

심리적, 신체적, 혈액학적 임상특성을 균형적으로 활용한 기계학습 기반 사상체질 진단모델의 개발

황동욱¹ · 장동엽² · 김호석³ · 정경식³ · 채한^{1,*}

¹부산대학교 한의학전문대학원, ²동일대학교 한의과대학, ³한국한의학연구원 한의약데이터부

Abstract

Development of a Sasang Type Diagnosis Model Using Machine Learning and Balanced Integration of Psychological, Physical, and Hematological Clinical Features

Dongwook Hwang¹ · Dongyeop Jang² · Hoseok Kim³ · Kyoungsik Jeong³ · Han Chae^{1,*}

¹Pusan National University School of Korean Medicine, Yangsan, 50612 Korea

²College of Korean Medicine, Dong-Eui University, Busan, 47340 Korea

³Korean Medicine Data Division, Korea Institute of Oriental Medicine, 34054 Korea

Objectives

Efforts to develop a Sasang type diagnostic method with high validity and clinical utility have continued, with growing interest in large-scale data and machine learning (ML). However, ML often suffers from reliance on input data, overfitting, and limited theoretical interpretability, reducing its clinical applicability. This study aimed to develop a diagnostic model with improved interpretability, predictive performance, and accessibility by systematically selecting clinical features and using an intuitive ML interface.

Methods

Psychological, physical, and hematological features were selected stepwise through literature review, ANCOVA, and correlation analysis. Data from 2,407 participants were split into training (70%) and test(30%) sets. Feature contribution was evaluated by information gain (IG), and the diagnostic performance of random forest (RF), naïve Bayes (NB), neural network (NN), and stochastic gradient descent (SGD) was assessed using accuracy, precision, recall, F1 score, Matthews correlation coefficient (MCC), ROC curve and AUC. Analyses were performed using Orange, an open-source Python-based Graphic User Interface.

Results

IG analysis showed notable contributions from BMI (0.531), RMRw (0.292), RMR (0.207), three SPQ subscales (0.329). Including BMI yielded high accuracy (0.817), but after its removal to balance features, accuracy remained at 0.755. Algorithm efficiency depended on interaction with input data.

Conclusions

Balanced integration of psychological, physical, and hematological features improved consistency with Sasang typology theory, clinical utility, and model generalizability. Appropriate feature selection had a greater impact on performance than algorithm choice. The proposed procedure and Orange workflow offer clinicians and researchers a practical foundation for ML-based analysis, contributing to the globalization of Sasang typology.

Key Words : Sasang typology, Machine Learning, Orange, Body Mass Index, Sasang Personality Questionnaire, Resting Metabolic Rate

Received 18, September 2025 Revised 20, September 2025 Accepted 06, November 2025

*Corresponding author: Han Chae

School of Korean Medicine, Pusan National University. 49, Busandaehak-ro, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do, 50612, Korea
Tel/Fax +82-51-510-8470, Email han@chaelab.org

© The Society of Sasang Constitution and Immune Medicine. All rights reserved. This is an open access article distributed under the terms of the Creative Commons attribution Non-commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>)

I. Introduction

세계의학에 있어서 임상적으로 효과적이면서 신속한 치료를 위해 다양한 유형론 또는 체질 이론이 제시되어 왔다^{1,2}. 서양의 히포크라테스와 갈렌의 사체액설^{3,5}, 크레즈머⁶와 셸던⁷의 체격 이론과 동양의 오행체질론⁸, 아우르베다 도샤 체질⁹ 등 다양한 체질론이 제시되었다.

이제마의 사상의학^{1,3,10}은 높은 유효성과 낮은 부작용을 지닌 맞춤형 침치료¹¹와 한약 처방¹²이라는 측면에서 높은 임상적 가치를 지니고 있다. 선행 연구에서는 사상체질별 고유한 심리적^{4,13,14}, 신체적^{10,15}, 병리적^{13,16-18} 특징이 보고되었으며^{3,19}, 우울증²⁰, 대사 질환¹⁶, 고혈압⁸, 그리고 골다공증²¹과 같은 질환의 치료에 임상적 유용성도 보고되었다.

사상의학 연구에 있어서, 사상체질 유형별 고유한 임상특징을 찾아내어 측정 지표로 개발하는 연구와 함께 이들 지표들을 활용하여 검사 및 임상 타당도가 높은 진단법을 제안하려는 연구들이 지속되고 있다. 심리적 측면에서는 NEO성격검사^{13,22}와 기질및성격검사²³의 유용성이 반복적으로 확인되어 왔으며, 기존의 임상 특성을 통합한 한의학적 생리심리이론과 사상성격검사^{13,14,22}가 개발되었다. 사상성격검사는 제프리 그레이(Jeffrey Gray)와 로버트 클로닝저(C. Robert Cloninger)의 행동활성체계를 자극되거나 활성화되거나 흥분되어 있는 생물심리사회적 기능을 의미하는 양(陽) 특성으로, 행동억제체계를 억제되거나 둔화되고 비활성화 되어 있는 음(陰)으로 재해석하여 사상의학의 성정(性情) 이론을 음양 생리심리학 및 표준화된 지표로 제시하였다².

신체적 측면에서도 체질량지수가 사상체질별 고유한 특성으로 확인되어 왔으니^{10,15,22}, 체지방량과 비만도 그리고 심혈관질환나 당뇨의 질환 예측지표임에도 체질별 소증(素證)으로서 과도한 가중치가 부여되지는 비판도 동시에 보고되었다²⁴. 이에, 새로운 신체지표로서 체형의 특성을 보다 정확하게 측정하는 폰더랄 지수^{10,22}와 에너지 대사와 체성분, 신체활성도를

반영하는 기초대사량²⁵이 사상체질의 생병리적 특성의 해석과 임상진단에 보다 유용하다고 제시되었다².

최근 컴퓨팅 능력의 폭발적인 향상에 따라 기계학습(Machine Learning, ML)과 빅데이터를 활용하여 사상체질 유형을 진단하려는 연구가 대두되었다^{12,26-28}. 그러나 이러한 데이터 기반 분석(Data-driven analysis)은 손쉽게 높은 성능의 진단법을 도출할 수 있다는 장점에도 불구하고²⁸ 기존의 이론적 토대 위에서 임상 현장에 이들을 적용함에 있어서는 명확한 한계를 보인다²⁶. 더구나, ML 분석을 시행하기 위해서는 복잡한 데이터 처리와 전문적인 python 코딩 능력도 요구된다²⁹.

이에, 본 연구에서는 사상의학 연구 및 임상 현장에서의 한계를 극복하고자 하였다. 다양한 임상특성을 체계적으로 추출한 다음 비전문가도 사용할 수 있는 ML 소프트웨어로 분석하고, 도출된 사상체질 진단모델의 성능을 종합적으로 분석하였다(Figure 1). 이러한 접근은 본 연구에서의 KS15를 비롯한 사상체질 진단법에서 BMI 등의 과대추정이나 분석 과정에 투입되는 임상 데이터의 구조적 연관성 등을 동시에 검토할 토대를 제공한다.

진단모델을 위한 임상데이터에 있어서 이론적 검증이 선행된 기존연구와 대규모 임상 데이터의 통계적 분석을 동시에 활용하는 단계적 추출법을 사용하였으며, 입력된 심리적, 신체적, 혈액학적 임상 데이터가 체질진단의 효율에 미치는 기여도는 정보이득³⁰ 분석으로 검토하였다. 체질진단 알고리즘으로는 다중 클래스 데이터에 적합하다고 알려진 Random Forest (RF), Naïve Bayes (NB), Neural Network(NN), Stochastic Gradient Descent (SGD)을 사용하였으며, 이들의 성능 평가에는 분류정확도, 정밀도, 재현율, F1 점수, 매튜스 상관계수, ROC 커브와 AUC를 사용하여 다면적으로 분석하였다.

이와 함께, 비 전문가도 ML 분석용 python 코드를 손쉽게 사용할 수 있도록 그래픽 사용자 인터페이스로 구현한 오픈소스 기계학습 소프트웨어인 Orange를 사용하여, 연구자 뿐만 아니라 임상가와 학부생도 본

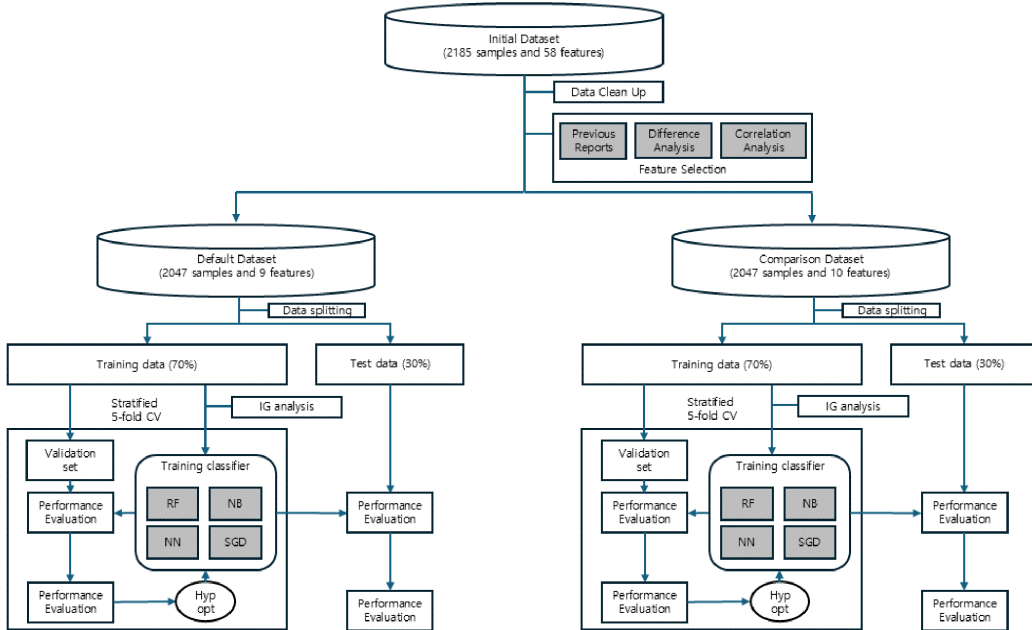


Figure1. The overall process for developing Sasang type classification model in current study

연구에서 직접적으로 사용된 기계학습 및 사상체질 진단모델 생성 과정(워크플로우, Figure 2, 3)을 쉽게 이해하고 자신의 연구에 직접 사용할 수 있도록 하였다²⁹.

사상체질 진단에서 심리적, 신체적, 혈액학적 특성의 상대적 기여도를 정보이득 분석으로 평가하는 것은 임상적 정확성과 맞춤형 치료전략을 수립함에 필수적이며, 기계학습으로 얻은 진단모델의 해석 가능성을 높일 수 있다. 이에 본 연구에서는 다양한 임상특성을 통합하고 체계적인 기계학습을 진행하여 이해 가능한 사상체질 진단모델을 개발하고자 하였으며, 유효한 진단모델의 단순 제시를 넘어서 분석에 사용되는 체질진단법과 투입 임상지표들의 한계성도 동시에 검토될 수 있도록 하였다.

본 연구는, 사용된 체계적인 분석절차(Figure 1)와 Orange 워크플로우(Figure 2, 3)를 통해 기계학습을 사용한 사상체질 진단모델 개발과정의 직관적 이해와 손쉬운 실행을 가능하게 하며, 임상 현장에서의 체질별 임상 특성의 이해도와 활용도를 높임으로써 과학

적인 사상의학 진단에 기여할 수 있을 것이다.

II. Methods

1. 연구 절차 및 피험자

본 연구에 활용된 피험자의 사상체질, 심리적, 신체적, 혈액학적 특성은 한국한의학연구원 한의임상정보은행(Korean Medicine Data Center, KDC) 자료를 활용하였으며, 서면동의서는 임상 데이터 수집 시 취득되었다. 본 연구는 임상 데이터를 분양 받기 이전에 기관 윤리심의위원회(KSU-2010001)의 승인을 받았으며, 인간 대상 의료 연구에 관한 헬싱키 선언의 윤리 원칙을 준수하였다. 본 연구에는 분양 받은 2,185명의 데이터 중에서 사상체질진단, 심리적, 신체적, 혈액학적 특성이 모두 있는 2,047명의 데이터를 분석에 활용하였다(Figure 1)¹⁹.

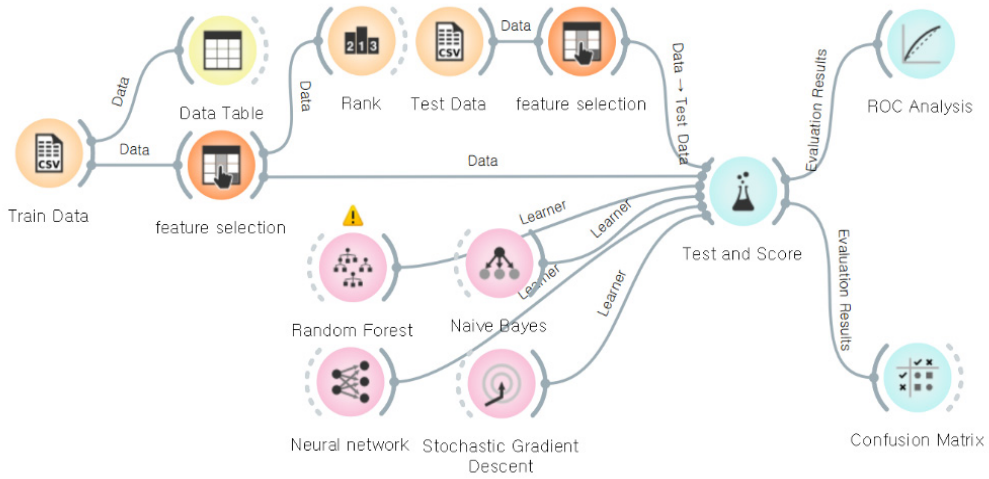


Figure 2. Visual workflow of orange software for machine learning analysis used in current study

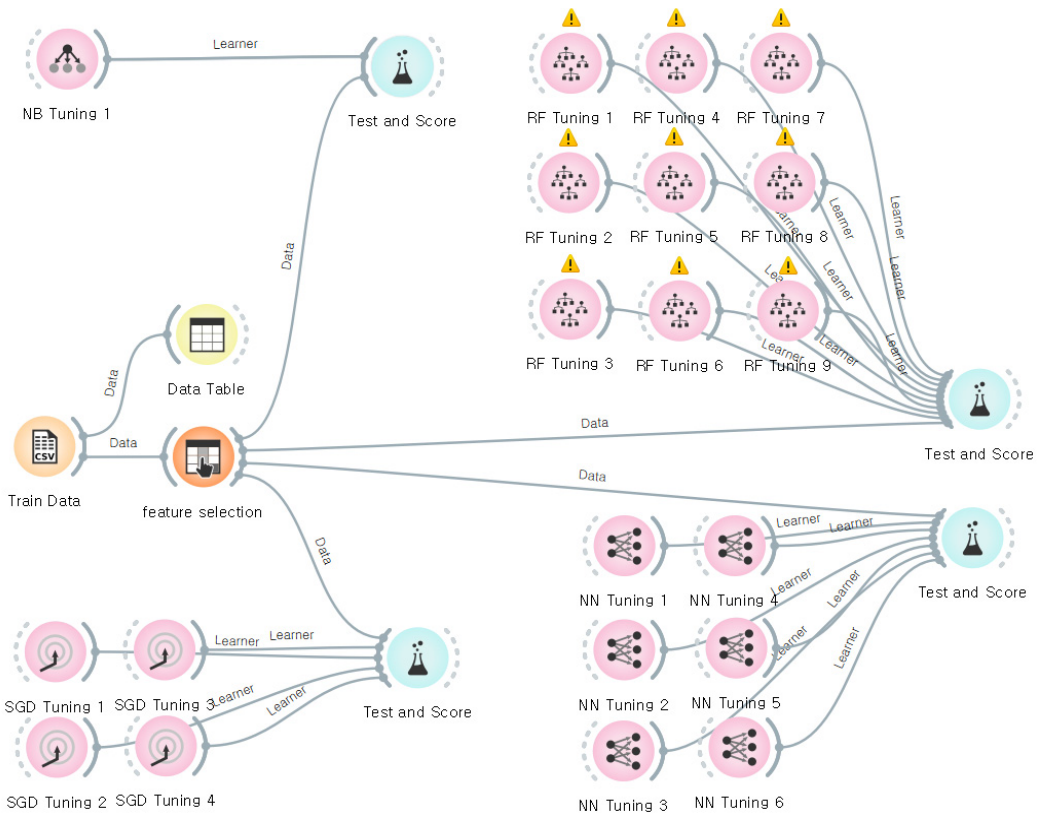


Figure 3. Workflow of hyperparameter tuning for diagnostic models used in current study

2. 연구에 사용된 임상특성

1) 사상체질 진단

본 연구에서의 사상체질의 진단에는 KS-15가 사용되는데, 체형(1문항, BMI)과 함께 성격(6문항)과 소증(8문항)에 대한 3점척도 자기보고식 설문으로 구성되어 있다. 보고된 응답을 바탕으로 태음인, 소음인, 소양인에 대한 확률 점수(체질 점수)가 산출되며, 가장 높은 점수에 해당하는 체질을 피험자의 체질로 결정한다³¹.

2) 심리적 특성

사상성격검사(Sasang Personality Questionnaire, SPQ)는 양양 기질 및 사상체질별 심리 특성을 측정하는 14 문항 자기보고형 임상검사³²로, 행동태도(SPQ-Behavior, SPQ-B), 인지방식(SPQ-Cognition, SPQ-C), 정서반응(SPQ-Emotion, SPQ-E)의 세가지 하위척도의 총합은 SPQ-total(SPQ-T)이다¹⁹. 높은 SPQ-T 점수는 적극적으로 행동하며, 자신의 의견을 표현하며 감정을 외부로 잘 드러내는 심리적 특징³²을 의미한다. 검사 문항들은 두 가지 상반된 성격 특성을 묘사하는 단어로 구성되어 있으며, 둘 중의 하나 또는 '불명확'을 답변으로 선택한다. 선행연구에서 SPQ-T, SPQ-B, SPQ-C, SPQ-E의 내적일치도는 각각 0.722, 0.769, 0.581, 0.641로 보고되었다¹⁹.

NEO-성격검사(NEO-Personality Inventory, NEO-PI)는 개인의 심리적 차이를 평가하기 위한 60 문항 자기보고형 검사로, 각 문항은 리커트 5점척도(1=전혀 아님, 5=매우 심함)로 구성되어 있다³³. NEO-PI는 5개의 하위척도로 구성되어 있는데, 불안, 분노, 우울 등의 부정적 감정인 신경증(Neuroticism, Neu), 사회성과 활기, 명랑함을 의미하는 외향성(Extraversion, Ext), 미적 감수성, 지적 호기심, 경험의 다양성을 의미하는 경험에 대한 개방성(Openness to new experience, Ogn), 신뢰, 이타주의 및 공감을 의미하는 동조성(Agreeableness, Agr), 목표를 향한 규칙적인 노력과 원칙의 준수를 의미하는 성실성(Conscientiousness, Con)이다³³. 한글판

NEO-PI의 신경증, 외향성, 경험의 개방성, 친화성, 성실성 척도의 내적 일치도는 각각 0.850, 0.766, 0.691, 0.644, 0.720로 보고되었다³⁴.

3) 신체적 특성

한의임상 정보은행에서 분양 받은 신체적 특성은 신장(m), 체중(kg), 그리고 8부위(이마, 목, 겨드랑이, 가슴, 갈비뼈, 허리, 골반, 엉덩이)의 신체 둘레(cm)이다. 체질량지수, 폰더랄 지수, 허리엉덩이비율, 기초대사량, 체중당기초대사량은 연령, 성별, 체중, 키, 허리 둘레, 엉덩이둘레를 활용하여 계산된다.

이마 둘레는 미간과 후두점을 연결하는 수평선, 목 둘레는 갑상선 연결과 후두 연결 사이를 통과하는 수평선, 겨드랑이 둘레는 오른쪽과 왼쪽 겨드랑이를 통과하는 수평선, 가슴 둘레는 오른쪽과 왼쪽 유두를 통과하는 수평선, 갈비뼈 둘레는 오른쪽과 왼쪽 7번째와 8번째 갈비 연결 관절 돌출 부위를 통과하는 수평선, 허리 둘레는 배꼽 주위를 통과하는 수평선, 골반 둘레는 오른쪽과 왼쪽 상전장골극(위앞엉덩뼈가시)을 통과하는 수평선, 엉덩이 둘레는 치골결합의 상단을 통과하는 수평선의 길이이다.

체질량지수(Body Mass Index, BMI, kg/m²)는 체지방의 양을 추정하기 위하여 체중을 키의 제곱으로 나눈 값³⁵이며, 폰더랄 지수(Ponderal Index, PI, kg/m³)는 신체 발달정도를 평가하기 위하여 체중을 키의 세제곱으로 나눈 값³⁶이다. 허리엉덩이비율(Waist Hip Ratio, WHR)은 허리둘레와 엉덩이둘레 간의 비율로 복부비만을 평가하는 지표³⁷이다.

기초대사량(Resting Metabolic Rate, RMR, kcal/day)은 신장, 체중, 성별을 이용하는 다음 세계보건기구의 공식을 사용하여 계산한다²⁵. 남성에 있어서 연령에 따른 공식은 다음과 같다. $15.4 * W - 0.27 * H + 717$ (나이 < 31), $11.3 * W + 0.16 * H + 901$ (30 < 나이 < 61), $8.8 * W + 11.28 * H - 1071$ (나이 > 60). 여성에 있어서 연령에 따른 공식은 다음과 같다. $13.3 * W + 3.34 * H + 35$ (나이 < 31), $8.7 * W - 0.25 * H + 865$ (30 < 나이 < 61), $9.2 * W + 6.37 * H - 302$

(나이 > 60). 체중당기초대사량(Resting Metabolic Rate per weight, RMRw, kcal/day/kg)은 체중 당 기초대사량으로, 피험자의 기초대사량에서 체중의 영향을 통제 한 것이다.

4) 혈액학적 특성

신장기능을 평가하는 Blood Urea Nitrogen (BUN)과 Creatinine, 간 기능을 평가하는 Direct bilirubin (D.Bili), Total bilirubin (T.Bili), Aspartate Aminotransferase (AST), Alanine Aminotransferase (ALT), Alkaline Phosphatase (ALP), 대사 및 지질량을 평가하는 Fasting Blood Sugar (FBS), Total Cholesterol (T.Chol), Triglycerides (Tg), High-Density Lipoprotein (HDL), Low-Density Lipoprotein (LDL), 혈구학적 특성을 평가하는 White Blood Cells (WBC), Red Blood Cells (RBC), Hemoglobin (Hb), Hematocrit (HCT), Mean Corpuscular Volume (MCV), Mean Corpuscular Hemoglobin (MCH), Mean Corpuscular Hemoglobin Concentration (MCHC), Red Cell Distribution Width (RDW), Platelet, Eosinophil이 사용된다.

3. 통계 분석

1) 인구학적 분석

성별간 연령, 교육 수준, 혼인 상태, 심리적 및 신체적 특성의 유의한 차이를 분석함에 있어서는 χ^2 검정과 t-test가 사용되었다.

2) 진단모델 입력특성의 선택

사상체질 진단모델에 입력될 심리적, 신체적, 혈액학적 임상 특성은, 유사변수를 제거해 예측 성능과 해석력을 높이고 과적합(overfitting)을 예방하기 위하여 다음과 같은 단계를 거쳐 선택되었다.

- 가. 선행연구 검토를 통해서 사상체질 간 유의한 차이가 보고된 심리적, 신체적, 혈액학적 특성을 확인한다.

- 나. 선행연구에서 유의한 차이가 보고된 특성들을 대상으로, 본 연구 데이터를 공분산분석(Analysis of Covariance, ANCOVA)을 시행하여 세 사상체질 그룹간 유의한 차이를 분석한다.

- 다. 세 체질간 유의한 차이가 확인된 특성을 대상으로, 심리, 신체, 혈액학적 특성별로 Pearson 상관 분석을 시행한다.

- 라. 상관 분석 결과, ANCOVA 분석에서 가장 높은 F값을 가진 특성 및 이들과의 상관계수의 절대값이 0.3 이하인 특성을 기준특성(reference feature)로는 선택한다.

- 마. 기준특성과 비교했을 때 Pearson 상관계수의 절대값이 0.6 미만인 특성들을 입력특성으로 선택한다.

3) 진단모델을 위한 데이터 세트의 구성

본 연구에서는 지지점(support point)을 활용하여 2,047명의 데이터를 2개로 분할하는데³⁸, 알고리즘들이 체질 진단을 학습하기위한 학습 데이터(전체의 70%, 1,435개)와 도출된 사상체질 진단모델의 성능평가를 위한 테스트 데이터(전체의 30%, 612개)로 사용한다(Figure 1).

이상에서 선택된 학습데이터(1,435명)의 심리적, 신체적, 혈액학적 임상정보를 입력특성으로 이들의 사상체질을 목표특성으로 구성하여 학습을 진행하고, 테스트 데이터를 사용하여 사상체질 진단모델의 성능평가를 진행한다(Figure 1).

4) 입력특성의 중요도 평가

입력특성의 중요도 평가를 위한 정보이득(Information Gain, IG)은 입력특성이 불확실성을 감소시키는 정도를 평가하는 지표³⁰, 또는 목표특성 예측에 대한 입력특성의 효과성을 의미한다. 정보이득(IG)은 목표특성(Y)의 엔트로피에서 입력특성(X)을 고려한 조건부 엔트로피를 뺀 값으로, $IG(X)=H(Y)-H(Y|X)$ 로 정의된

다. 정보이득은 데이터 분포에 따라 값이 달라지는 지표로³⁰, 본 연구에서는 학습데이터를 이용해 입력특성의 정보이득을 산출한다(Figure 1). 입력특성의 IG 값은 상대적 비교로 해석³⁹하기에, 체질진단에 기여하는 상대적 중요도를 심리, 신체, 혈액학적 특성별로 총합과 평균을 산출하여 비교한다.

5) 진단모델을 위한 학습 알고리즘

사상체질 진단모델 학습을 위한 ML의 알고리즘으로는 다중 클래스에 적합한 Random Forest (RF), Naïve Bayes (NB), Neural Network(NN), Stochastic Gradient Descent (SGD)를 사용한다.

RF는 여러 개의 결정 트리를 앙상블하여 예측의 정확도를 높이고 과적합을 방지한다. 다양한 데이터 패턴을 포착할 수 있으며, 특성의 중요도 평가에 유용하다. 데이터의 노이즈에 강하고, 특성 간의 상호작용을 잘 모델링할 수 있어 체질 분류 문제에 적합한 것으로 보고되었다⁴⁰. NB는 각 특성이 독립적이라는 가정 하에 확률적 분류를 수행하며, 계산적으로 효율적이고 대규모 데이터에서도 빠른 학습과 예측이 가능하다⁴¹. NN는 복잡한 비선형 관계를 모델링할 수 있는 능력을 갖추고 있으며, 고차원 데이터의 패턴을 효과적으로 학습한다⁴². SGD는 대규모 데이터에서 빠르고 효율적인 학습을 가능하게 하는 최적화된 알고리즘으로, 반복적인 파라미터 업데이트를 통해 모델의 수렴을 가속화한다. 메모리 사용을 최소화하며, 대규모 데이터 문제에서도 효율적으로 최적화를 수행할 수 있다⁴³.

6) 진단모델의 성능평가

사상체질 진단모델의 성능평가를 위한 혼동행렬 (Confusion Matrix)은 분류 성능을 평가하는 다음 지표들을 계산하기 위한 기초 자료로⁴⁴, 진단모델의 예측 결과를 실제 체질진단과 비교하여 표로 정리한 것이다.

최종 성능 평가는 테스트데이터(30%)를 대상으로 수행하며, 사상체질 진단모델의 성능평가에는 분류

정확도(Classification Accuracy, CA), 정밀도(Precision), 재현율(Recall), F1 점수(F1 score), 매튜스 상관계수(Matthews Correlation Coefficient, MCC)를 사용한다(Figure 1).

분류 정확도(CA)는 전체 예측 중 올바르게 예측된 샘플의 비율을 나타내며, 모델의 전반적인 성능을 평가하는 기본적인 지표이다⁴⁵. 정밀도(Precision)는 양성으로 예측된 샘플 중 실제로 양성인 샘플의 비율이다. 재현율(Recall)은 실제 양성 샘플 중에서 모델이 올바르게 예측한 양성 샘플의 비율이다. F1점수(F1 Score)는 정밀도와 재현율의 조화 평균으로, 정밀도와 재현율이 유사할수록 F1점수가 높아진다⁴⁵. 매튜스 상관계수(MCC)는 참 양성, 참 음성, 거짓 양성, 거짓 음성을 모두 반영하여 모델의 성능을 종합적으로 평가하는 지표로, 특정 클래스(체질 그룹)의 불균형 상황에서도 안정적인 평가가 가능하며⁴⁶, 다른 지표들(정확도, 정밀도, 재현율, F1 점수)과 달리 클래스 간 균형을 민감하게 반영한다. MCC는 -1에서 1 사이의 값을 가지며, 1은 완벽한 예측, 0은 무작위 예측, -1은 완전히 상반된 예측을 나타낸다.

민감도(Sensitivity)와 특이도(Specificity)을 분석함에 있어 수신자조작특성(Receiver Operating Characteristic, ROC) 커브와 곡선하면적(Area Under the Curve, AUC)을 활용한다. ROC 커브는 분류 임계값(threshold)에 따라 변화하는 민감도와 특이도를 시각화하여 모델의 판별력을 직관적으로 평가하며, AUC 값(0.5~ 1)은 1에 가까울수록 우수한 성능을 의미한다.

7) 알고리즘별 사전 설정 값의 최적화

알고리즘별로 최적화된 사전 설정 값은 다음 절차를 통해 선정한다. 알고리즘별로 다양한 사전 설정 값들을 탐색하는데 학습데이터(70%)를 활용한 5-겹 교차검증(5-fold cross validation) 시행을 시행하는데, 이들 탐색 값에서 가장 높은 평균 분류정확도(CA)를 보인 것을 본 연구에 사용한다.

RF에 있어서는 분할시 고려할 특성수(3)와 옵션(class balancing)을 사용한 상태에서 트리 수(100, 200, 300)와 최대 깊이(5, 10, 15)를 탐색한다. NB에서는

조정 가능한 설정값이 없었기에 기본값을 사용한다. NN에서는 활성화함수(ReLU), 최적화 알고리즘(Adam), 정규화 계수(0.0001)를 사용하며, 은닉층 구조(128 - 64 - 32, 64 - 32)와 최대 반복 횟수(100, 200, 300)를 탐색한다. SGD에 있어서는 로지스틱 손실(logistic loss)과 L2 정규화를 적용하며, 학습률(최적 스케줄), 반복 횟수(1,000), 허용오차(1×10^{-3})를 고정한 상태에서, 규제 강도 (10^{-5} , 10^{-4} , 10^{-3} , 10^{-2})를 탐색한다.

4. 통계 및 기계학습 도구

본 연구에서의 통계적 분석에는 친화적 그래픽 인터페이스로 R 언어를 사용하는 오픈소스 소프트웨어인 Jamovi 2.4.11 (The Jamovi Team, 22 October 2023, www.jamovi.org)를 사용하였다. 통계 분석의 결과는 평균±표준오차 또는 빈도(%)로 표기하였으며, 통계적 유의수준으로는 $p < 0.05$, $p < 0.01$ 및 $p < 0.001$ 이 사용되었다.

기계학습에는 데이터 분석, 데이터 마이닝, 기계학습과 데이터 시각화를 위한 시각적 프로그래밍을 사용하는 소프트웨어 패키지인 Orange 3.36.2 (university of Ljubljana, 31 October 2023, orangedatamining.com)를 활용하였다. 시각적 프로그래밍을 통해 사용자가 디자인한 위젯 요소들을 연결한 워크플로우를 만들게 되며, 데이터 분석은 이 워크플로우에 데이터를 투입하여 진행된다.

III. Results

1. 인구학적 분석

본 연구(2,047명)에서의 남녀 피험자의 인구통계학적, 심리학적, 신체적 특성은 Table 1에 제시되었다.

남녀 그룹 간 연령($t = -4.667$, $p < 0.001$), 교육 수준($\chi^2 = 56.336$, $p < 0.001$), 결혼($\chi^2 = 56.787$, $p < 0.001$)에서 유의미한 차이가 있었다.

심리 특성에 있어서, 남성이 NEO-PI 하위 척도에 서 여성보다 Ext 점수($t = 3.593$, $p < 0.001$)가 높았고, Neu 점수($t = -6.196$, $p < 0.001$)와 Agr 점수($t = -5.070$, $p < 0.001$)는 낮았다. 또한, 남성은 여성보다 더 높은 SPQ-C 점수($t = 2.872$, $p = 0.004$)를 보였다.

신체 특성에 있어서, 남성은 키($t = 50.378$, $p < 0.001$), 체중($t = 35.321$, $p < 0.001$), BMI($t = 10.801$, $p < 0.001$), WHR($t = 8.180$, $p < 0.001$), RMR($t = 79.801$, $p < 0.001$), RMRw($t = 12.079$, $p < 0.001$)에서 여성보다 컸다.

사상체질 유형 분포는 남녀 그룹 간에 유의미한 차이($\chi^2 = 49.153$, $p < 0.001$)가 있었다. 태양인은 제시되지 않았으며, 소양인, 태음인, 소음인의 분포는 남성에 있어서는 각각 164명, 361명, 106명 이었으나, 여성에 있어서는 각각 496명, 574명, 346명 이었다.

2. 선행연구 검토를 통한 입력특성의 선택

심리적, 신체적, 혈액학적 특성에 있어서 사상체질 그룹간 차이에 대한 선행 연구들을 검토하였다.

심리적 특성에 있어서, SPQ의 3가지 하위척도^{13,14,22}와 SPQ-T, NEO-PI의 외향성(Extraversion)^{13,22}, 기질및 성격검사(TCI)의 자극추구와 위험회피²³에 있어서 사상 체질 그룹간 유의한 차이가 보고되었다.

신체적 특성에 있어서, BMI^{10,15,22}, PI^{10,22}, WHR⁴⁷, RMRw²⁵, 늑갈각(Costal angle)²⁸, 목둘레⁴⁸, 가슴둘레⁴⁸, 허리둘레⁴⁹에 있어서 사상 체질 그룹간 유의한 차이가 보고되었다. 다만, 선행 연구를 통해 BMI가 사상 체질 그룹간 핵심적인 차이^{10,15,22}라는 연구와 과대평가되었다는 연구²⁴가 공존함을 확인하였기에, 본 연구에서는 피험자의 모든 특성들을 포함하는 기본 데이터셋(Default dataset)와, BMI 만을 제외한 비교 데이터셋(Comparison dataset)의 2 가지 데이터셋을 사용하여 체질 진단에 있어서의 BMI의 영향력을 분석하였다 (Figure 1).

혈액학적 특성에 있어서, 세 사상체질 그룹 간의 유의한 차이³⁰를 확인할 수 없었다.

Table 1. Demographic Features of the Participants in Current Study

	Male(n=631)	Female(n=1416)	
Age ^{***}	39.777±12.800	42.415±11.344	$t=-4.667, p<0.001$
Education ^{***}			$X^2=56.336, p<0.001$
None	0	3	
Elementary school	10	41	
Middle school	21	97	
High school	246	693	
Bachelor	277	503	
Ph.D.	77	80	
Marriage ^{***}			$X^2=56.787, p<0.001$
Single	207	280	
Married	421	1075	
Divorced	2	30	
Deceased	1	31	
Psychological characteristics			
SPQ-Total	28.471±5.412	28.322±5.533	$t=0.565, p=0.572$
SPQ-Behavior	10.700±2.633	10.772±2.610	$t=-0.570, p=0.569$
SPQ-Cognition ^{**}	9.929±2.268	9.606±2.383	$t=2.872, p=0.004$
SPQ-Emotion	7.842±2.043	7.944±2.084	$t=-1.036, p=0.300$
NEO-PI			
Neuroticism ^{***}	42.651±10.993	45.882±10.850	$t=-6.196, p<0.001$
Extraversion ^{***}	53.762±10.274	51.972±10.471	$t=3.593, p<0.001$
Openness	54.675±8.217	54.263±7.986	$t=1.067, p=0.286$
Agreeableness ^{***}	57.065±8.946	59.151±8.437	$t=-5.070, p<0.001$
Conscientiousness	56.761±8.809	56.449±8.596	$t=0.751, p=0.453$
Physical features			
Height ^{***}	172.006 ±5.860	158.791±5.303	$t=50.378, p<0.001$
Weight ^{***}	72.407±10.042	57.639±8.086	$t=35.321, p<0.001$
Anthropometrics			
BMI ^{***}	24.466±3.102	22.869±3.083	$t=10.801, p<0.001$
PI	14.246±1.910	14.429±2.086	$t=-1.881, p=0.060$
WHR ^{***}	0.899±0.050	0.878±0.056	$t=8.180, p<0.001$
RMR ^{***}	1745.269±152.498	1324.997±84.513	$t=79.801, p<0.001$
RMRw ^{***}	24.305±1.807	23.236±1.866	$t=12.079, p<0.001$
Sasang type ^{***}			$X^2=49.153, p<0.001$
So-Yang	164	496	
Tae-Eum	361	574	
So-Eum	106	346	

NEO-Personality Inventory, NEO-PI; Sasang Personality Questionnaire, SPQ; Body Mass Index, BMI; Ponderal Index, PI; Waist Hip Ratio, WHR; Resting Metabolic Rate, RMR; Resting Metabolic Rate per weight, RMRw

^{*}, $p < .05$; ^{**}, $p < .01$; ^{***}, $p < .001$

3. 사상체질간 차이 분석을 통한 입력특성의 선택

심리적, 신체적, 혈액학적 특성에 있어서 사상체질 그룹간 차이를 확인하기 위하여 학습데이터(1,435명)

에 대한 ANCOVA를 시행한 결과는 Table 2 및 Table 3과 같다.

Table 2. Estimated Biopsychological and Physical Features of Each Sasang Type Groups.

	So-Yang	Tae-Eum	So-Eum	Statistical analysis	Post-hoc
SPQ					
SPQ-Total	31.453±0.243	28.002±0.194	24.490±0.286	F=186.185, p<0.001	SY>TE>SE
SPQ-Behavior ^{***}	12.453±0.112	10.349±0.090	9.054±0.132	F=221.158, p<0.001	SY>TE>SE
SPQ-Cognition ^{***}	10.574±0.110	9.879±0.088	8.163±0.130	F=110.727, p<0.001	SY>TE>SE
SPQ-Emotion ^{***}	8.426±0.101	7.773±0.081	7.273±0.119	F=30.443, p<0.001	SY>TE>SE
NEO-PI					
Neuroticism ^{***}	42.156±0.523	44.510±0.418	45.506±0.616	F=10.624, p<0.001	TE>SY, SE>SY
Extraversion ^{***}	56.670±0.476	52.342±0.381	47.798±0.561	F=78.766, p<0.001	SY>TE>SE
Openness ^{**}	55.490±0.399	53.854±0.319	54.641±0.469	F=5.361, P=0.005	SY>TE
Agreeableness	58.522±0.413	57.850±0.330	58.053±0.487	F=0.847, P=0.429	
Conscientiousness ^{***}	58.087±0.420	55.986±0.336	56.983±0.495	F=7.950, P<0.001	SY>TE
Anthropometrics					
BMI ^{***}	22.409±0.101	25.904±0.080	20.672±0.118	F=804.227, P<0.001	TE>SY>SE
PI ^{***}	13.560±0.066	15.752±0.053	12.472±0.078	F=733.354, P<0.001	TE>SY>SE
WHR ^{***}	0.880±0.002	0.909±0.002	0.861±0.003	F=121.965, P<0.001	TE>SY>SE
RMR ^{***}	1496.800±4.589	1593.011±3.667	1456.767±5.403	F=270.634, P<0.001	TE>SY>SE
RMRw ^{***}	24.287±0.065	22.597±0.052	25.533±0.077	F=564.192, P<0.001	SE>SY>TE
Circumference					
Forehead ^{***}	56.926±0.086	57.489±0.068	56.443±0.101	F=40.433, P<0.001	TE>SY>SE
Neck ^{***}	35.447±0.090	37.184±0.072	34.539±0.106	F=256.462, P<0.001	TE>SY>SE
Axillary ^{***}	90.972±0.247	96.043±0.198	87.902±0.291	F=310.660, P<0.001	TE>SY>SE
Chest ^{***}	89.859±0.284	96.076±0.227	85.924±0.334	F=363.606, P<0.001	TE>SY>SE
Rib ^{***}	80.607±0.280	87.640±0.223	76.941±0.329	F=431.152, P<0.001	TE>SY>SE
Waist ^{***}	82.240±0.318	89.591±0.254	78.312±0.374	F=368.828, P<0.001	TE>SY>SE
Pelvic ^{***}	88.464±0.286	94.419±0.228	85.489±0.337	F=289.766, P<0.001	TE>SY>SE
Hip ^{***}	93.429±0.254	98.557±0.203	90.864±0.299	F=272.809, P<0.001	TE>SY>SE
Width					
Axillary ^{***}	31.500±0.090	33.045±0.072	30.686±0.106	F=203.134, P<0.001	TE>SY>SE
Chest ^{***}	29.014±0.107	31.070±0.086	28.023±0.126	F=241.904, P<0.001	TE>SY>SE
Rib ^{***}	27.883±0.095	29.781±0.076	27.015±0.112	F=256.908, P<0.001	TE>SY>SE
Waist ^{***}	27.973±0.106	30.282±0.084	26.783±0.124	F=324.434, P<0.001	TE>SY>SE
Pelvic ^{***}	30.190±0.120	31.348±0.096	29.537±0.142	F=65.550, P<0.001	TE>SY>SE

NEO-Personality Inventory, NEO-PI; Sasang Personality Questionnaire, SPQ; Body Mass Index, BMI; Ponderal Index, PI; Waist Hip Ratio, WHR; Resting Metabolic Rate, RMR; Resting Metabolic Rate per weight, RMRw
^{*}, p < .05; ^{**}, p < .01; ^{***}, p < .001

Table 3. Estimated Hemological Features of Each Sasang Type Groups.

	So-Yang	Tae-Eum	So-Eum	Statistical analysis	Post-hoc
BUN*	14.188±0.160	13.770±0.128	13.502±0.188	F=4.439, P=0.012	SY>SE
Creatinine	0.955±0.005	0.963±0.004	0.949±0.005	F=2.524, P=0.080	
FBS***	92.151±0.819	95.621±0.654	89.645±0.964	F=14.831, P<0.001	TE>SY, TE>SE
D.Bili**	0.272±0.005	0.256±0.004	0.273±0.006	F=4.753, P=0.009	SY>TE, SE>TE
T.Bili	0.876±0.016	0.842±0.013	0.872±0.019	F=1.770, P=0.171	
AST***	24.649±0.482	26.599±0.385	24.040±0.568	F=9.111, P<0.001	TE>SY, TE>SE
ALT***	22.499±0.817	28.306±0.653	21.083±0.962	F=26.345, P<0.001	TE>SY, TE>SE
ALP**	67.077±0.897	69.760±0.717	65.810±1.056	F=5.817, P=0.003	TE>SY, TE>SE
T.Chol***	184.285±1.553	190.505±1.241	181.636±1.829	F=9.965, P<0.001	TE>SY, TE>SE
Tg***	114.525±4.108	145.797±3.284	102.800±4.837	F=34.275, P<0.001	TE>SY, TE>SE
HDL***	56.596±0.597	53.219±0.477	58.387±0.703	F=21.971, P<0.001	SY>TE, SE>TE
LDL***	109.631±1.463	117.038±1.170	107.985±1.723	F=13.100, P<0.001	TE>SY, TE>SE
WBC***	5.689±0.073	5.997±0.058	5.547±0.086	F=11.558, P<0.001	TE>SY, TE>SE
RBC***	4.419±0.016	4.517±0.012	4.359±0.018	F=29.098, P<0.001	TE>SY>SE
Hb***	14.078±0.055	14.351±0.044	13.937±0.064	F=16.878, P<0.001	TE>SY, TE>SE
HCT***	41.642±0.119	41.642±0.119	40.477±0.176	F=17.074, P<0.001	TE>SY, TE>SE
MCV	92.684±0.261	92.246±0.209	92.941±0.308	F=2.042, P=0.130	
MCH	31.864±0.105	31.793±0.084	32.002±0.124	F=1.007, P=0.366	
MCHC	34.360±0.036	34.451±0.028	34.417±0.042	F=2.076, P=0.126	
RDW	13.156±0.071	13.109±0.056	13.153±0.083	F=0.180, P=0.835	
Platelet***	227.973±2.587	238.913±2.068	221.227±3.047	F=13.354, P<0.001	TE>SY, TE>SE
Eosinophil**	169.709±7.480	177.853±5.978	144.532±8.807	F=5.139, P=0.006	TE>SE

Blood Urea Nitrogen, BUN; Fasting Blood Sugar, FBS; Direct bilirubin, D.Bili; Total bilirubin, T.Bili; Aspartate Aminotransferase, AST; Alanine Aminotransferase, ALT; Alkaline Phosphatase, ALP; Total Cholesterol, T.Chol; Triglycerides, Tg; High-Density Lipoprotein, HDL; Low-Density Lipoprotein, LDL; White Blood Cells, WBC; Red Blood Cells, RBC; Hemoglobin, Hb; Hematocrit, HCT; Mean Corpuscular Volume, MCV; Mean Corpuscular Hemoglobin, MCH; Mean Corpuscular Hemoglobin Concentration, MCHC; Red Cell Distribution Width, RDW

*, $p < .05$; **, $p < .01$; ***, $p < .001$

심리적 특성에 있어서, 세 사상 체질간 유의미한 차이는 SPQ-T(F=186.185, $p<0.001$), SPQ-B(F=221.158, $p<0.001$), SPQ-C(F=110.727, $p<0.001$), SPQ-E(F=30.443, $p<0.001$), Ext(F=78.766, $p<0.001$)에서 확인할 수 있었다.

신체적 특성에 있어서, 세 사상 체질간 유의한 차이는 BMI(F=804.227, $p<0.001$), PI(F=733.354, $p<0.001$), WHR(F=121.965, $p<0.001$), RMR(F=270.634, $p<0.001$), RMRw(F=564.192, $p<0.001$)에 있었다. 신체 둘레에 있어서는 이마둘레(F=40.433, $p<0.001$), 목둘레(F=256.462, $p<0.001$), 겨드랑이둘레(F=310.660, $p<0.001$),

가슴둘레(F=363.606, $p<0.001$), 갈비뼈둘레(F=431.152, $p<0.001$), 허리둘레(F=368.828, $p<0.001$), 골반둘레(F=289.766, $p<0.001$), 엉덩이둘레(F=272.809, $p<0.001$)에서, 신체 너비에서 있어서는 겨드랑이너비(F=203.134, $p<0.001$), 가슴너비(F=241.904, $p<0.001$), 갈비뼈너비(F=256.908, $p<0.001$), 허리너비(F=324.434, $p<0.001$), 골반너비(F=65.550, $p<0.001$)에서 세 사상 체질간 유의한 차이가 있었다.

혈액학적 특성에 있어서, RBC(F=29.098, $p<0.001$)에서만 세 사상 체질간 유의한 차이를 확인할 수 있었다.

4. 상관분석을 통한 입력특성의 선택

심리적 특성에 대한 상관분석 결과는 Suppl. Table 2에 제시되었다. 심리적 특성에 있어서 선행연구 검토와 ANCOVA 분석 모두에서 사상체질 그룹간 유의한 차이가 확인된 것은 SPQ-T, SPQ-B, SPQ-C, SPQ-E, NEO-PI Ext였다. ANCOVA(Table 2)에서 F값이 가장 큰 것을 확인된 SPQ-B($F=221.158$)를 기준특성으로 사용하였는데, SPQ-B와의 Pearson 상관계수 절대값이 0.3 미만으로 확인된 특성은 없었다. 기준 특성과의 상관계수 절대값이 0.6 미만인 SPQ-C, SPQ-E를 입력 특성으로 추가하였다.

신체적 특성에서 대한 상관분석 결과는 Suppl. Table 3에 제시되었다. 신체적 특성에 대한 분석은 BMI를 포함하는 기본 데이터세트와 BMI를 제외한 비교 데이터 세트의 두가지 방식으로 진행되었다.

기본 데이터 세트에 있어서, 선행연구 검토와 ANCOVA 분석 모두에서 사상체질 그룹간 유의한 차이가 확인된 것은 BMI, PI, WHR, RMRw, 목둘레, 가슴둘레, 허리둘레이다. ANCOVA(Table 2)에서 F값이 가장 큰 것으로 확인된 BMI($F=804.227$)를 기준특성으로 사용하였는데, BMI와 상관계수 절대값이 0.3 이하인 특성은 없었다. 기준특성과의 상관계수 절대값이 0.6 미만인 WHR, RMR, 이마둘레, 골반너비를 입력특성으로 추가하였다.

비교 데이터 세트에 있어서, 선행연구 검토와 ANCOVA 분석 모두에서 사상체질 그룹간 유의한 차이가 확인된 것은 PI, WHR, RMRw, 목둘레, 가슴둘레, 허리둘레였다. BMI를 제외하기로 하였으므로, BMI와 Pearson 상관계수가 높은 PI(0.933)를 기준특성에서 제외하고, ANCOVA(Table 2)에서 F값이 가장 큰 RMRw($F=564.192$)를 기준특성으로 사용하였다. RMRw와 상관계수 절대값이 0.3 미만인 RMR, 이마둘레, 목둘레는 기준특성으로 사용하였다. 기준특성과의 상관계수 절대값이 0.6 미만인 WHR과 골반너비를 입력특성으로 추가하였다.

혈액학적 특성 중 ANCOVA 분석에서 세 체질 간

유의한 차이를 보인 것은 RBC가 유일하여, 상관분석은 진행하지 않았다.

5. 데이터 세트의 구성 및 활용

본 연구에서는 특성선택 결과에 따라 기본 데이터 세트와 비교 데이터세트를 각각 구성하여 사상체질 진단모델을 구축(Figure 1)하였으며, 본 연구의 ML 분석을 위해 Orange 워크플로우(Figure 2, 3)를 구성하여 사용하였다.

기본 데이터세트와 비교 데이터세트는 신체적 특성에서만 차이(BMI 제외)를 보일 뿐, 심리적 및 혈액학적 특성은 동일하게 사용되었다. 기본 데이터세트의 입력특성은 총 9개로, 심리적 특성(SPQ-B, SPQ-C, SPQ-E), 혈액학적 특성(RBC), 신체적 특성(BMI, WHR, RMR, 이마 둘레, 골반 너비)으로 구성되었다. 비교 데이터세트의 입력특성은 총 10개로, 심리적 특성과 혈액학적 특성은 기본 데이터세트와 동일하나 신체적 특성은 WHR, RMR, RMRw, 이마 둘레, 목 둘레, 골반 너비로 구성되었다.

6. 정보이득 분석을 통한 입력특성의 기여도 평가

사상체질 진단에 있어서 입력특성들의 상대적 중요도를 평가하기 위하여, 학습데이터를 대상으로 정보이득(IG) 분석을 수행한 결과는 Figure 4와 같다. 기본 데이터세트에 포함된 신체적 특성에 있어서 IG 값의 총합은 0.951, 평균은 0.190로, 가장 높은 IG 값을 보였다. 각 입력특성의 IG 값은 BMI(0.531), RMR(0.207), WHR(0.116), 골반너비(0.055), 이마둘레(0.042)이었다. BMI와 RMR이 평균보다 높은 IG 값을 보였는데, 특히 BMI(0.531)는 기본 데이터세트에 포함된 모든 신체특성의 IG 값(0.951)의 절반을 넘게 차지하는 것으로 확인되었다.

비교 데이터세트에 포함된 신체적 특성에 있어서 IG 값의 총합은 0.858, 평균은 0.143으로, 기본 데이터

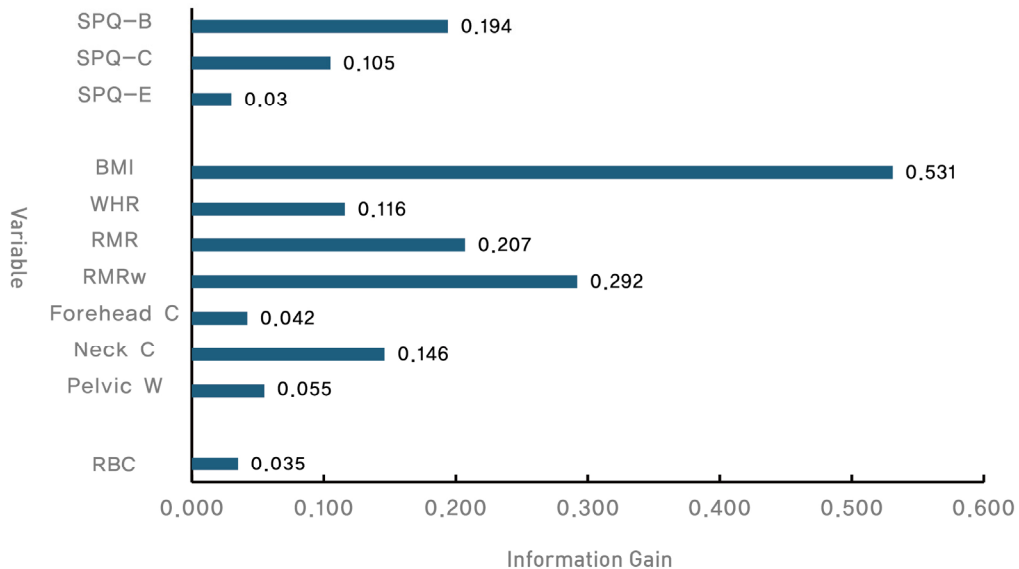


Figure 4. Information gain for selected clinical features in current study

Sasang Personality Questionnaire, SPQ; SPQ-Behavior, SPQ-B; SPQ-Cognition, SPQ-C; SPQ-Emotion, SPQ-E; Body Mass Index, BMI; Waist Hip Ratio, WHR; Resting Metabolic Rate, RMR; Resting Metabolic Rate per weight, RMRw; Forehead Circumference, Forehead C; Neck Circumference, Neck C; Pelvic Width, Pelvic W; Red Blood Cells, RBC

세트보다는 0.047 낮았다. 각 입력특성의 IG 값은 RMRw(0.292), RMR(0.207), 목둘레(0.146), WHR(0.116), 골반너비(0.055), 이마둘레(0.042)로 확인되었다. RMRw, RMR, 목둘레가 신체적 특성 전체 평균보다 높은 IG 값을 보였다.

심리적 특성에 있어서 IG 값의 총합은 0.329, 평균은 0.110로, 기본 데이터셋에 포함된 신체적 특성의 IG 값의 절반 정도인 것으로 확인되었다. 각 입력특성의 IG 값은 SPQ-B(0.194), SPQ-C(0.105), SPQ-E(0.030)으로, SPQ-B는 입력특성 중에서 3번째 높은 IG 값이었다.

혈액학적 특성의 경우, 입력특성으로 RBC만이 포함되었는데, IG 값은 0.035이었다.

7. 알고리즘 사전 설정 값의 최적화

기본 데이터셋과 비교 데이터셋 각각에 대해 학습데이터와 5-겹 교차검증을 통해 사전 설정 값을 선택하였으며, 해당 과정의 Orange 워크플로우는

Figure 4와 같다. RF의 트리 수(300)와 최대 깊이(15), NN의 은닉층 구조(64 - 32)와 최대 반복 횟수(100회), SGD의 규제 강도($\alpha=0.001$)가 두 데이터셋 모두에 동일하게 적용되었다.

8. 타당도 지표 분석을 통한 사상체질 진단모델의 성능 평가

사상체질 진단 모델의 성능은 테스트 데이터를 사용하여 평가하였으며, 그 결과는 Table 4(기본 데이터셋)와 Table 5(비교 데이터셋)에 제시되었다. 혼동 행렬은 Suppl. Table 3(기본 데이터셋)와 Suppl. Table 4(비교 데이터셋)에, 그리고 ROC 커브는 Figure 5에 제시되었다.

기본 데이터셋을 사용하여 사상체질 진단모델의 타당도를 분석하였을 때(Table 4), 모든 지표에서 RF가 가장 우수한 성능을 보였고, NN, SGD, NB의 순서를 보였다. 분류정확도, F1-score, Precision, Recall은 모두

Table 4. Diagnostic Performance of Sasang Type Classification Models Using Default Dataset

Classifier	CA	F1	Precision	Recall	MCC	AUC
RF	0.817	0.818	0.819	0.817	0.717	0.932
NB	0.783	0.782	0.782	0.783	0.663	0.913
NN	0.809	0.809	0.810	0.809	0.704	0.935
SGD	0.806	0.805	0.805	0.806	0.697	0.929

Random Forest, RF; Naïve Bayes, NB; Neural Network, NN; Stochastic Gradient Descent, SGD; Classification Accuracy, CA; Matthews Correlation Coefficient, MCC

Table 5. Diagnostic Performance of Sasang Type Classification Models Using Comparison Dataset

Classifier	CA	F1	Precision	Recall	MCC	AUC
RF	0.755	0.754	0.754	0.755	0.618	0.899
NB	0.742	0.740	0.741	0.742	0.598	0.888
NN	0.750	0.750	0.752	0.750	0.613	0.896
SGD	0.743	0.741	0.745	0.743	0.601	0.893

Random Forest, RF; Naïve Bayes, NB; Neural Network, NN; Stochastic Gradient Descent, SGD; Classification Accuracy, CA; Matthews Correlation Coefficient, MCC

RF(0.817~0.819)가 가장 높았고, NB(0.782~0.783)가 가장 낮은 값을 보였다. 이에 비해, MCC는 RF(0.717)와 NB(0.663)가 다른 지표들에 비해 상대적으로 낮았으며, 알고리즘 사이의 성능 차이도 더 뚜렷하게 나타났다. AUC는 NN(0.935)에서 가장 높게 나타났고, NB(0.913)에서 가장 낮게 나타났으나, 두 모델 간 차이는 0.024에 불과하였다(Figure 5).

비교 데이터셋을 사용한 사상체질 진단모델의 타당도(Table 5)는 기본 데이터셋과 유사한 것으로 확인되었으며, 모든 지표에서 RF, NN, SGD, NB의 순서로 성능이 낮아졌다. 분류 정확도, F1-score, Precision, Recall은 모두 RF(0.754~0.755)가 가장 높았으며, NB(0.740~0.742)이 가장 낮았다. MCC는 RF(0.618)와 NB(0.598)에서 각각 최대값과 최소값을 보였다. AUC는 RF(0.899)에서 가장 높게, NB(0.888)에서 가장 낮게 나타났으며, 두 모델 간 차이는 기본 데이터셋 보다 작은 0.011이었다(Figure 5).

알고리즘의 성능은 데이터셋의 종류와는 무관한 것을 알 수 있었는데, RF는 두 데이터셋 모두에서 안정적인고 우수한 성능을 보였으며 NB는 전반적으로

로 가장 낮은 수준의 성능을 보였다. NN과 SGD는 데이터셋에 따라 성능이 다소 달라졌으며, 상위권 내에서 서로 유사한 수준의 결과를 보여주었다.

IV. Discussion

본 연구에서는 사상의학 이론에 기반한 해석이 가능한 이론기반 분석과 임상 데이터와 오픈소스 ML 도구인 Orange를 활용한 데이터기반 분석을 동시에 활용하여 높은 실용성과 이론적 해석이 가능한 사상체질 진단 모델을 제안하고 그 성능을 다면적으로 분석하였다. 기존의 이론기반 분석은 예측력의 한계가, 데이터기반 분석은 해석의 어려움이 제기되었는데, 본 연구는 이론적 근거와 데이터 분석의 강점을 결합하여 이러한 한계를 보완하였다.

연구 결과, 체질진단에 중요한 임상특성은 신체적, 심리적, 혈액학적 특성의 순서로 기여도가 낮아지는 것을 확인할 수 있었다. 이와 함께, 신체적 특성에 있어서 (BMI가 포함된) 기본 데이터셋을 활용하였을

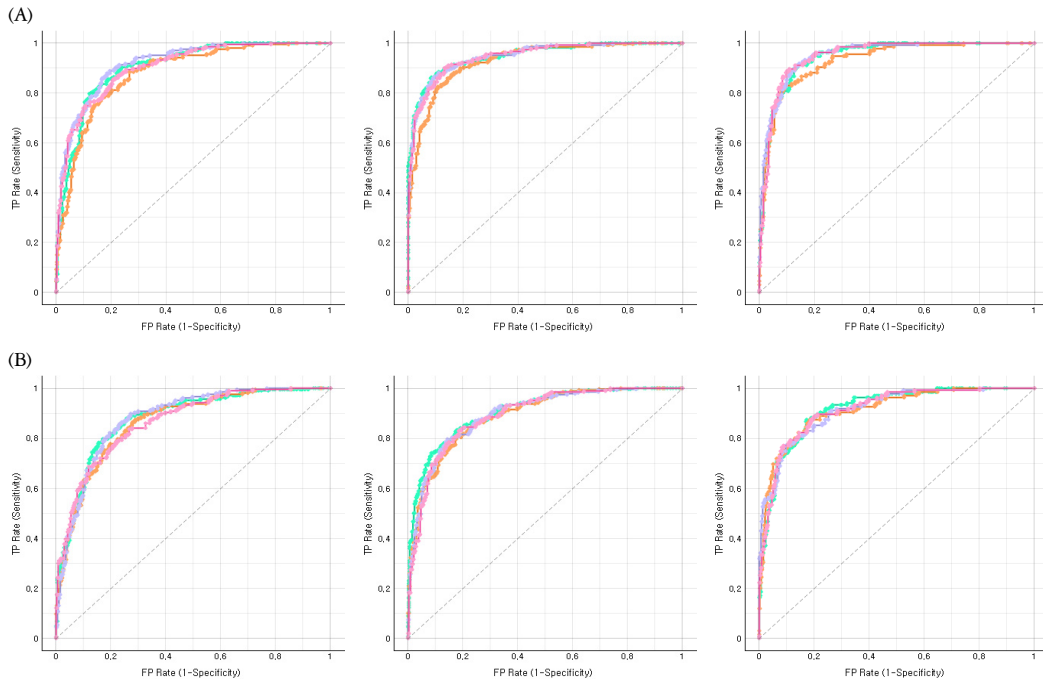


Figure 5. ROC curves of diagnostic models using default dataset (A) and comparison dataset (B). The left, middle, and right panels correspond to So-Yang, Tae-Eum, and So-Eum types, respectively.

For all panels, purple line denotes Neural Network (NN), green line Random Forest (RF), pink line Stochastic Gradient Descent (SGD), and orange line Naive Bayes (NB)

경우 BMI에 과도하게 의존하면서 가장 높은 RF를 비롯해 다양한 체질진단 알고리즘들이 전반적으로 높은 성능을 보였으나, 이에 비해 (BMI가 제외된) 비교 데이터세트를 활용하였을 경우에는 다양한 신체적 입력특성(RMRw, RMR, 목둘레)이 고르게 기여하였다.

이러한 결과를 통해 사상체질 진단에 사용된 심리적, 신체적, 혈액학적 임상특성들의 기여도(IG 값)를 선행 임상연구 결과들과 비교하는 과정에서 다음과 같은 지견을 재확인할 수 있었다.

첫째, 신체적 특성이 체질 진단에 가장 큰 기여도(IG 값)를 보였으며, BMI가 체질 진단에 과도한 영향력을 행사한 것을 재확인할 수 있었다.

신체적 특성의 IG 값(BMI 포함시 0.951, BMI 배제시 0.858)이, 심리적 특성(0.329)이나 혈액학적 특성(0.035)보다 매우 컸다. 또한, BMI(0.531) 단독으로 신

체적 특성의 절반 이상을 차지하며, 심리적(0.329) 및 혈액학적(0.035) 특성을 모두 합친 것을 넘는 영향력을 가졌다. 이러한 결과는 BMI가 사상체질 진단의 핵심적 특성^{10,15,22}이라는 선행 연구를 강조하기보다는 과대평가²⁴라는 주장을 지지하는 것으로, KS-15에 있어 BMI의 영향력을 축소할 필요가 있음을 제시하는 것으로 보인다.

이에 반해서, (BMI를 배제한) 비교 데이터세트에 있어서는 RMRw(0.292), RMR(0.207), 목둘레(0.146), WHR(0.116)와 같은 다양한 신체특성이 균형적으로 기여하는 것을 확인할 수 있었는데, 이는 PI^{10,22}, WHR⁴⁷, RMRw²⁵, 늑골각(Costal angle)²⁸, 목둘레⁴⁸, 가슴둘레⁴⁸, 허리둘레⁴⁹와 같은 신체적 특성이 중요하다고 제시된 선행 연구들을 포괄하는 것이며, 임상 진단에 있어 기초대사량을 측정하는 RMRw/RMR²⁵의 활

용이 임상적으로 높은 유용성을 지님을 시사하는 것으로 보인다.

둘째, 심리적 특성이 신체적 특성보다는 전반적인 중요도가 낮았음에도, 개인의 행동특성을 의미하는 SPQ-B(0.206)가 심리적 특성(0.329)의 절반 이상을 차지함과 동시에 RMRw(0.292), RMR(0.243)와 같은 신체적 특성들과 유사한 정도의 기여도를 보였다.

사상체질 진단에 있어서 신체적 특징인 체형기상(體形氣象) 뿐만 아니라 심리적인 용모사기(容貌詞氣)와 성질재간(性質材幹) 그리고 병증약리(病證藥理)와 소증(素證)을 종합적으로 고려해야 한다는 것을 고려한다면, 심리적 특성으로서 SPQ의 3가지 하위 척도(0.39)^{13,14,22}가 사상체질 진단에 있어 이론적, 임상적 중요성을 동시에 지니고 있음을 재확인한 것이다.

셋째, 혈액학적 특성에서는 세 체질 간 유의한 차이를 보인 RBC(0.035)만이 입력특성으로 포함되었으나, 그 기여도가 낮아 다른 신체적, 심리적 특성에 비해 사상체질 진단에 있어 임상적 활용이 제한적임을 확인할 수 있었다.

다만, 흥미로운 사실은, 신체 에너지대사 및 혈관건강에 중요한 지질인 Tg와 HDL는 대사증후군, 당뇨, 비만, 심혈관에서의 질환과의 관련성이 높는데, 건강한 대학생(평균22.59세)을 대상으로 QSCC2를 사용한 선행연구⁵⁰에서는 세 체질간 유의한 차이가 보고되지 않았었다는 점이다. 이에, 체지방량과 심혈관질환과 관련된 BMI²⁴의 영향력이 크게 설정된 KS-15에 있어서 이러한 지표들의 기여도가 축소되었을 것으로 추정할 수 있을 것이며, 임상군을 대상으로 사상체질의 학 전문의의 임상진단 또는 QSCC와 같은 다른 사상체질 진단법을 활용하여 소증 및 질환에 따른 체질별 혈액학적 특성의 영향력을 재확인해볼 필요가 있을 것이다.

넷째, 본 연구에서는, 선행연구를 통한 이론기반 접근과 IG 및 ML을 활용한 데이터기반 접근을 동시에 사용하여 사상체질 그룹간 현저한 차이를 지닌 입력 데이터를 추출함으로써 보다 높은 분류 정확도를 얻을 수 있었다.

본 연구에서 확인된 분류 정확도는 기본 데이터세트(0.817)와 비교 데이터세트(0.755) 모두에서 높은 수치를 보였는데, 이는 대규모(n=3,891) 임상 데이터를 활용한 선행 연구²⁸에서의 분류 정확도(0.604±0.060)를 상회하는 것이다.

이상에서 살펴본 사상체질 진단에 대한 심리적, 신체적, 혈액학적 임상특징의 영향의 고찰에 더해져, 사상체질 진단에 활용되었던 알고리즘과 입력특성 간 상호작용에 대해서는 다음과 같은 지견을 확인할 수 있었다.

첫째, 사상체질 진단을 위한 입력특성 선택에서 고른 기여도(IG 값)를 갖는 여러 개의 임상특성을 사용하는 것이, 큰 영향력을 갖는 소수의 임상특성을 사용하는 것보다 우수한 재현 가능성(reproducibility)을 갖고 있었다.

타당도 지표들(Table 4와 Table 5)에 있어서 MCC를 기준으로 최소값을 보인 알고리즘과 최대값을 보인 알고리즘 간의 차이를 비교하였을 때, 그 차이가 작은 것은 예측일관성(prediction consistency)이 높음을 의미한다. 기본 데이터세트에 있어서 RF(0.717)과 NB(0.663)의 MCC값의 차이는 0.054였으나, 비교 데이터세트에 있어서 RF(0.618)와 NB(0.598)의 MCC 값의 차이는 0.020에 불과했으므로, 고른 영향력을 가진 여러 임상특성을 사용하는 비교 데이터세트가 일반화가능성(generalizability)에서 우수하다고 할 수 있다.

둘째, 동일한 알고리즘이라도 사용되는 입력특성의 구성에 따라 타당도 지표가 상이함을 알 수 있었는데, 본 연구에서는 BMI라는 한가지 입력특성의 유무(기본 및 비교 데이터세트)만으로도 큰 차이를 보였다(Table 4와 Table 5).

본 연구에서, 기본 및 비교 데이터세트 간에 NN 알고리즘의 MCC(0.091)는 큰 차이를 보였지만, 이에 비해서 NB 알고리즘의 분류 정확도, F1 점수, 정밀도, 재현율은 상대적으로 작았다(0.041~0.042). 이러한 결과는, 고차원 데이터의 비선형적 패턴에 효과적인 NN(42)보다, 입력특성이 독립적일 것으로 가정하는 NB(41)가 사상체질 그룹의 진단에 보다 안정적인 결과

를 도출하는 것으로 사료된다.

셋째, 알고리즘간 성능의 확연한 차이는 없었지만, 미약한 차이를 확인할 수 있었다.

본 연구에 있어서, 기본 데이터셋을 사용한 경우에는 타당도 지표가 RF에서 가장 높고, NB는 가장 낮은 것으로 나타났으며(Table 4), 비교 데이터셋을 사용한 경우에도 동일한 결과를 확인할 수 있었다(Table 5). 이러한 연구 결과는, 선행연구에서 사상체질 진단에 가장 적합한 모델로 NB²⁷를 제안하고, 최저의 성능으로 RF³¹를 지목하였던 것과는 차이를 보인다. 또한, ROC 곡선과 AUC 값을 비교했을 때, 알고리즘간 차이는 기본 데이터셋에서 최대 0.024, 비교 데이터셋에서는 0.011로 매우 제한적이었다(Figure 5).

이러한 결과는 각 알고리즘이 전반적으로 유사한 수준의 성능을 보임을 의미하며, 따라서 사상체질 진단 모델의 성능은 특정 알고리즘의 우월성보다는 입력데이터의 선택과 구성에 의해 더 크게 좌우된다고 볼 수 있을 것이다.

본 연구는 기존의 이론기반 분석과 데이터 분석의 강점들을 결합하여 결과 해석의 어려움과 예측력의 한계를 동시에 보완하였다는 장점을 지니며, 과도한 영향력을 지닌 BMI보다는 의미 있는 영향력을 지닌 RMR, RMRw, 목둘레 및 SPQ라는 신체적, 심리적 임상 특성들을 균형적으로 고려하는 것이 임상에서 보다 높은 분류 정확도와 일반화가능성을 담보한다는 것을 데이터를 통해 확인할 수 있었다.

또한, 진단 알고리즘의 타당도는 입력 데이터의 특성에 따라 상반된 결과를 보일 수 있기에, 최선의 사상체질 진단 알고리즘을 찾는 것보다, 적절한 IG 값을 지니는 다양한 임상 특성들을 입력하는 것이 최선의 사상체질 진단모델이라는 것을 확인할 수 있었다.

사상의학 연구에 있어서 ML이나 인공지능, 빅데이터의 사용이 증가하고 있으나, 결과 해석의 어려움과 이론적 타당성의 부족, 낮은 재현가능성과 임상적 실용성 및 분석 방법에 대한 낮은 접근성 등이 여전히 중요한 한계로 지적되고 있다.

본 연구는 이론적인 고찰을 포함하는 선행연구와

함께 높은 접근성을 지닌 오픈소스 ML 패키지인 Orange를 사용한 분석을 통합하는 연구설계를 통해 이러한 한계를 극복하고자, 함께 제시된 본 연구의 전체적인 분석절차(Figure 1)와 Orange 워크플로우(Figure 2, 3)를 통해 동일한 ML 분석을 많은 사상의학 연구자와 임상인들이 그대로 재현할 수 있는 기반을 마련하였다.

제시된 Orange 기반의 시각적 워크플로우와 분석절차는 연구자 뿐 만 아니라 임상가와 학생들에게도 데이터 분석 과정을 이해하고 실제 적용하는 데 일정한 도움을 줄 수 있을 것으로 보인다. 이러한 점은 한의학적 데이터 리터러시를 높이고, 사상체질 관련 연구·진단·교육 영역에서 활용 가능성을 확장하는 데 기여할 수 있을 것으로 보인다³².

다만, 본 연구의 결과를 일반화하기에는 다음과 같은 한계성을 지니고 있을 수 있으므로, 이를 보완하기 위한 추가적인 후속 연구가 필요할 것이다.

첫째, 본 연구에서는 사상체질 진단에 있어 BMI를 첫 문항으로 사용하는 KS-15를 사용하였으므로, KS-15에 과적합된 사상체질의 진단으로 BMI 등의 과도한 영향^{24,31}과 순환 논증 편향의 위험성을 지니고 있다. 선행연구를 통해 본 연구의 BMI^{10,15,22}, PI^{10,22}, RMR과 RMRw²⁵, 목둘레⁴⁸ 및 SPQ^{13,14,22}의 중요성이 재확인되어 왔으니², 임상 전문가나 체질처방에 대한 반응, QSCC 등과 같은 다양한 사상체질 진단과 추가적인 독립 코호트를 대상으로 한 추가연구를 통해 본 연구 결과가 재확인되어야 할 것이다.

둘째, 사상체질 진단모델의 학습과정에서 심리적, 신체적, 혈액학적 특성들을 동일한 가중치로 반영하였기에, 성정(性情)으로 장부의 편차와 소증(素證)이 나타난다는 이론²³에 따른 심리적 특성 등의 중요도는 본 연구에서 고려하지 못하였다. 따라서, 후속연구에서는 이러한 입력변수의 가중치를 조절하는 방안도 고려해 볼 필요가 있을 것이며 이들 간의 복합적인 네트워크 분석이 필요할 것이다.

셋째, 분석과정에 한국인만을 대상으로 하였기에, 사상의학의 세계화를 위해서는 다양한 인종을 포함하

는 연구⁵³⁾가 요구될 수도 있다. 이러한 후속 다민족, 다문화 연구를 통해 사상체질 그룹별 심리적, 신체적, 혈액학적 차이에 대한 일반성과 타당도를 높일 수 있을 것이다.

V. Conclusion

본 연구에서는, 심리적, 신체적, 혈액학적 임상특성을 체계적으로 추출하여 기계학습 사상체질 진단모델에 대한 기여도를 분석하고, 사상체질 임상진단에 가장 타당한 알고리즘을 찾기 위한 다면적 분석을 시행하였다. 분석 결과, 최선의 사상체질 알고리즘을 찾는 것보다 균형 있는 영향력(IG 값)을 지닌 RMR, RMRw, 목둘레 및 SPQ 등과 같은 신체적, 심리적 임상특성을 동시에 사용할 때, 임상에서 보다 높은 사상체질 분류 정확도와 일반화가능성을 확보할 수 있다는 것을 결과로 확인할 수 있었다. 이러한 접근을 통해 사상의학 임상진단의 효율성과 일반화 가능성이 증대되어, 사상체질 진단의 표준화와 국제적 활용 가능성을 높이는 토대가 마련될 수 있을 것이다.

VI. Acknowledgement

이 연구는 한국한의학연구원 기본사업인 빅데이터 기반 한의 예방 치료 원천기술 개발의 지원을 받아 수행된 연구입니다(Grant No. KSN1739121).

VII. References

1. Lee SJ, Park SH, Cloninger CR, Kim YH, Hwang MW, Chae H. Biopsychological traits of Sasang typology based on Sasang Personality Questionnaire and body mass index. *BMC Complement Altern Med.* 2014;14:315. DOI: 10.1186/1472-6882-14-315
2. Chae H. The art of longevity and wellbeing. Busan: CALMTREE; 2025.
3. Chae H, Lyoo IK, Lee SJ, Cho S, Bae H, Hong M, et al. An alternative way to individualized medicine: psychological and physical traits of Sasang typology. *J Altern Complement Med.* 2003;9(4):519-28. DOI: 10.1089/107555303322284811
4. Chae H, Park SH, Lee SJ, Kim MG, Wedding D, Kwon YK. Psychological profile of Sasang typology: a systematic review. *Evid Based Complement Alternat Med.* 2009;6:21-9. DOI: 10.1093/ecam/nep079
5. Pailhez G, Bulbena A. Body shape and psychiatric diagnosis revisited. *Int J Psychiatry Clin Pract.* 2010;14(4):236-43. DOI: 10.3109/13651501.2010.505344
6. Campbell CM. Review of Physique and character: an investigation of the nature of constitution and of the theory of temperament. *J Abnorm Soc Psychol.* 1926;21(1):107-8.
7. Wile IS. Review of The varieties of human physique: An introduction to constitutional psychology [book review]. *American Journal of Orthopsychiatry.* 1941;11(1):182. DOI: 10.1037/h0097360
8. Veith I. The Yellow Emperor's classic of internal medicine. Baltimore: Williams & Wilkins; 1949.
9. Sumantran VN, Tillu G. Insights on personalized medicine from Ayurveda. *J Altern Complement Med.* 2013;19(4):370-5. DOI: 10.1089/acm.2011.0698
10. Chae H, Kwon YK. Best-fit index for describing physical perspectives in Sasang typology. *Integr Med Res.* 2015;4(1):20-8. DOI: 10.1016/j.imr.2014.11.001
11. Chae H, Lee JY, Jeon ES, Kim JK. Personalized

- acupuncture treatment with Sasang typology. *Integr Med Res.* 2017;6(4):329-36. DOI: 10.1016/j.imr.2017.07.002
12. Park SY, Kim YW, Song YR, Bak SB, Jang YP, Kim IK, et al. Compound-level identification of Sasang constitution type-specific personalized herbal medicine using a data science approach. *Heliyon.* 2023;9(2):e13692. DOI: 10.1016/j.heliyon.2023.e13692
 13. Han JH, Lee HS, Lee J, Lee SJ, Chae H. Systematic review of Sasang typology studies from a psychological perspective. *J Sasang Constitut Med.* 2019;31(1):36-63. DOI: 10.7730/JSCM.2019.31.1.36
 14. Chae H, Lee JY, Lee YJ. Study on the validation of Sasang Digestive Function Inventory (SDFI) for analyzing pathophysiological digestive symptoms and diagnosing Sasang types. *J Sasang Constitut Med.* 2021;33(3):1-5. DOI: 10.7730/JSCM.2021.33.3.1
 15. Lee MS, Sohn KW, Kim YH, Hwang MW, Kwon YK, Bae NY, et al. Digestive system-related pathophysiological symptoms of Sasang typology: systematic review. *Integr Med Res.* 2013;2(2):39-48. DOI: 10.1016/j.imr.2013.04.001
 16. Song KH, Yu SG, Kim JY. Prevalence of metabolic syndrome according to Sasang constitutional medicine in Korean subjects. *Evid Based Complement Alternat Med.* 2012;2012:646794. DOI: 10.1155/2012/646794
 17. Kim YY, Kim TY, Park SY, Jang ES. Sasang constitution may act as a risk factor for post-traumatic stress disorder. *J Orient Neuropsychiatry.* 2020;32(1):22-9. DOI: 10.7730/JSCM.2020.32.1.22
 18. Bae NY, Lee EJ. Clinical practice guideline for Sasang constitutional medicine: prevention and risk factors of Sasang in disease. *J Orient Neuropsychiatry.* 2015;27(1):82-109. DOI: 10.7730/JSCM.2015.27.1.082
 19. Chae H, Lee SW, Park SH, Jang ES, Lee SJ. Development and validation of a personality assessment instrument for traditional Korean medicine: Sasang Personality Questionnaire. *Evid Based Complement Alternat Med.* 2012;2012:657013. DOI: 10.1155/2012/657013
 20. Chae H. Comment on "Manifestations of Sasang typology according to common chronic diseases in Koreans." *Evid Based Complement Alternat Med.* 2020;2020:8706183. DOI: 10.1155/2020/8706183
 21. Lee SK, Yoon DW, Kim JY, Kim JK, Yi HR, Lee SH, et al. Association of Sasang constitutional type with bone mineral density, osteopenia, and osteoporosis. *J Orient Neuropsychiatry.* 2020;32(3):33-45. DOI: 10.7730/JSCM.2020.32.3.33
 22. Chae H, Lee SW, Lee SJ. Can data-driven analysis demonstrate the plausibility of traditional medical typology?. *J Orient Neuropsychiatry.* 2021;32:303-18. DOI: 10.7231/jon.2021.32.4.303
 23. Chae H, Cho YI, Lee SJ. The Yin-Yang personality from a biopsychological perspective using the revised Sasang Personality Questionnaire. *Integr Med Res.* 2021;10(1):100455. DOI: 10.1016/j.imr.2020.100455
 24. Kim SH. Reconsideration of body mass index for diagnosing Sasang constitution. *Korean J Orient Physiol Pathol.* 2024;38(4):81-8. DOI: 10.15188/kjopp.2024.08.38.4.81
 25. Park JE, Lee S, Lee YJ, Lee JY, Chae H. Resting metabolic rate for diagnosing Tae-Eum Sasang type and unraveling the mechanism of type-specific obesity. *Diagnostics.* 2023;13(4):672. DOI: 10.3390/diagnostics13040672
 26. Baek YH, Lee SW, Jeong KS, Jang ES. Sasang

- constitution type combined with general obesity may act as a risk factor for prediabetes mellitus. *Healthcare (Basel)*. 2022;10(11):2286. DOI: 10.3390/healthcare10112286
27. Hong JW, Kim YI, Kim BC, Eom IK, Hwang MW, Shin SW, et al. Data mining algorithms for the development of Sasang type diagnosis. *JPPKM*. 2009;23(6):1234-40.
 28. Park SY, Park MS, Lee WY, Lee CY, Kim JH, Lee SW, et al. Machine learning-based prediction of Sasang constitution types using comprehensive clinical information and identification of key features for diagnosis. *Integr Med Res*. 2021;10(3):100668. DOI: 10.1016/j.imr.2020.100668
 29. Lim HJ. A step-by-step guide to random forest model using Orange data mining in the field of periodontitis. *J Korean Acad Oral Health*. 2021;45(4):218-26. DOI: 10.11149/jkaoh.2021.45.4.218
 30. Quinlan JR. Induction of decision trees. *Mach Learn*. 1986;1(1):81-106. DOI: 10.1007/BF00116251
 31. Baek YH, Jang ES, Park KH, Yoo JH, Jin HJ, Lee SW. Development and validation of brief KS-15 (Korea Sasang constitutional diagnostic questionnaire) based on body shape, temperament and symptoms. *J Sasang Constitut Med*. 2015;27(2):211-21. (Korean) DOI: 10.7730/JSCM.2015.27.2
 32. Lee SJ, Park SH, Chae H. Biopsychological structure of Yin - Yang using Cloninger's temperament model and Carver and White's BIS/BAS scale. *PeerJ*. 2016;4:e2021. DOI: 10.7717/peerj.2021
 33. Costa PT Jr, McCrae RR. Personality disorders and the five-factor model of personality. *J Pers Disord*. 1990;4(4):362-71. DOI: 10.1521/pedi.1990.4.4.362
 34. Min BM, Lee KI, Jeong JC. NEO-PI-R. Seoul: PSI Consulting; 1997.
 35. Shetty PS, James W. Body mass index: a measure of chronic energy deficiency in adults. *FAO Food Nutr Pap*. 1994;56:1-57.
 36. Seltzer CC. Some re-evaluations of build and blood pressure study 1959 as related to ponderal index, somatotype, and mortality. *N Engl J Med*. 1966;274(5):254-9. DOI: 10.1056/NEJM196602032740505
 37. Neovius M, Linne Y, Rossner S. BMI, waist circumference and waist - hip ratio as diagnostic tests for fatness in adolescents. *Int J Obes*. 2005;29(2):163-9. DOI: 10.1038/sj.ijo.0802867
 38. Joseph VR, Vakayil A. SPlit: an optimal method for data splitting. *Technometrics*. 2022;64(2):166-76. DOI: 10.48550/arXiv.2012.10945
 39. Prasetyowati MI, Maulidevi NU, Surendro K. Determining threshold value on information gain feature selection to increase speed and prediction accuracy of random forest. *J Big Data*. 2021;8(1):84. DOI: 10.1186/s40537-021-00472-4
 40. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32. DOI: 10.1023/A:1010933404324
 41. Zhang H. The optimality of Naive Bayes. In: *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (FLAIRS 2004)*. 2004. p. 562-7.
 42. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323(6088):533-6. DOI:10.1038/323533a0
 43. Robbins H, Monro S. A stochastic approximation method. *Ann Math Stat*. 1951;22(3):400-7.
 44. Sokolova M, Lalpalme G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag*. 2009;45(4):427-37. DOI: 10.1016/

- j.ipm.2009.03.002
45. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In: AAAI Workshop - Technical Report. 2006. p. 24-9.
 46. Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim Biophys Acta*. 1975;405(2):442-51. DOI: 10.1016/0005-2795(75)90109-9
 47. Kim MJ, Lee DH, Ahn JY, Ha TY, Jang YJ, Do EJ, et al. A pilot study on characteristics of metabolomics and lipidomics according to Sasang constitution. *Evid Based Complement Alternat Med*. 2018;2018:9214960. DOI: 10.1155/2018/9214960
 48. Lee SJ, Park SH, Ko YS, Park SJ, Eom IK, Kim BC, et al. Analysis on physical traits of Sasang types using bioelectrical impedance analysis. *Korean J Orient Physiol Pathol*. 2009;23(2):433-7. (Korean)
 49. Jang ES, Baek YH, Park KH, Lee SW. Could the Sasang constitution itself be a risk factor of abdominal obesity? *BMC Complement Altern Med*. 2013;13(1):72. DOI: 10.1186/1472-6882-13-72
 50. Jeong MK, Youn SJ, Jun CY, Park JH, Choi YK. Evaluation of the relationship between the results of blood test and Sasang constitution. *Korean J Orient Physiol Pathol*. 2012;26(6):964-9. (Korean)
 51. Kim JS, Park SH, Jeong R, Lee ES, Kim YS, Sung HD, et al. Application of text-mining technique and machine-learning model with clinical text data obtained from case reports for Sasang constitution diagnosis: a feasibility study. *J Korean Med*. 2024;45(3):193-210. (Korean) DOI: 10.13048/jkm.24049
 52. Chae H. Data literacy and Yin - Yang (Eum - Yang) in Traditional Korean Medicine education. *J Korean Med Educ*. 2025;3(2):17-24. DOI: 10.23215/JKME.PUB.3.2.17
 53. Lee JY, Yim MH, Ang L. Facial features analysis of Sasang typology using facial images: a study on Caucasian American and native Korean populations. *Eur J Integr Med*. 2023;57:102210. DOI: 10.1016/j.eujim.2022.102210
- <저자들의 소속 및 직위>
 황동욱/부산대학교 한의학전문대학원/석사학생
 장동엽/동의대학교 한의과대학/조교수
 김호석/한국한의학연구원 한의약데이터부/
 기술연구원
 정경식/한국한의학연구원 한의약데이터부/
 기술연구원
 채 한/부산대학교 한의학전문대학원/교수

Suppl. Table 1. Correlation Coefficient among Selected Psychological Features

	SPQ-Total	SPQ-Behavior	SPQ-Cognition	SPQ-Emotion	Extraversion
SPQ-Total	—				
SPQ-Behavior	0.837 ^{***}	—			
SPQ-Cognition	0.813 ^{***}	0.553 ^{***}	—		
SPQ-Emotion	0.669 ^{***}	0.331 ^{***}	0.310 ^{***}	—	
Extraversion	0.481 ^{***}	0.608 ^{***}	0.341 ^{***}	0.124 ^{***}	—

Sasang Personality Questionnaire, SPQ
 *, $p < .05$; **, $p < .01$; ***, $p < .001$

Suppl. Table 3. Confusion Matrix of Sasang Type Classification Models Using Default Dataset

			Predicted label		
			TE	SE	SY
Actual label	TE	RF	85.9%	4.8%	9.3%
		NB	83.7%	4.1%	12.2%
		NN	84.1%	5.2%	10.7%
		SGD	85.9%	4.8%	9.3%
	SE	RF	6.7%	79.3%	14.1%
		NB	5.9%	79.3%	14.8%
		NN	5.2%	80.7%	14.1%
		SGD	8.9%	76.3%	14.8%
	SY	RF	11.6%	11.1%	77.3%
		NB	17.4%	12.1%	70.5%
		NN	11.1%	14.0%	74.9%
		SGD	15.9%	7.7%	76.3%

Bold represent the best performances in the categories.

Random Forest, RF; Naïve Bayes, NB; Neural Network, NN; Stochastic Gradient Descent, SGD

Suppl. Table 4. Confusion Matrix of Sasang Type Classification Models Using Comparison Dataset

			Predicted label		
			TE	SE	SY
Actual label	TE	RF	83.0%	6.3%	10.7%
		NB	82.2%	6.3%	11.5%
		NN	80.0%	7.8%	12.2%
		SGD	85.2%	5.6%	9.3%
	SE	RF	14.8%	69.6%	15.6%
		NB	13.3%	74.1%	12.6%
		NN	12.6%	74.8%	12.6%
		SGD	16.3%	67.4%	16.3%
	SY	RF	20.8%	11.6%	67.6%
		NB	25.1%	11.1%	63.8%
		NN	17.9%	13.5%	68.6%
		SGD	25.6%	7.7%	66.7%

Bold represent the best performances in the categories.

Random Forest, RF; Naïve Bayes, NB; Neural Network, NN; Stochastic Gradient Descent, SGD

Suppl. Table 2. Correlation Coefficient among Selected Physical Features

	BMI	PI	WHR	RMR	RMRw	Forehead C	Neck C	Axillary C	Chest C	Rib C	Waist C	Pelvic C	Hip C	Axillary W	Chest W	Rib W	Waist W	Pelvic W	
BMI	—																		
PI	0.930 ^{***}	—																	
WHR	0.544 ^{***}	0.513 ^{***}	—																
RMR	0.541 ^{***}	0.228 ^{***}	0.291 ^{***}	—															
RMRw	-0.728 ^{***}	-0.721 ^{***}	-0.384 ^{***}	-0.061 ^{***}	—														
Forehead C	0.341 ^{***}	0.138 ^{***}	0.172 ^{***}	0.595 ^{***}	-0.126 ^{***}	—													
Neck C	0.643 ^{***}	0.408 ^{***}	0.408 ^{***}	0.829 ^{***}	-0.255 ^{***}	0.578 ^{***}	—												
Axillary C	0.733 ^{***}	0.512 ^{***}	0.454 ^{***}	0.806 ^{***}	-0.379 ^{***}	0.584 ^{***}	0.824 ^{***}	—											
Chest C	0.814 ^{***}	0.672 ^{***}	0.511 ^{***}	0.616 ^{***}	-0.567 ^{***}	0.479 ^{***}	0.702 ^{***}	0.898 ^{***}	—										
Rib C	0.825 ^{***}	0.652 ^{***}	0.590 ^{***}	0.715 ^{***}	-0.515 ^{***}	0.455 ^{***}	0.782 ^{***}	0.855 ^{***}	0.867 ^{***}	—									
Waist C	0.814 ^{***}	0.700 ^{***}	0.773 ^{***}	0.530 ^{***}	-0.624 ^{***}	0.409 ^{***}	0.621 ^{***}	0.768 ^{***}	0.843 ^{***}	0.829 ^{***}	—								
Pelvic C	0.743 ^{***}	0.645 ^{***}	0.458 ^{***}	0.432 ^{***}	-0.638 ^{***}	0.373 ^{***}	0.520 ^{***}	0.688 ^{***}	0.790 ^{***}	0.730 ^{***}	0.843 ^{***}	—							
Hip C	0.739 ^{***}	0.594 ^{***}	0.265 ^{***}	0.537 ^{***}	-0.603 ^{***}	0.466 ^{***}	0.572 ^{***}	0.754 ^{***}	0.815 ^{***}	0.719 ^{***}	0.815 ^{***}	0.868 ^{***}	—						
Axillary W	0.657 ^{***}	0.445 ^{***}	0.437 ^{***}	0.758 ^{***}	-0.365 ^{***}	0.521 ^{***}	0.738 ^{***}	0.826 ^{***}	0.742 ^{***}	0.752 ^{***}	0.681 ^{***}	0.569 ^{***}	0.640 ^{***}	—					
Chest W	0.747 ^{***}	0.618 ^{***}	0.506 ^{***}	0.552 ^{***}	-0.547 ^{***}	0.322 ^{***}	0.603 ^{***}	0.661 ^{***}	0.688 ^{***}	0.723 ^{***}	0.659 ^{***}	0.545 ^{***}	0.536 ^{***}	0.734 ^{***}	—				
Rib W	0.749 ^{***}	0.573 ^{***}	0.507 ^{***}	0.684 ^{***}	-0.470 ^{***}	0.403 ^{***}	0.730 ^{***}	0.751 ^{***}	0.728 ^{***}	0.835 ^{***}	0.685 ^{***}	0.586 ^{***}	0.576 ^{***}	0.741 ^{***}	0.804 ^{***}	—			
Waist W	0.781 ^{***}	0.688 ^{***}	0.622 ^{***}	0.469 ^{***}	-0.621 ^{***}	0.330 ^{***}	0.561 ^{***}	0.670 ^{***}	0.746 ^{***}	0.763 ^{***}	0.818 ^{***}	0.708 ^{***}	0.677 ^{***}	0.657 ^{***}	0.738 ^{***}	0.773 ^{***}	—		
Pelvic W	0.429 ^{***}	0.382 ^{***}	0.235 ^{***}	0.199 ^{***}	-0.431 ^{***}	0.287 ^{***}	0.281 ^{***}	0.495 ^{***}	0.615 ^{***}	0.426 ^{***}	0.602 ^{***}	0.695 ^{***}	0.709 ^{***}	0.373 ^{***}	0.257 ^{***}	0.301 ^{***}	0.528 ^{***}	—	

Circumference, C; Width, W
 *, $P < .05$; **, $P < .01$; ***, $P < .001$