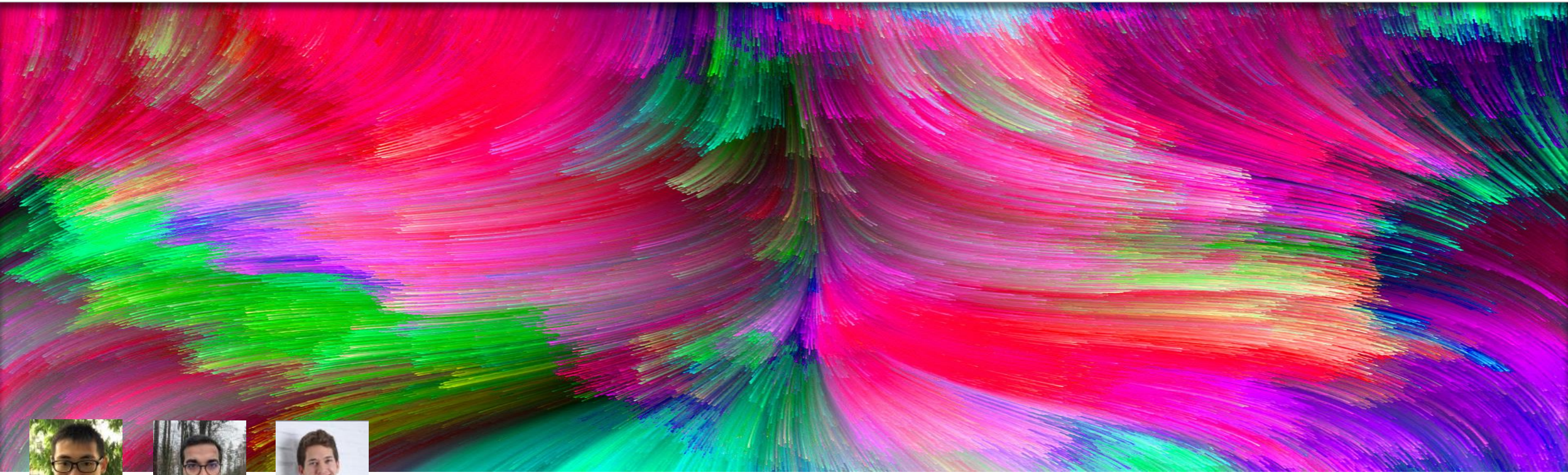
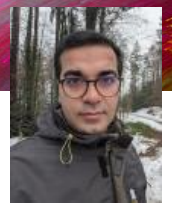



Spring Term ‘23 - 8,860,1.00 MCS Machine Learning

Lab 03 – Coding Challenge Kick-Off

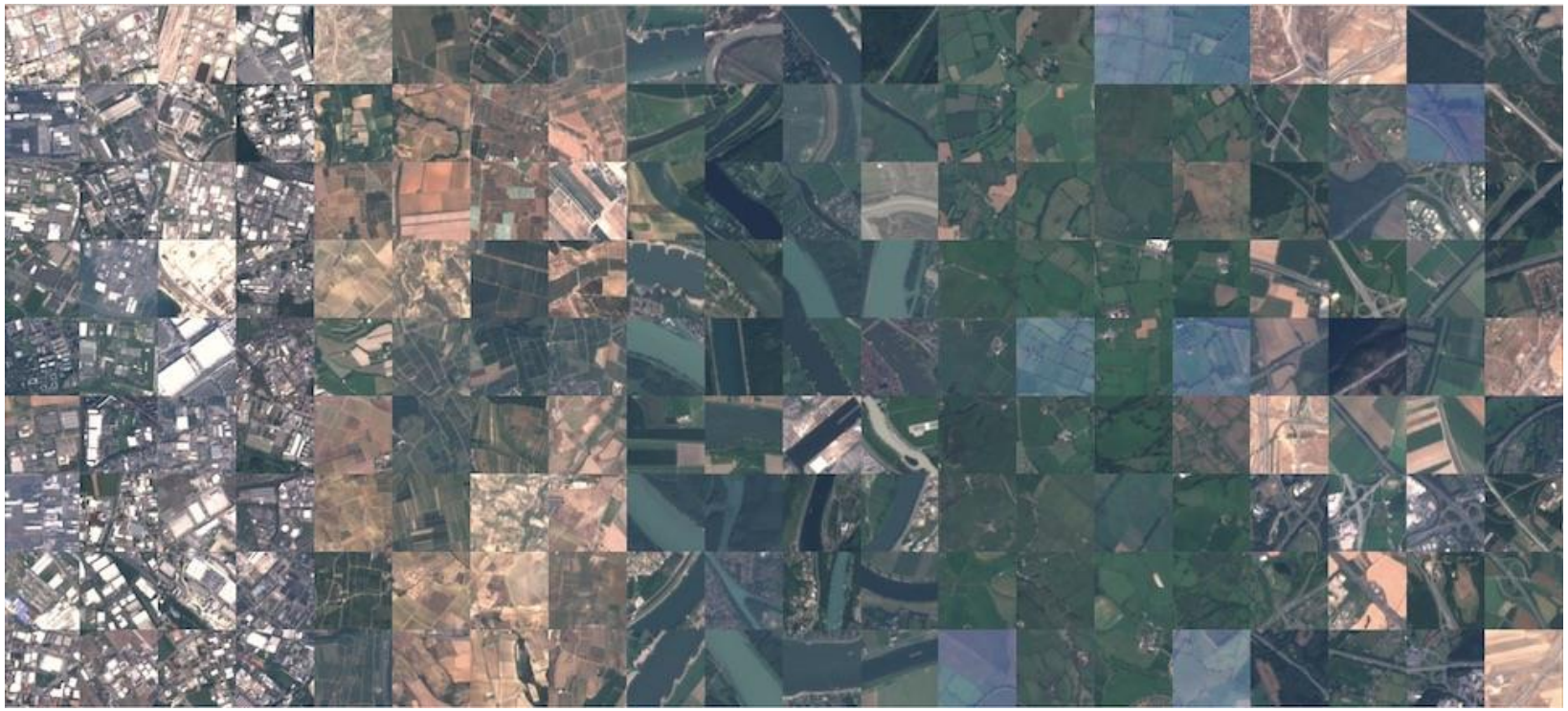
Shijun, Hamed, and Linus

From insight to impact.

Coding Challenge

kaggle

Satellite Image Classification - Eurosat



Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2217-2226.

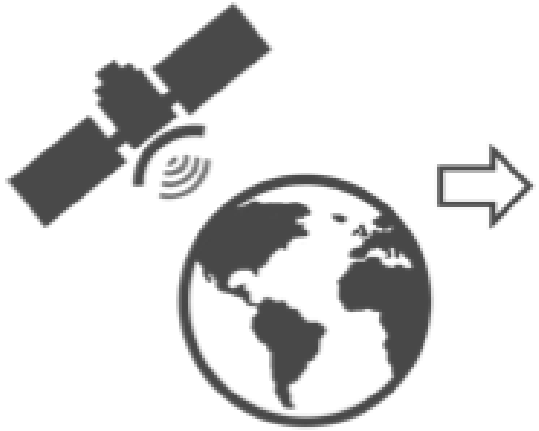
Course Logistics Updates

Weekly Updates and Questions

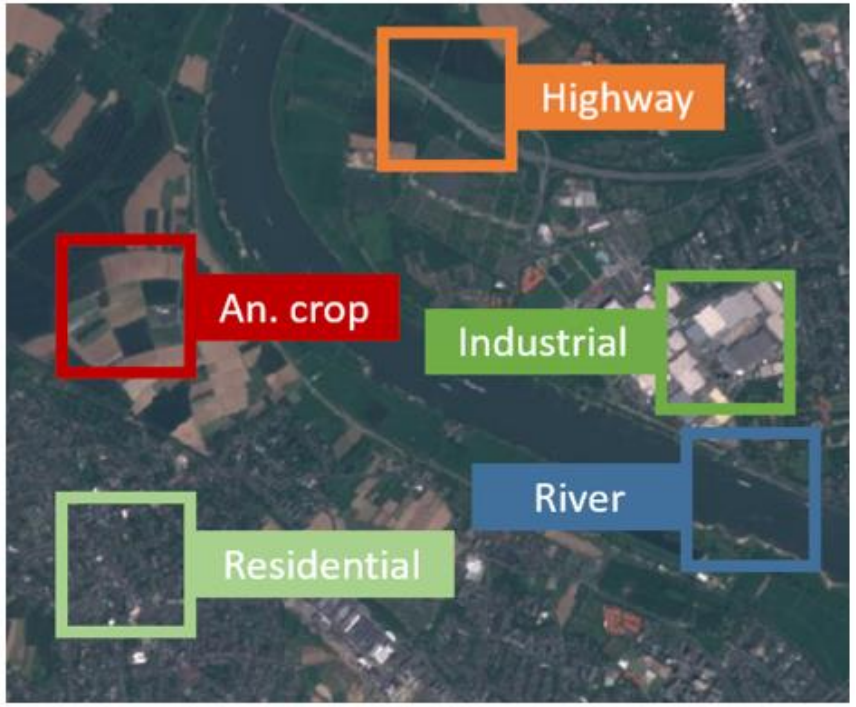
Course Labs and Challenge Roadmap

Event Type	Date	Description	Teaching Assistant	CC Phase
Lab 1	20 Feb 2023	Co-Lab Introduction	Shijun	Introduction
Lab 2	**27 Feb 2023**	**Support Vector Machine (SVM)**	**Shijun**	**Ramp-Up Phase**
CC 1	06 Mar 2023	**Coding Challenge - Kick-Off**	Linus	Coding Phase
Lab 3	13 Mar 2023	**Fully Connected Neural Networks**	Hamed	Ramp-Up Phase
Lab 4	20 Mar 2023	**Convolutional Neural Networks (CNNS)**	Hamed	Ramp-Up Phase
Lab 5	27 Mar 2023	**Long short-term memory (LSTM)**	Hamed	Ramp-Up Phase
Semester Break - Happy Easter!				

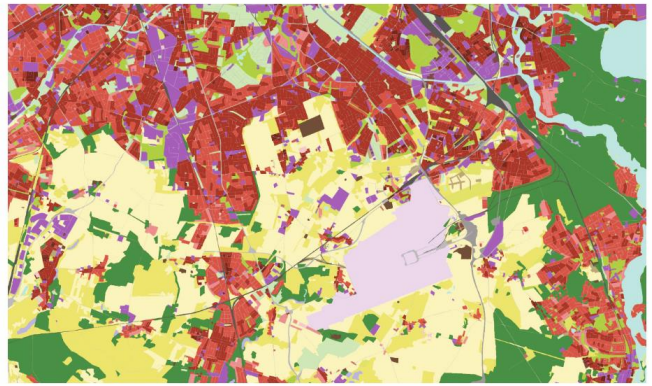
Eurosat



27,000 Sentinel-2 image patches



European Urban Atlas



South Berlin 2012

10 land-use classes

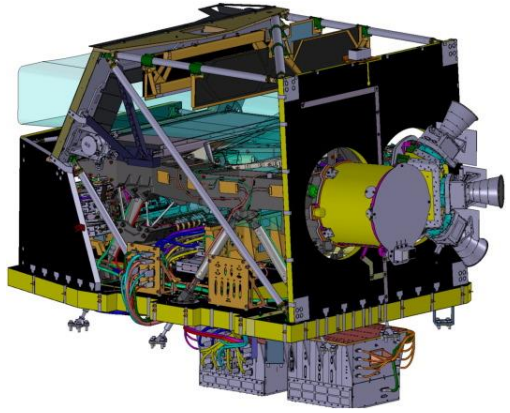
- Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2217-2226.
- Helber, P., Bischke, B., Dengel, A., & Borth, D. (2018, July). Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. In IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium (pp. 204-207). IEEE.
- https://land.copernicus.eu/local/urban-atlas

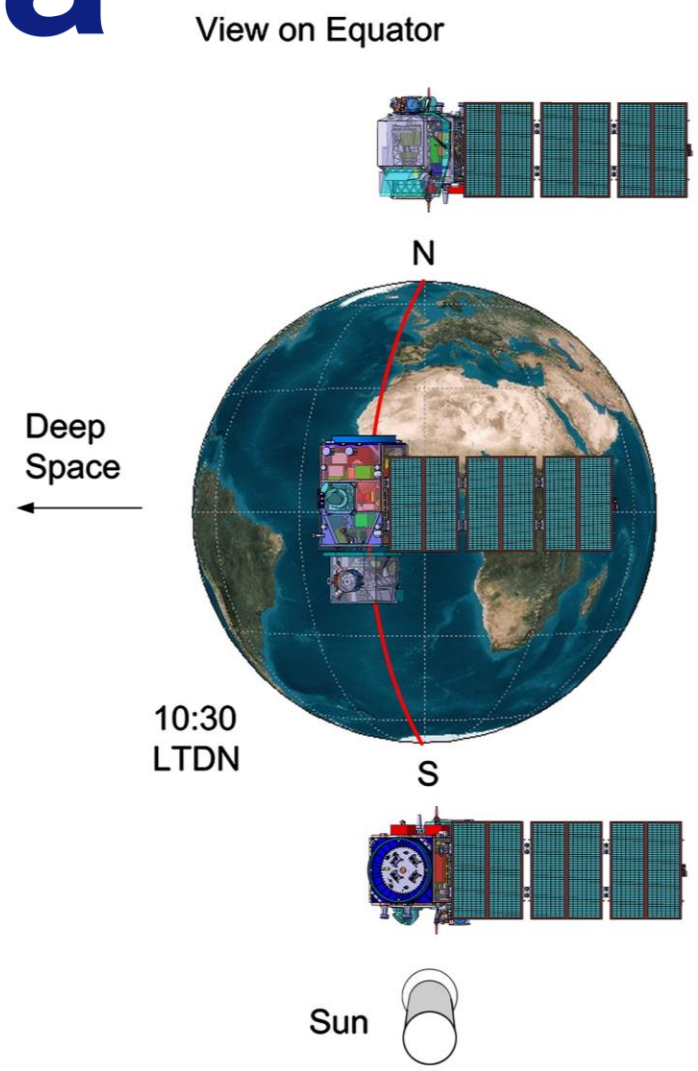
Sentinel-2

- ESA satellite mission with two satellites
- Launched 2015
- Sun-synchronous orbit
- 786km altitude
- 5-day revisit rate

Payload: **Multi Spectral Instrument**

- 290km field of view
- 13 spectral bands
- Visible, near- and short wave infra-red
- Spatial resolutions 10 to 60m

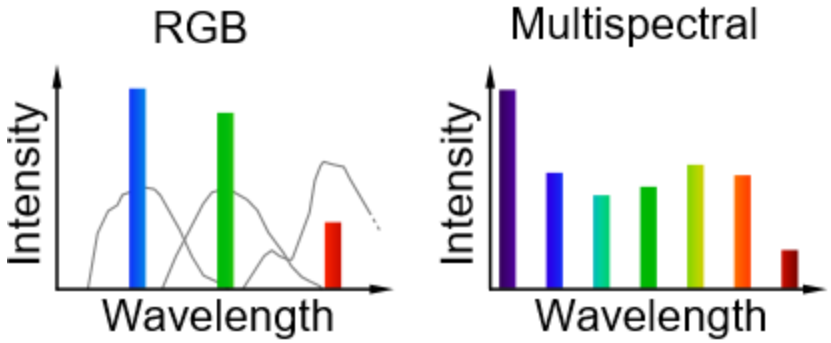




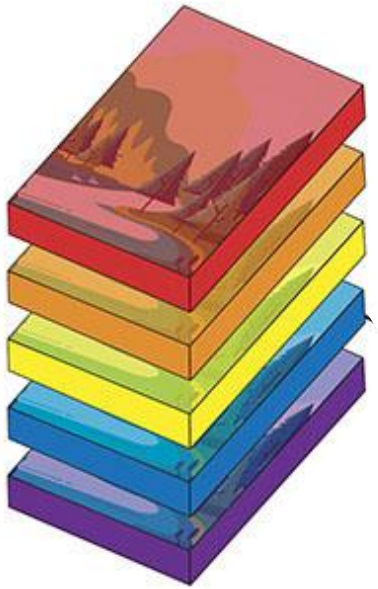
Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, *120*, 25-36.

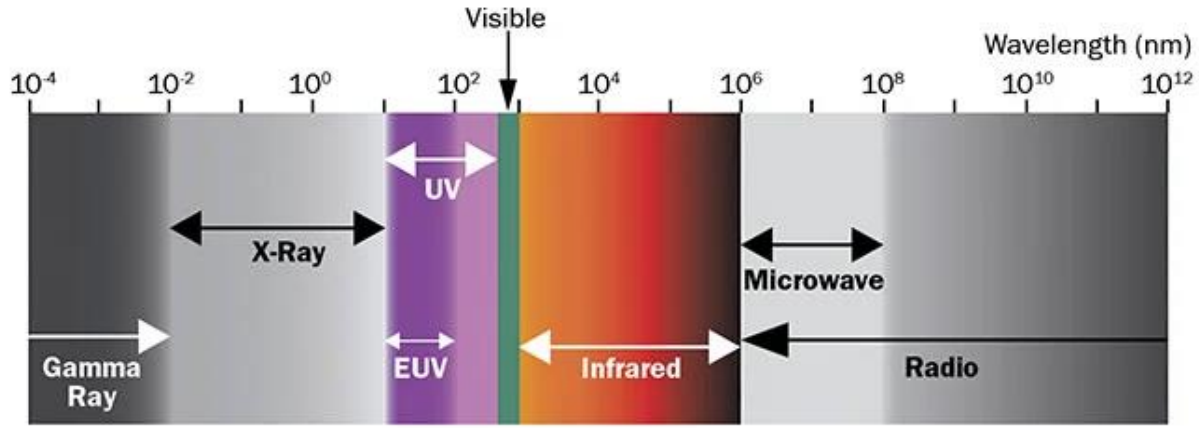
Multi Spectral Imaging

Measure electromagnetic radiation at wavelengths outside the visible spectrum



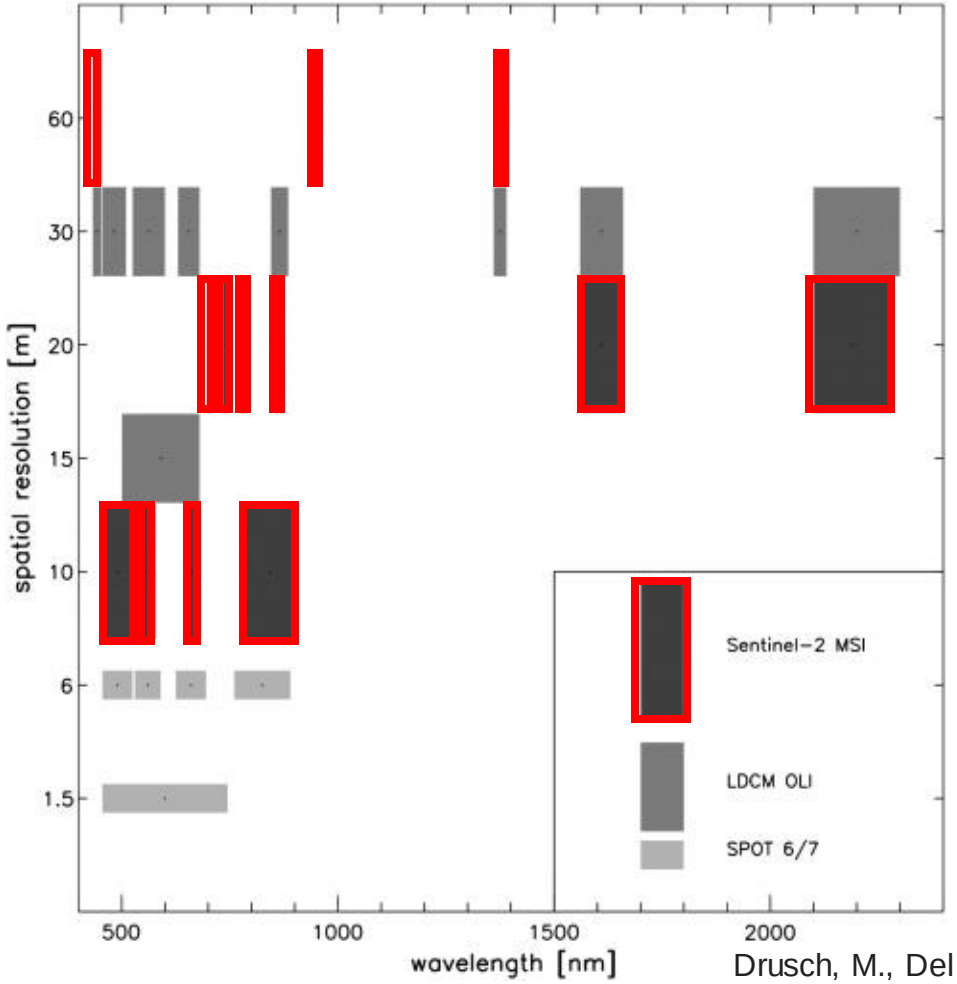
Multi-spectral image





- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, *120*, 25-36.
- https://www.photonics.com/Articles/Hyperspectral_and_Multispectral_Imaging/a65595
- https://wiki.tum.de/display/zfp/Hyper-+and+multispectral+imaging

Sentinel-2 Multi Spectral Imaging



13 spectral bands in the visible, near- and short wave infra-red

Spatial resolutions between 10 and 60m

Band	Spatial Resolution *m*	Central Wavelength *nm*
B01 - Aerosols	60	443
B02 - Blue	10	490
B03 - Green	10	560
B04 - Red	10	665
B05 - Red edge 1	20	705
B06 - Red edge 2	20	740
B07 - Red edge 3	20	783
B08 - NIR	10	842
B08A - Red edge 4	20	865
B09 - Water vapor	60	945
B10 - Cirrus	60	1375
B11 - SWIR 1	20	1610
B12 - SWIR 2	20	2190

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, *120*, 25-36

Sentinel-2 Multi Spectral Imaging

B4, B3, B2

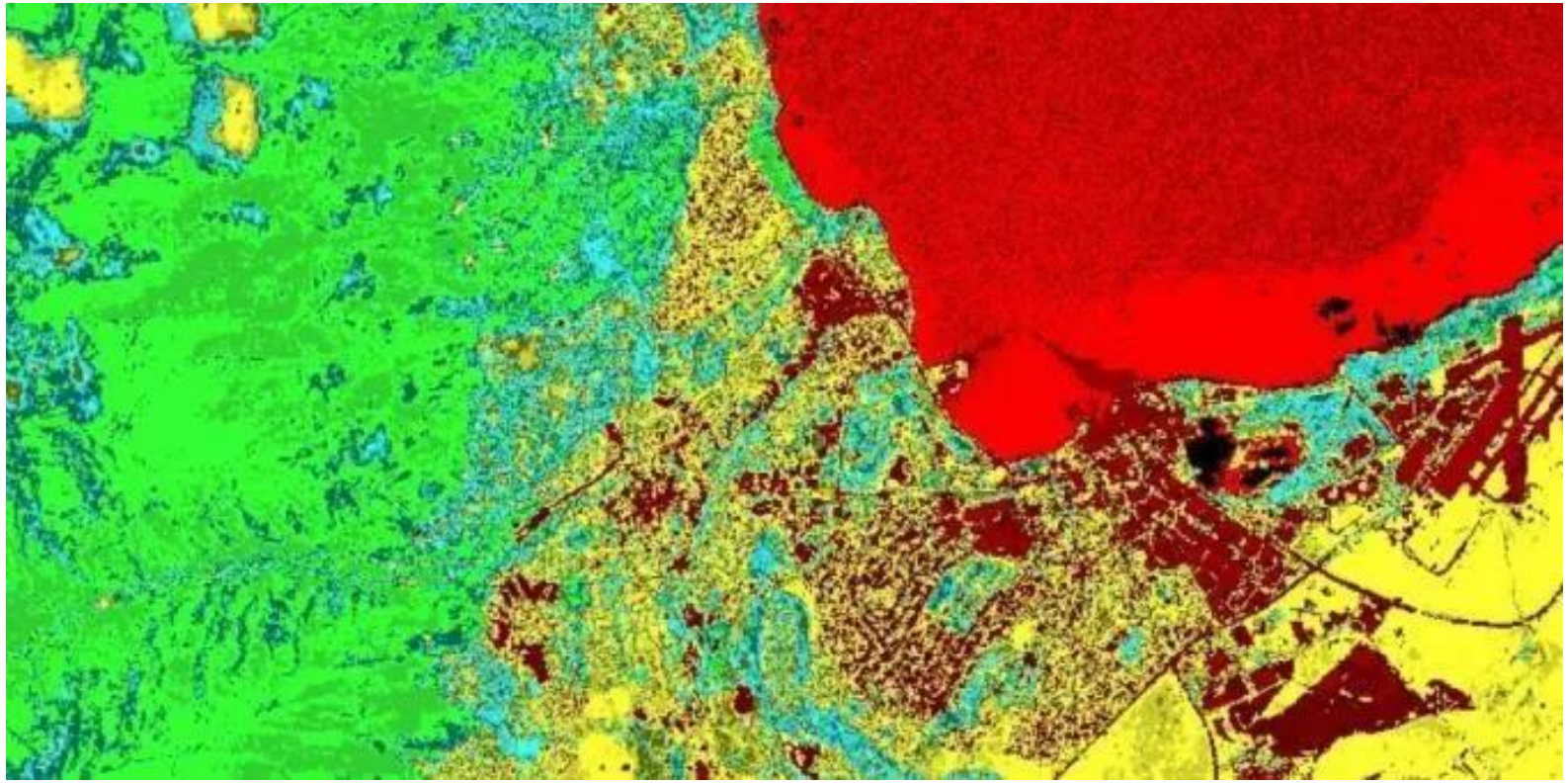
B11, B8, B2

B12, B8A, B5

B8, B4, B3

https://gisgeography.com/sentinel-2-bands-combinations/

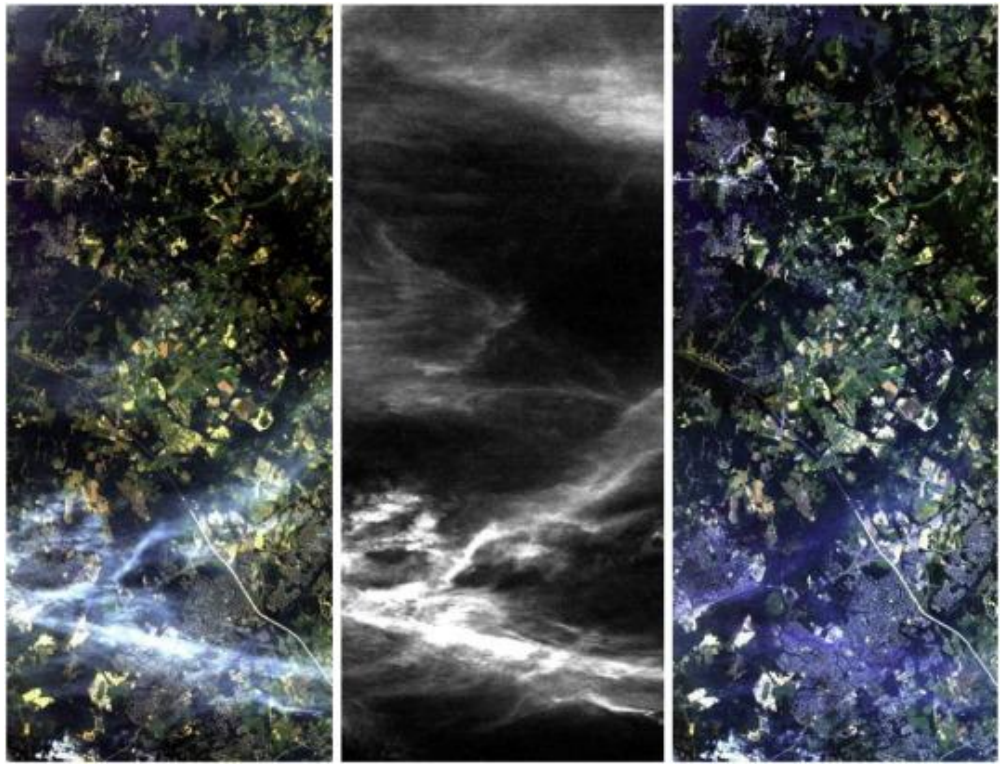
Normalized difference vegetation index (NDVI):
(B8-B4)/(B8+B4)



https://gisgeography.com/sentinel-2-bands-combinations/

Sentinel-2 Products

Level-1C
Top of Atmosphere reflectance



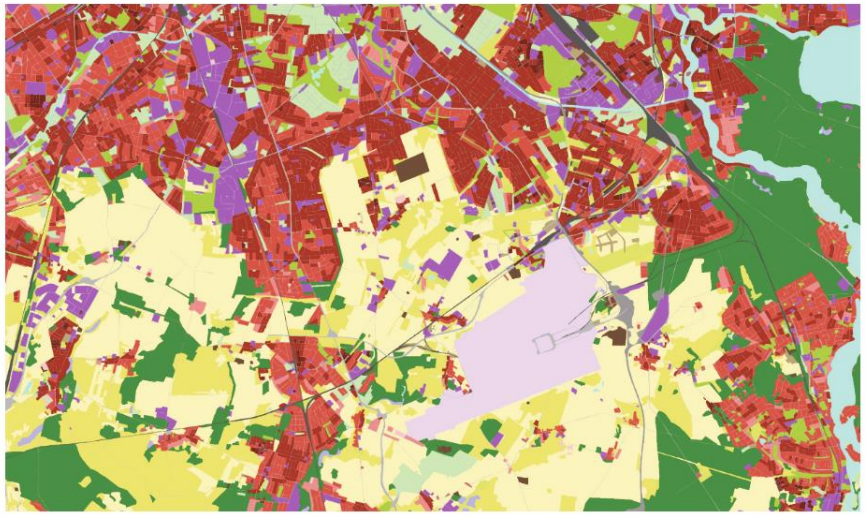
Cirrus cloud mask

Level-2A
Bottom of Atmosphere reflectance

European Urban Atlas

- Harmonized land-cover and land-use maps
- Hundreds of cities and their surrounding areas in Europe
- Initiative of the European Commission, European Environment Agency and ESA Copernicus
- 27 different land-use/cover classes

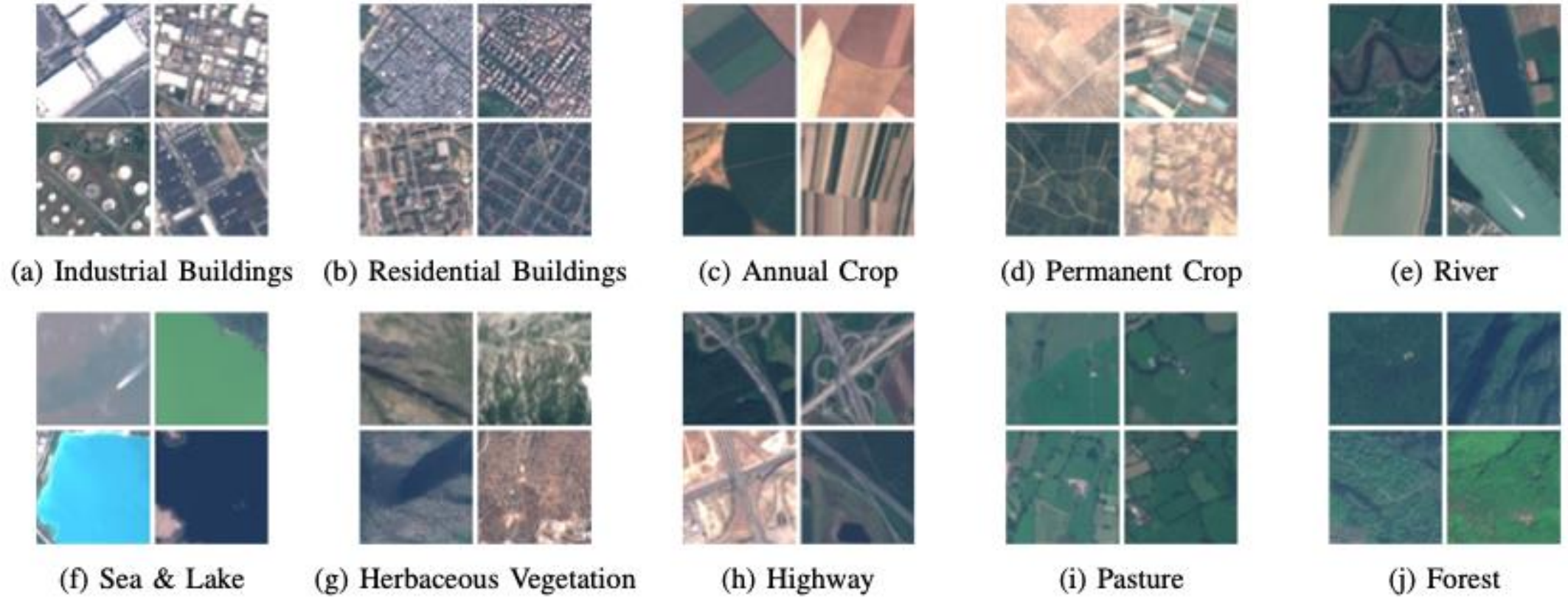
South Berlin 2012



Urban Atlas		2006	2012
Legende Code	Nomenclature		
11100	Continuous Urban Fabric (S.L.>80%)		
11210	Discontinuous Dense Urban Fabric (S.L.: 50% - 80%)		
11220	Discontinuous Medium Density Urban Fabric (S.L.: 30%-50%)		
11230	Discontinuous Low Density Urban Fabric (S.L.:10%-30%)		
11240	Discontinuous Very Low Density Urban Fabric (S.L.<10%)		
11300	Isolated Structures		
12100	Industrial, commercial, public, military and private units		
12210	Fast transit roads and associated land		
12220	Other roads and associated land		
12230	Railways and associated land		
12300	Port areas		
12400	Airports		
13100	Mineral extraction and dump sites		
13300	Construction sites		
13400	Land without current use		
14100	Green urban areas		
14200	Sports and leisure facilities		

- https://land.copernicus.eu/local/urban-atlas
- https://land.copernicus.eu/user-corner/publications/ua-flyer/at_download/file

Eurosat Classes



27,000 13x64x64 pixel images (2,000 – 3,000 per class)

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2217-2226.

Challenge: Land Cover Classification

Training Data

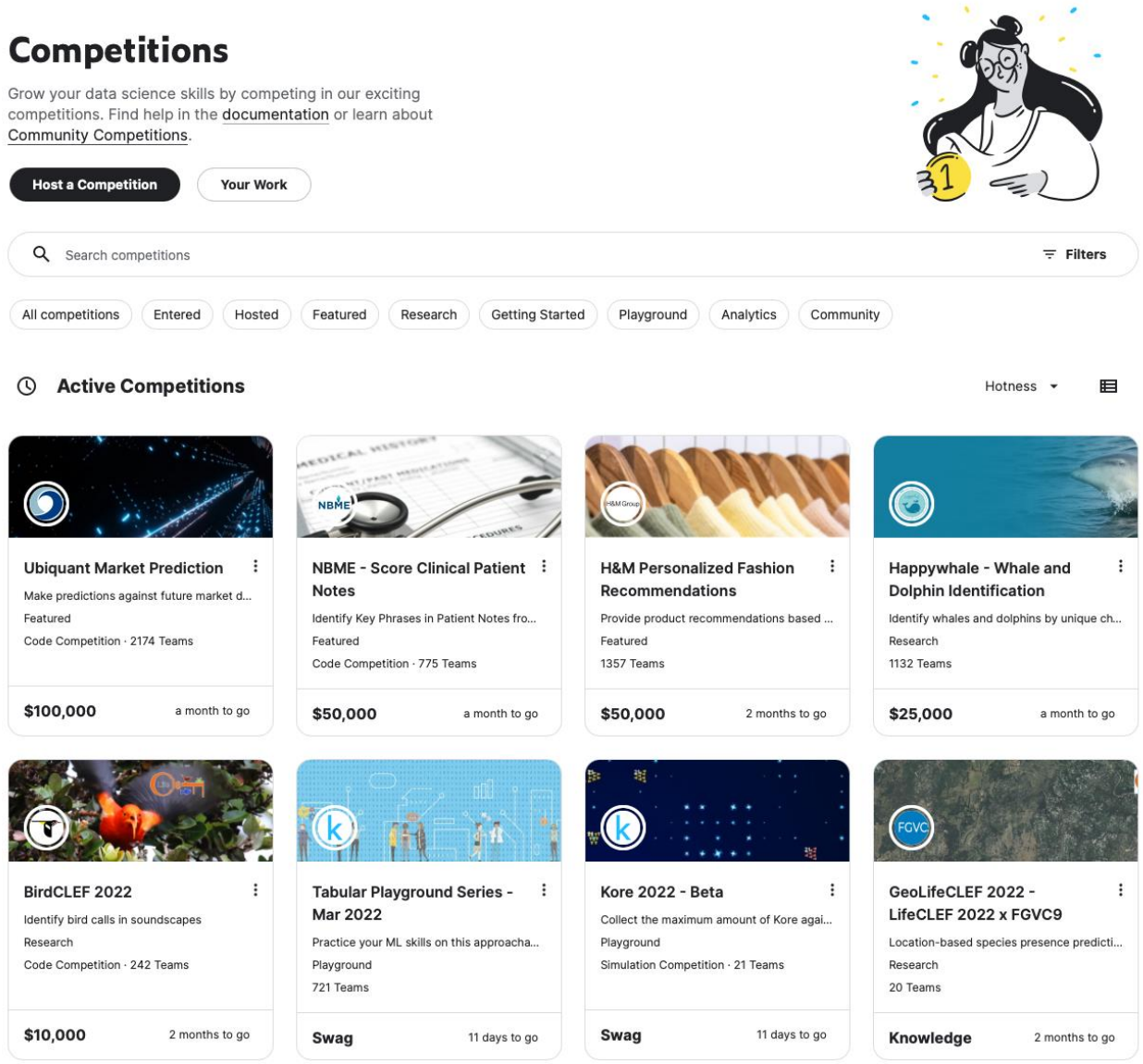
- Eurosat
- Level-1C
- 13 Bands
- GeoTiff

Test Data

- Atmospheric Correction
- Level-2A
- 12 bands
- Without B10 (Cirrus)
- npy files
- 4200 images

Kaggle Overview

- Platform for data science/machine learning challenges
- Challenge organizer provides data and task
- Participants can upload their solutions on test data
- Kaggle scores the solutions and creates a leaderboard
- Lots of additional functionality
  - Discussion board
  - Code sharing
  - Data hosting
  - …



www.kaggle.com

AIML Coding Challenge

- You need a Kaggle account to compete in the challenge
- Teams of (up to) two people:
  - Create teams on Kaggle in the `Team` tab
- Use the Kaggle discussion board to discuss with other students
  - We do not monitor conversations and might not respond there

Join challenge at:

- **https://www.kaggle.com/t/3054a89d9e9f451dacbf90bd47cb7f32**

Permanent challenge link:

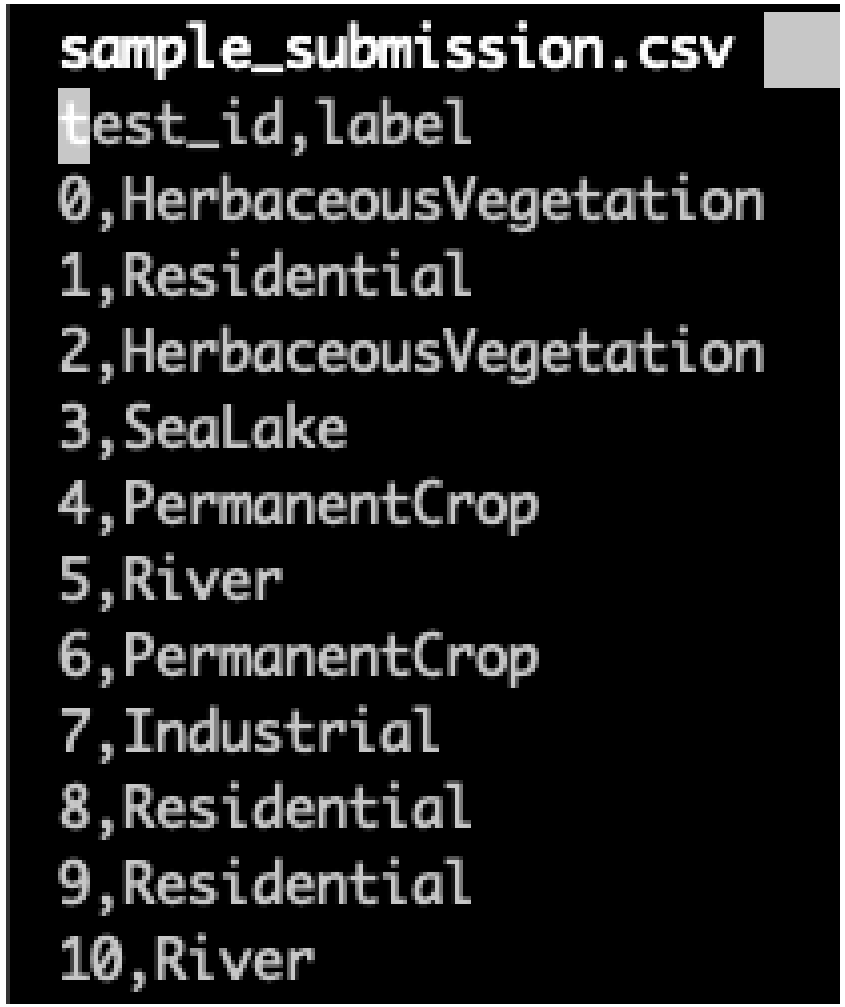
- **https://www.kaggle.com/c/aiml-coding-challenge-fs23**

Kaggle Submission Workflow

1. Download EuroSat and testset
2. Develop a model on EuroSat data
3. Apply the model on the testset
4. Upload a CSV with testset predictions to Kaggle
   - Mapping of test sample ids to predicted class label

Note: Kaggle provides **public** and **private** leaderboards

The final ranking will be based on the private leaderboard



Performance metric

Kaggle scores the submissions based on **accuracy.**

i.e., **the fraction of correctly classified samples**

Challenge Logistics

1. Team building: 11 teams
   - One team of 3 or 1
2. Lab on 17.04. coding challenge “mid-term”
   - Check-in/ Q&A
3. Last Kaggle submission: 21.05. 23:59pm
4. Final presentation:
   - 22.05.
   - 10 min. per team
   - Slide template will be provided
5. Grading:
   - 20% of course grade

Grading rubic

#	Performance Criteria	Max. Points	Min. Requirements for 2/4 pts	Min. Requirements for 3/6 pts	Min. Requirements for 4/8 pts	Min. Requirements for 5/10 pts
1	**Code Submission & Documentation**	5	Code for working model is submitted. Valid submission to Kaggle. Code reproduces csv file from Kaggle submission.	Code for working model is submitted with minimal documentation (e.g., cells have headers).	Code for working model is submitted with comprehensive documentation (code is wrapped in documented functions/classes).	Code for working model with excellent documentation (results are visualized e.g., accuracy over training epochs, confusion matrix, etc.)
2	**Classification Results** (private leaderboard)	10	Better than random guessing performance on test set (Kaggle > 15% accuracy).	Kaggle > 30% accuracy	Kaggle > 45% accuracy	Kaggle > 55% accuracy
6	**Final Presentation**	5	Presentation is submitted.	Presentation and slides are well structured.	Presentation and slides are well structured and individual slides are understandable.	Presentation and slides are well structured and individual slides are understandable. Provides details and visualizations are convincing.

Questions so far?

Course Labs and Challenge Roadmap

Event Type	Date	Description	Teaching Assistant	CC Phase
Lab 1	20 Feb 2023	Co-Lab Introduction	Shijun	Introduction
Lab 2	**27 Feb 2023**	**Support Vector Machine (SVM)**	**Shijun**	**Ramp-Up Phase**
CC 1	06 Mar 2023	**Coding Challenge - Kick-Off**	Linus	Coding Phase
Lab 3	13 Mar 2023	**Fully Connected Neural Networks**	Hamed	Ramp-Up Phase
Lab 4	20 Mar 2023	**Convolutional Neural Networks (CNNS)**	Hamed	Ramp-Up Phase
Lab 5	27 Mar 2023	**Long short-term memory (LSTM)**	Hamed	Ramp-Up Phase
Semester Break - Happy Easter!				
CC 2	17 Apr 2023	**Coding Challenge - Mid-Term**	Linus	Coding Phase
Lab 6	24 Apr 2023	**Attention**	Shijun	Coding Phase
Lab 7	01 May 2023	**K-Means, EM Clustering**	Shijun	Coding Phase
Lab 8	08 May 2023	**Autoencoder Anomaly Detection**	Hamed / Marco	Coding Phase
Lab 9	15 May 2023	**Transfer Learning / Self-Supervised Learning**	Shijun / Linus	Finalization Phase
CC 3	22 May 2023	**Coding Challenge – Final Presentation**	Linus	Finalization Phase

Getting Started

Let’s have a look at our data-loading and visualization notebook.

https://github.com/HSG-AIML-Teaching/ML2023-Lab/blob/main/cc_1/cc_01_getting_started.ipynb

Happy Coding! 🤓

Artificial Intelligence
and Machine Learning
Universität St.Gallen

MCS Machine Learning Class of Spring 2023

Chair of Artificial Intelligence
and Machine Learning [AI:ML]
Rosenbergstrasse 30, CH-9000 St. Gallen
aiml-teaching.ics@unisg.ch