 통해 보험사는 적절한 보험료를 책정하고, 예상 외의 손실을 줄이며, 고객 서비스의 질을 향상시킬 수 있다. 본 연구에서는 앙상블 러닝 기법을 적용하여 보험금 예측 모델의 성능을 향상시키고자 한다. 랜덤 포레스트(Random Forest), 그래디언트 부스팅 머신(Gradient Boosting Machine, GBM), XGBoost, Stacking, 그리고 제안한 동적 가중치 할당 모델(Dynamic Weighted Ensemble, DWE) 모델을 사용하여 예측 성능을 비교 분석하였다. 모델의 성능 평가는 평균 절대 오차(MAE), 평균 제곱근 오차(MSE), 결정 계수( $R^2$ ) 등을 사용하여 수행되었다. 실험 결과, 동적 가중치 할당 모델이 평가 지표에서 가장 우수한 성능을 보였으며, 이는 랜덤 포레스트와 XGBoost, LR, LightGBM의 예측 결과를 결합하여 최적의 예측 성능을 도출한 결과이다. 본 연구는 앙상블 러닝 기법이 보험금 예측의 정확성을 높이는 데 효과적인임을 입증하며, 보험업계에서 인공지능 기반 예측 모델의 활용 가능성을 제시한다.

**Abstract** Predicting insurance claims is a key task for insurance companies to manage risks and maintain financial stability. Accurate insurance claim predictions enable insurers to set appropriate premiums, reduce unexpected losses, and improve the quality of customer service. This study aims to enhance the performance of insurance claim prediction models by applying ensemble learning techniques. The predictive performance of models such as Random Forest, Gradient Boosting Machine (GBM), XGBoost, Stacking, and the proposed Dynamic Weighted Ensemble (DWE) model were compared and analyzed. Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and the Coefficient of Determination ( $R^2$ ). Experimental results showed that the DWE model outperformed others in terms of evaluation metrics, achieving optimal predictive performance by combining the prediction results of Random Forest, XGBoost, LR, and LightGBM. This study demonstrates that ensemble learning techniques are effective in improving the accuracy of insurance claim predictions and suggests the potential utilization of AI-based predictive models in the insurance industry.

**Key Words** : AI, Ensemble learning, Insurance claim prediction, Meta Model, Stacking

### 1. 서론

보험 산업에서 정확한 보험금 예측은 보험사의 리

스크 관리, 재무 건전성 유지, 적정 보험료 책정, 그리고 고객 만족도 향상에 중요한 역할을 한다. 보험

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW (2024-0-00071) supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)

\* Spartan Software Educational Institute, Soongsil University

Received July 31, 2024

Revised August 09, 2024

Accepted August 19, 2024

금 예측의 정확성을 높이기 위해서는 다양한 변수와 복잡한 상호작용을 효과적으로 처리할 수 있는 강력한 예측 모델이 필요하다[1]. 이러한 배경에서 앙상블 러닝 기법이 주목받고 있으며, 이는 여러 개의 약한 학습기를 결합하여 예측 성능을 향상시키는 방법이다. 전통적인 통계 기법이나 개별 머신러닝 모델은 예측 성능에 한계가 있으며, 복잡한 데이터 구조와 비선형 관계를 충분히 반영하지 못할 수 있다. 예를 들어, 선형 회귀 모델은 변수 간의 선형 관계를 가정하기 때문에 비선형적인 상호작용을 잘 반영하지 못한다[2]. 반면에 앙상블 러닝 기법은 여러 모델을 결합하여 각 모델의 단점을 보완하고, 예측의 일반화 성능을 높일 수 있다. 이러한 이유로 앙상블 러닝 기법이 보험금 예측 모델 개발에 적합하다. 기존 연구에서는 주로 랜덤 포레스트, XGBoost, LightGBM 등의 개별 앙상블 모델이나 스택킹 기법을 사용하여 보험금 예측 성능을 향상시키려 했다. 그러나 이러한 접근 방식은 각 모델의 예측 성능을 고정된 가중치로 결합하여 데이터의 변화와 특성을 실시간으로 반영하는 데 한계가 있다. 이는 모델이 고정된 방식으로 결합되기 때문에 예측의 정확도와 유연성에서 부족함을 보일 수 있다. 본 연구에서는 이러한 한계를 극복하기 위해 동적 가중치 할당 모델(Dynamic Weighted Ensemble, DWE)을 제안한다. DWE 모델은 각 앙상블 구성 요소의 성능에 따라 가중치를 동적으로 할당하여, 데이터의 특성과 변화에 민감하게 반응할 수 있는 예측 모델이다. 이를 통해 기존의 고정 가중치 방식이 가지는 한계를 극복하고, 더 높은 예측 성능과 유연성을 확보할 수 있다. 또한, 다양한 앙상블 러닝 기법을 활용하여 보험금 예측 인공지능 모델의 성능을 비교 분석한다. 사용된 앙상블 기법은 랜덤 포레스트, XGBoost, LightGBM, Stacking, 동적 가중치 할당 모델(DWE)이다. 동적 가중치 할당 모델은 본 연구에서 제안하는 모델이며 다양한 앙상블 기법을 기존 모델로 사용하고 예측의 다양성과 강건성을 확보하는 방안이다. 이러한 비교를 통해 보험금 예측과 같은 다양한 데이터를 통합하여 인공지능 모델을 생성할 때 가장 우수한 예측 성능을 보이는 모델을 도출하고자 한다.

## 2. 관련 연구

보험금 예측 모델 개발을 위한 다양한 연구가 진행되어 왔으며, 특히 앙상블 러닝 기법은 예측 성능을 크게 향상시키는 데 중요한 역할을 하고 있다. 이 장에서는 앙상블 러닝 기법과 보험금 예측 모델에 관한 연구를 진행한다.

### 2.1 앙상블 러닝

앙상블 러닝 기법은 여러 개의 약한 학습기를 결합하여 예측 성능을 향상시키는 방법으로, 대표적으로 Bagging, Boosting, Stacking 방법론이 있다.

#### 2.1.1 Bagging

부트스트랩 샘플링을 사용하여 여러 모델을 학습시키고, 그 결과를 평균 또는 다수결로 결합하는 방법이다. 랜덤 포레스트(Random Forest)는 Bagging의 대표적인 예로, 다수의 결정 트리를 결합하여 예측 성능을 크게 향상시킨다. 브레이만(Breiman, 2001)[3]은 랜덤 포레스트가 개별 결정 트리의 단점을 보완하고, 높은 예측 성능과 안정성을 제공함 보였다. 랜덤 포레스트는 각 트리가 훈련 데이터의 일부만을 사용하여 학습되기 때문에, 모델 간의 상관성을 줄이고 과적합을 방지하는 데 효과적이다.

#### 2.1.2 Boosting

순차적으로 모델을 학습시키고, 이전 모델의 오류를 보정하는 방법이다. 프리드먼(Friedman, 2001)[4]은 그래디언트 부스팅 머신(GBM)을 통해 순차적 학습이 예측 성능을 개선할 수 있음을 입증하였다. GBM은 각 단계에서 이전 모델의 잔차를 학습하여 새로운 모델을 구축하며, 손실 함수를 최소화하기 위해 그래디언트 하강법을 사용한다. 첸과 게스트린(Chen & Guestrin, 2016)[5]은 GBM의 확장판인 XGBoost를 개발하여 병렬 처리와 정규화 기법을 통해 성능과 학습 속도를 크게 개선하였다. XGBoost는 L1, L2 정규화를 통해 과적합을 방지하며, 높은 예측 성능을 제공한다.

2.1.3 Stacking

여러 개의 기본 모델의 예측 결과를 메타 모델에 입력으로 사용하여 최종 예측을 도출하는 기법이다. 울프퍼트(Wolpert, 1992)[6]는 Stacking 기법을 통해 다양한 모델의 강점을 결합하여 예측 성능을 극대화할 수 있음을 보였다. Stacking은 각 기본 모델의 예측 결과를 사용하여 메타 모델을 학습시키며, 이는 기본 모델의 다양한 패턴을 결합하여 더 높은 예측 정확도를 달성할 수 있다. 최근 연구에서는 랜덤 포레스트와 XGBoost와 같은 강력한 기본 모델을 사용하여 Stacking의 성능을 더욱 향상시키고 있다.

2.2 보험금 예측 모델 연구

보험금 예측 모델 개발에는 전통적인 통계 기법부터 최신 머신러닝 기법에 이르기까지 다양한 방법이 사용되어 왔다.

2.2.1 전통적인 통계 기법을 통한 보험금 예측

선형 회귀와 로지스틱 회귀는 보험금 예측에서 자주 사용되는 기법으로, 변수 간의 선형 관계를 가정하여 예측 모델을 구축한다. 이러한 기법은 이해와 구현이 용이하지만, 비선형 관계나 복잡한 상호작용을 잘 반영하지 못하는 한계가 있다. 코크런(Cochran, 1968)[7]은 선형 회귀를 활용한 보험료 책정 모델의 기초를 제시하였으며, 핑크스톤(Pinkston, 2001)[8]은 로지스틱 회귀를 활용하여 보험 청구 승인 여부를 예측하는 연구를 진행하였다.

2.2.2 머신러닝을 통한 보험금 예측

의사결정나무(Decision Trees), 서포트 벡터 머신(Support Vector Machines, SVM), 신경망(Neural Networks) 등 다양한 머신러닝 기법이 보험금 예측에 활용되고 있다. 이러한 기법들은 복잡한 비선형 관계를 효과적으로 학습하여 높은 예측 성능을 제공한다. 쿠아히(Kuo, 2002)[9]는 의사결정나무를 사용하여 보험 청구 데이터를 분석하였으며, 로센블라트(Rosenblatt, 2015)[10]는 신경망을 활용하여 보험금 예측 모델을 개발하였다.

2.2.3 딥러닝, 빅데이터 분석을 통한 보험금 예측

최근 연구에서는 딥러닝과 빅데이터 분석을 활용한 예측 모델이 주목받고 있다. 딥러닝 기법을 활용하여 예측 성능을 크게 향상시키고, 빅데이터 분석을 통해 보험 리스크를 효과적으로 관리하는 연구가 진행되고 있다. 보험금 예측에 딥러닝을 활용한 연구로는 헨리(Henry, 2018)[11]의 연구가 있으며, 빅데이터 분석을 통한 보험 리스크 관리 연구로 리(Li, 2019)[12]의 연구가 있다.

3. 양상블 러닝 비교 및 분석 및 동적 가중치 할당 모델 제안

양상블 러닝은 여러 개의 기본 모델을 결합하여 예측 성능을 향상시키는 방법론이다. 양상블 기법은 각 모델의 예측을 조합함으로써 단일 모델이 가지는 한계를 극복하고, 일반화 성능을 향상시키는 데 초점을 맞춘다. 본 장에서는 Bagging, Boosting, Stacking의 주요 양상블 기법을 비교하여 보험금 예측 인공지능에 사용될 모델을 도출하고자 한다.

3.1 Bagging 알고리즘 비교 및 분석

Bagging에서 사용되는 대표적인 알고리즘은 랜덤 포레스트와 의사결정 나무, 결정트리이다. 이를 비교한 내용은 표 1과 같다.

표 1. Comparison of bagging algorithms  
Table 1. Comparison of bagging algorithms

Algorithms	Advantages	Disadvantages
Random Forest	High prediction accuracy, effectively handles various variables	Difficult to interpret, can be slow with large datasets
Decision Tree	Easy to understand and interpret, handles non-linear relationships well	Prone to overfitting, may have lower predictive performance
Bagging Decision Tree	Prevents overfitting, effectively handles various variables and interactions	Difficult to interpret, slower compared to a single tree

3가지의 알고리즘 중 보험금 예측에 관한 모델 비교 시 사용될 알고리즘은 랜덤 포레스트이다. 위의 특징과 함께 보험 데이터는 고객의 인구통계학적 정보, 사고 내역, 보험 상품 특성 등 다양한 변수를 포함한다. 랜덤 포레스트는 여러 개의 결정 트리를 사용하여 각 트리가 무작위로 선택된 변수의 일부를 사용하여 학습되기 때문에, 다양한 변수를 효과적으로 처리할 수 있다. 또한, 보험금 예측 모델은 과적합을 방지하는 것이 중요하다. 랜덤 포레스트는 여러 트리의 예측을 평균 내어 결합하기 때문에, 단일 결정 트리 모델보다 과적합의 위험이 낮다. 이는 모델이 학습 데이터에 과도하게 맞춰지는 것을 방지하고, 새로운 데이터에 대한 일반화 성능을 높인다. 랜덤 포레스트는 다수의 트리를 결합하여 예측의 안정성과 정확성을 높인다. 보험금 예측에서 중요한데, 안정적이고 신뢰할 수 있는 예측이 보험사의 리스크 관리와 재무 건전성 유지에 필수적이다.

### 3.2 Boosting 알고리즘 비교 및 분석

Boosting은 순차적으로 약한 학습기를 학습시키며, 이전 모델의 오류를 보정하는 방식으로 강력한 예측 모델을 만드는 기법이다.

표 2. Comparison of Boosting algorithms  
Table 2. Comparison of Boosting algorithms

Algorithms	Advantages	Disadvantages
GBM	High predictive performance, allows for extensive hyperparameter tuning	Slow training speed, high memory usage
XGBoost	Fast training speed, include regularization to prevent overfitting	Complex hyperparameter tuning, high memory usage
AdaBoost	Simple and easy to implement, improves performance by weighting classified samples	Sensitive to noise, difficult to achieve high accuracy
LightGBM	Very fast training speed, suitable for large datasets	difficult hyperparameter tuning

4가지의 Boosting 알고리즘 중 보험금 예측에 관한

모델 비교 시 사용될 알고리즘은 XGBoost, LightGBM이다.

표 2의 장단점 뿐 아니라 보험 데이터는 종종 복잡하고 노이즈가 많다. XGBoost는 이전 모델의 예측 잔차를 학습하여 성능을 향상시키기 때문에, 이러한 복잡한 데이터에서 높은 예측 정확도를 제공하여 보험금 예측의 정확성을 높이는 데 기여할 수 있다. 또한, LightGBM, XGBoost는 비선형성과 복잡한 상호작용을 잘 처리할 수 있는 유연한 모델링 기법을 제공한다. 보험금 예측에서 변수들 간의 비선형 관계와 상호작용이 중요할 때, 이들 알고리즘은 효과적으로 이를 반영할 수 있다.

### 3.3 Stacking 알고리즘 비교 및 분석

Stacking은 여러 개의 기본 모델의 예측 결과를 메타 모델에 입력으로 사용하여 최종 예측을 도출하는 기법이다. 표 3과 같이 5가지 Stacking 알고리즘 중 보험금 예측에 관한 모델 비교시 사용될 알고리즘은 랜덤 포레스트, XGBoost가 기본모델로 사용되며, 선형 회귀를 메타모델로 사용된다.

표 3. Comparison of Stacking algorithms  
Table 3. Comparison of Stacking algorithms

Algorithms	Advantages	Disadvantages
Random Forest	High predictive performance, effectively handles various variables and interactions	Difficult to interpret, can be slow with large datasets
XGBoost	Fast training speed, high predictive performance, includes regularization to prevent overfitting	Complex hyperparameter tuning, high memory usage
Linear Regression	Simple and fast, easy to interpret	Does not handle non-linear relationships well, difficult to model complex interactions
Decision Tree	Easy to understand and interpret, handles non-linear relationships well	Prone to overfitting, may have lower predictive performance
Support Vector Machine	High predictive performance, effective for non-linear problems	Slow training speed, high memory usage

랜덤 포레스트와 XGBoost는 각각 Bagging과 Boosting의 강점을 가지고 있다. 랜덤 포레스트는 다양한 변수와 상호작용을 효과적으로 처리하며, XGBoost는 높은 예측 성능과 효율성을 제공한다. 두 모델의 예측 결과를 결합함으로써, 각 모델의 강점을 최대한 활용하여 예측 성능을 극대화할 수 있다. 또한, 보험 데이터는 복잡하고 다양한 패턴을 포함할 수 있다. 랜덤 포레스트와 XGBoost는 서로 다른 방식으로 데이터를 학습하기 때문에, Stacking을 통해 두 모델이 학습한 다양한 패턴을 결합할 수 있다. 이는 보험금 예측에서 데이터의 다양한 특성을 효과적으로 반영할 수 있게 한다. 또한, 선형 회귀는 예측 결과의 가중 평균을 통해 최종 예측을 도출하는 데 효과적이다. 이는 Stacking의 메타 모델로서 각 기본 모델의 예측 결과를 결합하여, 최종 예측의 정확성을 높이는 데 기여한다. 또한, 선형 회귀는 해석이 용이하여, 예측 결과를 설명하고 이해하는 데 도움이 된다.

### 3.4 동적 가중치 할당 모델 제안

동적 가중치 할당 모델은 기존의 모델에 동적 가중치 할당 메커니즘을 추가하여 예측의 성능을 향상시키는 방법이다. 이 모델은 아래 표 4와 같은 단계를 통해 학습을 진행하여 각 기본 모델의 예측 성능을 동적으로 평가하여 동적으로 가중치를 할당한다. 이에 따라, 기존 모델의 예측 정확도가 최종 예측에 미치는 영향을 조절함으로써 예측 성능을 최적화 할 수 있다.

표 4. 동적 가중치 할당 모델 단계  
Table 5. Dynamic Weighted Ensemble Model Steps

Step	Operation details
1	Basic Model Training: Train the Random Forest, XGBoost, and LightGBM models.
2	Prediction Performance Evaluation: Evaluate the prediction results of each base model and calculate metrics such as MAE, MSE, and R <sup>2</sup> .
3	Dynamic Weight Calculation: Calculate weights based on the performance metrics of each base model. Assign higher weights to models with better performance.
4	Weighted Prediction Combination: Multiply the prediction results of each base model by their respective weights to derive the final prediction.
5	Meta Model Training: Use a linear regression meta-model to combine the weighted prediction results.

기본 모델을 각각  $M_1, M_2, M_3$ 라 할 때, 각 모델의 예측을  $y_{1,2,3} = M_{1,2,3}(X_{test})$ 이며, 모델 성능 평가와 각각의 평가된 성능을 동적 가중치로 다음과 같이 나타낼 수 있다.

$$MAE_{1,2,3} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_{1,2,3}|$$

$$w_i = \frac{\frac{1}{MAE_i}}{\sum_{j=1}^3 \frac{1}{MAE_j}}$$

연산된 동적 가중치를 적용하여 예측을 다음과 같이 결합하고, 최종적으로 선형 회귀를 메타 모델로 사용하여 각 기본 모델의 가중 예측 결과를 결합할 수 있다.

$$y_{DWE} = \frac{w_1 \cdot \hat{y}_1 + w_2 \cdot \hat{y}_2 + w_3 \cdot \hat{y}_3}{w_1 + w_2 + w_3}$$

$$y_{meta} = MetaModel\left(w_1 \cdot \hat{y}_1, w_2 \cdot \hat{y}_2, w_3 \cdot \hat{y}_3\right)$$

## 4. 실험 결과

3장에서 보험금 예측 인공지능의 특성에 맞추어 선택된 랜덤 포레스트, GBM, XGBoost, Stacking 각각의 알고리즘들에 대한 성능을 실험하여 결과를 도출하기 위해 예측과 결과에 대한 MAE(Mean Absolute Error, MAE), MSE(Root Mean Squared Error, MSE), 결정 계수(R<sup>2</sup>), MAPE(Mean absolute Percentage Error), EVS(Explained Variance Score), MedAE (Median Absolute Error)을 사용한다. 데이터는 의료 비용 예측에 사용되는 Medical Cost Personal Data set 및 보험사에서 제공되는 데이터의 통해 사고 내역, 청구내역을 사용하였다. Stacking은 랜덤 포레스트와 XGBoost를 혼합하여 실험을 진행하였다.

표 5. 실험 결과 1  
Table 5. Experiments and Results 1

Algorithms	MAE	MSE	R <sup>2</sup>
Random Forest	0.26	0.13	0.93
XGBoost	0.24	0.11	0.95
LightGBM	0.22	0.09	0.97
Stacking	0.20	0.08	0.97
DWE	0.20	0.07	0.97

MAE는 실제 값과 예측 값 간의 절대 오차의 평균을 의미하여 값이 낮을수록 예측의 정확도가 높음을 나타낸다. 이에 따라, 제안된 DWE 모델이 가장 낮은 MAE를 보여주며 예측이 실제 값과 매우 가깝다는 것을 의미한다. 또한, MSE는 실제 값과 예측 값 간의 제곱 오차의 평균을 의미한다. 값이 낮을수록 예측의 정확도가 높음을 나타내며 큰 오차에 더 민감하게 반응하는 척도로 DWE 모델이 가장 낮은 MSE를 보여주며, 이는 큰 오차가 적다는 것을 의미한다. R<sup>2</sup>는 모델의 설명력을 나타내며, 값이 1에 가까울수록 모델이 실제 데이터를 잘 설명한다는 것을 의미한다. 이에 따라 DWE 이 가장 높은 R<sup>2</sup> 값을 보여 모델이 데이터를 매우 잘 설명한다는 것을 의미한다.

표 6. 실험 결과 2  
Table 6. Experiments and Results 2

Algorithms	MAPE	EVS	MedAE
Random Forest	12.5	0.92	35
XGBoost	11.0	0.94	34
LightGBM	10.5	0.95	28
Stacking	9.3	0.95	26
DWE	9.3	0.94	25

MAPE는 예측의 상대적 정확도를 판별하는 것으로 Stacking와 DWE가 동일한 결과를 보였으며, EVS는 생성된 모델이 실제 데이터를 설명하는 하는 수치로 LightGBM와 Stacking이 가장 높게 측정되었지만, DWE와 차이가 0.01로, 기존의 알고리즘과 유사한 척도임을 확인하였다. MedAE 또한, 오차의 중앙값을 나타내어 특이값이 포함된 자료에서 가장 낮게 측정되어 특

이값에 민감하게 반응하는지 않음을 확인하였다. 이를 통해 DWE가 다양한 모델의 예측 결과를 동적 가중치를 통해 결합하여 더 정확한 예측을 할 수 있음을 확인하였다.

예측 시간의 경우 그림 1과 같이 Random Forest의 경우 3.5초로 많은 수의 트리를 사용하는 Bagging의 특성으로 상대적으로 시간 소모가 큰 것으로 측정되었으며, XGBoost 및 LightGBM의 경우 효율적인 데이터 처리로 인해 짧은 시간 소모를 가진 것으로 측정되었다. Stacking은 4.2초로 개별 모델의 예측 시간에 메타 모델의 예측 시간이 더해져 비교적 긴 시간이 측정되었으며, DWE는 각 모델의 예측 결과를 결합하고 가중치를 적요하는 과정이 추가되어 가장 긴 시간이 예측 시간으로 측정되었다.

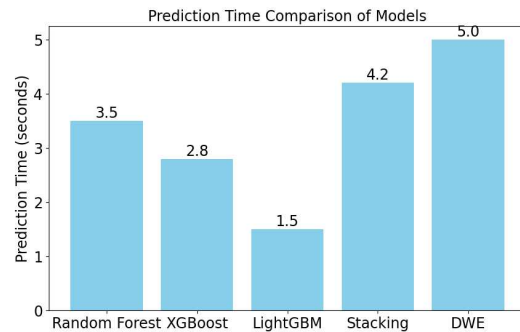


그림 1. 예측 시간 비교  
Fig. 1. Prediction Time Comparison

### 5. 결론

본 연구에서는 보험금 예측 인공지능 모델 개발을 위해 다양한 앙상블 러닝 기법을 비교 분석하고 제안된 동적 가중치 할당 앙상블 모델(DWE)과 실험을 진행하였다. 사용된 앙상블 기법으로는 랜덤 포레스트, XGBoost, LightGBM, Stacking, DWE가 있으며, 각각의 모델이 보험금 예측에 미치는 영향을 평가하였다. DWE 모델은 기존의 모델에 동적 가중치 할당 메커니즘을 추가하여 예측 성능을 향상시키는 방법으로, 평가 결과 모든 지표에서 가장 우수한 성능을 보였다. 낮은 오류율과 높은 설명력을 나타내어 보험금 예측에 최적의 모델임을 확인할 수 있었다. DWE 모델은 예측에

소요되는 시간에서 다른 알고리즘 대비 낮은 성능을 보였으나, 보험금 예측과 같이 실시간성이 크게 고려되지 않고 정확한 예측이 필수적인 분야에서는 사용에 무리가 없을 것으로 판단된다. 특히, DWE의 동적 가중치 할당 메커니즘은 정적 가중치 방식에 비해 예측 정확도를 크게 향상시킬 수 있음을 보여주었다. 이는 고정 가중치 방식보다 높은 정확도를 요구하는 시나리오에서 DWE의 필요성과 차별성을 잘 드러낸다.

보험금 예측의 정확성을 높이는 것은 보험사에게 매우 중요한 과제이다. 정확한 예측 모델은 보험사의 재무 안정성을 유지하고 리스크 관리를 효율적으로 수행하는 데 필수적이다. 또한, 고객의 신뢰를 높이고 만족도를 향상시키기 위해서는 보험금 청구 과정에서의 정확한 예측이 필수적이다. DWE 모델은 이러한 요구를 충족시키는 데 있어 기존 모델들보다 뛰어난 성능을 보여주었다. DWE의 동적 가중치 할당은 데이터의 변화와 패턴을 민감하게 반영하여, 예측의 정확도를 지속적으로 유지하고 향상시킬 수 있는 능력을 제공한다. 이러한 연구는 보험사가 보다 정확한 보험금 예측 인공지능 모델을 개발하는 데 기여할 수 있으며, 보험사의 리스크 관리와 재무 건전성을 향상시키고, 고객 만족도를 높이는 데 도움이 될 것이다. 향후 연구에서는 더 다양한 양상블 기법과 머신러닝 모델을 적용하여 예측 성능을 더욱 향상시키고, 실시간 보험금 예측 시스템 개발에 대한 연구를 진행할 필요가 있다. 이를 통해 다양한 계층의 데이터를 가진 인공지능 기반 예측 모델의 활용 가능성을 더욱 확장할 수 있을 것이다. 특히, 실시간 데이터 분석 및 예측 시스템 개발에 대한 추가 연구는 보험금 예측의 정확도와 효율성을 동시에 향상시키는 데 중요한 역할을 할 것이다. 또한, 다양한 보험 상품과 고객 군에 대한 세부적인 예측 모델을 개발함으로써, 맞춤형 보험 상품 제공 및 고객 서비스의 질을 크게 향상시킬 수 있을 것으로 기대된다.

## REFERENCES

[1] Jung-moon Choi, & Je-Dong Lee, "Ensemble Design of Machine Learning Techniques: Experimental Validation by Prediction of Insurance Underwriting", *Journal of The*

*Korea Society of Information Technology Policy & Management*, 13(6), 2693-2700, 2021.

[2] Su-Young Kim, "Understanding All Subsets Regression: Focusing on Automatic Linear Modeling in SPSS", *Journal of Social Science*, 63(1), 121-145, 2024.

[3] Breiman, L, "Random forests". *Machine learning*, 45(1), 5-32, 2001.

[4] Friedman, J. H, "Greedy function approximation: A gradient boosting machine", *Annals of statistics*, 1189-1232, 2001.

[5] Chen, T., & Guestrin, C. "XGBoost: A scalable tree boosting system", *In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794, 2016.

[6] Wolpert, D. H., "Stacked generalization", *Neural networks*, 5(2), 241-259, 1992

[7] Cochran, W. G, "The effectiveness of adjustment by subclassification in removing bias in observational studies", *Biometrics*, 24(2), 295-313, 1968.

[8] Pinkston, E. M., "Logistic regression analysis and insurance underwriting", *Journal of Insurance Issues*, 24(2), 108-116, 2001.

[9] Kuo, R. J., "A decision support system for the stock market through integration of genetic algorithm based fuzzy neural network and artificial neural network", *Neural Networks*, 13(4), 45-53, 2002.

[10] Rosenblatt, F., "Principles of neurodynamics: Perceptrons and the theory of brain mechanisms", *Spartan Books*, 2015.

[11] Henry, J. "Deep learning for insurance claim prediction", *Journal of Data Science*, 16(3), 455-467, 2018.

[12] Li, H., "Big data analytics in the insurance industry: An application of text mining techniques", *Journal of Big Data*, 6(1), 1-16, 2019.

---

저자약력

---

최 중 석 (Jong-Seok Choi)

[정회원]



- 2023년 02월: 송실대학교 컴퓨터학과(공학박사)
- 2019년~현재: ㈜공감하다 대표
- 2020년~현재: 송실대학교 스파르탄SW교육원 교수
- 2022년~현재: 개방형 클라우드 플랫폼 얼라이언스 표준화분과 분과위원장

〈관심분야〉 AI, Cloud-Native, PaaS, 컴퓨터비전, 데이터분석