

Machine Learning Project

Prof: Paola Velardi

Recurrent Neural Nets for music genre classification using Keras

[1. Introduction 3](#_Toc51319885)

[2. Recurrent neural-network 3](#_Toc51319886)

[2.1. Long short-term memory 3](#_Toc51319887)

[3. Implementation 4](#_Toc51319888)

[4. Features extracted 4](#_Toc51319889)

[4.1. MFCC 4](#_Toc51319890)

[4.2. Spectral Centroid 4](#_Toc51319891)

[4.3. Chroma 4](#_Toc51319892)

[4.4. Spectral Contrast 4](#_Toc51319893)

[5. Training 4](#_Toc51319894)

[6. Testing 4](#_Toc51319895)

[7. Conclusions 4](#_Toc51319896)

[8. References 4](#_Toc51319897)

# Introduction

# Recurrent neural-network

A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior. Derived from feedforward neural networks, RNNs can use their internal state (memory) to process variable length sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition.

The term “recurrent neural network” is used indiscriminately to refer to two broad classes of networks with a similar general structure, where one is finite impulse and the other is infinite impulse. Both classes of networks exhibit temporal dynamic behavior. A finite impulse recurrent network is a directed acyclic graph that can be unrolled and replaced with a strictly feedforward neural network, while an infinite impulse recurrent network is a directed cyclic graph that can not be unrolled.

Both finite impulse and infinite impulse recurrent networks can have additional stored states, and the storage can be under direct control by the neural network. The storage can also be replaced by another network or graph, if that incorporates time delays or has feedback loops. Such controlled states are referred to as gated state or gated memory, and are part of long short-term memory networks (LSTMs) and gated recurrent units. This is also called Feedback Neural Network.

# Long short-term memory

There are several architectures of LSTM units. A common architecture is composed of a cell (the memory part of the LSTM unit) and three "regulators", usually called gates, of the flow of information inside the LSTM unit: an input gate, an output gate and a forget gate. Some variations of the LSTM unit do not have one or more of these gates or maybe have other gates. For example, gated recurrent units (GRUs) do not have an output gate.

Intuitively, the cell is responsible for keeping track of the dependencies between the elements in the input sequence. The input gate controls the extent to which a new value flows into the cell, the forget gate controls the extent to which a value remains in the cell and the output gate controls the extent to which the value in the cell is used to compute the output activation of the LSTM unit. The activation function of the LSTM gates is often the logistic sigmoid function.

There are connections into and out of the LSTM gates, a few of which are recurrent. The weights of these connections, which need to be learned during training, determine how the gates operate.

# Implementation

# Features extracted

The features to be analyzed and exported are the following:

* MFCC
* Spectral Centroid
* Chroma
* Spectral Contrast

# MFCC

# Spectral Centroid

# Chroma

# Spectral Contrast

# Training

# Testing

# Conclusions

# References

GTZAN+: https://perso-etis.ensea.fr//sylvain.iloga/GTZAN+/