BLOG

Demand Forecasting 3: Neural Networks

By Semantive   August 20, 2018   No Comments

This post is a part of our series exploring different options for long-term demand forecasting. To better understand our journey, you might want to check out our **introductory blog post:** **Long-Term Demand Forecasting**

Today, we will cover another popular approach to forecasting – using Recurrent Neural Networks (RNNs), in particular LSTMs (Long Short-Term Memory) networks. We believed that the dataset is too small for this kind of machine learning model, but after initial experiments we achieved quite promising results.

regression model possible – consisting of 4 lines of Keras code:

```
basic_model = Sequential()
basic_model.add(LSTM(500, input_shape=(X_train_vals.shape[1], X_train_vals.shape[2])))
basic_model.add(Dense(500))
basic_model.compile(loss='mean_absolute_error', optimizer='adam')
```

RNNs should not require any complex features for this task, so we decided to use a simple vector of sales from previous day for an input, and train the network to predict sales on the following day. We have then fed the prediction back to the network in order to predict more days – this is the same approach as we used for step by step regression models.

After seeing that this model actually works, we added more features – one-hot encoded days of the week and months of the year, in order to make sure the network takes seasonality into account. The results were quite good: we achieved a mean SMAPE of 12.40 on our validation set and 14.58 on Kaggle which was one of the top scores we had so far. This led us to explore more complex RNN architectures.

Combining LSTMs and CNNs

One of the first ideas that came to our attention was using Convolutional Recurrent Neural Networks (CNNs). We found a paper describing such architecture in enough detail to be easily reproduced. Although the authors of the paper proposed to use the network for classification of time series, not forecasting, we decided to try adapting it to our case. Adding the convolutional layers should allow the network to learn about relations between sales values, and would therefore outperform basic LSTMs and other models that predict sales of every item in every store independently.

slightly worse than the first LSTM model. It scored 14.02 on Kaggle which was an improvement – it seems to better generalize to new data and is less prone to overfitting. Having a promising network architecture, we tried tuning its hyperparameters, ending up with 14.01 on Kaggle.
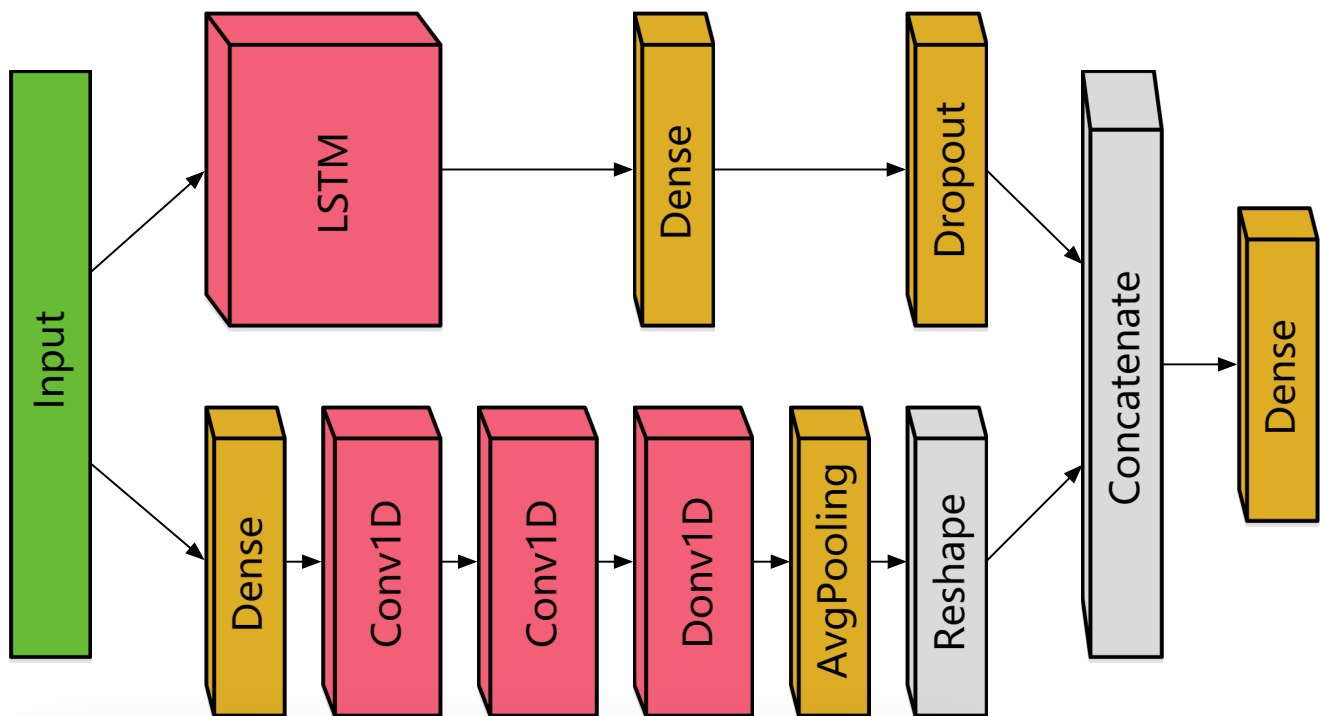
Complex model architecture (LSTM with Conv. layers)

We experimented with various other RNN architectures, but couldn't produce a better score. In particular, we tried double stacking LSTMs, using only CNNs and using them before LSTMs. Looking at what we managed to achieve with these models, it might be impossible to get any more accurate.
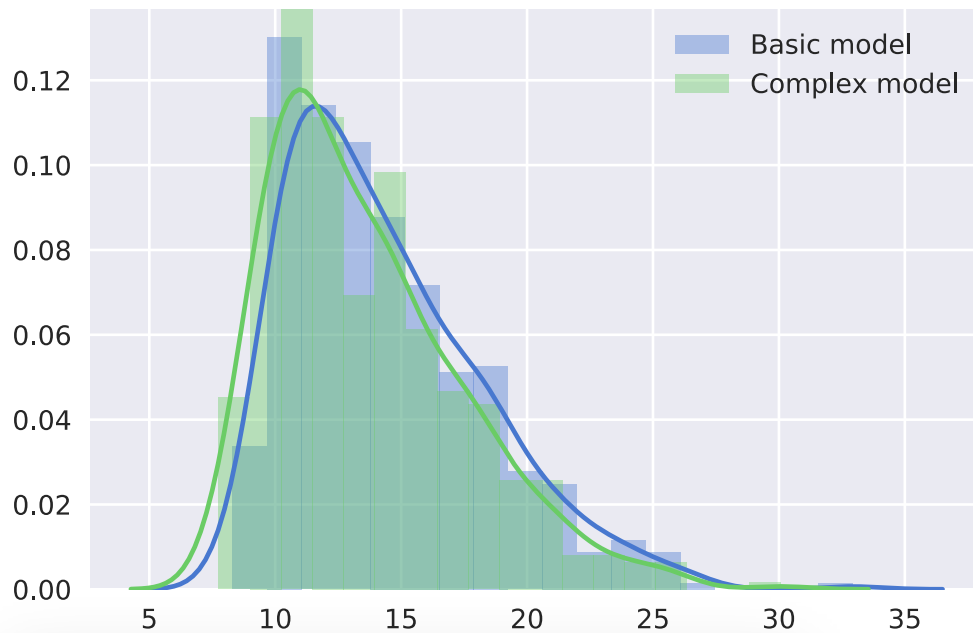
SMAPE Score Distribution on 500 time series

Code of both models is available on our Github.

Pros and cons of RNNs

Clearly, RNNs are the most powerful tools we can use for time predictions, and Keras makes them really easy to implement and train. Nonetheless, it is important to know about their limitations before actually using them in production. First of all, neural networks are not interpretable – this means that we have little to no idea about how their predictions are made. If the results need to be interpretable, it is probably best to stick with econometric models that we talked about at the very beginning of this series. For Kaggle contests, however, deep neural networks are clearly the best choice.

This is the end of our short series about forecasting demand. We invite you to still follow our blog, as there are more posts about machine learning coming soon. In the

Data Science professional services may give you the edge your business needs. Get in touch ➜

RECENT POSTS

4 modern AI solutions for manufacturing

Targi pracy IT 2019 w Warszawie. Co, gdzie, kiedy?

High-Performance computation in Python | NumPy

Text Summarization in Python

Data Science internship, and why Semantive program is worth it?

SEARCH

Search...   🔍

Previous Post

s with Databricks

Leave a Reply

My comment is..

Name *

Email *

Website

☐ Save my name, email, and website in this browser for the next time I comment.

SUBMIT COMMENT

CONTACT US

+48 510 002 513 | contact@semantive.com

ul. Nowogrodzka 42/41, 00-695 Warsaw, Poland

SEMANTIVE

Big data | AI & Data Science | Cloud

Services that make your organization data informed

Services   Workflow   Case studies   Training   Career   Blog
Contact   PL