

## 기계학습을 활용한 소재은행



과목명	창업연계융합종합설계1	담당교수	정경훈
팀이름	전공	학번	이름
5AM	신소재공학부	20200956	김세민
	기계금속재료전공		
	신소재공학부	20161080	이석모
	전자화학재료전공		
	신소재공학부	20190935	김설아
	기계금속재료전공		
	신소재공학부	20191044	유지선
	전자화학재료전공		
	신소재공학부	20191058	최희정
	전자화학재료전공		

## 목차

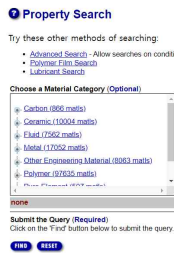
<b>1. 주제 선정 이유 .....</b>	<b>3</b>
(1) 주제 선정 배경 .....	3
(2) 선행 기술 조사 .....	4
<b>2. 목표 .....</b>	<b>4</b>
<b>3. 설계 .....</b>	<b>5</b>
(1) 데이터 범위 설정 .....	5
(2) 개발 과정 .....	6
1) 초기 데이터 수집 .....	6
2) 데이터 전처리 .....	7
3) 모델링 .....	9
4) 프로그램 구현 .....	11
<b>4. 결과 .....</b>	<b>14</b>
(1) 시스템 .....	14
(2) 합금 조성 예측 .....	14
(3) 3D 시각화 .....	15
(4) 소재은행의 독창성 .....	15
<b>5. 고찰 .....</b>	<b>16</b>
1) 기대효과 .....	16
2) 제품화 가능성 .....	16
3) 향후 계획 .....	16
<b>6. 부록 .....</b>	<b>17</b>

# 1. 주제 선정 이유

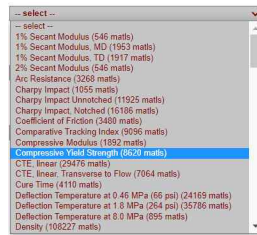
## (1) 주제 선정 배경

새로운 제품을 생성하기 위한 공학적인 설계, 제조 과정에서 적절한 소재를 찾는 것은 설계 아이디어의 구체화 및 확장에 매우 중요한 역할을 한다. 기존의 소재은행은 일반적으로 소재의 특성(밀도, 강도, 전기전도도 등)에 대한 데이터를 갖고있는 사이트다.

하지만 기존의 소재은행은 소재 검색에 있어 불편함이 있다. 아래는 대표적인 소재은행 matWeb b1)의 검색 과정이다. 카테고리 검색 항목을 열면 [그림 1]과 같은 화면이 나온다. [그림 1]에서 [그림 2]의 소재의 특성 3가지 선택하고, [그림 3]과 같이 특성의 범위와 단위를 선택하면, 해당 범위에 속하는 소재가 [그림 4]와 같이 나열된다.



[그림 1] 소재 검색창



[그림 2] 소재 특성 선택창



[그림 3] 범위 및 단위 설정

Select	Material Name	Compressive Modulus (GPa)	Density (g/cc)	Moisture Absorption at Equilibrium (%)
<input type="checkbox"/> 1	Overview of materials for Acetal Copolymer, Unreinforced	1.21 - 2.76	0.950 - 1.63	0.0200 - 0.800
<input type="checkbox"/> 2	Overview of materials for Acetal Homopolymer, Unreinforced	2.07 - 3.10	1.34 - 1.67	0.100 - 0.400
<input type="checkbox"/> 3	Overview of materials for Epoxy Molding Compound	1.03 - 35.0	0.550 - 2.45	0.230 - 0.600
<input type="checkbox"/> 4	Overview of materials for Epoxy Encapsulant, Unreinforced	0.0345 - 8.20	0.700 - 3.59	0.0700 - 2.60
<input type="checkbox"/> 5	Overview of materials for Epoxy Cure Resin	0.0552 - 168	0.860 - 2.60	2.80 - 9.40
<input type="checkbox"/> 6	Overview of materials for Polytetrafluoroethylene (PTFE), Glass Filled, Molded	0.430 - 2.14	1.40 - 3.10	0.0100 - 0.220
<input type="checkbox"/> 7	Overview of materials for Polyetheretherketone, Unreinforced	0.138 - 4.14	1.26 - 1.61	0.0700 - 0.500
<input type="checkbox"/> 8	Overview of materials for Polyetheretherketone, Glass Fiber Filled	3.45 - 55.0	1.30 - 1.93	0.0400 - 0.200
<input type="checkbox"/> 9	Overview of materials for Polyetheretherketone, Carbon Fiber Filled	2.28 - 152	1.32 - 1.90	0.0600 - 0.160
<input type="checkbox"/> 10	Overview of materials for Polyetheretherketone, PTFE Filled	2.28 - 5.45	1.32 - 1.63	0.0400 - 0.500
<input type="checkbox"/> 11	Overview of materials for Liquid Crystal Polymer (LCP), 30% Glass Fiber Filled	6.00 - 33.0	1.37 - 1.69	0.00200 - 0.0600
<input type="checkbox"/> 12	Overview of materials for Liquid Crystal Polymer (LCP), 40% Glass Fiber Filled	7.70 - 18.0	1.51 - 1.81	0.00200 - 0.0400
<input type="checkbox"/> 13	Overview of materials for Liquid Crystal Polymer (LCP), 50% Glass Fiber Filled	18.0 - 21.0	1.45 - 1.89	0.00600 - 1.10

[그림 4] 소재 검색 결과

이러한 검색과정은 비전문가에게 있어 원하는 소재 탐색이 어렵다.

첫 번째로, 검색이 어렵다. 전공 용어, 해당 특성의 범주에서 얼마만큼이 충족하고자 하는 범위인지, 그 단위는 무엇인지를 모두 알고 있어야 검색이 가능하기 때문이다.

두 번째로, 얻을 수 있는 데이터가 한정적이다. 등록된 소재만 검색이 가능하기 때문에, 원하는 소재의 검색이 불가능 할 수도 있고, 등록된 소재 안에서도 누락된 데이터들이 많다.

세 번째로, 정확한 데이터를 얻을 수 없다. 검색 결과로 범위형 데이터가 존재한다.

1) 데이터 출처 MatWeb <https://www.matweb.com/index.aspx>

## (2) 선행기술조사

새로운 소재은행을 설계하기 위해 기존의 대표적인 소재은행 사이트인 MatWeb, Total Materia의 특징과 한계를 분석하였다.

소재은행	특징	한계
MatWeb	<ul style="list-style-type: none"> <li>- 소재의 특성 뿐만 아니라 설명, 제조사 정보 등 추가적인 데이터 제공</li> <li>- 여러 소재를 비교하여 사용자의 요구에 맞는 소재를 찾을 수 있는 검색 및 분석 기능 제공</li> </ul>	<ul style="list-style-type: none"> <li>- 검색가능한 소재의 모든 특성이 기입된 것이 아닌 일부데이터만 존재하거나 아예 없는 경우 있음</li> <li>- 소재의 이름을 모르는 경우 검색 불가능</li> <li>- 정보 제공처가 다르기 때문에 동일한 공정을 거쳤더라도, 제조사에 따라 특성 결과가 다름</li> </ul>
Total Materia <sup>2)</sup>	<ul style="list-style-type: none"> <li>- 다양한 규격에 대한 비교 및 상호 변환 가능</li> <li>- 그래프와 같은 다양한 시각화 데이터 검색 가능</li> <li>- 소재 공급이 가능한 회사 정보 제공</li> <li>- 사용자가 검색한 소재와 비슷한 대체재 추천 기능</li> <li>- 한국어 지원</li> </ul>	<ul style="list-style-type: none"> <li>- 소재의 모든 특성이 등록되어 있지 않음</li> <li>- 연구되지 않은 소재는 검색 불가능</li> <li>- 조성만으로 소재의 특성 예측 불가능</li> <li>- 가공을 통한 특성이 반영되어 있지 않아 전문성 부족</li> </ul>

[표 1] MatWeb, Total Materia 특징 및 한계

## 2. 목표

기존 소재은행	기계학습을 활용한 소재은행
<ul style="list-style-type: none"> <li>- 범위형 결과값</li> <li>- 존재하지 않는 값</li> <li>- 한정된 소재 검색</li> <li>- 특성 범위를 지정하여 검색</li> </ul> <p>∴ 사용자가 원하는 소재를 찾는데 한계</p>	<ul style="list-style-type: none"> <li>- 특정한 값 검색</li> <li>- 연속적인 값               <ul style="list-style-type: none"> <li>- 결측치 없음</li> </ul> </li> <li>- 등록되지 않은 소재의 특성 제공</li> </ul> <p>∴ 더 효율적인 검색 + 더 다양한 데이터 제공</p>

[표 2] 기존 소재은행과 기계학습을 활용한 소재은행

이에 따라 우리는 기존 소재은행의 불편함을 개선한 새로운 소재은행을 제작하고자 하였다. 목표는 크게 3가지이다.

첫 번째로, 기존의 복잡하고 어려운 소재은행의 검색 방식을 개선한다. 이를 통해 합리적이고 효율적인 검색 기능을 제공할 수 있도록 한다.

두 번째로, 기계학습을 활용해 모델에 상관관계를 학습시켜 연속적인 데이터를 갖는 프로그램을 제작한다. 소재의 조성과 특성이 없는 데이터를 보완하고, 새로운 조성 비율 소재의 특성을 예측한다. 이를 통해 더 다양하고 연속적인 데이터를 제공할 수 있도록 한다.

세 번째로, 소재의 합금 조성에 따른 강도 특성의 상관관계를 3차원 시각화 데이터로 제공한다. 이를 통해 사용자가 조성과 인장강도(UTS) 사이의 관계를 쉽게 파악할 수 있도록 한다.

2) Total Materia <https://www.totalmateria.com/page.aspx?ID=Home&LN=>

### 3. 설계

#### (1) 데이터 범위 설정

소재은행을 제작하는 것에 있어서 모든 소재에 대한 데이터를 다룰 수 없었기 때문에, 데이터 범위를 알루미늄계 합금 소재로 한정하였다. 알루미늄계 합금은 용점이 낮고, 기체 용해도가 낮으며, 유동성이 크고, 표면처리가 우수하다는 장점이 있어, 상업적으로 많이 이용된다. 그만큼 다양하고 많은 양의 데이터를 확보할 수 있기에 소재은행의 메이저 소재를 알루미늄계 합금으로 설정하였다. 그 중에서도 5XXX계와 7XXX계 합금이 다른 시리즈랑 비교했을 때 데이터 품질이 좋았고, 데이터 사이즈가 적절하여 초기데이터로 선택하였다.

소재특성으로는 인장강도(UTS)를 선택하였다. 인장강도(UTS)는 합금의 성능을 잘 대표할 수 있는 특성이다. 합금 조성에 따른 상관관계에 있어 유의미한 분석 결과를 얻을 수 있을 것이라고 판단하였다.

합금을 만들 때, 전체 함량에서 차지하는 비율이 높은 순서대로 원소 7개를 범위로 설정하였다. [표 3]은 알루미늄계 금속의 대표적인 합금 조성 물질의 강화 효과에 대한 내용을 간략하게 기술한 것이다.

합금 원소	강화효과	합금 원소	강화효과
Si	Si는 Al보다 작아서 합금 내부의 격자 불규칙성과 결정구조 변화를 유발한다. 이는 인장강도(UTS)를 증가시키고 전위 이동을 어렵게 만든다. 따라서 Si 첨가는 합금 강도 향상에 기여한다.	Mg	합금 형성 과정에서 Mg <sub>2</sub> Al <sub>3</sub> 등의 화합물이 형성되어 격자 내부에서 강한 결합을 형성하여 인장강도(UTS)가 증가한다. Mg는 Al보다 크기가 크기 때문에 격자 불규칙성을 유발하여 인장강도(UTS)를 향상시킨다. Mg가 Al에 첨가되면 Mg <sub>2</sub> Al <sub>3</sub> 를 형성하고 전위의 이동을 방해하여 인장강도(UTS)가 증가한다.
Fe	합금 형성 과정에서 FeAl <sub>3</sub> , Fe <sub>3</sub> Al 등의 화합물이 형성되고 강한 결합을 이루어 인장강도(UTS)가 증가한다. 이는 Al과 Fe의 격자 상수 차이로 인해 합금 내부에 결합이 형성되고, 결정 격자의 규칙성을 유지하여 인장강도(UTS)를 증가시킨다. 또한, Fe가 Al에 고용되면 전위 이동이 어려워져 인장강도(UTS)가 증가한다. 따라서 Fe 첨가는 합금 구조를 강화하여 인장강도(UTS)를 향상시킨다.	Ni	합금 형성 과정에서 Al과 Ni가 결합하여 강한 구조를 형성하고, 이로 인해 인장강도(UTS)가 증가한다. Ni는 Al과 크기와 전하수가 다르기 때문에 합금 내부에 결합이 형성되어 인장강도(UTS)를 향상시킨다. Ni가 Al 격자 구조 내부에 들어가면서 전위 이동을 방해하여 인장강도(UTS)를 증가시킨다. 또한 Ni와 Al의 결합은 안정적이며, 구조의 변형 에너지를 높여 전위 이동을 어렵게하여 인장강도(UTS)를 향상시킨다.
Cu	Al <sub>2</sub> Cu, AlCu 등의 화합물이 형성되어 강한 결합을 이루어 인장강도(UTS)가 증가한다. Al과 Cu는 FCC 구조를 가지며, 합금 형성으로 새로운 격자 구조가 형성되어 인장강도(UTS)가 증가한다. 이는 Al과 Cu의 격자 상수 차이로 인해 선결합이 형성되고, 결정 격자의 규칙성을 유지하여 인장강도(UTS)를 향상시킨다.	Zn	합금 형성 과정에서 Al과 Zn이 결합하여 강한 구조를 형성하고, 이로 인해 인장강도(UTS)가 증가한다. Al과 Zn은 크기와 전하가 다르기 때문에 합금 내부에 결합이 형성되어 인장강도(UTS)를 향상시킨다. Zn의 크기가 크기 때문에 격자결합이 형성되고, 이는 확산을 억제하여 인장강도(UTS)를 증가시킨다. 또한 Zn은 Al 구조 내에서 불안정한 위치에 있어 전위 이동을 방해하고 상분리 경향성으로 전위와 상호작용하여 이동을 어렵게 한다.
Mn	합금 형성 과정에서 MnAl <sub>6</sub> , MnAl <sub>12</sub> 등의 화합물이 형성되어 강한 결합을 이루어 인장강도(UTS)가 증가한다. Mn은 Al과 크기 차이로 인한 격자 불규칙성을 유발하여 격자결합이 형성되어 인장강도(UTS)를 증가시킨다. Mn과 Al의 상호작용은 전위의 이동을 방해하여 인장강도(UTS)를 높인다. 따라서 Mn 첨가는 합금 구조를 강화하여 인장강도(UTS)를 향상시킬 수 있다.		

[표 3] 알루미늄 합금 원소의 강화효과

또한, 알루미늄계 합금은 여러가지 합금 원소를 다양하게 혼합하여 만드는 것이 특징이다. 이렇게 여러가지 합금 원소를 첨가하는 경우, 매트릭스 물질과 합금 원소간의 상호작용 뿐만 아니라, 다른 첨가 원소와의 상호작용으로 더 복합적인 강화 효과가 일어난다.

우리는 이러한 합금조성의 영향의 복합적인 강화 관계를 기계학습을 통해 파악하고, 이론적으로 나타나 있는 비선형관계를 연속적인 모델링을 통해 복합적 상관관계를 시각적, 수치적으로 나타내고자 하였다.

## (2) 개발 과정

데이터 전처리	모델링	프로그램 구현
<ul style="list-style-type: none"> <li>-결측치 처리: 인장강도(UTS), 조성 값이 Null인 경우 0으로 처리</li> <li>-이상치 처리: 가공 공정 처리된 합금 데이터 삭제</li> <li>-범위형 변수 처리: 평균값 처리</li> <li>-결측치 예측 코드 작성: 기존 data set의 Null 값을 예측하는 프로그램 작성</li> <li>-이를 기반으로 하여 새로운 Excel data 생성 (fixed data)</li> </ul>	<ul style="list-style-type: none"> <li>-모델 선정: 준지도 학습으로 모델 학습, 랜덤 포레스트 모델 선정</li> <li>-합금 조성 예측: 랜덤 포레스트 모델을 사용해 임의의 인장강도(UTS)값에 따른 합금 조성 예측 코드 작성</li> <li>-3D 시각화: 임의의 조성 2개 선택 시 인장강도(UTS)값을 시각화하는 그래프를 출력하는 코드 작성</li> </ul>	<ul style="list-style-type: none"> <li>-GUI 생성: 응용 프로그램으로 만들기 위한 코드 작성</li> <li>-합금 조성 예측 알고리즘과 3D 시각화 알고리즘 작성</li> <li>-사용자 인터페이스(UI) 디자인</li> <li>-응용 프로그램화</li> </ul>

[표 2] 설계 과정

### 1) 초기 데이터 수집

MatWeb 사이트를 이용하여 raw data를 수집하였다. [그림 5]와 같이 조성, 밀도, 인장강도(UTS)로 데이터를 기술하였다. 조성에 결측치가 포함되어있거나, 인장강도(UTS)가 나타나 있지 않은 데이터가 존재했다. 또한 조성, 인장강도(UTS)가 범위형으로 주어져 있는 데이터도 포함되어 있었으며, 특수 처리된 데이터가 포함되어 합금 조성 이외의 영향이 포함된 데이터가 존재했다.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Overview of materials for 3000 Series Aluminum Alloy	밀도	인장강도	Si	Fe	Cu	Mn	Mg	Cr	Ni	Zn	Ti	Sc	Zr
2	1 Aluminum 5054A Composition Spec	0.0160-2.8	110-150M	0.01-1.4	0.008-1.2	0.03-0.8	0.1-0.7	0.2-6.2	0.01-0.3	0.01-0.1	0.01-2.8	0.008-0.2	0.05-0.330	0.05-0.13
3	2 Aluminum 5054A Composition Spec	1.428	110-150M	0.705	0.004	0.415	0.066	1.2	0.255	0.055	1.405	0.104	0.3	0.175
4	3 Aluminum 5005 H12	2.69		0.3	0.45	0.05	0.15	0.7-1.1	0.1		0.2			
5	4 Aluminum 5005 H14	2.7	138	0.3	0.7	0.2	0.2	0.5-1.1	0.1		0.25			
6	10 Aluminum 5005 H38	2.7	179	0.3	0.7	0.2	0.2	0.5-1.3	0.1		0.25			
7	11 Aluminum 5005 O	2.7	200	0.3	0.7	0.2	0.2	0.5-1.4	0.1		0.25			
8	12 5006 Aluminum Composition Spec	2.7	124	0.3	0.7	0.2	0.2	0.5-1.5	0.1		0.25			
9	13 5010 Aluminum Composition Spec	2.71		0.4	0.8	0.1	0.4-0.8	0.8-1.3	0.1		0.25	0.1		
10	14 5016 Aluminum Composition Spec	2.71		0.4	0.7	0.25	0.1-0.3	0.2-0.6	0.15		0.3	0.1		
11	15 5017 Aluminum Composition Spec	2.7		0.25	0.6	0.2	0.1-0.7	1.4-1.9	0.1		0.15	0.05		
12	16 5018 Aluminum Composition Spec	2.69		0.4	0.7	0.18-0.28	0.6-0.8	1.9-2.2				0.09		
13	17 Aluminum 5018A Composition Spec	2.67		0.25	0.4	0.15	0.1	2.6-3.6	0.3		0.2	0.15		
14	18 5019 Aluminum Composition Spec	2.67		0.4	0.4	0.1	0.25	3.0-3.6	0.3		0.2	0.15		
15	19 Aluminum 5019A Composition Spec	2.65		0.4	0.5	0.1	0.2	4.5-5.6	0.3		0.2	0.2		
16	20 5021 Aluminum Composition Spec	2.65		0.2	0.35	0.15	0.2-0.5	4.4-5.4	0.1		0.25	0.1		
17	21 5022 Aluminum Composition Spec	2.66		0.4	0.5	0.15	0.1-0.5	2.2-2.8	0.15		0.15			
18	22 5023 Aluminum Composition Spec	2.66		0.25	0.4	0.2-0.5		0.2	3.5-4.9	0.1	0.25	0.1		
19	23 Aluminum 5024 H116 Al-Scandium Alloy	2.64		0.25	0.4	0.2-0.5		0.2	5.0-6.2	0.1	0.25	0.1		
20	24 5025 Aluminum Composition Spec	2.65	305	0.25	0.4	0.2		0.2	3.0-5.1	0.1	0.35	0.2	0.1-0.4	0.05-0.2
21	25 Aluminum 5026 Composition Spec	2.64		0.25	0.25	0.1	0.2	4.5-6.0	0.2		0.25	0.05-0.2	0.05-0.35	0.1-0.25
22	26 Aluminum 5027 Composition Spec	2.69		0.35-1.4	0.2-1.0	0.1-0.8	0.6-1.8	3.9-4.9	0.3		1	0.2		0.3
23	27 5040 Aluminum Composition Spec	2.65		0.05-0.2	0.2-0.4	0.05-0.15	0.4-0.8	4.7-5.4	0.1		0.25	0.15		
24	28 Aluminum 5041 Composition Spec	2.72		0.3	0.7	0.25	0.9-1.4	1.0-1.5	0.1-0.3		0.25			
25	29 Aluminum 5042 H19	2.67		0.4	0.4	0.1	0.3-1.0	3.0-4.0	0.5		0.1	0.2		
26	30 5043 Aluminum Composition Spec	2.67	360	0.2	0.35	0.15	0.2-0.5	3.0-4.0	0.1		0.25	0.1		
27	31 5049 Aluminum Composition Spec	2.72		0.4	0.7	0.05-0.35	0.7-1.2	0.7-1.3	0.05		0.25	0.1		
28	32 Aluminum 5050A Composition Spec	2.7		0.4	0.5	0.1	0.5-1.1	1.6-2.5	0.3		0.2	0.1		
29	33 Aluminum 5050C Composition Spec	2.69		0.4	0.7	0.2	0.3	1.1-1.8	0.1		0.25			
30	34 Aluminum 5050 H32	2.69		0.25	0.6	0.5	0.2	1.2-1.8	0.1		0.5	0.1		
31	39 Aluminum 5051A Composition Spec	2.69	145	0.4	0.7	0.2	0.1	1.1-1.8	0.1		0.25			
32	40 5051 Aluminum Composition Spec	2.69		0.3	0.45	0.05	0.25	1.4-2.1	0.3		0.2	0.1		
33	41 Aluminum 5052 H19 Foil	2.69		0.4	0.7	0.25	0.2	1.7-2.2	0.1		0.25	0.1		

[그림 5] raw data



## 2) 데이터 전처리

높은 수준의 학습 데이터를 확보하고자, 데이터 전처리 과정을 거쳤다. 데이터 전처리 과정은 최종 모델을 만들기 위한 데이터 준비 단계로서 관련 데이터끼리 관계 설정 및 데이터 이해, 데이터 병합 과정을 포함한다.

### 2-1) 1차 전처리 과정

1차 전처리 과정을 통해서 raw data를 모델학습을 시킬 수 있는 데이터로 전처리하였다. 이상치, 범위형 변수, 결측치를 처리하였다.

#### ① 이상치 처리

187 Constellium Alumold® 350 Aluminum	2.84 g/cc	>= 320 MPa
188 Constellium Alumold® 400 Aluminum	2.79 g/cc	>= 390 MPa
189 Constellium Alumold® 500 Rolled Aluminum	2.82 g/cc	>= 470 MPa
190 Constellium Alumold® 500 Forged Aluminum	2.82 g/cc	>= 410 MPa

[그림 6] 특수처리된 알루미늄 인장강도(UTS) 데이터

[그림 6]에서 표시한 데이터는 특수처리된 알루미늄으로, 공정상 소성 가공으로 인장강도(UTS)가 매우 높아진 것이다. 특수 처리된 합금의 경우, 합금 조성 외의 외부 요인으로 인장강도(UTS)가 증가한 것이므로, 합금조성과 인장강도(UTS) 사이의 상관관계를 파악하는데 어려움을 주고, 모델 학습에 방해가 될 수 있다. 이에 따라 특수 처리된 데이터는 학습 데이터에서 제외하고 데이터 세트를 생성했다. 이는 합금 조성과 인장강도(UTS) 외의 요인을 모델 학습에서 제외하기 위함이다.

#### ② 범위형 변수

Si	Fe	Cu	Mn	Mg	Cr	Ni
0.3	0.7	0.2	0.2	0.5-1.5	0.1	
0.4	0.8	0.1	0.4-0.8	0.8-1.3	0.1	
0.4	0.7	0.25	0.1-0.3	0.2-0.6	0.15	
0.25	0.6	0.2	0.1-0.7	1.4-1.9	0.1	
0.4	0.7	0.18-0.28	0.6-0.8	1.9-2.2		

[그림 7] 범위형 변수 데이터

[그림 7]은 범위형 변수 데이터를 표시한 것이다. 범위형 변수는 정확한 값을 파악하기 어렵기 때문에, 특정한 값으로 입력해야 하므로 평균값으로 처리하였다.

#### ③ 결측치 처리

	밀도	인장강도
5006 Aluminum Composition Spec	2.7	124
5010 Aluminum Composition Spec	2.71	
5016 Aluminum Composition Spec	2.71	
5017 Aluminum Composition Spec	2.7	
5018 Aluminum Composition Spec	2.69	

[그림 8] 결측치 데이터

[그림 8]은 인장강도(UTS) 값이 비어있는 데이터를 표시한 것이다. 인장강도(UTS), 조성값이 비어있는 데이터는 0으로 처리하였다.

## 2-2) 2차 전처리 과정

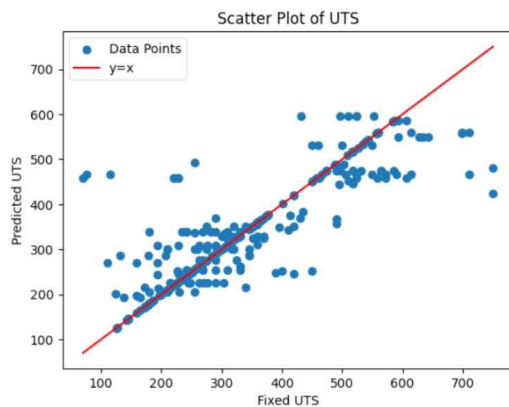
### ① 결측치 예측

1차 전처리를 한 data에서 더 정확한 데이터 제공을 위해 0으로 처리한 값 (null)을 예측하는 모델을 만들었다. 모델을 통해 비어 있는 데이터 세트에 예측한 값을 보완시켜 fixed data를 제작하였다.

### ② 오차율 측정

fixed data를 생성하기 위해 데이터 전처리 코드<sup>3)</sup>를 작성하였다. 이후 사용한 회귀 모델의 신뢰도를 평가하기 위해, fixed data의 인장강도(UTS)값을 모두 지우고, 모델 예측을 이용하여 predicted data를 만들어 두 데이터값을 비교해 오차를 측정했다.

그 결과, 우리가 예측에 사용한 모델이 약 85%의 정확도를 갖는다는 것을 확인할 수 있었다. [그림 9]는 fixed data와 predicted data의 인장강도(UTS)값을 비교한 것으로, y=x(빨간색 그래프)에 데이터(파란 점)가 가까울수록 모델이 정확하다는 것을 나타낸다. Data point가 y=x그래프에 몰려 있는 것을 확인할 수 있다.



$$\text{오차율} = \frac{\text{predict Data} - \text{fixed Data}}{\text{fixed Data}} * 100(\%)$$

[수식 1] 오차율 계산 수식

UTS	UTS2	오차율	절댓값		
315	328.902	4.413334	4.413334		오차율 평균
245	248.7161	1.516786	1.516786		15.0825
250	248.7161	-0.51355	0.51355		
165	193.0672	17.01041	17.01041		신뢰도
255	337.3927	32.31085	32.31085		약 84.9
215	225.5351	4.900032	4.900032		
215	225.5351	4.900032	4.900032		

[그림 9] predicted data, fixed data 비교 그래프 [그림 10] fixed data, predicted data 오차율 및 신뢰도

3) \*부록 별도 첨부 : 데이터 전처리 코드



### 3) 모델링

#### 3-1) 모델 선정

##### ① 모델 선정 기준

준지도 학습으로 모델 학습 진행하였다. 초기데이터는 인장강도(UTS)는 존재하나 조성이 나와있지 않은 데이터 혹은 조성은 나와있으나 인장강도(UTS)가 존재하지 않는 데이터가 존재했다. 준지도 학습은 레이블이 지정된 학습 데이터와 레이블이 지정되지 않은 비지도 학습 데이터를 함께 사용하여 모델을 학습하는 방법으로, 우리의 데이터를 분석하는데 가장 적합한 학습이라고 판단하였다.

##### ② 랜덤 포레스트 모델

랜덤 포레스트를 사용하여 합금 조성에 따른 인장강도(UTS)값을 예측하는 것은 [표 3]과 같이 비선형 관계 모델링, 변수 중요도 추정, 다중 상호작용 처리 등의 이점으로 인해 높은 예측 성능과 신뢰성을 제공할 수 있기에 채택하였다.

모델 특성	처리 방식의 장점
비선형 관계 처리	합금 조성과 인장강도(UTS) 값 사이의 관계는 일반적으로 비선형이다. 랜덤 포레스트는 다양한 결정 트리의 앙상블로 구성되어 비선형 관계를 모델링할 수 있다.
변수 중요도 추정	랜덤 포레스트는 각 조성 요소의 중요도를 추정할 수 있다. 중요도는 각 조성 요소가 인장강도(UTS) 값에 영향을 미치는 정도를 나타내며, 특정 조성 요소의 중요도가 높다면 해당 요소를 조정함으로써 인장강도(UTS) 값을 개선할 수 있다.
다중 상호작용 처리	합금 조성은 다양한 조합으로 이루어져 있으며, 조성 요소 간에 상호작용이 발생할 수 있다. 조성 요소 간의 복잡한 상호작용 패턴을 모델링할 수 있다. 이를 통해 조성 요소 간의 복합적인 영향을 포착하고 예측 정확도를 향상시킬 수 있다.

[표 3] 랜덤 포레스트 모델 특성 및 처리 방식의 장점

#### 3-2) 합금 조성 예측

랜덤 포레스트 모델을 사용해 임의의 인장강도(UTS) 값에 따른 합금 조성을 예측한다. 필요한 레이블과 데이터 세트 파일을 불러오고, 랜덤포레스트 모델을 학습시켜서 합금조성에 대한 예측을 수행한다. 이후에 [그림 11]과 같이 1열은 합금의 이름, 2열은 예측된 합금의 양을 나타낸 행렬 형식의 데이터를 출력한다.

```

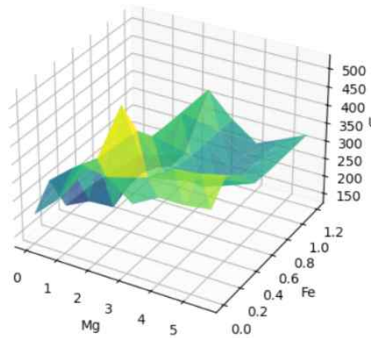
예측된 합금 조성:
Cu: 0.1308%
Fe: 0.1020%
Mg: 0.1000%
Mn: 0.1000%
Si: 0.1000%
Ti: 0.1050%
Zn: 0.1365%
  
```

[그림 11] 임의의 인장강도(UTS)값에 따른 합금 조성 예측 코드 실행 결과<sup>4)</sup>

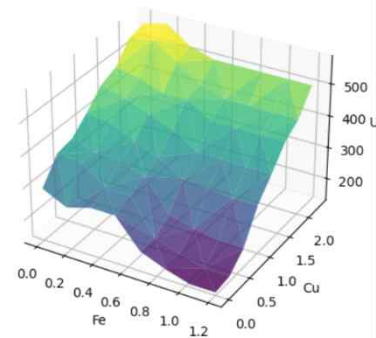
4) \*부록 별도 첨부 : 인장강도(UTS) 값에 따른 합금조성예측 모델링 코드

### 3-3) 3차원 시각화 그래프

조성을 2개 선택 하면, 그에 따른 인장강도(UTS) 값을 나타내는 그래프를 구현하였다. [그림 12], [그림 13]과 같이 축 인장강도(UTS) 설정 후, 조성 7개 중 2개를 선택해 차원 축소하여 모델을 시각화한다.



[그림 12] Mg, Fe, UTS 3차원 시각화 그래프



[그림 13] Fe, Cu, UTS 3차원 시각화 그래프

데이터를 불러와 학습에 사용할 특성 행렬  $x$ 와 목표 변수  $y$ 를 생성한다. 이 코드<sup>5)</sup>에서는 Mg, Fe 열을 선택하였다.  $x$ 는 선택한 2개 조성 열의 값들을,  $y$ 는 '인장강도(UTS)' 열의 값들을 갖는다. 이후, 랜덤포레스트회귀 모델을 생성하고, 생성한 데이터를 이용하여 모델을 학습시킨다. 선택한 조성 열의 최소값과 최대값을 이용하여 조합할 범위를 설정한다. 조합 가능한 모든 경우의 수를 생성하고, 조합된 값들을 저장한다. 모델을 사용하여 조합된 값들에 대한 '인장강도(UTS)' 값을 예측하여 3차원 곡면 그래프를 생성한다.

5) \*부록 별도 첨부 : 3차원 데이터 시각화 모델링 코드

#### 4) 프로그램 구현

파이썬 언어로 모델링한 합금 조성 예측 기능, 선택한 조성 간 관계를 나타낸 3D 시각화 기능을 사용자가 이용하기 쉽도록 하기 위해 응용 프로그램으로 만드는 과정이다.

##### 4-1) 응용 프로그램에 필요한 모듈 및 패키지 import

```
import tkinter as tk
from tkinter import *
from tkinter import messagebox
```

Tkinter는 파이썬에서 GUI(Graphical User Interface) 프로그래밍을 할 때 사용하는 모듈이다. GUI는 사용자가 그래픽을 통하여 컴퓨터와 상호작용할 수 있는 환경을 말한다. Tkinter는 Tcl/Tk에 대한 파이썬 Wrapper로서 Tcl/Tk를 파이썬에 사용할 수 있도록 한 Lightweight GUI 모듈이다. 본 프로젝트에서는 크로스 플랫폼에 사용되는 GUI 툴킷인 Tk를 import하여 사용한다.

##### 4-2) GUI 생성

```
# GUI 생성
root = Tk()
root.title("합금 조성 예측 및 3차원 시각화")
root.geometry("400x450")
root.option_add('*Font', '맑은고딕 12')
```

GUI 생성 시 사용한 root는 GUI 애플리케이션을 만들 때 사용되는 주요 객체로, Tkinter 애플리케이션의 메인 창이나 윈도우를 나타낸다. Tk() 함수를 호출하여 root 객체를 생성한 뒤, 윈도우의 제목, 크기, 폰트를 설정하는 과정이다.

##### 4-3) 합금 조성 예측을 위한 변수 정의

```
# 타이틀
title_lbl=Label(root, text='합금 조성 예측 및 3차원 시각화', font=('맑은고딕', 18, 'bold'))
title_lbl.grid(row=1, column=0, columnspan=3, pady=20)

def predict_composition():
    # 예측할 인장강도 입력
    input_strength = float(strength_entry1.get())

    # 합금 조성 예측
    predicted_composition = model.predict(data.iloc[:, 4:11])

    # 예측된 합금 조성 출력
    composition_names = data.columns[4:11] # 합금 조성 열의 이름들

    output = "예측된 합금 조성:\n"
    for composition_name, prediction in zip(composition_names, predicted_composition):
        output += f"{composition_name}: {prediction:.4f}%\n"

    messagebox.showinfo("결과", output)
```

Predict\_composition() 변수를 정의하고, 예측할 인장 강도를 입력하면 결과가 messagebox를 통해 출력되는 알고리즘이다.

```
def set_default():
    strength_entry1.delete(0, END)
```

Set\_default() 변수를 정의한다. 응용 프로그램 실행 후 예측할 인장강도(UTS) 값을 입력하는 화면에서 입력 값을 초기화하는 옵션을 삽입하기 위한 목적이다.

#### 4-4) 합금 조성 예측 기능 사용자 인터페이스(UI) 디자인

```
# UTS 입력 프레임
strength_lbl_frame=LabelFrame(root, text='조성 예측', font=('맑은고딕', 13, 'bold'))
strength_lbl_frame.grid(row=2, column=0, columnspan=3, padx=10)

# UTS 입력
strength_lbl1=Label(strength_lbl_frame, text='UTS 입력', width=10, font=('맑은고딕', 13))
strength_lbl1.grid(row=2, column=0, padx=10, pady=5)

strength_entry1=Entry(strength_lbl_frame, width=29, bg='powderblue')
strength_entry1.grid(row=2, column=1, sticky='w', padx=7, ipady=3)

# 버튼 생성
btn_frame1=Frame(strength_lbl_frame)
btn_frame1.grid(row=5, column=1, columnspan=3, padx=5)

btn_save1=Button(btn_frame1, text='조성 예측', width=10, font=('맑은고딕', 11), bg='lightblue', command=predict_composition)
btn_save1.grid(row=5, column=1, pady=20, padx=5)

btn_set_default=Button(btn_frame1, text='입력 초기화', width=10, font=('맑은고딕', 11), bg='lightblue', command=set_default)
btn_set_default.grid(row=5, column=2, pady=20, padx=5)
```

사용자가 합금 조성 예측 기능을 편리하게 사용할 수 있도록 인터페이스를 디자인하는 단계이다. 우선, 3D 시각화 기능과 시각적으로 분리하기 위해 프레임을 생성한 뒤 프레임 내에 'UTS 입력'이라는 이름을 가진 Label을 생성한다. 그 다음 인장강도(UTS) 값을 실제로 입력할 수 있도록 하는 위젯인 Entry를 정의한다. 마지막으로 3)에서 정의한 predict\_composition, set\_default를 command하기 위한 버튼을 생성한다.

#### 4-5) 3D 시각화를 위한 변수 정의

```
# 조합 선택 함수
def select_combination():
    selected_cols = []
    for var, val in vars_dict.items():
        if val.get() == 1:
            selected_cols.append(var)

    if len(selected_cols) < 2:
        messagebox.showerror("Error", "적어도 2개의 조성 열을 선택해야 합니다.")
        return

    visualize_3d(selected_cols)
```

2개 이상의 원소를 조합하여 상관 관계를 3D 시각화하기 위해 select\_combination() 변수를 정의한다. 사용자가 2개 미만의 원소를 선택하게 되면 시각화가 불가능하므로 messagebox를 통해 Error를 출력한다.

#### 4-6) 3D 시각화 기능 사용자 인터페이스(UI) 디자인

```
# 3차원 시각화 프레임
visualize_lbl_frame=LabelFrame(root, text='3차원 시각화', font=('맑은고딕', 13, 'bold'))
visualize_lbl_frame.grid(row=4, column=0, columnspan=3, padx=10, pady=10)

strength_lbl_frame.update_idletasks()

visualize_lbl_frame.grid_propagate(0)
visualize_lbl_frame.config(width=strength_lbl_frame.winfo_width(), height=strength_lbl_frame.winfo_height())

# 조성 열 선택 체크박스 생성
vars_dict = {}
checkbox_frame=Label(visualize_lbl_frame)
checkbox_frame.grid(row=2, column=1, columnspan=3, padx=10)

for i, col in enumerate(data.columns[4:11]):
    var = tk.IntVar()
    checkbox = tk.Checkbutton(checkbox_frame, text=col, variable=var)
    checkbox.grid(row=0, column=i)
    vars_dict[col] = var

# 버튼 생성
btn_frame2=Frame(visualize_lbl_frame)
btn_frame2.grid(row=10, column=3, pady=10, padx=5)

btn_save2=Button(btn_frame2, text='조합 선택', width=12, font=('맑은고딕', 11), bg='lightblue', command=select_combination)
btn_save2.grid(row=10, column=3, pady=10, padx=5)
```

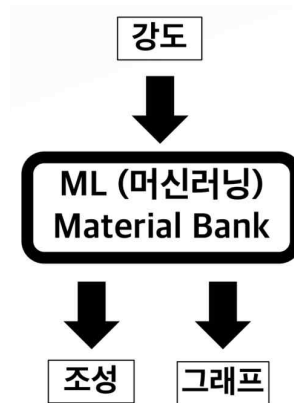
사용자가 3D 시각화 기능을 편리하게 사용할 수 있도록 인터페이스를 디자인하는 단계이다. 합금 조성 예측 기능과 시각적으로 분리하기 위해 프레임을 생성한다. Grid 함수 특성 상, 프레임이나 위젯을 원하는 위치에 배치하기 어려우므로 update\_idletask, grid\_propagate 함수를 사용하여 조성 예측 기능 프레임의 사이즈와 동일하게 만든다. 사용자가 원하는 조성을 선택할 수 있도록 tk 패키지를 사용해 체크박스를 생성한 뒤, 5)에서 정의한 select\_combination을 command하기 위한 버튼을 생성한다.

```
root.mainloop()
```

마지막으로 위에서 작성한 loop를 실행하기 위한 코드를 작성한다. Vscode 외부에서 프로그램을 사용할 수 있도록 window command 창에서 Pyinstaller를 설치한다. 'pyinstaller file\_name.py' 코드를 사용해 실행 파일로 변환하면 '내 PC' 내의 'dist' 폴더에서 저장된 실행 파일이 존재함을 확인할 수 있다.

## 4. 결과

### (1) 시스템

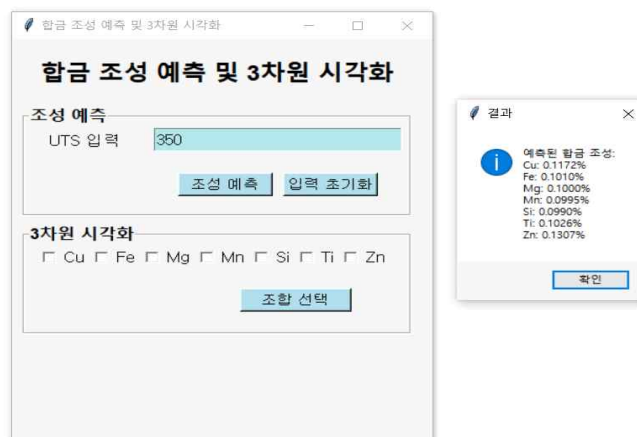


[그림 14] 기계학습을 활용한 소재은행 시스템 개략도

머신러닝 기반의 새로운 소재은행은 사용자가 특성 중 하나인 인장강도(UTS)를 입력하면 해당 특성에 맞는 적합한 소재의 조성을 추천하여 제시하는 기능을 가지고 있다. 초기 데이터인 raw data를 수집하고 합금 조성 이외의 결측치나 특수 처리 과정을 거친 데이터 등 결과에 영향을 주는 데이터를 처리하는 전처리 과정을 거쳐 높은 수준의 학습 데이터를 확보하였다. 또한 결측치가 없는 데이터 세트를 활용하여 모델을 학습시킨 후 예측된 값을 다시 채워 넣어 데이터세트를 보완하였다. 예측된 값의 신뢰도를 평가한 결과 약 85%의 높은 신뢰도가 도출되어 예측된 fixed data가 신뢰할 수 있다는 것을 확인하였다. 이후 준지도 학습을 활용하여 모델을 학습시킴으로써 보다 정교한 데이터 분석을 진행하였으며, 랜덤포레스트 모델을 활용하여 예측된 조성의 비율을 시각화하여 나타냄으로써 조성과 인장강도(UTS)의 관계를 직관적으로 파악할 수 있도록 하였다.

이를 통해 [그림 14]와 같이 사용자는 소재은행에서 추천받은 조성을 기반으로 특정 강도에 맞는 최적의 조성을 빠르게 찾을 수 있어 효율적으로 소재를 선택할 수 있다. 또한 기존 데이터에 없는 새로운 조성 비율을 가진 소재의 특성을 예측하고 추천해 주는 기능도 제공하여 사용자에게 보다 양질의 데이터를 제공한다.

### (2) 합금 조성 예측

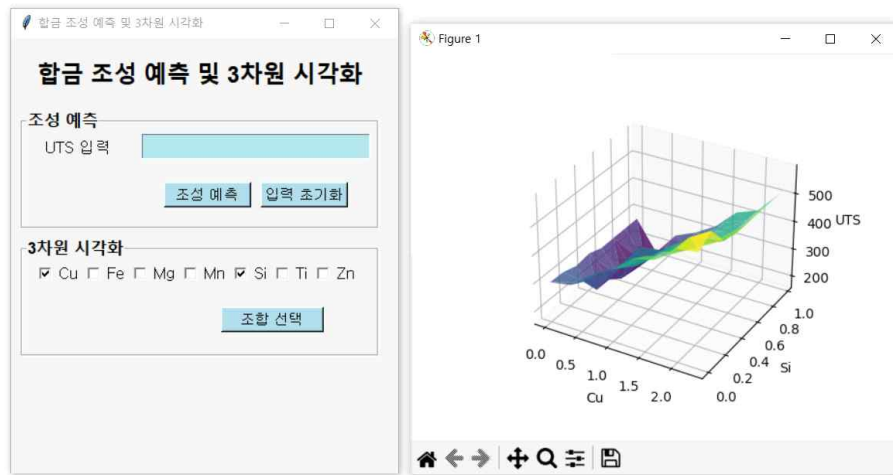


[그림 15] 합금 조성 예측 프로그램 실행 결과창

[그림 15]와 같이 사용자가 입력한 인장강도(UTS) 값에 따라 그에 맞는 7개의 조성 값을 예측하

여 보여주고, 그중 한 개의 조성 비율을 추천하여 제시한다. 위 사진에서 350이라는 인장강도(UTS) 값을 입력하였을 때 Cu, Fe, Mg, Mn, Si, Ti, Zn 총 7개의 조성을 예측한 결과를 확인할 수 있다. 위 결과를 통해 사용자는 특정 강도에 맞는 최적의 조성을 빠르게 찾을 수 있어 효율적으로 소재를 선택할 수 있다.

### (3) 3D 시각화



[그림 16] 3차원 시각화 프로그램 실행 결과창

랜덤 포레스트 모델을 사용하여 인장강도(UTS)와 조성을 3차원 그래프로 시각화하였다. X축과 Y축은 조성, Z축은 인장강도(UTS) 값으로, 조성 7개 중 2개를 선택해 차원을 축소하여 그래프를 도출하였다. [그림 16]에서 조성 Cu, Si 2개를 선택하여 그에 따른 시각화 그래프가 도출된 것을 확인할 수 있다. 위 그래프를 통해 사용자가 조성 and 인장강도(UTS) 사이의 관계를 쉽게 확인하고 파악할 수 있다.

### (4) 소재은행의 독창성

새로운 소재은행은 머신러닝을 기반으로 한 GUI 프로그램 설계되어, 기존의 복잡하고 어려운 소재은행의 검색 방식을 보완하고 개선하여 합리적이고 효율적인 검색 기능을 제공한다. 사용자가 직접 특성 값의 범위를 지정하여 검색할 필요 없이 강도만을 입력함으로써 간편한 검색이 가능하다. 또한 검색 결과값이 범위가 아닌 단일 값으로 제공되기 때문에 사용자가 보다 정확하고 효율적으로 소재를 선택하는 데 도움을 준다. 새로운 소재은행은 더 나아가 조성 and 인장강도(UTS)의 관계를 3차원 그래프로 시각화하여 제공함으로써 시각화된 그래프를 통해 사용자가 조성 and 인장강도(UTS) 사이의 관계를 쉽게 확인하고 파악할 수 있다. 이러한 3차원 그래프를 시각화하는 과정은 조성에 따른 연속적인 값을 모델이 학습함에 따라 이루어지기 때문에 기존에 존재하지 않은 새로운 조성 비율을 가진 소재도 추천해 줄 수 있어 보다 다양하고 신뢰성 있는 데이터를 제공한다. 결과적으로 새로운 소재은행은 복잡한 검색 과정을 단순화하고 사용자에게 직관적인 시각화 그래프를 제공함으로써 소재 선택의 효율성을 높일 뿐만 아니라 다양성도 향상시킨다. 또한 새로운 조성 비율을 가진 소재도 추천해줌으로써 사용자가 새로운 소재를 발견하고 창의적인 소재를 선택할 수 있는 기회를 제공한다.



## 5. 고찰

### (1) 기대효과

#### 1) 설계 및 제조 과정에서의 효율성 증대

새로운 소재은행에서 제공되는 강도에 따른 적합한 소재 추천 및 조성 정보를 통해 설계 및 제조 과정에서 보다 적합한 소재를 선택할 수 있으므로 제품의 품질 향상 및 생산성 향상을 기대할 수 있다.

#### 2) 시간 및 비용 절감 효과

새로운 소재은행은 기존 소재은행의 검색 방식에서 발생하는 불필요한 검색 시간 및 비용을 절감할 수 있으며, 적합한 소재의 선택으로 품질과 생산성이 향상됨에 따라 시간과 비용을 절감할 수 있다.

#### 3) 산업발전에 기여

새로운 소재은행을 통해 보다 쉽게 소재를 찾을 수 있게되어 사람들이 아이디어를 구현하기 용이해지고, 기존에 존재하지 않은 소재의 특성을 예측함으로써 새로운 소재를 제작하는 데 도움을 줄 수 있다. 이는 더욱 활발한 설계활동을 촉진시킴으로써 국가 과학기술 진흥산업의 발전에도 이바지할 수 있다.

#### 4) 소재에 대한 지식 습득

새로운 소재은행에서 제공되는 다양한 소재 정보와 추천 기능 및 시각화 그래프를 통해 사용자가 소재에 대한 전반적인 지식을 습득할 수 있어 일반인에게는 보다 쉽고 효율적으로 소재 선정을 가능하게 하며, 엔지니어 등의 관련 직무의 종사자에게는 보다 전문성있는 역할 수행에 도움을 줄 수 있다.

### (2) 제품화 가능성

소재를 판매하는 기업과 협업을 통해 실제로 웹사이트를 운영할 수 있다고 예상된다. 기업에게서 소재의 특성 data를 제공받아 데이터 전처리 과정을 더욱 효율적으로 운영할 수 있고, 더 많은 데이터를 사용자에게 제공이 가능하다.

### (3) 향후 계획

1) 현재 설계 단계에서는 합금의 특성을 나타내는 변수로 밀도와 인장강도(UTS)만 있지만, 더 많은 물성(전기전도도 등)을 추가하여 이를 바탕으로 조성을 예측할 수 있도록 한다.

2) 더 많은 합금 데이터를 지속적으로 업데이트하고, 인공지능망을 통해 전문적으로 학습시켜 용도에 적합한 소재를 추천해주는 기능을 추가하여 사용자들이 보다 정확하고 효율적으로 소재를 선택할 수 있도록 한다.

3) 현재 파이썬 환경 프로그램 형태로 제작된 소재은행이지만 사용자들의 접근성 향상을 위해 모바일 앱으로도 개발하여 휴대폰으로도 쉽게 소재를 검색할 수 있도록 한다.

궁극적으로 소재 관련 산업의 동향과 기술 발전에 대한 지속적인 관심과 연구를 통해 다양하고 혁신적인 소재 검색 및 분석 기능을 개발하여 산업 발전에 기여하는 소재 전문 플랫폼으로 성장한다. 이렇듯 추가 기능과 서비스로 사용자들의 편의성과 만족도를 높이고 산업발전에 기여하고자 한다.

## 6. 부록

### (1) 데이터 전처리

```
# 5000개 데이터 density, UTS 할형의 0값을 null로 바꾸기 <미리본 평균값 대체>

import pandas as pd
import numpy as np

# 데이터프레임 생성
df = pd.read_excel('/content/drive/MyDrive/Colab Notebooks/5000 전체 정리.xlsx')
# null 값을 0으로 대체
df.fillna(0, inplace=True)

# density, UTS 할형의 0값을 null로 바꾸기
df[['density', 'UTS']] = df[['density', 'UTS']].replace(0, np.nan)

# '-' 값을 평균값으로 대체하기
df['density'] = df['density'].astype(float)
df['UTS'] = df['UTS'].apply(lambda x: np.mean(list(map(float, str(x).split('-')))) if '-' in str(x) else float(x))

from google.colab import drive
import pandas as pd
```

### (2) 합금조성 예측 코드

```
#UTS에 따른 합금 조성 예측

import pandas as pd
from sklearn.ensemble import RandomForestRegressor

# 데이터 불러오기
data = pd.read_excel('/content/drive/MyDrive/Colab Notebooks/aluminum_predict.xlsx')

# 문자열 열 제거
data = data.select_dtypes(exclude='object')

# 학습 데이터 준비
X = data.iloc[:, 4:11] # 입력 특성 (합금 조성 데이터)
y = data.iloc[:, 8] # 타겟 변수 (인장강도)

# 모델 학습
model = RandomForestRegressor()
model.fit(X, y)

# 예측할 인장강도 입력
input_strength = [300]

# 합금 조성 예측
predicted_composition = model.predict(data.iloc[:, 4:11])

# 예측된 합금 조성 출력
composition_names = data.columns[4:11] # 합금 조성 열의 이름들

print("예측된 합금 조성:")
for composition_name, prediction in zip(composition_names, predicted_composition):
    print(f"{composition_name}: {prediction:.4f}x")
```

### (3) 3차원 데이터 시각화 모델링 코드

```
#3차원 데이터 시각화

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
from sklearn.ensemble import RandomForestRegressor

# 데이터 불러오기
data = pd.read_excel('/content/drive/MyDrive/Colab Notebooks/aluminum_predict.xlsx')

# 조성 선택
composition_cols = ['Mg', 'Fe'] # 선택할 조성 열 이름들을 리스트로 지정

# 모델 학습을 위한 특성 행렬 생성
X = data[composition_cols].values
y = data['UTS'].values

# RandomForestRegressor 모델 학습
model = RandomForestRegressor()
model.fit(X, y)

# 조성 범위 설정
composition_ranges = []
for col in composition_cols:
    col_range = np.linspace(min(data[col]), max(data[col]), 8)
    composition_ranges.append(col_range)

# 조합 생성
combinations = np.meshgrid(*composition_ranges)
combinations = np.vstack([combination.flatten() for combination in combinations]).T

# UTS 예측
predictions = model.predict(combinations)

## 3차원 곡면 그래프 생성
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')
ax.plot_trisurf(combinations[:, 0], combinations[:, 1], predictions, cmap='viridis', linewidth=0.2, alpha=0.8, label='Prediction')
ax.set_xlabel(composition_cols[0])
ax.set_ylabel(composition_cols[1])
ax.set_zlabel('UTS')
# ax.legend()

plt.show()
```