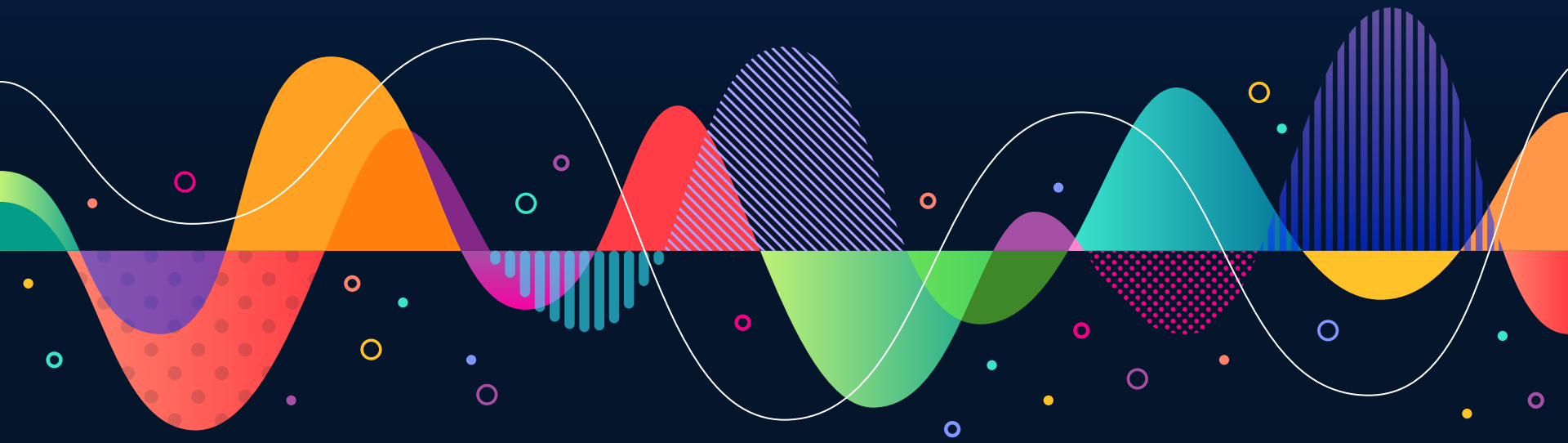
Developing a web-based Convolutional Neural Network based on Supervised Deep Learning for the rapid and accurate diagnosis of Pulmonary Tuberculosis via Chest X-rays



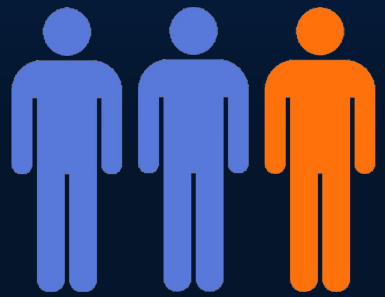
Background information

10 million **infected**
1.5 million **dead**
Every year

(Tuberculosis (TB), 2020)

Tuberculosis

- Bacterial disease with two variants - Latent TB and TB disease.
- Latent TB is an infection that can occur anywhere in the body, and it is very common.
- If Latent TB is left untreated or the patient is immunocompromised, it can progress.
- The situation is getting worse over time



(Latent tb infection, 2020)

Pulmonary Tuberculosis (PTB)

- Constitutes a significant portion of TB disease cases.
- **Highly** infectious
- Has been widely out of control for over **100** years, especially in the developing world.

(Millard, Mugarte-Gil, & Moore, 2015)

Issues with the method of testing for PTB

- Lack of modern, inexpensive, and accurate testing methods
- Smear Microscopy
  - Highly inaccurate - approx. 50% of results are wrong
  - Large time delay (at least 2 weeks)
  - Massive logistical challenges
  - Expensive

(Nardell, 2015)
(Hwang et al., 2018)

Deep Learning

- Deep Learning is a subset of Artificial Intelligence, and is a powerful form of data analysis.

- It can be used in the field of testing for PTB quite easily

(Chen et al., 2019)

Engineering Goal

- To build a Deep Learning model based on chest x-rays and integrate it into a web application so as to make it accessible.
- An 80% accuracy rate and 100% upload success rate are the goals of this project.
- The final product should allow users to get on, upload a chest x-ray, and get the PTB result with a click of a button

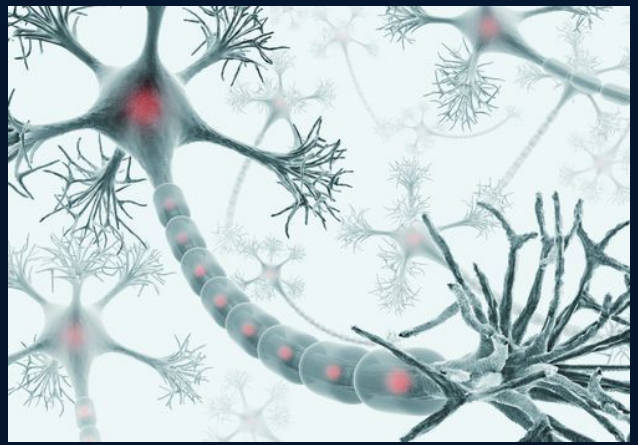
Shortcomings of previous research

- Using Deep Learning algorithms on numerical data rather than chest x-rays
- Not considering accessibility

Development

Convolutional Neural Network

- The deep learning algorithm used in this project was a Convolutional Neural Network



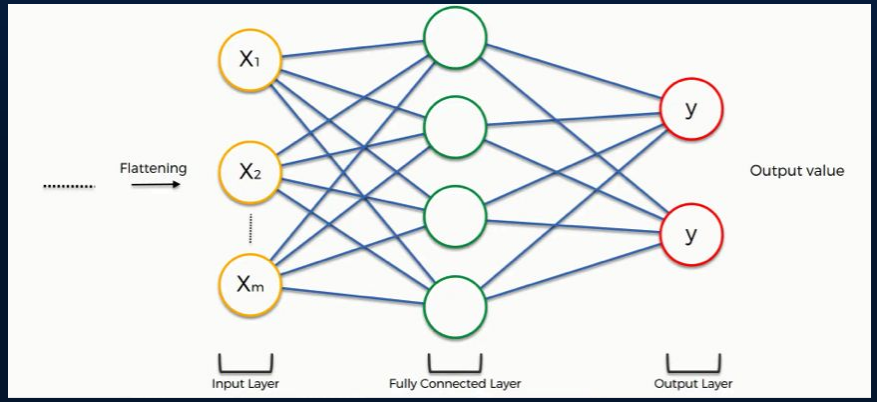
(Chen et al., 2019)

Training the algorithm

1) Convolution function
2) Activation Function (ReLU in this case)
3) Pooling
4) Flattening

Full Connection

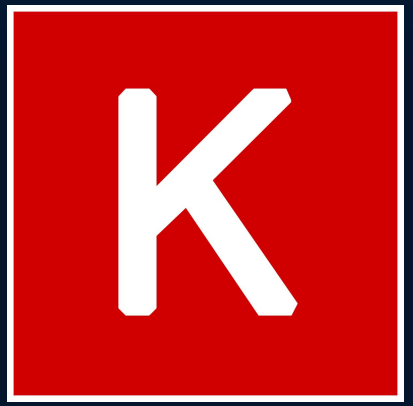
$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$



(Chen et al., 2019)

Training the algorithm (cont.)

- National Library of Medicine website
- 750 de-identified chest x-rays
- Python, Tensorflow, Keras



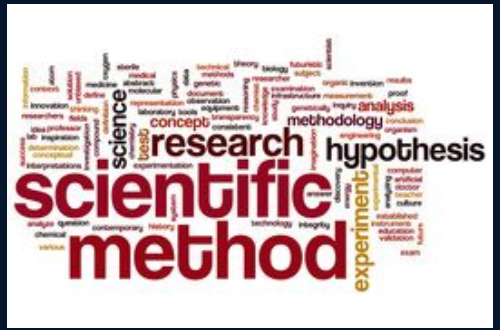
Building the web application

- To develop the overall application, I used HTML, CSS, Bootstrap.
- In addition, I used the Flask web framework to integrate the model within the web application
- I also integrated an addition algorithm to decide whether or not a chest x-ray was being uploaded.

Project details

Design and Method

- I used a Non-Experimental Engineering Design
- Engineering method for AI Development
  - A custom method developed to suit this project
  - Involves developing the AI component first and then any other components, and then integrating them.



(Hwang et al., 2019)

Data collection

- Data was collected and recorded in the format as shown below

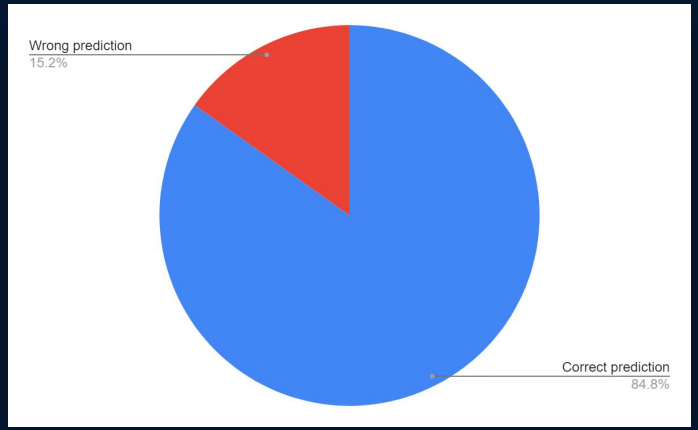
File name	Positive/Negative	Trial 1 result (Prediction & Upload Status)	Trial 2 result (Prediction & Upload Status)	Trial 3 result (Prediction & Upload Status)	Trial 4 result (Prediction & Upload Status)	Trial 5 result (Prediction & Upload Status)	Overall result
CHNCXR_0000_1	Positive	Positive	Positive	Positive	Positive	Positive	Positive
		Success	Success	Success	Success	Success	Success

A note about Statistical Analysis

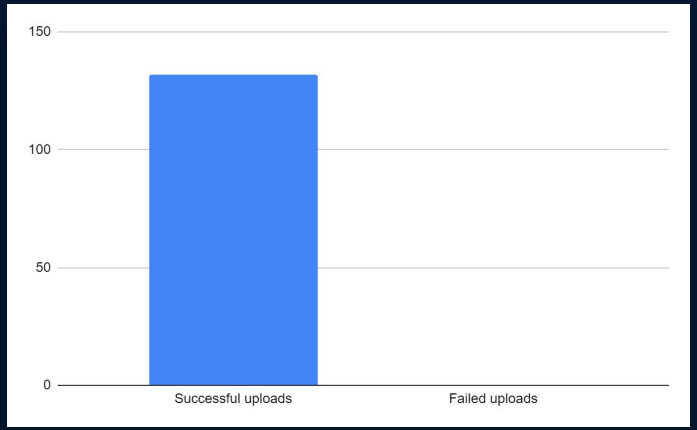
- My data is mostly qualitative as shown on the previous slide.
- No explicit statistics was performed

Results

- Accuracy: 85%
- Upload success rate: 100%
- False positives: 14% of errors, False negatives: 1% of errors
- Second algorithm accuracy: 98%



A pie chart showing correct vs wrong predictions



A bar graph of successful vs failed uploads

Conclusion

- The engineering goal was met
- Higher upload success rate than expected
- More false positives than false negatives

Website/Code Demonstration

Future directions

- Improving the accuracy of the model by training it on more data
- Making cross-platform versions of the app
- Making the app capable of operating completely offline.

Thank you for your time!

Works cited

- Barberis, I., Bragazzi, N. L., Galluzzo, L., & Martini, M. (2017). The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. Journal of preventive medicine and hygiene, 58(1), E9–E12.
- Chen, D., Liu, S., Kingsbury, P., Sohn, S., Storlie, C. B., Habermann, E. B., . . . Liu, H. (2019). Deep learning and alternative learning strategies for retrospective real-world clinical data. Npj Digital Medicine, 2(1). doi:10.1038/s41746-019-0122-0
- Hwang, E. J., Park, S., Jin, K., Kim, J. I., Choi, S. Y., Lee, J. H., . . . Kang, M. (2018). Development and Validation of a Deep Learning–based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. Clinical Infectious Diseases, 69(5), 739-747. doi:10.1093/cid/ciy967
- Islam, M. T., Aowal, M. A., Minhaz, A. T., & Ashraf, K. (2017). Abnormality detection and localization in chest x-rays using deep convolutional neural networks. arXiv preprint arXiv:1705.09850.
- Jo, T., Nho, K., & Saykin, A. J. (2019). Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data. Frontiers in Aging Neuroscience, 11. doi:10.3389/fnagi.2019.00220
- Millard, J., Ugarte-Gil, C., & Moore, D. (2015). Multidrug resistant tuberculosis. BMJ: British Medical Journal, 350. Retrieved October 25, 2020, from https://www.jstor.org/stable/26520021
- Nardell E. A. (2015). Transmission and Institutional Infection Control of Tuberculosis. Cold Spring Harbor perspectives in medicine, 6(2), a018192. https://doi.org/10.1101/cshperspect.a018192