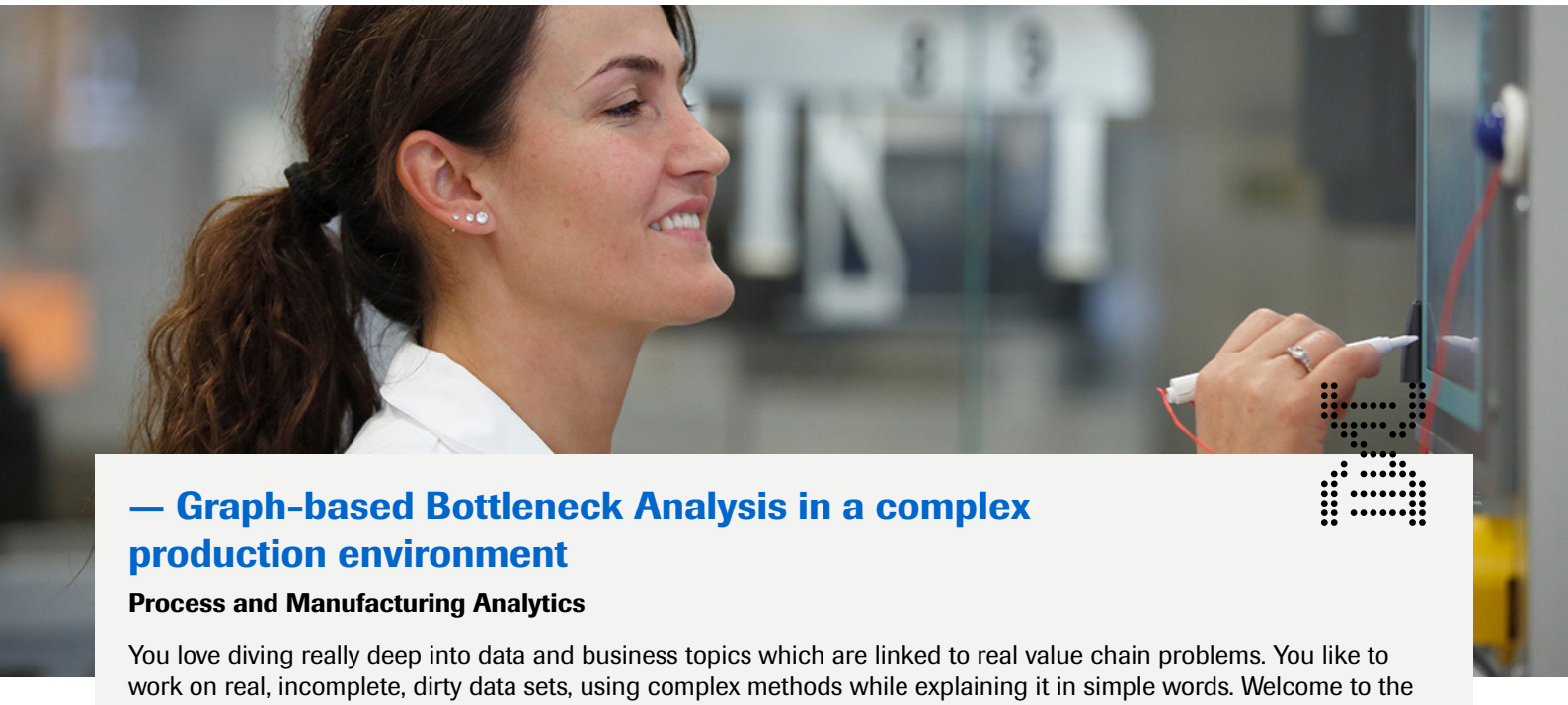
HEALTHCARE XPLORERS
Academic minds for innovation

Roche



— Graph-based Bottleneck Analysis in a complex production environment

Process and Manufacturing Analytics

You love diving really deep into data and business topics which are linked to real value chain problems. You like to work on real, incomplete, dirty data sets, using complex methods while explaining it in simple words. Welcome to the world of enterprise data science!

Apply until **April 18, 2021** / Xplorers Camp on **May 6, 2021**

Question to be solved

Early accurate diagnosis is essential to maintain health, prevent disease, as well as to find treatments and potential cures. The Roche Diagnostics products to be produced are mapped and scheduled on 100th of production resources (machines) with many constraints (capacity, money, dependencies). The problem can be seen as a sparse graph network with a continuous flow of work tasks to be scheduled under constraints (process industry). We are exploring novel methodologies on a graph-based approach to identify blocking elements (traffic) dynamically and near real-time scheduling possibilities to support business questions. Goal is to identify and resolve critical production bottlenecks in time and resources in a complex production network.

General Background

The overall topic lies in the field of value chain analytics f or diagnostics products . The key objective of value chain analytics is to source, produce, and deliver diagnostics products in highest quality and on time. We focus here on the sub-field of manufacturing operation of diagnostics products. A production task mapping is typically done within the process industry via a material requirements planning (MRP) run within enterprise resource planning systems (ERP). This run is very complex and typically lasts several hours, and is executed once or twice a week. Often this scheduling run is an optimization problem in time and resources. However, often incomplete information is available, which results in one possible outcome. Mostly it is not clear about the quality of the outcome since static information is used for scheduling. However, the world is dynamic, and information is uncertain (probabilistic). The research question can explore novel methods on sparse graphs to explore and identify the system's true dynamics and find new solutions for the so-called tactical business decision layer. The goal is to support data-driven decision making for business questions like:

- Which resource (machine) is most likely the bottleneck and why?
- What is the overall critical path, and what are options to reduce the overall lead time?
- How can we hierarchically decompose the problem space to obtain first a coarse grain solution?

Data Types & Technologies

The task is highly quantitative and the work will be carried out on historic production schedules data. Data sets are provided and partially cleaned, however the problem space requires a basic understanding of production scheduling problems. All proves and resulting visual prototypes are performed on data, backuped by theoretical ideas.

- Python programming know-how mandatory
- Usage of graph based libraries or databases favourable
- Visual prototyping via Tableau or other drill down tools.
- Data type: all about time series and relational data types.
- All approaches are welcome which taggle the problem in a massive parallelized way, e. g. problem mapping on TensorFlow

---

Supporting Material or Links

Time Continuity in Discrete Time Models: New Approaches for Production Planning in Process Industries, ISBN-10 : 3540245219

Needed Skills

- Good knowledge in Python
- Initial business understanding production scheduling

Mentor

Dr. Frank Kienle
Digital Strategy Manager, Materials & Business Process Management

---

Form of Cooperation

From master thesis to phd or long term cooperation

How to present your Idea

1-Pager formulation of your understanding of the problem and possible solution path (we do not expect a full solution overall, it is about problem break down and how you explain it on upcoming high level questions)
3-5 slides pitch to explain your approach,
Proof of your python / programming skills, e.g. Github Links