**Individual Learning Plan** - Yalin Chi

I’m taking this data science course to improve my ability in data processing and model training, which can be crucial for my thesis and future career. I’m also interested in collaborating and mutual learning with my colleagues during the journey of encountering and overcoming challenges.

My personal goal in machine learning is to improve proficiency in using powerful software tools such as sklearn, LightGBM, and CatBoost to practice my skills in the field of data science. These softwares, according to our group’s knowledge and research, have proved to be efficient for dealing with large datasets and powerful for predicting with high accuracy. Although I am currently unfamiliar with these tools, I am excited to learn and apply my knowledge and understanding to our project. By learning these machine learning tools, I aim to contribute effectively to our project's objective of predicting the energy pattern of prosumers in Estonia.

I also want to practice my coding skills. I am equipped with the fundamental knowledge of Python but haven’t had a chance to put the coding theories into practice that much. In our project, the first step would be data wrangling using Pandas to clean and select useful information from large datasets. I believe my Python skills will be helpful for the first phase of our project, which can be tedious but important for the further progression of our project. Meanwhile, I can get myself more familiar with data processing techniques, which could be useful for future studies.

Our project is a real-world case that Kaggle set up to solve the energy imbalance issue in Estonia. My motivation for the project is to move from simply learning theory to trying to solve real-world problems. Until now, I've mostly dealt with hypothetical situations and never got the chance to tackle genuine issues. It can be harder and more complicated, but I'm excited to apply what I've learned to practical situations. It is important to not just understand the concepts but also to be able to use that knowledge to solve actual problems.

I am keen to explore the ethical dimensions of data science. With the increasing prevalence of AI, ethical concerns surrounding data usage have become contentious. It is crucial to learn the appropriate methods of handling data, ensuring responsible usage without infringing on anyone's rights. Currently, I lack expertise in the ethical considerations related to data usage, therefore this course will be a great opportunity for me to gain insights and critically reflect on the ethical aspects of our upcoming project.