

Cadenza - generowanie kolejnych nut melodii za pomocą technik Machine Learningu

Wojciech Szlosek



Przypomnienie: temat projektu

Tematem było generowanie kolejnych nut muzycznych na bazie danej próbki początkowej.

Zaczniemy od przykładowych efektów, które udało się uzyskać:





Dataset

Fragment datasetu Maestro: 500 plików w formacie MIDI.

Instrument: Pianino. Klasyczny duch utworów.

Obróbka datasetu

1. Zmiana plików muzycznych midi na listę stringów (nut i akordów)
2. Stworzenie konwertera nuty na liczbę i liczby na nutę; konwersja listy nut na listę liczb
3. Zmiana każdej liczby na wektor binarny, gdzie pozycja 1 oznacza wartość liczbową nuty
4. Stworzenie zbiorów X i Y. $X[:20] \Rightarrow Y[20]$, $X[1:21] \Rightarrow Y[21]$ itd.

Przykład dla 22-elementowej listy nut/akordów:

$['2+7', 'B3', '11+2', \dots, 'B3', '2+7'] = [535, 1267, 391, \dots, 1267, 535] = [[0, 0, 0, \dots, [535]: 1, 0, \dots], \dots]$

$X[0] = ['2+7', 'B3', '11+2', '7+9', '2+7', 'D4', 'B3', 'G2', 'D3', 'G3', 'D4', 'D3', 'D4', '2+7', 'B3', '11+2', '7+9', '2+7', '0+4', 'C3']$

$Y[0] = 'B3'$

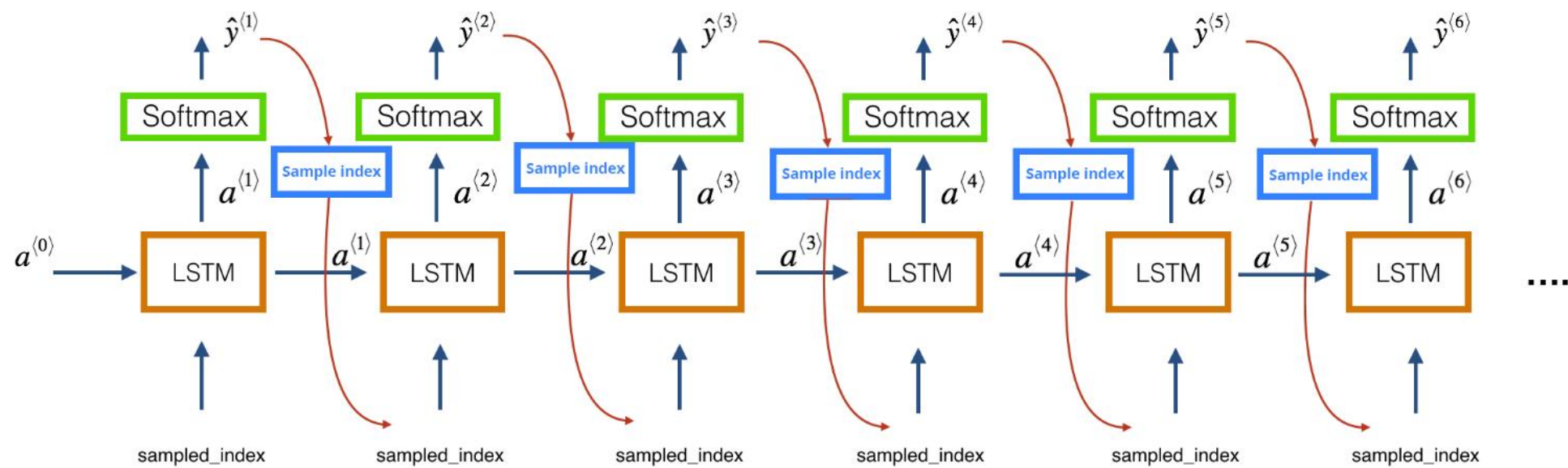
ltd.

Architektura modelu i użyte technologie:

Głównie: Python, TensorFlow, music21

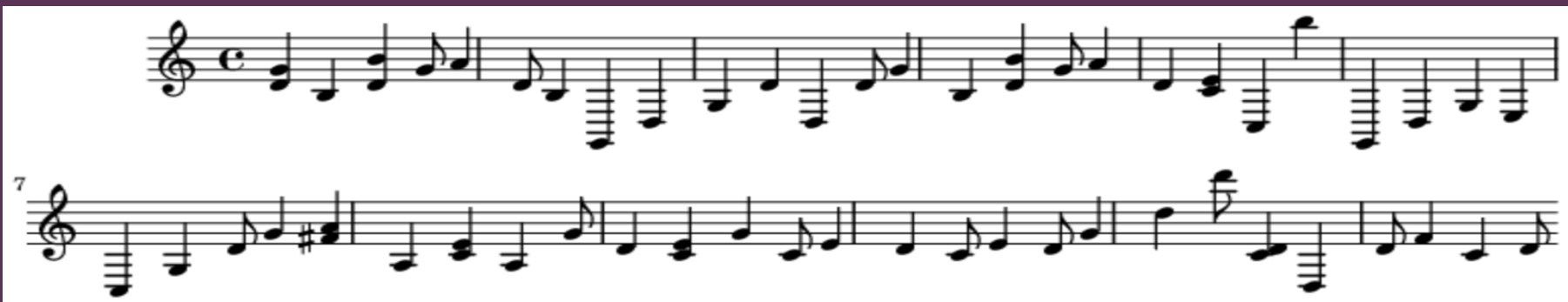
Model stanowi sieć warstw LSTM z końcową warstwą Dense. Architektura zgodna z planem z pierwszej prezentacji projektu.

```
model = Sequential()  
model.add(LSTM(128, return_sequences=True))  
model.add(LSTM(80))  
model.add(Dense(unique_notes, activation="softmax"))  
model.compile(loss='categorical_crossentropy', optimizer=tf.keras.optimizers.Adam())
```



Rezultaty predykcji są subiektywnie zadowalające.

Matematyczne określenie rezultatów nie jest możliwe, jednak w opinii mojej i znanych osób mających kontakt z muzyką, predykcja jest zaskakująco sensowna – najczęściej z zachowaną tonacją.



Napotkane trudności - i jak sobie z nimi poradziłem

- Obróbka plików midi - własne pomysły oraz czytanie plików midi za pomocą music21
- Jak odnaleźć regularność w muzyce? - założenie implikacji $X \Rightarrow Y$
- Mozolne próby konfiguracji architektury modelu: początkowy rezultat w postaci ciszy lub dźwięków spoza tonacji
- Problemy z pamięcią i złożonością - m.in. wykorzystanie biblioteki pickle

Wnioski. Czego się nauczyłem?

Był to mój pierwszy projekt z dziedziny Machine Learningu, więc każdy krok był jakąś nauką. Poznałem rolę obróbki danych do treningu oraz bibliotekę TensorFlow. Poza tym pogłębiłem wiedzę o sieciach RNN, szczególnie LSTM.

Praca na plikach muzycznych była fascynująca, a rezultaty okazały się bardzo ciekawe.

The header features a dark purple background with a series of overlapping semi-circular shapes. Each semi-circle is filled with a different pattern: some have concentric solid lines, some have concentric dashed lines, and some are solid dark purple.

**Więcej: na Githubie i w
sprawozdaniu.**

<https://github.com/wszlosek/Cadenza>

The top of the slide features a decorative header with a dark purple background. It contains several overlapping semi-circular shapes. Some of these shapes are filled with a lighter shade of purple, while others are outlined with concentric arcs or a pattern of small, radiating lines.

Questions?