

Project teams - Machine Learning 2024

Read this before proceeding but do not modify this part (big brother is watching!):

TEMPLATE (Please, don't edit this part):

Title: short title

TXX is a global project identifier for each Team. Please, iterate it yourselves (T01, T02, ...)
If you choose a Project from a partner (e.g. P16) you should specify the project ID (P16) in a description of your project. If you sign up for a project from a partner, make sure that your team does not exceed the limit of teams accepted by the partner, otherwise, choose a different project.

Team: Member 1; Member 2; Member 3; Member 4; (up to four team members)

Project type: a project from a partner (what project ID?) / self-proposed / Kaggle competition

Description: a short description of your project (even if you take a project from a partner) - write in your own words.

Link: provide a link to the code repository.

Preferred presentation day: Monday/Tuesday/Wednesday

EXAMPLE PROJECT (Please, do not edit):

T00: Building artificial general intelligence using GPT3 and U-Net

Team: Dmytro Fishman; Hasan Tanvir; Dzvinka Yarish; ...;

Project type: Self-proposed project
wei

Description: We will use a combination of NLP and computer vision algorithms (U-Net + GPT3) to design the first-ever model to recognize common geometric primitives and describe them in the rrrrrrrhuman-understandable language (English). We will use transfer learning to make our model mimic all models, to achieve artificial general intelligence (AGI)
We are very sure that this will work out as we are using the latest smartphone to make calculations. Also, we are ready to invest in buying another smartphone to exploit the wisdom of crowds.

Link: <https://github.com/skyfallen/AGI> (this is an example)

Preferred presentation day: Monday

Add your project and the team below:

Title: T01 - Automatic labeling of traffic light detection dataset

Team: Leyla Eminova, Vusal Abdullayev, Eiko Keeman, Vinayak Khavasi

Project type: Partner Project - P03

Description: ...

Link: ...

Preferred presentation day: (Monday/Tuesday/Wednesday)

Title: T02 - Honey Fraud Detection through DNA-Based Taxonomical Composition Analysis

Team: Sofiya Babok, Stepan Strelchenko

Project type: based on P02 - Honey Fraud Detection through DNA-Based Taxonomical Composition Analysis

Description: ...

Link: [private git repository](#)

Preferred presentation day: (Monday/Tuesday/Wednesday)

Title: T03 - Predicting which businesses are unlikely to pay their fines

Team: Erle Maidu, Kersti Mikkov, Karl Soosalu, Ants Adamson

Project type: based on P09 - Predicting which businesses are unlikely to pay their fines

Description:

We will build an ML model to predict which businesses registered in the Estonian business register are unlikely to pay their fines and to make this prediction even before the fine has been issued. With this model we will give our input into improving the Estonian business environment and making the playground more fair for all participants.

Link: <https://github.com/erlemaido/ML-prediction-of-fine-payment-RIK>

Preferred presentation day: Tuesday

Title: T04 - Smart invoice classification with Trigon

Team: Mihkel Roomet; Marilyn Ahvenainen; Veronika Kukk; Taavi Eistre

Project type: based on P10 - Smart invoice classification with Trigon

Description: Helping accountants do manual invoice categorization quicker by building a model that predicts default values.

Link: https://github.com/mihkelroomet/ML_Invoices_Trigon

Preferred presentation day: Monday

Title: T05 - Factors of success in abstract theoretical courses

Team: Peeter Pissarenko, an-Julius Laak, Aksel Õim, Kirill Rymak

Project type: P12 - Factors of success in abstract theoretical courses

Description: Analysing course data from a theoretical maths course to determine whether there are some specific actions that students do (or events) that lead to high performance.

Link: ...

Preferred presentation day: Wednesdays

P02 and P10 have already been taken, please, do not choose it.

Title: T06 - Automatic labeling of traffic light detection dataset

Team: Shivam Maheshwari, Darren Gallois, Lamouille Dorian, Venkata Narayana B

Project type: based on Partner Project - 03

Description: ...

Link: <https://github.com/chunkydonut21/automatic-labelling-of-traffic-light/>

Preferred presentation day: (Monday/Tuesday/Wednesday)

Title: T07 - Time-series forecasting: renewable electricity production

Team: Julius Välja, Kaur Kullamäe, Martin Hõbemägi, Phasha Davrishev

Project type: based on Partner Project - P07

Description: Predicting the hourly production of renewable electricity in the Baltic region

Link: <https://github.com/JuliusValja/ML-2024-Renewable-energy-project>

Preferred presentation day: (Monday/Tuesday/Wednesday)

P02, P07 and P10 have already been taken, please, do not choose it.

Title: T08 - Predicting which businesses are unlikely to pay their fines

Team: Robert Israel, Liis Harjo, Mary-Ann Kubre, Ahmed Ruhul Quddos Joyon

Project type: P09 - Predicting which businesses are unlikely to pay their fines

Description: ...

Link: ...

Preferred presentation day: Monday or Wednesday

Title: T09 - Rexplorer solar energy ML development

Team: Ija Rõõs, Marie-Johanna Perli, Anton Katsuba, Helen Kustavus

Project type: P01 - Rexplorer solar energy ML development

Description: ...

Link: ...

Preferred presentation day: Monday

Title: T10 - Time series forecasting: hourly national electricity consumption

Team: Zeyad Kenawi, Markus Rene Pae, Kelem Negasi Amare, Mohamed Nagy

Project type: P06 - day ahead prices for electricity

Description: ...

Link: ...

Preferred presentation day: (Monday/Tuesday/Wednesday)

P02, P06, P07 and P10 have already been taken, please, do not choose it.

Title: T11 - Segmentation Smackdown: Battle of the Models

Team: Andero Raava, Kristel Saul, Tarvi Tamm, Pirje Õunpuu

Project type: P11

Description: Improving the segmentation approach implemented so far by testing state-of-the-art segmentation approaches and then fine-tuning these models.

Link: ...

Preferred presentation day: Monday

Title: T12 - Factors of success in abstract theoretical courses

Team: Pirjo Vainjärv, Anna Maria Tammin, Kaidi Tootmaa, Hanna-Maria Kukk

Project type: P12 - Factors of success in abstract theoretical courses

Description: The goal is to find activity patterns to separate students by efficiency and thus predicting the students final score.

Link: [GitHub repository](#)

Preferred presentation day: Tuesday

Title T13: Kaggle - Quick, Draw! Doodle Recognition Challenge

Team: Iryna Hurova, Ivan Gorbachenko, Kirils Gončarovs, Arthur Alexander Kanoshin

Project type: Kaggle - Quick, Draw! Doodle Recognition Challenge

Description: The goal is to predict the correct "word" of the drawing.

Link: <https://www.kaggle.com/competitions/quickdraw-doodle-recognition/data>

Preferred presentation day: (Monday/Tuesday/Wednesday)

P02, P06, P07, P09 and P10 have already been taken, please, do not choose it.

Title: T14 - Time-series forecasting: electricity prices

Team: Tarlan Soltanov, Akshin Mashadiyev, Adil Musali, Orkhan Hasanzade

Project type: P08

Description: ...

Link: ...

Preferred presentation day: Wednesday

P02, P06, P07, P08, P09 and P10 have already been taken, please, do not choose it.

Title: T15 - Child Mind Institute — Problematic Internet Use

Team: Andres Caceres, Jonathan Lopez, Ahmed Soliman

Project type: [Kaggle Challenge](#)

Description: develop a predictive model that analyzes children's physical activity and fitness data to identify early signs of problematic internet use

Link: [Github Repository](#)

Preferred presentation day: Tuesday

Title: T16 - P12 - Factors of success in abstract theoretical courses

Team: Phil de Guzman, Jannis Rosenbaum, Mislav Rukonic, Jerin George

Project type: Partner

Description: Courses containing a lot of mathematics are traditionally hard for students. However, there are still some students who excel in them. The objective is to find out what these students do differently and whether there are activities or behavioral patterns that could be recommended to the other students to make such courses easier for them.

The dataset consists of coded activities of 119 students in the Theoretical Computer Science course each day throughout the whole semester. The activities are, for example, viewing study materials, submitting homework, participating in a lecture online, etc. The goals are: 1) predict the final grades of students based on their activity pattern; 2) discover activity patterns that separate high-performing students from low-performing students.

Link:

Preferred presentation day: (**Monday**)

Title: T17 - Predicting Estonian Stock Market Movements Using Sentiment Analysis of News

Team: Oliver Vaht, Kirill Anohin, Märten Mikk, Henri Annilo

Project type: Self-proposed project

Description: We aim to predict Estonian stock market movements by analyzing sentiment from Estonian news articles. Using natural language processing techniques, we'll perform sentiment analysis on collected news data and correlate the results with stock market trends to build a predictive model.

Link: ...

Preferred presentation day: (Tuesday)

Title: T18 - Lidar-only self-supervised 3D object detection network

Team: Buraq Khan, Joshua Hamilton-Brown, Markkus Koddala, Greete Aava

Project type: P05

Description: Train lidar-only self-supervised 3D object detection network. Replicate the LISO paper: <https://baurst.github.io/liso/>

Link: [Private GitHub Repo](#)

Preferred presentation day: (Monday/Tuesday/Wednesday)

P02, P06, P07, P08, P09 and P10 have already been taken, please, do not choose it.

Title: T19 - Automatic labeling of vehicle blinker detection dataset

Team: Sten Marcus Nelson, Salme Ussanov, Eidi Paas, Noman Ashraf

Project type: P04

Description: Automatic labeling of vehicle blinker detection dataset using Segment Anything 2 model. Knowing the state of blinkers is important for the prediction module, to estimate if the vehicle is going to change lane or turn. Detecting blinkers from single video frame is not possible, as they are, you know, blinking. Bonus is to train Yolo detection model that outputs in addition to vehicle bounding boxes also the blinker state as class.

Link: [GitHub Repo](#)

Preferred presentation day: (Wednesday)

Title: T20 - Classification with an Academic Success Dataset

Team: Juan Gonzalo Quiroz Cadavid, Naiara Alonso Montes, Fidan Suleymanova

Project type: Kaggle

Description: Based on multiple information about students in higher education levels, the aim is to predict if students will graduate/drop their studies.

Link: [Kaggle competition](#)

Preferred presentation day: Monday

T21: Identifying Key Factors for High-Speed Production Line Efficiency

Team: Marek Bankiir; Anjish Panta, Famil Babayev, Marten Kuusmann

Project type: Self-proposed project (in collaboration with a software company)

Description:

The goal of this project is to identify key factors that influence the speed of a production line using data collected from line sensors at one-minute intervals. The project will involve binary classification to distinguish between conditions that lead to high-speed vs. low-speed production. The data is pre-labeled with high-speed and low-speed categories, allowing us to apply classification models (such as decision trees and random forests) to uncover patterns in the data. Additionally, classification trees will help visualise complex relationships between variables, providing insights into how various factors affect production speed. We also plan to carry out extensive feature engineering to optimise model performance and extract meaningful insights from the sensor data. This project presents a real-world challenge with direct implications for improving production efficiency.

Link: https://github.com/MarekBEst/ML_T21/

Preferred presentation day: Tuesday

T22: Diabetes Risk Factor Analysis

Team: Yang Sheng; Zhou Han; William Vaask

Project type: Kaggle

Description: Predict the onset of diabetes based on diagnostic measures

Link: [Pima Indians Diabetes Database \(kaggle.com\)](#)

Preferred presentation day: except Wednesday and Friday

T2

T23: Classification with an Academic Success

Team: Andres Alumets, Anton Zakatov, Pihla Järv, Sohaib Anwer

Project type: Kaggle

Description: Develop a ML model that would enable predicting academic success based on students' data. The goal of this competition is to predict academic risk (categorically) of students in higher education. This is a Kaggle competition that closed on 1 July 2024.

Link:

- Competition: <https://www.kaggle.com/competitions/playground-series-s4e6/overview>
- Repo: <https://github.com/azakatov/ut-ml-project-2024> (private)

Preferred presentation day: Monday, preferably during the lecture and not the practice session

T24: Skin Cancer Detection with 3D-TBP

Team: Minnie Zhou , Zekang Chen , Ashley Williams, Fedir Kyrychenko

Project type: Kaggle

Description: To develop a method for image-based algorithms to identify histologically confirmed skin cancer cases with single-lesion crops from 3D total body photos (TBP). The detection algorithms can help the diagnosis and initialise its treatment in a timely manner

Link: <https://www.kaggle.com/competitions/isic-2024-challenge/overview/evaluation>

Kaggle: <https://www.kaggle.com/code/fedirkrychenko/testcancer/edit/run/213276772>

Preferred presentation day: (Monday/Tuesday/Wednesday)

T25: P03 - Automatic labeling of traffic light detection dataset

Team: Ilmar Möls, Mark-Erik Aan, Toomas Kangur, Jaan Otter

Project type: Partner (P03)

Description: The project involves creating an automatic labeling system for a traffic light detection dataset using the Segment Anything Model (SAM2). The dataset will include annotations for image sequences, which are crucial for detecting blinking lights - a task that cannot be accomplished using single images alone. The challenge includes training a lightweight Yolo model on the created dataset for efficient traffic light detection. Artificial data generation may be employed, and the complexity of labeling and training on video sequences presents a significant challenge. Successful implementation could lead to advancements in autonomous driving systems.

Link: <https://github.com/JAAN555/Machine-Learning-Project.git>

Preferred presentation day: Wednesday

P02, P03, P06, P07, P08, P09 and P10 have already been taken, please, do not choose it.

T26: Detection and labelling of Traffic signs

Team: Muhammad Zain Razzaq, Asier Mandiola and Borja Palacios

Project type: Kaggle

Description: We are going to train model based on dataset of traffic signs and then detect and classify them in a full road picture

Link: ...

Preferred presentation day: TBD

...

T27: Threat actor grouping and possible attribution based on known indicators of compromise using IP addresses metadata

Team: Marvin Uku, Taavi Kase, Joosep Orasmäe, Karl-Ingmar Adamson

Project type: Self-proposed project

Description:

Goal is to group IP addresses based on the similarities in the metadata field values and group them together so that possible infrastructure groups would emerge, which are being used by some groups for attacking or hosting command and control servers. Based on some known indicators of compromise (intelligence), these groups could be tied to some kind of actor group,

without doing the usual pivoting from one attribute to actually finding something. This model should be able to predict (if you have enough intel for that (it can be simulated with made up groups)), which group it is with probabilities based on importance of feature analysis.

Link: [private](#)

Preferred presentation day: Monday/Tuesday/Wednesday

T28: Child Mind Institute - Problematic Internet Use

Team: Vita Komel, Hanna Junk, Bipul Gyawali

Project type: [Kaggle](#)

Description: The project is to develop a predictive model that can identify early signs of problematic internet use in children and adolescents based on their physical activity and fitness data.

Link: https://github.com/vicquad/ml24_cmi

Preferred presentation day: Monday

Title: T29 - Skin Cancer Detection with 3D-TBP

Team: Anastasiia Petrovych, Liliana Hotsko, Zhexin Rao, Zhen Xu

Project type: [Kaggle](#)

Description: This project is focused on developing image-based algorithms to identify histologically confirmed skin cancer cases with single-lesion crops from 3D total body photos

Link: [GitHub](#)

Preferred presentation day: Wednesday

Title: T30 - Coupled seq2seq training

Team: Swagata Datta, Syed Muhammad Hamza Bacha, Silvestro Matteo Malara, Rohan Iyer

Project type: P13

Description: This project focuses on extending a novel method for training sequence-to-sequence (seq2seq) models, where a pre-trained model (e.g., Whisper, mT5, NLLB) serves as a guide for training a new model. This approach ensures the new model's vector space remains compatible with the guide model, allowing their encoders and decoders to be recombined for flexible use. The method has been tested on machine translation, but the goal is to expand its application to other modalities like speech, images, music generation, and additional NLP tasks. The project involves identifying more seq2seq models on HuggingFace and updating the method to support these models.

Link:

Preferred presentation day: (Monday/Tuesday)

Title: T31 - Lidar-only self-supervised 3D object detection network

Team: Oliver Vainikko, Mart-Mihkel Aun, Evaldas Petnunas, Felix Aschpurwis

Project type: P05

Description: Train lidar-only self-supervised 3D object detection network. Replicate the [LISO paper](#)

Link: ...

Preferred presentation day: Monday

Title: T32 - House Prices - Advanced Regression Techniques

Team: Natig Agharzayev, Deniss Ruder, Markus Ilves

Project type: [Kaggle](#)

Description: Ask a home buyer to describe their dream house, and they probably won't begin with the height of the basement ceiling or the proximity to an east-west railroad. But this playground competition's dataset proves that much more influences price negotiations than the number of bedrooms or a white-picket fence.

With 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, this competition challenges you to predict the final price of each home.

Link: [here](#)

Preferred presentation day: Tuesday

Title: T33 - Molecular Classification and Property Prediction

Team: Lucas Ferrando Plo, Iris Hättestrand, Alice Kollnitz

Project type: Self-proposed project

Description: Several molecular properties like the boiling point will be predicted. Several possibilities for the input features will be explored, such as fingerprints or molecular graphs. Additionally, a comparison between a global and compound-type-specific model will be made. Firstly, the train molecule set will be clustered and a model will be fit to each one; Secondly, another model will be trained on the whole dataset. The performances will be compared.

Link: ...

Preferred presentation day: Tuesday

Title: T34 - Factors of success in abstract theoretical courses

Team: Jared Wildermuth, Kalju Jake Nekvasil, Mihkel Hani, Edvin Martin Andrejev

Project type: P12 **Factors of success in abstract theoretical courses**

Description: Courses containing a lot of mathematics are traditionally hard for students. However, there are still some students who excel in them. The objective is to find out what these students do differently and whether there are activities or behavioural patterns that could be recommended to the other students to make such courses easier for them.

Link: [here](#)

Preferred Presentation Day: Tuesday

Title: T35 - Segmentation Smackdown: Battle of the Models

Team: Maria Anett Kaha, Argo Aljand, Artur Eksi, Raimo Kõidam

Project type: P11

Description: Segmenting fibrous structures in scanning electron microscopy images is challenging due to the diverse fiber shapes. This project aims to enhance the segmentation methods used in FiBar (<https://fiбар.elixir.ut.ee/>). We will test advanced segmentation techniques, including the Segment Anything Model by Meta and U-Net models, and refine them to achieve better results.

Link: ...

Preferred presentation day: Monday

P02, P03, P06, P07, P08, P09 and P10 have already been taken, please, do not choose it. Please, do not take P12 (factors of success) as well, that is enough teams for this project :)

Team: T36

Title: Loan Approval Prediction

Team: Gregor Nepste, Markus Müüripeal, Vladimirs Racejevs

Project type: Kaggle

Description: The goal for this is to predict whether an applicant is approved for a loan. For this the competition has provided a dataset.

Link: <https://www.kaggle.com/competitions/playground-series-s4e10>

Preferred presentation day: Wednesday

Title: T37 - Segmentation Smackdown: Battle of the Models

Team: Lenards Skrodelis, Georgs Narbutis, Bekarys Toleshov, Pablo Ramirez

Project type: P11

Description: The main goal is to improve the segmentation approach implemented thus far in FiBar. The idea is to test state-of-the-art segmentation approaches and fine-tune these models to improve segmentation.

Link: Gitlab TBA

Preferred presentation day: Monday

Title: T39

Team: Fidan Karimova, Naveen Kumar

Project type: Grain humidity prediction

Description: Use any ML tools to predict the amount of humidity remaining in the grain during the drying process and to forecast future humidity (in order to estimate the process finishing time).

To be stored safely, grain must be dried below a certain level of humidity. Right now, one can set the machine to dry for a number of hours, predicting the time it needs to dry based on gut feeling and experience. This results in either a waste of energy due to running the machine for too long, or needing to launch the machine again, as the needed level was not reached (waste of time at the busiest period of the year — harvest).

Measuring the humidity of grain during the drying process would allow stopping the process at the right moment, but it necessitates a costly sensor. Instead, the remaining humidity amount should be predicted based on temperature measures (at different locations along the machine), environmental variables, and the original humidity of the grain at $T=0$.

Link: TBA

Preferred presentation day: Monday

Title: T40

Team: Gustav Nikopensius, Lauri Lopp, Martin Leissoo

Project type: Kaggle - Solar Power Generation Data

Description: This data has been gathered at two solar power plants in India over a 34 day period.

Can we predict the power generation for the next couple of days? - this allows for better grid management

Can we identify the need for panel cleaning/maintenance?

Can we identify faulty or suboptimally performing equipment?

Link: [Solar Power Generation Data \(kaggle.com\)](https://www.kaggle.com/datasets/ashishpatel26/solar-power-generation-data)

Preferred presentation day: Monday

Title: T41

Team: Jaagup Tamme, Belinda Lepmets, Juhan Oskar Hennoste, Ronald Judin

Project type: Kaggle competition

Description: Predict energy usage of prosumers, or the ones who both consume and generate energy. Data we will have to work with is the weather data, energy prices and records of the installed photovoltaic capacity. In order to submit to Kaggle we have to use the time series API.

Link: [Enefit - Predict Energy Behavior of Prosumers \(kaggle.com\)](https://www.kaggle.com/datasets/enefit/enefit-predict-energy-behavior-of-prosumers)

Preferred presentation day: Monday

Title: T42 -**Team:** Riki-Taavi Nurm, Kevin Sults, Rait Robert Suurmets, Stefan Ehin**Project type:** Kaggle - Solar Power Generation Data**Description:** This data has been gathered at two solar power plants in India over a 34 day period. It has two pairs of files - each pair has one power generation dataset and one sensor readings dataset. The power generation datasets are gathered at the inverter level - each inverter has multiple lines of solar panels attached to it. The sensor data is gathered at a plant level - single array of sensors optimally placed at the plant.

There are a few areas of concern at the solar power plant -

- 1) Can we predict the power generation for the next couple of days? - this allows for better grid management
- 2) Can we identify the need for panel cleaning/maintenance?
- 3) Can we identify faulty or suboptimally performing equipment?

Link: <https://www.kaggle.com/datasets/anikannal/solar-power-generation-data/data>**Preferred presentation day:** (Monday)

Title: T43**Team:** Mark Edvard Oliver Oja, Murad Magomedkerimov, Sander Saska, Teet Egert Vaabel**Project type:** Self-proposed**Description:** Using LiDAR it is hard to differentiate between hay bales and burial mounds. For that we will build a image recognition model.**Link:** TBA**Preferred presentation day:** Wednesday

Title: T44**Team:** an Madis Kask, Alfred Kraft**Project Type:** Self-proposed based on Kaggle - Top Hits Spotify from 2000-2019 dataset**Description:** Predict song similarity. Clustering songs based on their features (BPM, key, energy etc.) and suggesting similar songs.**Link:** https://github.com/KMKgit/ML_project_T44 (private)**Preferred presentation day:** Monday

Team: T45**Team:** Märten Josh Peedimaa, Kert Karsna, Karl Gustav Loog**Project type:** Kaggle competition - [Binary Prediction of Poisonous Mushrooms](#)**Description:** Try to predict whether a mushroom is edible or poisonous based on its physical characteristics. 20 features, 3+ million data points.

Link: [UT GitLab](#)

Preferred presentation day: Monday

Team: T46

Team: Inga Tallinn, Agnes Kala, Priit Peterson, an Variksoo

Project type: Kaggle competition - Mushroom Classification

Description:

Link: <https://www.kaggle.com/datasets/uciml/mushroom-classification/data>

Preferred presentation day: Monday/Tuesday/Wednesday

Title: T47 - Natural Language Processing with Disaster Tweets

Team: Taavi Vestel, Remi-Marcus Maiste, Sofia Kriuchkova, Simon Fox Kuuse

Project type: Kaggle competition

Description: Creating a machine learning model which is capable of filtering tweets of disasters. It's able to filter out disaster tweets which are not really about the disaster but similar to it. Data contains 10 000 tweets of which the model needs to be able to filter out false disaster tweets.

Link: <https://www.kaggle.com/competitions/nlp-getting-started>

Preferred presentation day: (Tuesday)

Title: T48

Team: Nikita Umov, Sviatoslav-Oleh Savchak

Project type: Kaggle competition

Description: In this competition your task is to predict whether a passenger was transported to an alternate dimension during the *Spaceship Titanic*'s collision with the spacetime anomaly. To help you make these predictions, you're given a set of personal records recovered from the ship's damaged computer system. (For learning process)

Link: <https://www.kaggle.com/competitions/spaceship-titanic>

Preferred presentation day: Monday/Tuesday/Wednesday

Title: T49 - Rexplorer solar energy ML development

Team: Maryne Dey, Loann Rio, Lucas Marais, Anton Vykhovanets

Project type: UT ML project proposals - P01

Description: Rexplorer is helping to automate solar power planning on buildings to improve the energy efficiency of our homes, offices and other real-estate.

The goal of the AI development project is to further develop our Gitlab library AI models used to detect installed PV & roof obstacles in solar project planning.

The library currently has 2 image recognition models, set of jupyter notebook that takes as input a geojson Polygon or Multipolygon features of areas of interest. The pipeline is built so that it is

run locally but all source files, running scripts and results are stored at separated protected server and could be accessed by SSH.

The task for students is to train the models by annotation of orthomosaic images highlighting feature borders and detected objects, measuring the test dataset IoU (intersection over union), polishing the scripts and scaling the model to new areas.

Link: Rexplorer OÜ (<https://www.replorer.ee/>)

Preferred presentation day: Wednesday

This is the final number of projects, please, do not add more.
