Machine Learning in Business Process Monitoring – A Comparison of Deep Learning and Classical Approaches used for Outcome Prediction

Wolfgang Kratsch, Jonas Manderscheid, Maximilian Röglinger, Johannes Seyfried

Business & Information Systems Engineering (2020)

Appendix (available online via http://link.springer.com)

Appendix A: Overview over Accuracy and F-Score Measures per Log and per Prediction Point

Table A1: Accuracy LSTM

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std.
BPIC11	0.7789	0.7741	0.7597	0.7652	0.7585	0.7325	0.7352	0.7103	0.7321	0.7512	0.7498	0.0206
BPIC13	0.8118	0.8251	0.8864	0.9139	0.9583	0.9641	0.9603	0.9839	0.9817	0.9840	0.9269	0.0620
RTFM	0.9963	0.9934	0.9937	0.9971	0.9881	0.8854					0.9757	0.0405
PL	0.9091	0.9000	0.9250	0.9918	0.8396	0.8065	0.9682	0.9905	0.7561	0.9949	0.9082	0.0796
RL	0.5068	0.5068	0.6143	0.6664	0.7372	0.7644	0.6849	0.6831	0.7396	0.7788	0.6682	0.0932
									Aggregated		**0.8144**	**0.0495**

Table A2: Accuracy DNN

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std.
BPIC11	0.7761	0.7881	0.8368	0.7680	0.7483	0.7478	0.7183	0.7000	0.7263	0.7288	0.7539	0.0400
BPIC13	0.7939	0.8004	0.8434	0.8652	0.8766	0.8873	0.9166	0.9118	0.9206	0.9451	0.8761	0.0511
RTFM	0.9963	0.9934	0.9932	0.9964	0.9739	0.8508					0.9673	0.0577
PL	0.7855	0.7943	0.8500	0.7918	0.7646	0.7652	0.9636	0.7381	0.8976	0.7795	0.8130	0.0700
RL	0.5068	0.5068	0.62708	0.66092	0.69968	0.68764	0.68924	0.69272	0.68704	0.68712	0.6445	0.0755
									Aggregated		**0.8011**	**0.0589**

Table A3: Accuracy RF

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std.
BPIC11	0.6506	0.7926	0.6285	0.5959	0.6030	0.6132	0.7135	0.5637	0.6735	0.6381	0.6473	0.0627
BPIC13	0.4503	0.4836	0.6953	0.7037	0.7832	0.8574	0.8419	0.8305	0.8191	0.8339	0.7299	0.1419
RTFM	0.9963	0.9934	0.9932	0.9939	0.9634	0.8278					0.9613	0.0608
PL	0.7763	0.7500	0.7610	0.7538	0.7632	0.7446	0.7069	0.6886	0.7267	0.7161	0.7387	0.0266
RL	0.5000	0.5000	0.5674	0.6408	0.6648	0.6639	0.6629	0.6664	0.6474	0.6273	0.6141	0.0635
									Aggregated		**0.7383**	**0.0711**

Table A4: Accuracy SVM

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std.
BPIC11	0.7647	0.7519	0.7401	0.7279	0.7155	0.7055	0.6941	0.6866	0.6783	0.6724	0.7137	0.0318
BPIC13	0.7944	0.7980	0.8544	0.8323	0.8124	0.7888	0.7465	0.7091	0.6866	0.7027	0.7725	0.0580
RTFM	0.8964	0.9104	0.9099	0.9609	0.9533	0.7587					0.8983	0.0731
PL	0.7763	0.7689	0.7610	0.7538	0.7474	0.7446	0.7356	0.7246	0.7205	0.7097	0.7442	0.0217
RL	0.5068	0.5068	0.6145	0.6482	0.6537	0.6469	0.6435	0.6435	0.6435	0.6348	0.6142	0.0576
									Aggregated		**0.7486**	**0.0484**

Table A5: F-Score LSTM

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std. Dev.
BPIC11	0.8722	0.8677	0.8558	0.8571	0.8542	0.8347	0.8398	0.8187	0.8313	0.8381	0.8470	0.0162
BPIC13	0.7456	0.7591	0.8832	0.9171	0.9628	0.9698	0.9675	0.9871	0.9860	0.9879	0.9166	0.0881
RTFM	0.9981	0.9967	0.9968	0.9986	0.9939	0.9352					0.9866	0.0230
PL	0.9435	0.9391	0.9528	0.9947	0.8950	0.8765	0.9783	0.9934	0.8401	0.9963	0.9410	0.0517
RL	0.0000	0.0000	0.5375	0.6225	0.7269	0.7605	0.6763	0.6612	0.7387	0.7685	0.5492	0.2824
										Aggregated	**0.8481**	**0.0923**

Table A6: F-Score DNN

Prefix/ Log	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Mean*	*Std. Dev.*
BPIC11	0.8707	0.8747	0.8978	0.8564	0.8452	0.8422	0.8210	0.7917	0.8214	0.8192	0.8440	0.0318
BPIC13	0.7069	0.7172	0.8414	0.8737	0.8901	0.9024	0.9315	0.9290	0.9385	0.9578	0.8688	0.0894
RTFM	0.9981	0.9967	0.9966	0.9982	0.9864	0.9171					0.9822	0.0322
PL	0.8773	0.8775	0.9053	0.8735	0.8497	0.8419	0.9759	0.8070	0.9341	0.8537	0.8796	0.0485
RL	0.0000	0.0000	0.5828	0.6393	0.6967	0.6589	0.6679	0.6836	0.6810	0.6589	0.5269	0.2795
										Aggregated	**0.8203**	**0.0963**

Table A7: F-Score RF

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std. Dev.
BPIC11	0.7120	0.8749	0.7233	0.6914	0.6950	0.7125	0.8206	0.6420	0.7909	0.7395	0.7403	0.0656
BPIC13	0.4318	0.4596	0.7391	0.7576	0.8293	0.8884	0.8809	0.8755	0.8712	0.8825	0.7616	0.1658
RTFM	0.9982	0.9967	0.9966	0.9970	0.9814	0.9058					0.9792	0.0334
PL	0.8581	0.8446	0.8547	0.8491	0.8480	0.8437	0.7975	0.7847	0.8111	0.8039	0.8295	0.0258
RL	0.5581	0.5581	0.5673	0.6170	0.6452	0.6553	0.6485	0.6534	0.6317	0.5883	0.6123	0.0385
										Aggregated	**0.7846**	**0.0658**

Table A8: F-Score SVM

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std. Dev.
BPIC11	0.8667	0.8583	0.8506	0.8425	0.8342	0.8273	0.8194	0.8141	0.8083	0.8041	0.8326	0.0216
BPIC13	0.7254	0.7294	0.8596	0.8532	0.8512	0.8443	0.8259	0.8100	0.8040	0.8154	0.8118	0.0485
RTFM	0.9380	0.9490	0.9486	0.9795	0.9755	0.8569					0.9412	0.0444
PL	0.8740	0.8693	0.8642	0.8596	0.8553	0.8535	0.8476	0.8401	0.8374	0.8300	0.8531	0.0143
RL	0.0000	0.0000	0.5771	0.6412	0.6447	0.6388	0.6343	0.6343	0.6343	0.6209	0.5026	0.2656
										Aggregated	**0.7883**	**0.0789**

Table A9: ROC AUC LSTM

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std. Dev.
BPIC11	0.7078	0.7187	0.6951	0.7741	0.7668	0.7488	0.7301	0.7298	0.7166	0.7051	0.7293	0.0264
BPIC13	0.7920	0.8030	0.8876	0.9170	0.9598	0.9637	0.9599	0.9836	0.9785	0.9812	0.9226	0.0688
RTFM	0.5000	0.5072	0.5404	0.7904	0.8386	0.6696					0.6410	0.1355
PL	0.8367	0.7851	0.8579	0.9833	0.7681	0.6939	0.9599	0.9833	0.6491	0.9964	0.8514	0.1204
RL	0.5000	0.5000	0.6122	0.6650	0.7368	0.7643	0.6847	0.6825	0.7397	0.7784	0.6664	0.0956
										Aggregated	**0.7324**	**0.0892**

Table A10: ROC AUC DNN

Prefix/ Log	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Mean*	*Std. Dev.*
BPIC11	0.7761	0.7881	0.8368	0.7680	0.7483	0.7478	0.7183	0.7000	0.7263	0.7288	0.7539	0.0400
BPIC13	0.7655	0.7729	0.8443	0.8656	0.8767	0.8897	0.9143	0.9098	0.9154	0.9430	0.8697	0.0600
RTFM	0.5000	0.5000	0.5000	0.7375	0.8327	0.5725					0.6071	0.1439
PL	0.5324	0.6048	0.7508	0.6313	0.6403	0.6902	0.9438	0.7167	0.8177	0.6891	0.7017	0.1163
RL	0.5000	0.5000	0.6260	0.6602	0.6998	0.6867	0.6886	0.6926	0.6871	0.6863	0.6427	0.0782
										Aggregated	**0.7150**	**0.0877**

Table A11: ROC AUC RF

Prefix/ Log	1	2	3	4	5	6	7	8	9	10	Mean	Std. Dev.
BPIC11	0.7588	0.6805	0.6004	0.5591	0.5805	0.6177	0.5876	0.5737	0.5820	0.5879	0.6128	0.0581
BPIC13	0.4585	0.4875	0.7336	0.7792	0.8440	0.9335	0.8931	0.8950	0.8760	0.9071	0.7807	0.1644
RTFM	0.6779	0.7312	0.7438	0.9177	0.9522	0.6535					0.7794	0.1145
PL	0.6955	0.6821	0.6695	0.7293	0.6919	0.7025	0.6989	0.6967	0.7272	0.7455	0.7039	0.0221
RL	0.5057	0.5059	0.5912	0.6948	0.7194	0.7273	0.7259	0.7262	0.7165	0.6865	0.6599	0.0861
										Aggregated	**0.7074**	**0.0890**

Table A12: ROC AUC SVM

Prefix/ Log	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Mean*	*Std. Dev.*
BPIC11	0.7450	0.6452	0.6004	0.5591	0.5805	0.6177	0.5876	0.5737	0.5820	0.5879	0.6079	0.0511
BPIC13	0.7817	0.7686	0.8699	0.8672	0.8835	0.8870	0.8850	0.8624	0.8386	0.8380	0.8482	0.0423
RTFM	0.5000	0.5232	0.5254	0.7574	0.8114	0.4835					0.6001	0.1446
PL	0.5276	0.4736	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5001	0.0127
RL	0.5000	0.5000	0.6149	0.6585	0.6482	0.6382	0.6335	0.6335	0.6335	0.6240	0.6084	0.0584
										Aggregated	**0.6329**	**0.0618**

Appendix B: Illustrative Sequence Encoding

Initial Log:

```
<trace>
    <int key="TA0" val="6"/>
    <float key="TA1" val="3,12"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:ressource" val="res_c"/>
        <int key="EA1" val="4"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <int key="EA2" val="7"/>
        <int key="EA3" val="13"/>
    </event>
</trace>
<trace>
    <int key="TA0" val="8"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:ressource" val="res_a"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <int key="EA2" val="17"/>
        <int key="EA3" val="5"/>
    </event>
    <event>
        <string key="concept:name" value="ev_c"/>
        <int key="EA4" val="4"/>
        <int key="EA5" val="16"/>
    </event>
</trace>
<trace>
    <int key="TA1" val="14"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:ressource" val="res_a"/>
        <int key="EA1" val="1"/>
    </event>
</trace>
```

Log after Step 1:
- Labels are added
- Float-attribute is transformed to an int-attribute

```
<trace>
    <int key="TA0" val="6"/>
    <int key="TA1" val="312"/>
    <int key="label" val="0"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_c"/>
        <int key="EA1" val="4"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <int key="EA2" val="7"/>
        <int key="EA3" val="13"/>
    </event>
</trace>
<trace>
    <int key="TA0" val="8"/>
    <int key="label" val="1"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_a"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <int key="EA2" val="17"/>
        <int key="EA3" val="5"/>
    </event>
    <event>
        <string key="concept:name" value="ev_c"/>
        <int key="EA4" val="4"/>
        <int key="EA5" val="16"/>
    </event>
</trace>
<trace>
    <int key="TA1" val="14"/>
    <int key="label" val="0"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_a"/>
        <int key="EA1" val="1"/>
    </event>
</trace>
```

Log after Step 2:
- Trace 3 is removed due to too few events
- Event 3 is removed from trace 2 due to too many events

```
<trace>
    <int key="TA0" val="6"/>
    <int key="TA1" val="312"/>
    <int key="label" val="0"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_c"/>
        <int key="EA1" val="4"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <int key="EA2" val="7"/>
        <int key="EA3" val="13"/>
    </event>
</trace>
<trace>
    <int key="TA0" val="8"/>
    <int key="label" val="1"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_a"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <int key="EA2" val="17"/>
        <int key="EA3" val="5"/>
    </event>
</trace>
```

Log after Step 3:
- Trace-attribute „TA1“ is added to trace 2 with a replacement value
- Event-attribute „EA1“ is added to event 2 in trace 2 with a replacement value

```
<trace>
    <int key="TA0" val="6"/>
    <int key="TA1" val="312"/>
    <int key="label" val="0"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_c"/>
        <int key="EA1" val="4"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <string key="concept:resource" val="res_x"/>
        <int key="EA2" val="7"/>
        <int key="EA3" val="13"/>
    </event>
</trace>
<trace>
    <int key="TA0" val="8"/>
    <int key="TA1" val="4"/>
    <int key="label" val="1"/>
    <event>
        <string key="concept:name" value="ev_a"/>
        <string key="concept:resource" val="res_a"/>
        <int key="EA1" val="16"/>
    </event>
    <event>
        <string key="concept:name" value="ev_b"/>
        <string key="concept:resource" val="res_x"/>
        <int key="EA2" val="17"/>
        <int key="EA3" val="5"/>
    </event>
</trace>
```

Log after step 4 (before applying One-Hot-Encoding):
- Trace-attributes are added to each event.
- Log transformed from an XML-format to a tabular form

TA0	TA1	concept:name	concept:resource	EA1
6	312	ev_a	res_c	4
8	4	ev_a	res_a	16

Event 1

concept:name	concept:resource	EA2	EA3	label
ev_b	res_x	7	13	0
ev_b	res_x	17	5	1

Event 2

Figure B1: Illustrative sequence encoding for an example process log for prediction point two

Appendix C: Hyper-parameterization of the used classifiers

Table C1: Hyper-parameterization of the used classifiers

	Parameter	*Description*	*Value*
DNN	Learning Rate	Size of the step taken at each iteration	$10^{-[1,2,3,4,5,6]}$
	Maximum number of epochs	Maximum number of times the training data is processed[1]	50
	Early Stopping: $min\ \delta$	Minimum change in the loss value to qualify as an improvement	10^{-6}
	Early Stopping: patience	Number of epochs with no improvements after which the training process will stop	15
	Dropout	Proportion of units to drop for the input's linear transformation	uniform(0,0.3)
	Optimizer	Optimizes the loss function	$[Adam, SGD, RMSprop]$
	Number of hidden layers	Number of hidden layers	[2,3,4]
	Hidden layer specification	Number of input neurons in the respective hidden layer	[64,128,256,512,1024]
	Activation function of output layer	Function that specifies when an output neuron is activated	$Sigmoid$
	Loss function	Function that is minimized during training	Binary Crossentropy
LSTM	Recurrent Dropout	Proportion of units to drop for recurrent state's linear transformation	uniform(0,0.3)
RF	Number of decision trees	Number of decision trees that are built during the training process	[50,100,200,400,800, 1600]
	Number of maximum Features	Maximum number of features used for a split	uniform(0,numFeatures)
	Minimum number of leaves	The minimum number of samples that must be at a leaf node	rand(1,11)
	Minimum number of samples	Number of cases required to build a new decision criterion	rand(2,11)
	Maximum depth	Limit of the maximum depth; increasing the depth can increase the precision but tends to cause overfitting	$[3, Indefinite]$
	Bootstrap	Defines whether bootstrap samples are used when building trees	$[true, false]$
	Criterion	Defines the function to measure the quality of a split	$[gini, entropy]$
SVM	Kernel	Kernel type used in the algorithm	rbf
	C	Penalty parameter of the error term; weights the misclassification of training examples against simplicity of the classifier	rand(1,100)
	Gamma	Kernel coefficient defining how much influence a single training example has	$\text{rand}(10^{-4}, 10^{3})$

[1] We used this parameter as an upper limit in order to save computing time. Additionally, we applied an early stopping approach to terminate the training process as long as there were no further improvements. On average, the training was stopped after the 30th epoch.