

## Development of an intelligent skin condition diagnosis information system based on social media

Hyung-Hoon Kim\*, Seung-Ho Ohk\*\*

\*Professor, Dept. of Cosmetic Science, Kwangju Womens University, Gwangju, Korea

\*\*Professor, School of Dentistry, Chonnam National University, Gwangju, Korea

### [Abstract]

Diagnosis and management of customer's skin condition is an important essential function in the cosmetics and beauty industry. As the social media environment spreads and generalizes to all fields of society, the interaction of questions and answers to various and delicate concerns and requirements regarding the diagnosis and management of skin conditions is being actively dealt with in the social media community. However, since social media information is very diverse and atypical big data, an intelligent skin condition diagnosis system that combines appropriate skin condition information analysis and artificial intelligence technology is necessary.

In this paper, we developed the skin condition diagnosis system SCDIS to intelligently diagnose and manage the skin condition of customers by processing the text analysis information of social media into learning data. In SCDIS, an artificial neural network model, AnnTFIDF, that automatically diagnoses skin condition types using artificial neural network technology, a deep learning machine learning method, was built up and used. The performance of the artificial neural network model AnnTFIDF was analyzed using test sample data, and the accuracy of the skin condition type diagnosis prediction value showed a high performance of about 95%. Through the experimental and performance analysis results of this paper, SCDIS can be evaluated as an intelligent tool that can be used efficiently in the skin condition analysis and diagnosis management process in the cosmetic and beauty industry. And this study can be used as a basic research to solve the new technology trend, customized cosmetics manufacturing and consumer-oriented beauty industry technology demand.

▶ **Key words:** Skin condition diagnosis, beauty industry, cosmetics industry, artificial intelligence convergence, big data convergence

---

• First Author: Hyung-Hoon Kim, Co-Author: Seung-Ho Ohk, Corresponding Author: Hyung-Hoon Kim  
\*Hyung-Hoon Kim (hhkim@kwu.ac.kr), Dept. of Cosmetic Science, Kwangju Womens University  
\*\*Seung-Ho Ohk (shohk@jnu.ac.kr), School of Dentistry, Chonnam National University  
• Received: 2022. 07. 08, Revised: 2022. 08. 19, Accepted: 2022. 08. 26.

## [요 약]

화장품 및 뷰티산업에서 고객의 피부상태 진단과 관리는 중요한 필수기능이다. 소셜미디어 환경이 사회 전 분야에 확산되고 일반화되면서 피부 상태의 진단과 관리에 대한 다양하고 섬세한 고민과 요구 사항의 질문과 답변의 상호작용이 소셜미디어 커뮤니티에서 활발하게 다루어지고 있다. 그러나 소셜미디어 정보는 매우 다양하고 비정형적인 방대한 빅데이터이므로 적절한 피부상태 정보분석과 인공지능 기술을 접목한 지능화된 피부상태 진단 시스템이 필요하다.

본 논문에서는 소셜미디어의 텍스트 분석정보를 학습데이터로 가공하여 고객의 피부상태를 지능적으로 진단 및 관리하기 위한 피부상태진단시스템 SCDIS를 개발하였다. SCDIS에서는 딥러닝 기계학습 방법인 인공신경망 기술을 사용하여 자동적으로 피부상태 유형을 진단하는 인공신경망 모델 AnnTFIDF을 빌드업하여 사용하였다. 인공신경망 모델 AnnTFIDF의 성능은 테스트샘플 데이터를 사용하여 분석되었으며, 피부상태 유형 진단 예측 값의 정확성은 약 95%의 높은 성능을 나타내었다. 본 논문의 실험 및 성능분석결과를 통하여 SCDIS는 화장품 및 뷰티산업 분야의 피부상태 분석 및 진단 관리 과정에서 효율적으로 사용 가능한 지능화된 도구로 평가할 수 있다. 본 논문에서 제안된 시스템은 소셜미디어 기반의 새로운 환경에서 화장품 및 피부미용에 대한 사용자의 요구를 체계적으로 파악하고 진단하는 기초 기술로 사용 가능할 것이다. 그리고 이 연구는 새로운 기술 트렌드인 맞춤형 화장품제조와 소비자중심의 뷰티산업기술 수요를 해결하기 위한 기초 연구로 사용될 수 있을 것이다.

▶ **주제어:** 피부 상태, 뷰티산업, 화장품 산업, 인공지능 융합, 빅데이터 융합

## I. Introduction

최근 화장품산업과 뷰티산업의 동향은 기능과 성분으로 중시하며 기능성 화장품과 개인 맞춤형 뷰티산업 방향으로 발전되어 가고 있다. 화장품과 뷰티산업에서 개인 맞춤형 기능성 화장품과 미용 서비스를 제공하기 위해서는 무엇보다 화장품 및 미용 소비자의 니즈를 파악하는 것이 가장 중요하다. 뷰티산업 분야에서 소비자의 니즈 파악 과정에서 가장 기본적인 내용은 소비자의 피부 상태 파악과 요구 사항 파악이라 할 수 있다. 이러한 관점에서 화장품과 학 및 미용과학 분야의 뷰티산업 종사자와 연구개발자에게 피부 상태의 파악과 진단은 중요한 관심사이다. 일반적으로 뷰티산업 종사자들이 피부 미용을 목적으로 피부 상태를 진단하는 경우에 대부분 육안 관찰 또는 스마트폰 등의 보편적으로 보급된 관찰 장비를 사용하여 관찰하고, 피부미용을 원하는 고객의 피부 상태에 대한 의견을 고려하여 피부 미용적 진단과 관리를 진행하고 있다[1, 2].

4차 산업혁명시대의 도래와 함께 화장품기술 환경과 뷰티산업 전반의 환경에 큰 변혁이 일어나고 있다. 화장품과 뷰티산업분야의 많은 기능과 역할이 온라인 비대면 환경으로 급속하게 발전하고 있다. 화장품 및 뷰티산업분야에 대한 고객의 요구 수준은 더욱더 높아지고 있으며 이에 대

한 다양하고 섬세한 고민과 요구 사항은 소셜미디어 커뮤니티와 같은 비대면 온라인 환경에서 더욱더 활발하게 다루어지고 있다. 따라서 뷰티산업종사자는 피부미용 관리 및 피부 상태 진단에 있어서 피부미용 고객이 소셜미디어 커뮤니티와 같은 다양한 온라인 환경에서 자신의 피부미용 및 피부 상태와 관련한 사용자의 요구 사항을 파악하고 고객의 피부 상태 진단 및 관리의 상호작용을 할 필요가 있다.

그러나 소셜미디어에서 고객의 피부 상태에 대한 요구는 매우 비정형적이고 다양한 방식으로 표현되는 빅데이터의 특성을 가지고 있으므로 적절한 전처리와 빅데이터 분석이 요구된다. 그리고 소셜미디어의 고객 의견은 고객 중심의 매우 주관적이고 비전문적인 내용으로 표현되어 뷰티산업분야의 전문가적 표현과 차이가 존재한다. 피부 상태 진단 및 관리에 소셜미디어 환경을 이용하고자 할 때 이와 같은 피부 상태 평가와 표현에 대한 고객과 뷰티산업 종사자 사이의 차이로 인하여 결과적으로 피부 상태 진단과 분석이 객관적이지 못하거나 올바르지 못하게 되는 문제가 발생할 수 있다.

본 논문에서는 소셜미디어를 기반한 고객의 정확한 피

부 상태 진단 및 관리를 위하여 적절한 빅데이터 분석을 통하여 피부 상태 분석정보를 획득하고, 인공지능 기반의 인공지능 기술로 구현한 지능화된 피부 상태 진단시스템을 개발하였다. 즉, 소셜미디어에서 수집하여 분석된 피부 상태 진단 및 관리 정보를 기계학습의 훈련데이터로 가공하여 지능적으로 고객의 피부 상태를 진단할 수 있는 인공지능 기반의 피부 상태 진단시스템을 개발하고 성능을 분석하였다.

## II. Related Research

### 2.1 Skin condition analysis method

일반적으로 뷰티산업 종사자들이 피부 미용을 목적으로 피부 상태 진단을 하는 경우에 대부분 육안 관찰, 스마트폰 등의 보편적으로 보급된 관찰 장비를 사용한 관찰과 피부미용을 원하는 일반인들의 피부 상태에 대한 의견 등을 통하여 피부 미용적 진단과 관리를 진행하고 있다. 정확한 진단을 위해 사용되는 피부 상태 측정 장비는 현재까지는 고가의 장비 가격 및 측정기기의 크기 등으로 특별한 설치 운영 환경이 요구되어 대부분의 뷰티산업 종사자에게는 현실적으로 사용하기 쉽지 않고 일반 피부 미용 이용자에게도 피부상태 진단의 불편함과 비현실적 비용 부담으로 대중화되어 사용하기가 매우 어려운 실정이다.

일반적으로 뷰티산업종사자에 의해 진행되는 피부 상태 평가시 피부 표피 최상부의 각질층에 대한 평가가 매우 중요하다. 피부 상태 최상부 각질층에 대한 평가방법으로는 시진, 촉진, 문진, 기기측정 등의 방법이 사용되고 있다. 피부 미용 고객들을 대상으로 현재까지 일반적으로 많이 사용하고 있는 방법은 시진, 즉 육안으로 관찰하는 방법이다. 최근에는 기존의 대면 방식의 육안관찰 및 문진방법과 함께 온라인 시스템을 기반한 다양한 비대면 방법도 사용하고 있다. 대면 및 비대면을 통한 육안관찰은 제시된 피부 사진 이미지를 사용하여 진행된다. 피부 각질층의 표면 상태는 피부 각질층 하부의 다양한 변화를 반영하고 있으므로 육안으로 피부상태를 관찰하여 피부상태 유형을 판단하는 것은 타당성 있는 판별 방법이라 할 수 있다. 뷰티산업분야에서 피부상태를 육안으로 관찰하여 판단하는 방법을 사용하고 있으며 피부표면에 증상으로 나타나는 형태적 요소와 색채적 요소에 의해 판단되어지고 있다. 즉, 피부 각질층에 대한 피부 유형의 종류는 크게 피부표면의 형태적 요소인 피부결(skin texture), 모공(pore), 각질(scale), 주름(winkle)과 색채적 요소인 피부톤(skin

tone), 홍반(홍조, redness), 여드름(흑색면포, 백색면포, black head, white head), 색소침착(pigmentation)으로 나누어진다[1].

### 2.2 Artificial intelligence technology in skin care

피부미용과 화장품과학 분야의 기술 추세에 의하면 포괄적 고객을 대상으로 한 보편적 기능을 갖는 화장품과 보편적 피부미용 및 관리 방식에서 개인 맞춤형 화장품과 개인에게 최적화된 피부미용을 위한 방향으로 발전되어 가고 있다. 최근에 화장품 산업에서는 피부미백, 주름개선, 자외선차단 등 기능성 화장품이 개발되어 사용되고 있으며, 피부표면의 외형을 관리하는 동시에 건강한 피부를 위한 영양소를 제공하여 피부 조직 등을 향상 시키는 고기능성 화장품에 해당하는 코스메티컬(cosmeceutical)의 품목이 지속적으로 확대되고 있다.

이와 같은 소비자 중심의 맞춤형 화장품 조제와 소비자 중심의 피부 상태 진단 및 관리의 뷰티산업 서비스 공급을 위하여 빅데이터분석 및 인공지능기술을 접목한 체계적이고 지능화된 화장품 및 뷰티산업 개발 및 운영 환경을 구축하기 위한 연구가 진행되고 있다. 인공지능 기계학습 기술이 급속도로 발전되면서 화장품 및 뷰티산업분야에서 필수적으로 요구되는 피부 상태 유형 판단에 인공지능 기술을 활용하기 위한 연구도 다양하게 시도되고 있다. [4]에서는 소비자들의 피부관리를 위하여 소비자의 피부 이미지로부터 ORB(Oriented FAST and Rotated BRIEF)를 사용하여 특징점들을 추출하고 이 데이터를 기계학습 분류방법 SVM을 사용하여 기계학습을 진행하고, 학습된 시스템을 사용하여 입력된 피부 이미지의 피부 상태 유형을 건성과 지성으로 분류하는 방법이 연구되었다. 이때 사용한 기계학습 분류방법 SVM은 선형 분리 모델에 적합한 대표적인 기계학습 방법이다. [5]에서는 피부주름에 의한 피부진단을 위하여 광각렌즈를 탑재한 현미경으로 피부영상을 촬영하고, 이 피부영상으로부터 인공지능 딥러닝을 적용한 자동 주름 추출에 대한 연구가 진행되었다. 피부영상을 취득하기 위하여 고가의 광각렌즈를 탑재한 현미경과 특별한 운영 환경을 갖출 수 있는 곳에서 사용 가능하다.

피부관리 등의 뷰티산업분야 뿐만 아니라 화장품 제거나 성분분석 및 평가 과정에서도 인공지능기술이 다양하게 활용되었다. 위생 및 화장품 상용품 제품에 방부제로 첨가된 항균제 가운데 일부는 인간의 건강에 해로운 것으로 의심되고 있어 이에 대한 성분분석 및 평가 과정에 많은 연구와 관심이 집중되고 있다. 사용된 보존제들은 상이한 부류의 화학종에 속하며 일반적으로 그들의 혼합물로

사용된다. 따라서 품질관리분석에서 상이한 화학구조를 갖는 종을 동시에 결정할 수 있는 다중성분방법이 많이 요구된다. 이때 사용 가능한 크로마토그래피 방법에서는 동시에 모든 항균제를 분리 추출하기 위해 복잡도가 높은 성분 분석 모델의 개발 및 다양한 조건의 최적화가 요구된다. 크로마토그래피 방법의 모델 개발과 다양한 조건의 최적화를 위해 인공지능망을 사용하는 연구가 진행되었다.

화장품 및 뷰티산업 분야는 소비환경과 트렌드에 민감한 분야이므로 다양한 방법으로 고객의 니즈를 파악할 필요가 있다. 이에 따라 화장품 및 뷰티산업 분야에서 다양한 소셜미디어로부터 수집된 빅데이터를 이용하여 고객의 니즈를 파악하고자 하는 연구가 진행되고 있다.

### III. The Proposed System

#### 3.1 Overview

화장품 및 뷰티산업 분야의 기술 트렌드는 맞춤형 제품과 서비스를 지향하고 있으며, 이에 대한 고객의 요구 수준은 더욱더 고급화되어 가고 있다. 맞춤형 제품과 서비스를 제공하기 위해 가장 기본적이고 필수적인 기능은 고객의 피부 상태에 대한 진단이다. 기존 뷰티산업 분야에서 피부 미용 및 관리 목적의 피부 상태 진단은 다양한 형태의 고객 요구 의견 파악과 피부의 육안 관찰을 통하여 이루어졌다.

정보통신기술의 발전과 소셜미디어 서비스가 일반화되면서 피부 상태 및 관리와 관련하여 사용자들의 다양한 질문과 응답으로 구성되는 정보교환과 상호작용이 소셜미디어 커뮤니티에서 자발적으로 이루어지고 있다. 따라서 화장품 및 뷰티산업 분야의 기술 트렌드인 맞춤형 제품 및 서비스를 구현하고 높아진 고객의 요구를 충족시키기 위해 소셜미디어 상에서 표현되는 고객의 요구 사항을 파악하고 진단할 필요가 있다. 그러나 소셜미디어에서 표현되고 교류되는 피부 상태에 대한 요구 및 관리 내용은 매우 비정형적이며 다양한 표현을 포함하고 있고 주관적 판단과 의견이 포함하고 있으므로 이에 대한 적절한 분석과 지능적인 진단을 할 수 있는 시스템이 필요하다. 즉, 고객이 소셜미디어에서 표현한 피부미용 및 피부 상태와 관련한 요구 사항을 파악하고 피부 상태 진단이 가능한 지능적이며 체계적인 도구와 환경을 이용할 필요성이 요구되고 있다.

본 논문에서는 우리가 앞서 연구한 [6]의 연구 결과로 소셜미디어에서 피부 상태 분석을 위한 빅데이터 분석 결과를 인공지능망을 훈련하기 위한 소스 정보로 활용하였

다. 본 논문에서는 소셜미디어를 통하여 제시된 고객의 피부 상태에 대한 요구 내용인 텍스트를 중심으로 분석한 결과를 사용하여 피부 상태를 지능적으로 진단하기 위한 소셜미디어기반의 지능형 피부 상태 진단정보시스템 SCDIS(Skin Condition Diagnosis Information System)를 개발하였다. 소셜미디어 커뮤니티에서 교류되는 정보는 매우 비정형적이고 데이터 사용 목적에 부합되지 않는 다양한 내용과 형태를 포함하고 있으므로 이러한 특성을 고려할 수 있는 빅데이터기술과 딥러닝 기계학습 기술을 사용하여 지능화된 시스템으로 개발하였다.

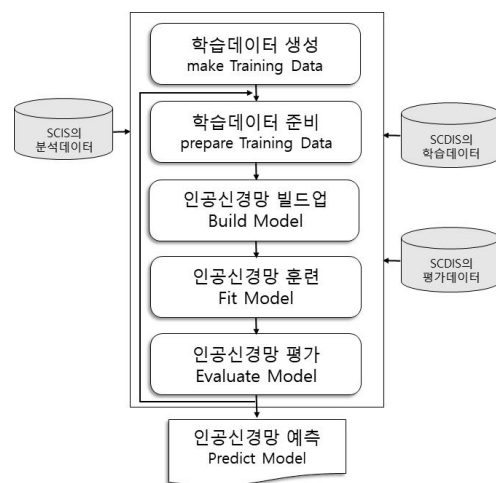


Fig. 1. Overview of SCDIS

SCDIS는 그림 1과 같이 훈련데이터생성단계, 훈련데이터준비단계, 인공지능망빌드업단계, 인공지능망훈련단계, 인공지능망평가단계, 인공지능망예측단계로 구성된다.

#### 3.2 Training data generation

SCDIS의 첫 번째 단계인 훈련데이터생성단계(makeTrainData)는 기계학습의 첫 번째 단계인 훈련 데이터 수집 단계에 해당하는 것으로 SCDIS의 최종 예측 모델의 성능을 결정하는 중요한 단계이다. 피부 상태 진단을 위한 훈련데이터를 생성하기 위하여 빅데이터 분석 결과를 훈련데이터생성단계의 소스 데이터로 사용하였다. 빅데이터 분석 결과인 핵심 단어에 대한 분석정보인 상대빈도수(TFIDF) 결과를 가공하여 SCDIS의 학습데이터로 생성하였다. 일반적으로 기계학습에 사용되는 훈련데이터는 입력값에 해당하는 특징들과 정답에 해당하는 라벨로 구성된 데이터를 말한다.

표 1은 훈련데이터생성단계에서 생성된 훈련데이터의 구조를 나타내고 있다.

Table 1. Structure of training data

like	skinType	word 1	word 2	word i
300	주름	0.123	0.543	...
10	모공	0.321	0.789	...
...	...	...	...	...

표 1에서 like는 소셜미디어로부터 수집된 각 게시물에 대한 “좋아요” 값을 나타내고, “word i”는 게시물 텍스트로부터 분석 추출된 단어들을 나타내고 있다. 표 1에서 word i(i = 핵심단어 개수) 컬럼은 빅데이터 분석결과인 상대빈도수 값이고, skinType은 상대빈도수 값 등을 근거로 결정된 각 게시물에 대한 피부상태 진단 유형 결과를 나타낸다. SCDIS의 훈련데이터의 입력 값에 해당하는 특징으로 word i 컬럼의 상대빈도수 값을 사용하였고, 정답에 해당하는 라벨로 skinType 컬럼의 피부상태 유형 결과를 사용하였다. 훈련데이터로 사용할 게시물 데이터는 게시물의 좋아요 값에 대한 임의의 값(>10) 이상을 갖는 게시물로 제한하여 처리하였다. 소셜미디어 이용 환경에서 좋아요 값은 여러 사용자들로부터 해당 게시물 내용에 대하여 긍정 또는 지지를 나타내는 역할을 하는 것이므로 이를 해당 게시물 내용의 정확성과 객관성을 판단하는 정보로 활용하였다. 물론 좋아요 값이 높다 하더라도 정확성과 객관성이 떨어지는 경우가 있을 수 있지만, 빅데이터의 방대한 분량의 정보를 사용하기 때문에 이와 같은 특별한 예외 상황으로 인한 영향은 받지 않는다.

표 1의 분석된 피부상태 진단 정보를 활용하여 기계학습의 훈련데이터셋을 생성하였다. 훈련데이터셋은 핵심단어(word i)에 대한 상대빈도수값들과 피부 상태 유형(skinType)값으로 구성된다. 훈련데이터생성단계에서의 skinType은 이미 분석된 상대빈도수 값이 높은 단어 가운데 피부상태 진단 유형과 일치하거나 포함되는 피부상태 유형을 해당 게시물의 피부상태 유형으로 결정하였다. skinType 값으로 결정되는 피부유형은 뷰티산업분야에서 피부상태 진단에 일반적으로 사용하는 “피부결”, “모공”, “각질”, “주름”, “피부톤”, “홍반”, “홍조”, “여드름”, “색소침착”의 9개 피부 상태 유형 가운데 상대빈도수가 높은 유형으로 결정된다. 인공지능망 훈련과정에서 skinType이 훈련데이터에 정답 라벨로 사용될 때에는 인공지능망의 출력값과의 차이를 정량화하기 위해 원핫인코딩(one-hot-encoding)을 적용하여 사용하였다.

### 3.3 Training data preparation

SCDIS의 두 번째 단계인 훈련데이터준비단계(prepareTrainData)는 설계된 인공지능망을 훈련과정에

서 신뢰성 있는 피부 상태 진단 도구로 사용될 수 있도록 훈련데이터를 준비하는 과정이다. 일반적으로 기계학습에서 사용되는 데이터는 훈련데이터와 테스트데이터의 두 부분으로 분할되어 사용된다. 훈련데이터는 설계된 인공지능망을 훈련 또는 학습하는 데 사용되고, 테스트데이터는 훈련된 인공지능망 모델의 성능을 평가하는 데 사용된다. 훈련데이터에 포함되어 있지 않는 테스트데이터로 훈련된 인공지능망 모델을 평가하는 것이 실제 현장에서 발생하는 새로운 내용의 피부상태 진단에 대한 예측 정확성 평가에 적합하기 때문이다.

```
# create a ANN object from class AnnTFIDF
ann = AnnTFIDF()

# Training data preparation
X_train, y_train, X_test, y_test, df_srcIdx
= ann.prepareTrainData(inFile)
```

Fig. 2. The process of the training data preparation stage

본 논문에서는 훈련데이터생성단계에서 가공된 데이터를 반복적인 훈련 과정에서 매번 랜덤하게 샘플링하였으며, 훈련데이터는 전체 데이터의 80%, 테스트데이터는 전체 데이터의 20%로 분할하여 사용하였다. 그림 2에 나타난 것과 같이 훈련데이터준비단계의 결과는 훈련데이터인 X\_train, y\_train과 테스트데이터인 X\_test, y\_test이다. 여기에서 X\_train과 x\_test는 입력값이고, y\_train과 y\_test는 정답 레이블이다.

### 3.4 Artificial neural network model buildup

SCDIS의 세 번째 단계인 인공지능망모델빌드업단계(buildModel)는 사용할 인공지능망 모델 AnnTFIDF를 설계하고 구현하는 단계이다. 인공지능망모델빌드업단계에서는 텐서플로우의 딥러닝 프레임워크와 케라스 라이브러리를 사용하였다. 그리고 반복적인 훈련과 평가 과정을 거쳐 가장 높은 성능을 나타내는 입력층, 은닉층, 출력층의 인공지능망 구조를 모델링하고, 은닉층의 개수, 뉴런의 개수, 활성화함수, 학습률 등의 하이퍼 파라미터를 최적화하였다. 또한 1개의 은닉층으로 구성되는 단층인공지능망으로 된 선형적 모델이 아닌 여러 개의 은닉층으로 구성된 비선형적 모델을 사용하였다.

그림 3은 인공지능망모델빌드업단계에서 빌드업하고자 하는 AnnTFIDF 인공지능망 모델에 대한 개념적 구조를 나타내고 있다. 입력층(Input Layer)은 분석된 피부상태 진단 및 관리에 대한 핵심 단어의 상대빈도수 값을 입력받을 수 있도록 최대 핵심 단어의 개수를 고려한 충분한 개수의 뉴런으로 구성한다. 출력층(Output Layer)은 피부상

태 유형 진단의 종류인 "피부결", "모공", "각질", "주름", "피부톤", "홍반", "홍조", "여드름", "색소침착"를 출력할 수 있도록 9개의 뉴런으로 구성한다. "홍반"과 "홍조"는 동일한 피부상태 유형을 나타내고 있으나 본 논문에서는 소셜 미디어에서 구별되어 수집하고 분석되어 이에 맞추어 2개의 유형을 구분하여 출력할 수 있도록 구성하였다.

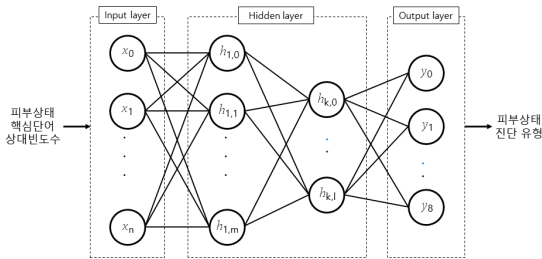


Fig. 3. Architecture of build-up model AnnTFIDF

은닉층(Hidden Layer)은 반복적인 훈련 과정을 통하여 적절한 은닉층 개수, 은닉층 유형, 뉴런의 개수 등이 결정되어 구성된다. 은닉층의 구성 내용은 인공지능망의 정확도와 견고성에 많은 영향을 미친다.

인공지능망에서는 각 층에 포함된 뉴런의 출력 값을 계산하기 위하여 활성화함수가 사용되며, 인공지능망의 성능과 적합성을 높이기 위해 인공지능망 모델을 빌드업하는 과정에서 적절한 활성화함수를 결정하게 된다. 일반적으로 사용되는 활성화함수는 계단함수, 시그모이드함수, ReLU 함수 등이 있다. 본 논문에서 사용되는 훈련데이터는 소셜 미디어로부터 수집, 분석된 방대한 정보를 근간으로 하고 있으므로, 완성된 인공지능망의 예측 성능을 높일 수 있는 활성화함수의 선택이 필요하다. 활성화함수 ReLU함수는 미분도 간단하고 심층신경망에서 나타나는 그라디언트 감쇠가 일어나지 않아서 본 논문의 annTFIDF 인공지능망 모델 구성에 적합한 특성을 갖는다. 인공지능망의 예측 성능과 효율성을 고려하여 입력층과 은닉층의 활성화함수는 ReLU함수를 사용하여 설계하였으며, 출력층의 활성화함수는 피부상태 유형 결정에 적합한 Softmax 활성화함수를 사용하여 설계하였다.

### 3.5 Artificial neural network training

SCDIS의 네 번째 단계인 인공지능망훈련단계(fitModel)는 학습률 등의 하이퍼 파라미터 수정을 통한 모델의 예측 능력 개선과 최종 훈련된 모델의 획득 과정이다. 즉, 훈련 데이터로 주어진 특정 피부상태 게시물이 어떤 피부상태 유형인지를 예측하도록 한다. 그리고 예측 결과가 정답인 경우 결정 과정에 사용된 인공지능망의 가중치와 바이어스

매개변수를 강화하고, 오답인 경우 정답이 나오도록 가중치와 바이어스 매개변수를 수정하여 점진적으로 예측의 정확도를 향상시킨다. 인공지능망의 훈련은 케라스에서 제공되는 API 함수에 다양한 모델의 파라미터를 설정하여 인공지능망의 성능을 높일 수 있도록 진행한다. 성공적으로 훈련된 인공지능망을 획득하기 위해서는 모델의 하이퍼 파라미터인 학습률, 배치크기(batch size), 에포크(epochs)수, 손실함수(loss function), 옵티마이저(optimizer) 등을 적절하게 설정하여 반복 훈련을 진행한다.

그림 4는 인공지능망모델빌드업단계에서 설계된 AnnTFIDF 인공지능망 모델에 하이퍼 파라미터인 손실함수(loss='categorical\_crossentropy'), 옵티마이저(optimizer='adam'), 배치크기(batch\_size=900), 에포크(epochs=200)를 설정하여 훈련하는 과정에 대한 코드의 일부이다. 손실함수 "categorical\_crossentropy"는 인공지능망이 예측한 피부상태 유형 값과 정답에 해당하는 피부상태 유형 값의 차이를 비교하는 함수이다. 인공지능망을 훈련하는 과정은 손실함수에 의해 평가된 오류를 최소화 시키기 위해 사용되며 인공지능망 모델의 성능을 높이기 위하여 여러 가지 손실함수를 선택하여 사용할 수 있다.

```
# Build AnnTFIDF model
model = ann.buildModel(modelType=modelType)

# compile AnnTFIDF model
model.compile(loss='categorical_crossentropy',
              optimizer='adam', metrics=['accuracy'])

# Training AnnTFIDF model
hist = model.fit(X_train, y_train, batch_size=900,
                epochs=200, verbose=1, validation_split=0.1)
```

Fig. 4. The process of artificial neural network model training stage

본 논문에서는 피부상태 유형 진단에 손실함수 CategoricalCrossentropy(categorical\_crossentropy)를 사용하였다. 손실함수 CategoricalCrossentropy는 훈련중인 모델의 예측 값인 피부상태 유형과 정답 레이블인 피부상태 유형 간의 교차 엔트로피 손실을 계산하는 함수이다.

옵티마이저는 손실함수를 미분하여 최적의 학습률과 모멘텀으로 훈련중인 모델의 가중치를 변경하는 파라미터이다. 본 논문에서는 현재 가장 많이 사용되는 옵티마이저인 Adam 옵티마이저를 사용하였다. Adam 옵티마이저는 최근의 그라디언트만 사용함으로써 전역적 최저점에 도착하기 전에 지역적 최저점에서 알고리즘이 멈추는 문제를 해결한 옵티마이저이다. 인공지능망훈련단계에서는 높은 성능을 나타내는 에포크, 배치크기 등 하이퍼 파라미터를 검토하여 최종적으로 훈련된 인공지능망을 완성하였다. 훈련

된 인공지능망 AnnTFIDF의 가중치 값을 저장하여 이후 예측 과정에 사용하였다.

#### IV. Analysis on Proposed System

피부상태 진단정보시스템 SCDIS는 소셜미디어 인스타그램으로부터 도출된 분석 정보를 인공지능망기술의 훈련데이터로 활용한 지능적인 피부상태 진단 및 관리 시스템이다. 본 논문에서는 훈련데이터생성단계와 훈련데이터준비단계를 통하여 피부상태 진단에 사용할 훈련데이터를 생성 및 준비하고, 인공지능망빌드업단계, 인공지능망훈련단계, 인공지능망평가단계, 인공지능망예측단계를 인공지능망의 다양한 파라미터를 반복적으로 실험 및 평가하여 적절한 성능을 갖는 인공지능망 모델을 빌드업하였다. 최종적으로 빌드업된 인공지능망 AnnTFIDF 모델을 사용하여 테스트데이터로 준비된 소셜미디어의 게시물에 나타내고 있는 피부 상태 유형을 예측하고 이에 대한 예측 결과를 분석하였다.

SCDIS 시스템 개발을 위해 기계학습 및 데이터분석 등에 사용 가능한 아나콘다3(Anaconda3), 텐서플로우(Tensor Flow), 케라스(Keras) 개발환경과 파이썬(Python) v3.8 프로그래밍 언어를 사용하였다.

##### 4.1 Training data generation and preparation result analysis

본 논문에서는 빅데이터 분석정보인 상대빈도수(TFIDF) 결과를 가공하여 SCDIS의 훈련데이터로 생성하였다. 훈련데이터는 핵심단어에 대한 상대빈도수 값을 입력값에 해당하는 특징들로 사용하였고, 게시물 항목의 좋아요 값과 분석된 핵심단어 정보를 사용하여 결정된 피부상태 유형 값을 정답에 해당하는 라벨로 구성하였다. 그림 5는 표 1에 설명된 구조로 훈련데이터생성단계에 의해 생성된 훈련데이터를 엑셀파일로 저장한 결과 내용이다.

	A	B	C	D	E	F	G	H	I	J
1	Index	keyword	srcidx	like	skinType	크림	아드름	강종	피부	드러블
2	0	피부색소증	0	131	색소침착	0.0029747	0	0	0.067411	0
3	1	피부색소증	1	753	색소침착	0.114347	0	0	0.117708	0
4	2	피부색소증	2	131	색소침착	0.0234	0	0	0.175704	0
5	3	피부색소증	3	108	색소침착	0.24335	0	0	0.099667	0
6	4	피부색소증	5	807	색소침착	0.35232	0	0	0.120247	0
7	5	피부색소증	6	548	색소침착	0.335398	0	0	0.032706	0
8	6	피부색소증	7	347	색소침착	0.301768	0	0	0.068662	0
9	7	피부색소증	8	403	색소침착	0.168148	0	0	0.038259	0
10	8	피부색소증	10	892	색소침착	0.144237	0	0	0.035902	0
11	9	피부색소증	15	20	색소침착	0	0.126887	0.606035	0.127799	0
12	10	피부색소증	16	20	색소침착	0	0.145801	0.696372	0.14685	0

Fig. 5. Results of training data generation

훈련데이터생성단계에서 저장된 훈련데이터는 인공지능망을 훈련 및 평가할 때마다 랜덤하게 샘플링되어 사용된다. 그림 6은 훈련데이터준비단계에서 생성된 전체 훈련데이터 1,739개의 데이터를 훈련샘플데이터 X\_train, y\_train 80%(1,392개), 테스트샘플데이터 X\_test, y\_test 20%(347개)로 분할하여 준비한 훈련데이터 구성 내용을 나타내고 있다.

훈련 샘플 X\_train 데이터 크기: (1392, 6183)  
 훈련 샘플 y\_train 레이블 크기: (1392, 9)  
 테스트 샘플 X\_train 데이터 크기: (347, 6183)  
 테스트 샘플 y\_test 레이블 크기: (347, 9)

Fig. 6. Overview of prepared training data

훈련샘플데이터는 설계된 인공지능망 모델을 훈련 또는 학습하는 데 사용되고, 테스트샘플데이터는 훈련된 인공지능망 모델의 성능을 평가하는 데 사용된다. 훈련샘플 데이터와 테스트샘플데이터로 구분하여 사용하는 것은 훈련된 인공지능망의 평가 결과를 객관적인 관점에서 분석할 수 있게 하는 훈련데이터 생성 및 준비 방법이라고 평가할 수 있다. 즉, 훈련샘플 데이터에 포함되어 있지 않은 테스트 샘플 데이터로 평가한 모델의 정확성 및 성능은 매우 높다고 평가할 수 있다.

##### 4.2 Analysis of artificial neural network AnnTFIDF build-up result

인공지능망빌드업단계에서는 다양한 파라미터를 적용하고 반복적으로 실험을 진행하여 인공지능망 모델 AnnTFIDF를 구축하였다. 표 2는 빌드업한 AnnTFIDF 인공지능망 모델에 대한 요약 내용을 나타내고 있다.

Table 2. Summary of built-up model AnnTFIDF

Layer(type)	Output shape	Param #
Input Layer	(None, 6183)	0
Hidden Layer1(Dense)	(None, 3000)	18,552,000
Hidden Layer2(Dense)	(None, 1000)	3,001,000
Hidden Layer3(Dense)	(None, 500)	500,500
Hidden Layer4(Dense)	(None, 250)	125,250
Hidden Layer5(Dense)	(None, 128)	32,128
Output Layer	(None, 9)	1,161
Total		22,212,039

입력층(Input Layer)에 포함되는 뉴런의 개수는 분석된 피부상태 진단 및 관리에 대한 핵심 단어의 상대빈도수 값을 입력 받을 수 있도록 최대 핵심단어 개수인 6,183개의 뉴런으로 구성하였다. 출력층(Output Layer)에 포함되는 뉴런의 개수는 피부상태 유형 진단의 종류인 "피부결", "모

공, "각질", "주름", "피부톤", "홍반", "홍조", "여드름", "색소침착"를 나타낼 수 있도록 9개의 뉴런으로 구성하였다. 은닉층(Hidden Layer)은 조밀하게 연결된 형태의 Keras의 Dense 유형의 레이어 5개(Hidden Layer 1 ~ Hidden Layer 5)로 구성하였다.

빌드업된 인공신경망 모델인 AnnTFIDF의 은닉층 개수, 유형, 각 은닉층의 뉴런의 개수는 인공신경망의 정확도와 견고성을 높이기 위하여 반복적인 실험을 통하여 인공신경망의 다양한 파라미터와 함께 결정되었다. 활성화함수는 인공신경망 빌드업과정에서 일반적으로 효율성과 성능이 우수하다고 평가되는 ReLU함수를 적용하였다. 빌드업된 인공신경망 AnnTFIDF에서 훈련과정을 통하여 학습 결과를 나타내는 가중치 등의 파라미터는 총 22,212,039개로 표 2와 같다.

그림 7은 완성된 AnnTFIDF 인공신경망 모델의 구조를 시각화하여 나타내고 있다. 각 은닉층은 Keras의 Dense 유형으로 구성되어 있고 활성화함수는 ReLU함수를 사용하고 있는 것을 보여주고 있다.

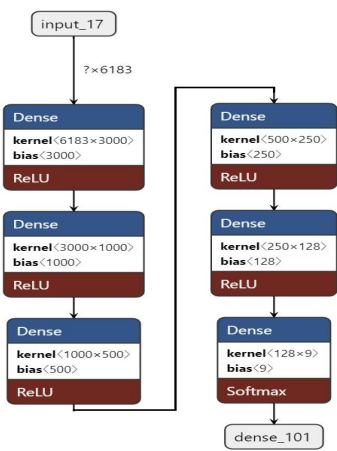


Fig. 7. Architecture of built-up model AnnTFIDF

인공신경망 AnnTFIDF의 출력층은 피부상태 유형 진단 결과를 나타내는 계층으로 피부상태 유형 진단을 위해 Softmax 활성화함수를 사용하였다. 출력층의 Softmax 활성화함수는 가장 높은 예측 값을 갖는 피부상태 유형을 결과로 출력할 수 있도록 한다. 즉, 9개의 피부상태 진단 종류 가운데 가장 높은 예측 결과를 갖는 피부상태 유형 종류를 AnnTFIDF 진단 결과로 출력하도록 하였다.

SCDIS의 네 번째 단계인 인공신경망훈련단계는 모델의 성능과 예측 능력을 높이기 위해 학습률, 손실함수 등 파라미터를 수정하고, 반복 훈련을 통하여 적절한 모델의 성능을 갖는 인공신경망 훈련을 완료하는 단계이다. 그림 8

은 훈련데이터와 검증데이터를 사용한 훈련과정의 훈련 정확도(train accuracy), 검증 정확도(valid accuracy), 훈련 손실값(train loss), 검증 손실값(val loss) 등의 분석된 성능 결과를 보여주고 있다. 그림 8에서 가로축은 epoch를 나타내고, 세로축은 정확도(accuracy)와 손실값(loss)를 나타낸다.

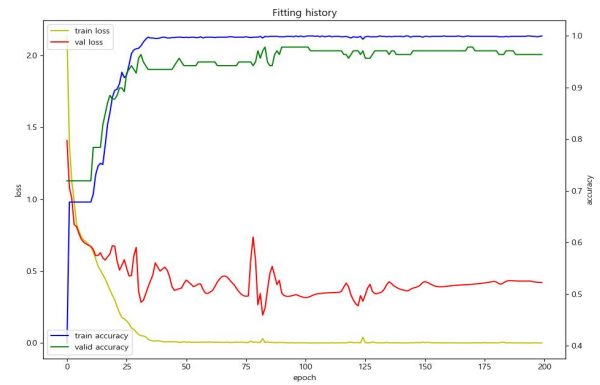


Fig. 8. Fitting history of built-up model AnnTFIDF

훈련샘플 데이터에 대해서는 약 25 에포크에서 훈련 정확도(train accuracy)는 1에 가까운 높은 정확도를 유지하고, 훈련 손실 값(train loss)은 0에 가까운 좋은 성과를 나타내고 있다. 검증샘플 데이터에 대해서는 약 25 에포크까지 검증 정확도(valid accuracy)가 빠른 속도로 상승하고, 검증 손실 값(valid loss)은 0으로 수렴하는 좋은 성과를 보여주었으나 이후 에포크에서 검증 손실 값은 불안정한 변화를 나타내면서 줄어드는 방향성을 보여주었다. 검증샘플 데이터에 대한 검증 정확도는 약 25 에포크까지 급격하게 상승하고 이후 평균적으로 90% 이상의 정확도를 나타내었다.

인공신경망 모델의 훈련과정에는 훈련샘플 데이터의 10%를 검증샘플 데이터(validation data)로 사용하였다. 반복 훈련은 사용할 각 단계의 데이터 크기인 배치크기 batch\_size=600, 훈련 반복 회수인 에포크 epochs=200으로 진행하였다. 이러한 인공신경망훈련 과정에서의 훈련샘플 데이터에 대한 성과와 이후 단계인 테스트샘플데이터에 대한 성능 평가 결과를 반복적으로 분석하고 훈련함으로써 적절한 수준의 최종적인 훈련된 인공신경망 모델 AnnTFIDF를 완성하였다.

SCDIS의 다섯 번째 단계인 인공신경망평가단계는 테스트샘플 데이터를 사용하여 모델의 예측값을 계산하고 정답 레이블과 비교 분석하는 과정이다. 인공신경망평가단계에 사용되는 테스트샘플 데이터는 훈련샘플데이터에 포함되지 않는 데이터로서, 매 평가단계마다 랜덤하게 샘플링되어 모델이 공정하게 평가 받을 수 있도록 준비된 데이터



이다. 테스트샘플 데이터는 전체 훈련데이터의 20%(347개)를 사용하였다.

훈련된 인공지능망 AnnTFIDF의 테스트샘플 데이터 347개에 대한 평가분석 결과는 그림 9와 같다. 피부상태 유형 진단 예측 값의 정확성(test accuracy)은 약 95%의 높은 성능을 보이고 있고, 손실 값(test loss)은 약 0.44로 0에 수렴하는 비교적 낮은 오차를 나타내고 있다.

```
Evaluate Results
batch_size = 69
score = [0.4426054060459137, 0.9481267929077148]
test loss = 0.4426054060459137
test accuracy = 0.9481267929077148
```

Fig. 9. Evaluate results of AnnTFIDF

표 3은 테스트샘플 데이터를 사용하여 평가하고, 테스트샘플 데이터의 개별적 게시물의 텍스트 내용에 대한 훈련된 모델 AnnTFIDF의 피부상태 유형에 대한 진단 결과값(predict)을 나타내고 있다.

Table 3. Results of model AnnTFIDF prediction

content	show imgFile	label	predict
20211122첫 호스트 경험날설고 떨리고 ㅋㅋ 그래도 너무 재미있었던 하루 기회주셔서 감사합니다		피부톤	피부톤
22 할인중 한마디로 50할인 저는 2년동안 그렇기 때문에 더더욱 강한자극 없이 트러블 없이 매일		피부결	모공
기미잡티 관리와 주름개선까지 도움이 되니넘나 좋은것별질이 쉬운 비타민C를 최대한 변질없이 사		피부결	피부결
혼자 카밍젤팩 하면서 2030분 뒤에빠지고 싫으세요어서오세요사 인이가 함께 해요		각질	각질

표 3에서 content는 소셜미디어 게시물 텍스트를 나타내고, label은 훈련데이터세트의 정답인 피부유형, predict는 훈련된 모델 AnnTFIDF로 예측된 피부유형 결과를 나타낸다. 표 3과 같이 모델의 예측 결과값 predict을 게시물의 콘텐츠(content), 이미지 출력 화면(show imgFile), 피부상태 유형 정답(label)과 함께 나타내어 피부진단의 결과를 확인할 수 있도록 하였다.

## V. Conclusions

컴퓨터 및 정보통신기술의 발전과 함께 4차 산업혁명기술이 다양한 분야와 실생활에 적용되어 가고 있으며, 화장품 및 피부미용 분야에서도 혁명과 같은 큰 변화와 발전이 일어나고 있다. 최근 화장품산업과 뷰티산업의 동향은 기능과 성분을 중시하며 기능성 화장품과 개인 맞춤형 뷰티 산업 방향으로 발전되어 가고 있다. 화장품과 뷰티산업에서 개인 맞춤형 기능성 화장품과 미용 서비스를 제공하기 위해서는 무엇보다 화장품 및 미용 소비자의 니즈를 파악하는 것이 가장 중요하고, 가장 필수적인 기능은 소비자의 피부 상태 및 요구 사항 파악이라 할 수 있다. 이러한 관점에서 화장품과학 및 미용과학 분야의 뷰티산업 종사자와 연구개발자에게 고객 피부상태의 파악과 진단은 중요한 관심사이다.

4차 산업혁명시대로 사회 전반이 발전되어 가면서 기존의 많은 활동이 온라인 비대면 환경으로 급속하게 발전하고 온라인을 기반으로 한 새로운 세대가 나타나고 있다. 따라서 다양한 온라인 환경에서 자신의 피부미용 및 피부 상태와 관련된 사용자의 요구 사항을 이해 및 파악하고 이를 참고하여 고객의 피부미용 관리 및 피부 상태 진단의 상호작용을 할 필요가 있다.

피부상태 진단에 대한 기존 연구에서는 준비된 피부상태 이미지를 활용하여 선형적 분류 알고리즘인 SVM(Support Vector Machine)으로 기계학습을 진행하고 이를 근거로 피부 이미지를 지성과 건성으로 분류하는 제한된 방식의 피부상태 진단 모델에 관한 연구를 진행하였다. 기존 연구에서는 피부 이미지로부터 정형적인 형태의 기하학적 특징점을 추출하는 제한점과 양질의 훈련데이터를 준비해야한다는 제한점이 있다. 따라서 화장품 및 뷰티산업분야에서 고객의 피부상태 진단 및 관리에 있어서 변화되는 화장품 및 피부미용에 대한 고객들의 요구 사항을 지속적이고 실시간적으로 파악하고 대응할 빅데이터 분석시스템과 지능화된 피부 상태 유형 진단 및 관리 체계 시스템의 연구가 필요하다.

소셜미디어 상의 피부 상태 요구 표현은 매우 비정형적이고 다양한 표현의 방대한 규모의 빅데이터로 발생되므로 인공지능기술이 접목된 지능화된 솔루션을 제공하는 것이 필요하다. 즉, 소셜미디어에 표현된 고객의 피부 상태 요구에 대하여 최신의 빅데이터 분석기술과 딥러닝 기계학습 기술을 사용하여 피부상태 유형을 진단하고 관리하는 지능화된 도구 및 환경이 필요하다.

본 논문에서 개발한 SCDIS는 소셜미디어의 피부 상태

분석정보를 피부 상태 유형 진단을 위한 훈련 데이터로 가공하여 인공지능망 AnnTFIDF 시스템의 훈련데이터로 사용함으로써 고객의 피부 상태를 지능적으로 진단 및 관리할 수 있다. SCDIS는 훈련데이터생성단계, 훈련데이터준비단계, 인공지능망빌드업단계, 인공지능망훈련단계, 인공지능망평가단계, 인공지능망예측단계로 구성된다. 훈련데이터생성단계에서는 소셜미디어의 피부 상태 게시물에 대한 빅데이터 분석정보인 좋아요 값, 핵심 단어 분석정보인 상대빈도수 결과를 활용하여 피부 유형 진단에 사용할 훈련데이터를 생성한다. 훈련데이터준비단계는 설계된 인공지능망에서 사용할 데이터를 훈련샘플데이터와 테스트샘플데이터로 구성하여 준비하는 과정이다. 인공지능망빌드업단계 및 훈련단계는 인공지능망 모델에 대한 입력층, 은닉층, 출력층의 구조를 모델링하고 피부상태 진단 성능을 높일 수 있도록 최적의 하이퍼 파라미터를 결정하는 과정이다. 인공지능망 평가 및 예측단계에서는 소셜미디어로부터 수집된 테스트샘플 데이터를 사용하여 평가하고 분석하였다. 인공지능망 모델 AnnTFIDF의 성능 분석 결과에서 피부상태 유형 진단 예측 값의 정확성이 약 95%의 높은 성능을 나타내었고, 손실 값은 약 0.44의 낮은 오차값을 나타냄으로써 화장품 및 뷰티산업분야의 피부상태 진단 및 관리에 있어서 지능화된 도구 및 환경으로 사용 가능한 것으로 분석되었다.

본 논문에서는 빅데이터의 텍스트정보를 중심으로 분석된 피부 상태 분석정보를 활용하여 피부 상태를 지능적으로 진단하는 시스템을 구축하였다. 향후에는 빅데이터의 이미지 정보를 활용하여 더욱더 발전된 지능적인 피부 상태 진단이 가능한 추가적인 연구가 필요할 것이다. 빅데이터 분석 결과를 훈련데이터로 활용한 지능화된 피부 상태 진단 도구 SCDIS는 급변하는 4차 산업환경에 대응하고 맞춤형화장품 등 소비트렌드를 충족시킬 수 있는 솔루션이 될 것으로 기대된다. 또한 SCDIS는 화장품 및 뷰티산업분야의 발전과 향후 보다 정교한 피부 상태 진단시스템을 구축하는데 필요한 기초 연구자료가 될 것이다.

## REFERENCES

- [1] Eui-Hyang Kim, Hyun-joo Kim, "A Study on the Analysis Method of Skin Condition through Visual Confirmation of Skin Surface", *Journal of Convergence for Information Technology*, Vol. 11. No. 5, pp.267-275, May. 2021
- [2] J. S. Koh, "Observation of Sebum, Moisture content and Microtopography of Different aged Skin for classification of facial Skin Types", *Journal of the Society of Cosmetic Scientists of Korea*, Vol. 15. No. 1, pp.63-70, 1989
- [3] Salton G. and McGill, M. J., "Introduction to modern information retrieval", McGraw-Hill, ISBN 0-07-054484-0., 1983
- [4] Kim Ki Young, "A Study on Model of Skin Type Judgment Tool Using Machine Learning Technique", *The Treatise on The Plastic Media*, Vo4. No. 21, pp.115-121, Nov. 2018
- [5] Hyeon-yeong Choi, Jae-pil Ko, "Deep Learning-based Automatic Wrinkles Segmentation on Microscope Skin Images for Skin Diagnosis", *Journal of Advanced Navigation Technology*, Vol. 24. No. 2, pp.148-154, Apr., 2020
- [6] Hyung-Hoon Kim, Jeong-Ran Cho, "Development of big data based Skin Care Information System SCIS for skin condition diagnosis and management", *Journal of The Korea Society of Computer and Information*, Vol. 27. No. 3, pp.137-147, Mar. 2022
- [7] Zheng, Tong, Zha. Dongmiao, Wang, Xingtian, Gao, Weijun, Miao, Sheng, "Architecture Descriptions Analysis Based on Text Mining and Crawling Technology", 2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer, pp.138-143, Nov. 2021
- [8] Hyung-Woo Lee, "Development of Supervised Machine Learning based Catalog Entry Classification and Recommendation System", *Journal of Internet Computing and Services*, Vol. 20. No. 1, pp.57-66, Feb. 2019
- [9] Dae-Chul Kim, Cheol\_hee Lee, Myong-Hui Choi, Yeong-Ho Ha, "Skin Detection Method using Color Space based Methods and Focus Region", *Journal of Korean Society for Imaging Science & Technology*, Vol. 18. No. 4, pp.16-22, Dec. 2012
- [10] Jeonghoon Kang, Hojung Lim, Wonsik Ko, "Visualization of Text Mining Technology based on Big Data", *The Korean Society of Visualization*, Dec. 2015
- [11] Won-Jo Lee, "A Study on Word Cloud Techniques for Analysis of Unstructured Text Data", *The Journal of the Convergence on Culture Technology*, Vol. 6. No. 4, pp.715-720, Nov.. 2020
- [12] M. Han, Y. Kim, C. Lee, "Analysis of News Regarding New southeastem Airport Using Text Mining Techniques", *Smart Media Journal*, Vol. 6. No. 1, 2017
- [13] Joon-Hwan Kim, Hyung-Jin Mun, Hang Lee, "A Study on Trend Analysis in Convergence Research Applying Word Cloud in Korea", *Journal of Digital Convergence*, Vol. 19. No. 2, pp.33-38, 2021
- [14] Min Ah Jang, Jung Min Lee, "Research on domestic and international industrial trends of functional cosmetics", *Journal of Digital Convergence*, Vol. 38. No. 2, pp.618-627, Apr., 2021
- [15] E. Marengo, V. Gianotti, S. Angioi, M.C. Gennaro, "Optimization by experimental design and artificial neural networks of the ion-interaction reversed-phase liquid chromatographic separation of twenty cosmetic preservatives", *Journal of Chromatography A*,

1029, pp.57-65, 2004

- [16] Hyung-Hoon Kim, Jeong-Ran Cho, “Sign Language Image Recognition System Using Artificial Neural Network”, Journal of The Korea Society of Computer and Information, Vol. 24. No. 2, 2018
- [17] Sin-Hae Hwang, Dong-Young Ku, Jeoung-Kun Kim, “Application of Social Big Data Analysis for CosMedical Cosmetics Marketing : H Company Case Study”, Journal of Digital Convergence, Vol. 17., No. 7, pp.35-41, 2019
- [18] Lim, Hee Suk, Shin, Jae Wook, “A Study on the Trends of Cosmetics through Big Data Analysis-Focusing on text mining and semantic network analysis”, The Korean Society of Illustration Research, Vol. 66. pp.85-95, 2021

## Authors



Hyung-Hoon Kim received the B.S. degrees in Computer Science from Chonnam National University, Korea, in 1986. He received the M.S. degrees in Computer Science from KAIST, Korea, in 1988.

He received the Ph.D. degrees in Computer Science from Hanyang University, Korea, in 2007. Dr. Kim joined the faculty of Kwangju Womens University, Gwangju, Korea, in 1994. He is currently a Professor in the Department of Cosmetic Science at Kwangju Womens University. He is interested in web programming, medical information system, information security, and AI.



Seung-Ho Ohk received the B.S., M.S., and Ph.D. degree in Microbiology from Yonsei University, Seoul, Korea in 1991, 1993 and 1997. From 1997 to 2000, he was a Postdoctoral Researcher with Department of

Oral Biology, SUNY at Buffalo, USA. Since 2003, he has been a Professor with the School of Dentistry, Chonnam National University, Gwangju, Korea. He also was an exchange scholar with the Department of Food Science, Purdue University, USA from 2007 to 2009. Now, he is trying to develop fast and specific ways to detect oral infections with several biochemical substances. Dr. Ohk has been elected as a President of the Korean Academy of Oral Biology from 2021.