

kNN과 앙상블을 이용한 항결핵약 캡슐제의 용출 예측 연구

김선호¹, 최낙훈², 오종석^{3*}

¹동아ST, ²공주대학교 대학원 미래융합공학과, ³공주대학교 미래자동차공학과

A Study on Dissolution Prediction of Anti-Tuberculosis Capsules Using kNN and Ensemble Algorithm

Sun-Ho Kim¹, Nak-Hun Choi², Jong-Seok Oh^{3*}

¹Dong-A ST

²Dep. of Future Convergence Engineering, Graduate School, Kongju National University

³Dep. of Future Automotive Engineering, Kongju National University

요약 이 연구에서는 항결핵약 캡슐 제제에 대하여 머신러닝을 통해 원료의 물리적, 화학적 실험 결과만을 갖고 완성된 제품의 용출률 예측을 진행하였다. 적용된 방법은 제약분야 QbD(Quality by Design)에 적용되는 회귀분석을 이용하여 용출에 중요한 인자를 도출한 후, 머신러닝 기법을 이용하여 학습하고 예측 결과를 도출하여 실제값과 비교 분석하고자 하였다. 우선 시험한 데이터를 입력 조건으로 하고 용출과 정량적으로 연결하는 행위를 통하여 변수들과의 상관관계를 확인하였다. 회귀분석을 통해 변수의 중요도를 파악한 뒤 중요도에 따라서 약물 용출에 영향이 높은 변수를 도출하였다. 다음으로 500번 반복을 통한 최적화 과정을 거친 머신러닝 알고리즘의 결과와 비교하였으며, 사용된 데이터는 수년간의 제약 생산 과정에서 기록된 로트 데이터 중 512개의 데이터를 활용한다. 회귀분석 방법을 통하여 중요 원료 품질특성 18종류를 입력하여 상관성 분석 결과, 예측에 영향을 미치지 않는다고 판단되는 인자 10가지를 제외하였다. 위 결과를 토대로, 머신러닝을 통한 용출률 예측 결과 95% 예측 정확도가 도출되었으며, 이는 현장에서 실용적으로 즉시 사용 가능한 결과이다. 본 연구를 통하여 원료 품질특성만을 이용해 완제품의 용출을 예측할 수 있다면 초기 개발 시 연구 기간을 단축시킬 수 있으며 더 낮은 비용으로 의약품 생산하여 더 나은 환자 치료에 기여할 수 있을 것이다.

Abstract In this study, the dissolution of the finished product was predicted using only the physical and chemical experimental results of the raw material by machine learning for an anti-tuberculosis drug capsule. The method used was to derive important factors for elimination using regression analysis in the pharmaceutical field QbD (Quality by Design) and then to derive learning and prediction results using machine learning techniques to compare and analyze them with actual values. First, the correlation between variables was confirmed by quantitatively connecting dissolution as an input condition with the tested data. After determining the importance of variables through regression analysis, variables with strong effects on drug dissolution were derived according to their importance. Next, it was compared with the results of a machine learning algorithm that has undergone an optimization process through 500 iterations. The data used utilized 512 data points among lot data recorded during the years of pharmaceutical production. Through the regression analysis method, 18 important raw material quality characteristics were input. Correlation analysis revealed 10 factors that did not affect the prediction and were excluded. The prediction accuracy of 95% for the dissolution rate was derived as a result of prediction through machine learning, which can be used practically in the field. Drug production and better patient treatment can be accomplished in less time and at a lower cost during initial development if the dissolution of the finished product can be predicted using only the raw material quality characteristics.

Keywords : CMA, Cycloserine, Dissolution Predict, Machine Learning, Pharmaceutical Manufacturing, QbD

본 논문은 2022년 중소기업부 산학연 Collabo R&D 예비연구과제 (과제번호 S3246206)의 지원을 받아 연구하였음.

*Corresponding Author : Jong-Seok Oh(Kongju National Univ.)

email: jongseok@kongju.ac.kr

Received December 14, 2022

Revised December 29, 2022

Accepted January 6, 2023

Published January 31, 2023

1. 서론

현재 제약 산업에서 신약개발은 시간과 비용이 많이 필요하며, 약물 개발 프로세스에서 보다 효율적이고 체계적인 방법이 요구된다[1]. 전통적으로 제품 품질 검사를 위해서는 공정 중 일부 샘플을 채취하여 함량, 용출 등의 QC(Quality Control) 시험이 진행되기 때문에 완제품 품질을 판단하기 위해서 상당한 시간과 노력이 소비되어 원가 상승의 원인이 된다. 또한 사후적 QC시험으로 인해 공정이 끝난 뒤에야 품질을 확인할 수 밖에 없어 완제품에 대한 사전 혹은 실시간 품질 예측 분석이 필요한 시점이다.

이에 제품 품질을 예측하여 할 수 있는 다변량 분석을 채택하고 구현하는 방법이 중요해졌으며[2], 그 방법 중 하나가 QbD 이다. QbD는 가능한 한 적은 인력, 원료 소비로 우수한 품질을 예측하는 것을 목표로 한다[3]. 하지만 컨트롤 가능한 변수에 한계가 있으며 함량, 용출과 같은 핵심 품질 특성에 영향을 미치는 프로세스를 변경에 할 시에 결과 값 예측에 있어 3개 이상의 공정 변수를 동시에 변경하기는 어렵다.

이를 해결하는 방법으로, 약제학적 예측을 위한 알고리즘과 데이터에 적합한 평가 기준을 적용한 머신러닝과 같은 인공지능이 제약산업에 적용되고 있다. 약학과 인공지능의 통합으로 제약 연구의 패러다임을 경험 의존 연구에서 데이터 중심 방법론으로 전환하는 연구가 추진 중이다[4]. 머신러닝 및 딥러닝의 방법으로 제약 제제의 각 변수간의 복잡한 상관관계를 찾을 수 있으며, 이를 의약품 완제품 예측에 적용할 수 있다. 복잡한 구조에서 직접 학습할 수 있고 신제품에 적용할 수 있는 자동화된 프로세스로 기존방법을 대체할 방법이 학계에 제시되었다[5-7]. 머신러닝을 이용한 빅 데이터를 처리하는 프로세스는 궁극적으로 제품 품질을 향상시켜 환자에게 효과적이고 안전한 의약품 사용에 도움을 줄 수 있다고 알려져 있다[8,9].

본 연구에서는 항결핵약 캡슐 제제에 대하여 원료의 물리 및 화학적 실험 데이터만을 이용하여 완제품의 용출률을 예측 할 수 있는 머신러닝 알고리즘을 제시한다. 분석대상으로 완제품 품질 평가 항목 중 임상과 관련성이 많은 용출을 연구하였으며 선정 사유는 다음과 같다. 용출이란 약물입자가 시간에 따라 용해되는 속도를 의미하며, 약물 제제로부터 약물입자가 용출될 때 표면의 약물분자가 용액 중에 녹아 생체막에 접하고 그 결과 체내에 흡수가 일어나게 된다. 결론적으로 정제가 체내에 흡

수되기 위해서는 용출 과정이 필수적인데 제형 및 공정 과학 관점에서, 정제 용출이라는 물리, 화학적 특성이 생물학적으로 관련이 있으며 환자의 생체 흡수율 등 제품의 효능 및 안전성에 지대한 영향을 미칠 수 있으므로 용출은 주요 완제품 품질 평가 항목으로 인식된다[10]. 또한 용출에 영향을 미치는 원료의 품질 매개변수는 pH, 함량, 순도 등 다양한 변수가 있으며, 물리 화학적 특성에 약간의 변화라도 영향을 미칠 수 있어 여러 가지 변수들로부터 용출을 예측하기 위해 본 연구를 진행하였다.

기존에 머신러닝을 이용한 제약 공정 연구는 부족한 상황이며 앞으로 제형 및 공정 과학 관점에서 임상적으로 중요한 용출 패턴을 예측하는 것은 대단히 중요한 연구이다. 기존의 용출 예측 연구는 주성분이 아닌 일부 종류의 부형제를 이용해 진행되었다. 또한 원료 품질 특성뿐 아니라 공정변수를 모두 포함하여 용출 예측이 진행되었다. 또한 대규모 제조단위에서 다수의 원료 품질 특성을 이용한 캡슐제에 대한 용출 예측 연구는 진행되지 않았다.

따라서 이 연구의 목표는 대규모 제조단위에서 항결핵약 주성분을 함유한 캡슐제에 대하여 다수의 원료의 물리적, 화학적 특성을 시험한 결과값 만을 가지고 완제품 용출률을 예측 할 수 있는 방법을 제시하고자한다.

2. 연구 방법

2.1 제조 공정

캡슐제 정제 공정은 원료 무게측정, 혼합을 거쳐 캡슐 충전 공정으로 이루어졌다. 해당 공정은 철저하게 온도도가 관리되는 공간에서 진행하였다. 정확하게 무게가 측정된 주원료 Cycloserine과 Talc가 혼합기에 차례로 투입되어 혼합 속도 5rpm으로 25분 동안 혼합되었다. 해당 과정을 통해 혼합된 원료가 공캡슐에 충전되었으며, 캡슐 충전기에 공캡슐을 장착 후 100~140cpm의 속도로 충전되었다.

2.2 용출 시험 방법

완제품의 용출 시험에서, 캡슐 정제의 용출시험은 USP(United States Pharmacopeia) 회전검체통법을 사용하여 수행하였다. 약물 용해용 배지는 900mL의 완충액(pH 6.8, 100rpm, 37±0.5℃)으로 구성하였다. 각 샘플에 대한 실험 횟수는 6회였다. 미리 결정된 시점에

서 용해 매질로부터 각각 분취량을 수집하였다. 약물 농도는 시린지 필터로 시료 용액을 여과한 후 HPLC-UV 방법으로 구하였다. 컬럼 온도, 주입 부피 및 이동상의 유속은 각각 30 ℃, 10 μL 및 1.0 mL/min으로 유지되었다.

2.3 데이터 전처리

입력 변수로 원료의 속성을 검사하는 함량, 함습도, 순도 및 기타 물리적 특성을 선정하였다. Table 1과 같이 원료의 조성분 및 부형제에 대한 핵심재료특성 18 종류를 모두 포함하였다. 데이터 취합은 캡슐 공정이 종료된 후 각 로트 단위별로 시험을 한 결과 값으로, 시험은 USP, KP(Korean Pharmacopeia) 및 IP(International Pharmacopeia) 시험법에 의거하여 진행되었다[11,12].

Table 1. Input data of materials attributes for analysis

Name	Stage	Material attributes
S01 -S10	Cyclo serine	Residue on ignition, Loss on drying, Unknown related substances, Specific optical rotation, API contents, Total impurity, pH, condensation product[USP], condensation product[IP], Contents[CAI]
S11 -S18	Talc	Loss on ignition[KP], Magnesium content, Water-soluble substances[KP], Acid-soluble substances, Acidity, Calcium, Water-soluble substances[USP], Loss on ignition[USP]

앞선 과정을 통해 용출률과 관련된 인자만을 활용하여 바로 머신러닝에 적용할 수 있지만, 각 인자가 갖는 값의 범위나 단위가 서로 상이하여 원활한 학습이 이루어지지 않을 수 있다. 따라서 본 연구에서는 이러한 문제를 해결하기 위해 Eq. (1)을 통해 데이터 정규화 과정을 진행하였다.

$$M = \frac{C - (C_{\min} - W_n \times D)}{(C_{\max} + W_n \times D) - (C_{\min} + W_n \times D)} \quad (1)$$

Eq. (1)에서, M은 정규화 값, C는 정규화 하고자 하는 값, Cmax과 Cmin는 각 인자의 최대 최소값, D는 Cmax과 Cmin의 차이값, Wn인자는 정규화 범위를 조절하는 인자가 들어가게 된다. 일반적인 정규화 범위는 0 ~ 1사이의 범위로 정규화 되지만, 머신러닝을 위한 데이터 정규화 진행 시 0~1의 범위보다 축소하여 진행할 경우, 좀 더 나은 분류 성능을 기대할 수 있다는 선행 연

구를 바탕으로, Wn 인자에 0.125를 적용하여 0.1 ~ 0.9사이의 범위로 정규화 과정을 진행하였다[13].

2.4 머신러닝 알고리즘

머신러닝에 사용된 데이터는 앞선 과정에서 정규화된 512개의 데이터 중 412개의 데이터를 학습용 데이터로, 나머지 100개의 데이터를 학습된 모델을 검증하는 테스트 데이터로 구분하여 진행하였다. 본 연구에 사용된 알고리즘은 kNN(k-Nearest Neighbor), 앙상블(Ensemble) 2가지 알고리즘을 사용하여 예측을 진행하였다.

kNN 알고리즘은, 학습 또는 예측하고자 하는 값 X로부터 가장 가까운 거리에 위치한 k개 만큼의 레이블을 확인 후, 가장 많이 도출된 레이블을 값으로 취하는 방식으로 작동되어 다른 알고리즘과 비교 시 간단하면서 훈련 속도가 빨라 사용이 쉽다는 장점이 있다[14,15].

앙상블 알고리즘은 다양한 학습 모델을 하나의 모델로 사용하는 알고리즘이다[14]. 2가지 알고리즘에 사용될 하이퍼파라미터를 찾기 위해 Random Search를 사용하였으며, 다양한 조합으로 총 500번 반복 진행하여 최적 값을 도출하였다. 최적화한 하이퍼 파라미터의 범위 및 최적화된 값은 Table 2와 같다.

Table 2. Hyper Parameter Optimization Range and Optimized Values by Each Algorithm

Algorithm	Hyper Parameter	Range	Optimized Values
kNN	Number of Neighbor	1 ~ 206	16
	Measuring Distance	Cityblock, Chebychev, Correlation, Cosine, Euclidean, Hamming, Jaccard, Mahalanobis, Minkowski, Seclidean, Spearman	Correlation
	Distance Weight	Equal, Inverse, Squared Inverse	Squared Inverse
	Data Standardization	True, False	False
Ensemble	Ensemble Method	Bagging, Ada Boost, Rus Boost	Ada Boost
	Number of Learners	10 ~ 500	329
	Learning Rate	0.001 ~ 1	0.1932
	Number of Predictors to Sample	1 ~ 411	60

3. 실험 결과

3.1 회귀분석

3.1.1 다변량 회귀분석

독립변수와 종속변수의 상관관계를 분석하여 변수 간 회귀방정식을 구축하고 변수 중요도 및 결과값 예측을 위해 회귀모형을 사용하였다. 회귀 적합 방정식은 Table 1과 같이 18개의 인자를 사용하였다.

Table 3. Abbreviations for Each Variable in Study

Predictor	Name
A	Residue on ignition
B	Loss on drying
C	Unknown related substances
D	Specific optical rotation
E	API contents
F	Total impurity
G	pH
H	Condensation Product[USP]
J	Contests[CA]
K	Condensation Product[IP]
L	Loss on ignition
M	Manesium content
N	Water-soluble substances[KP]
O	Acid-soluble substances
P	Acidity
Q	Calcium
R	Loss on ignition[1]
S	Water-soluble substances[USP]

3.1.2 중요도 분석

용출에 영향을 미치는 중요한 요소를 식별하고 정량화하고자 Pareto 차트를 이용하였다. Pareto 차트에서는 반응 변수에 대한 독립변수 항의 상대적 중요도를 나타내며 Fig. 1과 같이 나타내었다. 차트에서 막대는 내림차순으로 정렬되며 막대가 길수록 반응변수에 미치는 영향이 큰 것을 의미한다. 해당 결과를 기반으로, 회귀분석 및 머신러닝 모델 품질을 높이기 위해 중요하지 않은 요소를 제거하고자 하였다. Fig. 1에서 통계적으로 유의한 효과를 나타내기 위해 붉은색 참조선을 표시하였으며, 참조선보다 왼쪽에 있는 요인들은 다변량 모델링 출력에 큰 영향을 미치지 않았다고 판단하였다. 해당 결과, 총 18개의 변수 중에 8개의 요소가 약물 용출에 유의미한 영향을 미치는 것으로 인식되었으며 나머지 10개의 변수는 분석 시에 제외하도록 한다.

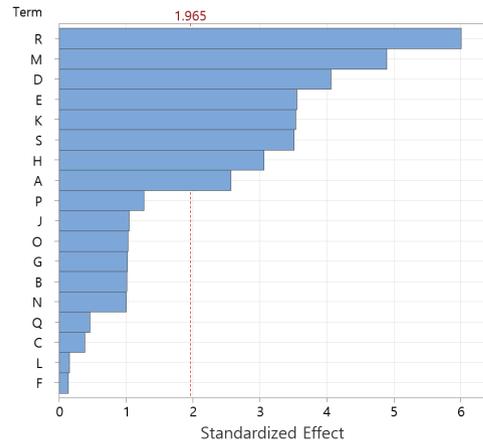


Fig. 1. Pareto chart used to obtain the significant factors that affect the API Dissolution

3.1.3 정확도 분석

용출률 예측 분석 결과에 따라 Pareto 차트를 이용하여 중요 인자를 선별하였고 예측 품질을 더욱 향상시키기 위해 통계적으로 유의하지 않은 변수들을 제외하여 Eq. (2)와 같은 회귀 방정식을 도출하였다. 해당 수식의 약어에 대한 항목은 Table 3를 참조한다.

$$Dissolution = 7.7 - 1.55[A] + 0.1646[D] + 0.03018[E] + 1.978[H] - 3.09[K] + 1.007[M] + 4.388[R] + 0.3099[S] \quad (2)$$

도출된 Eq. (2)에 테스트 데이터를 적용하여 실제값과 예측값이 얼마나 일치하는지 정확도를 확인하였다. 해당 과정에서 일반적인 의약품의 동등성 검증을 위해서는 용출률 10% 미만의 오차를 가지는 것이 일반적이나, 본 연구에서는 예측값과 실제값 사이의 오차범위로 ±1% 이하를 적용하였으며 해당 과정은 Fig. 2와 같이 표현할 수 있다. Fig. 2와 같이 분류된 데이터를 Eq. (3)에 적용하여 예측 정확도를 도출한 결과, 64%의 예측 정확도를 확인할 수 있었다.

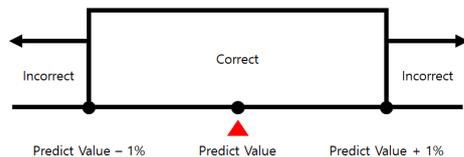


Fig. 2. Criteria for judging correct or incorrect

$$Accuracy = \frac{Correct}{Correct + Incorrect} \quad (3)$$

3.2 머신러닝 해석 결과

3.2.1 kNN 알고리즘

kNN 알고리즘을 통해 용출률 예측 학습 결과의 혼동 행렬은 Fig. 3과 같다. 학습된 모델에 테스트 데이터를 적용한 결과, 88%의 정확도로 예측 가능한 것을 확인할 수 있었으며, 실제 용출률과 예측결과의 산포도를 작성하면 Fig. 4와 같이 작성할 수 있다.

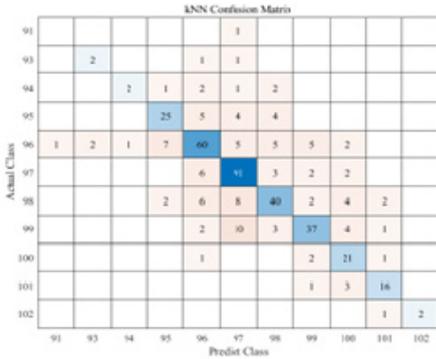


Fig. 3. Confusion Matrix of Dissolution Train Result by kNN Algorithm

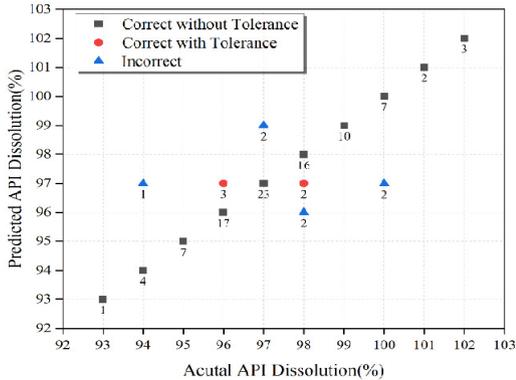


Fig. 4. Scatter Plot of Dissolution Test Data by kNN Algorithm

Fig. 4의 산포도는, 회색 사각형은 정확하게 맞춘 결과를, 빨간색 원은 오차범위를 감안 하였을 때 옳은 결과를, 파란색 삼각형은 오 분류된 값을 의미하며, 각 위치마다 중복된 값이 존재하여 해당 위치에 중복된 값의 개수를 표시하였다. 해당 용출률 예측을 진행에 사용된 알고리즘은 kNN 알고리즘으로, 학습에 사용된 레이블로만 예측값을 도출하는 알고리즘의 한계가 존재한다. 따라서 이로 인해 발생하는 분류 오차를 줄이기 위해 회귀 분석

의 정확도 분석에 적용된 $\pm 1\%$ 의 오차범위를 적용한 결과, 93%의 정확도로 용출률 예측이 가능함을 확인하였다.

3.2.2 앙상블 알고리즘

앙상블 알고리즘을 통해 용출률 예측 학습 결과의 혼동 행렬은 Fig. 5와 같다. 학습된 모델에 테스트 데이터를 적용한 결과, 88%의 정확도로 kNN 알고리즘과 동일한 예측성능을 확인할 수 있었으며, 실제 용출률과 예측결과의 산포도를 작성하면 Fig. 6과 같이 작성할 수 있다.

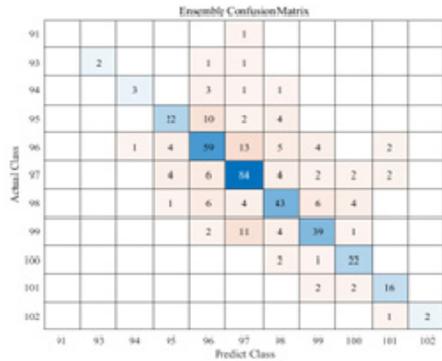


Fig. 5. Confusion Matrix of Dissolution Train Result by Ensemble Algorithm

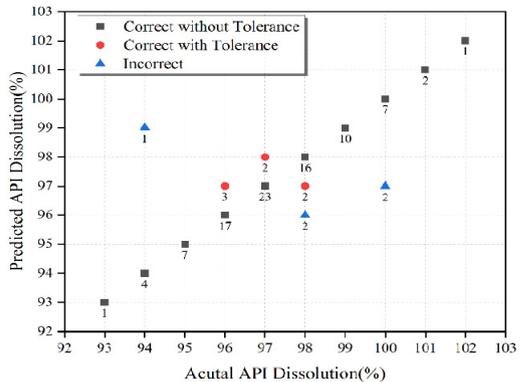


Fig. 6. Scatter Plot of Dissolution Train Result by Ensemble Algorithm

Fig. 6의 산포도 또한, Fig. 4와 같은 구성으로 되어있다. 앙상블 알고리즘을 통한 용출률 예측결과에도 $\pm 1\%$ 의 오차범위를 적용한 결과, 95%의 정확도로 동일한 조건에서 kNN 알고리즘 보다 2% 높은 정확도로 용출률 예측이 가능함을 확인할 수 있다.

3.3 회귀분석과 머신러닝 비교

Eq. (2)를 통해 도출된 회귀 방정식을 이용한 용출률 예측결과와 머신러닝 결과를 비교 진행하였다. 회귀 분석을 통한 예측값의 산포도는 Fig. 7과 같으며, 머신러닝을 통한 예측값과의 비교하기 위한 표는 Table 4와 같다.

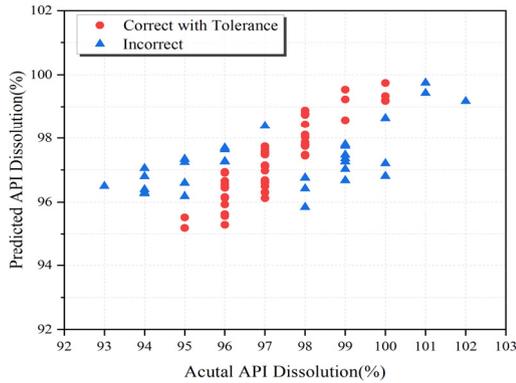


Fig. 7. Scatter Plot of Dissolution Test Data by Regression Analysis

Table 4. Comparison of Accuracy by Analysis Method

Method		Accuracy (%)
Regression Analysis	Without Tolerance	-
	With Tolerance	64
kNN Algorithm	Without Tolerance	88
	With Tolerance	93
Ensemble Algorithm	Without Tolerance	88
	With Tolerance	95

Table 4에서 볼 수 있듯이 회귀분석을 통해 예측한 결과보다, 머신러닝 알고리즘을 사용한 예측결과가 최대 31% 높게 도출됨을 확인할 수 있다. 이러한 결과로 미루어 보아, 머신러닝의 방법을 사용하면 용출률 예측 모델 정확도를 크게 향상시키며, 결과값 예측에 상당한 이점이 있다는 것을 확인할 수 있다.

4. 결론

본 연구에서는 머신러닝을 통한 항결핵약 캡슐제의 제조공정 데이터를 활용하여 용출률 예측을 진행하고자 하였다. 해당 과정에서 모든 인자를 머신러닝에 적용하는 것이 아닌, 선정된 인자를 머신러닝 알고리즘에 적용하고자 하였다. 이후 선정된 인자를 활용하여 머신러닝을

통한 예측이 우수함을 확인하기 위해 기존 제약 분야에서 사용된 회귀분석을 통해 도출된 용출률과 비교하고자 하였다. 이러한 비교는 학습에 사용된 레이블을 통해 결과값을 도출하는 머신러닝의 특징으로 인해 머신러닝에 있어 좀 더 유리한 결과가 도출될 수 있으나, 이러한 문제를 해결하기 위해 통상적으로 사용되는 캡슐제의 동등성 검증을 위한 10%의 미만의 오차범위보다 낮은 $\pm 1\%$ 의 오차범위를 적용하였다.

1. 적용 후 결과 비교 시, 회귀분석을 통한 용출률 예측 결과는 64%, kNN 알고리즘은 93%, 앙상블 알고리즘의 결과는 95%의 정확도로 도출되었다.
2. 용출률 예측이 상당한 정확도로 가능한 것으로 미루어 보아, 기존의 회귀분석보다 머신러닝을 통한 예측이 효과적이며, 다양한 알고리즘 중 앙상블 알고리즘이 가장 높은 정확도로 예측할 수 있음을 확인하였다.

회귀분석보다 머신러닝이 좀 더 예측성능이 높은 이유로, 회귀분석의 경우 분류를 위한 경계가 선형성을 갖고 있으나 머신러닝 알고리즘의 경우 비선형성의 결정경계를 가질 수 있어 성능 차이가 발생한 것으로 사료된다. 나아가, 같은 머신러닝 중 다른 알고리즘 대비 앙상블 알고리즘에서 가장 높은 예측 정확도가 도출된 이유는, 앙상블 알고리즘이 다수의 알고리즘을 통한 결과를 바탕으로 도출되는 특성이 있어 두 알고리즘 사이의 성능 차이가 발생한 것으로 사료된다.

References

- [1] W. Zhang, Q. Zhao, J. Deng, Y. Hu, Y. Wang, D. Ouyang, "Big data analysis of global advances in pharmaceuticals and drug delivery 1980-2014", Drug Discovery Today, vol. 22, no. 8, pp. 1201-1208, Aug. 2017. DOI: <https://doi.org/10.1016/j.drudis.2017.05.012>
- [2] Guideline, ICH Harmonised Tripartite. "Pharmaceutical Development Q8 (R2)." Current step 4, 2009.
- [3] L. X. Yu, "Pharmaceutical quality by design: product and process development, understanding, and control", Pharm. Res. vol. 25, no. 4, pp. 781-791, Apr. 2008. DOI: <https://doi.org/10.1007/s11095-007-9511-1>
- [4] Y. Yang, Z. Ye, Y. Su, Q. Zhao, X. Li, D. Ouyang, "Deep learning for in vitro prediction of pharmaceutical formulations", Acta pharmaceutica sinica B, vol. 9, no. 1, pp. 177-185, Jan. 2019. DOI: <https://doi.org/10.1016/j.apsb.2018.09.010>

- [5] S. Faghih-Roohi, S. Hajizadeh, A. Núñez, R. Babuska, B. De Schutter, "Deep convolutional neural networks for detection of rail surface defects", 2016 International joint conference on neural networks (IJCNN), IEEE, BC, Canada, pp.2584-2589, Jul. 2016. DOI: <https://doi.org/10.1109/IJCNN.2016.7727522>
- [6] D. Tabernik, S. Šela, J. Skvarč, D. Skočaj, "Segmentation-based deep-learning approach for surface-defect detection", Journal of Intelligent Manufacturing, vol. 31, pp.759-776, Mar. 2020. DOI: <https://doi.org/10.1007/s10845-019-01476-x>
- [7] P. H. Chen, S. S. Ho, "Is overfeat useful for image-based surface defect classification tasks?", 2016 IEEE international conference on image processing (ICIP), IEEE, AZ, USA, IEEE, pp.749-753, Sep. 2016. DOI: <https://doi.org/10.1109/ICIP.2016.7532457>
- [8] S. J. Qin, L. H. Chiang, "Advances and opportunities in machine learning for process data analytics", Computers & Chemical Engineering, vol. 126, pp. 465-473, July 2019. DOI: <https://doi.org/10.1016/j.compchemeng.2019.04.003>
- [9] V. Venkatasubramanian, "The promise of artificial intelligence in chemical engineering: Is it here, finally?", AIChE Journal, vol. 65, no. 2, pp.466-478, Feb. 2019. DOI: <https://doi.org/10.1002/aic.16489>
- [10] G. L. Amidon, H. Lennernäs, V. P. Shah, & J. R. Crison, "A theoretical basis for a biopharmaceutical drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability", Pharmaceutical research, vol. 12, no. 3, pp.413-420, Mar. 1995. DOI: <https://doi.org/10.1023/A:1016212804288>
- [11] The United States Pharmacopeia: Cycloserine; United States Pharmacopoeial convention, Inc.: Rockville, MD, USA, 2019; Volume 42, pp. 1187-1188.
- [12] World Health Organization. Monographs: Pharmaceutical substances: Cycloserine (Cycloserinum). The International Pharmacopoeia, vol. 9, 2019. Available From: <http://apps.who.int/phint/en/p/docf/> (accessed March 2022).
- [13] T. N. Phyu, "Survey of classification techniques in data mining", Proceedings of the international multiconference of engineers and computer scientists, Vol. 1, No. 5, Citeseer. Mar. 2009.
- [14] Z. Zhang, "Introduction to machine learning: k-nearest neighbors", Annals of translational medicine, vol. 4, no. 11, pp. 218, Jun. 2016. DOI: <https://doi.org/10.21037/atm.2016.03.37>
- [15] T. G. Dietterich, "Ensemble methods in machine learning", International workshop on multiple classifier systems, Springer, Berlin, Heidelberg, pp. 1-15, Jun. 2000. DOI: https://doi.org/10.1007/3-540-45014-9_1

김 선 호(Sun-Ho Kim)

[정회원]



- 2016년 2월 : 단국대학교 약학과
- 2018년 3월 ~ 현재 : 단국대학교 약학과 박사과정
- 2016년 1월 ~ 현재 : 동아ST 품질보증1팀 (책임)

〈관심분야〉

제약, 스마트 팩토리, 불량 예측

최 낙 훈(Nak-Hun Choi)

[준회원]



- 2021년 2월 : 공주대학교 기계자동차공학부 (공학사)
- 2021년 3월 ~ 현재 : 공주대학교 미래융합공학과 석사과정

〈관심분야〉

스마트 팩토리, 머신러닝, 불량 예측

오 증 석(Jong-Seok Oh)

[정회원]



- 2009년 2월 : 인하대학교 기계공학과 (기계공학사)
- 2011년 2월 : 인하대학교 기계공학과 (기계공학석사)
- 2015년 8월 : 인하대학교 기계공학과 (기계공학박사)
- 2017년 2월 : 현대모비스 미래기술연구팀 (책임연구원)
- 2017년 3월 ~ 2021년 2월 : 공주대학교 기계자동차공학부 조교수
- 2021년 3월 ~ 현재 : 공주대학교 미래자동차공학과 부교수

〈관심분야〉

기계 및 자동차 계측 제어 시스템 설계