R. David Dunphy
MEng Computer and Electronic Systems
Supervisor: Marc Roper

University of Strathclyde Glasgow

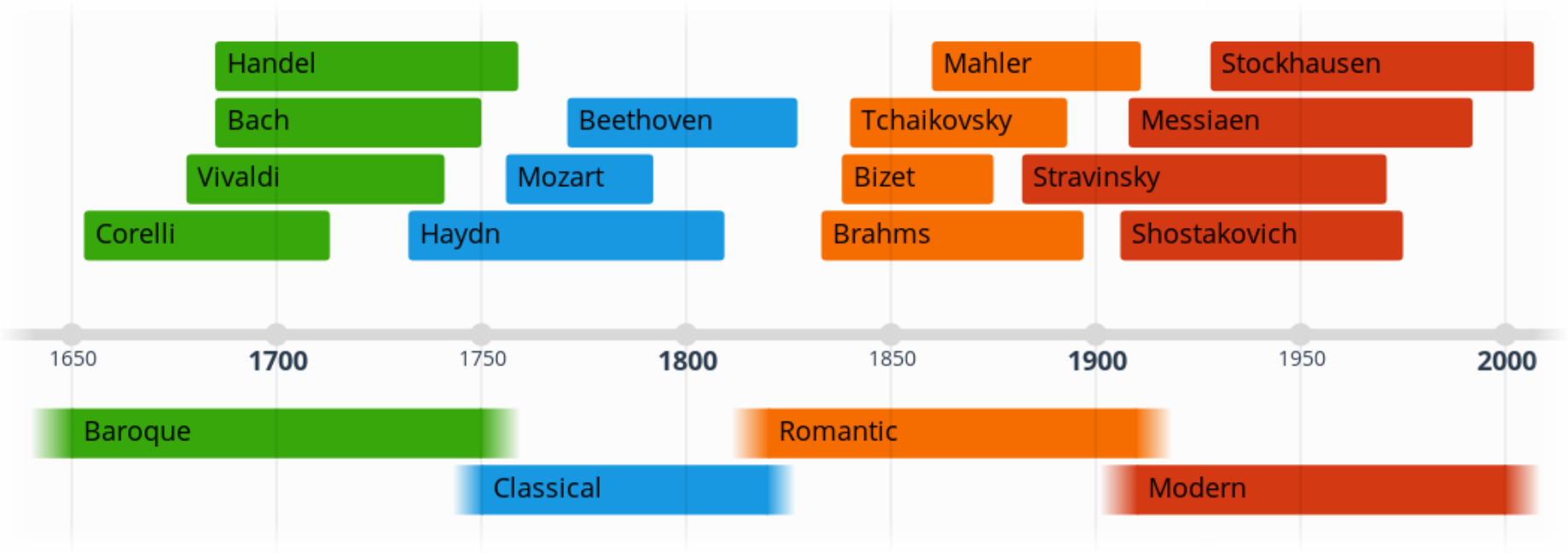
Applying Machine Learning Techniques to the Classification of Classical Orchestral Music

Objectives

- Analyse musical extracts in order to identify features useful for automatic classification
- Create an application capable of classifying music by composition period using a deep learning algorithm
- Compare the performance of deep learning algorithms to traditional machine learning algorithms
- Extend the application to attempt to identify the composer of a piece of music

Overview

- Orchestral music can be categorised into compositional periods. Although the exact dates of these periods are somewhat subjective, most music can be broadly identified as either Baroque, Classical, Romantic, or Modern.
- Trained musicians are generally able to identify the period or even the composer of a piece of music based on stylistic features, such as harmonic and rhythmic complexity or melodic "mannerisms".
- Classification relies on identifying subtle and often subjective features of music, making this a non-trivial problem to solve programmatically. This project aims to explore which modern machine learning techniques can usefully be applied to this domain.



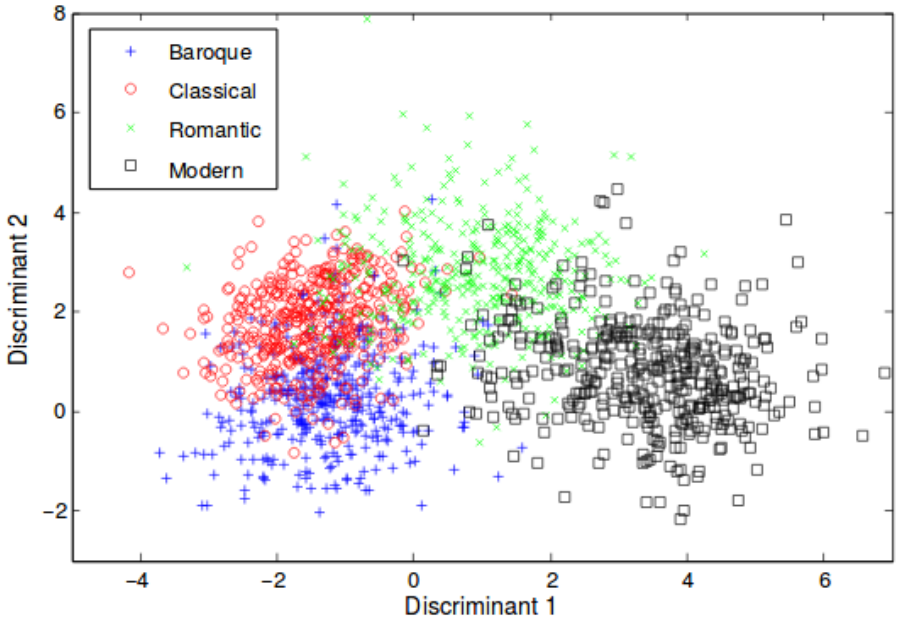
Existing Research

The following represent a sample of existing research which can be built upon, as well as an overview of the current state of the art in music classification.

- Weihs *et al.*, "Classification in Music Research" (2007): A detailed overview of a number of preprocessing techniques useful for music classification.
- Fujinaga, "Automatic Genre Classification Using Large High-Level Musical Feature Sets" (2004): Achieved over 90% success rate classifying music into genres "classical", "jazz", and "pop".
- Weiss *et al.*, "Timbre-invariant Audio Features for Style Analysis of Classical Music" (2014): Achieved around 82% success classifying orchestral and piano music using support vector machines to analyse triads.

While significant progress has been made in this field, further research is required to determine which features and algorithms are best suited to this task.

This could facilitate a more fine-grained approach capable of identifying composers, as well as increase success rates to near-human levels.



Results from Weiss et al.

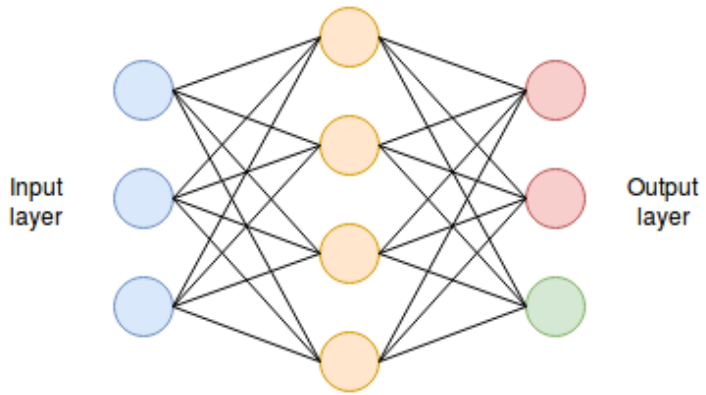
Machine Learning Approaches

For this project supervised learning algorithms will be used as classifiers. These algorithms analyse features of labelled training data and then try to predict the labels of test data.

Conventional Algorithms

The simplest classifiers are traditional algorithms, such as k-nearest neighbours (k-NN) and support vector machines (SVMs).

- Have the advantage of speed and simplicity.
- Require features to be extracted manually.



Symbolic diagram of a simple neural network

Deep Learning Algorithms

Deep learning algorithms such as neural networks offer a modern and more powerful alternative.

- Convolutional neural networks (CNNs) are powerful tools for analysing images, which could be used if tracks can be visualised.
- Recurrent neural networks (RNNs) have been extremely successful at tasks with temporal components, such as speech recognition.
- Slower than traditional approaches, as a large corpus of data is required.
- Corpus needs to be manually assembled and labelled.
- In order to speed up training, a simple representation of the data is required.
- LSTMs can take into account higher-level structure as well as low-level features.

Data Processing and Feature Extraction

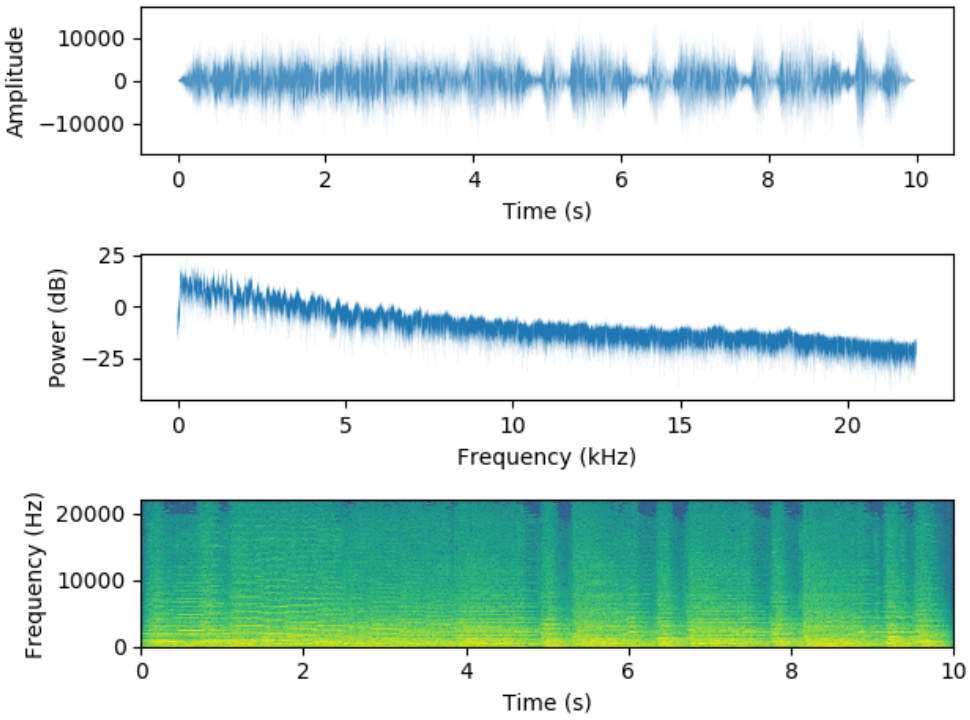
In order to classify music, the raw audio data has to be processed and features suitable for training extracted. Low-level features that can be extracted using signal processing techniques include:

- Timbre (used to identify instruments)
- Pitch range and pitch classes
- Rhythm, tempo, and metre
- Dynamic range (i.e. signal amplitude)

All of these can be visualised in a spectrogram, which can be supplied to a CNN. This would likely result in a timbre-driven approach.

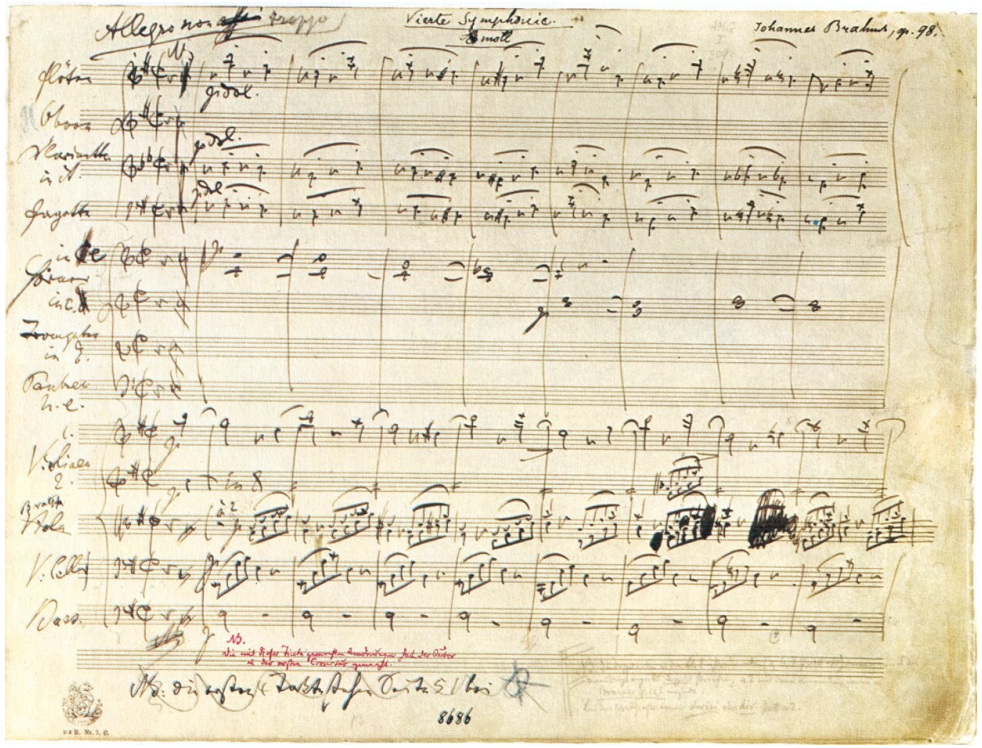
For the purposes of an RNN, a sequence of temporal events that represent the essential features of the music need to be extracted, similar to notes in a MIDI file.

In order to do this, pitch classes must be identified. This can be done using a combination of onset detection and fast Fourier transforms (FFTs) across overlapping windows.



10 seconds from Beethoven's Symphony No. 5, represented in time domain, frequency domain, and as a spectrogram.

Conclusion and Work Plan



- Several ML approaches to classifying music are possible, including deep learning algorithms that were not feasible for complex tasks until very recently.
- This project aims to attempt and compare several of these approaches and evaluate which methods are well-suited to this task.
- A large corpus of music tracks by more than 20 composers is being assembled and labelled.
- A core part of the project consists of extracting musical features from raw audio data.