

ОСОБЛИВОСТІ ЗАСТОСУВАННЯ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ (RNN) ДЛЯ КЛАСИФІКАЦІЇ МУЗИКИ

Вінницькій національний технічний університет

Анотація

У статті досліджено застосування рекурентних нейронних мереж (RNN) для класифікації музики. Розглянуто їхню архітектуру, метод зворотного поширення помилки в часі (BPTT).

Ключові слова: RNN, класифікація музики, машинне навчання, часові залежності.

Abstract

The article explores the application of recurrent neural networks (RNN) for music classification. Their architecture and the backpropagation through time (BPTT) method are discussed.

Keywords: RNN, music classification, machine learning, temporal dependencies.

Рекурентні нейронні мережі (RNN) [1] є потужним інструментом для обробки послідовних даних, що робить їх особливо корисними для задач класифікації музичних творів. Музика за своєю природою є послідовністю звуків, які розвиваються в часі, тому моделі, здатні враховувати часові залежності, мають перевагу в аналізі таких даних.

На відміну від звичайних нейронних мереж, RNN мають зворотні зв'язки, що дозволяє їм "пам'ятати" інформацію з попередніх кроків. Це досягається шляхом введення прихованого стану h_t , який залежить від вхідних даних x_t та попереднього прихованого стану h_{t-1} :

$$h_t = \phi(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (1.1)$$

де:

ϕ – нелінійна функція активації (наприклад, tanh);

W_{xh} та W_{hh} – матриці ваг;

b_h – вектор зміщень.

Вихід мережі y_t визначається як:

$$y_t = \sigma(W_{hy}h_t + b_y) \quad (1.2)$$

де σ — функція активації (наприклад, softmax для класифікації).

Для того щоб навчити RNN ефективно розпізнавати музичні жанри, використовується метод зворотного поширення помилки в часі (Backpropagation Through Time, BPTT) [2]. Цей метод дозволяє обчислити похибку на кожному часовому кроці та оновити ваги мережі з урахуванням всіх попередніх станів. Функція втрат у цьому випадку зазвичай обирається як крос-ентропія:

$$L = - \sum_{t=1}^T \sum_{k=1}^K y_{tk} \log(\hat{y}_{tk}) \quad (1.3)$$

де:

T – кількість часових кроків;

K – кількість класів;

y_{tk} – істинна мітка класу k у момент часу t ;

\hat{y}_{tk} – прогнозована ймовірність приналежності до класу k у момент часу t .

Рекурентні нейронні мережі є надзвичайно потужним інструментом для класифікації музики завдяки їхній здатності враховувати часові залежності в аудіосигналі. Їхнє поєднання з іншими сучасними методами машинного навчання, такими як CNN та механізм уваги, дозволяє створювати ефективні та гнучкі системи. Ці гібридні моделі підвищують точність класифікації та відкривають нові

можливості в аналізі музичних творів, що є надзвичайно актуальним у сучасному світі цифрової музики.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Рекурентна нейронна мережа. Вікіпедія. Електронний ресурс. Режим доступу: https://uk.wikipedia.org/wiki/Рекурентна_нейронна_мережа.

2. Graves, A., et al. "Supervised sequence labelling with recurrent neural networks." Studies in Computational Intelligence, 2012. Електронний ресурс. Режим доступу: <https://link.springer.com/book/10.1007/978-3-642-24797-2>.

Каспров Владислав Олександрович – студент групи 2ПІ-23м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, vlad.kasprov@gmail.com.

Kasprov Vladyslav Oleksandrovych – a student of group 2PI-23m, Faculty of Information Technology and Computer Engineering, Vinnytsia National Technical University, vlad.kasprov@gmail.com.