

머신 러닝 회귀 방안을 이용한 인공지능기공 크기 예측모델 성능에 관한 연구

이승연*·허용정**†

*한국기술교육대학교대학원 메카트로닉스공학과,
**한국기술교육대학교 메카트로닉스공학부

A Study on Prediction Model Performance of Scaffold Pore Size Using Machine Learning Regression Method

Song-Yeon Lee * and Yong Jeong Huh **†

*Mechatronics Engineering, Graduate School of Korea University of Technology and Education,
**†Department of Mechatronics Engineering, Korea University of Technology and Education

ABSTRACT

In this paper, We need to change all print factors when which print scaffold with 400 μm pore using FDM 3d printer. Therefore the print quantity is 10 billion times, So we are difficult to print on workplace. To solve the problem, we used the prediction model based machine learning regression. We preprocessed and learned the securing print condition data, and we produced different kinds of prediction models. We predicted the pore size of scaffolds not securing with new print condition data using prediction models. We have derived the print conditions that satisfy the pore size of 400 μm among the predicted print conditions of pore size. We printed the scaffolds 5 times on the condition. We measured the pore size of the printed scaffold and compared the average pore size with the predicted pore size. We confirmed that error was less than 1%, and we were identify the model with the highest pore size prediction performance of scaffold.

Key Words : 3D Printer, Machine Learning, Precision, Regression, Scaffold

1. 서 론

반도체 분야에서는 체내에 삽입되는 생체센서 제작에 관한 연구가 진행되고 있다. 그 중에서 FDM방식의 3D프린터를 이용하여 생체센서를 구성하는 인공지능기공 제작이 활발히 진행되고 있다[1]. FDM 방식의 3D프린터는 노즐과 베드가 이동하면서 한 줄씩 적층하는 구조이기 때문에 인공지능기공을 제작하기에 적합하다.

생체 센서를 제작하기 위하여 인공지능기공은 생체 적합성 소재로 구성 되어야 한다. 그리고 체액의 원활한 투과

를 위하여 일정한 기공크기를 가져야 한다. 여러 크기의 기공 중에서 골세포 배양을 위한 인공지능기공은 400 μm 의 기공크기를 가져야 한다[2].

3D프린터는 여러 가지 출력 요인이 이동, 변화하여 출력물을 제작한다. 3D프린터의 출력 요인은 노즐 온도, 베드온도, 채움 밀도, 출력 속도, 토출량 등 7개의 3D프린터 내부 요인과 기온, 기압 등 3개의 외부 요인으로 구성되어 있다[3].

인공지능기공의 기공크기를 400 μm 로 제작하기 위하여 3D프린터의 10가지 출력 요인을 10개의 수준으로 나누고, 모두 변경하여 출력한다면 출력을 100억 회 실행해야 한다. 이러한 출력 횟수는 현장에서 수행하기 어렵기 때

†E-mail: yjhuh@koreatech.ac.kr

문에 400 μm 의 기공크기를 갖는 인공지지체의 출력 조건 도출이 어렵다. 이러한 문제점을 해결하기 위하여 데이터를 기반으로 하는 머신 러닝에 적용하여 해결방안을 도출하였다. 머신 러닝은 기존의 실험, 연구를 통하여 수집된 학습 데이터로 다양한 모델들을 제작한다[4]. 다양한 모델들 중에서 인공지지체의 기공 크기를 예측할 수 있는 모델을 제작한다. 예측모델은 새로운 데이터를 이용하여 기공크기를 예측할 수 있기 때문에 예측모델을 이용하여 확보되지 않은 조건의 기공크기를 예측이 가능하다. 따라서 목표 기공크기를 맞추기 위하여 여러 출력요인을 변경 한 뒤, 인공지지체를 출력하는 번거로움이 없다는 장점을 갖고 있다. 기존의 연구에서는 다중 선형회귀 방안을 이용한 1개의 예측모델을 제작하여 인공지지체의 기공 크기를 예측하였다. 하지만 머신 러닝에는 여러 종류의 모델들이 있다. 다중 선형회귀(Multi linear regression), 라소회귀(Lasso regression), 릿지회귀(Ridge regression) 방안이 있다. 이 3개의 회귀방안을 이용하여 머신 러닝 회귀 기반의 예측모델을 제작한 뒤, 인공지지체의 기공크기를 예측하는 성능이 가장 우수한 모델을 선정한다.

본 논문에서는 FDM방식의 3D프린터를 이용하여 기공 크기가 400 μm 인 인공지지체를 제작할 때, 출력 요인들을 일일이 변경해야 한다. 따라서 출력 횟수가 10억 회가 되어 현장에서 수행하기 어려운 문제를 갖는다. 이러한 문제에 대한 해결방안을 도출하기 위하여 머신 러닝의 회귀 기반 예측모델을 이용하였다[5].

기존에 확보된 출력조건 데이터를 전처리하여 학습시킨 뒤, 다중 선형회귀, 라소회귀, 릿지회귀를 사용한 3개의 예측모델들을 제작하였다. 제작된 예측모델들을 이용하여 새로운 출력조건 데이터로 확보되지 않은 인공지지체의 기공크기를 예측하였다. 예측된 여러 기공크기의 출력조건 중에서 400 μm 의 기공 크기를 만족하는 출력 조건을 도출하였다. 해당 조건으로 인공지지체를 5회 출력하였다. 출력된 인공지지체의 기공 크기를 측정하여 평균 기공 크기와 예측한 인공지지체의 기공 크기를 비교하였다. 오차가 1%미만임을 확인하고 인공지지체의 기공크기 예측 성능이 가장 높은 모델을 확인할 수 있었다.

2. 연구 목적 및 방법

2.1 연구 목적

인공지지체의 기공 크기는 세포의 종류에 따라 매우 다양하다. 세포배양 연구에서 400 μm 의 기공 크기를 갖는 인공지지체가 가장 많이 사용된다.

인공지지체의 10개 출력 요인을 10수준으로 나눈다면 경우의 수가 10억회가 되기 때문에 출력해야 할 횟수가

너무 많다. 따라서 현장에서 400 μm 의 기공 크기를 갖는 인공지지체의 출력 조건을 도출하기 어렵다는 문제가 있다. 이러한 문제의 해결방안을 도출하기 위하여 머신 러닝 회귀 기반의 다중 선형회귀, 라소회귀, 릿지회귀를 이용한 3개의 예측모델들을 제작한다. 제작된 기공 크기를 예측 가능한 모델들로 확보되지 않은 출력조건으로 인공 지지체의 기공 크기를 예측한다. 출력 조건 중에서 400 μm 의 기공 크기를 갖는 인공지지체 출력 조건을 도출하여 성능이 가장 우수한 모델을 선정하는 것이 본 연구의 목적이다.

2.2 출력 장비

일반적인 FDM 방식의 3D프린터는 챔버(Chamber)가 없어서 공기의 온도, 습도 등의 3D프린터 외부요인을 제어하기가 어렵다. 따라서 챔버가 있어 3D프린터 외부 요인도 제어할 수 있는 3D프린터를 선정하여 연구를 진행하였다.

본 연구에서는 큐비콘에서 제작한 스타일이라는 FDM 방식의 3D 프린터를 사용하였다.

Fig 1은 인공지지체 기공 크기의 데이터 수집에 사용된 3D프린터를 도시하였다.



Fig. 1. FDM 3d printer using data acquisition.

2.3 인공지지체 모델

인공지지체의 모양은 세포배양 연구에 많이 사용되는 웰 플레이트(Well plate) 중에서 24구로 되어있는 규격에 맞

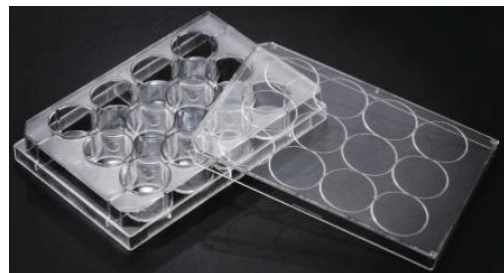


Fig. 2. Well plate for scaffold standard.

추었다. 24 웰 플레이트에 맞는 인공지지체는 지름 20mm, 두께 2mm규격의 원판모양이다. Fig. 2는 인공지지체의 규격에 참고된 웰 플레이트를 도시하였다.

2.4 출력조건과 기공크기 측정

출력 조건은 인공지지체를 출력하기 위하여 Cubicreator 4라는 프로그램의 설정 요인을 이용하였다. Cubicreator 4는 노즐온도, 베드온도, 토출량 등 7개의 3D프린터 내부요인과 3개의 외부요인을 조절할 수 있다. 각각의 설정 값을 인공지지체의 출력조건으로 설정하였다.

출력된 인공지지체의 기공크기는 1,000배율의 USB 현미경을 이용하여 측정하였다. Fig. 3은 기공 크기 측정을 위해 사용한 USB 현미경을 도시하였다.



Fig. 3. USB microscope for measuring pore size.

3. 데이터 수집과 전처리

3.1 데이터 수집

지름 20 mm, 두께 2 mm 인 원판모양의 인공지지체를 제작할 때, 3D프린터의 각 출력 요인에 대한 조건을 사용하였다. 3D프린터의 출력 요인 은 노즐 온도, 베드온도, 채움 밀도, 출력 속도, 토출량 등 7개의 내부 요인과 기온, 습도 등 3개의 외부 요인으로 구성되어 있다. 10개의 출력 요인은 3D프린터 프로그램인 Cubicreator 4로 모두 조절 가능하다.

3.2 데이터 전처리

예측모델의 Training dataset을 구성하기 위하여 데이터를 전처리 하였다. Full dataset은 수집된 모든 데이터를 이용하여 예측모델을 구축하는데 사용되었다. 그리고 Selected dataset은 예측 성능향상을 위하여 데이터를 축소하여 구성한다[6].

Fig. 4는 인공지지체 제작에 사용되는 요인들의 특징과 상관관계를 분석하여 Dataset을 형성하는 과정을 도시하였다.

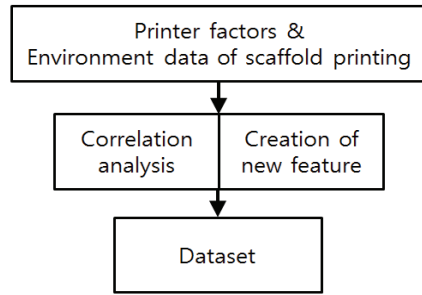


Fig. 4. Correlation analysis for dataset creating.

인공지지체의 기공크기 예측모델을 제작하기 위한 3D 프린터의 모든 출력 요인의 데이터인 Full dataset을 Table 1에 도시하였다.

Table 1. Full dataset used scaffold printing

No	Attribute of data(unit)	Contents
1	Nozzle temperature(°C)	Temperature of printer nozzle
2	Bed temperature(°C)	Surface temperature of printer bed
3	Bed roughness(α)	Surface roughness of printer bed
4	Output speed(mm/s)	Move speed of printer nozzle
5	Move precision(μm)	Minimum move distance of nozzle
6	Fill density(%)	Lamination density of printing layer
7	Discharge rate(%)	Discharge rate of filament
8	Air temperature(°C)	Air temperature in printing place
9	Air pressure(atm)	Air pressure in printing place
10	Air humidity(%)	Air humidity in printing place

데이터의 축소를 위하여 상관관계 분석은 필터 방식을 선택하였다. Filter 방식은 제작된 최종 모델을 이용하여 실제로 인공지지체를 제작할 경우 발생하는 다중공선성 문제를 해결하기 위하여 사용되는 방식이기 때문이다.

필터 방식은 Pearson correlation coefficient 방법을 사용하였다. 측정되는 데이터 중 중복되는 데이터는 제외하였다.

그리고 상관계수의 값을 분석하여 강한 음의 상관관계와 강한 양의 상관관계에 있는 변수들을 선택하였다. 그 다음 조건을 만족하지 못하는 변수는 제외하였다. 10개의 변수 중 3개의 변수가 선택되었다. 상관관계분석을 통하여 이동속도, 채움밀도, 토출량으로 축소된 3개의 데이터를 Table 2에 도시하였다.

Table 2. Selected dataset with correlation analysis

No	Attribute of data(unit)	Contents
1	Output speed(mm/s)	Move speed of printer nozzle
2	Fill density(%)	Lamination density of printing layer
3	Discharge rate(%)	Discharge rate of filament

3.3 Training dataset 구성

Training dataset 은 입력 변수를 모두 반영한 Full dataset과 분석을 통한 데이터 축소를 수행하여 Selected dataset으로 구성되었다. Full dataset은 인공지능기공체를 제작할 때 사용되는 3D프린터의 모든 출력 요인인 10개의 요인으로 구성되었다. 구성된 출력 요인은3D프린터의 각 요인에 대한 속성들로 나누어 총 10개의 요인으로 구성되었다. 사용된 데이터의 수량은 각 요인 별로 545 Sample을 이용하였다. Select dataset의 각 요인 별로 사용된 샘플의 수량을 Table 3에 도시하였다.

Table 3. Selected dataset used scaffold printing

No	Attribute of data(unit)	Sample quantity
1	Output speed(mm/s)	545 ea
2	Fill density(%)	545 ea
3	Discharge rate(%)	545 ea

4. 머신러닝기반 예측모델 제작

4.1 머신 러닝

머신 러닝은 기존에 확보된 데이터들을 이용하여 학습한 뒤 분류, 회귀 등의 과정을 수행하는 것을 의미한다[6].

학습 데이터의 종류에 따라 지도학습, 비지도학습, 강화학습으로 나뉜다. 본 연구에서는 지도학습에서의 회귀 방안을 사용하였다. 회귀방안은 데이터를 학습하여 회귀 모델을 제작한 뒤, 데이터를 예측한다.

Fig. 5는 훈련 데이터를 이용하여 예측모델을 제작한 뒤, 기공 크기를 예측하는 과정을 플로우차트(Flowchart)로 도시하였다.

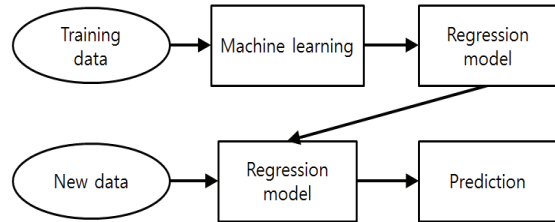


Fig. 5. Flowchart of data prediction used regression.

4.2 사용된 머신러닝 방안

예측모델을 제작하기 위하여 머신 러닝 회귀 방안을 사용하였다. 머신 러닝 회귀 방안은 다중 선형회귀, 라소회귀, 릿지회귀가 있다.

본 연구에서는 3가지의 회귀 방안 모두를 사용하여 인공지능기공 크기 예측모델을 제작하였다.

4.3 예측모델 제작

인공지능기공체의 기공크기와 3D프린터의 출력요인을 이용하여 예측모델을 제작하였다. 제작된 예측모델들은 비교분석을 위하여 다중 선형회귀, 라소회귀, 릿지회귀 모두 선형방정식을 이용하여 인공지능기공체의 기공크기를 예측하였다. 모든 모델의 학습횟수는 5,000 회였으며 학습 간격은 0.01, 가중치(Weight)는 10으로 설정하였다. 릿지회귀 기반 예측모델 제작을 위한 알파 값(Alpha value)은 0.5로 설정하였다. 입력 데이터는 이동속도, 채움밀도, 토출량으로 3 개이기 때문에 오버피팅을 방지하기 위한 추가적인 정규화(Regression)는 설정하지 않았다. 데이터 구성은 학습 70%, 테스트 30%의 비율로 진행하였다.

5. 예측모델 검증

5.1 예측모델 비교

인공지능기공 크기의 예측모델들을 다중 선형회귀, 라소회귀, 릿지회귀를 이용하여 제작하였다. 제작한 예측 모델들의 성능을 비교하였다. 그 결과 다중 선형회귀 방식을 이용한 예측모델의 성능이 99%로 가장 높았다. 라소회귀를 이용한 예측모델은 96% 였으며, 릿지회귀를 이용한 예측모델은 95%로 가장 낮았다. 예측모델들의 성능을 Table 4에 도시하였다.

Table 4. Accuracy of prediction models performance

No	Algorithm	True positive (%)	True negative (%)
		400 μm	400 μm
1	Multi linear regression	99	66
2	Lasso regression	96	69
3	Ridge regression	95	69

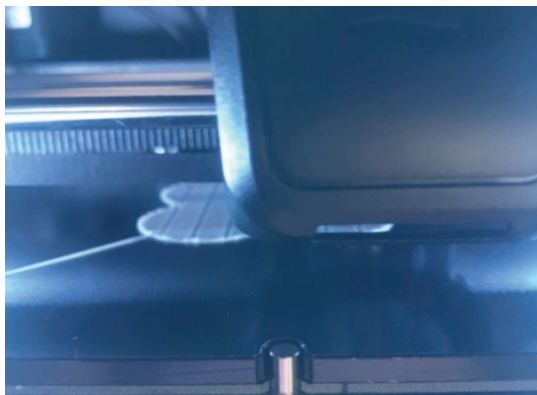
5.2 실험을 통한 예측모델 검증

모델들의 예측 성능을 검증하기 위하여 각 예측모델 별로 400 μm 의 기공크기에 맞는 출력 조건으로 인공지지체를 5개씩 출력하였다. 각각의 회귀방안에서의 400 μm 기공 크기 출력조건을 Table 5에 도시하였다.

Table 5. Accuracy of prediction models performance

No	Algorithm	Output speed(mm/s)	Fill density(%)	Discharge rate(%)
1	Multi linear regression	3	85	100
2	Lasso regression	4	85	97
3	Ridge regression	3	86	95

예측모델로 도출된 출력조건으로 인공지지체를 5회 출력한 뒤, 기공 크기를 측정하였다. 각 모델 별 인공지지체의 평균 기공크기를 도출하였다. 예측된 기공크기와 측정

**Fig. 6.** Scaffold printing using prediction data.

된 평균 기공 기공 크기의 오차를 비교하였다. 비교한 결과 각 모델들의 오차는 다중 선형회귀가 1%, 라소회귀가 3%, 릿지 회귀가 4% 인 것을 확인하였다. 따라서 다중 선형회귀 모델을 이용하여 인공지지체 기공 크기의 예측모델은 다중 선형회귀 방안을 이용한 예측모델이 가장 성능이 우수한 것으로 확인되었다.

예측모델로 도출된 출력 조건으로 3D프린터를 이용하여 인공지지체를 제작하는 모습을 Fig. 6에 도시하였다.

6. 결 론

본 논문에서 FDM 방식의 3D프린터로 400 μm 의 기공 크기를 갖는 인공지지체를 제작할 때 발생하는 문제에 대해 연구하여 다음과 같은 결론을 도출하였다.

1. 데이터 기반의 머신 러닝을 이용하여 해결방안을 도출할 수 있었다.
2. 머신 러닝의 다중 선형회귀, 라소회귀, 릿지회귀를 이용하여 3개의 예측모델들을 제작하고 성능을 수치화한 뒤, 예측모델의 성능을 정량적으로 비교하여 다중 선형회귀를 이용한 예측모델의 성능이 가장 우수한 것을 확인할 수 있었다.
3. 예측모델들을 이용하여 목표인 400 μm 의 기공크기에 적합한 각 모델의 출력조건을 도출할 수 있었다.
4. 예측모델로부터 도출된 출력 조건들을 이용하여 3D 프린터로 인공지지체를 5개씩 제작한 뒤, 예측된 기공크기와 측정된 평균 기공 기공 크기의 오차를 비교하였다. 비교한 결과 오차범위가 1% 미만인 것을 확인하여 예측모델의 실제 성능을 확인할 수 있었다.

참고문헌

1. Seung-Hyeok Choi, Min-Woo Sa and Jong-Young Kim, "New Fabricatio Method of Bio-Ceramic Scaffolds Based on Mold using a FDM 3D Printer", J. of The Korean Society of Precision Engineering, Vol.18, pp. 957-963, 2018.
2. Song-Yeon Lee and Yong-Jeong Huh, "A Study on Prediction Model of Scaffold Pore Size Using Machine Learning", J. of The Korean Society of Semiconductor & Display Technology, Vol.18, pp. 46-50, 2019.
3. Young-Woo Park and Sang-Won You, "Direction of Improvement of Reproducibility through Shape Distortion of Fused Deposition 3D Printing", J. of Basic Design & Art, Vol.19, pp.195-204, 2018.
4. Woong-Sup Lee, Jong-Yeol Ryu, Tae-Won Ban, Seong-

-
- Hwan Kim, Sang-Kee Kang, Young-Hwa Ham and Hyun-June Lee, "Estimation of Body Core Temperature of Cow using Neck Sensor based on Machine Learning", J. of Institute of Information and communication Engineering, Vol.22, pp.1611-1617, 2018.
5. Yeon-Ho Chu and Young-Kyu Choi, "A Deep Learning based IOT Device Recognition System" J. of The Korea Society of Semiconductor & Display Technology, Vol.18, pp.01-05, 2019.
6. Yong-Beom Park, Dong-Bin Choi and In-Soo Cho, "Taxation Analysis Using Machine Learning", J. of The Korean Society of Semiconductor & Display Technology, Vol.18, pp. 73-77, 2019.
-
- 접수일: 2020년 3월 5일, 심사일: 2020년 3월 11일,
게재확정일: 2020년 3월 18일