

# 기침 소리의 다양한 변환을 통한 코로나19 진단 모델

김민경

(주)데이터월드  
(zizipong@dataworld.co.kr)

김건우

한밭대학교 융합경영학과  
(gkim@hanbat.ac.kr)

최근호

한밭대학교 융합경영학과  
(keunho@hanbat.ac.kr)

2019년 11월 중국 우한시에서 발병한 코로나19는 2020년 중국을 넘어 세계로 퍼져나가 2020년 3월에는 전 세계적으로 확산되었다. 코로나19와 같이 전염성이 강한 바이러스는 예방과 확진시 적극적인 치료도 중요하지만 우선 전파 속도가 빠른 바이러스인 점을 감안할 때, 확진 사실을 재빠르게 파악하여 전파를 차단하는 것이 더욱 중요하다. 그러나 감염 여부를 확인하기 위한 PCR검사는 비용과 시간이 많이 소요되고, 자가키트검사 또한 접근성은 쉽지만 매번 수시로 받기에는 키트의 가격이 부담이 될 수밖에 없는 실정이다. 이러한 상황에서 기침 소리를 기반으로 코로나19 양성 여부를 판단할 수 있게 된다면 누구나 쉽게 언제, 어디서든 확진 여부를 체크할 수 있어 신속성과 경제성 측면에서 큰 장점을 가질 수 있을 것이다. 따라서 본 연구는 기침 소리를 기반으로 코로나19 확진 여부를 식별할 수 있는 분류 모델을 개발하는 것을 목적으로 하였다. 이를 위해, 본 연구에서는 먼저 MFCC, Mel-Spectrogram, Spectral contrast, Spectrogram 등을 통해 기침 소리를 벡터화 하였다. 이 때, 기침 소리의 품질을 위해 SNR을 통해 잡음이 많은 데이터는 삭제하였고, chunk를 통해 음성 파일에서 기침 소리만 추출하였다. 이후, 추출된 기침 소리의 feature를 이용하여 코로나 양성과 음성을 분류하기 위한 모델을 구축하였으며, XGBoost, LightGBM, FCNN 알고리즘을 통해 모델 학습을 수행하고 각 알고리즘별 성능을 비교하였다. 또한, 기침 소리를 다차원 벡터로 변환한 경우와, 이미지로 변환한 경우에 대해 모델 성능에 대한 비교 실험을 수행하였다. 실험 결과, 건강상태에 대한 기본정보와 기침 소리를 MFCC, Mel-Spectrogram, Spectral contrast, 그리고 Spectrogram을 통해 다차원 벡터로 변환한 feature를 모두 활용한 LightGBM 모델이 0.74의 가장 높은 정확도를 보였다.

**주제어** : 코로나19, 기침 소리, LightGBM, MFCC, SNR

논문접수일 : 2023년 4월 25일  
원고유형 : Regular Track

논문수정일 : 2023년 7월 17일  
교신저자 : 최근호

게재확정일 : 2023년 8월 1일

## 1. 서론

2019년 11월 중국 우한시에서 발병한 코로나19는 2020년 중국을 넘어 세계로 퍼져 나가 2020년 3월에는 전 세계적으로 확산되었다. 우리나라도 2020년 2월 감염병 위기경보를 ‘심각’으로 격상하고 ‘코로나 바이러스감염증-19 중앙재난안전대책본부’를 신설하고 대응을 시작하였다. 코로나 팬데믹은 중국에서 한국, 유럽, 미국으로 번져 나갔고 코로나 19는 우리 일상의 풍경부터 경제, 정치 등 모든

분야에 큰 영향을 주고 있다(김가혜, 이소현, 2022; 유소연, 임규건, 2021; 정연길 et al., 2022). 즉, 전 세계의 생활 패턴에 획기적인 변화를 초래하였다. 코로나19의 장기화에 따른 국가 경제 상황 등을 고려하여 현재는 엔데믹 시대를 선언하였지만, 여전히 확진자가 발생하고 있으며 요양병원, 노약자 등 고위험군과 기저질환자 감염시 위중증률이 높아 사망에까지 이르고 있다. 또한, 앞으로 또 다른 신규 질병이 유행할 가능성도 적지 않기 때문에 여전히 코로나19의 광범위한 영향과 진단

법을 연구하고 향후 다른 질병에 응용할 수 있는 기법을 제시하는 것은 중요하다고 할 수 있다.

코로나19는 전파 속도가 매우 빠른 바이러스인 점을 감안할 때, 미리 예방하고, 확진 시 적극적으로 치료하는 것도 중요하지만 확진 사실을 빠르게 파악하여 전파를 차단하는 것이 더욱 중요하다. 코로나19는 감염 시 기침, 인후통, 발열 등의 증상을 동반한다. 코로나19 검출을 위한 표준 방법은 역전사 중합효소 연쇄 반응(RT-PCR) 검사로서, 코로나 19 초기에는 증상 발현자 또는 확진자와 접촉한 자는 PCR검사를 통해 감염 여부를 확인할 수 있었다. 그러나 PCR검사는 비용이 많이 들고, 지정된 선별진료소를 방문하여 검사를 진행하며 결과를 받기까지 최소 몇 시간에서 하루이상 소요된다. 하지만, PCR검사를 할 수 있는 선별진료소와 인력은 한정되어 있고 재유행으로 인해 감염자가 다시 급증할 경우, 의료체계가 무너질 가능성이 있기 때문에 좀 더 신속한 검사를 위해 지정된 병원에서 신속항원검사를 받거나 각자 개인별로 자가 검사키트를 이용해 검사를 진행할 수 있게 되었다. 그러나 전 세계적으로 코로나19가 확산되면서 PCR검사와 자가키트 등 검사 물품 확보에 어려움을 겪기도 했는데, 이처럼 PCR검사나 신속항원 검사는 검사 물품이 확보되어야 하고 수시로 진행하기에는 진료비나 키트의 가격이 부담이 될 수밖에 없는 실정이다. 이에 누구나 쉽게 사용할 수 있도록 대규모로 배포할 수 있으며, 기존 진단 방식의 한계점을 보완할 수 있는 대체 진단 도구가 필요한 실정이다. 코로나19의 여러 증상 중 가장 두드러진 증상은 기침과 호흡 곤란을 포함하는데 감염자의 67.7%가 이러한 증상을 보이는 것으로 나타났다. 기침 소리를 이용한 분류는 최근 호흡기 질환을 빠르게 진단하기 위한 연구 분야로 떠오르고 있다 (Pramono et al., 2016; Alqudaihi et al., 2021).

따라서, 인공지능 기술을 활용하여 기침 소리로 부터 코로나19 양성 여부를 판단하는데 유용한 정보를 얻을 수 있다면 접근성이 용이해 누구든 쉽게 확진 여부를 진단할 수 있게 되어 신속성과 경제성 측면에서 큰 장점을 가질 수 있을 것이다.

사회는 점차 비대면 환경으로 확장되고 있다. 이제 직접 병원에 방문하지 않고 스마트폰, 스마트 워치 등 웨어러블 기기를 통해 질병을 예측하는 것은 점차 확대되어질 것이다. 특히 호흡기 관련 질환은 목소리를 이용한 예측하는 연구가 계속 진행되고 있다. 이에 본 연구는 코로나19 감염 여부를 분류하는 연구를 통해 좀 더 쉽고 빠르게 코로나19 확진 여부를 진단할 수 있는 모델을 개발하고자 한다.

## 2. 선행 연구

소리의 특성을 기반으로 한 예측은 이전부터 다양한 분야에서 활발하게 이루어져 왔다(오상우, 2019). 특히 요즘 기침, 목소리 등 소리로 질환을 예측하는 기술은 여러 나라에서 연구가 진행되고 있다. 폐렴과 같은 호흡기 질환 뿐만 아니라 목소리로 우울, 불안 등 감정을 알아내는 분야의 관심도 높아지고 있다. 선행 연구는 기침 소리를 기반으로 호흡기 질환을 예측하는 연구, 특히 코로나19 감염 여부를 예측하는 연구에 중점을 두어 살펴보았다.

### 2.1. 기침 소리를 이용한 호흡기 질환 진단 연구

사람의 기침이나 잠긴 목소리, 발음으로도 이 사람이 무슨 질병인지를 대략적으로 파악할 수 있는 경우가 많다. 목소리는 폐, 성대, 구강, 입술 등 여러 가지 발성기관을 거치면서 만들어지기 때문에 사람의 목소리 변화를 통해서 그 사람의 건강을 체크할 수 있다. 폐의 호흡이 성대를 진동

시켜 만들어낸 음파가 목과 입술을 지나면서 발음으로 완성된다. 이스라엘의 스타트업 업체인 비온드 버벌사는 음성과 인공지능 모델을 통해 파킨슨 병을 앓고 있는지 여부를 98.6% 정확도로 진단이 가능하고 아동 ADHD 진단을 90% 정확성으로 식별해 낼 수 있다. 또한 이스라엘 스타트업 업체인 보칼리스헬스는 목소리 데이터에서 512가지 특성을 도출해 음성 데이터를 시각화하여 폐질환을 모니터링하는 AI스마트폰 앱 ‘보칼리스체크’를 개발했다. 이 앱은 AI로 환자의 음성 데이터를 분석하여 코로나19, 만성 폐쇄성 폐질환(COPD), 울혈성심부전(CHF), 폐고혈압(PH) 등의 질병 여부를 식별하고 상태를 진단하는 것이 가능하다.

마찬가지로 기침은 무수히 많은 호흡기 질환의 증상이며, 기침의 소리와 유형은 질병을 진단할 때 유용하게 고려할 수 있는 특징이다(Alqudaihi et al., 2021). 기침 소리를 이용한 호흡기 질환 진단에 관한 연구를 살펴보면 다음과 같다.

Kumar et al.(2022)는 기침 소리를 분석하여 10종의 중증 폐질환을 분류하는 방법을 제시하였고, Pramono et al.(2016)는 기침 및 콧바람 소리를 분석하여 백일해를 진단하는 알고리즘을 개발하였는데, 이는 92%의 정확도를 보였다. Amrulloh et al.(2015)는 기침 소리를 통해서 소아의 폐렴과 천식을 진단하는 기법을 연구하였다. 이 방법은 비가우시안성(non-Gaussianity) 등의 음향학적 특성을 추출하여 분석하였으며, 각 질병을 진단할 뿐만 아니라 원격 지역에서도 유용한 도구로 개발할 수 있는 가능성을 제시하였다. 마찬가지로 소아를 대상으로 연구한 Abeyratne et al.(2013)는 폐렴, 기관지염 및 천식과 같은 급성 호흡기 질환을 앓는 91명 소아 환자의 기침 소리를 분석하였다.

코로나19는 폐와 기도에 영향을 미치는 호흡기 질환으로 공기가 호흡기를 통해 이동하는 방식

에 영향을 미칠 수 있으므로 건강한 사람과 환자 사이에 소리의 차이가 발생할 가능성이 있다. 이에 기침 소리를 기반으로 코로나19 확진자를 판별하는 연구들이 진행되어 왔다(Orlandic et al., 2021; Brown et al., 2020; 이예진, 2022; 손명진, 2022; 최혜진, 2022; 유성주 et al., 2022; 이혜빈 et al., 2021). 확진자를 판별하는 모델 구축에 앞서 가장 중요한 것은 코로나19 확진자들의 기침 소리 데이터 확보인데, 기존 연구에서 많이 활용되고 있는 코로나19 기침 데이터는 CAMBRIDGE, COSWARA, COUGHVID 등이 있는데, 이 데이터들은 모두 애플리케이션을 이용한 crowd-sourcing 방식으로 수집되었다. 이 데이터에는 기침 소리, 지역, 성별, 호흡기 질환 여부 등이 기본적인 정보로 구성되어 있다. 기침 소리 학습에 앞서 데이터의 품질을 확보하기 위해 Orlandic et al.(2021)는 기침 감지 머신러닝 모델을 생성하고 기침이 아닌 데이터를 추출 후 제거하였다. 또한 각 기침 데이터의 신호 대 잡음비(SNR) 값을 사용하여 원하는 품질의 데이터만 추출할 수 있었다. 이렇게 필터링된 기침 데이터를 기반으로 음성인식 분야에서 가장 많이 사용되고 있는 MFCC를 활용하여 음성을 벡터화하였고, 기침 소리에서 Spectrogram 이미지를 추출하였다. Brown et al.(2020)은 기침 및 호흡 소리를 분석하여 COVID-19와 천식을 진단하였다. 프레임당 MFCC 계수를 구한 후, 그 값의 각 프레임 간의 차이인  $\Delta$ -MFCC 계수를 구하였으며,  $\Delta$ -MFCC 값의 각 프레임간의 차이인  $\Delta^2$ -MFCC 계수도 구해 모델에 활용하였다. Ahmed et al.(2021)는 기침 소리에서 추출한 spectrogram 이미지를 기반으로 ResNet-50 알고리즘을 활용하여 모델을 훈련하였다. Imran et. al.(2020)는 AI4COVID-19이라는 애플리케이션을 개발하여 코로나19를 진단하는 기법을 제시하였다. 이는 3초

동안의 기침 소리 3개를 클라우드에서 실행되는 AI 엔진에 기록하고 전송한 후, 2분 이내에 결과를 도출하는 방식을 적용하였다.

하지만, 기존의 연구들은 양성판정자의 데이터 건수가 매우 적었으며, 데이터 균형화 작업을 수행하지 않고 모델을 학습한 경우가 많아 모델의 정확성에 대한 신뢰도 확보 측면에서 다소 부족한 모습을 보였다. 또한, 기침 소리를 활용한 분석 측면에서 다양한 방법들에 대한 비교가 없이 단일 또는 몇 가지 방법만을 이용하여 모델을

개발하였다는 점에서, 보다 많은 양성판정자 데이터를 이용하여 다양한 방법으로 기침 소리를 변환한 후, 모델의 정확도를 높일 수 있는 최적의 조합을 찾고자 하는 본 연구와 차별화된 모습을 갖는다. 기침 소리를 이용한 호흡기 질환 진단 연구들을 요약하여 정리하면, <Table 1>과 같다.

호흡기 질환은 기도 상태의 변화로 기침 소리가 변한다. 이런 기침 소리의 차이를 판별하기 위해서는 소리의 특성 추출이 중요하다. 소리의 특성을 추출하기 위해 SNR, MFCC, Spectrogram

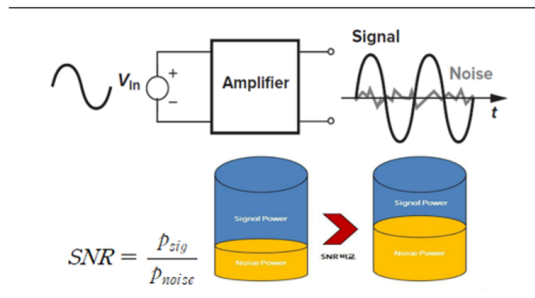
<Table 1> Summary of literature review

Authors	Purpose	Data	Methods	Results & Implications
Kumar et al. (2022)	Analyzing Cough Sounds for Classification of 10 Severe Respiratory Diseases	Vitals signs of 108 Indian Patients	Tree Ensemble Model	Provide a model for classifying the top 10 serious respiratory diseases commonly found among Indian teenagers
Pramono et al. (2016)	Development of Diagnostic Technique for Pertussis Based on Cough and Nasal Sound Analysis	Audio recordings of 38 patients	Mel Frequency Cepstral Coefficients (MFCCs), Logistic Regression Model	Automatically detects with 92% accuracy and a positive predictive value of 97%. Useful for rapid early diagnosis and infection control
Amrulloh et al. (2015)	Development of Diagnostic Technique for Pneumonia and Asthma through Cough Sound Analysis	18 pediatric patients at Sardjito Hospital in Indonesia	Artificial Neural Network Classifiers	Attained sensitivity of 89%, specificity of 100%, and a Kappa ( $\kappa$ ) value of 0.89. Presenting the possibility of developing a diagnostic method to differentiate between pneumonia and asthma
Imran et al. (2020)	Developing diagnostic techniques for COVID-19 through an application called AI4COVID-19	1,838 cough sounds and 3,597 non-cough sounds data 96 cases of bronchitis, 130 cases of whooping cough, 70 cases of COVID-19, 247 samples of normal cough sounds data	Deep Transfer Learning-based Multi Class classifier (DTL-MC), Classical Machine Learning-based Multi Class classifier (CML-MC), Deep Transfer Learning-based Binary Class classifier (DTL-BC)	AI4COVID-19 application provides a quick, fast, and accurate diagnostic tool to differentiate COVID-19 cough from various other types of non-COVID-19 coughs
Brown et al. (2020)	Diagnosing COVID-19 and asthma through analysis of cough and respiratory sounds	Data of coughing and respiratory sounds from 141 COVID-19 infected individuals and 350 non-infected individuals	Logistic Regression, Gradient Boosting Trees, Support Vector Machines (SVMs)	Distinguishing the cough sounds of COVID-19 and asthma from cough sounds of other diseases. Proposing the potential for medical applications in cough and respiratory sound analysis

등 소리의 특성값을 추출하는 방식에 대한 선행 연구들을 살펴보았다.

### 2.1.1. SNR(Signal to Noise Ratio)

SNR은 신호대 잡음비로 신호와 노이즈 비율을 알고 싶을 때 사용하는 물리량이다. 오디오 애플리케이션에 대한 신호 대 잡음비(SNR)을 계산하는 방법으로 많이 활용되고 있다. <Figure 1>과 같이, 정보를 지닌 신호의 세기를 잡음의 세기로서 나눈 값으로 신호에 동반하는 잡음 성분의 양을 수치로 표시한 것으로 신호(Signal, S)와 잡음(Noises, N)의 비율을 표시한다. 단위는 dB이고 수치가 클수록 노이즈가 작다는 것을 의미한다. 선행연구로는 Mishra & Sahu(2011)의 압의 유형별 진단, Gaurav et al.(2021) 및 Rezaee et al.(2022)의 파킨슨병 관련 연구 등이 있다.



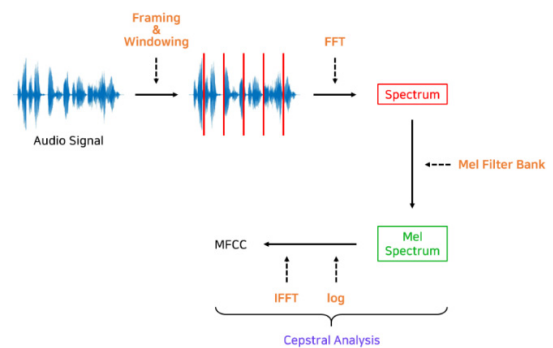
<Figure 1> An example of SNR

### 2.1.2. MFCC(Mel-Frequency-Cepstral-Coefficient)

MFCC는 오디오 신호에서 추출할 수 있는 feature로 소리의 고유한 특징을 나타내는 수치이다. 주로 음성인식, 화자인식, 음성 합성, 음악장르 분류 등 오디오 도메인의 문제를 해결하는데 가장 널리

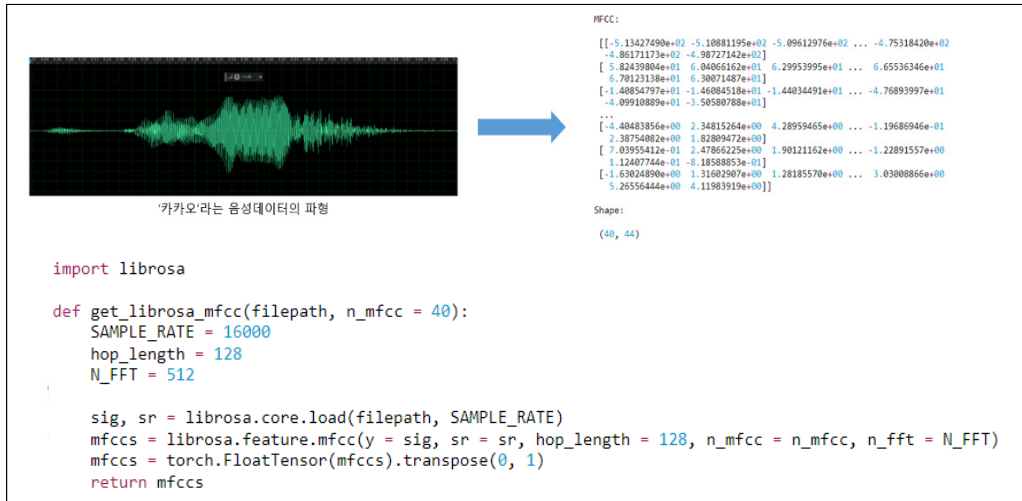
사용되고 있다. Muda et al. (2010)는 MFCC와 DTW 기법을 이용한 목소리 인식 알고리즘을 제안하였고, Tiwari(2010)는 화자 인식 시스템을 설계하기 위해 MFCC를 통해 추출한 feature를 활용하였다. Ittichaichareon et al.(2012)와 Dave(2013)는 음성을 텍스트로 변환해주는 음성 인식에 MFCC 기법을 활용하였다. 의학적으로는 Sadi & Hassan (2020)이 MFCC 및 CNN 모델을 결합하여 폐 진단 기법을 개발하였고, Doulah & Fattah(2014)의 신경근육질환 진단법을 제시한 연구가 있다.

MFCC 추출과정은 <Figure 2>와 같다. 첫째, 오디오 신호를 프레임별(보통 20ms~40ms)로 나누어 FFT(푸리에 변환)를 적용해 스펙트럼을 구한다. 둘째, 이 스펙트럼에 Mel Filter Bank를 적용하여 Mel Spectrum을 구한다. 마지막으로, Mel Spectrum에 log와 역푸리에 변환을 해주면 MFCC값이 추출된다. MFCC는 음성데이터를 특징벡터(Feature)화해주는 알고리즘이다. 머신러닝에서 음성데이터는 벡터화를 통해 학습이 가능해진다. 이때 데이터에서 Feature를 어떤 방법으로 추출하느냐에 따라 모델의 성능이 달라지므로 이는 굉장히 중요한 과정이다.

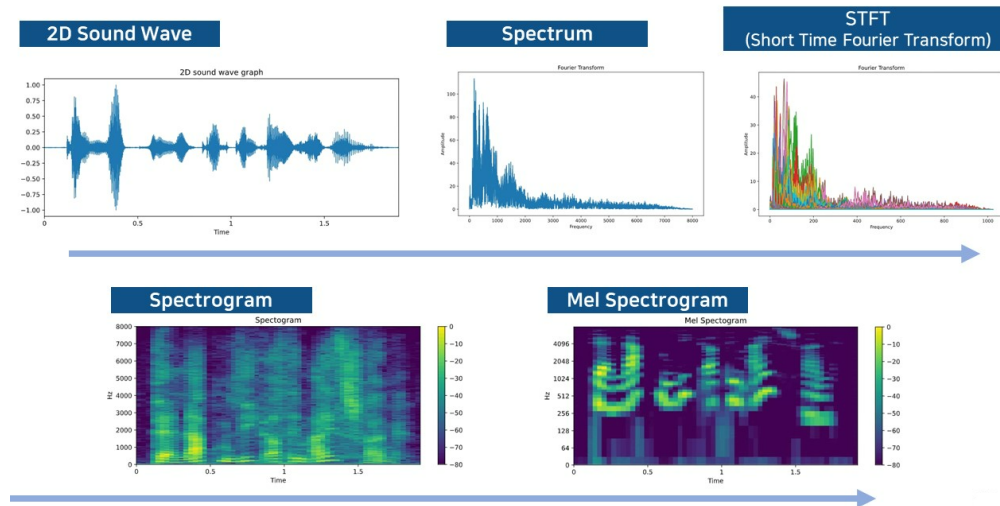


<Figure 2> MFCC extraction process (Source: haythamfayek<sup>1)</sup>)

1) <https://haythamfayek.com/2016/04/21/speech-processing-for-machine-learning.html>



<Figure 3> An example of librosa library



<Figure 4> Process of Spectrogram Transformation (Source: 6 o'clock liberal arts<sup>2</sup>)

MFCC Feature는 <Figure 3>과 같이 python에서 제공되는 librosa 라이브러리를 이용해서 추출이 가능하다.

### 2.1.3. Spectrogram

Spectrogram은 음성신호 처리에서 아주 기본적인 feature로 활용되고 있다. <Figure 4>와 같이, spectrogram은 소리나 파동을 시각화하여 파악하기

2) <https://blog.naver.com/hannaurora/222412280068>

위한 도구로 파형(waveform)과 스펙트럼(spectrum)의 특징이 조합되어 있다. Spectrogram은 소리의 스펙트럼을 시각화하여 그래프로 표현하는 기법으로 시간상 진폭축의 변화를 시각적으로 볼 수 있는 파형과 주파수상 진폭 축의 변화를 시각적으로 볼 수 있는 스펙트럼의 특징이 모두 결합된 구조라고 볼 수 있다. 선행연구로는 Jana et al.(2019) 및 Jana & Nath(2022)의 말초동맥질환 진단과 Tadesse et al.(2020)의 멀티모달 질병 진단법 개발 등이 있다.

#### 2.1.4. Spectral contrast

Spectral contrast은 스펙트럼 피크, 스펙트럼 밸리 및 각 주파수 하위대역에서의 차이를 고려한다. 최상위 분위수의 평균 에너지(피크 에너지)를 하단 분위수(밸리 에너지)의 평균 에너지와 비교하여 추정하며, 고대비 값은 선명한 협대역 신호에 해당하고 낮은 대비 값은 광대역 잡음에 해당된다. 최근 연구에서는 물리적 환경에서 발생하는 소리나 음성 분석에서 이 기법을 활용한 바 있다(Issa et al., 2020; Singh & Biswas, 2022; Su et al., 2019).

## 2.2. 분류 알고리즘

### 2.2.1. XGBoost

XGBoost는 Extreme Gradient Boosting의 약자로 여러 개의 Decision Tree를 조합해서 사용하는 Ensemble 알고리즘 중 하나이다. Boosting 기법을 이용하여 구현한 가장 대표적인 알고리즘인 Gradient Boost를 병렬 학습이 지원되도록 구현한 알고리즘이 XGBoost이다(Chen et al., 2016). Ogunleye & Wang(2020)은 XGBoost를 활용하여 만성신부전증을 진단하는 모델을 개발하였고,

Dhaliwal et al.(2018)은 효과적인 네트워크 침입 탐지 시스템을 개발하기 위해 XGBoost를 사용하였다.

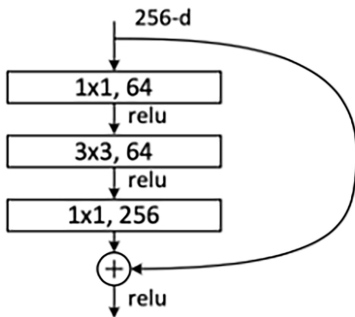
### 2.2.2. LightGBM

LightGBM은 XGBoost의 단점을 개선한 알고리즘으로 XGBoost의 속도와 Overfitting 문제를 개선한 방법론이다. LightGBM의 장점은 빠른 속도와 정확도이다. LightGBM은 Gradient Boosting 프레임워크로 Tree 기반 학습 알고리즘이다. 기존의 tree기반 알고리즘은 트리의 수평적 확장인 level wise(균형 트리 분할) 방식을 사용했다. 이는 트리의 균형을 잡기위해 분할시 트리의 깊이는 최소화되지만 트리의 균형을 맞추기 위해 시간이 소요된다는 의미이다. LightGBM은 트리의 수직적 확장 방식인 leaf wise tree분할 방식이다. tree의 균형을 고려하지 않고 최대 손실 값(max data loss)를 가지는 leaf node를 지속적으로 분할하면서 tree의 깊이가 깊어지고 비대칭적인 tree가 생성된다. 이와 같이 최대 손실값을 가지는 leaf node를 반복 분할하는 방식은 level wise tree분할 방식보다 예측 오류 손실을 최소화할 수 있다(Ke et al., 2017). Daoud(2019)는 대출금 상환 여부를 예측하기 위해, LightGBM, XGBoost, CatBoost를 활용한 모델을 개발하였는데, LightGBM이 가장 우수한 성능을 보였다. Wang et al.(2017)은 유방암 환자에서 miRNA를 식별하기 위해 LightGBM을 활용한 모델을 개발하였다.

### 2.2.3. ResNet-50

ResNet은 Residual Learning 이라는 개념을 통해 모델의 층이 깊어져도 학습이 잘되도록 구현한 모델로 Convolution Layer(1X1) 2개를 사용하기

때문에 파라미터 수가 감소하여 연산량이 줄어들었다. 또한 Layer가 많아짐에 따라 Activation Function이 증가하여 더 많은 non-linearity가 들어갔다. ResNet은 Skip Connection을 이용한 Shortcut과 Bottleneck 구조를 이용하여 더 깊게 층을 쌓을 수 있다. Wen et al.(2020)은 이미지 데이터를 활용한 고장 진단 모델을 개발하기 위해 ResNet-50을 활용하였고, Rezende et al.(2017)은 악성소프트웨어 분류를 위해 ResNet-50을 활용하였다.



<Figure 5> A residual block(He et al., 2015)

### 3. 연구방법

#### 3.1. 데이터 및 전처리

##### 3.1.1. 학습 데이터

본 논문에서 실험을 위해 사용한 데이터셋은 DACON에서 제공하는 코로나19 검사자료로 검사자의 음향 데이터(기침 소리: 오디오 파일)와 건강상태에 대한 기본정보(나이, 성별, 호흡기 질환 여부, 발열, 근육통 증상 여부)로 구성되어 있다.

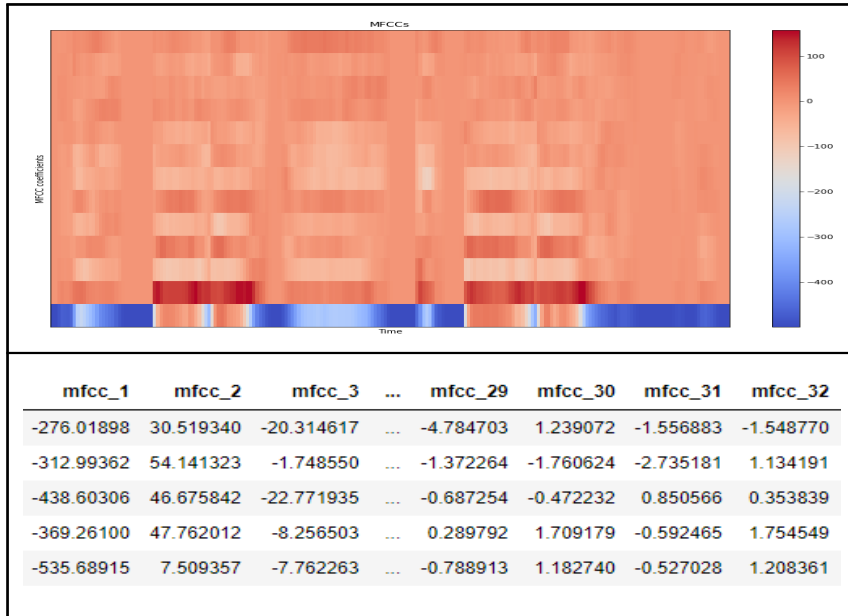
DACON은 국내 최초의 인공지능 경진대회 플랫폼으로 인공지능 분석을 위한 다양한 데이터를 제공하고 경진대회를 개최하고 있다. DACON에서

제공하는 기침 소리 데이터는 총 3,805건으로 성별은 남성이 67.2%, 여성이 31.8%의 비율을 보였고, 83.5%가 기존에 호흡기 질환을 가지고 있지 않았으며, 코로나 음성판정 3,499건과 양성판정 306건으로 구성되어 있어 Target Label인 검사 결과는 음성판정과 양성판정의 비율이 굉장히 불균형한 구조였다. 본 연구에서는 양성판정 데이터의 확보를 위해 COUGHVID 데이터 셋에서 양성 데이터 441건을 추가하여 총 4,246건으로 최종 데이터셋을 구성하였다. COUGHVID는 Orlandic et al.(2021)가 구축한 대규모 코로나19 관련 기침 소리 데이터셋으로, 2020년 4월 1일부터 2020년 12월1일까지 온라인을 통해 전세계 125개국에서 기침 소리 데이터를 수집하였다. COUGHVID 기침 소리 데이터셋의 성별 비율은 남성이 66.0%, 여성이 33.5%였고, 83.6%가 기존에 호흡기 질환을 가지고 있지 않았다.

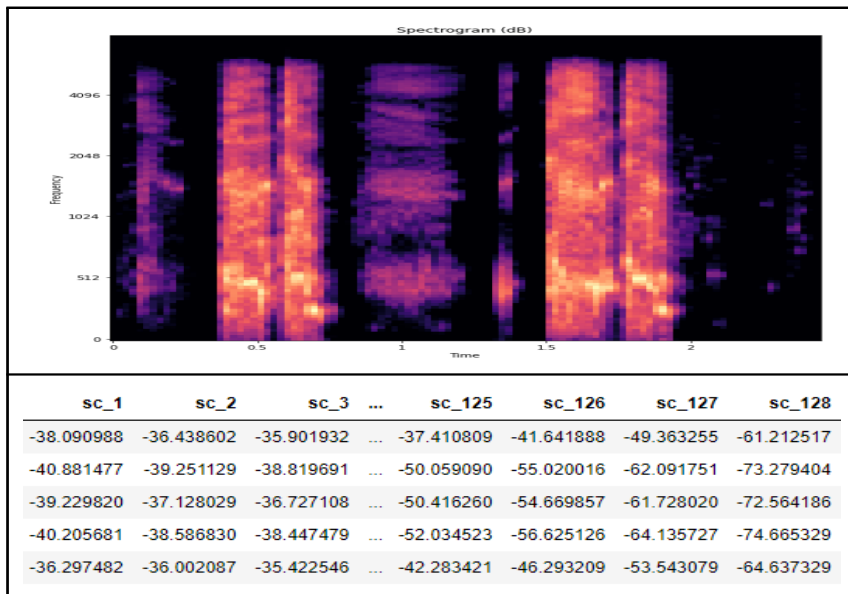
#### 3.1.2. 데이터 전처리

본 논문의 실험에서는 음성 데이터 품질을 위해 기침 소리의 SNR(신호대 잡음비)를 구하고 SNR 계수가 0, 15, 25 이상인 데이터를 추출하여 사용하는 방식으로 활용하였다. 그 후 기침 소리인 음성데이터를 활용하기 위해 MFCC를 통해 벡터화하여 32개의 변수를 추출하였다. 또 음성데이터를 Mel-Spectrogram, Spectral contrast, Spectrogram을 통해 feature로 추출하였다(McFee et al., 2015). 또한, 음성데이터 원본을 그대로 변수화하는 방법과 음성데이터에서 기침 소리만 잘라 변수화하는 방법(chunk, segmentation)의 두가지 방법으로 feature를 추출하였다. <Figure 6> ~ <Figure 10>은 MFCC, Mel-Spectrogram, Spectral contrast, chunk, segmentation에 대한 예시이다.

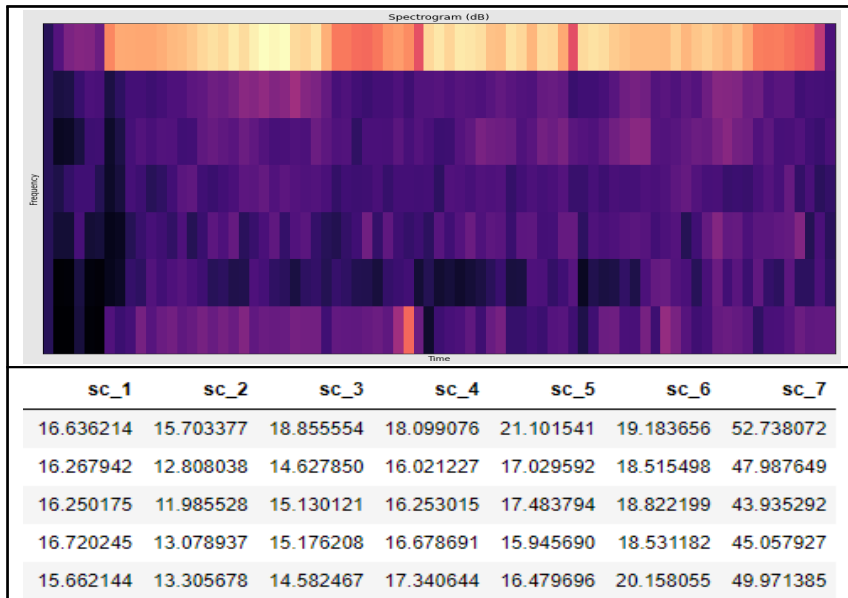




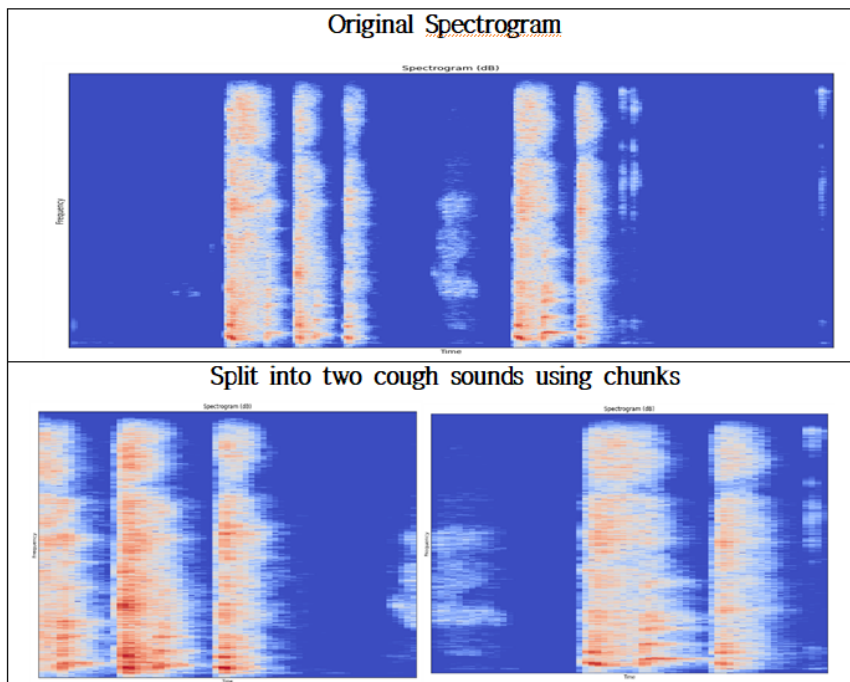
〈Figure 6〉 An example of MFCC



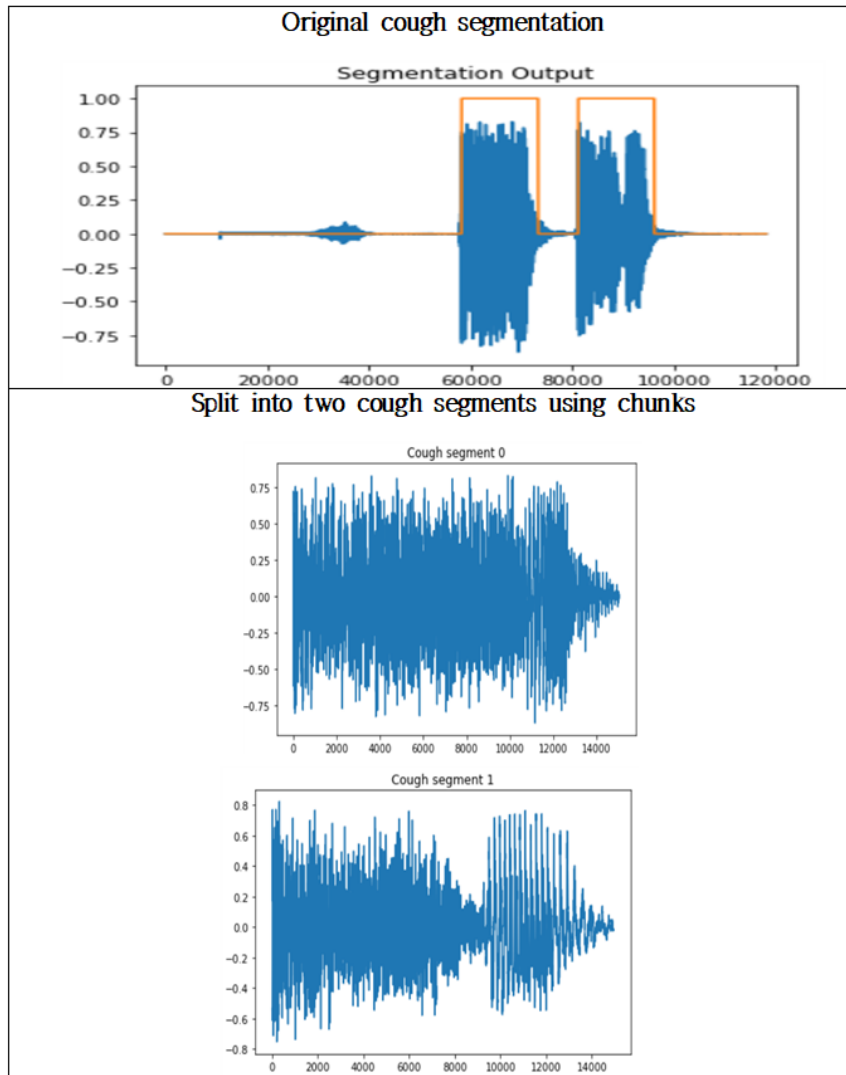
〈Figure 7〉 An example of Mel-Spectrogram



〈Figure 8〉 An example of Spectral contrast



〈Figure 9〉 An example of partitioning of audio chunk



〈Figure 10〉 An example of segmentation

최종 변수는 건강상태에 대한 기본정보와 MFCC 32개 feature, Mel-Spectrogram 128개 feature, Spectral contrast 7개 feature, Spectrogram 1,025개 feature로 구성하여 각각 feature별로 사용, 그리고 결합하여 사용하였다. 데이터 품질을 위해 SNR 계수값으로 필터를 주었다.

위와 같이, 본 연구에서는 기침 소리를 다차원 벡터로 변환하여 feature를 추출하였고, 연령, 성별 등 건강상태에 대한 기본정보와 다차원 벡터를 기반으로 분류에서 많이 활용되고 있는 XGBoost, LightGBM, FCNN 알고리즘을 사용하여 학습결과를 확인하였다. XGBoost 등 분류 알고리즘은

under-sampling을 통해 target label 비율을 1:1로 맞추어 학습을 진행하였다. 기침 소리를 이미지로 변환한 경우, ResNet-50 알고리즘을 활용해 모델을 학습하였다.

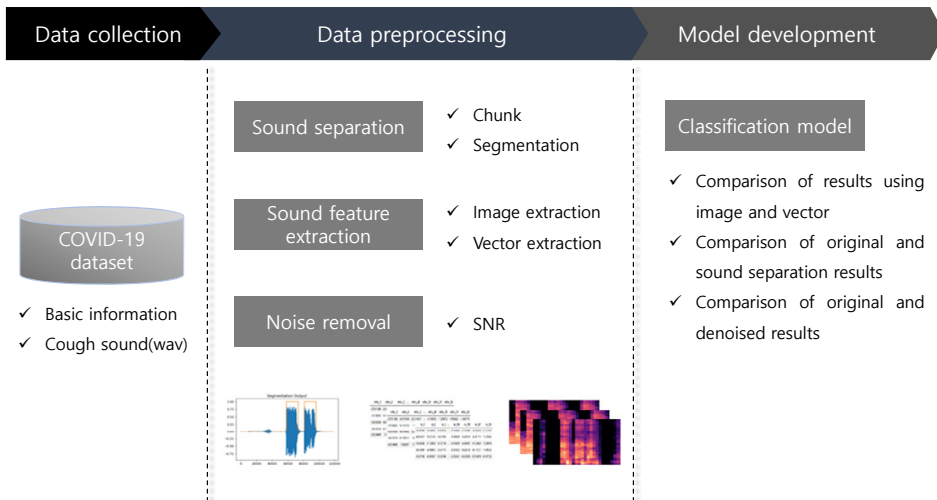
### 3.1.3. 연구 아키텍처

본 연구에서는 기침 소리에서 feature 추출 시, 다차원 벡터로 변환하는 방법과 이미지로 변환하는 방법을 수행하였다. 이에 연령, 성별 등의 기본정보 feature와 다차원 벡터를 기반으로 분류 문제에서 많이 활용되고 있는 XGBoost, LightGBM, FCNN 알고리즘을 사용하여 학습결

과를 확인하였고, 이미지는 ResNet50 알고리즘을 활용해 학습하였다. XGBoost 등 분류 알고리즘은 under-sampling을 통해 Target Label 비율을 1:1로 맞추어 학습을 진행하였고, 학습용 데이터와 검증용 데이터는 7:3의 비율로 분할하였다.

### 3.2. 모델 생성

본 연구에서는 전처리에서 추출된 다양한 feature를 기반으로 알고리즘별 예측 모델을 수행하였다. <Figure 12>는 본 연구에서 사용된 각 알고리즘의 하이퍼파라미터를 보여준다.



<Figure 11> Research architecture

XGBoost	LightGBM	FCNN	ResNet-50
n_estimators=650 learning_rate=0.01 max_depth=20	max_depth = 10 min_child_samples=45 subsample=0.8 num_leaves=30 n_estimators=300 learning_rate=0.1	learning_rate=0.001 batch_size=32 epochs=100 activation='relu' kernel_initializer='glorot_uniform' kernel_regularizer=l2(0.01) bias_regularizer=l2(0.01)	include_top=True weights=None input_shape=(64, 64, 3)

<Figure 12> Hyper-parameter values for each algorithm

### 3.2.1. 기본정보를 활용한 모델

#### 1) 분류 알고리즘

학습데이터 중 기본정보인 나이, 성별, 호흡기 질환 여부, 발열 및 근육통 증상 여부 항목만 활용하여 학습을 수행하였다. 알고리즘은 분류에 많이 활용되고 있는 XGBoost, LightGBM, FCNN을 사용하여 학습을 진행하였다. 학습데이터는 검사결과 중 양성 데이터 건수에 맞추어 under-sampling을 수행 후 진행하였다. 기본정보만 활용한 학습 결과는 <Table 2>와 같다.

<Table 2> Comparisons of results using basic information

Algorithm	Dataset	Accuracy
XGBoost	Basic information	0.67
LightGBM	Basic information	0.66
FCNN	Basic information	0.66

### 3.2.2. 기본정보와 기침 소리를 활용한 모델

#### 1) 기침 소리를 벡터로 변환

본 연구에서는 코로나 진단에 대한 기침 소리의 효과를 평가하기 위해, 기본정보에 기침 소리를 추가하여 학습을 수행하였다. 먼저, 기침 소리를 벡터로 변환했을 때의 효과를 평가하기 위해, 기침 소리를 MFCC를 통해 벡터화하여 32개 변수를 추출하였다. 이후, 연령 등의 기본정보 4개 feature와 MFCC 32개 feature로 총 36개의 feature를 사용하여 학습하였다. 학습결과는 <Table 3>과 같다. LightGBM이 가장 좋은 성능을 보여 주었으며, 기본정보만 활용하였을 경우와 기침 소리를 추가하여 활용하였을 경우의 결과 성능을 비교해 보면, 기침 소리를 활용한 경우 분류

성능이 조금 더 높음을 확인하였다. 이에 기침 소리 변수를 좀 더 다양한 방법으로 활용하여 실험을 진행하였다.

<Table 3> Comparisons of results using basic information and MFCC features

Algorithm	Dataset	Accuracy
XGBoost	Basic information + MFCC32	0.67
LightGBM	Basic information + MFCC32	0.69
FCNN	Basic information + MFCC32	0.63

#### 2) 기침 소리를 이미지로 변환

이어서 본 연구에서는 기침 소리를 이미지로 변환했을 때의 효과를 평가하기 위해, 기침 소리를 Mel-Spectrogram, Spectrogram을 통해 이미지로 변환하였다. 기본정보와 기침 소리에서 추출한 이미지를 기반으로 ResNet-50 알고리즘으로 학습을 수행한 결과는 <Table 4>와 같다. 기침 소리를 이미지로 변환 후 사용한 경우 벡터로 변환하여 사용한 경우보다 분류 정확도가 현저히 낮게 나타났다.

<Table 4> Comparison of ResNet-50 results using image data

Dataset	Accuracy
Spectrogram	0.57
Mel-Spectrogram	0.48

#### 3) MFCC feature와 SNR 필터 적용

이번 실험에서는 기본정보와 MFCC 32개 feature에서 SNR 필터를 적용하였다. SNR이 0이하 건인 699건을 제거, SNR이 15미만인 2,483건 제거,

SNR이 25미만인 3,413건을 제거하며 학습을 수행하였다. 학습결과는 <Table 5>와 같다. SNR이 15미만인 경우를 제거하였을 경우가 0.71로 가장 좋은 성능을 보였다.

<Table 5> Comparison of LightGBM results using Basic information, SNR, and MFCC32

Dataset	Accuracy
Basic information + MFCC 32 + SNR(>0)	0.69
Basic information + MFCC 32 + SNR(>=15)	0.71
Basic information + MFCC3 2 + SNR(>=25)	0.61

4) Chunk로 기침구간 분할 후 MFCC feature 활용

이번에는 chunk로 기침을 구간별로 분할하여 대상 데이터 4,246건을 총 14,484건으로 생성하고 분할한 데이터로 MFCC를 통해 32개 feature로 추출 후 LightGBM 알고리즘으로 학습을 수행하였다. 기본정보와 MFCC 32개 feature를 사용한 경우, MFCC 32개 feature에 SNR 또는 Spectral flatness로 필터를 적용한 경우를 살펴보았다. 학습결과는 <Table 6>과 같다. Chunk로 기침구간을 분할한 경우, SNR을 적용하지 않을 때, 0.72로 가장 좋은 성능을 보였다.

<Table 6> Comparison of LightGBM results using basic information, SNR, and MFCC32(chunk)

Dataset	Accuracy
Basic information + MFCC32(chunk)	0.72
Basic information + MFCC32(chunk) + SNR(>0)	0.71
Basic information + MFCC32(chunk) + SNR(>=15)	0.69
Basic information + MFCC32(chunk) + SNR(>=25)	0.63

5) Chunk로 기침구간 분할 후 Mel-Spectrogram feature 활용

이번 실험에서는 기본정보와 chunk로 분할한 데이터로 Mel-Spectrogram을 통해 128개 feature를 추출한 후, LightGBM 알고리즘으로 학습을 수행하였다. SNR로 필터를 적용한 결과도 살펴 보았다. 학습결과는 <Table 7>과 같다. Chunk로 기침구간을 분할한 경우, SNR을 적용하지 않을 때, 0.7로 가장 좋은 성능을 보였다.

<Table 7> Comparison of LightGBM results using basic information, SNR, and Mel-Spectrogram(chunk)

Dataset	Accuracy
Basic information + Mel-Spectrogram128(chunk)	0.7
Basic information + Mel-Spectrogram128(chunk) + SNR(>0)	0.66
Basic information + Mel-Spectrogram128(chunk) + SNR(>=15)	0.63

6) Chunk로 기침구간 분할 후 Spectrogram feature 활용

이번 실험은 기본정보와 chunk로 분할한 데이터로 Spectrogram을 통해 1,025개 feature를 추출한 대상 데이터를 LightGBM 알고리즘으로 학습을 수행하였다. 학습결과는 <Table 8>과 같다.

<Table 8> Comparison of LightGBM results using basic information and Spectrogram(chunk)

Dataset	Accuracy
Basic information + Spectrogram1025(chunk)	0.7

7) Chunk로 기침구간 분할 후 Spectral contrast feature 활용

이번 실험은 기본정보와 chunk로 분할한 데이

터로 Spectral contrast를 통해 7개 feature를 추출한 후, LightGBM 알고리즘으로 학습을 수행하였다. 학습결과는 <Table 9>와 같다.

<Table 9> Comparison of LightGBM results using basic information and Spectral contrast(chunk)

Dataset	Accuracy
Basic information + Spectral contrast7(chunk)	0.68

8) Chunk로 기침구간 분할 후 다양한 feature 결합

마지막으로 앞 실험에서 추출한 기침 소리 관련 feature를 결합 후 학습을 진행하였다. 기본정보와 chunk 기반의 MFCC 32개 feature, Mel-Spectrogram 128개 feature, Spectral contrast 7개 feature, Spectrogram 1025개 feature를 결합 후 LightGBM 알고리즘으로 학습을 수행하였다.

먼저 이전 실험에서 좋은 성능을 보였던 MFCC, Mel-Spectrogram feature를 결합 후 기본정보와 결합한 feature를 기반으로 학습하였다. 학습결과는 <Table 10>과 같다.

<Table 10> Comparison of LightGBM results using basic information, MFCC32(chunk), and Mel-Spectrogram(chunk)

Dataset	Accuracy
Basic information + MFCC32(chunk) + Mel-Spectrogram128(chunk)	0.72

<Table 11> Comparison of LightGBM results using basic information and all feature from cough sound

Dataset	Precision	Recall	Accuracy
Basic information + all features from cough sound(chunk)	0.76	0.72	0.74
Basic information + all features from cough sound(chunk)+ SNR(>0)	0.72	0.66	0.70
Basic information + all features from cough sound(chunk)+ SNR(>=15)	0.68	0.69	0.69

다음으로 기본정보와 모든 기침 소리 feature를 결합한 후 학습을 수행하였다. 즉, 소리변수는 MFCC32, Mel-Spectrogram128, Spectral contrast7, Spectrogram1025 벡터값을 모두 활용하였다. 학습결과는 <Table 11>과 같다.

본 연구에서 수행한 결과를 살펴보면 기본정보와 chunk로 기침을 구간별로 분할한 데이터로 MFCC 32개 feature, Mel-Spectrogram 128개 feature, Spectral contrast 7개 feature, Spectrogram 1025개를 모두 결합한 후 LightGBM 알고리즘으로 학습을 수행한 경우가 정확도 0.74로 가장 좋은 성능을 보였다.

이상의 실험 결과들을 모두 요약하여 정리하면 <Table 12>와 같다.

3.3. 분석 결과에 대한 논의

본 연구의 실험 결과를 종합해보면 다음과 같다. 첫째, 성별, 나이, 호흡기 질환 여부, 발열, 근육통 증상 여부 등과 같은 기본 정보 이외에 기침 소리 데이터를 함께 사용할 경우, 코로나19 진단 정확도가 7%p 향상되는 것으로 나타났고, 이를 통해 코로나19 진단 시, 기침 소리의 활용 가능성을 확인하였다. 둘째, 음성 데이터에서 기침 소리만 잘라 변수화 하는 것은 모델의 정확도 개선에 좋은 영향을 미친 반면, 잡음의 비율은 정확도 개선에 큰 영향을 미치지 못한다는 것을 확인할 수 있었다. 셋째, 기침 소리를 다차원 벡

〈Table 12〉 Summary of experimental results

Type of Cough Sound	Dataset	Algorithm			
		ResNet-50	LightGBM	XGBoost	FCNN
	Basic information	-	0.66	0.67	0.66
Image	Basic information+Spectrogram	0.57	-	-	-
	Basic information+Mel-Spectrogram	0.48	-	-	-
Vector	Basic information+MFCC32	-	0.69	0.67	0.63
	Basic information+MFCC32+SNR(>0)	-	0.69	-	-
	Basic information+MFCC32+SNR(>=15)	-	0.71	-	-
	Basic information+MFCC32+SNR(>=25)	-	0.61	-	-
	Basic information+MFCC32(chunk)	-	0.72	-	-
	Basic information+MFCC32(chunk)+SNR(>0)	-	0.71	-	-
	Basic information+MFCC32(chunk)+SNR(>=15)	-	0.69	-	-
	Basic information+MFCC32(chunk)+SNR(>=25)	-	0.63	-	-
	Basic information+Mel-Spectrogram128(chunk)	-	0.7	-	-
	Basic information+Mel-Spectrogram128(chunk)+SNR(>0)	-	0.66	-	-
	Basic information+Mel-Spectrogram128(chunk)+SNR(>=15)	-	0.63	-	-
	Basic information+Spectrogram1025(chunk)	-	0.7	-	-
	Basic information+Spectral contrast7(chunk)	-	0.68	-	-
	Basic information+MFCC32(chunk)+Mel-Spectrogram128(chunk)	-	0.72	-	-
	Basic information+MFCC32(chunk)+Mel-Spectrogram128(chunk)+Spectrogram1025(chunk)+Spectral contrast7(chunk)	-	<b>0.74</b>	-	-
	Basic information+MFCC32(chunk)+Mel-Spectrogram128(chunk)+Spectrogram1025(chunk)+Spectral contrast7(chunk)+SNR(>0)	-	0.70	-	-
Basic information+MFCC32(chunk)+Mel-Spectrogram128(chunk)+Spectrogram1025(chunk)+Spectral contrast7(chunk)+SNR(>=15)	-	0.69	-	-	

터로 변환하여 활용하는 것이 기침 소리를 이미지로 변환하여 활용하는 것 보다 더 좋은 성능을 보여준다는 것을 확인하였다. 마지막으로, 다양한 방법을 통해 기침 소리로부터 추출한 feature를 모두 함께 사용하는 것이 기침 소리의 다양한 측면을 반영할 수 있어 가장 우수한 정확도를 보

이는 것을 확인하였다.

앞서 언급한 바와 같이, 선행 연구들은 양성판정자의 데이터 건수가 매우 적었고, 데이터 균형화 작업을 수행하지 않고 모델을 학습한 경우가 많아, 본 연구의 분석 결과와 선행 연구의 결과를 직접 비교하는 데에는 어려움이 있다고 판단



된다. 다만, 본 연구는 보다 많은 양성판정자 데이터를 이용하여 다양한 방법으로 기침 소리를 변환한 후, 모델의 정확도를 높일 수 있는 최적의 조합을 찾았다는 점에서 선행 연구들과 차이가 있다.

#### 4. 결론

본 연구에서는 기침 소리 데이터를 이용해 코로나19 확진 여부를 판별하기 위해 기침 소리 데이터에 대한 전처리를 통해 feature를 추출하고 다양한 알고리즘을 통해 예측 모델을 개발하였다. 예측 모델의 성능 향상을 위해 MFCC, Spectrogram, Mel-Spectrogram, Spectral contrast 등 다양한 방법을 통해 다차원벡터로 feature를 추출하였다. 이렇게 추출된 feature를 이용하여 지도학습 기반의 다양한 분류 알고리즘을 활용한 코로나 확진 여부 진단 모델을 개발하였고 각 모델의 성능을 비교하였다. 그 결과 건강상태에 대한 기본정보와 다양한 방법으로 변환한 기침 소리 정보를 모두 활용한 LightGBM 모델이 가장 높은 0.74의 정확도를 보였다.

본 연구는 다음과 같은 학문적 시사점을 지닌다. 먼저, 본 연구는 기침 소리를 다양한 방법으로 변환하고, 각 변환 방법의 분류 정확도를 제시하였다. 이를 통해, 음성데이터에서 기침 소리만 잘라 변수화 하는 것은 성능 개선에 좋은 영향을 미치나, 잡음의 비율은 성능 개선에 큰 영향을 미치지 못한다는 것을 제시하였다. 둘째, 모델의 분류 성능을 높이기 위해, 변환 방법들의 다양한 조합별 실험을 진행하여 최적의 조합을 제안하였다. 이는 향후 모델의 정확도 개선을 위한 추가 연구에 참고자료로 활용될 수 있을 것으로 기대된다. 셋째, 기침 소리를 다차원 벡터로 변환한 경우와 이미지로 변환한 경우에 따른 모델의

성능 비교를 통해, 코로나 확진 여부 진단의 경우 기침 소리를 이미지로 변환하기 보다는 다차원 벡터로 변환하여 활용하는 것이 보다 좋은 성능을 얻을 수 있다는 것을 제안하였다. 넷째, 다양한 분류 알고리즘을 이용한 모델 개발을 통해, 분류 알고리즘별 성능을 비교하였고, LightGBM이 다른 분류 알고리즘들에 비해 높은 분류 정확도를 보인다는 것을 제시하였다. 마지막으로 보다 많은 양성판정자 데이터를 이용하고, 데이터 균형을 통해 모델을 개발하고 평가함으로써, 모델의 분류 정확성에 대한 신뢰도를 제고하였다.

코로나19는 이제 감기처럼 토착화되어가는 분위기이다. 코로나19는 비말 등에 의해 전파가 빠른 바이러스로 확진 시 빠른 격리나 마스크 착용 등의 조치가 필요하다. 그러므로 일상에서 코로나19 의심 증상이 있을 경우 빠르고 쉽게 접할 수 있는 검사는 분명 필요하다. 이에 본 연구는 다음과 같은 실무적 시사점을 지닌다. 첫째, 코로나19 검사 비용에 따른 경제적 부담 없이 누구나 코로나19 감염 여부를 진단해 볼 수 있는 추가적인 진단 도구 개발에 기여하였다. 둘째, 원격 환경에서 코로나19 진단이 가능하기 때문에, 의료 접근성이 낮은 환경에서 유용하게 활용될 수 있다. 셋째, 현재의 코로나19 검사 방법은 면봉으로 코를 깊숙이 찔러 코 안쪽을 긁는 방법을 사용하고 있어 많은 사람들이 고통을 호소하고 있으나, 기침 소리를 이용할 경우 이러한 고통 없이 코로나19 진단이 가능하게 된다. 넷째, 모델을 통한 기침 소리 분석 후 코로나19 진단까지의 시간이 다른 방법들에 비해 매우 짧아, 쉽고 빠르게 진단이 가능하다. 마지막으로 공공장소에서 활용할 경우, 발열체크를 통한 의심환자 식별 보다 우수한 식별 능력을 보여줄 수 있어, 감염 확산을 사전에 저지하는데 기여할 수 있다.

본 연구에서 사용한 기침데이터는 chunk를 통해 침소리만 분할하거나 SNR로 소음을 제외하더라도 기침 소리가 아닌 말소리 등이 섞여 있었는데, 이로 인해 학습의 성능이 저하된 점은 본 연구의 한계점으로 여겨진다. 향후 연구에서는 기침 소리만 포함되도록 자료 정제를 통해 좀 더 좋은 성능의 모델을 제공하고자 한다. 또한 확진자의 확진 전/후 기침 소리 수집이 가능하다면 더욱 좋은 학습결과를 얻을 수 있을 것으로 기대된다. 이에 기침데이터 수집을 체계적으로 수행하여 품질 좋은 기침데이터 셋이 마련되면 좀 더 실용적인 연구로 진행될 수 있을 것이다.

## 참고문헌(References)

- 김가혜, 이소현. (2022). 코로나 전후 행복 이슈 변화 분석 및 행복 증진 방안 연구. *지능정보 연구*, 28(3), 81-103.
- 손명진. (2022). 기침소리 데이터를 이용한 COVID-19 감염자 진단. 상명대학교 석사학위논문.
- 오상우. (2019). 인공지능 기반의 음성분석을 통한 우울증, 불안증, 조기치매, 또는 자살 징후 조기판별 시스템. 특허등록번호: 10-2041848.
- 유성주, 김재운. (2022). 기침 파형 패턴을 활용한 COVID-19 확진자 식별 딥러닝 모델. *한국통신학회 동계종합학술발표회*, 1299-1300.
- 유소연, 임규건. (2021). 텍스트 마이닝과 의미 네트워크 분석을 활용한 뉴스 의제 분석: 코로나 19 관련 감정을 중심으로. *지능정보 연구*, 27(1), 47-64.
- 이예진. (2022). 기침 소리를 이용한 COVID-19 진단에 최적화된 feature들 제안. 상명대학교 석사학위논문.
- 이혜빈, 박현진. (2021). 기침 음성 데이터를 이  
 용한 COVID-19 분류. *대한전자공학회 하계 학술대회논문집*, 1701-1703.
- 정연길, 서수민, 강현민. (2022). 토픽 모델링을 활용한 코로나19 초기 생활체육 이슈 분석. *지능정보연구*, 28(3), 57-79.
- 최혜진. (2022). VGG16모델을 활용한 코로나 진단에 관한 연구. 한양대학교 석사학위논문.
- Abeyratne, U. R., Swarnkar, V., Setyati, A., & Triasih, R. (2013). Cough sound analysis can rapidly diagnose childhood pneumonia. *Annals of Biomedical Engineering*, 41(11), 2448-2462.
- Alqudaihi, K. S., Aslam, N., Khan, I. U., Almuhaideb, A. M., Alsunaidi, S. J., Ibrahim, N. M. A. R., ... & Alshahrani, M. S. (2021). Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities. *Ieee Access*, 9, 102327-102344.
- Amrulloh, Y., Abeyratne, U., Swarnkar, V., & Triasih, R. (2015). Cough sound analysis for pneumonia and asthma classification in pediatric population. In *2015 6th International Conference on Intelligent Systems, Modelling and Simulation*, Kuala Lumpur, Malaysia, 127-131.
- Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombath, A., Spathis, D., Xia, T., Cicuta, P., & Mascolo, C. (2020). Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3474-3484.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785-794.
- Daoud, E. (2019). Comparison between XGBoost,

- LightGBM and CatBoost Using a Home Credit Dataset. *International Journal of Computer and Information Engineering*, 13(1), 6 - 10.
- Dave, N. (2013). Feature extraction methods LPC, PLP, and MFCC in speech recognition. *International Journal for Advance Research in Engineering and Technology*, 1(6), 1-5.
- Dhaliwal, S. S., Nahid, A. A., Abbas, R.(2018) Effective Intrusion Detection System Using XGBoost. *Information*, 9(7), 149.
- Doulah, A. B. M. S. U., & Fattah, S. A. (2014). Neuromuscular disease classification based on mel frequency cepstrum of motor unit action potential. In *2014 International Conference on Electrical Engineering and Information & Communication Technology*, 1-4.
- Fakhry, A., Jiang, X., Xiao, J. Chaudhari, G., Han, A., & Khanzada, A. (2021). Virufy: A multi-branch deep learning network for automated detection of COVID-19. *Interspeech*.
- Gaurav, R., Yahia-Cherif, L., Pyatigorskaya, N., Mangone, G., Biondetti, E., Valabrègue, R., ... & Lehericy, S. (2021). Longitudinal changes in neuromelanin MRI signal in Parkinson's disease: a progression marker. *Movement Disorders*, 36(7), 1592-1602.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Patatern Recognition*, 770-778.
- Imran, A., Posokhova, I., Qureshi, H. N., Masood, U., Riaz, M. S., Ali, K., John, C. N., Hussain, M. I., & Nabeel, M. (2020). AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Informatics in Medicine Unlocked*, 20, 100378.
- Issa, D., Demirci, M. F., & Yazici, A. (2020). Speech emotion recognition with deep convolutional neural networks. *Biomedical Signal Processing and Control*, 59, 101894.
- Ittichaichareon, C., Suksri, S., & Yingthawornsuk, T. (2012). Speech recognition using MFCC. *International Conference on Computer Graphics, Simulation and Modeling*, Pattaya, Thailand.
- Jana, B., & Nath, P. K. (2022). A Single-Chip Solution for Diagnosing Peripheral Arterial Disease. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 30(5), 671-675.
- Jana, B., Oswal, K., Mitra, S., Saha, G., & Banerjee, S. (2019). Detection of peripheral arterial disease using Doppler spectrogram based expert system for Point-of-Care applications. *Biomedical Signal Processing and Control*, 54, 101599.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30.
- Kumar, A., Abhishek, K., Ghalib, M. R., Nerurkar, P., Shah, K., Chandane, M., Bhirud, S., Patel, D., & Busnel, Y. (2022). Towards cough sound analysis using the internet of things and deep learning for pulmonary disease prediction. *Transactions on Emerging Telecommunications Technologies*, 33(10), e4184.
- McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa:audio and music signal analysis in python, In *Proceedings of the 14th Python in Science Conference*. Austin, TX, 18 - 24.
- Mishra, D., & Sahu, B. (2011). Feature selection for cancer classification: a signal-to-noise ratio approach. *International Journal of Scientific*

- & *Engineering Research*, 2(4), 1-7.
- Muda, L., Begam, M., & Elamvazuthi, I. (2010). Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *Journal of Computing*, 2(3), 138-143.
- Ogunleye, A., and Wang, Q. G.(2020). XGBoost Model for Chronic Kidney Disease Diagnosis. *In IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 17(6), 2131-2140.
- Orlandic, L., Teijeiro, T., & Atienza, D. (2021). The COUGHVID crowdsourcing dataset: A corpus for the study of large-scale cough analysis algorithms. *Science Data* 8, 156, <https://doi.org/10.1038/s41597-021-00937-4>.
- Pramono, R. X. A., Imtiaz, S. A., & Rodriguez-Villegas, E. (2016). A cough-based algorithm for automatic diagnosis of pertussis. *PLoS One*, 11(9), e0162128.
- Rezaee, K., Savarkar, S., Yu, X., & Zhang, J. (2022). A hybrid deep transfer learning-based approach for Parkinson's disease classification in surface electromyography signals. *Biomedical Signal Processing and Control*, 71, 103161.
- Rezende, E., Ruppert, G., Carvalho, T., Ramos, F., & Geus, P.(2017). Malicious Software Classification Using Transfer Learning of ResNet-50 Deep Neural Network. *In Proceedings of 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Cancun, Mexico, 1011-1014.
- Sadi, T. M., & Hassan, R. (2020). Development of classification methods for wheeze and crackle using mel frequency cepstral coefficient (MFCC): A deep learning approach. *International Journal on Perceptive and Cognitive Computing*, 6(2), 107-114.
- Singh, Y., & Biswas, A. (2022). Robustness of musical features on deep learning models for music genre classification. *Expert Systems with Applications*, 199, 116879.
- Su, Y., Zhang, K., Wang, J., & Madani, K. (2019). Environment sound classification using a two-stream CNN based on decision-level fusion. *Sensors*, 19(7), 1733.
- Tadesse, G. A., Javed, H., Thanh, N. L. N., Thi, H. D. H., Thwaites, L., Clifton, D. A., & Zhu, T. (2020). Multi-modal diagnosis of infectious diseases in the developing world. *IEEE Journal of Biomedical and Health Informatics*, 24(7), 2131-2141.
- Tiwari, V. (2010). MFCC and its applications in speaker recognition. *International Journal on Emerging Technologies*, 1(1), 19-22.
- Wang, D., Zhang, Y., & Zhao, Y.(2017). LightGBM: An effective miRNA classification method in breast cancer patients. *In Proceedings of the 2017 International Conference on Computational Biology and Bioinformatics*, 7 - 11.
- Wen, L., Li, X., & Gao, L. A.(2020). Transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Computing and Applications*, 32, 6111 - 6124.

Abstract

## A COVID-19 Diagnosis Model based on Various Transformations of Cough Sounds

Minkyung Kim\* · Gunwoo Kim\*\* · Keunho Choi\*\*\*

COVID-19, which started in Wuhan, China in November 2019, spread beyond China in 2020 and spread worldwide in March 2020. It is important to prevent a highly contagious virus like COVID-19 in advance and to actively treat it when confirmed, but it is more important to identify the confirmed fact quickly and prevent its spread since it is a virus that spreads quickly. However, PCR test to check for infection is costly and time consuming, and self-kit test is also easy to access, but the cost of the kit is not easy to receive every time. Therefore, if it is possible to determine whether or not a person is positive for COVID-19 based on the sound of a cough so that anyone can use it easily, anyone can easily check whether or not they are confirmed at anytime, anywhere, and it can have great economic advantages. In this study, an experiment was conducted on a method to identify whether or not COVID-19 was confirmed based on a cough sound. Cough sound features were extracted through MFCC, Mel-Spectrogram, and spectral contrast. For the quality of cough sound, noisy data was deleted through SNR, and only the cough sound was extracted from the voice file through chunk. Since the objective is COVID-19 positive and negative classification, learning was performed through XGBoost, LightGBM, and FCNN algorithms, which are often used for classification, and the results were compared. Additionally, we conducted a comparative experiment on the performance of the model using multidimensional vectors obtained by converting cough sounds into both images and vectors. The experimental results showed that the LightGBM model utilizing features obtained by converting basic information about health status and cough sounds into multidimensional vectors through MFCC, Mel-Spectrogram, Spectral contrast, and Spectrogram achieved the highest accuracy of 0.74.

**Key Words** : COVID-19, Cough sound, LightGBM, MFCC, SNR

Received : April 25, 2023 Revised : July 17, 2023 Accepted : August 1, 2023

Corresponding Author : Keunho Choi

---

\* DataWorld Co., Ltd.

\*\* Department of Business Administration, Hanbat National University

\*\*\* Corresponding author: Keunho Choi

Department of Business Administration, Hanbat National University, 125, Dongseo-daero, Yuseong-gu, Daejeon, Korea, 34158  
Tel: +82-42-821-1468, Fax: +82-42-821-1597, E-mail: keunho@hanbat.ac.kr

## 저 자 소개



**김민경**

국립한밭대학교에서 경영학 석사학위를 수여하였으며, 현재 ㈜데이터월드에서 재직 중입니다. 관세청 등 국가공공기관 관련 시스템 구축 및 유지 사업에 참여하고 있으며, 주요 관심분야는 머신러닝, 딥러닝 등입니다.



**김건우**

고려대학교에서 경영학 박사학위를 수여하였으며, 현재 국립한밭대학교 융합경영학과에서 교수로 재직 중입니다. 주요 관심분야는 비즈니스 온톨로지 모델, 빅데이터 분석, 핀테크 기술 및 전략 등입니다.



**최근호**

고려대학교에서 경영학 박사학위를 수여하였으며, 현재 국립한밭대학교 융합경영학과에서 부교수로 재직 중입니다. 주요 관심분야는 추천시스템, 의료 빅데이터 분석, 딥러닝, 머신러닝 등입니다.