**Thesis First Draft**

Introduction

* Background
* Problem Statement
* Research Question
* Outlook

Literature Review/Theoretical Background

* Business Process Management
* ML in Predictive process monitoring
  + SVM & RF
* DL in Predictive process monitoring
  + DL & LSTM
* Event Log properties
* Performance Evaluation

Research Method/ Study Design

* Data Collection
  + Description of used event log
  + Sampling method
* Data Preprocessing
  + ?
* Implementing Classifier

Discussion/Results interpretation

* Accuracy/F1 Score
* Comparison to Kratsch’s outcome prediction

Conclusion

* Summary
* Implementation
* Limitation and Future research

Question:

Preprocessing encoding method: LSTM without categorical attributes/ one hot/ embedded etc.

2019 Cambgo paper/ 2019 Li Lin

LSTM using NEW embedded dimension instead of One-Hot Encoding

To do List

* Having drafted code for prediction. SVM/RF/DT
* Set down all used dataset.
* If having time, try out prediction task.

Preprocess and ML

<https://github.com/irhete/predictive-monitoring-benchmark>

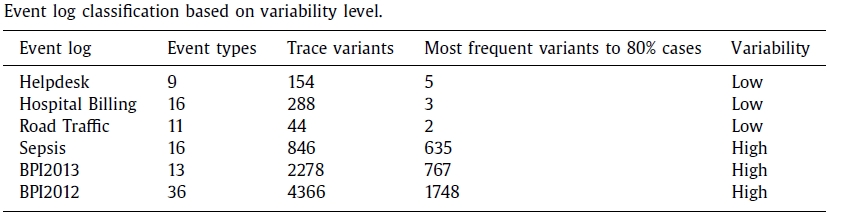
Verenich,2016 SVM

Knockout problem

Unuvar,2016 DT

Comparision:

Update variability: number of variants that covers 80% cases. Help desk vs. 2012.



2017 LSTM Tax paper for simple prediction

<https://github.com/verenich/ProcessSequencePrediction>

Preprocessing

|  |  |  |
| --- | --- | --- |
| Nominal | case\_id | case |
|  | etype\_1 | Event type |
|  | etype\_2 |  |
|  | etype\_3 |  |
| Numerical | duration | Duration b/t fist and last event in the considered time windown |
| Label | Next\_etype | Next event type to be predicted |

