Jheronimus Academy of Data Science

EINDHOVEN UNIVERSITY OF TECHNOLOGY

On Deep Learning

Practical Session

João Pereira

PDEng Data Science Talks

December 2019

Outline

- Remarks on previous sessions
- Introduction to TensorFlow
- Computational graph: tensors and operations
- Implementing neural networks in Python with TensorFlow:
  - Fraud Detection
  - Time Series Classification
  - Time Series Forecasting

When should we use deep learning?

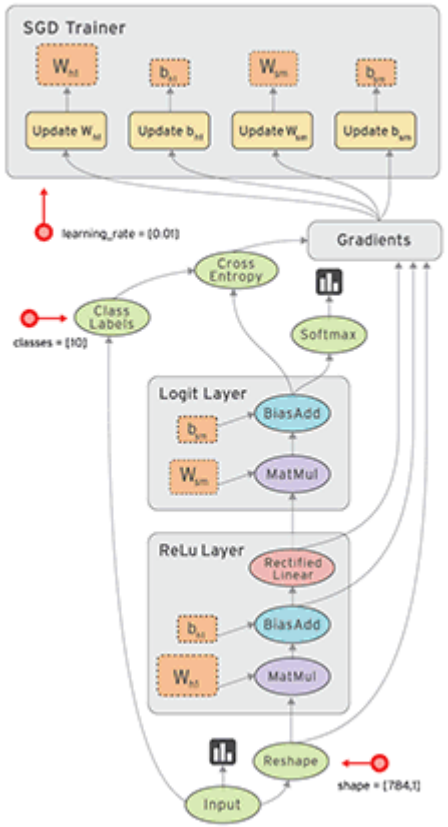
- For tackling **complex** problems
- We have **lots** of data available
- Highly **scalable** solutions
- **High-dimensional** and **non-linear** data: vision, speech, text, …
- Model **performance** is crucial

What is TensorFlow?

- Machine learning framework
- Computes gradients
- Developed by Google Brain for internal use
- Open-sourced in 2015
- Widely used in research and industry
- Allows deployment in diverse platforms (CPUs, GPUs, and TPUs)

Computational Graph

- Numeric computation is represented as a **graph**
- Nodes are **operations**
- Edges are **tensors** (multi-dimensional arrays)



Colab Environment

- Similar to Jupyter Notebooks
- Runs online in Google servers
- Available for free!
- GPU and TPU support up to 12h in row

(Tesla K80 GPU, 12GB of RAM)

Let’s get started!

Code available on Github 

github.com/jpcpereira/pdeng__tutorial_deep_learning

Fraud Detection

Using autoencoders UNSUPERVISED

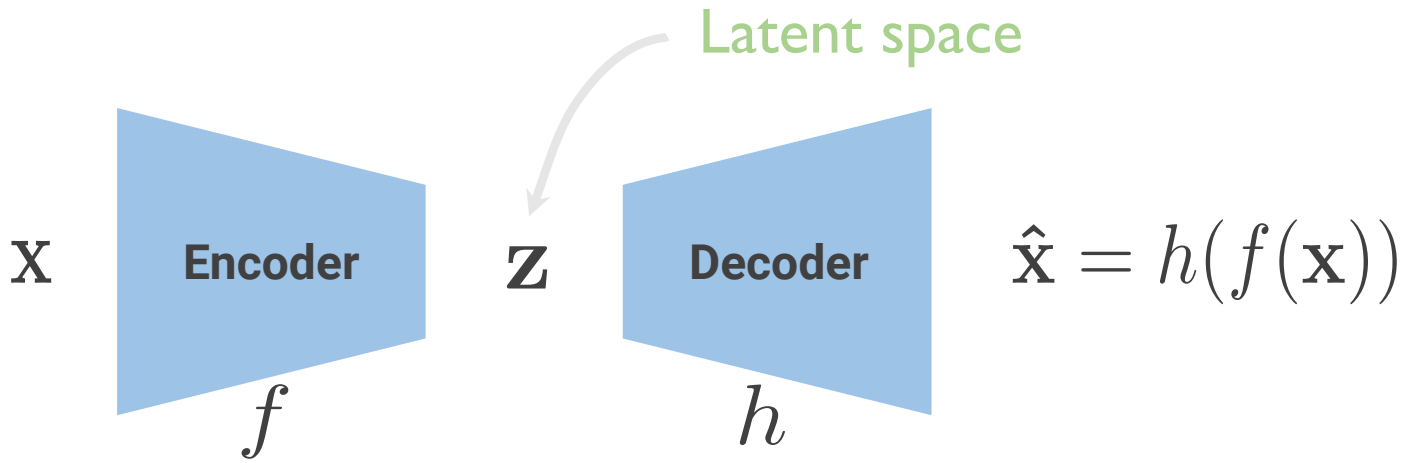
`fraud_detection.ipynb`

Fraud Detection

- Highly imbalanced problem
- Scalability is key
- Inference time has to be very short (<20ms)

Transaction #	Location	Store	…	Amount	
1	Netherlands	Albert Heijn		15.75 €	Not Fraud
…	…	…			
N	Portugal	CR7 Museum		35.00 €	Fraud

Autoencoders (Recap)



$$\mathcal{L}(\mathbf{x}; \mathbf{W}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$$

Fraud Detection Approach

- Train it on data with mostly normal transactions.
- Learns to reconstruct normal transactions.
- Measure the reconstruction error
- At test time, **fraud transactions have high reconstruction error**.

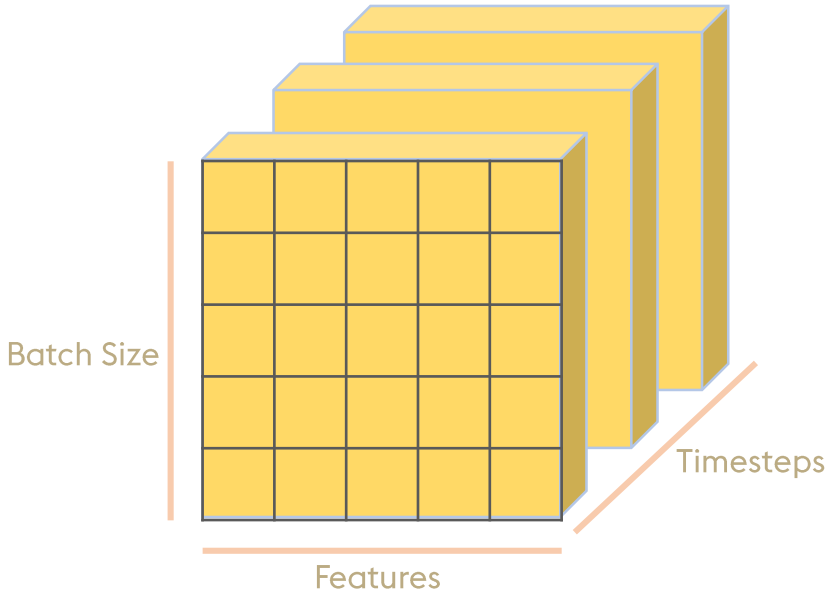
Time Series Classification

Using Recurrent Neural Networks SUPERVISED

`time_series_classification.ipynb`

Training RNNs in TensorFlow

Data Format




```
training_data is a tensor with shape:

(BATCH_SIZE, TIMESTEPS, DIMENSIONS)
```

Time Series Forecasting

Using Recurrent Neural Networks (Supervised)

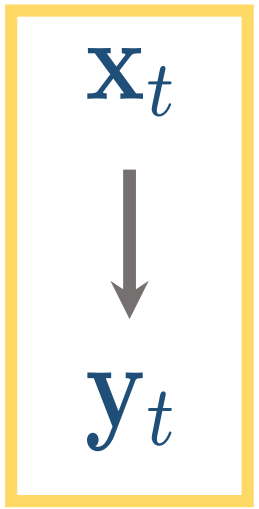
`time_series_forecasting.ipynb`

Task

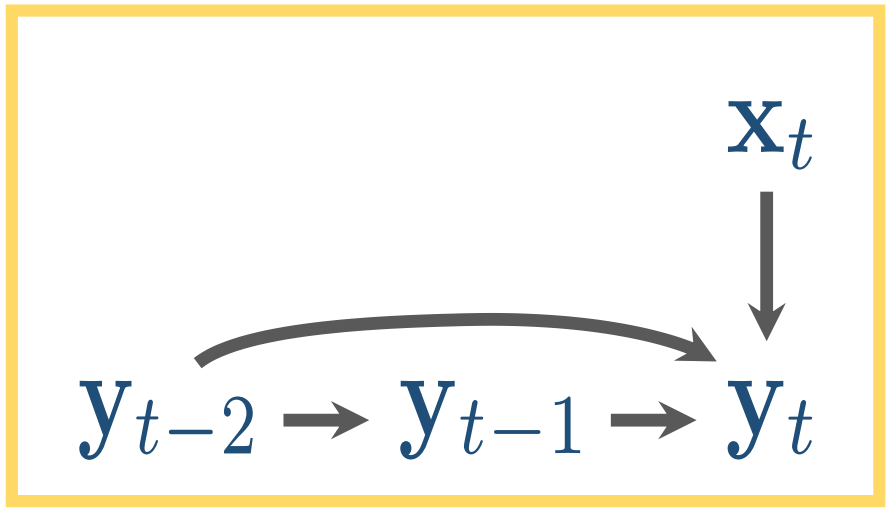
- Time series forecasting can be seen as a **regression task**
- Very often the **order** of the observations matter.
- Let's use a Recurrent Neural Network (RNN)!

Time Series Forecasting Models

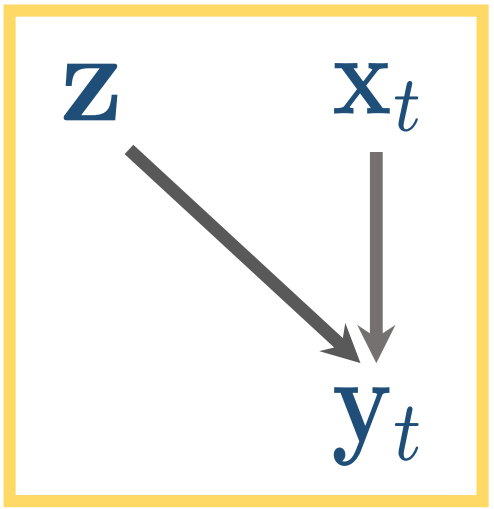
Features
(exogeneous variables)



Autoregressive



Latent Variables



TensorFlow References

TensorFlow 2.0

Tweets by Francois Chollet (inventor of Keras)

https://twitter.com/fchollet/status/1105139360226140160

Questions?

✉ mail@joao-pereira.pt

Thank you for your attention!