**Name of student:** Mickael LIU

**Name of your Level:** L0

**Source (e.g. scholars.google.com):** <https://www.researchgate.net/publication/352811334_Improving_the_In-Hospital_Mortality_Prediction_of_Diabetes_ICU_Patients_Using_a_Process_MiningDeep_Learning_Architecture>

**Paper title:** Improving the In-Hospital Mortality Prediction of Diabetes ICU Patients Using a Process Mining/Deep Learning Architecture

**Keywords specific to the paper:**

Process mining, deep learning, in-hospital mortality, risk assessment, diabetes, intensive care

**Summary of the main contributions:**

With 26.9 million diagnosed patients in the United States and accounting for 45% of ICU patients above the age of 65, diabetes mellitus (DM) is a widespread illness that requires unique attention. There are two main types of DM. Type I is controlled with diet and insulin. Type II, the more common, is controlled with diet, exercise, and a multitude of medicines. However, even a small infection can lead to severe patient outcomes and difficulties to control the disease. Hence, DM is rarely a standalone reason for severe patient outcomes but a factor that increases the likelihood of potentially life-threatening outcomes. In the aftermath, hospitalized DM patients require significantly more healthcare resources compared to other chronic disease populations. One way to allocate healthcare resources more efficiently and to lower the rate of mortality is a precise in-hospital mortality risk assessment of diabetes ICU patients.

The calculation of risk and mortality scores in hospital settings has a long-standing history and is well-studied in the literature. The first methods were developed in the 1980s and are used in healthcare facilities for decades. Commonly, such methods target the general patient population and include comorbidity assessment methods and specific risk assessments like the probability for organ failure. Risk and mortality scores are calculated mostly based on patient information of the current admission and provide a snapshot of the patient’s condition. With the ongoing adoption of Electronic Health Records (EHRs), healthcare facilities are building empiric patient data repositories. This enables the application of data mining and machine learning approaches. However, recent methods often neglect EHR data of a patient's past hospital encounters when assessing their risk, or require substantial financial investment to be applicable in the real-world.

A patient’s health records can be understood as a sequence of observations. Such observations may include performed services, diagnoses, or lab measurements, and are also known as careflows. Process mining is a comparatively young research discipline that aims to extract knowledge from such sequences. Applications of process mining can be found across many industries, mainly to analyze and optimize applied processes, such as in business process management, automation, manufacturing, and recently in healthcare. Healthcare organizations increasingly acknowledge process mining and the use of empirical data to improve processes. However, there is a lack of studies that include the patient’s past hospital encounters using process mining to predict outcomes.

In this paper, a novel process mining/deep learning architecture is proposed to enhance established risk calculation methods of in-hospital mortality of diabetes ICU patients by incorporating patient care flows from earlier hospital admissions. More specifically, the paper demonstrates a strategy to convert electronic health records to a careflow format suitable for process mining and how this information can be used to predict the patient outcome. The proposed approach leverages demographic information, diagnoses and procedures, diabetes-related health measurements, and existing risk scores that are calculated after 24 hours after being admitted in the hospital .

The contributions of this paper are two-fold :

* First, a process mining/deep learning architecture is proposed to transform careflows embedded in EHRs into event logs suitable for process mining.
* Second, the proposed architecture is successfully demonstrated to improve the in-hospital mortality prediction of diabetes ICU patients by enhancing established risk scoring methods. The manuscript accentuates the non-negligible importance of modeling past patient careflows for outcome prediction and highlights process mining as a prospective set of tools for future research directions.

This paper has demonstrated one of the first process mining based approaches to model historical EHR data of diabetes ICU patients in combination with severity scores to predict in-hospital mortality. Specifically, an approach has been introduced that converts past medical records prior to the index hospital admission to event logs that are suitable for process mining. Then, a combination of existing risk scoring methods and Decay Replay Mining is used to predict the probability of mortality of a patient. In this way, established methodologies are combined with the advantages of incorporating historical information that provides an increased holistic view of the patients’ conditions. The paper demonstrates significant performance improvements in predicting the in-hospital mortality of diabetes ICU patients that have a patient history in the hospital of the MIMIC-III database compared to established risk assessment scores and machine learning approaches.

However, the current methodology also has certain limitations :

1. This approach addresses hospitals with a strong longitudinal patient record and may not be useful for ones with minimal longitudinal patient history. Yet, this limitation can be overcome if information exchange is occurring between the index and other medical centers.
2. The approach focuses on DM patients only and is not easily transferable to other chronic diseases. Each chronic disease may require its own process mining based model to maintain a satisfying objective dependent predictive performance.
3. Studying the MIMIC-III database for analysis can be challenging because the timing of patients' procedures and diagnoses is not always included. This information is often recorded for billing reasons without specific timestamps. To enhance the reliability of the findings and confirm the assumptions, it's suggested to extend the research to other hospital databases with more comprehensive data.