**6CCE3EEP/7CCEMEEP**

**Individual Project Submission 2021/22**

**Name: Gustavo Masson Auriemo**

**Student Number: 19017464**

**Degree Programme: Electronic Engineering with Management**

**Project Title:**

**Supervisor: Dr. Bipin Rajendran**

**Word count:**

**RELEASE OF PROJECT**

Following the submission of your project, the Department would like to make it publicly available via the library electronic resources. You will retain copyright of the project.

I **agree** to the release of my project

I **do not** agree to the release of my project

**RELEASE OF VIDEO**

Following the submission of your project demonstration, the Department would like to make it publicly available via youtube. You will retain copyright of the project.

I **agree** to the release of my project video

I **do not** agree to the release of my project video

**Signature:** (An e-signature is acceptable, i.e., attach your signature as a picture, sign using the draw function)

**Date:03/04/2022**

**Originality Avowal**: "I verify that I am the sole author of this report, except where explicitly stated to the contrary."

**Abstract or Executive Summary**: This is a half to one-page summary of your report (rather than your project) listing your problem statement and main findings.

**Contents Page**: It is very useful to have a high-level contents page, list of figures and tables in your report as well as a list of appendices. It will also help you when you structure the report.

**Introduction**: Your introduction really sets the scene for the report. It should be a clear statement of what the project is about, a summary of the background and context and summarise what you set out to achieve.

**Background**: The background is not just a literature review but explains why your project is important and explores the theory and previous work. A good literature review synthesises and critically evaluates the existing literature identifying gaps in current knowledge that the project will fill rather than simply describing previous work.

**Report**: The next part of the report might be then broken into chapters such as methodology, findings, specification, requirements, design etc. The aim of the section as a whole is to describe the work you've done, justify your approach and explain how you arrived at your conclusion. You might also analyse the technical findings of your work.

**Conclusion**: here you can summarise the project again, make any conclusions, statements or assertions that you believe your project has achieved and offer some ways that the project might be taken forward in future.

**Professionalism and Responsibility:**It is important that you consider the professional influences on your project such as standards and competencies. You can also discuss general ethics, sustainability, cyber-security or other issues applicable to the project.

**Bibliography**: a list of all your references following the acceptable [college citation format.](https://libguides.kcl.ac.uk/reference)   
(Martin et al., 2020) or (Salas & D’Agostino, 2020) for two authors. For every in-text citation, there must be a reference list one: APA

**Appendix**: additional useful information that won't be marked but provides some completeness E.g. tables of data, additional graphs etc

# **Abstract**

This report sets out a scalable method for harmony recognition utilizing the Neural Engineering Framework (NEF) developed by Eliasmith & Anderson.

A proof-of-concept is shown, along with results utilizing a restricted dataset with four different learning rates over 10 epochs.

Three different approaches to consider the seriousness of error were taken

Table of Contents

[Abstract 3](#_Toc99905030)

[Introduction 3](#_Toc99905031)

[Background 3](#_Toc99905032)

[Report 6](#_Toc99905033)

[Conclusion 6](#_Toc99905034)

[Professionalism and Responsibility 7](#_Toc99905035)

[Bibliography 7](#_Toc99905036)

[Appendix 7](#_Toc99905037)

# **Introduction**

# **Background**

The nature of this project is interdisciplinary, and as such, basic knowledge of each area will benefit the reader to better understand how each aspect connects to the other in a meaningful way.

Key words will be in bold and will be referenced throughout the report.

## Harmony theory

A musical **note (or tone)** is, in its most simple terms, a frequency. For example, when a string in a well-tuned guitar is plucked, it oscillates at this frequency and produces a sound. Now, if two strings are plucked at the same time, the interaction of such frequencies creates new patterns of oscillations, and subsequently new sounds. Such sounds are dependent on how far apart the frequencies between these two notes are.

It was defined that every note with a frequency ratio of 2 (or 0.5) to a subsequent note would be denoted as an **octave**, as due to their closely related harmonics they sound extremely similar (to the human ear).For example, given a note with a frequency of 440Hz, both 880Hz and 220Hz would be **octaves** of this frequency.

Every octave is denoted by the same letter, accompanied by a number to denote the frequency. Considering the example above, 440Hz is denoted “A4”, whereas 880Hz and 220Hz are “A5” and “A3”, respectively.

In the Western world, the most common way to divide the range of frequencies between two octaves is in 12 equally spaced frequencies on the logarithmic scale. This frequency ratio can be written as , where is the number of notes between the two frequencies, this is called an **interval** ( can also be called as the number of **semitones** between the two notes).

Figure 1 below shows **octaves** for a range of frequencies. Note that the “accidentals” (# and b) accompanying some of the notes are a matter of how the labelling system was devised, but matters not for the purpose of this report, as frequencies and ratios will be mostly used.

Table

Description automatically generated

Figure 1: Notes, their octaves, and their frequencies[[1]](#footnote-1)

## For example, between C1 and D1 we have two **semitones** (), which in terms of the frequency ratio is equivalent to the ratio between G4 and A4: in both cases their ratio is .

When two or more notes are played simultaneously, we call it **harmony**. Naturally, the more frequencies played at a time, the more complex their interactions become, and the harder is it to distinguish individual notes. **Harmony** and **chords** in this context will be used interchangeably as only one instrument is considered.

Convention has set names for combinations of notes: one may call the chord made up of the notes C4, E4, and G4, a C Major chord. However, the same name would be attributed to a chord made up of C3, E4, and G5, with notes in different octaves. The latter is called a **voicing** of C Major and is extremely important in adding flair to music, something musicians often aim to do. The importance of voicings will become clearer once the literature is discussed.

## Neural Networks

The field of Machine Learning at large is focused on developing efficient pattern recognition methods that can scale well with the size of the problem domain and of the data sets (Simeone, 2017). For this reason, many model classes, algorithms, and training methods have been devised, and are applied according to the purpose desired. One such model class is called Neural Network. In simple terms, Neural Networks aim to extract the most useful features of the data and utilise such features to infer a result. This model is mostly used for situations in which it is difficult to determine good features of the data a priori. For example, if the input is a text, the features may be determined as a vector with the number of occurrences of given words. If the input was an image, however, such decision as to what good features are becomes harder; and that is the main role of the Neural Network model.

Through a supervising signal, which determines whether the output of the model was correct or not, the parameters of the model (also referred to as weights) are updated to reflect such error, and hopefully yield a better result. The method in which they are updated is usually by taking the gradient of the loss function over a small set of data points (called mini batch) and changing the weights according to the negative direction of such gradient (i.e., minimizing the loss). The loss function is set a priori and depends on the nature of the problem. It may look something like 𝟙, which means that if prediction , the loss will zero, and otherwise it will be one.

Say, it is hard to

The way in which this model class extracts such features from data is trainable.

are non-linear models used for

Diagram, schematic

Description automatically generated

The binary nature of spikes means that for any given input for the neurons, the output neuron is only required to sum the weights of the corresponding neurons that spiked (Rajendran, 2021)

Figure 2 - Computations in ANN VS SNN

## Neural Engineering Framework & Nengo

The Neural Engineering Framework (**NEF**) is a methodology that allows the construction of large-scale, biologically plausible neural models of cognition (Eliasmith & Anderson, 2003). Instead of setting connection weights between **neurons** manually or through some learning rule, the NEF solves for these according to the function you desire to compute. For example, say one wants to compute , the framework will determine the connection **weights** between two ensembles of neurons that best approximates this function (the way in which it does will be discussed later). In the case where such function is not known, traditional methods can still be utilised.

*Why model according to the NEF?*

Considering that the **NEF** uses computational units based on biological neurons, any computation done by the model is forced to adhere to the basic operations that are available to neurons (Stewart, 2012). This allows insight into of what sort of algorithms can and cannot be implemented in the human brain. Further, the representation of signals is done so as match the neuron ensemble behaviour seen in the brain, which is a direct effect of how neurons behave.

In 2014, Bekolay et al. developed an open-source Python package called **Nengo**[[2]](#footnote-2), implementing the NEF for building and simulating such models in a computer environment. Given the goal of this project, all the models in this report were built on **Nengo**, following the **NEF**.

why spiking

For a given signal vector , the NEF uses an encoding vector to represent neuron activity in a given neuron . Assuming that the input current of the neuron is a linear function of the actual values being represented

Consequently, the learning methods within the framework are distinct from that of traditional machine learning (Stochastic Gradient Descent generally). Discuss PES learning rule

## Literature and General Remarks

The current literature in the subject of harmony recognition is vast, but so is the complexity of the problem. Considering the several steps inherent to the problem; from the processing of the audio, to design choices of classification algorithms, to the scope of the solution, and the choice of databases, there were many unexplored aspects which were considered and utilised.

Most importantly, this model differs in its scope and approach. That is, it is focused solely on the piano, in utilising a biologically plausible approach for analysing the data and learning, and in being able to classify harmony within a significantly higher number of possible outputs. Further to the last point, contrary to most of the research which focuses on pre-determined structures of musical chords, which can be given by 24 different names/values (Cheng et al., 2008), the model focuses on any possible combination of notes within a piano. Given that a piano has 88 notes, of which three are selected at a time to make up the chord, there are combinations possible.

Further, the main audio processing technique used by researchers, Pitch Class Profiles (Fujishima, 1999), reduces the dimensionality of audio data by restricting the representation of the signal to a 12-bin vector (i.e., one **octave**). In order words, after mapping the audio from time domain to frequency domain (through Fourier transform), the frequencies are separated at their note-boundaries (i.e., ), and the sum of their respective magnitudes is taken and assigned to a note-bin. All the bins for a given note are then summed, losing any information about octaves, and subsequently, about **voicings**.

Chart

Description automatically generated

Figure 3 – 88-bin representation of a C chord with an alternate voicing

Insert two images here: one with the 12-bin and one of my model with 88 bins

Inspired by the Cochlea in the human ear, this report chose to maintain the logarithm scale format, but adapt it to range of a piano, namely with 88 bins ranging from 27.5Hz to 4186Hz, one for each note of the piano. This allows for a representation of the position of the notes within the instrument and addresses the issue of representing **voicings**.

only recently have machine learning methods (rather than relying on signal processing techniques and pattern matching), began to be used.

# **Report**

Things I tried

What I succeeded in, and substantiate with results (and mention how significant these results are! Show why they are important) —> look at the poster for the CHD class and the questions

What I failed in and why

## Methodology

### The dataset

## Framework

## Design

Findings

# **Conclusion**

# **Professionalism and Responsibility**

# **Bibliography**

Eliasmith, C., & Anderson, C. H. (2003). Neural engineering: Computation, representation and dynamics in neurobiological systems. Cambridge, MA: MIT Press.

Stewart, T. C. (2012). (tech.). *A Technical Overview of the Neural Engineering Framework*.

Rajendran, B. (2021). *Brain-Inspired Computing & Hardware Design.*

Osvaldo Simeone (2017). A Brief Introduction to Machine Learning for Engineers. Available at: <https://nms.kcl.ac.uk/osvaldo.simeone/notesMLSimeone.pdf>

Heng-Tze Cheng, Yi-Hsuan Yang, Yu-Ching Lin, I-Bin Liao and H. H. Chen, (2008). "Automatic chord recognition for music classification and retrieval," 2008 IEEE International Conference on Multimedia and Expo, pp. 1505-1508, doi: 10.1109/ICME.2008.4607732.

Engineering ToolBox, (2003). *Notes, Octaves and Frequencies*. [online] Available at: <https://www.engineeringtoolbox.com/note-frequencies-d_520.html>

Takuya Fujishima (1999). Realtime Chord Recognition of Musical Sound: A System Using Common Lisp Music, ICMC Proceedings.

# **Appendix**

1. Engineering ToolBox, (2003). *Notes, Octaves and Frequencies*. [online] Available at: https://www.engineeringtoolbox.com/note-frequencies-d\_520.html. [↑](#footnote-ref-1)
2. <https://www.nengo.ai/> [↑](#footnote-ref-2)