**Distinguish COVID-19 by coughing**

**Team8(Brno Ting Ting)**

**Victor Mutinda Mutuku**

**Ban Sangsue**

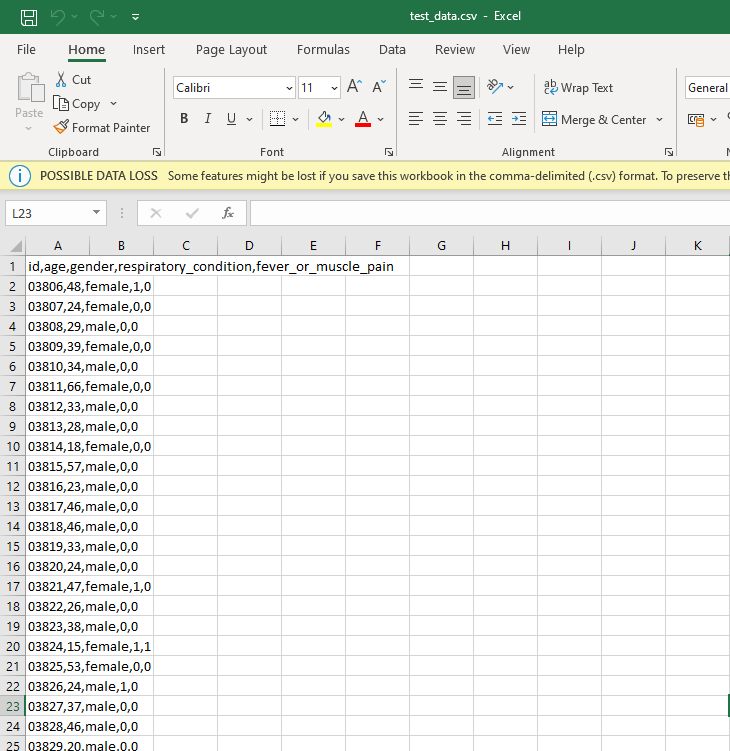
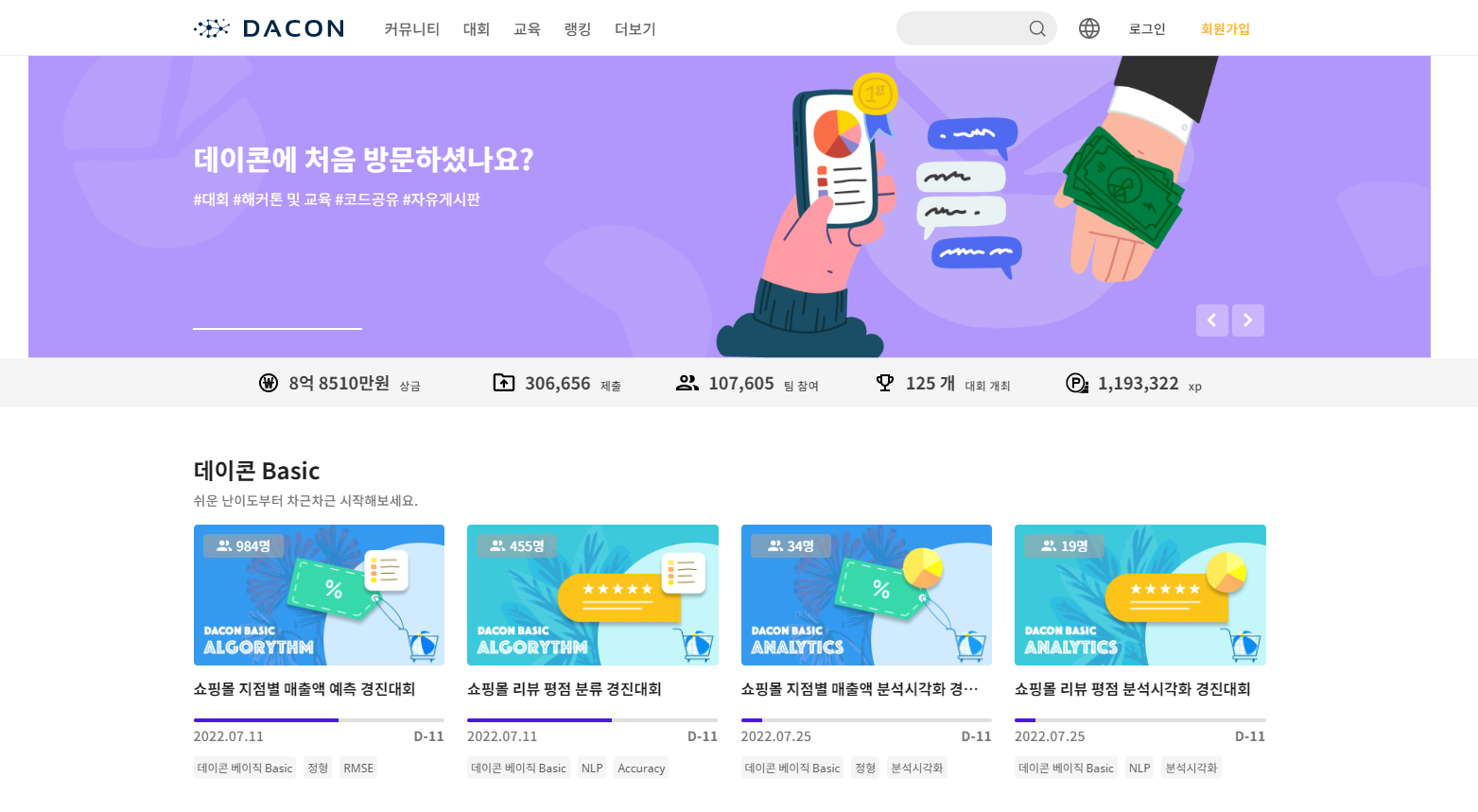
**Noh Ha young**

**Jang Suyoon**

1. **Intro**

COVID-19 is a disease characterized by a dry cough and fever, among other symptoms. These symptoms are normally identified manually by a doctor to determine the presence of the disease. To ease the process, we have trained a machine learning model to be able to identify the disease from the cough sound of the patient. This can be used in conjunction with a microphone to be able to detect the presence of the disease, as early as possible, in public spaces, thus helping to mitigate the spread.

1. **Used data set**

****

* Use Data Set from ‘Dacon’
* Found train(3806 files), test(5733 files), unlabeled(1868 files) data of various cough sound, gender, fever or muscle pain and respiratory condition.

1. **Code link**

[**https://drive.google.com/drive/folders/1pJnKNQ\_5WKZ2FJ3AADgxcEzTanzlbMev**](https://drive.google.com/drive/folders/1pJnKNQ_5WKZ2FJ3AADgxcEzTanzlbMev)

1. **How to approach & Task**

1) Collecting data on cough sounds from people with COVID-19 and non-COVID-19 - Dacon(mentioned above)

2) Data preprocessing : Data Labeling, Data Classification

* Enables machine learning through data preprocessing by Data import, labeling.
* Used librosa, Numpy, Pandas.

3) Feature extraction(MFCC (librosa))

* Extract specific values of covid-19 cough sounds through MFCC encoding
* Used MFCC (librosa), sklearn preprocessing OneHotEncoder

4) Making Training model : [sklearn.neural\_network](https://scikit-learn.org/stable/modules/classes.html#module-sklearn.neural_network).MLPClassifier, semi supervised

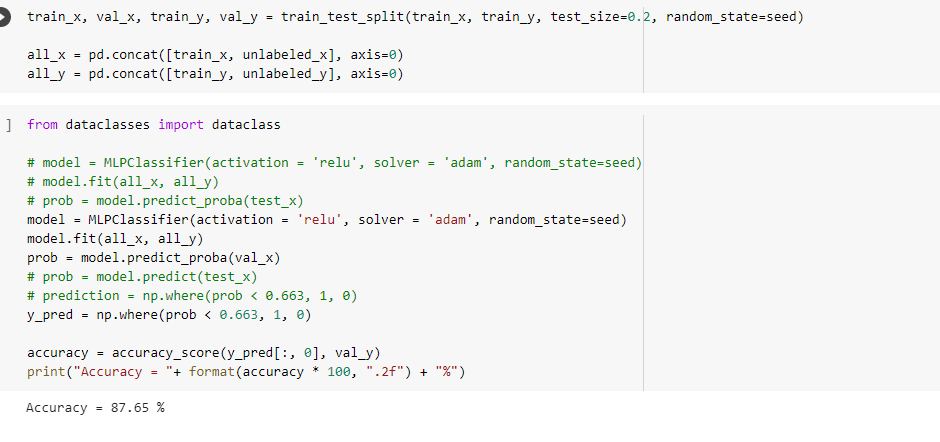
* training model used Sklearn Neural Network MLPClassifier
* activation = relu, solver = adam
* Improve accuracy with semi-supervised learning
* Determining the sound of a specific value being detected as covid-19
* Results 0 is negative and 1 is positive of Covid-19

5) Accuracy check, Input new data and test the model

1. **Library**

-> sklearn, librosa, numpy, pandas, audiomentations

1. **Accuracy**

**our project accuracy is 87.65%**

1. **Result**

The topic of determining whether it was COVID-19 through the sound of coughing was quite interesting. The standard method for detecting COVID-19 is PCR testing, which is expensive and time-consuming, so if the amount of testing increases due to the re-pandemic, the medical system may collapse. If COVID-19 testing is possible through this method, I think it is an alternative diagnostic tool that can be solved with existing limitations. The accuracy of 87.65% was predicted by conducting a more precise examination through fever and respiratory symptoms as well as coughing. It's unfortunate that it's not more accurate, but it's my first time dealing with sound data, so I was satisfied with this level of achievement, and I think I had fun studying related methodologies through good topics, so I was able to finish the project happily.