

МАТМАТИЧНІ МОДЕЛІ І МЕТОДИ ПРОГНОЗУВАННЯ ЗАБРУДНЕННЯ ПОВІТРЯ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Кит М. О.

Науковий керівник – доц., Есілевський В.С
Харківський національний університет радіоелектроніки
(61166, Харків, пр. Науки, 14, каф. Прикладної математики,
тел. (057) 702-14-36), e-mail: mykyta.kyt@nure.ua

The work considers the problem of building an air pollution prediction model. Air quality control is important for achieving good environmental conditions. This type of control is performed by means of air monitoring. The solution of the problem is based on recurrent neural networks, or more precisely LSTM networks. The experiment demonstrated that recurrent networks are designed to solve this problem. LSTM networks are networks with cyclic links that allow storing information.

В роботі розглядається методи побудови рекурентної нейронної мережі для прогнозування забруднення повітря. Основою розв'язку проблеми є LSTM-мережа[1]. В цій роботі в якості входних змінних обрані дрібнодисперсні частинки PM1.0 (мкг/м³), дрібнодисперсні частинки PM2.5(мкг/м³), дрібнодисперсні частинки PM10.0(мкг/м³), оксид азоту NO2 (мкг/м³), аміак NH3 (мг/м³), монооксид вуглецю (мг/м³), температура (°C), вологість (%). Але у ході проведення аналізу даних, було з'ясовано, що частинки PM1.0, PM2.5 та PM10.0 мають високу кореляцію між собою також частинки NO2 та CO мають високу кореляцію, тому було прийнято рішення видалити зайві ознаки. LSTM в даний час є стабільною технікою для розв'язання проблем прогнозування послідовності, тобто підходять для вирішення нашої проблеми.

Мережі довго-короткострокової пам'яті (Long Short Term Memory) — зазвичай просто називають «LSTM», які здатні до навчання довгостроковим залежностям. Вони працюють неймовірно добре на великій різноманітності проблем і в даній ситуації добре підходять. LSTM спеціально спроектовані таким чином, щоб уникнути проблеми довгострокових залежностей. Запам'ятовувати інформацію на тривалий період часу — це практично їх поведінка за замовчуванням, а не щось таке, що вони тільки намагаються зробити. Все рекурентні нейронні мережі мають форму ланцюга з рекурентних модулів нейронної мережі.

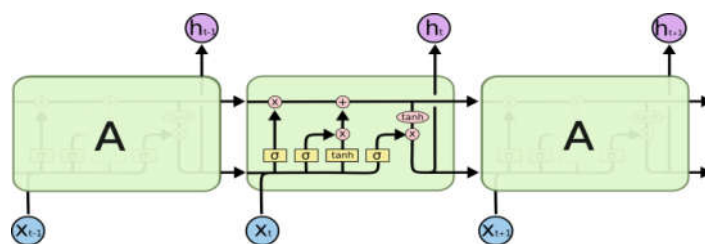


Рисунок 1. — Рекурентний модуль в LSTM

У цій роботі були реалізовані Recurrent Neural Network і дві різні архітектури для LSTM. Я використовував дані яке були отримані з моніторингових датчиків з проміжком в 30 хвилин. Найкращу продуктивність забезпечує stacked LSTM, що складається з декількох прихованих шарів.

Обчислення, пов'язані з LSTM:

$$\begin{aligned}
 i_{(t)} &= \sigma(W_{xi}^T \cdot x_{(t)} + W_{hi}^T \cdot h_{(t-1)} + b_i) \\
 c_{(t)} &= f_{(t)} \otimes c_{(t-1)} + i_{(t)} \otimes g_{(t)} \\
 y_{(t)} = h_{(t)} &= o_{(t)} \otimes \tanh(c_{(t)}) \\
 f_{(t)} &= \sigma(W_{xf}^T \cdot x_{(t)} + W_{hf}^T \cdot h_{(t-1)} + b_f) \\
 o_{(t)} &= \sigma(W_{xo}^T \cdot x_{(t)} + W_{ho}^T \cdot h_{(t-1)} + b_o) \\
 g_{(t)} &= \tanh(W_{xg}^T \cdot x_{(t)} + W_{hg}^T \cdot h_{(t-1)} + b_g)
 \end{aligned}$$

де $W_{xi}, W_{xf}, W_{xo}, W_{xg}$ — матриці ваг кожного з чотирьох слоїв для їх зв'язку з вхідним вектором $x_{(t)}$.

$W_{hi}, W_{hf}, W_{ho}, W_{hg}$ — матриці ваг кожного з чотирьох слоїв для їх зв'язку з попереднім короткостроковим станом $h_{(t-1)}$.

b_i, b_f, b_o, b_g — члени зміщення для кожного з чотирьох слоїв.

$h_{(t)}$ — короткостроковий стан (яке дорівнює виходу для даного тимчасового кроку $y_{(t)}$).

f — шлюз забування.

i — вхідний шлюз.

o — вихідний шлюз.

Проведений обчислювальний експеримент показав збіжність процесу навчання і принципову можливість застосування запропонованого підходу в прогнозуванні забруднення повітря.

Безумовно є ряд речей, які заслуговують подальшого вивчення, які можуть поліпшити продуктивність моделі. Наприклад можна використовувати дані про інші види забруднення в якості характеристик для прогнозування забруднення, наприклад SO₂. Але, можливо, інші види забруднення будуть корелювати із забрудненням SO₂ або іншими.

Власне таким чином, застосування нейронних мереж дозволяє побудувати прогностичну модель забруднення атмосферного повітря, що дає можливість своєчасного реагування та прийняття управлінських рішень.

Список використаних джерел:

1. Жерон Орельєн «Прикладне машинне навчання за допомогою Scikit-Learn та TensorFlow» Вільямс, 2018 рік, 688 стр.