**Laporan Akhir Deep Learning**



Kelompok 7:

Nouchka Indra Dewa / C14170021

Jason Lucky Harrysiwanto / C14170129

Leonardo Yurion Tungribali / C14170153

Eric Pranoto / C14180139

**Universitas Kristen Petra**

**SURABAYA**

**2021**

1. **Judul Proyek**

Judul : Detecting Infectious Regions of COVID-19 in X-ray Images and CT scans

Using Different Pretrained Densenet Models

1. **Latar belakang**

Di masa sekarang ini kita semua sedang merasakan ketidaknyamanan akibat adanya pandemi yang melanda seluruh dunia. Hal ini disebabkan karena adanya virus COVID-19 yang telah menyebar dengan gejala dan akibat yang tak menentu. Salah satu akibatnya berdampak pada paru-paru dan mungkin ada beberapa area lain yang terdampak jika terjangkit COVID-19 ini. Oleh karena itu kami ingin mencoba menggali informasi dengan cara mendeteksi area mana yang terinfeksi virus COVID-19 melalui gambar X-ray dan CT-scan menggunakan beberapa pretrained Densenet Model yang berbeda.

1. **Tujuan**

Untuk mendeteksi dan mengetahui efek yang ditimbulkan saat seseorang terjangkit virus COVID-19 pada bagian tertentu dalam paru-paru, sehingga mungkin untuk masa mendatang dapat membantu tim medis untuk menindaklanjuti kasus-kasus dengan dampak berbeda dan dapat dengan tepat dan cepat mengobati pasien berdasarkan kasus tersebut.

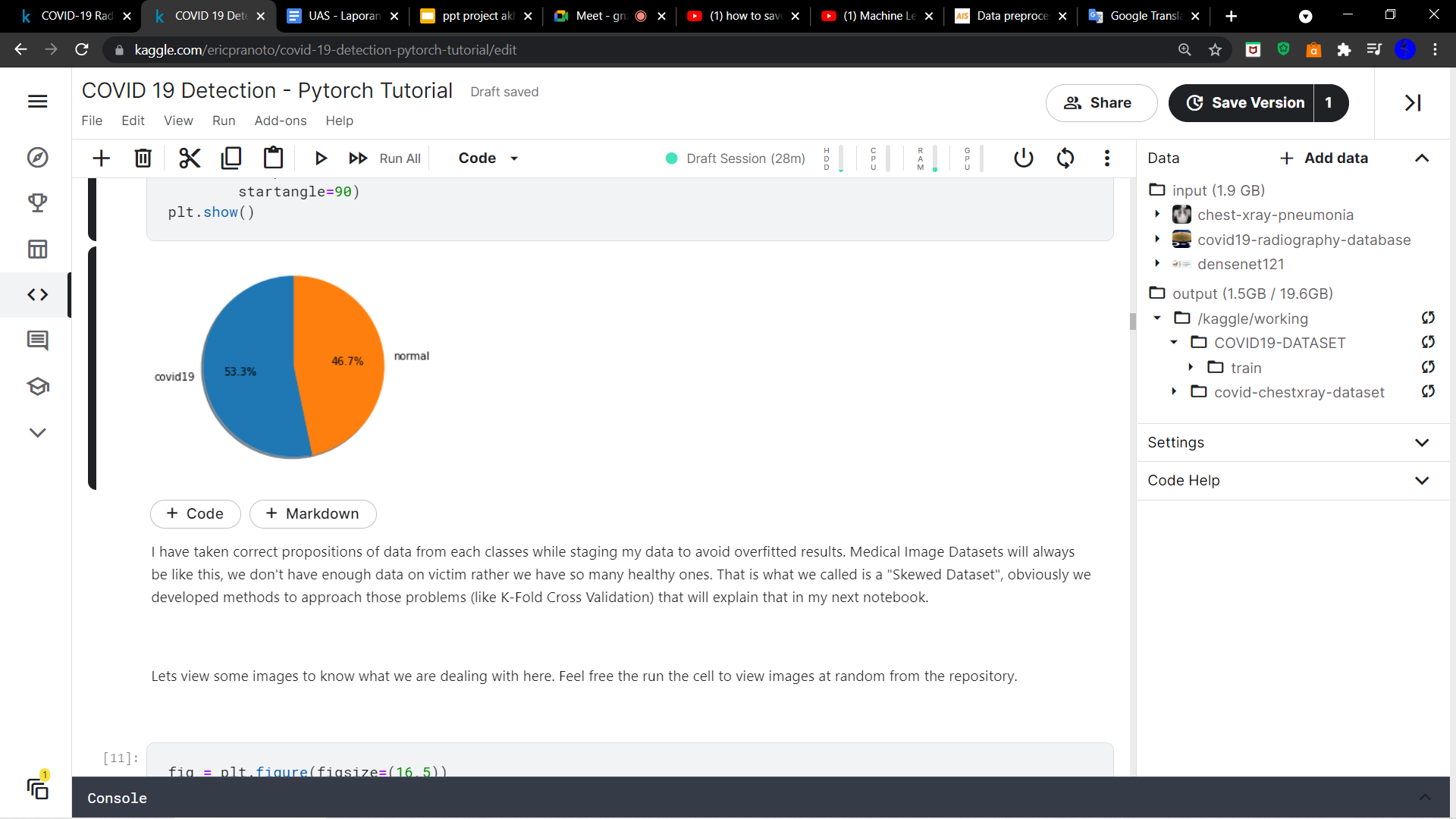
1. **Sumber Data**

Kami menggunakan Dataset yang berjudul “**COVID-19\_Radiography\_Dataset**“ dari kaggle. Dataset ini dibuat oleh tim peneliti dari Qatar University, Qatar dan University of Dhaka, Bangladesh beserta kolaborator dari Pakistan dan Malaysia yang juga bekerja sama dengan dokter medis.

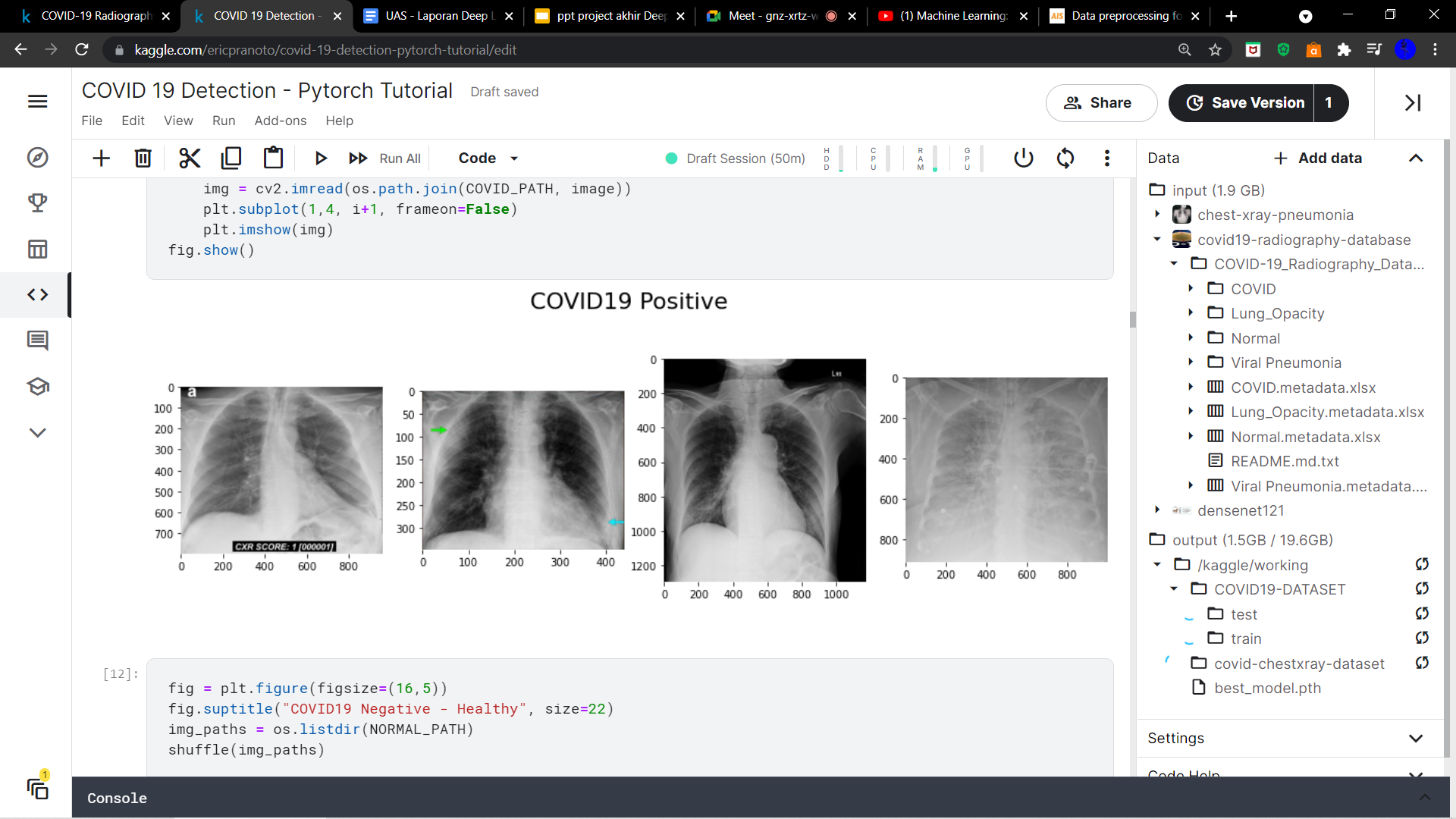
Pertama kali dirilis dataset ini berisi 219 kasus positif COVID-19, 1341 normal, dan 1345 gambar rontgen dada virus pneumonia (CXR). Dalam pembaruan pertama, telah ditingkatkan kelas COVID-19 menjadi 1200 gambar CXR. Dalam pembaruan ke-2, telah ditingkatkan database menjadi 3616 kasus positif COVID-19 bersama dengan 10.192 Gambar Normal, 6012 Opacity Paru (Infeksi paru-paru Non-COVID) dan 1345 gambar Viral Pneumonia. Dan kabarnya masih akan terus diupdate.

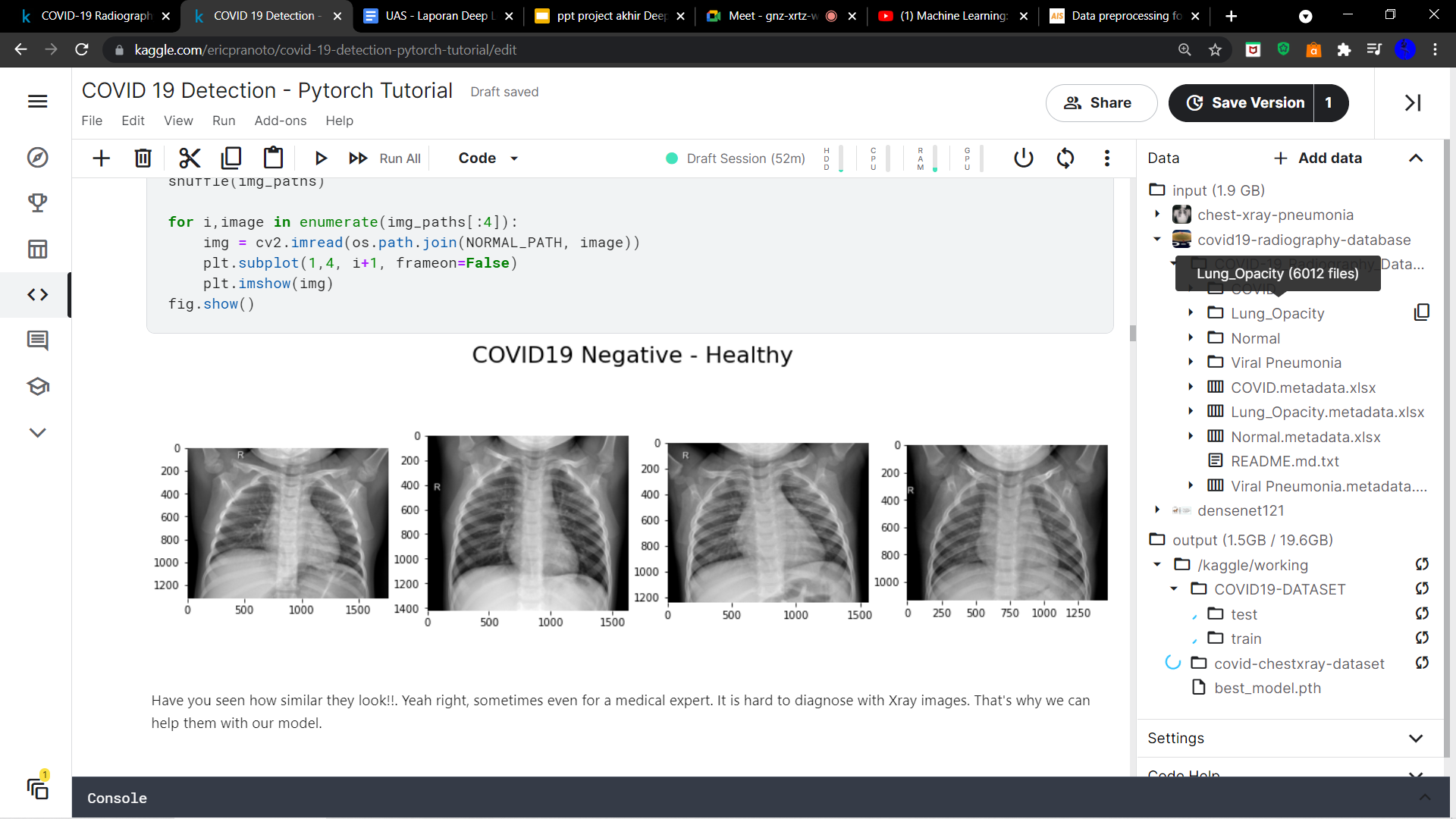
1. **Pengerjaan**

# EKSPERIMEN 1

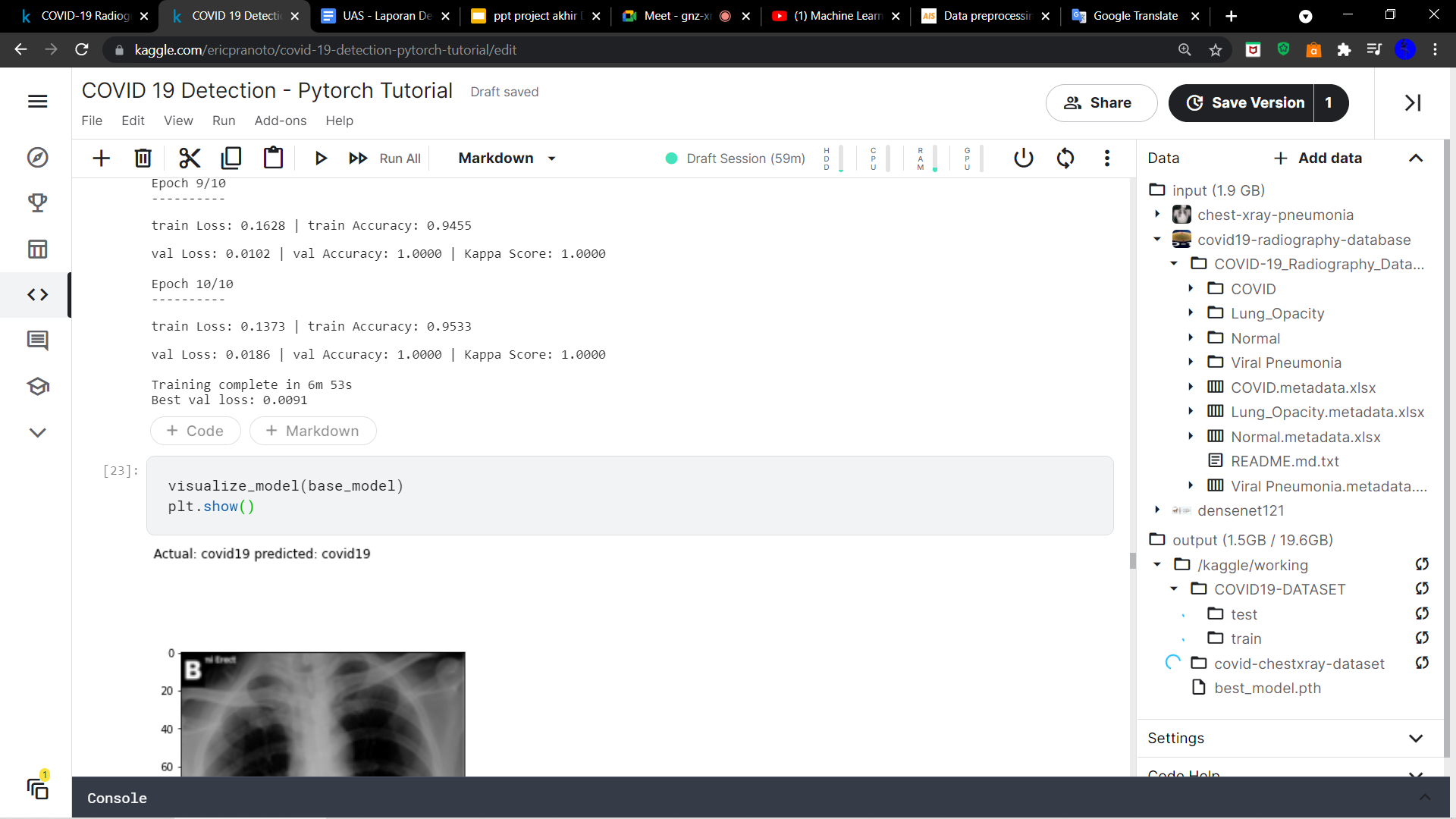
Kami mencoba menggunakan coding densenet dengan notebook dari kaggle. Kami menggunakan arsitektur Densenet-121 sebagai core dari model ini. Kami juga menggunakan pre-trained weights sebagai sarana untuk Transfer Learning untuk mencapai akurasi yang lebih tinggi pada model ini. Untuk data pre-processingnya kami menggunakan data transformation. 

Pertama-tama kami train sebuah CNN classifier yang bisa classify gambar Xray paru-paru yang terinfeksi covid dari gambar Xray paru-paru sehat menggunakan Cohen's [COVID Chest X-ray Dataset](https://github.com/ieee8023/covid-chestxray-dataset) dan Paul Mooney's [Chest X-ray Dataset (Pneumonia)](https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia). Kami load dataset untuk training dan transfer ke directory baru, lalu plot distribusi dari tiap class nya, dan hasilnya seperti di gambar kanan.

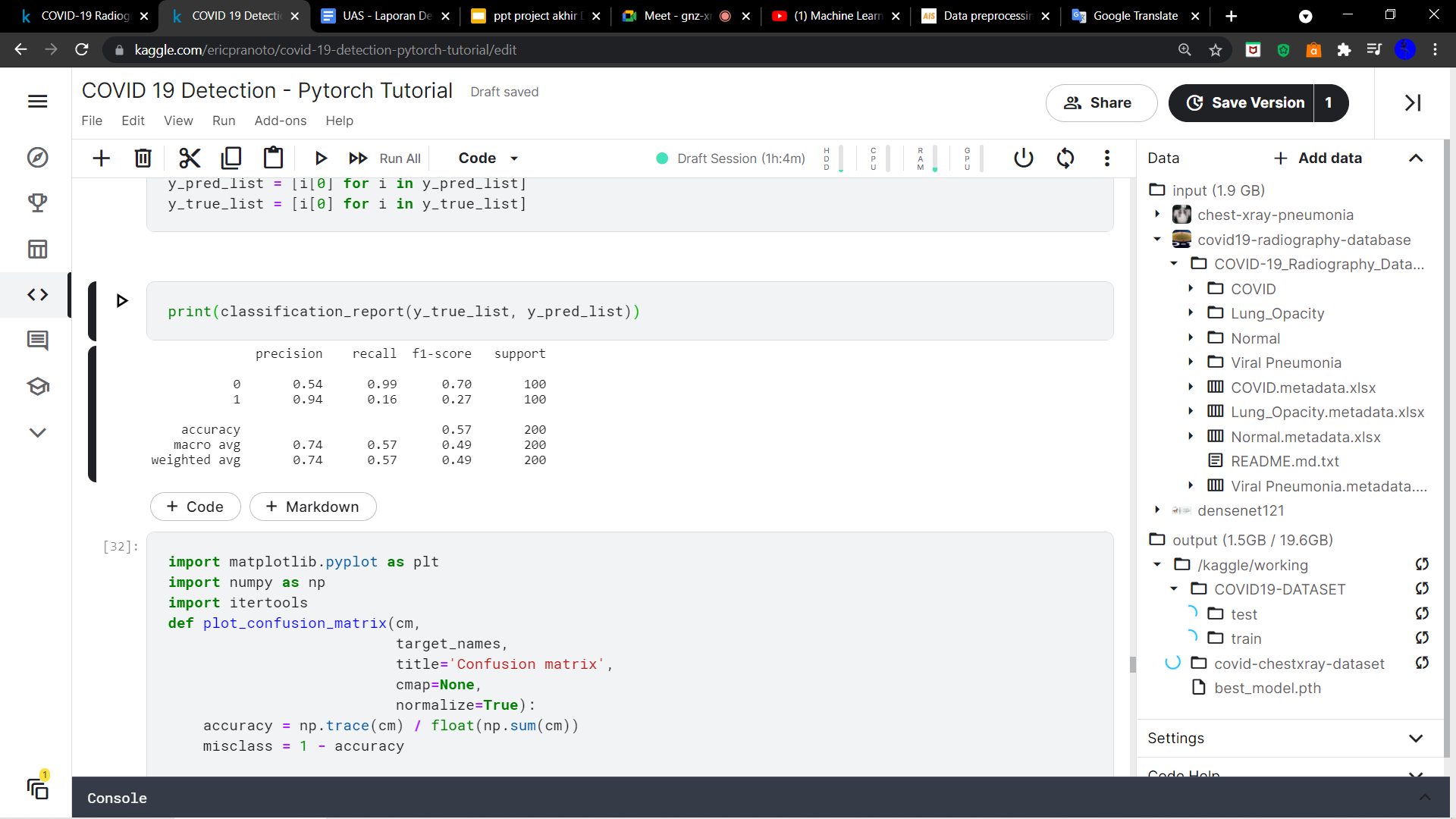
Kemudian kita coba lihat contoh gambar paru-paru yang mengidap COVID-19 dan yang sehat. Berikut gambarnya.

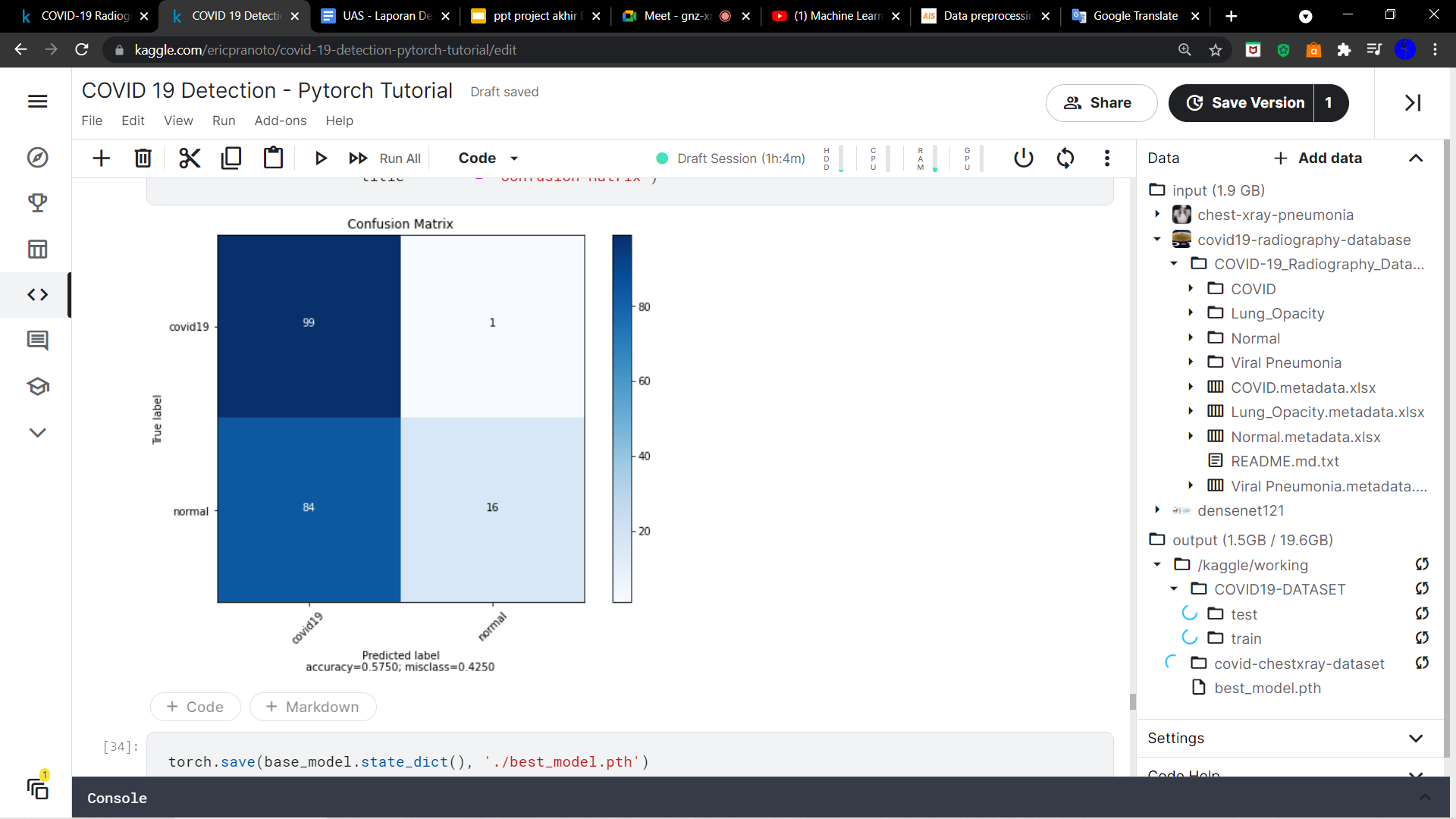


Setelah itu kita transform datanya, lalu kita split datanya menjadi train dan validation set untuk mentrain dan memvalidasi model dengan dataset terpisah. Kemudian barulah kita train data trainnya dengan CNN model. Kita akan menggunakan metode early stopping untuk memilih model dengan performa terbaik, kita menggunakan 10 ephoc, dimana dalam percobaan pertama semua ephoc berhasil diselesaikan dalam waktu 29 menit, lalu percobaan kedua berjalan sekitar 7 menit dengan hasil sebagai berikut:

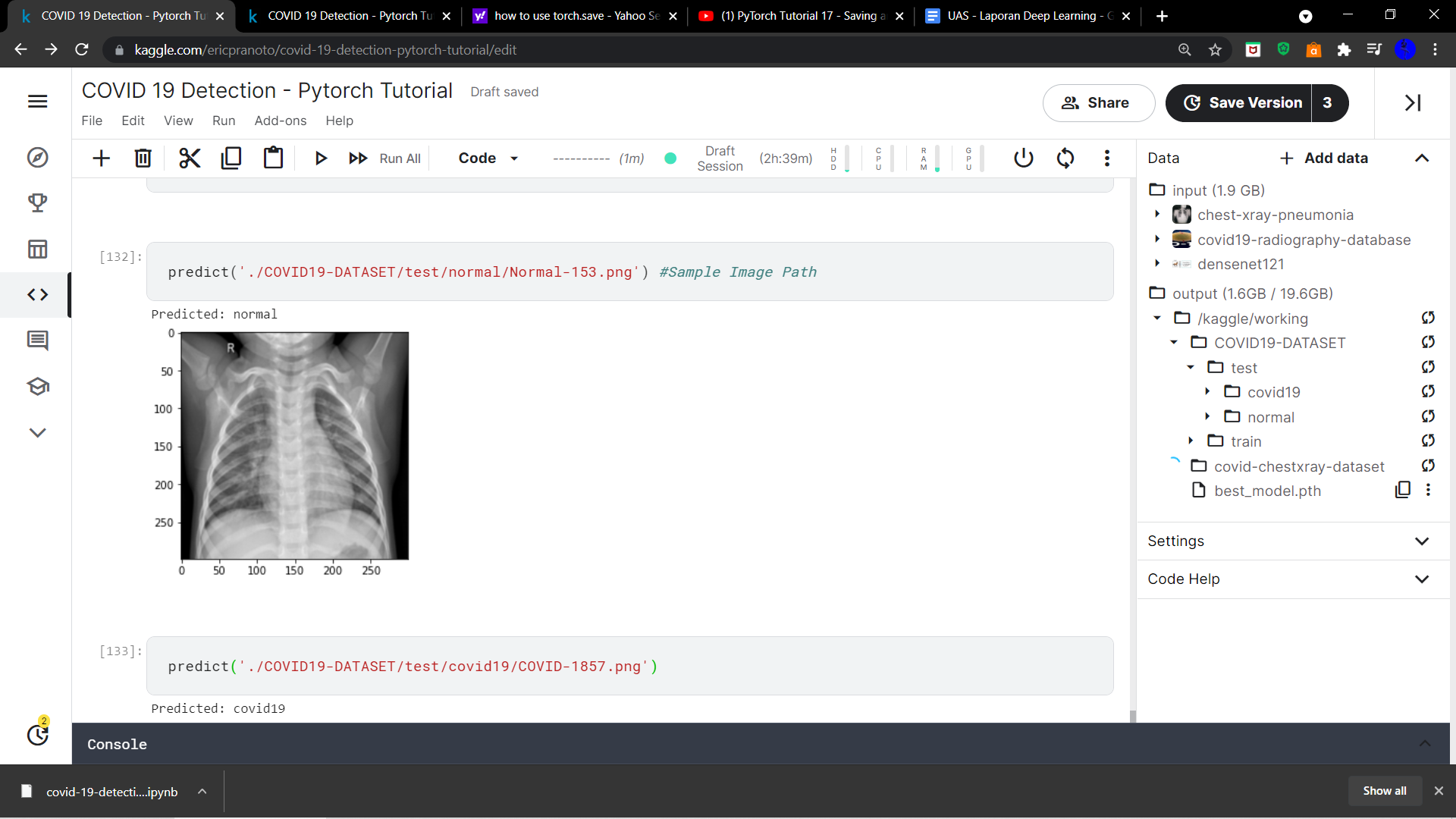


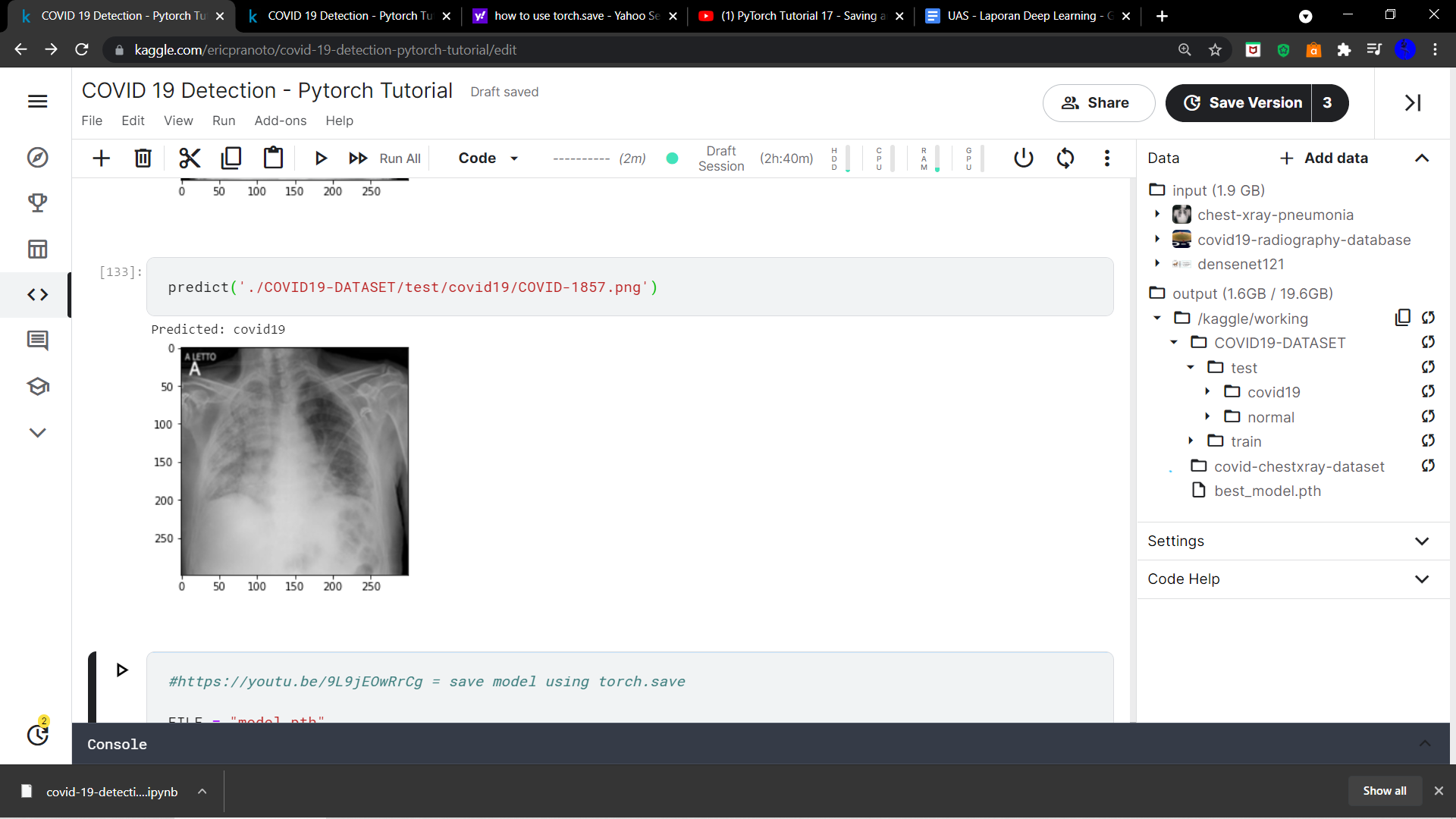
Kemudian kita melakukan model evaluation, barulah kita disini melakukan test terhadap **COVID-19\_Radiography\_Dataset** dengan mengambil 100 gambar dari tiap class dan setelah diprediksi, hasil reportnya sebagai berikut



Dan confussion matrixnya sebagai berikut

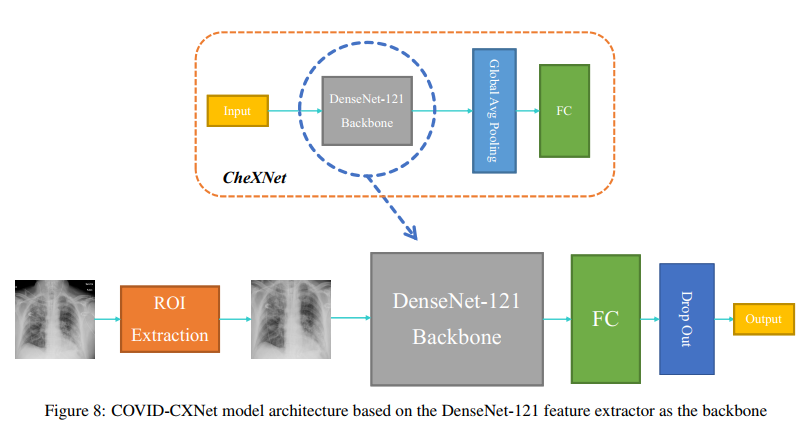
Setelah kita melakukan training dan testing, selanjutnya kita akan predict sebuah contoh gambar dan hasilnya sebagai berikut.





# EKSPERIMEN 2

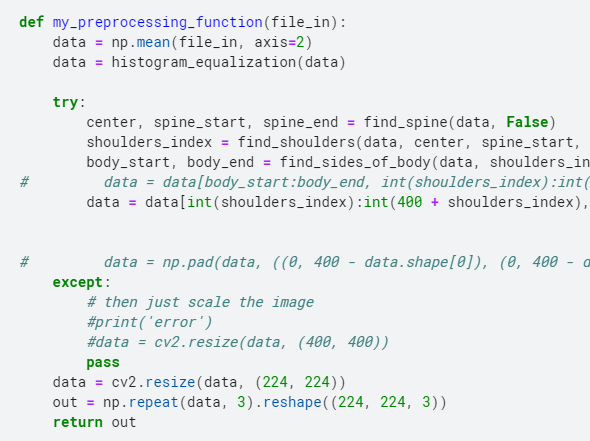
Setelah eksperimen sebelum-nya, kami mencari metode lain yang berhubungan dengan Densenet, dan pada akhirnya mengambil inspirasi arsitektur dari paper [COVID-CXNet: Detecting COVID-19 in Frontal Chest X-ray Images using Deep Learning](https://arxiv.org/abs/2006.13807). Pada waktu pengerjaan laporan ini, paper riset belum di *peer-review* seperti yang dikatakan pada halaman web paper.



Model yang diusulkan dalam paper menggunakan Densenet121 sebagai basis-nya. Selain Densenet121 tersebut, terdapat sebuah Fully Connected Layer yang memiliki 10 node dan dilanjut dengan sebuah dropout layer dengan rate 0.2 untuk mencegah overfitting. Activation function yang digunakan adalah sigmoid karena tujuan dari model ini adalah binary classification.

Pada proses preprocessing eksperimen ini, sebuah gambar di segmentasi untuk mencari bagian-bagian tertentu dalam bagian tubuh x-ray dan proses penghitungan histogram image.

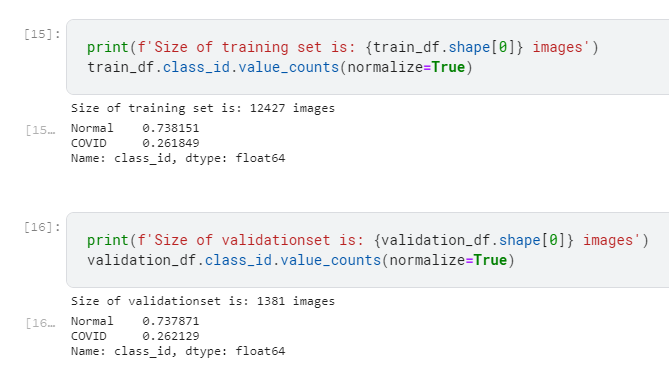




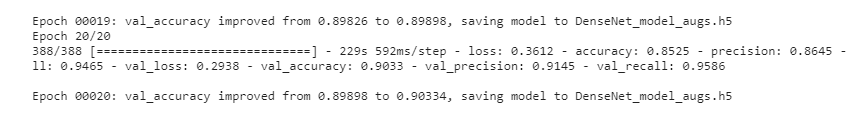
Model ditrain dengan dataset [COVID-19 Radiography Database](https://www.kaggle.com/tawsifurrahman/covid19-radiography-database) yang tersedia di kaggle. Jumlah image yang tersedia kurang lebih 14000 dengan perbandingan sebagai berikut:



2 class yang digunakan dalam classification binary ini adalah image paru-paru yang memiliki COVID dan paru-paru yang tidak memiliki COVID. Pembagian akhir dari dataset dapat dilihat pada gambar dibawah:

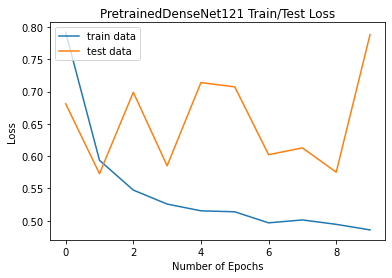


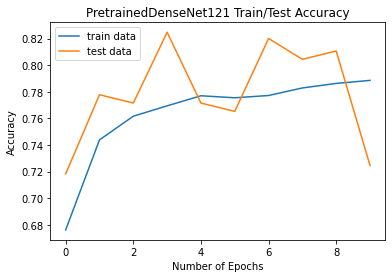
Model di train selama 20 epoch dengan hasil akhir akurasi 0.9033, 20 epoch tersebut membutuhkan waktu sekitar satu setengah jam dengan cloud GPU yang disediakan oleh kaggle.



# EKSPERIMEN 3

Kami mencari lagi suatu metode yang sama dengan menghasilkan output yang berbeda dari dataset yang lain dan densenet 121 digunakan untuk melakukan proses pencocokan dan mencari tahu lokasi dari covid, viral dan pneumonia didalam paru-paru dengan tampilan heat view.



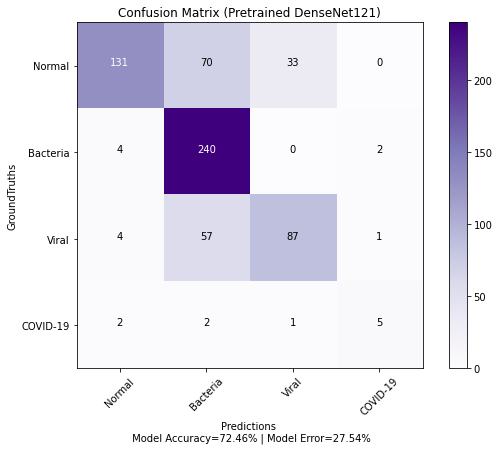


Gambar diatas adalah hasil dari pretrain model pada data set class yang ada dengan densenet 121. Menggunakan epoch sebanyak 10 kali untuk melakukan pretrain model dan mendapatkan akurasi sebesar 82,7 persen. Waktu yang dibutuhkan per epoch melakukan proses sekitar 8,5 menit dan membutuhkan waktu sekitar sejam setengah untuk menyelesaikan tugasnya.

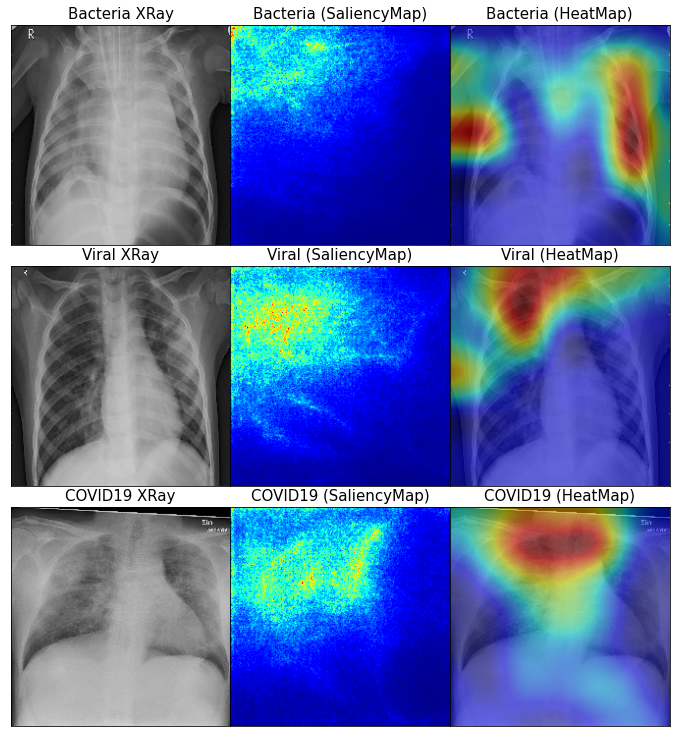


Ini adalah hasil perhitungan epochnya yang sudah dilakukan train sebanyak 10 kali.

Kemudian dilakukan model evaluation. Dan kemudian akan melakukan test terhadap data set yang ada. Dan kita membuat confusion matrix nya.



Setelah kita melakukan training dan test pada data set maka kita akan mendapatkan hasil seperti berikut.



Dari hasil output tersebut didapatkan lokasi dimana penyakit yang didalam paru-paru tersebut berada. Dan juga bisa membedakan yang mana terkena bakteri, viral ataupun covid19.

1. **Kesimpulan**

DenseNet sangat berguna untuk membantu menyelesaikan masalah sehari-hari dan dapat diimplementasikan dengan berbagai metode dengan hasil yang beragam pula.

1. **Hasil Refleksi**
2. Jason Lucky Harrysiwanto: Ilmu yang saya dapat dari mengerjakan proyek ini sebagian besar ilmu teknis dan tambahan mengenai model densenet dan ilmu tambahan mengenai macam-macam metode preprocessing. Menurut paper COVID-CXnet network Densenet memiliki keunggulan dalam task yang berhubungan dengan radiologi dengan contoh utamanya berupa
3. Eric Pranoto: hal yang saya dapat dari mengerjakan proyek ini lebih ke cara pengolahan data menggunakan kaggle notebook, cara import data, cara connect ke daabase, juga cara penggunaan densenet yang mana saya termasuk sangat awal di bidang densenet ini, sehingga saya mendapat ilmu baru walaupun ibaratnya hanya permukaannya saja. Juga selain itu saya mendapat pengetahuan lain mengenai perbedaan kondisi paru-paru yang tidak maupun telah terinfeksi covid-19.

Untuk hambatannya mungkin lebih ke teknis penggunaan kaggle, lalu untuk pencarian coding densenet dan dataset agak susah menemukan yang sesuai.

1. Nouchka Indra: Hal yang saya pelajari selama mengerjakan proyek ini adalah saya bisa tau cara kerja densenet yang sebelum nya saya tidak tau, saya juga belajar bagaimana proses metode yang kami lakukan untuk proyek akhir, lalu dengan apa yang kami kerjakan saya juga menjadi belajar tentang covid-19 yang ternyata lebih mengerikan dari yang saya bayangkan, saya juga jadi tau apa perbedaan paru-paru sehat dan tidak.
2. Leonardo Yurion: hal yang saya pelajari selama mengerjakan proyek ini, saya bisa tau berbagai macam implementasi dengan dense net yang sebelumnya terdengar asing bagi saya. Dari projek ini saya tau bahwa ngetrain data itu lama dan cepatnya tergantung dari besar data yang harus ditrain jika ingin mendapatkan hasil yang maksimal butuh banyak percobaan dan waktu. Dari projek ini saya belajar mengetahui gambaran paru-paru terkena penyakit dan bagaimana yang normal.