Operationalize deep learning: How to deploy and consume your LSTM networks for predictive maintenance scenarios

Francesca Lazzeri & Fidan Boylu Uz, Data Scientists

Microsoft, AI & Research

@frlazzeri & @FidanBoyluUz

Strata DATA CONFERENCE

Session Goals

- Understand
  - How deep learning can be used in Predictive Maintenance
- Compare
  - Traditional predictive maintenance based on feature engineering
  - Deep Learning approach
- Apply
  - Introduction to Azure Machine Learning Workbench
  - How to operationalize your Deep Learning models

What is Predictive Maintenance?

Predict the possibility of failure of an asset in the near future so that the assets can be monitored proactively to take action before the failures occur.

Aerospace	Utilities	Manufacturing	Transportation & Logistics
What is the likelihood of delay due to mechanical issues?	When is my solar panel or wind turbine going to fail next?	Will the component pass the next stage of testing on factory floor or do I need to rework?	Should I replace the break disks in my car or can I wait for another month?
When is this aircraft component likely to fail next?	Which circuit breakers in my system are likely to fail in the next month?	What is the root cause of the test failure?	What maintenance task should I perform on my elevator?

Example Scenario

- Predictive Maintenance Template

https://gallery.cortanaintelligence.com/Collection/Predictive-Maintenance-Template-3

Sample training data

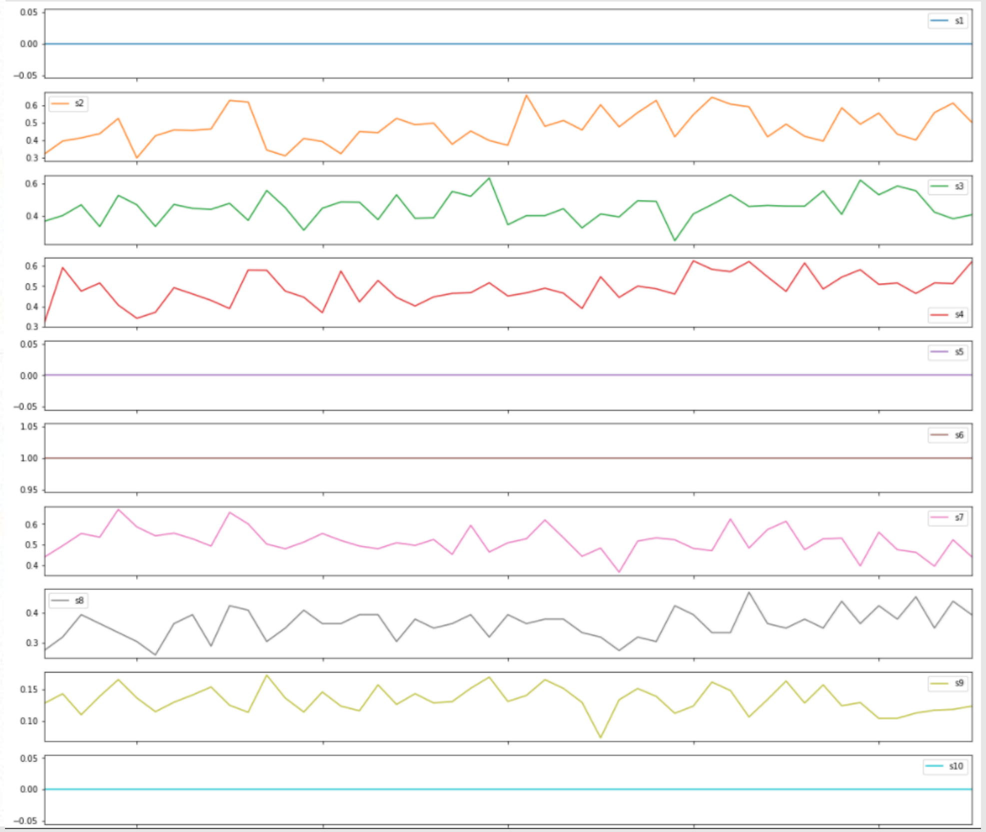
~20k rows,
100 unique engine id

id	cycle	setting1	setting2	setting3	s1	s2	s3	...	s19	s20	s21
1	1	-0.0007	-0.0004	100	518.67	641.82	1589.7		100	39.06	23.419
1	2	0.0019	-0.0003	100	518.67	642.15	1591.82		100	39	23.4236
1	3	-0.0043	0.0003	100	518.67	642.35	1587.99		100	38.95	23.3442
...	...										
1	191	0	-0.0004	100	518.67	643.34	1602.36		100	38.45	23.1295
1	192	0.0009	0	100	518.67	643.54	1601.41		100	38.48	22.9649
2	1	-0.0018	0.0006	100	518.67	641.89	1583.84		100	38.94	23.4585
2	2	0.0043	-0.0003	100	518.67	641.82	1587.05		100	39.06	23.4085
2	3	0.0018	0.0003	100	518.67	641.55	1588.32		100	39.11	23.425
...	...										
2	286	-0.001	-0.0003	100	518.67	643.44	1603.63		100	38.33	23.0169
2	287	-0.0005	0.0006	100	518.67	643.85	1608.5		100	38.43	23.0848

Sample testing data

~13k rows,
100 unique engine id

id	cycle	setting1	setting2	setting3	s1	s2	s3	...	s19	s20	s21
1	1	0.0023	0.0003	100	518.67	643.02	1585.29		100	38.86	23.3735
1	2	-0.0027	-0.0003	100	518.67	641.71	1588.45		100	39.02	23.3916
1	3	0.0003	0.0001	100	518.67	642.46	1586.94		100	39.08	23.4166
...	...										
1	30	-0.0025	0.0004	100	518.67	642.79	1585.72		100	39.09	23.4069
1	31	-0.0006	0.0004	100	518.67	642.58	1581.22		100	38.81	23.3552
2	1	-0.0009	0.0004	100	518.67	642.66	1589.3		100	39	23.3923
2	2	-0.0011	0.0002	100	518.67	642.51	1588.43		100	38.84	23.2902
2	3	0.0002	0.0003	100	518.67	642.58	1595.6		100	39.02	23.4064
...	...										
2	48	0.0011	-0.0001	100	518.67	642.64	1587.71		100	38.99	23.2918
2	49	0.0018	-0.0001	100	518.67	642.55	1586.59		100	38.81	23.2618
3	1	-0.0001	0.0001	100	518.67	642.03	1589.92		100	38.99	23.296
3	2	0.0039	-0.0003	100	518.67	642.23	1597.31		100	38.84	23.3191
3	3	0.0006	0.0003	100	518.67	642.98	1586.77		100	38.69	23.3774
...	...										
3	125	0.0014	0.0002	100	518.67	643.24	1588.64		100	38.56	23.227
3	126	-0.0016	0.0004	100	518.67	642.88	1589.75		100	38.93	23.274



- More realistic scenario

https://gallery.cortanaintelligence.com/Collection/Predictive-Maintenance-Modelling-Guide-1

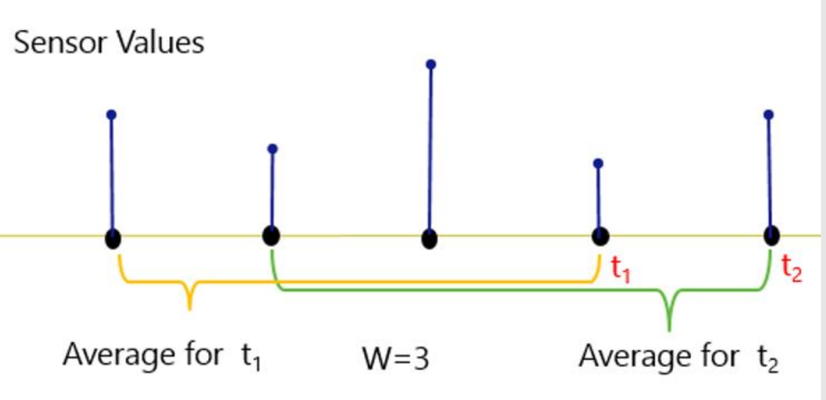
Traditional Predictive Maintenance

Feature Engineering

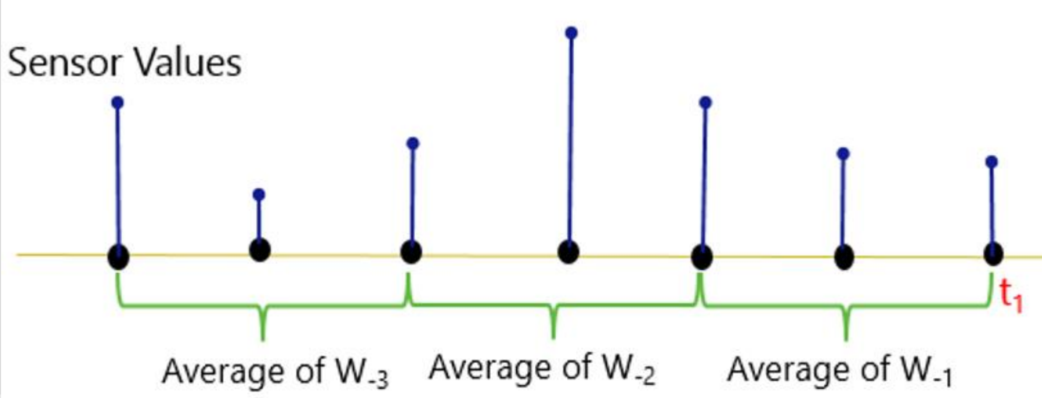
Time Series Data:

Operation conditions of a machine, e.g. data collected from sensors.

Rolling Aggregates



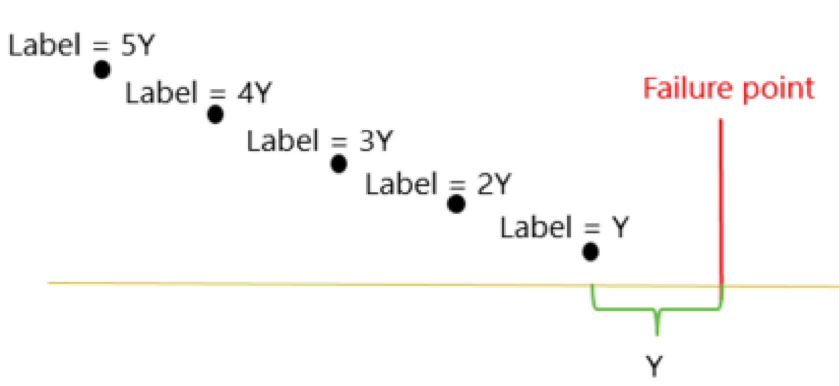
Tumbling Aggregates



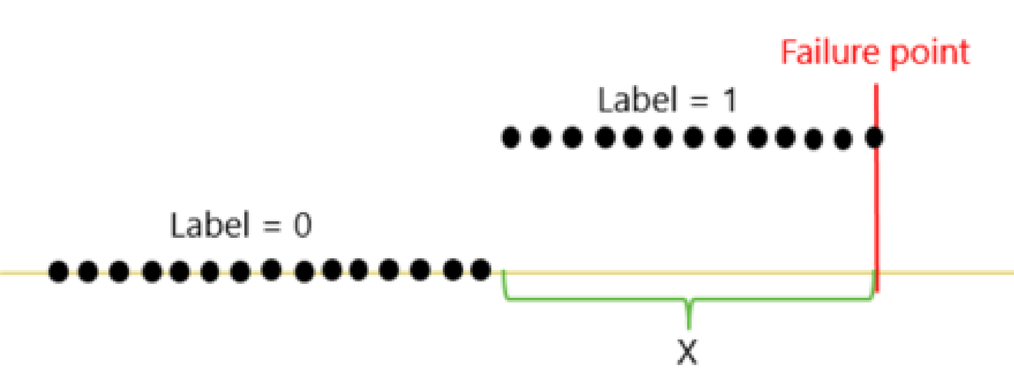
Modelling

- Regression
- Classification
  - Binary
  - Multi-Class

Regression



Binary Classification



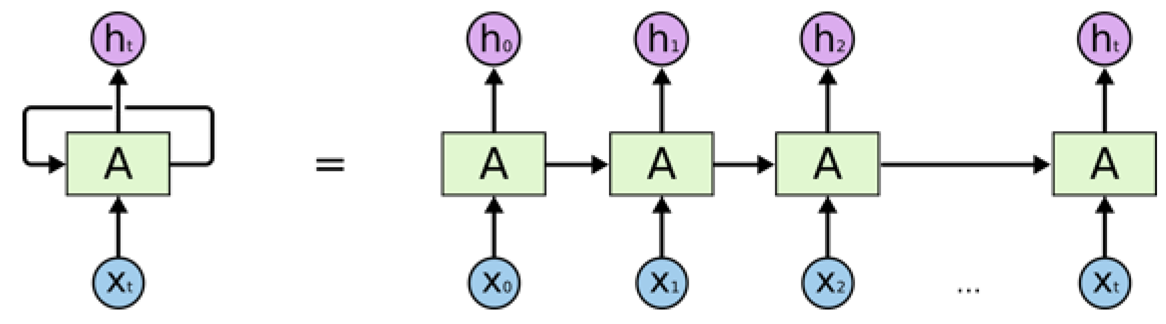
- https://azure.microsoft.com/en-us/documentation/articles/cortana-analytics-playbook-predictive-maintenance/

Deep Learning for Predictive Maintenance

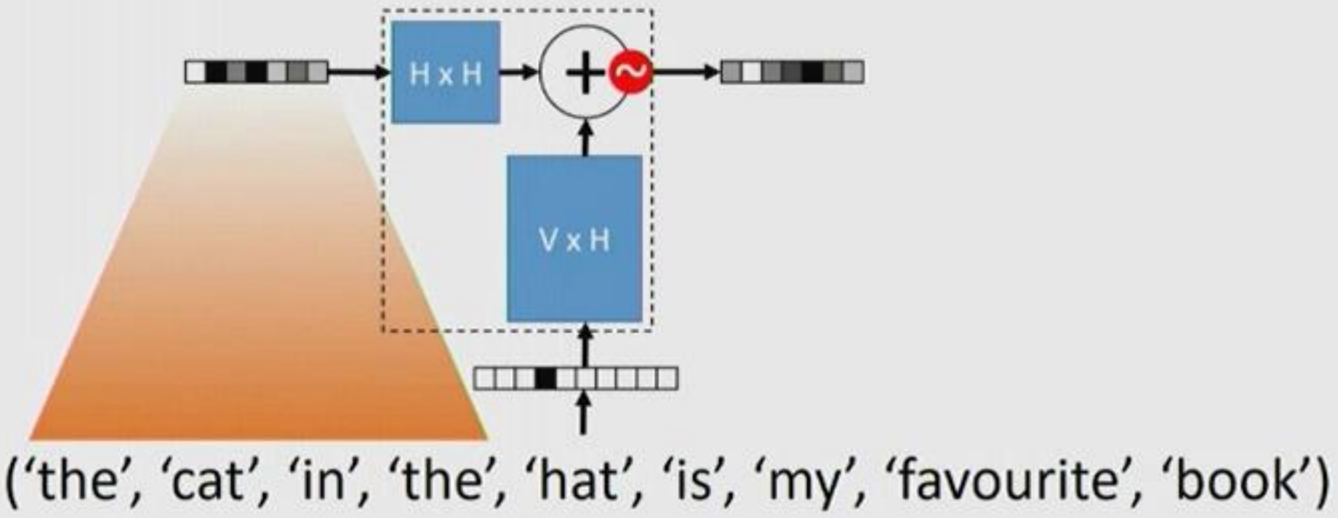
- Problems with traditional approach:
  - Manual construction of features
  - Look back period/window – How long?
  - What type of aggregation? Std., min., max., avg., etc.
  - Which ones work better?
  - Aggregation over long time periods – information loss
  - Hard to reuse the model since it won't apply to a different data set.

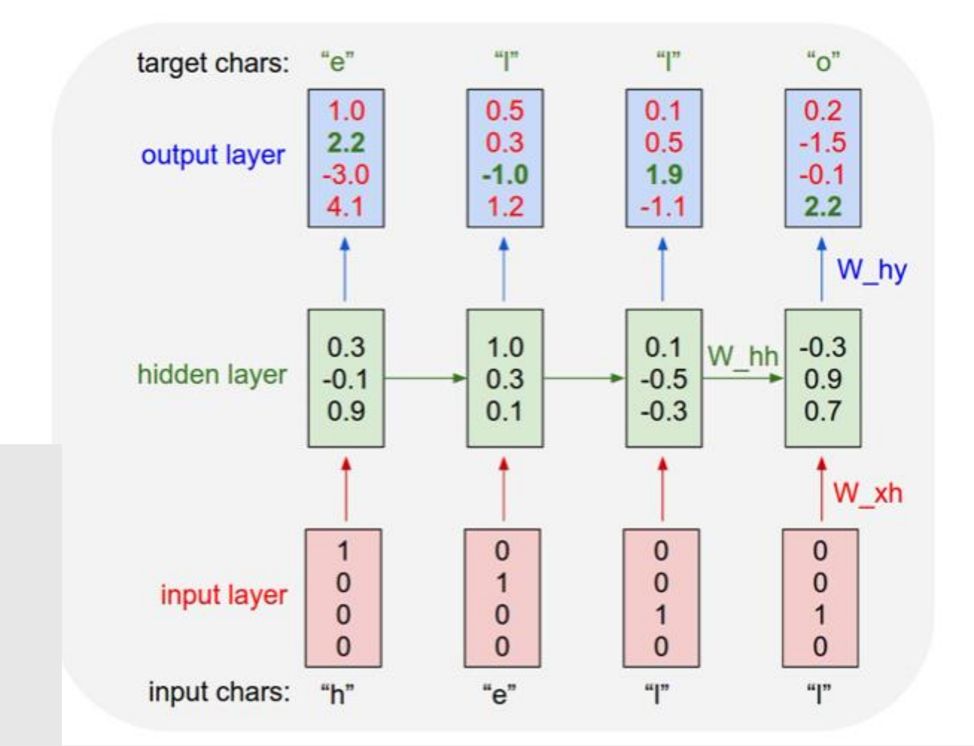
- Can we abstract the window in an automatic way?
  - Deep learning to extract the right features
  - Recurrent Neural Nets
    - Long Short Term Memory (LSTM) Networks

LSTMs



An unrolled recurrent neural network.





- http://karpathy.github.io/2015/05/21/rnn-effectiveness/
- http://colah.github.io/posts/2015-08-Understanding-LSTMs/
- https://resnet.microsoft.com/video/38264

LSTMs for Predictive Maintenance

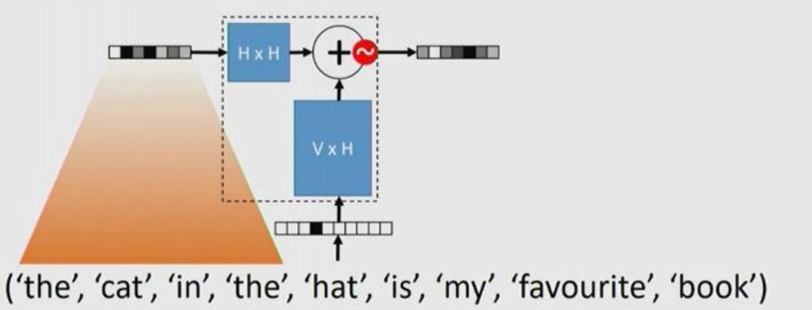
- Deep Learning for Predictive Maintenance Tutorial
  - https://github.com/Azure/lstms_for_predictive_maintenance
  - Uses same dataset with Predictive Maintenance Template
  - Same data preparation and labeling steps except the feature engineering
  - Uses label1 for binary classification

id	cycle	setting1	setting2	setting3	s1	s2	s3	s4	s5	...	s16	s17	s18	s19	s20	s21	RUL	label1
1	1	0.459770	0.166667	0.0	0.0	0.183735	0.406802	0.309757	0.0	...	0.0	0.333333	0.0	0.0	0.713178	0.724662	191	0
1	2	0.609195	0.250000	0.0	0.0	0.283133	0.453019	0.352633	0.0	...	0.0	0.333333	0.0	0.0	0.666667	0.731014	190	0
1	3	0.252874	0.750000	0.0	0.0	0.343373	0.369523	0.370527	0.0	...	0.0	0.166667	0.0	0.0	0.627907	0.621375	189	0
1	4	0.540230	0.500000	0.0	0.0	0.343373	0.256159	0.331195	0.0	...	0.0	0.333333	0.0	0.0	0.573643	0.662386	188	0
1	5	0.390805	0.333333	0.0	0.0	0.349398	0.257467	0.404625	0.0	...	0.0	0.416667	0.0	0.0	0.589147	0.704502	187	0
1	189	0.465517	0.666667	0.0	0.0	0.894578	0.547853	0.772451	0.0	...	0.0	0.583333	0.0	0.0	0.263566	0.301712	3	1
1	190	0.344828	0.583333	0.0	0.0	0.731928	0.614345	0.737677	0.0	...	0.0	0.833333	0.0	0.0	0.271318	0.239299	2	1
1	191	0.500000	0.166667	0.0	0.0	0.641566	0.682799	0.734639	0.0	...	0.0	0.500000	0.0	0.0	0.240310	0.324910	1	1
1	192	0.551724	0.500000	0.0	0.0	0.701807	0.662089	0.758778	0.0	...	0.0	0.666667	0.0	0.0	0.263566	0.097625	0	1

LSTMs for Predictive Maintenance

- LSTMs in time series
  - Pick a sequence length (similar to picking window size for feature engineering) (e.g. sequence_length = 50)
  - Prepare sequences for Keras:
    - 3D tensor (samples, time steps, features)
      - samples: the number of training sequences
      - time steps: the look back window or sequence length
      - features is the number of features of each sequence at each time step
  - For each engine id, for each cycle (time step), use 25 features for the last 50 cycles.
    -

```
seq_array.shape
(15631, 50, 25)
```



LSTMs for Predictive Maintenance

- Keras with CNTK backend.

```
nb_features = seq_array.shape[2]
nb_out = label_array.shape[1]

model = Sequential()

model.add(LSTM(
         input_shape=(sequence_length, nb_features),
         units=100,
         return_sequences=True))
model.add(Dropout(0.2))

model.add(LSTM(
          units=50,
          return_sequences=False))
model.add(Dropout(0.2))

model.add(Dense(units=nb_out, activation='sigmoid'))
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
```

LSTMs for Predictive Maintenance

- Fit the network

```
%%time
# fit the network
model.fit(seq_array, label_array, epochs=10, batch_size=200, validation_split=0.05, verbose=1,
          callbacks = [keras.callbacks.EarlyStopping(monitor='val_loss', min_delta=0, patience=0, verbose=0, mode='auto')])
```

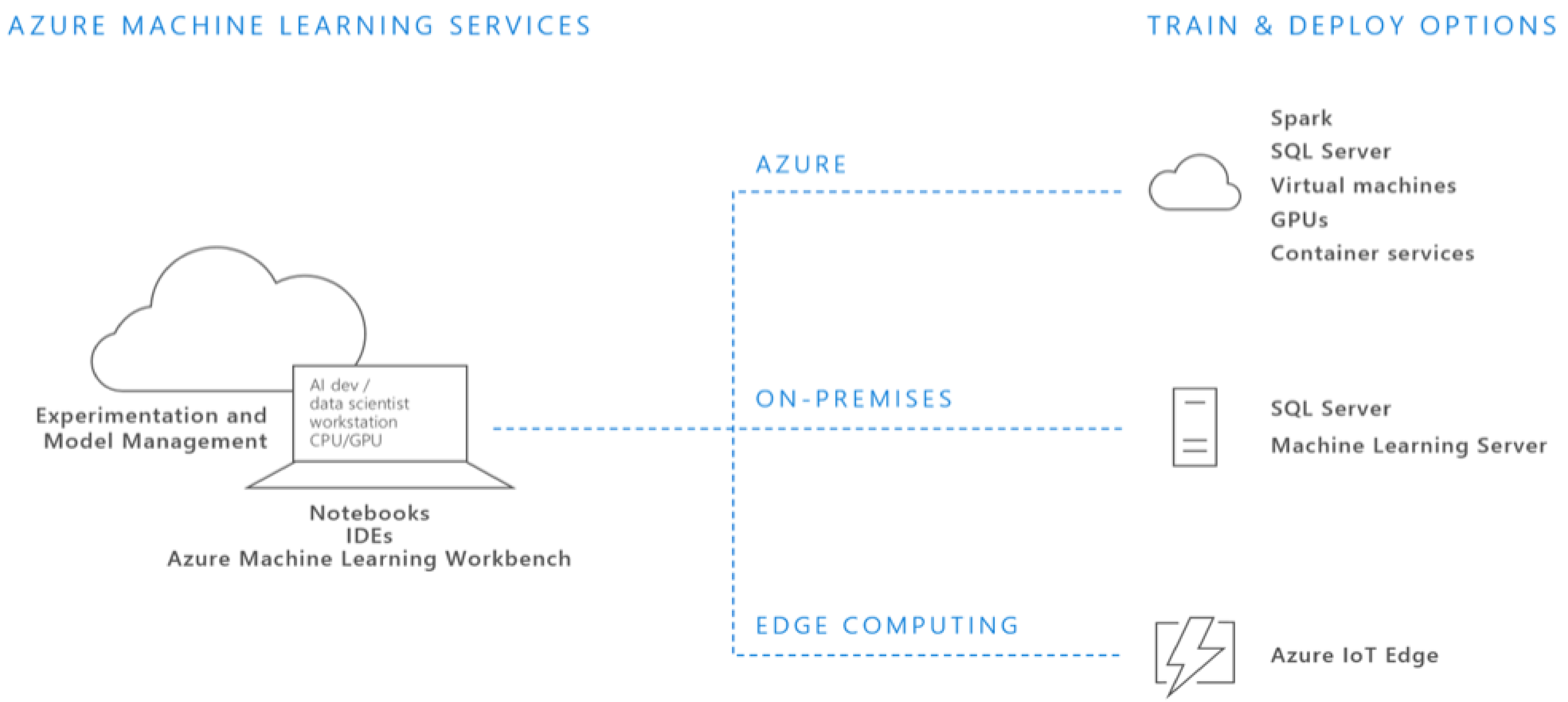
```
Train on 14849 samples, validate on 782 samples
Epoch 1/10
14849/14849 [==============================] - 7s - loss: 0.2733 - acc: 0.8819 - val_loss: 0.1307 - val_acc: 0.9412
Epoch 2/10
14849/14849 [==============================] - 6s - loss: 0.1187 - acc: 0.9525 - val_loss: 0.0906 - val_acc: 0.9578
Epoch 3/10
14849/14849 [==============================] - 7s - loss: 0.0849 - acc: 0.9648 - val_loss: 0.0750 - val_acc: 0.9629
Epoch 4/10
14849/14849 [==============================] - 6s - loss: 0.0786 - acc: 0.9675 - val_loss: 0.0538 - val_acc: 0.9872
Epoch 5/10
14849/14849 [==============================] - 6s - loss: 0.0632 - acc: 0.9748 - val_loss: 0.0626 - val_acc: 0.9744
CPU times: user 25.8 s, sys: 9.09 s, total: 34.9 s
Wall time: 35.7 s
```

- Compare test set performance between traditional and LSTM

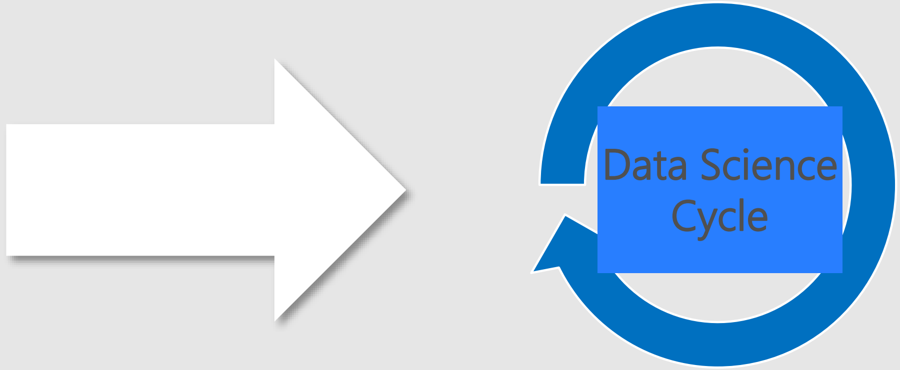
	Accuracy	Precision	Recall	F1-score
LSTM	0.978495	0.960000	0.96	0.960000
Template Best Model	0.940000	0.952381	0.80	0.869565

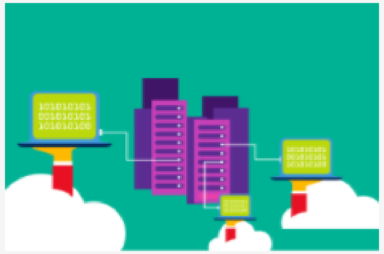
How can you operationalize your LSTM networks?

AZURE MACHINE LEARNING



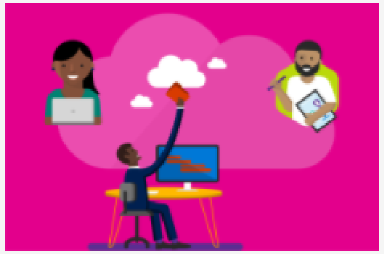
Azure Machine Learning Workbench





Distributed Tuning of Hyperparameters
Scale out tuning of hyperparameters using Docker container and Spark cluster

Classify US Incomes - TDSP Example
Predict annual income of individuals, following Team Data Science Process



Team Data Science Process Template
An agile, iterative, data science methodology to improve team collaboration and learning.

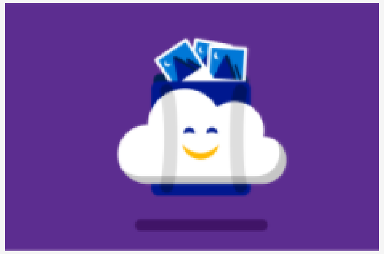
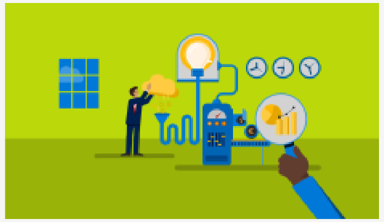
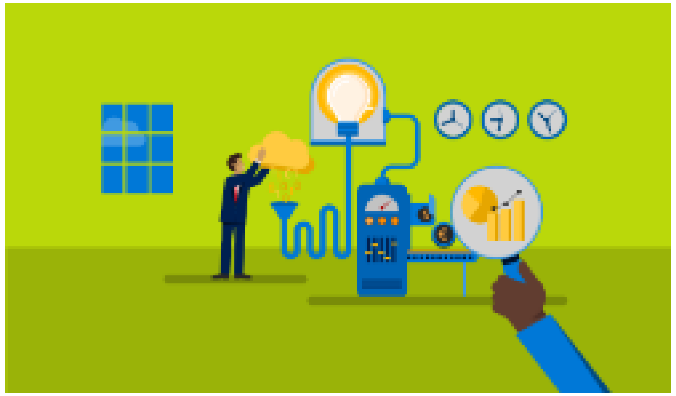


Image Classification with CNTK
Train, evaluate, and deploy your own image classification model using CNTK.



Deep Learning for Predictive Maintenance
Build a predictive maintenance modeling using simulated aircraft engine run-to-failure events.

Deep Learning for Predictive Maintenance



Deep Learning for Predictive Maintenance Scenario

Filter

› Use IDE extensions
Use Azure IoT Edge AI Toolkit
› Configure compute environment
› Acquire and understand data
› Develop models
› Operationalize models
⌄ Real-world examples
Document collection analysis
Q & A matching
Predictive maintenance
Aerial image classification
Server workload forecasting on terabytes of data
Energy demand time series forecasting
Distributed tuning of hyperparameters
Customer churn prediction
Sentiment analysis with deep learning
Biomedical entity recognition - TDSP project
Classify US incomes - TDSP project
Image classification using CNTK
Deep Learning for Predictive Maintenance

Deep learning for predictive maintenance real-world scenarios

11/22/2017 • 6 minutes to read • Contributors

Deep learning is one of the most popular trends in machine learning, with applications to many areas including driverless cars, speech and image recognition, robotics and finance. Also referred to as Artificial Neural Networks (ANN), these methods are inspired by the individual neurons within the brain (biological neural networks).

Predictive maintenance uses machine learning methods to determine the condition of equipment in order to preemptively perform maintenance and avoid adverse machine performance. In these scenarios, data is collected over time to monitor the state of the machine with the final goal of finding patterns to predict failures. Long Short Term Memory (LSTM) networks are especially appealing for predictive maintenance for the ability to learning from sequences of data. LSTMs are designed for application to time series data to detect temporal patterns that could lead to machine failures.

Use case overview

This tutorial uses the example of simulated aircraft engine run-to-failure events to demonstrate the predictive maintenance modeling process. The scenario is described at Predictive Maintenance

The implicit assumption of the scenario is the engine has progressive degradation pattern. The pattern signal is reflected in sensor measurements and a machine learning algorithm can learn the relationship between the changes in these sensor values and the historical failures. The model can then Predict engine failures in the future based on the current state of sensor measurements.

This scenario creates an LSTM network for the data to predict remaining useful life of aircraft engines using historical aircraft sensor values. This scenario uses the Keras with Tensorflow deep learning framework as a back end to train and test the LSTM network.

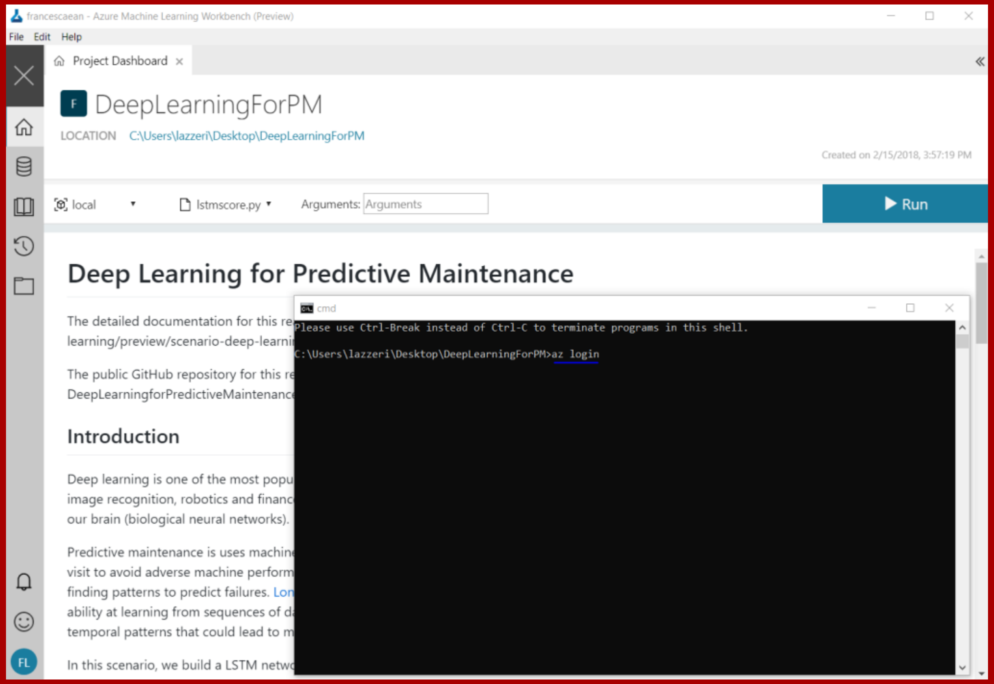
A public GitHub repository for this scenario is located athttps://github.com/Azure/MachineLearningSamples-DeepLearningforPredictiveMaintenance for issue reports and contributions.

Prerequisites

- An Azure account (free trials are available).
- Azure Machine Learning Workbench, with a workspace created.

- https://docs.microsoft.com/en-us/azure/machine-learning/preview/scenario-deep-learning-for-predictive-maintenance
- https://github.com/Azure/MachineLearningSamples-DeepLearningforPredictiveMaintenance

DEMO - Azure ML Workbench and Operationalization



Conclusion

- How deep learning can be used in Predictive Maintenance
- Traditional predictive maintenance based on feature engineering
- Working with this notebook, we have completed:
  - Operationalization asset generation and model deployment in the Code/3_operationalization.ipynb notebook.
- This scenario is intended to help guide you through the predictive maintenance model development process with your own data:
  https://github.com/Azure/MachineLearningSamples-DeepLearningforPredictiveMaintenance

Thank you!

Francesca Lazzeri & Fidan Boylu Uz, Data Scientists

Microsoft, AI & Research

@frlazzeri & @FidanBoyluUz

Strata

DATA CONFERENCE