

УДК 004.81

ИССЛЕДОВАНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ПОКАЗАНИЙ ПРИБОРОВ УЧЕТА РЕСУРСОВ.

Долженко А.И.

доктор экономических наук, доцент, РГЭУ (РИНХ)

alexdoljenko@mail.ru

Состин И.В.

магистрант, РГЭУ (РИНХ)

vanessence@rsue.ru

Аннотация: Приведено описание архитектур нейронных сетей. Представлены их структуры и примеры реализации с использованием библиотеки Tensorflow и надстройкой Keras.

Ключевые слова: сеть прямого распространения, рекуррентная нейронная сеть, свёрточная нейронная сеть, долгая краткосрочная память, управляемые рекуррентные блоки.

THE STUDY OF MACHINE LEARNING ALGORITHMS FOR RECOGNIZING THE READINGS OF RESOURCE METERING DEVICES

Dolzhenko A.I.

Sostin I.V.

Abstract: A brief description of the architecture of neural networks is provided. Their structures and implementation examples are presented using the Tensorflow library and the Keras add-in.

Keywords: direct distribution network, recurrent neural network, convolutional neural network, long-term short-term memory, controlled recurrent blocks.

Основные задачи, решаемые при помощи методов машинного обучения, относятся к двум разным видам: обучение с учителем либо без него [1]. Однако этим учителем вовсе не обязательно является сам программист, в терминах машинного обучения – это само вмешательство в процесс обработки информации.

Целью исследования является применение алгоритма машинного обучения с учителем в разных моделях нейронных сетей при различных параметрах обучения моделей и сравнение их точности при схожих условиях. В работе были рассмотрены сеть прямого распространения (feedforward neural network, FNN), рекуррентная нейронная сеть (recurrent neural network, RNN) и свёрточная нейронная сеть (convolutional neural network, CNN) [2].

Было исследовано влияние таких параметров, как размер партии обучающих объектов (batch_size), количество эпох обучения (epochs), количество слоев и количество связей (нейронов) в этих слоях на точность обучения модели нейронной сети.

В сети прямого распространения [3] (рисунок 1) процесс передачи входных значений в нейронную сеть и получения выходных данных, которые называются прогнозируемым значением.

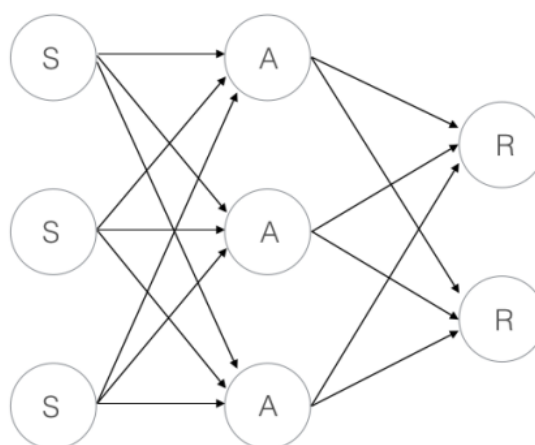


Рисунок 1. Схема FNN.

В ходе экспериментов было выявлено, что сигналы по слоям проходят с высокой скоростью и каждая эпоха обучается за малый промежуток времени, что позволяет простым увеличением числа параметров и самих эпох обучения добиться повышения точности. Но бесконечно увеличивать параметры не удастся, поскольку возникает эффект переобучения и сеть теряет точность.

Листинг 1.

```
model = tf.keras.models.Sequential([
```

```

tf.keras.layers.Flatten(input_shape = (28, 28)),
tf.keras.layers.Dense(256, activation = 'relu'),
tf.keras.layers.Dropout(0.3),
tf.keras.layers.Dense(128, activation = 'relu'),
tf.keras.layers.Dropout(0.2),
tf.keras.layers.Dense(10, activation = 'softmax')

```

)

Сеть с обратными связями — искусственная нейронная сеть, в которой выход нейрона может вновь подаётся на его вход [2]. В более общем случае это означает возможность распространения сигнала от выходов к входам. Это означает, что выход какого-нибудь нейрона определяется не только его весами и входным сигналом, но еще и предыдущими выходами (так как они снова вернулись на входы).

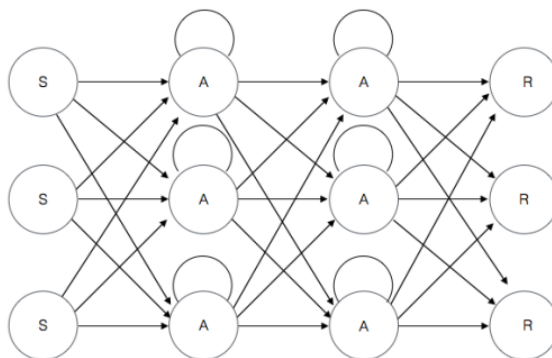


Рис.2. Схема основных подтипов RNN.

Подсети этой архитектуры показывают отличные друг от друга результаты.

В первом случае архитектура SimpleRNN подразумевает однослойную модель и даже значительное увеличение параметров дает результат не выше 98%.

Листинг 2.

```

model = tf.keras.models.Sequential([
    tf.keras.layers.SimpleRNN(512,input_shape=(28,28),activation='relu'),

```

```
tf.keras.layers.Dropout(0.1),  
tf.keras.layers.Dense(10,activation='softmax')
```

)

Архитектура с долгой краткосрочной памятью (long short-term memory, LSTM) не ограничена в количестве слоёв и показывает лучшие результаты даже при меньшем количестве параметров [3].

Листинг 3.

```
model = tf.keras.models.Sequential([  
    tf.keras.layers.LSTM(128,input_shape=(28, 28),  
activation="relu", return_sequences=True),  
    tf.keras.layers.Dropout(0.3),  
    tf.keras.layers.LSTM(128, activation='relu'),  
    tf.keras.layers.Dropout(0.2),  
    tf.keras.layers.Dense(32, activation='relu'),  
    tf.keras.layers.Dropout(0.2),  
    tf.keras.layers.Dense(10, activation='softmax')
```

)

Архитектура с управляемым рекуррентным блоком (gated recurrent units, GRU) самая современная из рассматриваемых рекуррентных моделей. Она показывает наиболее высокую точность даже при меньшем числе параметров [4].

Листинг 4.

```
model = tf.keras.models.Sequential([  
    tf.keras.layers.GRU(128, input_shape = (28, 28),  
activation="relu", return_sequences=True),  
    tf.keras.layers.Dropout(0.1),  
    tf.keras.layers.GRU(64, activation='relu'),  
    tf.keras.layers.Dense(10, activation='softmax')
```

Д)

Свёрточная нейронная сеть ориентирована на эффективное распознавание образов [5].

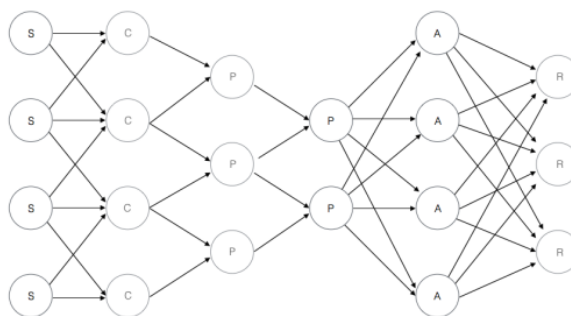


Рис.3. Схема CNN.

В архитектуре свёрточной нейронной сети выходы промежуточных слоев образуют матрицу или набор матриц. Так, например, на вход свёрточной нейронной сети можно подавать три слоя изображения (R-, G-, B-каналы изображения).

Листинг 5.

```
model = Sequential()
model.add(Conv2D(filters=32,kernel_size=(3,3),activation='relu', input_shape=input_shape))
model.add(Conv2D(64, (3,3), activation='relu'))
model.add(MaxPool2D(2,2))
model.add(Dropout(0.25))
model.add(Flatten())
model.add(Dense(32, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(10, activation='softmax'))
```

В результате проведенной работы были разработаны несколько моделей искусственных нейронных сетей, а результаты их обучения сведены в таблицу.

Model	Epochs time	Loss	Accuracy	Params
FNN	7s	0.065	0.981	235,146
SimpleRNN	135s	0.101	0.976	282,122
LSTM	~248s	0.066	0.981	216,426
GRU	148s	0.030	0.991	97,994
CNN	129s	0.035	0.990	314,090

Таблица 1. Результаты и параметры обучения моделей нейронных сетей.

В итоге наиболее эффективной оказалась архитектура GRU рекуррентной нейронной сети. При незначительно большем времени обучения каждой эпохи, сеть показывает наибольшую точность и при этом ей требуется наименьшее количество параметров.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Николенко С., Кадурин А. Архангельская Е.Н. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.
2. Хайкин С. Нейронные сети: Полный курс. – 2-е издание / пер. с англ. – М. Издательский дом «Вильямс», 2006. – 1104 с.
3. Recurrent Neural Network Tutorial, Part 4 – Implementing a GRU/LSTM RNN with Python and Theano – WildML.
4. Sepp Hochreiter; Jürgen Schmidhuber. Long short-term memory (англ.) // Neural Computation: journal. — 1997. — Vol. 9, No. 8. — Pp. 1735—1780.
5. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel: Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation.