FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

ASSIGNMENT OF MASTER'S THESIS

Title:	Utilizing AI/ML methods for measuring data quality
Student:	Bc. Michael Mikuš
Supervisor:	Ing. Tomáš Pajurek
Study Programme:	Informatics
Study Branch:	Knowledge Engineering
Department:	Department of Applied Mathematics
Validity:	Until the end of summer semester 2020/21

Instructions

Traditional data quality measurement methods need significant manual effort and domain expert experience. Apart from being a time-consuming process, these traditional methods are often based on manual definition of rules or thresholds and are therefore subject to human error. The utilization of AI/ML methods offers an alternative data-driven approach that might overcome some of the drawbacks of traditional methods.

1. Describe key aspects of data quality and review the state-of-the-art methods for monitoring and measuring data quality with special emphasis on AI/ML based approaches.
2. Experimentally compare traditional (non-AI) methods for measuring data quality with methods using AI/ML techniques.
3. Propose directions for further improvements in the application of AI/ML in data quality.

References

Will be provided by the supervisor.

Ing. Karel Klouda, Ph.D.
Head of Department

doc. RNDr. Ing. Marcel Jiřina, Ph.D.
Dean

Prague September 23, 2019