



University of Reading
Department of Computer Science

An AI-assisted decision making system for thyroid nodule classification

Stefanos Stefanou

Supervisor: Huizhi Liang

A report submitted in partial fulfilment of the requirements of
the University of Reading for the degree of
Bachelor of Science in *Computer Science*

April, 2021

Declaration

I, Stefanos Stefanou, of the Department of Computer Science, University of Reading, confirm that all the sentences, figures, tables, equations, code snippets, artworks, and illustrations in this report are original and have not been taken from any other person's work, except where the works of others have been explicitly acknowledged, quoted, and referenced. I understand that if failing to do so will be considered a case of plagiarism. Plagiarism is a form of academic misconduct and will be penalised accordingly.

Stefanos Stefanou
April, 2021

Abstract

Deep learning has found numerous applications in the health care community. Recently, a massive explosion of research on the relevant field, driven by large amounts of available data, has generated important disease prevention and identification results. Fine Needle Aspiration (FNA) is the dominant procedure for thyroid nodule classification. FNA has associated risks and expenses, and in this project, we will try to reduce both using the recent advancements in Artificial Intelligence and Deep Learning. Our primary goal is to bring closer the radiologists 'on the field' with those complex algorithms and provide value to real patients by providing an interface, in the form of a web application, for probabilistically predicting the severity and the category of a given module.

Contents

Abstract	ii
1 Introduction	2
2 Literature Review	3
2.1 Brief Table of books	3
2.2 Brief Table of papers	3
2.3 Analytic Report: Books	4
2.3.1 Learning Path Visualised	4
2.3.2 (Probability/Statistics)ST1PS-18-9A: Probability and Statistics (2018/19)	4
2.3.3 (Linear Algebra)Linear Algebra and Its Applications	4
2.3.4 (Image Analysis)Digital Image Processing Author(s): Rafael C. Gonzalez, Richard E. Woods + CS3IA16-20-1A: Image Analysis (2020/21)	4
2.3.5 (Machine Learning)Deep Learning Book by Aaron Courville, Ian Goodfellow, and Yoshua Bengio	4
2.3.6 (Techniques)Deep Learning for Medical Image Analysis	4
2.3.7 (ToolStack)Hands-On Machine Learning with Scikit-Learn and TensorFlow:	4
2.3.8 Papers	5
3 Users Perspective	6
3.1 Introduction	6
3.2 Our Users	6
3.3 Our Users	6
4 Abstracted View of the System	7
4.1 Frontend Web App	7
5 The Frontend	8
5.0.1 Technology Stack	8
6 The Backend	10
6.1	10
6.2	10
6.2.1	10
6.3 Summary	10

7	The Service	11
7.1	...	11
7.2	...	11
7.2.1	...	11
7.3	Summary	11
8	Discussion, Conclusion and Future work	12
8.1	...	12
8.2	...	12
8.2.1	...	12
8.3	Summary	12
9	Reflection	13
9.1	...	13
9.2	...	13
9.2.1	...	13
9.3	Summary	13
	Appendices	15
A	Technology Stack	15
A.1	...	15
A.2	...	15
A.2.1	...	15
A.3	Summary	15

Acknowledgments

Acknowledgments section is optional. You may like to acknowledge the support and help of your supervisor(s), friend(s), or any other person(s), department(s), institute(s).

Chapter 1

Introduction

Deep learning has found numerous applications in the health care community. Recently, a massive explosion of research on the relevant field, driven by large amounts of available data, has generated important disease prevention and identification results. Fine Needle Aspiration (FNA) is the dominant procedure for thyroid nodule classification. FNA has associated risks and expenses, and in this project, we will try to reduce both using the recent advancements in Artificial Intelligence and Deep Learning. Our primary goal is to bring closer the radiologists 'on the field' with those complex algorithms and provide value to real patients by providing an interface, in the form of a web application, for probabilistically predicting the severity and the category of a given module.

Abbreviations

FNA(Fine Needle Aspiration), AI(Artificial Intelligence), DP(Deep Learning)Abbreviations

Keywords

FNA(Fine Needle Aspiration), AI(Artificial Intelligence), DP(Deep Learning)Abbreviations

Chapter 2

Literature Review

2.1 Brief Table of books

ISBN	Name	Type
N/A	ST1PS-18-9A: Probability and Statistics (2018/19)	Module Lectures
9780030105678	Linear Algebra and Its Applications	Book
9780131687288	Digital Image Processing	Book
9780262035613	Deep Learning	Book
9780128104088	Deep Learning for Medical Image Analysis	Book
9781491962244	Hands-on machine learning with scikit-learn and tensorflow	Book

2.2 Brief Table of papers

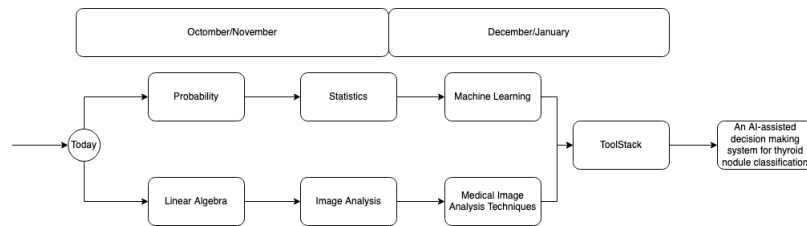
- Ye, H., Hang, J., Chen, X. et al. An intelligent platform for ultrasound diagnosis of thyroid nodules. Sci Rep 10, 13223 (2020). <https://doi.org/10.1038/s41598-020-70159-y>
- Nguyen DT, Pham TD, Batchuluun G, Yoon HS, Park KR. Artificial Intelligence-Based Thyroid Nodule Classification Using Information from Spatial and Frequency Domains. J Clin Med. 2019;8(11):1976. Published 2019 Nov 14. doi:10.3390/jcm8111976
- Manivannan T, Ayyappan N. Classification of thyroid nodules using ultrasound images. Bioinformation. 2020;16(2):145-148. Published 2020 Feb 29. doi:10.6026/97320630016145
- Nguyen DT, Kang JK, Pham TD, Batchuluun G, Park KR. Ultrasound Image-Based Diagnosis of Malignant Thyroid Nodule Using Artificial Intelligence. Sensors (Basel). 2020;20(7):1822. Published 2020 Mar 25. doi:10.3390/s20071822
- Chen J, You H, Li K. A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images. Comput Methods Programs Biomed. 2020 Mar;185:105329. doi: 10.1016/j.cmpb.2020.105329. Epub 2020 Jan 9. PMID: 31955006.

2.3 Analytic Report: Books

2.3.1 Learning Path Visualised

The reader should note that many of the learning nodes for the following diagram already existed in the student's program in parts 1 and 2. The reference here is solely for pointing out the need for a complete revision over the aforementioned topics and for clarity.

Figure 2.1: Learning Path



2.3.2 (Probability/Statistics)ST1PS-18-9A: Probability and Statistics (2018/19)

In part 1, my module *Probability and statistics* covers everything essential regarding my statistical background for this project.

2.3.3 (Linear Algebra)Linear Algebra and Its Applications

This excellent book will supplement my knowledge of linear algebra, covers almost anything that I will need later in the image analysis part.

2.3.4 (Image Analysis)Digital Image Processing Author(s): Rafael C. Gonzalez, Richard E. Woods + CS3IA16-20-1A: Image Analysis (2020/21)

This book, recommended by the lecturer in CS3IA16, will supplement my knowledge in image analysis and basic image transformation algorithms needed for the preprocessing part of the Machine learning service.

2.3.5 (Machine Learning)Deep Learning Book by Aaron Courville, Ian Goodfellow, and Yoshua Bengio

This excellent book is an introduction to Machine learning with Deep Learning techniques, much needed in the AI-analy

2.3.6 (Techniques)Deep Learning for Medical Image Analysis

This book has plenty of industry-standard techniques for medical image analysis. It will help me catch up with the latest research methods in this field.

2.3.7 (ToolStack)Hands-On Machine Learning with Scikit-Learn and TensorFlow:

This book will teach me the basic AI-toolkit stack, as well as the practical techniques for writing intelligent systems.

2.3.8 Papers

Around that time, i will start learning specialized techiques for this domain. The aforementioned papers will be my source of information.

Chapter 3

Users Perspective

3.1 Introduction

In this Section we will provide a useful manual for the end-users of our system.

3.2 Our Users

In Chapter 1 we briefly mentioned our system's purpose as ...

Definition 1. *An AI-assisted decision making system for thyroid nodule classification.*

As we are going to see later, various scientific methods are applied, taking into account multiple parameters, in order to produce a probabilistic result. This fact implies that the end-user will be an expert on Radiology, to understand the terms, and carefully interpret the results. From now on we assume that our system users will be experienced Radiologists

3.3 Our Users

In Chapter 1 we briefly mentioned our system's purpose as ...

Definition 2. *An AI-assisted decision making system for thyroid nodule classification.*

As we are going to see later, various scientific methods are applied, taking into account multiple parameters, in order to produce a probabilistic result. This fact implies that the end-user will be an expert on Radiology, to understand the terms, and carefully interpret the results. From now on we assume that our system users will be experienced Radiologists

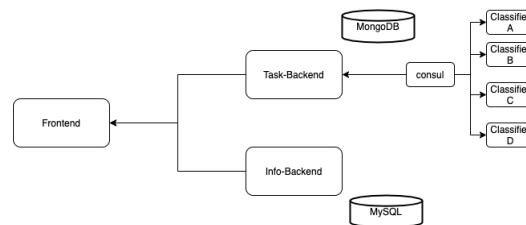
...

Chapter 4

Abstracted View of the System

In this chapter, we will introduce the architecture of our system, explaining the essential elements that is composed of, and their interactions.

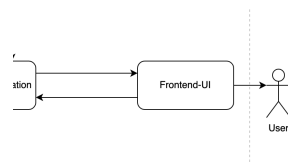
Figure 4.1: Essential System Components



4.1 Frontend Web App

The Frontend component has the responsibility of being the edge in our system.

Figure 4.2: Our system's edge



Every action from our users, should be channeled through the frontend. Our frontend is a web-based application, and as a consequence of that, a design decision is that the API between the web app and the backend application will not be a public one. This decision will increase the security of our system, as the process of writing spam bots will be significantly harder without a known API. More information about the API (Application programming interface) will be given below.

4.2 Backend

There is a number of design choices that we have made on our Backend System, in order to increase security, and decrease complexity. Our Backend System follows the design principles of the microservice pattern. Microservice pattern tries to decrease complexity and increase safety by splitting the internal logic of a system into several components called 'Microservices'.

Each microservice is essentially a server that handles a small portion of the systems logic. As opposed to the monolithic services, microservices have a number of advantages such as

- Highly maintainable and testable
- Loosely coupled
- Independently deployable
- No Single Point of Failure

4.2.1 Information Backend

The first of our services is the Information Backend. This service will have the responsibility to handle the information related to a scan, as well as its statistics and associations between scans and patients. The majority of the models composing our systems will be available through this service, via a well designed API.

4.2.2 Task Backend

This microservice will have the responsibility to trigger prediction and classification tasks for our system. The whole procedure, due to its CPU Intensive nature, will have to be asynchronous and to be executed in the background. The Frontend will send a request for a given task, and the server will have to return a token, associated for that particular task. Later, The frontend may request to learn the progress of its task or its results(if completed) by using the relevant token. This design choice is unavoidable given that the HTTP protocol has embedded the notion of 'timeout', it is just impossible and impractical to wait until a given task is complete. Another great advantage of this asynchronous design is the fact that multiple users may request Tasks without eliminating the server's resources, such as CPU time and amount of RAM available. Independent of the number of requests, the server will implement a queue FIFO (First-In-First-Out) strategy and it will inform its users when the task is ready to be seen.

4.2.3 Classification Backend

By using multiple classification techniques, our system will have eliminated the

Chapter 5

The Frontend

5.0.1 Technology Stack

In the construction of our system, we will need a number of open source technologies, libraries and standards to support our development. An exhaustive list is given below

- HTML5
- CSS3
- Javascript
- React.Js
- Bootstrap

HTML5

HTML5 is a markup language mainly used for structuring content on the World Wide Web. The its last major version(version 5.0) it is recommended by the World Wide Web Consortium (W3C). The responsible organisation WHATWG (Web Hypertext Application Technology Working Group) is a consortium of the major browser vendors(Apple, Google, Mozilla, and Microsoft)Gudeliauskas (2019).

CSS3

CSS stands for Cascading Style Sheets with an emphasis placed on “Style.” While HTML is used to structure a web document, CSS comes through and specifies your document’s style—page layouts, colors, and fonts are all determined with CSSMorris (2020). We will use CSS, version 3, to make our frontend application aesthecaly pleasing and easy-to-use for our end-users.

Javascript

longside HTML and CSS, JavaScript is one of the major technologies of the World Wide Web. JavaScript makes possible interactive web pages and is an integral part of web applications.

React

React (also known as React.js or ReactJS) is an open-source, front end, JavaScript library[3] for building user interfaces or UI components. It is maintained by Facebook and a community of individual developers and companies.[4][5][6] React can be used as a base in the development of single-page or mobile applications. However, React is only concerned with state management and rendering that state to the DOM, so creating React applications usually requires the use of additional libraries for routing.[7][8] React Router[9] is an example of such a library.

Bootstrap

Bootstrap is a free and open-source CSS framework directed at responsive, mobile-first front-end web development. It contains CSS- and (optionally) JavaScript-based design templates for typography, forms, buttons, navigation, and other interface components. Bootstrap is among the most starred projects on GitHub, with more than 142,000 stars, behind freeCodeCamp (almost 312,000 stars) and marginally behind Vue.js framework.[2]

Chapter 6

The Backend

...

6.1 ...

....

6.2 ...

...

6.2.1 ...

6.3 Summary

...

Chapter 7

The Service

...

7.1 ...

....

7.2 ...

...

7.2.1 ...

7.3 Summary

...

Chapter 8

Discussion, Conclusion and Future work

...

8.1 ...

....

8.2 ...

...

8.2.1 ...

8.3 Summary

...

Chapter 9

Reflection

...

9.1 ...

....

9.2 ...

...

9.2.1 ...

9.3 Summary

...

References

Gudeliauskas, D. (2019), 'The basics of hypertext markup language'.

URL: [*https://www.hostinger.co.uk/tutorials/what-is-html*](https://www.hostinger.co.uk/tutorials/what-is-html)

Morris, S. (2020), 'What is css, how does it work and what is it used for?'.

URL: [*https://skillcrush.com/blog/css/*](https://skillcrush.com/blog/css/)

Appendix A

Technology Stack

...

A.1 ...

....

A.2 ...

...

A.2.1 ...

A.3 Summary

...