MASTER‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Ambroz Jakub** Personal ID number: **499162**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Computer Science**

Study program: **Open Informatics**

Specialisation: **Data Science**

II. Master’s thesis details

Master’s thesis title in English:

Anomaly Detection in Metrans Transport Orders

Master’s thesis title in Czech:

Detekce anomálií v transportních objednávkách firmy Metrans

Guidelines:

Metrans, a leading logistics company, handles a high volume of transportation orders daily. Each order includes details such as the client companies involved, the transportation route, and the types of wagons used for freight. Historical data will be analyzed to proactively identify anomalies or irregularities in orders, helping to prevent errors in future order creation and mitigate potential issues or financial losses.
1. Analyze the data from historical orders
2. Research existing methods for anomaly detection
3. Choose appropriate methods and apply them to this problem
4) Compare and measure their effectiveness on testing data that contains both valid and anomalies entries that will be provided by Metrans
5. Create synthetic data if necessary, especially for rare anomalous data
6. Develop a component that generates events to flag irregular or anomalous data points in real-time
During development use microservices patterns. Test the application continuously and deploy the final application into CodeNOW together with documentation.

Bibliography / sources:

[1] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly detection: A survey. ACM Comput. Surv. 41, 3, Article 15 (July 2009)
[2] Domingues, Remi & Filippone, Maurizio & Zouaoui, Jihane. (2017). A comparative evaluation of outlier detection algorithms: Experiments and analyses. Pattern Recognition. 74. 10.1016/j.patcog.2017.09.037.
[3] Liu, Fei Tony & Ting, Kai & Zhou, Zhi-Hua. (2009). Isolation Forest. 413 - 422. 10.1109/ICDM.2008.17.
[4] Chalapathy, Raghavendra, and Sanjay Chawla. "Deep learning for anomaly detection: A survey." arXiv preprint arXiv:1901.03407 (2019).
[5] ASA-SIAM Series on Statistics and Applied Mathematics, Data Clustering: Theory, Algorithms, and Applications, Guojun Gan, Chaoqun Ma, and Jianhong Wu

